RESEARCH   **Open Access**

Uncertainty and filtering of hidden Markov models in discrete time

Samuel N. Cohen

Abstract We consider the problem of filtering an unseen Markov chain from noisy observations, in the presence of uncertainty regarding the parameters of the processes involved. Using the theory of nonlinear expectations, we describe the uncertainty in terms of a penalty function, which can be propagated forward in time in the place of the filter. We also investigate a simple control problem in this context.

Keywords Filtering · Optimal control · Robustness · Nonlinear expectation

AMS Subject Classification 62M20 · 60G35 · 93E11

Abbreviations

BSDE: Backward stochastic difference equation
i.i.d.: Independent and identically distributed
MAP extimator: Maximum a posteriori estimator
DR-expectation: Either "data-driven robust estimator" or "divergence robust estima-tor" (the acronym is deliberately ambiguous)
StaticUP: Static generators, uncertain prior framework
DynamicUP: Dynamic generators, uncertain prior framework
StaticDR: Static generators, DR-expectation framework
DynamicDR: Dynamic generators, DR-expectation framework

S. N. Cohen (✉)
Mathematical Institute, University of Oxford, Woodstock Road, Oxford, UK
e-mail: samuel.cohen@maths.ox.ac.uk

1 Introduction

Filtering is a common problem in many applications. The essential concept is that there is an unseen Markov process, which influences the state of some observed process, and our task is to approximate the state of the unseen process using a form of Bayes' theorem. Many results have been obtained in this direction, most famously the Kalman filter (Kalman 1960; Kalman and Bucy 1961), which assumes the underlying processes considered are Gaussian, and gives explicit formulae accordingly. Similarly, under the assumption that the underlying process is a finite-state Markov chain, a general formula to calculate the filter can be obtained (the Wonham filter Wonham (1965)). These results are well known, in both discrete and continuous time (see Bain and Crisan (2009) or Cohen and Elliott (2015) Chapter 21 for further general discussion).

In this paper, we consider a simple setting in discrete time, where the underlying process is a finite-state Markov chain. Our concern is to study uncertainty in the dynamics of the underlying processes, in particular, its effect on the behaviour of the corresponding filter. That is, we assume that the observer has only imperfect knowledge of the dynamics of the underlying process and of their relationship with the observation process, and wishes to incorporate this uncertainty in their estimates of the unseen state. We are particularly interested in allowing the level of uncertainty in the filtered state to be endogenous to the filtering problem, arising from the uncertainty in parameter estimates and process dynamics.

We model this uncertainty in a general manner, using the theory of nonlinear expectations, and concern ourselves with a description of uncertainty for which explicit calculations can be carried out, and which can be motivated by considering statistical estimation of parameters. We then apply this to building a dynamically consistent expectation for random variables based on future states, and to a general control problem, with learning, under uncertainty.

1.1 Basic filtering

Consider two stochastic processes, $X = \{X_t\}_{t \geq 0}$ and $Y = \{Y_t\}_{t \geq 0}$. Let Ω be the space of paths of (X, Y) and $\mathbb{P}$ be a probability measure on Ω. We denote by $\{\mathcal{F}_t\}_{t \geq 0}$ the (completed) filtration generated by X and Y and denote by $\mathcal{Y} = \{\mathcal{Y}_t\}_{t \geq 0}$ the (completed) filtration generated by Y. The key problem of filtering is to determine estimates of $\phi(X_t)$ given $\mathcal{Y}_t$, that is, $\mathbb{E}_{\mathbb{P}}[\phi(X_t)|\mathcal{Y}_t]$, where ϕ is an arbitrary Borel function.

Suppose that X is a Markov chain with (possibly time-dependent) transition matrix $A_t^\top$ under $\mathbb{P}$ (the transpose here saves notational complexity later). Without loss of generality, we assume that X takes values in the standard basis vectors $\{e_i\}_{i=1}^N$ of $\mathbb{R}^N$ (where N is the number of states of X), and so we write

$$X_t = A_t X_{t-1} + M_t,$$

where $\mathbb{E}_{\mathbb{P}}[M_{t+1}|\mathcal{F}_t] = 0$, so $\mathbb{E}_{\mathbb{P}}[X_t|\mathcal{F}_{t-1}] = A_t X_{t-1}$.

We suppose the process Y is multivariate real-valued[1]. The law of Y depends on X, in particular, the $\mathbb{P}$-distribution of Y_t given $\{X_s\}_{s \leq t} \cup \{Y_s\}_{s<t}$ (that is, given all past observations of X and Y and the current state of X) is

$$Y_t \sim c(y; t, X_t) d\mu(y)$$

for μ a reference measure on $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$, where $\sim$ is used to indicate the density of the distribution of a random variable.

For simplicity, we assume that $Y_0 \equiv 0$, so no information is revealed about X_0 at time 0. It is convenient to write $C_t(y) = C(y; t)$ for the diagonal matrix with entries $c(y; t, e_i)$, so that

$$C_t(y) X_t = c(y; t, X_t) X_t.$$

Note that these assumptions, in particular the values of A and C, depend on the choice of probability measure $\mathbb{P}$. Conversely, as our space Ω is the space of paths of (X, Y), the measure $\mathbb{P}$ is determined by A and C. We call A and C the *generators* of our probability measure.

As we have assumed X_t takes values in the standard basis in $\mathbb{R}^N$, the expectation $\mathbb{E}_{\mathbb{P}}[X_t | \mathcal{Y}_t]$ determines the entire conditional distribution of X_t given $\mathcal{Y}_t$. In this discrete time context, the filtering problem can be solved in a fairly simple manner: Suppose we have already calculated $p_{t-1} := \mathbb{E}_{\mathbb{P}}[X_{t-1} | \mathcal{Y}_{t-1}]$. Then, by linearity and the dynamics of X, using the fact

$$\mathbb{E}_{\mathbb{P}}\left[M_t | \mathcal{Y}_{t-1}\right] = \mathbb{E}_{\mathbb{P}}\left[\mathbb{E}_{\mathbb{P}}\left[M_t | \mathcal{F}_{t-1}\right] \big| \mathcal{Y}_{t-1}\right] = 0,$$

we can calculate

$$\mathbb{E}_{\mathbb{P}}\left[X_t | \mathcal{Y}_{t-1}\right] = \mathbb{E}_{\mathbb{P}}\left[A_t X_{t-1} + M_t | \mathcal{Y}_{t-1}\right] = A_t p_{t-1}.$$

Bayes' theorem then states that, with probability one, with $\propto$ denoting equality up to proportionality,

$$\mathbb{P}\left(X_t = e_i | \mathcal{Y}_t\right) = \mathbb{P}\left(X_t = e_i | \{Y_s\}_{s<t}, Y_t\right) \propto c(Y_t; t, e_i) \mathbb{P}(X_t = e_i | \mathcal{Y}_{t-1}),$$

which can be written in a simple matrix form, as given in the following theorem which summarizes the classical filter.

Theorem 1 *For X a hidden Markov chain with transition matrix $A_t^\top$, and Y an observation process with conditional density (given X_t) given by*

$$Y_t | X_t \sim c(y; t, X_t) d\mu(y) = \mathbf{1}^\top C_t(y) X_t d\mu(y),$$

the conditional distribution $\mathbb{E}[X_t | \mathcal{Y}_t] = p_t$ satisfies the recursion

$$p_t = \mathbb{G}(p_{t-1}, A_t, C_t, Y_t) := \frac{C_t(Y_t) A_t p_{t-1}}{\mathbf{1}^\top C_t(Y_t) A_t p_{t-1}}, \tag{1}$$

where $\mathbf{1}$ denotes a vector with all components 1.

---

[1]This assumption can easily be relaxed, to allow Y to take values in an appropriate Polish or Blackwell space. We restrict to the real setting purely for simplicity.

We call p_t the "filter state" at time t. Note that, if we assume the density c is positive, A_t is irreducible and p_{t-1} has all entries positive, then p_t will also have all entries positive.

In practice, the key problem with implementing these methods is the requirement that we know the underlying transition matrix $A^\top$ and the density C. These are generally not known perfectly, but need to be estimated prior to the implementation of the filter. Uncertainty in the choice of these parameters will lead to uncertainty in the estimates of the filtered state, and the aim of this paper is to derive useful representations of that uncertainty.

As variation in the choice of A and C corresponds to a different choice of measure $\mathbb{P}$, we see that using an uncertain collection of generators corresponds naturally to uncertainty regarding $\mathbb{P}$. This type of uncertainty, where the probability measure is not known, is commonly referred to as "Knightian" uncertainty (with reference to Knight (1921), related ideas are also discussed by Keynes (1921) and Wald (1945)).

Effectively, we wish to consider the propagation of uncertainty in Bayesian updating (as the filter is simply a special case of this). Huber and Ronchetti 2009 p. 331 briefly touch on this, however (based on earlier work by Kong) argue that this propagation is computationally infeasible. However, their approach was based on Choquet integrals, rather than nonlinear (convex) expectations in the style of Peng (2010) and others. In the coming sections, we see how the structure of nonlinear expectations allows us to derive comparatively simple rules for updating.

Remark 1 *While we will present our theory in the context where X is a finite state Markov chain, our approach does not depend in any significant way on this assumption. In particular, it would be equally valid, mutatis mutandis, if we supposed that X followed the dynamics of the Kalman filter, and our uncertainty was on the coefficients of the filter. We specialize to the Markov chain case purely for the sake of concreteness.*

The aim of this paper is to provide, with a minimum of technical complexity, the basic structures which underlie this approach to filtering with a nonlinear expectation. It proceeds as follows: In Section 2 we give some key facts about the measures which naturally appear in a filtering context. In Section 3, we introduce the theory of nonlinear expectations, and a means of connecting these with statistical estimation from Cohen (2017). Section 4 unites these expectations with filtering, giving recursive equations which replace the filtering equation in Theorem 1, it also outlines some concrete simplifications of this general structure, depending on whether the underlying parameters can vary through time and how new information is to be incorporated. In Section 5 we consider dynamic properties of this nonlinear expectation when looking at future events, and some connections with the theory of (discrete time) BSDEs. Finally, Section 6 considers a generic control problem in this context.

2 Conditionally Markov measures

In order to incorporate learning in our nonlinear expectations and filtering, it is useful to extend slightly from the family of measures previously described. In particular, we

wish to allow the dynamics to depend on past observations, while preserving enough Markov structure to enable filtering. The following classes of probability measures will be of interest.

Definition 1 *We write $\mathcal{M}_1$ for the space of probability measures equivalent to a reference measure $\mathbb{P}$.*

Let $\mathcal{M}_M \subset \mathcal{M}_1$ denote the probability measures under which

- *X is a Markov chain, that is, for all t, X_{t+1} is independent of $\mathcal{F}_t$ given X_t;*
- *$\{Y_s\}_{s\geq t+1}$ is independent of $\mathcal{F}_t$ given X_{t+1};*
- *both X and Y are time homogeneous, that is, the conditional distributions of $X_{t+1}|X_t$ and $Y_t|X_t$ do not depend on t.*

Let $\mathcal{M}_{M|\mathcal{Y}} \subset \mathcal{M}_1$ denote the probability measures under which

- *X is a conditional Markov chain, that is, for all t, X_{t+1} is independent of $\mathcal{F}_t$ given X_t and $\{Y_s\}_{s\leq t}$; and*
- *$\{Y_s\}_{s\geq t+1}$ is independent of $\mathcal{F}_t$ given $\{X_{t+1}\} \cup \{Y_s\}_{s\leq t}$.*

We note that, if we consider a measure in $\mathcal{M}_{M|\mathcal{Y}}$, there is a natural notion of the generators A and C. In particular, $\mathcal{M}_M$ corresponds to those measures under which the generators A and C are constant, while $\mathcal{M}_{M|\mathcal{Y}}$ corresponds to those measures under which the generators A and C are functions of time and $\{Y_s\}_{s\leq t}$ (i.e. $\{\mathcal{Y}_t\}_{t\geq 0}$-predictable processes).

For each t, these generators determine the measure on $\mathcal{F}_t$ given $\mathcal{F}_{t-1}$, and (together with the distribution of X_0) this determines the measure at all times. It is straightforward to verify that our filtering equations hold for all measures in $\mathcal{M}_{M|\mathcal{Y}}$, with the appropriate modification of the generators.

Definition 2 *For a measure $\mathbb{Q} \in \mathcal{M}_{M|\mathcal{Y}}$, we write $\left(A^{\mathbb{Q}}, C^{\mathbb{Q}}(\cdot)\right)$ for the generator of (X, Y) under $\mathbb{Q}$, recalling that $C_t^{\mathbb{Q}}(y) = \text{diag}\left(\left\{c_t^{\mathbb{Q}}(y; e_i)\right\}_{i=1}^N\right)$, and that $A_t^{\mathbb{Q}}$ and $C_t^{\mathbb{Q}}$ are now allowed to depend on $\{Y_s\}_{s<t}$. For notational convenience, we shall typically not write the dependence on $\{Y_s\}_{s<t}$ explicitly.*

Similarly, for a $\{\mathcal{Y}_t\}_{t\geq 0}$-predictable process $(A_t, C_t(\cdot))_{t\geq 0}$ taking values in the product of the space of transition matrices and the space of diagonal matrix-valued functions, where each diagonal element is a probability density on $\mathbb{R}^d$, and p_0 a probability vector in $\mathbb{R}^N$, we write $\mathbb{Q}(A, C, p_0)$ for the measure with generator $(A_t, C_t(\cdot))_{t\geq 0}$ and initial distribution $\mathbb{E}_{\mathbb{Q}}[X_0] = p_0$.

In what follows, we will be variously wishing to *restrict* a measure $\mathbb{Q}$ to a σ-algebra, and to *condition* a measure on a σ-algebra. To prevent notational confusion, we shall write $\mathbb{Q}\|_{\mathcal{F}}$ for the restriction of $\mathbb{Q}$ to $\mathcal{F}$, and $\mathbb{Q}|_{\mathcal{F}}$ for $\mathbb{Q}$ conditioned on $\mathcal{F}$.

In our setting, our fundamental problem is that we do not know what measure is "true", and so work instead under a family of measures. In this context, measure changes can be described as follows.

