Adv. Appl. Probab. **54,** 625–655 (2022)
*doi:*10.1017/apr.2021.44

DISCRETE-TIME RISK-AWARE OPTIMAL SWITCHING WITH NON-ADAPTED COSTS

RANDALL MARTYR,* *Kingston University London*
JOHN MORIARTY,** *Queen Mary University of London*
MAGNUS PERNINGE,*** *Linnaeus University*

Abstract

We solve non-Markovian optimal switching problems in discrete time on an infinite horizon, when the decision-maker is risk-aware and the filtration is general, and establish existence and uniqueness of solutions for the associated reflected backward stochastic difference equations. An example application to hydropower planning is provided.

Keywords: Infinite horizon; optimal switching; risk measures; reflected backward stochastic difference equations; hydropower planning

2020 Mathematics Subject Classification: Primary 93E20
Secondary 60G40; 91B08; 49N30

1. Introduction

1.1. Optimal switching problems

Optimal switching problems involve an agent controlling a system by successively switching an operational mode between a discrete set of choices. Time may be either continuous or discrete, and in all cases the latter is useful for numerical work (see for example [3]). In related contexts, risk sensitivity with respect to uncertain costs has been modelled using nonlinear expectations; see [1] for example. This feature is particularly appropriate in data-driven settings where models themselves may be uncertain. Examples include when the probability model is derived from numerical weather predictions depending on unknown physical parameters, or, alternatively, in model-free reinforcement learning. In the latter context, recent work has applied a general analytic framework for risk sensitivity [10].

Taking a probabilistic approach, in this paper we consider a general filtration, which interacts with the nonlinear expectation. More precisely, let $\mathbb{T}$ be a subset of $\mathbb{N}_0 := \{0, 1, \dots\}$, and let $\left\{\tilde{g}_{\xi_{t-1},\xi_t}(t)\right\}_{t\in\mathbb{T}}$ be a sequence of random costs dependent on a switching strategy ξ, i.e. a random sequence $(\xi_t)_{t\in\{-1\}\cup\mathbb{T}}$ taking values in a finite set $\mathcal{I} := \{1, \dots, m\}$, representing the set of operating modes. We do not require that every cost be observable, which, for example, enables study of the interaction between delayed or missing observations and risk sensitivity. The time horizon is either infinite ($\mathbb{T} = \mathbb{N}_0$) or finite ($\mathbb{T} = \{0, 1, \dots, T\}$ for some finite $T \geq 0$),

Received 30 July 2020; revision received 10 August 2021.
* Postal address: River House, 5357 High Street, Surrey, UK. Email address: r.martyr@kingston.ac.uk
** Postal address: School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London, UK. Email address: jmoriarty@qmul.ac.uk
*** Postal address: PG Vejdes väg 7, 352 52 Växjö, Sweden.

and the value of the switching problem is defined under a nonlinear expectation (cf. Equations (2.1) and (3.2) below). Optimal stopping problems (see, for example, [1]) are recovered in the special case of two modes (i.e. $m = 2$), when optimisation is performed over strategies ξ with a single jump.

1.2. Setup and related work

We have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\mathbb{G} = \{\mathcal{G}_t\}_{t\in\mathbb{T}}$ of sub-σ-algebras of $\mathcal{F}$. Given operating modes $\mathcal{I} := \{1, \ldots, m\}$ and essentially bounded random variables $g = \{g_i(t)\colon i \in \mathcal{I}\}_{t\in\mathbb{T}}$ and $c = \{c_{i,j}(t)\colon i, j \in \mathcal{I}\}_{t\in\mathbb{T}}$ on $(\Omega, \mathcal{F}, \mathbb{P})$, we are interested in solving an optimal switching problem with running costs g and switching costs c when the information available to the decision-maker is given progressively according to $\mathbb{G}$, and where a dynamic measure of risk sensitivity is used which generalises the usual sequence $\{\mathbb{E}[\,\cdot\,|\mathcal{G}_t],\ t \in \mathbb{T}\}$ of conditional expectations with respect to $\mathbb{G}$. For the following discussion we set

$$\tilde{g}_{\xi_{t-1},\xi_t}(t) := g_{\xi_t}(t) + c_{\xi_{t-1},\xi_t}(t). \tag{1.1}$$

Note that we are considering a setting where each of the costs $\left\{\tilde{g}_{i,j}(t)\colon i, j \in \mathcal{I}\right\}_{t\in\mathbb{T}}$ is measurable with respect to the σ-algebra $\mathcal{F}$ and $\mathbb{G}$ is any filtration with $\mathcal{G}_t \subset \mathcal{F}$ for all $t \in \mathbb{T}$. We thus may have, but do not limit ourselves to, the situation where $\mathbb{G}$ is the natural filtration generated by $\{\tilde{g}_{i,j}(t)\colon i, j \in \mathcal{I}\}_{t\in\mathbb{T}}$. To our knowledge, the necessary and sufficient conditions we provide for an optimal switching strategy in this infinite-horizon setting under general filtration are novel and extend results in, for example, [1, 4, 8, 11, 15].