Proposition 1 *Let $\bar{\mathbb{P}}$ be a reference probability measure under which X is a sequence of i.i.d. uniform random variables from the basis vectors $\{e_1, ...e_N\} \subset \mathbb{R}^N$*

and $\{Y_t\}_{t\geq0}$ is independent of X, with i.i.d. distribution $Y_t \sim d\mu$. The measure $\mathbb{Q}(A, C, p_0) \in \mathcal{M}_{M|\mathcal{Y}}$ has Radon–Nikodym derivative (or likelihood)

$$\frac{d\mathbb{Q}(A, C, p_0)\|_{\mathcal{F}_T}}{d\bar{\mathbb{P}}\|_{\mathcal{F}_T}} = N\left(X_0^\top p_0\right) \prod_{t=1}^T \left(\left(X_t^\top A_{t-1} X_{t-1}\right)\left(\mathbf{1}^\top C(Y_t) X_t\right)\right).$$

Remark 2 *The requirement that $\bar{\mathbb{P}}$ is a probability measure is unnecessary (i.e., in Bayesian parlance, the reference distribution may be improper). For example, we can use Lebesgue measure on $\mathbb{R}$ as the marginal reference measure μ for Y_t without difficulty, in which case $c_t(y)$ is the usual (Lebesgue) density of the distribution of Y_t.*

Proof A simple verification of this result is possible by factoring the proposed Radon–Nikodym density as the product of three terms:

$$\frac{d\mathbb{Q}(A, C, p_0)\|_{\mathcal{F}_T}}{d\bar{\mathbb{P}}\|_{\mathcal{F}_T}} = \frac{X_0^\top p_0}{1/N} \cdot \prod_{t=1}^T \left(X_t^\top A_{t-1} X_{t-1}\right) \cdot \prod_{t=1}^T \left(\mathbf{1}^\top C(Y_t) X_t\right).$$

The first term, $(X_0^\top p_0)/(1/N)$, changes the distribution of X_0 from uniform to p_0, as is seen from the calculation

$$E_{\bar{\mathbb{P}}}\left[X_0\left(X_0^\top p_0\right)/(1/N)\right] = N E_{\bar{\mathbb{P}}}\left[X_0 X_0^\top\right] p_0 = N N^{-1} I_N p_0 = p_0.$$

This is clearly a probability density with respect to $\bar{\mathbb{P}}$.

The second term, $\prod_{t=1}^T \left(X_t^\top A_{t-1} X_{t-1}\right)$, changes the conditional distribution of $X_t | \{X_s\}_{s<t}$ (for each t), from a uniform distribution to the probability vector $A_{t-1} X_{t-1}$; as can be demonstrated by a calculation similar to that for X_0 in the first term. As the columns of A_{t-1} sum to one, it is easy to verify that this product has expectation 1 (conditional on X_0) and is nonnegative, that is, it is a probability density which does not modify the distribution of X_0.

The third term changes the conditional distribution of $Y_t | \{X_t, X_s, Y_s\}_{s\neq t}$ from μ to $C(y) X_t d\mu(y) = c(y; X_t) d\mu(y)$. This is most easily seen by calculating

$$E_{\mathbb{Q}}[g(Y_t)|\{X_t, X_s, Y_s\}_{s\neq t}] = \int g(y) \mathbf{1}^\top C(y) X_t d\mu(y) = \int g(y) c(y; X_t) d\mu(y)$$

for a general bounded Borel function g. As c is defined to be a density, it is again easy to verify that the product $\prod\left(\mathbf{1}^\top C(Y_t) X_t\right)$ is a probability density with respect to $\bar{\mathbb{P}}$, and that it does not modify the distribution of X.

As we are on a canonical space, the measure $\mathbb{Q}$ is determined by the laws of X and Y, and so we have the result. $\square$

The above proposition gives a Radon–Nikodym derivative adapted to the full filtration $\{\mathcal{F}_t\}_{t\geq0}$. In practice, it is also useful to consider the corresponding Radon–Nikodym derivative adapted to the observation filtration $\{\mathcal{Y}_t\}_{t\geq0}$. As this filtration is generated by the process Y, it is enough to multiply together the conditional

distributions of $Y_t|\mathcal{Y}_{t-1}$, leading to the following convenient representation. Recall that

$$\sum_i p_i c_t(y; e_i) = \mathbf{1}^\top C_t(y) p.$$

Proposition 2 *For* $\mathbb{Q}(A, C, p_0) \in \mathcal{M}_{M|\mathcal{Y}}$, *the Radon–Nikodym derivative restricted to* $\mathcal{Y}_T$ *is given by*

$$L^{obs}(\mathbb{Q}(A, C, p_0)|\mathbf{y}) := \frac{d\mathbb{Q}(A, C, p_0)\|_{\mathcal{Y}_T}}{d\bar{\mathbb{P}}\|_{\mathcal{Y}_T}} = \prod_{t=1}^{T} \mathbf{1}^\top C_t(Y_t) A_t p_{t-1}^{(A,C), p_0},$$

where $p_{t-1}^{(A,C), p_0}$ *is the solution to the filtering problem in the measure* $\mathbb{Q}(A, C, p_0)$, *as determined by* (1) *(and so includes further dependence on* $\{Y_s\}_{s<t}$).

Proof The distribution of $Y_t|\mathcal{Y}_{t-1}$ is determined by (for Borel sets B)

$$
\begin{aligned}
E_{\mathbb{Q}}[1_{Y_t \in B}|\mathcal{Y}_{t-1}] &= E_{\mathbb{Q}}\Big[E_{\mathbb{Q}}[1_{Y_t \in B}|\mathcal{Y}_{t-1} \vee \sigma(X_t)]\Big|\mathcal{Y}_{t-1}\Big] \\
&= E_{\mathbb{Q}}\Big[\int_B \mathbf{1}^\top C_t(y) X_t d\mu(y)\Big|\mathcal{Y}_{t-1}\Big] \\
&= \int_B \mathbf{1}^\top C_t(y) E_{\mathbb{Q}}[X_t|\mathcal{Y}_{t-1}] d\mu(y) \\
&= \int_B \mathbf{1}^\top C_t(y) A_t p_{t-1}^{(A,C), p_0} d\mu(y).
\end{aligned}
$$

As $Y_t|\mathcal{Y}_{t-1}$ has distribution μ under $\bar{\mathbb{P}}$, it follows that $\mathbf{1}^\top C_t(Y_t) A_t p_{t-1}^{(A,C), p_0}$ is the Radon–Nikodym density of the conditional law of $Y_t|\mathcal{Y}_{t-1}$ under $\mathbb{Q}$ with respect to its law under $\bar{\mathbb{P}}$. As $\{Y_s\}_{s \leq T}$ generates $\mathcal{Y}_T$, the result follows by induction.   □

In order to apply classical statistical methods, we take a (generic) parameterization of this family of measures. This will also allow us to encode which parts of the generators we believe are static (and so can be learnt from observations), and which are dynamic (and so will violate stationarity).

Assumption 1 *For fixed* $m > 0$, *we assume we are given a (Borel measurable) function* $\Phi : \mathbb{N} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{A}$ *such that the generators satisfy*

$$\Phi(t, \mathfrak{S}, \mathfrak{D}_t) = (A_t, C_t),$$

where $\mathfrak{S}$ *is a static parameter, and* $\mathfrak{D}_t$ *is a parameter which may vary at each point in time. We write* $\mathcal{Q}$ *for the family of measures in* $\mathcal{M}_{M|\mathcal{Y}}$ *induced by this parameterization, and typically omit to write the argument t of* Φ.

With a slight abuse of notation, if $(A_t, C_t) = \Phi(t, \mathfrak{S}, \mathfrak{D}_t)$, *we write* $\mathbb{Q}(\mathfrak{S}, \mathfrak{D}, p_0)$ *as an alias for the measure* $\mathbb{Q}(A, C, p_0)$, *and* $\mathbb{G}(t, \mathfrak{S}, \mathfrak{D}_t, p)$ *as an alias for the function* $\mathbb{G}(t, A, C, p)$ *defined in* (1).

3 Nonlinear expectations

In this section, we introduce the concepts of nonlinear expectations and convex risk measures, and discuss their connection with penalty functions on the space of measures. These objects provide a technical foundation with which to model the presence of uncertainty in a random setting. This theory is explored in some detail in Föllmer and Schied (2002b). Other key works which have used or contributed to this theory, in no particular order, are Hansen and Sargent (2008) (see also Hansen and Sargent (2005, 2007) for work related to what we present here); Huber and Roncetti (2009), Peng (2010), El Karoui et al. (1997), Delbaen et al. (2010), Duffie and Epstein (1992), Rockafellar et al. (2006), and Riedel (2004) and Epstein and Schneider (2003). We base our terminology on that used in Föllmer and Schied (2002b) and Delbaen et al. (2010).

We here present, without proof, the key details of this theory as needed for our analysis.

Definition 3 *For a σ-algebra $\mathcal{G}$ on Ω, let $L^\infty(\mathcal{G})$ denote the space of essentially bounded $\mathcal{G}$-measurable random variables. A nonlinear expectation on $L^\infty(\mathcal{G})$ is a mapping*

$$\mathcal{E} : L^\infty(\mathcal{G}) \to \mathbb{R}$$

satisfying the assumptions

- *Strict Monotonicity: for any $\xi_1, \xi_2 \in L^\infty(\mathcal{G})$, if $\xi_1 \geq \xi_2$ a.s., then $\mathcal{E}(\xi_1) \geq \mathcal{E}(\xi_2)$ and, if in addition $\mathcal{E}(\xi_1) = \mathcal{E}(\xi_2)$, then $\xi_1 = \xi_2$ a.s.;*
- *Constant triviality: for any constant k, $\mathcal{E}(k) = k$;*
- *Translation equivariance: for any $k \in \mathbb{R}$, $\xi \in L^\infty(\mathcal{G})$, $\mathcal{E}(\xi + k) = \mathcal{E}(\xi) + k$.*

A "convex" expectation also satisfies

- *Convexity: for any $\lambda \in [0, 1]$, $\xi_1, \xi_2 \in L^\infty(\mathcal{G})$,*

$$\mathcal{E}(\lambda\xi_1 + (1 - \lambda)\xi_2) \leq \lambda\mathcal{E}(\xi_1) + (1 - \lambda)\mathcal{E}(\xi_2).$$

If $\mathcal{E}$ is a convex expectation, then the operator defined by $\rho(\xi) = \mathcal{E}(-\xi)$ is called a convex risk measure. A particularly nice class of convex expectations is those which satisfy

- *Lower semicontinuity: For a sequence $\{\xi_n\}_{n \in \mathbb{N}} \subset L^\infty(\mathcal{G})$ with $\xi_n \uparrow \xi$ pointwise (and $\xi \in L^\infty(\mathcal{G})$), we have $\mathcal{E}(\xi_n) \uparrow \mathcal{E}(\xi)$.*

The following theorem (which was expressed in the language of risk measures) is due to Föllmer and Schied (2002a) and Frittelli and Rosazza Gianin (2002).

Theorem 2 *Suppose $\mathcal{E}$ is a lower semicontinuous convex expectation. Then there exists a "penalty" function $\mathcal{R} : \mathcal{M}_1 \to [0, \infty]$ such that*

$$\mathcal{E}(\xi) = \sup_{\mathbb{Q} \in \mathcal{M}_1} \left\{ \mathbb{E}_{\mathbb{Q}}[\xi] - \mathcal{R}(\mathbb{Q}) \right\}.$$

Provided $\mathcal{R}(\mathbb{Q}) < \infty$ for some $\mathbb{Q}$ equivalent to $\mathbb{P}$, we can restrict our attention to measures in $\mathcal{M}_1$ equivalent to $\mathbb{P}$ without loss of generality.

Remark 3 *This result gives some intuition as to how a convex expectation can model "Knightian" uncertainty. One considers all the possible probability measures on the space, and then selects the maximal expectation among all measures, penalizing each measure depending on its plausibility, as measured by $\mathcal{R}(\cdot)$. As convexity of $\mathcal{E}$ is a natural requirement of an "uncertainty averse" assessment of outcomes, Theorem 2 shows that this is the only way to construct an "expectation" $\mathcal{E}$ which penalizes uncertainty, while preserving monotonicity, translation equivariance, and constant triviality (and lower semicontinuity).*

In order to relate our penalty function with the temporal structure of filtering, we focus our attention on measures in our parametric family $\mathcal{Q}$, defined in Assumption 1. We also allow the penalty $\mathcal{R}$ to depend on time. In the analysis of this paper, the following definition allows us to obtain a (forward) recursive structure in our nonlinear expectations, as one might expect in a filtering context.

Definition 4 *We say a family of penalty functions $\{\mathcal{R}_t\}_{t \geq 0}$, is additive if it can be written in the form*

$$\mathcal{R}_t(\mathbb{Q}) = \begin{cases} \left(\frac{1}{k} \alpha_t (\mathbb{Q}, \{Y_s\}_{s \leq t}) \right)^{k'} & \text{if } \mathbb{Q} \in \mathcal{Q} \\ +\infty & \text{otherwise,} \end{cases}$$

where, if $\mathbb{Q} = \mathbb{Q}(\mathfrak{S}, \mathfrak{D}, p_0)$ and $p_t^{\mathbb{Q}}$ is the solution of the filtering Eq. 1 under $\mathbb{Q}$, the function α_t is of the form

$$\alpha_t \left(\mathbb{Q}, \{Y_s\}_{s \leq t} \right) = \kappa_{prior}(p_0, \mathfrak{S}) + \sum_{s \leq t} \gamma_s \left(\mathfrak{S}, \mathfrak{D}_s, \{Y_n\}_{n \leq s}, p_{s-1}^{\mathbb{Q}} \right) + m_t.$$

Here k and k' are positive constants, κ_{prior} and $\{\gamma_t\}_{t \geq 0}$ are known real functions bounded below, and m_t is a $\mathcal{Y}_t$-measurable scalar random variable which ensures the normalization condition $\inf_{\mathbb{Q}} \alpha \left(\mathbb{Q}, \{Y_s\}_{s \leq t} \right) = 0$ holds for almost all observation sequences $\{Y_s\}_{s \geq 0}$.

3.1 DR-expectations

From the discussion above, it is apparent that we can focus our attention on calculating the penalty function $\mathcal{R}$, rather than working with the nonlinear expectation directly. This penalty function is meant to encode how "unreasonable" a probability measure $\mathbb{Q}$ is as a model for our outcomes.

In Cohen (2017), we have considered a framework which links the choice of the penalty function to statistical estimation of a model. The key idea of Cohen (2017) is to use the negative log-likelihood function for this purpose, where the likelihood is taken against an arbitrary reference measure, and evaluated using the observed data. This directly uses the statistical information from observations to quantify our uncertainty.

In this paper, we make a slight extension of this idea, to explicitly incorporate prior beliefs. In particular, we replace the log-likelihood with the log-posterior density, which in turn gives an additional term in the penalty.

Definition 5 *For $\mathbb{Q} \in \mathcal{Q}$, the observed likelihood $L^{obs}(\mathbb{Q}|\mathbf{y})$ is given in Proposition 2. Inspired by a "Bayesian" approach, we augment this by the addition of a prior distribution over $\mathcal{Q}$. Suppose a (possibly improper) prior is given, with density in terms of the parameters $(\mathfrak{S}, \{\mathfrak{D}_t\}_{t \geq 0})$*

$$\exp\left(-\kappa_{\text{prior}}(\mathfrak{S}, p_0) - \sum_t \gamma_{\text{prior}}(\mathfrak{S}, \mathfrak{D}_t)\right).$$

The posterior relative density is given by the product

$$L(\mathbb{Q}|\mathbf{y}) = L^{obs}(\mathbb{Q}|\mathbf{y}) \exp\left(-\kappa_{\text{prior}}(\mathfrak{S}) - \sum_t \gamma_{\text{prior}}(\mathfrak{S}, \mathfrak{D}_t)\right).$$

The "$\mathcal{Q}|\mathbf{y}_t$-divergence" is defined to be the normalized negative log-posterior relative density

$$\alpha_{\mathbf{y}}(\mathbb{Q}) := -\log\left(L(\mathbb{Q}|\mathbf{y})\right) + \sup_{\tilde{\mathbb{Q}} \in \mathcal{Q}} \left\{\log\left(L\left(\tilde{\mathbb{Q}}|\mathbf{y}\right)\right)\right\}. \tag{2}$$

Remark 4 *The right-hand side of (2) is well defined whether or not a maximum a posteriori (MAP) estimator exists. Given a MAP estimate $\hat{\mathbb{Q}} \in \mathcal{Q}$, we would have the simpler representation*

$$\alpha_{\mathbf{y}}(\mathbb{Q}) := -\log\left(\frac{L(\mathbb{Q}|\mathbf{y})}{L(\hat{\mathbb{Q}}|\mathbf{y})}\right).$$

Definition 6 *For fixed observations $\mathbf{y}_t = (Y_1, Y_2, ..., Y_t)$, for uncertainty aversion parameters $k > 0$ and $k' \in [1, \infty]$, we define the convex expectation*

$$\mathcal{E}_{\mathbf{y}_t}^{k,k'}(\xi) := \sup_{\mathbb{Q} \in \mathcal{Q}} \left\{\mathbb{E}_{\mathbb{Q}}[\xi|\mathbf{y}_t] - \left(\frac{1}{k}\alpha_{\mathbf{y}_t}(\mathbb{Q})\right)^{k'}\right\}, \tag{3}$$

where we adopt the convention $x^{\infty} = 0$ for $x \in [0, 1]$ and $+\infty$ otherwise[2].

We call $\mathcal{E}_{\mathbf{y}_t}^{k,k'}$ the "DR-expectation" (with parameter k, k'). We may omit to write k, k' for notational simplicity.

With deliberate ambiguity, the acronym "DR" can either stand for "divergence robust" or "data-driven robust".

Remark 5 *By construction, $\mathcal{Q}$ is parameterized by $\mathfrak{S}$ and $\{\mathfrak{D}_t\}_{t \geq 0}$, which lie in $\mathbb{R}^m$ for some m. The divergence and conditional expectations given $\mathbf{y}_t$ are continuous*

---

[2]The convention $1^{\infty} = 0$ simply ensures that x^{∞} is lower semicontinuous, and also that $x \mapsto x^{\infty}$ is the convex conjugate of $x \mapsto x^1$.

with respect to this parameterization and can be constructed to be Borel measurable with respect to $\mathbf{y}_t$. Consequently, measure theoretic concerns which arise from taking the supremum will not cause difficulty, in particular, the DR-expectation defined in (3) is guaranteed to be a Borel measurable function of $\mathbf{y}_t$ for every ξ. (This follows from Filippov's implicit function theorem, see, for example, Cohen and Elliott (2015) Appendix 10.)

Remark 6 *The choice of parameters k and k' determines much of the behaviour of the nonlinear expectation. The role of k is simple, as it acts to scale the uncertainty aversion—a higher value of k results in smaller penalties, and hence the DR-expectation will lie further above the MAP expectation. The parameter k' determines the "curvature" of the uncertainty aversion. Taking $k' = \infty$ results in the DR-expectation being positively homogeneous, that is, it is a coherent expectation in the sense of Artzner et al. (1999). In Cohen (2017), the asymptotic behaviour of the DR-expectation is studied, under the assumption of iid observations. For $k' = 1$, the DR-expectation of corresponds (for large samples) to the expected value under the maximum likelihood model plus $k/2$ times the sampling variance of the expectation, while for $k' = \infty$, the DR-expectation corresponds to the expected value under the maximum likelihood model plus $\sqrt{2k}$ times the sampling standard error of the expectation. In this paper, we will not be considering such an asymptotic result, so the values of k and k' will not play a significant role. Their presence nevertheless gives a more general class of penalty functions in Definition 4, and they are kept for notational consistency with other papers considering the DR-expectation.*

Remark 7 *In principle, we could now apply the DR-expectation framework to a filtering context as follows: Take a collection of models $\mathcal{Q}$. For a random variable ξ, and for each measure $\mathbb{Q} \in \mathcal{Q}$, compute $E_{\mathbb{Q}}[\xi|\mathbf{y}_t]$ and $\alpha_{\mathbf{y}_t}(\mathbb{Q})$. Taking a supremum as in (3), we obtain the DR-expectation. However, this is generally not computationally tractable in this form.*

Lemma 1 *Let $\{\mathcal{F}_t\}_{t \geq 0}$ be a filtration such that Y is adapted. For $\mathcal{F}_t$-measurable random variables, the choice of horizon $T \geq t$ in the definition of the penalty function α is irrelevant. In particular, for $\mathcal{F}_t$-measurable ξ, if $\mathcal{Q}\|_{\mathcal{F}_t} = \{\mathbb{Q}\|_{\mathcal{F}_t} : \mathbb{Q} \in \mathcal{Q}\}$, we know*

$$\mathcal{E}_{\mathbf{y}_t}(\xi) = \sup_{\mathbb{Q} \in \mathcal{Q}\|_{\mathcal{F}_t}} \left\{ \mathbb{E}_{\mathbb{Q}}[\xi|\mathbf{y}_t] - \left(\frac{1}{k}\alpha_{\mathbf{y}_t}(\mathbb{Q}\|_{\mathcal{Y}_t}) \right)^{k'} \right\},$$

where $\alpha_{\mathbf{y}_t}(\mathbb{Q}\|_{\mathcal{Y}_t})$ is defined as above, in terms of the restricted measure $\mathbb{Q}\|_{\mathcal{Y}_t}$.

Proof By construction, $\alpha_{\mathbf{y}_t}$ is obtained from the posterior relative density, which is determined by the restriction of $\mathbb{Q}$ to $\mathcal{Y}_t \subseteq \mathcal{F}_t$, while the expectation depends only on the restriction of $\mathbb{Q}$ to $\mathcal{F}_t$. As these are the only terms needed to compute the DR-expectation, the result follows.   □

Theorem 3 *The penalty in the DR-expectation is additive, in the sense of Definition 4.*

Proof From Proposition 2, we have the likelihood

$$L^{\mathrm{obs}}(\mathbb{Q}(A, C, p_0)|\mathbf{y}_t) = \prod_{s=1}^{t} \mathbf{1}^{\top} C_s(Y_s) A_s p_{s-1}^{(A,C),p_0},$$

where $p_s^{(A,C),p_0}$ is the solution to the filtering Eq. 1. By Lemma 1, the penalty in the DR-expectation is given by

$$\begin{aligned}
\alpha_{\mathbf{y}_t}(\mathbb{Q}\|\mathcal{Y}_t) \\
= -\log\left(L^{\mathrm{obs}}(\mathbb{Q}(A, C, p_0)|\mathbf{y}_t) \cdot e^{-\kappa_{\mathrm{prior}}(\mathfrak{S}, p_0) - \sum_t \gamma_{\mathrm{prior}}(\mathfrak{S}, \mathfrak{D}_t)}\right) - m_t \\
= \kappa_{\mathrm{prior}}(\mathfrak{S}, p_0) + \sum_{s \le t}\left[\gamma_{\mathrm{prior}}(\mathfrak{S}, \mathfrak{D}_s) - \log\left(\mathbf{1}^{\top} C_s(Y_s) A_s p_{s-1}^{(A,C),p_0}\right)\right] - m_t,
\end{aligned}$$

where m_t is chosen to ensure $\inf_{\mathbb{Q}}(\alpha_{\mathbf{y}_t}(\mathbb{Q}) \equiv 0$. As $(A_t, C_t(\cdot)) = \Phi(\mathfrak{S}, \mathfrak{D}_t)$, we obtain the desired form by setting

$$\gamma_t(\mathfrak{S}, \mathfrak{D}_t, \{Y_s\}_{s \le t}, p) = \gamma_{\mathrm{prior}}(\mathfrak{S}, \mathfrak{D}_t) - \log\left(\mathbf{1}^{\top} C_t(Y_t) A_t p\right).$$

$\square$

Remark 8 *The purpose of the nonlinear expectation is to give an "upper" estimate of a random variable, accounting for uncertainty in the underlying probabilities. This is closely related to robust estimation in the sense of Wald (1945). In particular, one can consider the robust estimator given by*

$$\arg\inf_{\hat{\xi} \in \mathbb{R}^N} \mathcal{E}_{\mathbf{y}_t}\left(\left\|\xi - \hat{\xi}\right\|^2\right),$$

which gives a "minimax" estimate of ξ, given the observations $\mathbf{y}_t$ and a quadratic loss function. The advantage of the nonlinear expectation approach is that it allows one to construct such an estimate for every random variable/loss function, giving a cost-specific quantification of uncertainty in each case.

We can also see a connection with the theory of H^∞ filtering (see, for example, Grimble and El Sayed (1990) or more recently Zhang et al. (2009) and references therein, or the more general H^∞-control theory in Başar and Bernhard (1991)). In this setting, we look for estimates which perform best in the worst-situation, where "worst" is usually defined in terms of a perturbation to the input signal or coefficients. In our setting, we focus not on the estimation problem directly, but on the "dual" problem of building an upper expectation, i.e. calculating the "worst" expectation in terms of a class of perturbations to the coefficients (our setting is general enough that perturbation to the signal can also be included through shifting the coefficients).

Remark 9 *There are also connections between our approach and what is called "risk-sensitive filtering", see, for example, James et al. (1994) and Dey and Moore (1995); or the review of Boel et al. (2002) and references therein (from an engineering perspective); or Hansen and Sargent (2007) and Hansen and Sargent (2008)*

(from an economic perspective). In their setting, one uses the nonlinear expectation defined by

$$\mathcal{E}(\xi|\mathcal{Y}_t) = -k \log \mathbb{E}_{\mathbb{P}}\big[\exp(-\xi/k)\big|\mathcal{Y}_t\big],$$

for some choice of robustness parameter $1/k > 0$. This leads to significant simplification, as dynamic consistency and recursivity is guaranteed in every filtration (see Graf (1980) and Kupper and Schachermayer (2009), and further discussion in Section 5). The corresponding penalty function is given by the conditional relative entropy,

$$\mathcal{R}_t(\mathbb{Q}) = k\mathbb{E}_{\mathbb{Q}}[\log(d\mathbb{Q}/d\mathbb{P})|\mathcal{Y}_t],$$

which is additive (Definition 4) and the one-step penalty can be calculated accordingly. In this case, the optimization defining the nonlinear expectation could also be taken over $\mathcal{M}_1$, so this approach has a claim to be including "nonparametric" uncertainty, as all measures are considered, rather than purely Markov measures or measures in a parametric family (however, the optimization can be taken over conditionally Markov measures, and one will obtain an identical result!).

The difficulty with this approach is that it does not allow for easy incorporation of knowledge of the error of estimation of the generators (A, C) in the level of robustness—the only parameter available to choose is k, which multiplies the relative entropy. A small choice of k corresponds to a small penalty, hence a very robust expectation, but this robustness is not directly linked to the estimation of the generators (A, C). Therefore, the impact of statistical estimation error remains obscure, as k is chosen largely exogenously of this error. For this reason, our approach, which directly allows for the penalty to be based on the statistical estimation of the generators, has advantages over this simpler method.

4 Recursive penalties

The DR-expectation provides us with an approach to including statistical estimation in our valuations. However, the calculations suggested by Remark 7 are generally intractable in their stated form. In this section, we shall see how the assumption that the penalty is additive (Definition 4) can be used to simplify our calculations.

Our arguments will be based on dynamic programming techniques. For the sake of precision and brevity, we here state a (forward in time) abstract "dynamic programming principle" which we can call on in later arguments.

Theorem 4 *Let U be a topological space which is equal to the countable union of compact metrizable subsets of itself. For some $m > 0$, let $g_t : \mathbb{R}^m \times U \to \mathbb{R}^m$ be a sequence of Borel measurable functions. For any sequence $\mathbf{u} = (u_1, ..., u_T)$ in U, let the sequence $Z_t^{\mathbf{u},z_0}$ be defined by the recursion*

$$Z_t^{\mathbf{u},z_0} = g_t\left(Z_{t-1}^{\mathbf{u},z_0}, u_t\right) \qquad and \qquad Z_0^{\mathbf{u},z_0} = z_0.$$

For each $u \in U$, we write $g^{-1}(\cdot, u)$ for the (set-valued) inverse of $g(\cdot, u)$.

Suppose we have a sequence of Borel measurable maps $\mathcal{A}_t : \mathbb{R} \times U \times \mathbb{R}^m \to \mathbb{R}$ such that $v \mapsto \mathcal{A}(v, u, z)$ is nondecreasing and continuous (uniformly in u, z) and

$\mathcal{A}(v, u, z) \to \infty$ as $v \to \infty$. For each $\mathbf{u}$ and z_0, we define the sequence of values at each time t by

$$V_t(\mathbf{u}, z_0) = \mathcal{A}_t \left[V_{t-1}(\mathbf{u}, z_0), u_t, Z_{t-1}^{\mathbf{u}, z_0} \right] \qquad and \qquad V_0(\mathbf{u}, z_0) = v_0(z_0).$$

Then, the minimal value,

$$V_t^*(z) := \inf_{\left\{ \mathbf{u}, z_0 : Z_t^{\mathbf{u}, z_0} = z \right\}} V_t(\mathbf{u}, z_0),$$

satisfies the recursion

$$V_t^*(z) = \inf_{u \in U} \inf_{y \in g_t^{-1}(z, u)} \left\{ \mathcal{A}_t \left[V_{t-1}^*(y), u, y \right] \right\} \qquad and \qquad V_0^*(z_0) = v_0(z_0), \quad (4)$$

(with the convention that the infimum of the empty set is $+\infty$).