The rest of the paper is structured as follows. Section 2 presents our main results in the finite-horizon setting, and these are extended to infinite horizon in Section 3. In both cases, the solution to the optimal switching problem is used to establish the existence of solutions to the associated reflected backward stochastic difference equations, and we also prove uniqueness of the solution. We close the paper with two examples. Section 4 briefly confirms that the approach taken to missing or delayed observations is capable of changing both the value process and the optimal strategy. In Section 5 we apply neural networks to obtain numerical solutions to a non-Markovian hydropower planning problem with non-adapted costs and examine the risk sensitivity of the solutions.

2. Finite-horizon risk-aware optimal switching under general filtration

For the rest of the paper, we establish the following notation:

- Let $m\mathcal{F}$ denote the space of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.
- Let $L^\infty_{\mathcal{F}}$ be the subspace of essentially bounded random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.
- Let $\mathbb{G} = \{\mathcal{G}_t\}_{t\in\mathbb{T}}$ be a filtration, with $\mathcal{G} = \bigvee_{t\in\mathbb{T}} \mathcal{G}_t$ the σ-algebra generated by all $\mathcal{G}_t$ and $\mathcal{G} \subset \mathcal{F}$.
- Let $T < \infty$ be a finite time horizon, and for $0 \le t \le T$, let $\mathscr{T}_{[t,T]}$ (resp. $\mathscr{T}_t$) denote the set of $\mathbb{G}$-stopping times with values in $t, \ldots, T$ (resp. $t, t+1, \ldots$).
- Let ρ be a $\mathbb{G}$-conditional risk mapping: a family of mappings $\{\rho_t\}_{t\in\mathbb{T}}$, $\rho_t\colon L^\infty_{\mathcal{F}} \to L^\infty_{\mathcal{G}_t}$, satisfying normalisation, conditional translation invariance, and monotonicity (see Appendix A.1).

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

- For $s, t \in \mathbb{T}$ with $s \leq t$, let $\rho_{s,t}$ be the finite-horizon aggregated (or nested) risk mapping generated by ρ ([5, 14, 15, 20, 21]; see also [2]): that is, $\rho_{t,t}(W_t) = \rho_t(W_t)$ and

$$\rho_{s,t}(W_s, \ldots, W_t) = \rho_s\big(W_s + \rho_{s+1}\big(W_{s+1} + \cdots + \rho_{t-1}\big(W_{t-1} + \rho_t(W_t)\big) \cdots\big)\big), \;\; s < t.$$

- All inequalities are interpreted in the $\mathbb{P}$-almost-sure sense.

For the finite-time-horizon setting of Section 2 we also set $\mathbb{T} = \{0, 1, \ldots, T\}$.
The value process for the finite-horizon optimal switching problem is

$$V_t^i := \operatorname*{ess\,inf}_{\xi \in \mathcal{U}_t^i} \rho_{t,T}\big(\tilde{g}_{\xi_{t-1},\xi_t}(t), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}(T)\big), \tag{2.1}$$

where $\tilde{g}_{i,j}(t) := g_j(t) + c_{i,j}(t)$, $\mathcal{U}_t^i$ is the set of $\mathbb{G}$-adapted strategies ξ with $\xi_{t-1} = i$, and the infimum of the empty set is taken to be ∞. Since for each t the costs $c_{i,i}(t)$ depend only on i and may therefore be accounted for in the term $g_i(t)$, without loss of generality we may make the following assumption.