Proof We proceed by induction. Clearly, the result holds at $t = 0$, as does the (at $t = 0$ empty) statement

$$V_t(z) = +\infty \quad \text{for all } z \notin \bigcup_{\mathbf{u}, z_0} \left\{ Z_t^{\mathbf{u}, z_0} \right\}. \tag{5}$$

Suppose then that (4) and (5) hold at $t = n - 1$. For every $\epsilon > 0$, there exists $(\mathbf{u}, z_0)$ such that

$$\mathcal{A}_n \left[V_{n-1}(\mathbf{u}, z_0), u_n, Z_{n-1}^{\mathbf{u}} \right] = V_n(\mathbf{u}, z_0) = \mathcal{A}_n \left[V_{n-1}(\mathbf{u}, z_0), u_n, Z_{n-1}^{\mathbf{u}} \right]$$
$$\leq \mathcal{A}_n \left[V_{n-1}^* \left(Z_{n-1}^{\mathbf{u}} \right) + \epsilon, u_n, Z_{n-1}^{\mathbf{u}} \right].$$

Taking the infimum over $\left\{ \mathbf{u}, z_0 : Z_t^{\mathbf{u}} = z \right\}$ (which can be done measurably with respect to z, given Filippov's implicit function theorem, see, for example, Cohen and Elliott (2015) Appendix 10, and sending $\epsilon \to 0$ gives

$$\inf_{\left\{ \mathbf{u}, z_0 : Z_n^{\mathbf{u}} = z \right\}} V_n(\mathbf{u}, z_0) = \inf_{\left\{ \mathbf{u}, z_0 : Z_n^{\mathbf{u}} = z \right\}} \mathcal{A}_n \left[V_{n-1}^* \left(Z_{n-1}^{\mathbf{u}} \right), u_n, Z_{n-1}^{\mathbf{u}} \right].$$

From the definition of g, we know that

$$\left\{ \mathbf{u}, z_0 : Z_n^{\mathbf{u}, z_0} = z \right\} = \left\{ \mathbf{u}, z_0 : Z_{n-1}^{\mathbf{u}, z_0} \in g_n^{-1}(z, u_n) \right\}$$

from which we derive

$$V_n^*(z) = \inf_{\left\{ \mathbf{u}, z_0 : Z_{n-1}^{\mathbf{u}, z_0} \in g_n^{-1}(z, u_n) \right\}} \mathcal{A}_n \left[V_{n-1}^* \left(Z_{n-1}^{\mathbf{u}, z_0} \right), u_n, Z_{n-1}^{\mathbf{u}, z_0} \right].$$

The right side of this equation depends on $\mathbf{u}, z_0$ only through the values of $Z_{n-1}^{\mathbf{u}, z_0}$ and u_n. In particular, considering the set of attainable y, that is, for $y \in \bigcup_{\mathbf{u}, z_0} \left\{ Z_n^{\mathbf{u}, z_0} \right\}$, we change variables to write

$$V_n^*(z) = \inf_{u_n \in U} \inf_{\substack{y \in g_n^{-1}(z, u_n) \\ y \in \bigcup_{\mathbf{u}, z_0} \left\{ Z^{\mathbf{u}, z_0} \right\}}} \mathcal{A}_n \left[V_{n-1}^*(y), u_n, y \right].$$

As the infimum on the empty set is $+\infty$, we also obtain (5) at time n, and simplify to give (4) for $t = n$, completing the inductive proof. $\qquad \square$

Corollary 1 *Suppose, instead of Z being defined by a forward recursion, for some Borel measurable function $\tilde{g}$ we had the backward recursion*

$$\tilde{g}_t \left(Z_t^{\mathbf{u}, z_0}, u_t \right) = Z_{t-1}^{\mathbf{u}, z_0}.$$

The result of Theorem 4 still holds (with effectively the same proof), where we write $\tilde{g}_t$ instead of g_t^{-1}, so the second infimum in (4) is unnecessary (as $\tilde{g}(z, u)$ is single-valued).

For practical purposes, it is critical that we refine our approach to provide a recursive construction of our nonlinear expectation. In classical filtering, one obtains a recursion for expectations $\mathbb{E}[\phi(X_t)|\mathcal{Y}_t]$, for Borel functions ϕ; one does not typically consider the expectations of general random variables. Similarly, we will consider the expectations of random variables $\phi(X_t)$.

Proposition 3 *For each t, there exists a $\mathcal{Y}_t \otimes \mathcal{B}(\mathbb{R})$-measurable function κ_t such that, for every Borel function ϕ,*

$$\begin{aligned}
\mathcal{E}_{\mathbf{y}_t}(\phi(X_t)) &:= \sup_{\mathbb{Q} \in \mathcal{Q}} \left\{ \mathbb{E}_{\mathbb{Q}}[\phi(X_t)|\mathbf{y}_t] - \mathcal{R}_t(\mathbb{Q}) \right\} \\
&= \sup_{q \in S_N^+} \left\{ \sum_i q_i \phi(e_i) - \left(\frac{1}{k} \kappa_t(\omega, q) \right)^{k'} \right\},
\end{aligned} \tag{6}$$

where S_N^+ denotes the probability simplex in $\mathbb{R}^N$, that is, $S_N^+ = \{x \in \mathbb{R}^N : \sum_i x_i = 1, \quad x_i \geq 0 \quad \forall i\}$.

Proof Fix the observations $\mathbf{y}_t$. Taking Ω_X to be the possible states of X_t, that is, the basis vectors in $\mathbb{R}^N$ with the discrete topology and corresponding Borel σ-algebra, the space of measures on Ω_X is represented by the probability simplex S_N^+. We consider the map

$$\phi \mapsto \mathcal{E}'(\phi(\omega_X)) := \mathcal{E}_{\mathbf{y}_t}(\phi(X_t))$$

as a nonlinear expectation with underlying space Ω_X. By Theorem 2, it follows that there exists a penalty function $\kappa_{\mathbf{y}_t}(q)$ such that

$$\mathcal{E}'(\phi(\omega_X)) = \sup_{q \in S_N^+} \left\{ \sum_i q_i \phi(e_i) - \left(\frac{1}{k} \kappa_{\mathbf{y}_t}(q) \right)^{k'} \right\}.$$

Taking a regular version of this penalty function (which by convex duality exists as $\mathcal{E}$ is measurable in $\mathbf{y}_t$), we can write $\kappa_t(\omega, q) = \kappa_{\mathbf{y}_t}(q)$ as desired. □

Our aim is to find a recursion for κ_t, for various choices of $\mathcal{R}$. Our constructions will depend on the following object.

Definition 7 *Recall from Theorem 1 and Assumption 1 that, given a generator $(A_t, C_t(\cdot)) = \Phi(\mathfrak{S}, \mathfrak{D}_t) \in \mathbb{A}$ at time t, our filter dynamics are described by the recursion (up to proportionality)*

$$p_t = \mathbb{G}(p_{t-1}, \mathfrak{S}, \mathfrak{D}_t, Y_t) \propto C(Y_t) A p_{t-1}.$$

We correspondingly define the (set-valued) inverse

$$\mathbb{G}^{-1}(p; \mathfrak{S}, \mathfrak{D}_t, Y_t) = \left\{ p' \in S_N^+ : \mathbb{G}\left(p', \mathfrak{S}, \mathfrak{D}_t, Y_t\right) = p \right\}.$$

For notational simplicity, we will omit the argument Y_t when this does not lead to confusion.

The set $\mathbb{G}^{-1}(p; \mathfrak{S}, \mathfrak{D}_t, Y_t)$ represents the filter states at time $t - 1$ which evolve to p at time t, assuming the generator of our process (at time t) is given by $\Phi(\mathfrak{S}, \mathfrak{D}_t)$ and we observe Y_t. This set may be empty, if no such filter states exist. As the matrix A is generally not invertible (even accounting for the restriction to S_N^+), the set $\mathbb{G}^{-1}(p; \mathfrak{S}, \mathfrak{D}_t, Y_t)$ is not generally a singleton.

4.1 Filtering with uncertainty

We now show that if we assume our penalty is additive, then the function κ appearing in (6) can be obtained in a recursive manner.

Theorem 5 *Suppose $\mathcal{R}_t$ is additive, in the sense of Definition 4. Then, a function κ satisfying (6) is given by*

$$\kappa_t(p) = \inf_{\mathfrak{S}} K_t(p, \mathfrak{S}),$$

where K_t satisfies the recursion

$$K_t(p, \mathfrak{S}) = \inf_{\mathfrak{D}_t} \left\{ \inf_{p' \in \mathbb{G}^{-1}(p, \mathfrak{S}, \mathfrak{D}_t, Y_t)} \left\{ K_{t-1}\left(p', \mathfrak{S}\right) + \gamma_t\left(\mathfrak{S}, \mathfrak{D}_t, \{Y_s\}_{s \geq 0}, p'\right) \right\} \right\} - m_t',$$

with initial value $K_0(p_0, \mathfrak{S}) = \kappa_{prior}(p_0, \mathfrak{S})$, where m' is chosen to ensure we have the normalization $\inf_{p, \mathfrak{S}} K_t(p, \mathfrak{S}) \equiv 0$.

Proof As we know that $\mathcal{R}_t$ is additive, we have the representation $\mathcal{R}_t(\mathbb{Q}) = \left(k^{-1}\alpha_t(\mathbb{Q}, \{Y_s\}_{s \leq t})\right)^{k'}$, where

$$\alpha_t(\mathbb{Q}, \{Y_s\}_{s \leq t}) = \kappa_{\text{prior}}(p_0, \mathfrak{S}) + \sum_{s \leq t} \gamma_s\left(\mathfrak{S}, \mathfrak{D}_s, \{Y_n\}_{n \leq s}, p_{s-1}^{\mathbb{Q}}\right) + m_t.$$

As $\mathbb{E}_{\mathbb{Q}}[\phi(X_t)|\mathbf{y}_t]$ depends only on the conditional law of $X_t|\mathcal{Y}_t$ under $\mathbb{Q}$, it is easy to see that (6) is satisfied when

$$\left(\frac{1}{k}\kappa_t(p)\right)^{k'} = \inf_{\{\mathbb{Q} : \mathbb{E}_{\mathbb{Q}}[X_t|\mathcal{Y}_t] = p\}} \mathcal{R}_t(\mathbb{Q}) = \inf_{\{\mathbb{Q} : \mathbb{E}_{\mathbb{Q}}[X_t|\mathcal{Y}_t] = p\}} \left(\frac{1}{k}\alpha_t(\mathbb{Q})\right)^{1/k'}, \quad (7)$$

We wish to write the minimization in (7) as a recursive control problem, to which we can apply Theorem 4. Given p_0, $\mathfrak{S}$, and $\{\mathfrak{D}_s\}_{s \leq t}$, the law of $X_t|\mathcal{Y}_t$ is given by the solution to the filtering Eq. 1. Write $Z_t = (p_t, \mathfrak{S})$, and $u_t = \mathfrak{D}_t$, so that Z is a state process defined by $Z_0 = z_0 = (p_0, \mathfrak{S})$ and the recursion (controlled by $\mathbf{u} = \{u_s\}_{s \leq t} = \{\mathfrak{D}_s\}_{s \leq t}$)

$$Z_t = \hat{\mathbb{G}}(Z_{t-1}, \mathfrak{D}_t) := \left(\mathbb{G}\left(p_{t-1}, \mathfrak{S}, \mathfrak{D}_t, Y_t\right), \mathfrak{S}\right).$$

Omitting the constant m_t from the definition of α_t, we define

$$V_t(z_0, \mathbf{u}) = \kappa_{\text{prior}}(p_0, \mathfrak{S}) + \sum_{s \leq t} \gamma_s\left(\mathfrak{S}, \mathfrak{D}_s, \{Y_n\}_{n \leq s}, p_s\right).$$

Taking $\mathcal{A}_t$ to be the operator

$$\mathcal{A}_t\left[V_{t-1}, u_t, Z_t^{\mathbf{u}, z_0}\right] = V_{t-1} + \gamma_t(\mathfrak{S}, u_t, \{Y_s\}_{s \leq t}, p_{t-1}),$$

we see that V satisfies the structure assumed in Theorem 4. Therefore, its minimal value satisfies

$$V_t^*(z) = \inf_{\left\{z_0, \mathbf{u}: Z_t^{z_0, \mathbf{u}} = z\right\}} V_t(z_0, \mathbf{u})$$

$$= \inf_{\mathfrak{D}_t}\left\{\inf_{(p', \mathfrak{S}) = z' \in \hat{\mathbb{G}}^{-1}(z)} \left\{V_{t-1}^*(z') + \gamma_t\left(\mathfrak{S}, u_t, \{Y_s\}_{s \geq 0}, p'\right)\right\}\right\}$$

with initial value $V_0(z) = \kappa_{\text{prior}}(p_0, \mathfrak{S})$. We renomalize this by setting $m_t' = \inf_z V_t^*(z)$ and $K_t(p, \mathfrak{S}) := V_t^*(z) - m_t'$, and so obtain the stated dynamics for K. By construction, we know

$$K_t(p, \mathfrak{S}) = \inf_{\{\mathfrak{D}_s\}_{s \leq t}} \left\{\alpha_t\left(\mathbb{Q}, \{Y_s\}_{s \leq t}\right) : \mathbb{E}_{\mathbb{Q}}[X_t | \mathcal{Y}_t] = p, \mathbb{Q} = \mathbb{Q}\left(p_0, \mathfrak{S}, \{\mathfrak{D}_s\}_{s \leq t}\right)\right\}.$$

It follows that (7), and hence (6), are satisfied by taking $\kappa_t(p) = \inf_{\mathfrak{S}} K_t(p, \mathfrak{S})$, as desired. $\square$

4.2 Examples

In this section, we will seek to outline a few key settings where this theory can be applied.

4.2.1 Static generators, uncertain prior (StaticUP)

We first consider the case where uncertainty is given over the prior inputs to the filter. In particular, this "prior uncertainty" is not updated given new observations, and $\mathcal{R}$ will not change through time.

Framework 1 (StaticUP) *In a StaticUP setting, the inputs to the filtering problem are the initial filter state p_0 and the generator $(A, C(\cdot))$, which we parameterize solely using the static parameter $\mathfrak{S}$. in particular, we exclude dependence on the "dynamic" parameters $\{\mathfrak{D}_t\}_{t \geq 0}$. To represent our uncertain prior, we take a penalty*

$$\mathcal{R}(\mathbb{Q}) = \left(\frac{1}{k}\alpha(\mathbb{Q})\right)^{k'} = \left(\frac{1}{k}\kappa_{\text{prior}}(p_0, \mathfrak{S})\right)^{k'}$$

for some prescribed penalty κ_{prior}.

We now apply Theorem 5 (omitting dependence on $\mathfrak{D}_t$, as we are in a purely static setting) to see that a dynamic version κ_t of the penalty function, satisfying (6), can

⌂ Springer Open

be computed as

$$\kappa_t(p) = \inf_{\mathfrak{S}} K_t(p, \mathfrak{S}),$$

where K_t satisfies the recursion

$$K_t(p, \mathfrak{S}) = \inf_{p' \in \mathbb{G}^{-1}(p, \mathfrak{S}, Y_t)} \left\{ K_{t-1}(p', \mathfrak{S}) \right\} - m'_t.$$

Assuming $\inf_{(p_0, \mathfrak{S})} \kappa_{prior}(p_0, \mathfrak{S}) = 0$, we further compute $m'_t \equiv 0$. This completely characterizes the penalty function, and hence the nonlinear expectation.

Remark 10 *Inspired by the DR-expectation, a possible choice of penalty function κ_{prior} would be the negative log-density of a prior distribution for the inputs $(p_0, \mathfrak{S})$, shifted to have minimal value zero. Alternatively, taking an empirical Bayesian perspective, κ_{prior} could be the log-likelihood from a prior calibration process. In this case, we are incorporating our prior statistical uncertainty regarding the parameters in the filtering problem.*

Remark 11 *We emphasize that there is no learning of the generator being done in this framework—the penalty applied at time $t = 0$ is simply propagated forward; our observations do not affect our opinion of the plausible generators. In particular, if we assume no knowledge of the initial state (i.e., a zero penalty), then we will have no knowledge of the state at time t (unless the observations cause the filter to degenerate).*

Example 1 *For a concrete example of the StaticUP framework, we take the class of models in $\mathcal{M}_M$ where A and C are perfectly known and $A = I$, so $X_t = X_0$ is constant (but X_0 is unknown). We take $N = 2$, so X takes only one of two values. For the observation distribution C, we assume that*

$$Y_t | (X_t = e_1) \sim Bernoulli(a), \qquad Y_t | (X_t = e_2) \sim Bernoulli(b),$$

where $a, b \in (0, 1)$ are fixed constants. Effectively, in this example we are using filtering to determine which of two possible parameters is the correct mean for our observation sequence. It is worth emphasising that the filter process p corresponds to the posterior probabilities, in a Bayesian setting, of the events that our Bernoulli process has parameter a or b.

It is useful to note that, from classical Bayesian statistical calculations[3], for a given p_0, one can see that the corresponding value of p_t is determined from the log-odds ratio,

$$\log\left(\frac{p_t^1}{p_t^2}\right) = \log\left(\frac{p_0^1}{p_0^2}\right) + t\bar{Y}_t \log\left(\frac{a}{b}\right) + t(1 - \bar{Y}_t) \log\left(\frac{1-a}{1-b}\right),$$

where $\bar{Y}_t = t^{-1} \sum_{s \leq t} Y_s$ is the average number of successes observed at time t.

---

[3]One can derive the stated formula using the filtering equations, for the vector $p_t = \left(p_t^1, p_t^2\right)^\top$. However, the closed-form solution given here is more easily obtained using alternative methods for Bayesian hypothesis testing (which is effectively what this problem encodes).

To write down the StaticUP penalty function, let the (known) dynamics be described by $\mathfrak{S}^*$. Consequently, we can write $K(p, \mathfrak{S}) = \infty$ for all $\mathfrak{S} \neq \mathfrak{S}^*$. We initialize with a known penalty $\kappa_{prior}(p, \mathfrak{S}^*) = \kappa_0(p)$ for all $p \in S_N^+$. As $\mathfrak{S}^*$ is known, there is no distinction between K and κ, that is,

$$\kappa_t(p) = \inf_{\mathfrak{S}} K_t(p, \mathfrak{S}) = K_t(p, \mathfrak{S}^*) = \inf_{\left\{ p_0 : \mathbb{E}_{\mathbb{Q}(\mathfrak{S}^*, p_0)}[X_t | \mathcal{Y}_t] = p \right\}} \{ \kappa_0(p_0) \}.$$

In this example, we can express our penalty in terms of the log-odds, for the sake of notational simplicity given the closed-form solution to the filtering problem, and hence can explicitly calculate the (unique) initial distribution p_0 which would evolve to a given p at time t. In particular, the time-t penalty is given by a shift of the initial penalty:

$$\kappa_t \left(\log \left(\frac{p_t^1}{p_t^2} \right) \right) = \kappa_0 \left(\log \left(\frac{p_t^1}{p_t^2} \right) - t \bar{Y}_t \log \left(\frac{a}{b} \right) - t(1 - \bar{Y}_t) \log \left(\frac{1-a}{1-b} \right) \right).$$

Remark 12 *This example demonstrates the following behaviour:*

- *If the initial penalty is zero, then the penalty at time t is also zero—there is no learning of which state we are in.*
- *When parameterized by the log-odds ratio, there is no variation in the curvature of the penalty (and so no change in our "uncertainty"), we simply shift the penalty around, corresponding to our changing posterior probabilities.*
- *The update of κ is done purely using the tools of Bayesian statistics, rather than having any direct incorporation of our uncertainty.*

Remark 13 *We point out that this is, effectively, the model of uncertainty proposed by Walley (1991) (see, in particular, Walley (1991) section 5.3, although there he takes a model where the unknown parameter is Beta distributed). See also Fagin and Halpern (1990).*

4.2.2 Dynamic generators, uncertain prior (DynamicUP)

If we model the generator (A, C) as fixed and unknown (i.e., it depends only on $\mathfrak{S}$), calculation of $K_t(p, \mathfrak{S})$ suffers from a curse of dimensionality—the dimension of $\mathfrak{S}$ determines the size of the domain of K_t. On the other hand, if we suppose the generator at time t depends only on the dynamic parameters $\mathfrak{D}_t$, we can use dynamic programming to obtain a lower-dimensional problem.

Framework 2 (DynamicUP) *In the DynamicUP setting, for an initial penalty on the initial hidden state, $\kappa_{prior}(p_0)$, and a penalty on the time-t generator, $\gamma_t(\mathfrak{D}_t)$, our total penalty is given by $\mathcal{R}(\mathbb{Q}) = \left(\frac{1}{k} \alpha(\mathbb{Q}) \right)^{k'}$, where we now have*

$$\alpha(\mathbb{Q}) = \kappa_{\text{prior}}(p_0) + \sum_{s=1}^{\infty} \gamma_s(\mathfrak{D}_s).$$

In this case, as we ignore the static parameter $\mathfrak{S}$, we simplify Theorem 5 through the identity $\kappa_t(p) = K_t(p, \mathfrak{S})$. This yields the recursion

$$\kappa_t(p) = \inf_{\mathfrak{D}_t} \left\{ \inf_{p' \in \mathbb{G}^{-1}(p, \mathfrak{D}_t, Y_t)} \left\{ \kappa_{t-1}(p') + \gamma_t(\mathfrak{D}_t) \right\} \right\} - m'_t$$

and again, if we assume $\inf_{\mathbb{Q}} \alpha(\mathbb{Q}) \equiv 0$, we then conclude $m'_t \equiv 0$.

This formulation of the uncertain filter allows us to use dynamic programming to solve our problem forward in time. In the setting of Example 1, as the generator is perfectly known, there is no distinction between the dynamic and static cases.

A continuous-time version of this setting (for a Kalman–Bucy filter) is considered in detail in Allan and Cohen (2019a).

4.2.3 Static generators, DR-expectation (StaticDR)

In the above examples, we have regarded the prior as uncertain and used this to penalize over models. We did not use the data to modify our penalty function $\mathcal{R}$. The DR-expectation gives us an alternative approach in which the data guides our model choice more directly. In what follows, we apply the DR-expectation in our filtering context and observe that it gives a slightly different recursion for the penalty function. Again, we can consider models where our generator is constant (i.e., depends only on $\mathfrak{S}$) or changes dynamically (i.e., depends only on $\mathfrak{D}_t$).

Framework 3 (StaticDR) *As in the StaticUP framework, we assume that the generator (A, C) is determined by the static parameter $\mathfrak{S}$. For each $\mathfrak{S}$, with $\mathbb{Q} = \mathbb{Q}(A, C, p_0)$ and $(A, C) = \Phi(\mathfrak{S})$, we have a penalty $\mathcal{R}(\mathbb{Q}) = \left(\frac{1}{k} \alpha(\mathbb{Q}) \right)^{k'}$ given by the log-posterior density*

$$\alpha(\mathbb{Q} \| \mathcal{Y}_t) = \kappa_{\text{prior}}(p_0, \mathfrak{S}) - L^{obs}\left(\mathbb{Q}(A, C, p_0) | \mathbf{y}_t \right) + m_t$$

which is additive, as shown in Theorem 3. Applying Theorem 5, we see that the penalty can be written $\kappa_t(p) = \inf_{\mathfrak{S}} K_t(p, \mathfrak{S})$, where $K_0(p, \mathfrak{S}) = \kappa_{prior}(p, \mathfrak{S})$ and K satisfies the recursion

$$K_t(p, \mathfrak{S}) = \inf_{\mathfrak{D}_t} \left\{ \inf_{p' \in \mathbb{G}^{-1}(p, \mathfrak{S}, Y_t)} \left\{ K_{t-1}(p', \mathfrak{S}) + \log c^{\mathfrak{S}}\left(Y_s; A^{\mathfrak{S}} p' \right) \right\} \right\} - m'_t.$$

Unlike in the uncertain prior cases, we cannot typically claim that $m'_t \equiv 0$, instead it is a random process dependent on our observations.

Remark 14 *Comparing Framework 1 (StaticUP) with Framework 3 (StaticDR), we see that the key distinction is the presence of the log-likelihood term. This term implies that observations of Y will affect our quantification of uncertainty, rather than purely updating each model.*

Example 2 *In the setting of Example 1, recall that X is constant, so we know $(A, C(\cdot))$. One can calculate the StaticDR penalty either directly, or through solving*

the stated recursion using the dynamics of p. As in the StaticUP case, the result is most simply expressed by first calculating p_0 from p_t through

$$\log\left(\frac{p_0^1}{p_0^2}\right) = \log\left(\frac{p_t^1}{p_t^2}\right) - t\bar{Y}_t \log\left(\frac{a}{b}\right) - t(1-\bar{Y})_t \log\left(\frac{1-a}{1-b}\right)$$

and then

$$\kappa_t(p_t) = \kappa_0(p_0) - \log\left(p_0^1 a^{t\bar{Y}_t}(1-a)^{t(1-\bar{Y}_t)} + p_0^2 b^{t\bar{Y}_t}(1-b)^{t(1-\bar{Y}_t)}\right) - m_t,$$

where m_t is chosen to ensure $\inf_p \kappa_t(p) = 0$. From this, we see that the likelihood modifies our uncertainty directly, rather than us simply propagating each model via Bayes' rule. A consequence of this is that if we start with extreme uncertainty ($\kappa_0 \equiv 0$), then our observations teach us what models are reasonable, thereby reducing our uncertainty (i.e., we will find $\kappa_t(p) > 0$ for $p \in (0, 1)$ when $t > 0$).

Remark 15 *It is interesting to ask what the long-term behaviour of these uncertain filters will be. In Cohen (2017), the long run behaviour of the DR-expectation based on i.i.d. observations is derived and, in principle, a similar analysis is possible here. Using the asymptotic analysis of maximum likelihood estimation for hidden Markov models in Leroux (1992) or Douc et al. (2011), we know that the MLE will converge with probability one to the true parameter, under appropriate regularity conditions. Here, the presence of the prior influences this slightly, however, this impact vanishes as $t \to \infty$. With further regularity assumptions, one can also show that the log-likelihood function, divided by the number of observations, almost surely converges to the relative entropy between a proposed model and the true model (see, for example, Leroux (1992) section 5). If one also knew that the relative entropy is smooth and convex, the analysis of Cohen (2017) Theorems 4 and 5 is possible, showing that the DR-expectation corresponds to adding a term related to the sampling variance of the hidden state[4]. In particular, as the number of observations increases, the DR-expectation will converge to the expected value under the filter with the true parameters.*

4.2.4 Dynamic generators, DR-expectation (DynamicDR)

As in the uncertain prior case, it is often impractical to calculate a recursion for $K(p, \mathfrak{S})$ given the high dimension of $\mathfrak{S}$. We therefore consider the case when (A, C) depends only on the dynamic parameters $\mathfrak{D}_t$.

Framework 4 (DynamicDR) *As before, for each $\{\mathfrak{D}_s\}_{s\geq 0}$, with $\mathbb{Q} = \mathbb{Q}(A, C, p_0)$ and $(A_t, C_t) = \Phi(\mathfrak{D}_t)$, we have a penalty $\mathcal{R}(\mathbb{Q}) = \left(\frac{1}{k}\alpha(\mathbb{Q})\right)^{k'}$ given by the log-posterior density*

$$\alpha(\mathbb{Q}\|\mathcal{Y}_t) = \kappa_{\text{prior}}(p_0) - \log L^{obs}\left(\mathbb{Q}(A, C, p_0)\big|\mathbf{y}_t\right) + m_t.$$

---

[4]One difficulty is that the analysis of Cohen (2017) considers the divergence to the MLE model, rather than the true model. This allows a slightly finer control over the asymptotic behaviour, which would need to be replicated in the filtering setting.