Assumption 2.1. *For all $i \in \mathcal{I}$ we have $c_{i,i}(t) = 0$ for all $t \in \mathbb{T}$.*

2.1. Dynamic programming equations

The use of aggregated risk mappings provides sufficient structure for dynamic programming. In our non-Markovian setting, appropriate equations are

$$\begin{cases} \hat{V}_T^i = \min_{j \in \mathcal{I}} \rho_T\big(\tilde{g}_{i,j}(T)\big), \\ \hat{V}_t^i = \min_{j \in \mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + \hat{V}_{t+1}^j\Big), \quad \text{for } 0 \leq t < T. \end{cases} \tag{2.2}$$

(The random fields $\{\hat{V}_t^i : i \in \mathcal{I}, t \in \mathbb{T}\}$ coincide with Snell envelopes; see Remark 2.3.) We note by induction that $\hat{V}_t^i \in L^\infty_{\mathcal{G}_t}$ for each $i \in \mathcal{I}$ and $t \in \mathbb{T}$.

Remark 2.1. For comparison, in a Markovian framework with full observation and the linear expectation, randomness stems from an $\mathbb{R}^k$-valued Markov chain $X^{s,x} := \{X_t^{s,x}\}_{s \leq t \leq T}$, where $(s, x) \in \mathbb{T} \times \mathbb{R}^k$ is fixed and $X_r^{s,x} = x$ for $0 \leq r \leq s$ almost surely under $\mathbb{P}^{(s,x)}$, and $\mathbb{G}$ is the natural filtration of $X^{s,x}$. In the Markovian case, by virtue of each strategy ξ being adapted to $\mathbb{G}$, for every $t \geq 0$ there exists a function $\Xi_t : \big(\mathbb{R}^k\big)^{t+1} \to \mathcal{I}$ such that $\xi_t = \Xi_t(X_0, \ldots, X_t)$. The Bellman equation is then the appropriate formulation for dynamic programming: for any $i \in \mathcal{I}$ and $(s, x) \in \mathbb{T} \times \mathbb{R}^k$,

$$\begin{cases} v^i(T, x) = \min_{j \in \mathcal{I}} \tilde{g}_{i,j}(T, x), \\ v^i(s, x) = \min_{j \in \mathcal{I}} \Big(\tilde{g}_{i,j}(s, x) + \mathbb{E}^{(s,x)}\big[v^j\big(s+1, X_{s+1}\big)\big]\Big), \;\; s < T, \end{cases} \tag{2.3}$$

where the v^i and $\tilde{g}_{i,j}$ are deterministic functions on $\mathbb{T} \times \mathbb{R}^k$.

Theorem 2.1. *The random field $\{\hat{V}_t^i : i \in \mathcal{I}, t \in \mathbb{T}\}$ consists of value processes for the optimal switching problem, in the sense that*

$$\hat{V}_t^i = V_t^i \quad \forall\, i \in \mathcal{I}, \;\; t \in \mathbb{T}.$$

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Moreover, starting from any $0 \le t \le T$ *and* $i \in \mathcal{I}$, *an optimal strategy* $\xi^* \in \mathcal{U}_t^i$ *can be defined as follows:*

$$\begin{cases} \xi^*_{t-1} = i, \\ \xi^*_s \in \arg\min_{j\in\mathcal{I}} \rho_s\Big(\tilde{g}_{\xi^*_{s-1},j}(s) + \hat{V}^j_{s+1}\Big), \quad t \le s < T, \\ \xi^*_T \in \arg\min_{j\in\mathcal{I}} \rho_T\Big(\tilde{g}_{\xi^*_{T-1},j}(T)\Big). \end{cases} \tag{2.4}$$

Proof. Note that the result holds trivially for $t = T$. We will apply a backward induction argument and assume that for $s = t+1, t+2, \ldots, T$ and all $i \in \mathcal{I}$ we have $\hat{V}^i_s = V^i_s = \rho_{s,T}\left(\tilde{g}_{i,\xi_s}(s), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}(T)\right)$, where $\xi_{t-1} = i$ and

$$\xi_s \in \arg\min_{j\in\mathcal{I}} \rho_s\Big(\tilde{g}_{\xi_{s-1},j}(s) + V^j_{s+1}\Big),$$

with $V^j_{T+1} := 0$ for all $j \in \mathcal{I}$.