From Theorem 3, we know that the log-posterior density is additive. Applying Theorem 5, and the identity $\kappa_t(p) = K_t(p, \mathfrak{S})$, we conclude that the penalty $\kappa_t(p)$ in (6) can be computed from the recursion

$$\kappa_t(p) = \inf_{\mathfrak{D}_t} \left\{ \inf_{p' \in \mathbb{G}^{-1}(p, \mathfrak{D}_t, Y_t)} \left\{ \kappa_{t-1}(p') + \gamma_{\text{prior}}(t, \mathfrak{D}_t; \{Y_s\}_{s<t}) \right. \right.$$
$$\left. \left. - \log \left(c^{\mathfrak{D}_t}(Y_t; A_t^{\mathfrak{D}} p_{t-1}) \right) \right\} \right\} - m'_t,$$

with initial value $\kappa_0(p) = \pi(p)$, where m'_t is chosen to ensure $\inf_{p \in S_N^+} \kappa_t(p) = 0$ for all t.

Remark 16 *We expect that there will be less difference between the dynamic uncertain prior and dynamic DR-expectation settings than between the static uncertain prior and static DR-expectation settings. This is because there is only limited learning possible in the dynamic DR-expectation, as $\mathfrak{D}_t$ may vary independently at every time, so the DR-expectation has only one value with which to infer the value of each $\mathfrak{D}_t$. This increases the relative importance of the prior term γ_{prior}, which describes our understanding of typical values of the generator. In practice, the key distinction between the dynamic DR-expectation and uncertain prior models appears to be when the initial penalty is near zero—in this case, the DR-expectation regularizes the initial state quickly, while the uncertain prior model may remain near zero indefinitely.*

Example 3 *In the setting of Example 2, as the dynamics are perfectly known, there is again no difference between the dynamic and static generator DR-expectation cases.*

A continuous-time version of this setting (for a Kalman–Bucy filter) is considered in Allan and Cohen (2020).

5 Expectations of the future

The nonlinear expectations considered above do not consider how the future will evolve. In particular, we have focussed our attention on calculating $\mathcal{E}_{\mathbf{y}_t}(\phi(X_t))$, that is, on the expectation of functions of the *current* hidden state. In other words, we can consider our nonlinear expectation as a mapping

$$\mathcal{E}_{\mathbf{y}_t} : L^\infty(\sigma(X_t) \otimes \mathcal{Y}_t) \to L^\infty(\mathcal{Y}_t).$$

If we wish to calculate expectations of *future* states, then we may wish to consider doing so in a filtration-consistent manner. This is of particular importance when considering optimal control problems.

Definition 8 *For a fixed horizon $T > 0$, suppose that for each $t < T$ we have a mapping $\mathcal{E}(\cdot|\mathcal{Y}_t) : L^\infty(\mathcal{Y}_T) \to L^\infty(\mathcal{Y}_t)$. We say that $\mathcal{E}$ is a $\mathcal{Y}$-consistent convex expectation if $\mathcal{E}(\cdot|\mathcal{Y}_t)$ satisifes the following assumptions, analogous to those above,*

- *Strict Monotonicity: for any $\xi_1, \xi_2 \in L^\infty(\mathcal{Y}_T)$, if $\xi_1 \geq \xi_2$ a.s., then $\mathcal{E}(\xi_1|\mathcal{Y}_T) \geq \mathcal{E}(\xi_2|\mathcal{Y}_t)$ a.s., and if, in addition, $\mathcal{E}(\xi_1|\mathcal{Y}_t) = \mathcal{E}(\xi_2|\mathcal{Y}_t)$ then $\xi_1 = \xi_2$ a.s.;*
- *Constant triviality: for $b \in L^\infty(\mathcal{Y}_t)$, $\mathcal{E}(b|\mathcal{Y}_t) = b$;*
- *Translation equivariance: for any $b \in L^\infty(\mathcal{Y}_t)$, $\xi \in L^\infty(\mathcal{Y}_T)$, $\mathcal{E}(\xi + b|\mathcal{Y}_t) = \mathcal{E}(\xi|\mathcal{Y}_t) + b$;*
- *Convexity: for any $\lambda \in [0, 1]$, $\xi_1, \xi_2 \in L^\infty(\mathcal{Y}_T)$,*

$$\mathcal{E}(\lambda\xi_1 + (1 - \lambda)\xi_2|\mathcal{Y}_t) \leq \lambda\mathcal{E}(\xi_1|\mathcal{Y}_t) + (1 - \lambda)\mathcal{E}(\xi_2|\mathcal{Y}_t);$$

- *Lower semicontinuity: for a sequence $\{\xi_n\}_{n\in\mathbb{N}} \subset L^\infty(\mathcal{Y}_T)$ with $\xi_n \uparrow \xi \in L^\infty(\mathcal{Y}_T)$ pointwise, $\mathcal{E}(\xi_n|\mathcal{Y}_t) \uparrow \mathcal{E}(\xi|\mathcal{Y}_t)$ pointwise for every $t < T$;*

and the additional asssumptions

- *$\{\mathcal{Y}_t\}_{t\geq 0}$-consistency: for any $s < t < T$, any $\xi \in L^\infty(\mathcal{Y}_T)$,*

$$\mathcal{E}(\xi|\mathcal{Y}_s) = \mathcal{E}(\mathcal{E}(\xi|\mathcal{Y}_t)|\mathcal{Y}_s);$$

- *Relevance: for any $t < T$, any $A \in \mathcal{Y}_t$, $\mathcal{E}(I_A\xi|\mathcal{Y}_t) = I_A\mathcal{E}(\xi|\mathcal{Y}_t)$.*

The assumption of $\mathcal{Y}$-consistency is sometimes simply called recursivity, time consistency, or dynamic consistency (and is closely related to the validity of the dynamic programming principle), however, it is important to note that this depends on the choice of filtration. In our context, consistency with the *observation* filtration $\mathcal{Y}$ is natural, as this describes the information available for us to make decisions.

Remark 17 *Definition 8 is equivalent to considering a lower semicontinuous convex expectation, as in Definition 3, and assuming that for any $\xi \in L^\infty(\mathcal{Y}_T)$ and any $t < T$, there exists a random variable $\xi_t \in L^\infty(\mathcal{Y}_t)$ such that $\mathcal{E}(I_A\xi) = \mathcal{E}(I_A\xi_t)$ for all $A \in \mathcal{Y}_t$. In this case, one can define $\mathcal{E}(\xi|\mathcal{Y}_t) = \xi_t$ and verify that the definition given is satisfied (see Föllmer and Schied (2002b) and Cohen and Elliott (2010)).*

Much work has been done on the construction of dynamic nonlinear expectations (see, for example, Epstein and Schneider (2003), Duffie and Epstein (1992), El Karoui et al. (1997), and Cohen and Elliott (2010); and references therein). In particular, there have been close relations drawn between these operators and the theory of BSDEs (for a setting covering the discrete-time examples we consider here, see Cohen and Elliott (2010) and Cohen and Elliott (2011)).

Remark 18 *The importance of $\mathcal{Y}$-consistency is twofold: First, it guarantees that, when using a nonlinear expectation to construct the value function for a control problem, an optimal policy will be consistent in the sense that (assuming an optimal policy exists) a policy which is optimal at time zero will remain optimal in the future. Second, $\{\mathcal{Y}_t\}_{t\geq 0}$-consistency allows the nonlinear expectation to be calculated recursively, working backwards from a terminal time. This leads to a considerable simplification numerically, as it avoids a curse of dimensionality in intertemporal control problems.*

Remark 19 *One issue in our setting is that our lack of knowledge does not simply line up with the arrow of time—we are unaware of events which occurred in the past, as well as those which are in the future. This leads to delicacies in questions of dynamic consistency. Conventionally, this has often been considered in a setting of "partially observed control", and these issues are resolved by taking the filter state p_t to play the role of a state variable, and solving the corresponding "fully observed control problem" with p_t as underlying. In our context, we do not know the value of p_t, instead we have the (even higher dimensional) penalty function K_t as a state variable.*

In the following sections, we will outline how our earlier approach can be extended to provide a dynamically consistent expectation, and how enforcing dynamic consistency will modify our perception of risk.

5.1 Asynchronous expectations

We will focus our attention on constructing a dynamically consistent nonlinear expectation for random variables in $L^\infty(\sigma(X_T) \otimes \mathcal{Y}_T)$, given observations up to times $t < T$. Throughout this section, we will use the following construction:

Definition 9 *Suppose we have a nonlinear expectation*

$$\mathcal{E}_{\mathbf{y}_T} : L^\infty(\sigma(X_T) \otimes \mathcal{Y}_T) \to L^\infty(\mathcal{Y}_T)$$

constructed for our nonlinear filtering problem, as above, and are given a a $\mathcal{Y}$-consistent family of maps

$$\bar{\mathcal{E}}(\cdot|\mathcal{Y}_t) : L^\infty(\mathcal{Y}_T) \to L^\infty(\mathcal{Y}_t).$$

We then extend $\bar{\mathcal{E}}$ to variables in $L^\infty(\sigma(X_T) \otimes \mathcal{Y}_T)$ by the composition

$$\bar{\mathcal{E}}(\cdot|\mathcal{Y}_t) := \bar{\mathcal{E}}(\mathcal{E}_{\mathbf{y}_T}(\cdot)|\mathcal{Y}_t).$$

Given this definition, our key aim is to construct the $\mathcal{Y}$-consistent family $\bar{\mathcal{E}}(\cdot|\mathcal{Y}_t)$, in a way which "agrees" with our uncertainty in the underlying filter. As we are in discrete time, we can construct a $\mathcal{Y}$-consistent family through recursion, if we have its definition over each single step. The definition of the DR-expectation can be applied to generate these one-step expectations in a natural way.

Definition 10 *For $\mathcal{R}$ an additive penalty function (Definition 4), we define the one-step expectation, for $\xi \in L^\infty(\mathcal{Y}_{t+1})$, by*

$$\bar{\mathcal{E}}(\xi|\mathcal{Y}_t) = \operatorname{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \left\{ \mathbb{E}_{\mathbb{Q}} \left[\xi - \mathcal{R}_{t+1}(\mathbb{Q})|\mathcal{Y}_t \right] \right\},$$

where the essential supremum is taken among the bounded $\mathcal{Y}_t$-measurable random variables. Using this, we define a $\mathcal{Y}$-consistent expectation $L^\infty(\sigma(X_T) \otimes \mathcal{Y}_T) \to L^\infty(\mathcal{Y}_t)$ by recursion,

$$\bar{\mathcal{E}}(\xi|\mathcal{Y}_t) = \bar{\mathcal{E}}(\bar{\mathcal{E}}(\xi|\mathcal{Y}_{t+1})|\mathcal{Y}_t) = \bar{\mathcal{E}}(\cdots \bar{\mathcal{E}}(\mathcal{E}_{\mathbf{y}_T}(\xi)|\mathcal{Y}_{T-1})\cdots|\mathcal{Y}_t).$$

Remark 20 *It is necessary to use the penalty $\mathcal{R}_{t+1}$ in this definition, as our penalty should include the behaviour of the generator $C_{t+1}(\cdot)$, which determines the distribution of $Y_{t+1}|X_{t+1}$.*

Recall that, as $\mathcal{Y}$ is generated by Y, the Doob–Dynkin lemma states that any $\mathcal{Y}_{t+1}$-measurable function ξ is simply a function of $\{Y_s\}_{s\leq t+1}$, so we can write

$$\xi(\omega) = \hat{\xi}(Y_{t+1}, \{Y_s\}_{s\leq t}). \tag{8}$$

For any conditionally Markov measure $\mathbb{Q}$, if $\mathbb{Q}$ has generator $(A_t, C_t(\cdot))_{t\geq 0}$, it follows that

$$\mathbb{E}_{\mathbb{Q}}[\xi|\mathcal{Y}_t] = \int_{\mathbb{R}^d} \hat{\xi}(y, \{Y_s\}_{s\leq t}) \left(\mathbf{1}^\top C_{t+1}(y) A_t p_t\right) d\mu(y).$$

In particular, we apply this to our penalty function to define the function $\hat{\mathcal{R}}$ such that

$$\hat{\mathcal{R}}_{t+1}(Y_{t+1}, p, \mathfrak{S}, \mathfrak{D}_{t+1}, \{Y_s\}_{s\leq t}) = \left(\frac{1}{k}\left(K_t(p_t, \mathfrak{S}) + \gamma_{t+1}\left(\mathfrak{S}, \mathfrak{D}_{t+1}, \{Y_s\}_{s\leq t+1}, p_t\right)\right)\right)^{k'}. \tag{9}$$

Applying this to our definition of $\bar{\mathcal{E}}$, we obtain the following representation.

Lemma 2 *The one-step expectation $\bar{\mathcal{E}}$ can be written*

$$\bar{\mathcal{E}}(\xi|\mathcal{Y}_t) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \left\{\mathbb{E}_{\mathbb{Q}}[\xi - \mathcal{R}_{t+1}(\mathbb{Q})|\mathcal{Y}_t]\right\}$$

$$= \operatorname*{ess\,sup}_{\mathfrak{S}, \mathfrak{D}_{t+1}, p} \left\{\int_{\mathbb{R}^d} \left(\hat{\xi}\left(y, \{Y_s\}_{s\leq t}\right) - \hat{\mathcal{R}}_{t+1}\left(y, p, \mathfrak{S}, \mathfrak{D}_{t+1}, \{Y_s\}_{s\leq t}\right)\right)\right.$$

$$\left.\left(\mathbf{1}^\top C_{t+1}(y) A_{t+1} p\right) d\mu(y)\right\},$$

where K is the dynamic penalty constructed in Theorem 5 and $(A_{t+1}, C_{t+1}(\cdot)) \equiv \Phi(\mathfrak{S}, \mathfrak{D}_{t+1})$.

Proof Write

$$\bar{\mathcal{E}}(\xi|\mathcal{Y}_t) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \left\{\mathbb{E}_{\mathbb{Q}}[\xi|\mathcal{Y}_t] - \mathbb{E}_{\mathbb{Q}}[\mathcal{R}_{t+1}(\mathbb{Q})|\mathcal{Y}_t]\right\}.$$

We know that

$$\mathbb{E}_{\mathbb{Q}}[\xi|\mathcal{Y}_t] = \int_{\mathbb{R}^d} \hat{\xi}(y, \{Y_s\}_{s\leq t}) \left(\mathbf{1}^\top C_{t+1}(y) A_{t+1} p_t\right) d\mu(y),$$

which depends on $\mathbb{Q}$ only through A_{t+1}, C_{t+1} and p_t, or equivalently, through the parameters $\mathfrak{S}$, $\mathfrak{D}_{t+1}$ and p_t. In particular, as $\mathcal{R}$ is additive, we can substitute in its structure and simplify using the definition of K in Theorem 5 to obtain

$$\bar{\mathcal{E}}(\xi|\mathcal{Y}_t) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \left\{\mathbb{E}_{\mathbb{Q}}[\xi|\mathcal{Y}_t]\right.$$

$$\left. - \mathbb{E}_{\mathbb{Q}}\left[\left(\frac{1}{k}\left(K(p_t, \mathfrak{S}) + \gamma_{t+1}(\mathfrak{S}, \mathfrak{D}_{t+1}, \{Y_s\}_{s\leq t+1}, p_t)\right)\right)^{k'} \middle| \mathcal{Y}_t\right]\right\}.$$

Using the definition of $\hat{\mathcal{R}}$, we change these conditional expectations to integrals, and obtain the desired representation.   □

Remark 21 *There is a surprising form of double-counting of the penalty here. To see this, let's assume ϕ does not depend on Y. If we consider $\xi_{t+1} = \mathcal{E}_{\mathbf{y}_{t+1}}(\phi(X_{t+1}))$, then we have included a penalty for the proposed model at $t + 1$, that is,*

$$\xi_{t+1} = \mathcal{E}_{\mathbf{y}_{t+1}}(\phi(X_{t+1})) = \sup_{p \in S_N^+} \left\{ \sum_i p^i \phi(e_i) - \left(\frac{1}{k}(K_{t+1}(p)) \right)^{k'} \right\},$$

where $K_{t+1}(p)$ is the penalty associated with the filter state at time $t + 1$, which includes the penalty γ_{t+1} on the parameters $\mathfrak{S}$ and $\mathfrak{D}_{t+1}$.

When we calculate $\bar{\mathcal{E}}(\xi_{t+1}|\mathcal{Y}_t)$, we do so by using the penalty $K(p_t, \mathfrak{S}) + \gamma_{t+1}(\mathfrak{S}, \mathfrak{D}_{t+1}, \{Y_s\}_{s \leq t+1}, p_t)$, which again includes the term γ_{t+1} which penalizes unreasonable values of the parameters $\mathfrak{S}$ and $\mathfrak{D}_{t+1}$. This "double counting" of the penalty corresponds to us including both our "uncertainty at time $t + 1$" (in $\mathcal{E}_{\mathbf{y}_{t+1}}$), and also our "uncertainty at t about our uncertainty at $t + 1$" (in $\bar{\mathcal{E}}(\cdot|\mathcal{Y}_t)$).

Remark 22 *One should be careful in this setting, as the recursively-defined non-linear expectation will be optimized for a different value of $\mathfrak{S}$ at every time. As $\mathfrak{S}$ is considered to be a static penalty, this is an internal inconsistency in the modelling of our uncertainty—we always estimate assuming that $\mathfrak{S}$ has never changed, but evaluate the future by considering our possible future opinions of the value of $\mathfrak{S}$.*

5.2 Review of BSDE theory

While it is useful to give a recursive definition of our nonlinear expectation, a better understanding of its dynamics is of practical importance. In what follows, for the dynamic generator case, we consider the corresponding BSDE theory, assuming that Y_t can take only finitely many values, as in Cohen and Elliott (2010). We now present the key results of Cohen and Elliott (2010), in a simplified setting.

In what follows, we suppose that Y takes d values, which we associate with the standard basis vectors in $\mathbb{R}^d$. For simplicity, we write $\mathbf{1}$ for the vector in $\mathbb{R}^d$ with all components 1.

Definition 11 *Write $\bar{\mathbb{P}}$ for a probability measure such that $\{Y_t\}_{t \geq 0}$ is an i.i.d. sequence, uniformly distributed over the d states, and M for the $\bar{\mathbb{P}}$-martingale differ-ence process $Y_t - d^{-1}\mathbf{1}$. As in Cohen and Elliott (2010), M has the property that any $\mathcal{Y}$-adapted $\bar{\mathbb{P}}$-martingale L can be represented by $L_t = L_0 + \sum_{0 \leq s < t} Z_s M_{s+1}$ for some Z (and Z is unique up to addition of a multiple of $\mathbf{1}$).*

Remark 23 *The construction of Z in fact also shows that, if L is written $L_t = \tilde{L}(Y_1, ..., Y_{t-1}, Y_t)$, then $e_i^\top Z_t = L(Y_1, ..., Y_{t-1}, e_i)$ for every i (up to addition of a multiple of $\mathbf{1}$).*

We can then define a BSDE (Backward Stochastic Difference Equation) with solution (ξ, Z):

$$\xi_t(\omega) - \sum_{t \leq u < T} f(\omega, u, \xi_u(\omega), Z_u(\omega)) + \sum_{t \leq u < T} Z_u(\omega) M_{u+1}(\omega) = \xi_T(\omega), \quad (10)$$

where T is a finite deterministic terminal time, f a $\mathcal{Y}$-adapted map $F : \Omega \times \{0, ..., T\} \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, and ξ_T a given $\mathbb{R}$-valued $\mathcal{Y}_T$-measurable terminal condition. For simplicity, we henceforth omit the ω argument of ξ, Z, and M.

The general existence and uniqueness result for BSDEs in this context is as follows:

Theorem 6 *Suppose f is such that the following two assumptions hold:*

(i) *For any ξ, if $Z^1 = Z^2 + k\mathbf{1}$ for some k, then $f\left(\omega, t, \xi_t, Z_t^1\right) = f\left(\omega, t, \xi_t, Z_t^2\right)$, $\bar{\mathbb{P}}$-a.s. for all t.*
(ii) *For any $z \in \mathbb{R}^d$, for all t, for $\bar{\mathbb{P}}$-almost all ω, the map*

$$\xi \mapsto \xi - f(\omega, t, \xi, z)$$

is a bijection $\mathbb{R} \to \mathbb{R}$.

Then, for any terminal condition ξ_T essentially bounded, $\mathcal{Y}_T$-measurable, and with values in $\mathbb{R}$, the BSDE (10) has a $\mathcal{Y}$-adapted solution (ξ, Z). Moreover, this solution is unique up to indistinguishability for ξ and indistinguishability up to addition of multiples of $\mathbf{1}$ for Z.

In this setting, we also have a comparison theorem:

Theorem 7 *Consider two discrete-time BSDEs as in (10), corresponding to coefficients f^1, f^2, and terminal values ξ_T^1, ξ_T^2. Suppose the conditions of Theorem 6 are satisfied for both equations, let $\left(\xi^1, Z^1\right)$ and $\left(\xi^2, Z^2\right)$ be the associated solutions. Suppose the following conditions hold:*

(i) $\xi_T^1 \geq \xi_T^2$ $\bar{\mathbb{P}}$-a.s.
(ii) $\bar{\mathbb{P}}$-a.s., for all times t and every $\xi \in \mathbb{R}$ and $z \in \mathbb{R}^d$,

$$f^1(\omega, t, \xi, z) \geq f^2(\omega, t, \xi, z).$$

(iii) $\bar{\mathbb{P}}$-a.s., for all t, f^1 satisfies

$$f^1\left(\omega, t, \xi_t^2, Z_t^1\right) - f^1\left(\omega, t, \xi_t^2, Z_t^2\right) \geq \min_{j \in \mathbb{J}_t}\left\{\left(Z_t^1 - Z_t^2\right)\left(e_j - d^{-1}\mathbf{1}\right)\right\},$$

where $\mathbb{J}_t := \{i : \bar{\mathbb{P}}(X_{t+1} = e_i | \mathcal{F}_t) > 0\}$.
(iv) $\bar{\mathbb{P}}$-a.s., for all t and all $z \in \mathbb{R}^d$, $\xi \mapsto \xi - f^1(\omega, t, \xi, z)$ is an increasing function.

It is then true that $\xi^1 \geq \xi^2$ $\bar{\mathbb{P}}$-a.s. A driver f^1 satisfying (iii) and (iv) will be called "balanced".

Finally, we also know that all dynamically consistent nonlinear expectations can be represented through BSDEs:

Theorem 8 *The following two statements are equivalent.*

(i) $\bar{\mathcal{E}}(\cdot|\mathcal{Y}_t)$ *is a $\mathcal{Y}_t$-consistent, dynamically translation invariant, nonlinear expectation.*

(ii) *There exists a driver f which is balanced, independent of ξ, and satisfies the normalisation condition $f(\omega, t, \xi, 0) = 0$, such that, for all ξ_T, the value of $\xi_t = \bar{\mathcal{E}}(\xi_T|\mathcal{Y}_t)$ is the solution to a BSDE with terminal condition ξ_T and driver f.*

Furthermore, these two statements are related by the equation

$$f(\omega, t, \xi, z) = \bar{\mathcal{E}}(zM_{t+1}|\mathcal{Y}_t).$$

5.3 BSDEs for future expectations

By applying the above general theory, we can easily see that our nonlinear expectation has a representation as the solution to a particular BSDE.

Theorem 9 *Write*

$$\xi_t := \bar{\mathcal{E}}\left(\phi\left(X_T, \{Y_t\}_{t \le T}\right) \middle| \mathcal{Y}_t\right).$$

The dynamically consistent expectation satisfies the BSDE

$$\xi_{t+1} = \xi_t - f(Z_t; \kappa_t) + Z_t M_{t+1}$$

with driver

$$
\begin{aligned}
&f\left(Z_t; \hat{\mathcal{R}}_{t+1}\right) \\
&:= \operatorname*{ess\,sup}_{p, \mathfrak{S}, \mathfrak{D}_{t+1}} \sum_i \left\{ \left(Z^i - \hat{\mathcal{R}}_{t+1}(e_i, p, \mathfrak{S}, \mathfrak{D}_{t+1}, \{Y_s\}_{s \le t})\right) \left(\mathbf{1}^\top C_{t+1}(e_i) A_{t+1} p\right) \right. \\
&\qquad\qquad\qquad \left. - d^{-1} Z^i \right\},
\end{aligned}
$$

where $(A_{t+1}, C_{t+1}(\cdot)) \equiv \Phi(\mathfrak{S}, \mathfrak{D}_{t+1})$.

Proof As ξ_{t+1} is $\mathcal{Y}_{t+1}$-measurable, by the Doob–Dynkin lemma there exists a Borel measurable function $\hat{\xi}_{t+1}$ such that $\xi_{t+1} = \hat{\xi}_{t+1}(Y_{t+1})$ (omitting to write $\{Y_s\}_{s \le t}$ as an argument). We write Z_t for the vector containing each of the values of this function. From the definition of M, as in the proof of the martingale representation theorem in Cohen and Elliott (2010), it follows that

$$\xi_{t+1} - \mathbb{E}_{\bar{\mathbb{P}}}[\xi_{t+1}|\mathcal{Y}_t] = Z_t M_{t+1} \qquad \text{and} \qquad \mathbb{E}_{\bar{\mathbb{P}}}[\xi_{t+1}|\mathcal{Y}_t] = \sum_i d^{-1} Z^i.$$

We then calculate, using Lemma 2 (simplified to our finite-state setting and omitting $\{Y_s\}_{s \le t}$ as an argument),

$$\xi_t - \mathbb{E}_{\bar{\mathbb{P}}}[\xi_{t+1}|\mathcal{Y}_t]$$

$$= \bar{\mathcal{E}}(\xi_{t+1}|\mathcal{Y}_t) - \mathbb{E}_{\bar{\mathbb{P}}}[\xi_{t+1}|\mathcal{Y}_t]$$

$$= \operatorname*{ess\,sup}_{p,\mathfrak{S},\mathfrak{D}_{t+1}} \left\{ \sum_i \left(Z^i - \hat{\mathcal{R}}_{t+1}(e_i, p, \mathfrak{S}, \mathfrak{D}_{t+1}) \right) \left(\mathbf{1}^\top C_{t+1}(e_i) A_{t+1} p \right) - \mathbb{E}_{\bar{\mathbb{P}}}[\xi_{t+1}|\mathcal{Y}_t] \right\}$$

$$= f(Z_t; \kappa_t).$$

The answer follows by rearrangement. $\qquad\square$

6 A control problem with uncertain filtering

In this final section, we consider the solution of a simple control problem under uncertainty, using the formal structures previously developed. In some ways, this approach is similar to those considered by Bielecki et al. (2017), where the DR-expectation is replaced by an approximate confidence interval. (Taking $k' = \infty$ in a StaticDR model would give a very similar problem to the one they consider.) A key complexity in doing this is that our uncertainty does not agree with the arrow of time—at time t, we do not know the future values of $\{Y_s, X_s\}_{s>t}$ (as is typical for stochastic control), but we also do not know the values of $\{X_s\}_{s\leq t}$, even though these have an indirect impact on our costs.

Suppose a controller selects a control u from a set U, which we assume is a countable union of compact metrizable sets. Controls are required to be $\mathcal{Y}$-predictable (i.e., u_t is $\mathcal{Y}_{t-1}$-measurable), and we write $\mathcal{U}$ for the space of such controls, and $\mathbf{u} = (u_1, ..., u_T)$ for the vector of controls at every time.

A control has an impact on the generator of X, Y, through modifying the penalty function γ, which describes the "reasonable" models for the transition matrix A and the distribution of observations C. In particular, for a given u the term γ_t in the additive structure of $\mathcal{R}_t$ is permitted to depend on u_t. We assume $\gamma_t(\cdots ; u_t)$ is continuous in u_t for every value of its other arguments. This is a variant on a standard weak formulation of the control problem—our agent no longer selects the generators $(A, C(\cdot))$ directly, but instead modifies the penalty determining which values of $(A, C(\cdot))$ are 'reasonable models'.

In order to separate the effects of past controls (which determine the agent's understanding of the present hidden state), and future controls (which modify the future dynamics), we write $\bar{\mathcal{E}}_{\mathbf{u}, K}(\cdot|\mathcal{Y}_t)$ for the $\mathcal{Y}$-consistent expectation generated by the single-step expectations (omitting $\{Y_s\}_{s\leq t}$ as an argument)

$$\bar{\mathcal{E}}_{u_{t+1}, K_t}(\xi|\mathcal{Y}_t)$$

$$= \operatorname*{ess\,sup}_{p,\mathfrak{S},\mathfrak{D}_{t+1}} \int_{\mathbb{R}^d} \left(\hat{\xi}(y) - \hat{\mathcal{R}}_{t+1}(y, p, \mathfrak{S}, \mathfrak{D}_{t+1}; u_{t+1}, K_t) \right)$$

$$\cdot \left(\mathbf{1}^\top C_{t+1}(y) A_{t+1} p \right) d\mu(y),$$

where

$$\hat{\mathcal{R}}_{t+1}(Y_{t+1}, p, \mathfrak{S}, \mathfrak{D}_{t+1}; u_{t+1}, K_t) = \left(\frac{1}{k}\left(K_t(p_t, \mathfrak{S}) + \gamma_{t+1}(Y_{t+1}, \mathfrak{S}, \mathfrak{D}_{t+1}, p_t; u_{t+1})\right)\right)^{k'}.$$

(11)

We then observe that $\bar{\mathcal{E}}_{\mathbf{u}, K}(\cdot | \mathcal{Y}_t)$ is formally independent of $(u_1, ..., u_{t-1})$. Nevertheless, the effective value of K_t depends on $(u_1, ..., u_{t-1})$, as these now appear in the γ terms appearing in Theorem 5.

The controller wishes to minimize an expected cost

$$\bar{\mathcal{E}}_{\mathbf{u}, K_0}\left(\mathfrak{C}(X_T, \{Y_s\}_{s \le T}) + \sum_{t < T} \mathfrak{L}_t(\{Y_s\}_{s \le t}, u_{t+1})\right),$$

where $K_0 = \kappa_{\text{prior}}$ is the uncertainty before the control problem begins. Here $\mathfrak{C}$ is a terminal cost, which may depend on the hidden state X_T, and $\mathfrak{L}$ is a running cost, which depends on the control u_{t+1} used at time t. We assume $\mathfrak{C}$ and $\mathfrak{L}$ are continuous in u (almost surely). We think of the cost $\mathfrak{L}_t$ as being paid at time t, depending on the choice of control u_{t+1} (which will affect the generator at time $t+1$). For notational simplicity, we omit Y as an argument when unnecessary.

Remark 24 *We do not allow $\mathfrak{L}_t$ to depend on X_t, as this may lead to paradoxes (as the agent could learn information about the hidden state by observing their running costs).*

For a given control process $\mathbf{u}$, we define the remaining cost

$$J(\omega, t, K, \mathbf{u}) = \bar{\mathcal{E}}_{\mathbf{u}, K}\left(\mathfrak{C}(X_T) + \sum_{t \le s < T} \mathfrak{L}_s(u_{s+1}) \Big| \mathcal{Y}_t\right)$$

and hence the value function

$$V(\omega, t, K) = \inf_{u \in \mathcal{U}} \bar{\mathcal{E}}_{\mathbf{u}, K}\left(\mathfrak{C}(X_T) + \sum_{t \le s < T} \mathfrak{L}_s(u_{s+1}) \Big| \mathcal{Y}_t\right).$$

Remark 25 *We define our expected cost using the $\mathcal{Y}$-consistent expectation $\bar{\mathcal{E}}_{\mathbf{u}, K}$, rather than the (inconsistent) DR-expectation $\mathcal{E}_{\mathbf{y}_t}$, as this leads to time-consistency in the choice of controls.*

Remark 26 *We see that the calculation of the value function is a "minimax" problem, in that V minimizes the cost, which we evaluate using a maximum over a set of models. However, given the potential for learning, the requirement for time consistency, and the uncertainties involved, it is not clear that one can write V explicitly in terms of a single minimization and maximization of a given function.*

Remark 27 *As the filter-state penalty K is a general function depending on the control, and Y only takes finitely many states, it is not generally possible to express the effect (on K) of a control through a change of measure relative to some reference*

dynamics. In particular, we face the problem that controls u_s for times $s < T$ have an impact on the terminal cost $V_T = \bar{\mathcal{E}}_{\mathbf{u}, K_T}(\mathfrak{C}(X_T)|\mathcal{Y}_T)$, through their impact on the uncertainty K_T. Unlike in a traditional control problem, V_T is not independent of u given $\mathcal{Y}_T$; this is problem of the 'arrow of time' mentioned at the start of this section. For this reason, even though we model the impact of a control through its effect on the generator, we cannot give a fully "weak" formulation of our control problem, and are restricted to a "Markovian" setting with K as a state variable.

Theorem 10 *The value function satisfies a dynamic programming principle, in particular, if an optimal control u^* exists, then for every $t \le T$,*

$$V(\omega, t - 1, K_{t-1}) = \bar{\mathcal{E}}_{\mathbf{u}, K}\left(V_t\left(\omega, t - 1, K_t^{(u^*, K_{t-1})}\right)\Big|\mathcal{Y}_{t-1}\right) + \mathfrak{L}_{t-1}\left(u_t^*\right),$$

where $K_t^{(u^*, K_{t-1})}$ is the one-step solution of the recursion of Theorem 5 using the control u^*.

  A similar result also holds if we only assume an ϵ-optimal control exists for every $\epsilon > 0$.

  Proof This effectively falls into the setting of our abstract dynamic programming principle (in particular, Corollary 1), with a time reversal. For any control u, using the recursivity of $\bar{\mathcal{E}}$ and writing $\kappa_t = \kappa_t^{(u, \kappa_{t-1})}$ for simplicity, we have

$$J(\omega, t - 1, K_{t-1}, \mathbf{u})$$

$$= \bar{\mathcal{E}}_{\mathbf{u}, K_{t-1}}\left(\mathfrak{C}(X_T) + \sum_{t-1 \le s \le T} \mathfrak{L}_s(u_{s+1})\Big|\mathcal{Y}_{t-1}\right)$$

$$= \bar{\mathcal{E}}_{\mathbf{u}, K_{t-1}}\left(J(\omega, t, \kappa_t, \mathbf{u})\big|\mathcal{Y}_{t-1}\right) + \mathfrak{L}_{t-1}(u_t).$$

  In reversed time $\tau = T - t$, we have the state variable $z_\tau = K_{T-t}$, which is defined using a backward recursion (in terms of τ) by Theorem 5. The operator $\mathcal{A}_\tau$ is then given by

$$\mathcal{A}_\tau[J(\omega, T - \tau - 1, K_{T-\tau-1}, \mathbf{u}), u, K_{T-\tau-1}]$$
$$:= \bar{\mathcal{E}}_{\mathbf{u}, K_{T-\tau-1}}\left(J(\omega, T - \tau, K_{T-\tau}, u)\big|\mathcal{Y}_{t-1}\right) + \mathfrak{L}_{T-\tau-1}(u_{T-\tau}),$$

which is monotone and continuous in its first argument. The result then follows from Corollary 1. □

Remark 28 *The appeal to Corollary 1 is slightly complicated by the fact that V is a random variable, rather than simply a scalar value. The reader can verify that this does not affect the proof of Theorem 4 significantly, as we are in a finitely generated probability space (so measurable selection arguments remain straightforward), and the operator $\mathcal{A}$ forces the solution to have the desired measurability through time.*

  By combining this dynamic programming property with the definition of the one-step expectation, we can write down a difference equation which V must solve.

Theorem 11 *The value function of the control problem satisfies the recursion*

$$V(\omega, t, K_t)$$

$$= \inf_{u \in U} \operatorname{ess\,sup}_{p, \mathfrak{S}, \mathfrak{D}_{t+1}} \left\{ \int_{\mathbb{R}^d} \left(V\left(\omega, t, K_t^{u, K_{t-1}}\right) + \mathfrak{L}_{t+1}(y, u) \right. \right.$$

$$\left. \left. - \hat{\mathcal{R}}_{t+1}\left(y, p, \mathfrak{S}, \mathfrak{D}_{t+1}; u_{t+1}, K_t\right) \right) \left(\mathbf{1}^\top C_{t+1}(y) A_{t+1} p \right) d\mu(y) \right\}$$

and terminal value

$$V(\omega, T, K_T) := \mathcal{E}_{\mathbf{y}_T}\left(\mathfrak{C}\left(X_T, \{Y_s\}_{s \leq T}\right) \right).$$

Here, $K_t^{u, K_{t-1}}$ is the value of K_t starting at K_{t-1}, with control u at time t, and evolving following Theorem 5 and $\hat{\mathcal{R}}$ is as in (11).

Corollary 2 *A control is optimal if and only if it achieves the infimum in the formula for V above.*

Remark 29 *If we assume that the terminal cost depends only on X_T (and not on Y), and the running cost does not depend on Y, then one can observe a Markov property to the control problem, that is, V_s is conditionally independent of Y given K_s. The corresponding optimal controls can then also be taken only to depend on K_s.*

Authors' contributions

The author(s) read and approved the final manuscript.

Funding

Availability of data and materials

Not applicable.

Competing interests

The author declares that he has no competing interests.

References

Allan, A.L. and S.N. Cohen. (2019a). *Parameter uncertainty in the Kalman–Bucy filter*, SIAM J. Control Optim. **57**, no. 3, 1646–1671.

Allan, A.L. and S.N. Cohen. (2020). *Pathwise Stochastic Control with Applications to Robust Filtering*, Ann. Appl. Prob. arXiv::1902.05434.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath. (1999). *Coherent measures of risk*, Math. Finan. **9**, no. 3, 203–228.

Başar, T. and P. Bernhard. (1991). *H^∞-Optimal Control and Related Minimax Design Problems, A Dynamic Game Approach*, Birkhäuser, Basel.

Bain, A. and D. Crisan. (2009). *Fundamentals of Stochastic Filtering*, Springer, Berlin–Heidelberg–New York.

Bielecki, T.R., T. Chen, and I. Cialenco. (2017). *Recursive construction of confidence regions*, Electron. J. Stat. **11**, no. 2, 4674–4700.

Boel, R.K., M.R. James, and I.R. Petersen. (2002). *Robustness and risk-sensitive filtering*, IEEE Trans. Autom. Control **47**, no. 3, 451–461.

Cohen, S.N. and R.J. Elliott. (2010). *A general theory of finite state backward stochastic difference equations*, Stoch. Process. Appl. **120**, no. 4, 442–466.

Cohen, S.N. and R.J. Elliott. (2011). *Backward stochastic difference equations and nearly-time-consistent nonlinear expectations*, SIAM J. Control Optim. **49**, no. 1, 125–139.

Cohen, S.N. and R.J. Elliott. (2015). *Stochastic Calculus and Applications*, 2nd ed., Birkhäuser, New York.

Cohen, S.N. (2017). *Data-driven nonlinear expectations for statistical uncertainty in decisions*, Electron. J. Stat. **11**, no. 1, 1858–1889.

Delbaen, F., S. Peng, and E. Rosazza Gianin. (2010). *Representation of the penalty term of dynamic concave utilities*, Finan. Stochast. **14**, no. 3, 449–472.

Dey, S. and J.B. Moore. (1995). *Risk-sensitive filtering and smoothing for hidden Markov models*, Syst. Control Lett. **25**, 361–366.

Douc, R., E. Moulines, J. Olsson, and R. van Handel. (2011). *Consistency of the maximum likelihood estimator for general hidden Markov models*, Ann. Stat. **39**, no. 1, 474–513.

Duffie, D. and L.G. Epstein. (1992). *Asset pricing with stochastic differential utility*, Rev. Finan. Stud. **5**, no. 3, 411–436.

El Karoui, N., S. Peng, and M.C. Quenez. (1997). *Backward stochastic differential equations in finance*, Math. Finan. **7**, no. 1, 1–71.

Epstein, L.G. and M. Schneider. (2003). *Recursive multiple-priors*, J. Econ. Theory **113**, 1–31.

Fagin, R. and J. Halpern. (1990). *A new approach to updating beliefs*, AUAI Press, Corvallis.

Föllmer, H. and A. Schied. (2002a). *Convex measures of risk and trading constraints*, Finan. Stochast. **6**, 429–447.

Föllmer, H. and A. Schied. (2002b). *Stochastic Finance: An Introduction in Discrete Time. Studies in Mathematics 27*, de Gruyter, Berlin-New York.

Frittelli, M. and E. Rosazza Gianin. (2002). *Putting order in risk measures*, J. Bank. Financ. **26**, no. 7, 1473–1486.

Graf, S. (1980). *A Radon–Nikodym theorem for capacities*, J. für die reine und angewandte Mathematik **320**, 192–214.

Grimble, M.J. and A. El Sayed. (1990). *Solution of the H_∞ optimal linear filtering problem for discrete-time systems*, Trans. Acoust. Speech Sig. Process. IEEE **38**, no. 7.

Hansen, L.P. and T.J. Sargent. (2005). *Robust estimation and control under commitment*, J. Econ. Theory **124**, 258–301.

Hansen, L.P. and T.J. Sargent. (2007). *Recursive robust estimation and control without commitment*, J. Econ. Theory **136**, no. 1, 1–27.

Hansen, L.P. and T.J. Sargent. (2008). *Robustness*, Princeton University Press, Princeton.

Huber, P.J. and E.M. Roncetti. (2009). *Robust Statistics, 2nd edn.*, Wiley, Hoboken.

James, M.R., J.S. Baras, and R.J. Elliott. (1994). *Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems*, Trans. Autom. Control IEEE **39**, no. 4, 780–792. https://doi.org/10.1109/9.286253.

Kalman, R.E. (1960). *A new approach to linear filtering and prediction problems*, J. Basic Eng. ASME **82**, 33–45.

Kalman, R.E. and R.S. Bucy. (1961). *New results in linear filtering and prediction theory*, J. Basic Eng. ASME **83**, 95–108.

Keynes, J.M. (1921). *A Treatise on Probability*, Macmillan and Co., New York. Reprint BN Publishing, 2008.

Knight, F.H. (1921). *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston. reprint Dover 2006.

Kupper, M. and W. Schachermayer. (2009). *Representation results for law invariant time consistent functions*, Math. Financ. Econ. **2**, no. 3, 189–210.

Leroux, B.G. (1992). *Maximum-likelihood estimation for hidden Markov models*, Stoch. Process. Appl. **40**, 127–143.

Peng, S. (2010). *Nonlinear Expectations and Stochastic Calculus under Uncertainty*, arxiv::1002.4546v1.

Riedel, F. (2004). *Dynamic coherent risk measures*, Stochast. Process. Appl. **112**, no. 2, 185–200.

Rockafellar, R.T., S. Uryasev, and M. Zabarankin. (2006). *Generalized deviations in risk analysis*, Finan. Stochast. **10**, 51–74.

Wald, A. (1945). *Statistical decision functions which minimize the maximum risk*, Ann. Math. **46**, no. 2, 265–280.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, London.

Wonham, W.N. (1965). *Some applications of stochastic differential equations to optimal nonlinear filtering*, SIAM J. Control **2**, 347–369.

Zhang, J., Y. Xia, and P. Shi. (2009). *Parameter-dependent robust H_∞ filtering for uncertain discrete-time systems*, Automatica **45**, 560–565.