The induction hypothesis implies that

$$\begin{aligned} \hat{V}^i_t &= \min_{j\in\mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + V^j_{t+1}\Big) \\ &= \min_{j\in\mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^j_{t+1}} \rho_{t+1,T}\Big(\tilde{g}_{j,\xi_{t+1}}(t+1), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}(T)\Big)\Big). \end{aligned}$$

For any $\xi' \in \mathcal{U}^i_t$ we note that by monotonicity and conditional translation invariance we have

$$\begin{aligned} \hat{V}^i_t &\le \min_{j\in\mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + \rho_{t+1,T}\Big(\tilde{g}_{j,\xi'_{t+1}}(t+1), \ldots, \tilde{g}_{\xi'_{T-1},\xi'_T}(T)\Big)\Big) \\ &\le \sum_{j=1}^m \mathbf{1}_{\{\xi'_t=j\}} \rho_t\Big(\tilde{g}_{i,j}(t) + \rho_{t+1,T}\Big(\tilde{g}_{j,\xi'_{t+1}}(t+1), \ldots, \tilde{g}_{\xi'_{T-1},\xi'_T}(T)\Big)\Big) \\ &= \rho_t\left(\sum_{j=1}^m \mathbf{1}_{\{\xi'_t=j\}} \Big\{\tilde{g}_{i,j}(t) + \rho_{t+1,T}\Big(\tilde{g}_{j,\xi'_{t+1}}(t+1), \ldots, \tilde{g}_{\xi'_{T-1},\xi'_T}(T)\Big)\Big\}\right) \\ &= \rho_{t,T}\Big(\tilde{g}_{i,\xi'_t}(t), \ldots, \tilde{g}_{\xi'_{T-1},\xi'_T}(T)\Big). \end{aligned}$$

Taking the infimum over all $\xi' \in \mathcal{U}^i_t$ we conclude that $\hat{V}^i_t \le V^i_t$. However, letting $\xi'_{t-1} = i$ and defining

$$\xi'_s \in \arg\min_{j\in\mathcal{I}} \rho_t\Big(\tilde{g}_{\xi'_{s-1},j}(s) + \hat{V}^j_{s+1}\Big)$$

for $s = t, \ldots, T$, with $\hat{V}^j_{T+1} := 0$ for all $j \in \mathcal{I}$, we find that

$$\begin{aligned} \hat{V}^i_t &= \rho_t\Big(\tilde{g}_{i,\xi'_t}(t) + \operatorname*{ess\,inf}_{\xi\in\mathcal{U}_{t+1}} \rho_{t+1,T}\Big(\tilde{g}_{\xi'_t,\xi_{t+1}}(t+1), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}(T)\Big)\Big) \\ &= \rho_{t,T}\Big(\tilde{g}_{i,\xi'_t}(t), \ldots, \tilde{g}_{\xi'_{T-1},\xi'_T}\Big) \\ &\ge V^i_t. \end{aligned}$$

□

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

2.2. Relation to systems of RBSΔEs

We now introduce a *reflected backward stochastic difference equation* (RBSΔE), which is a class of equations relevant to both optimal stopping and switching problems and is studied systematically in [1] for finite-state processes. Let $\mathcal{L}^{\infty}_{\mathbb{G},T} := \otimes_{t=0}^{T} L^{\infty}_{\mathcal{G}_t}$. To avoid excessive notation, some notation for scalar-valued processes will be reused for vector-valued ones, with the interpretation that all components are in the same space. Similarly, inequalities and martingale properties will be understood componentwise, and given $i \in \mathcal{I}$ we write $\mathcal{I}^{-i} := \mathcal{I} \setminus \{i\}$.

Definition 2.1 *(Finite horizon RBSΔEs).* With $\mathbb{T} = \{0, \ldots, T\}$, where $0 \le T < \infty$, let $Y = \{Y_t\}_{t\in\mathbb{T}}$, $M = \{M_t\}_{t\in\mathbb{T}}$, and $A = \{A_t\}_{t\in\mathbb{T}}$ be $\mathbb{G}$-adapted $\mathbb{R}^m$-valued processes satisfying

$$\begin{cases} Y^i_t = \min_{j\in\mathcal{I}} \rho_T\big(\tilde{g}_{i,j}(T)\big) + \sum_{s=t}^{T-1} \rho_s\Big(g_i(s) + \Delta M^i_{s+1}\Big) \\ \qquad -\big(M^i_T - M^i_t\big) - \big(A^i_T - A^i_t\big), \quad \forall\, t \in \mathbb{T}, \\ Y^i_t \le \min_{j\in\mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + Y^j_{t+1}\Big), \quad \forall\, t \in \mathbb{T}, \\ \sum_{t=0}^{T-1} \Big(Y^i_t - \min_{j\in\mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + Y^j_{t+1}\Big)\Big)\Delta A^i_{t+1} = 0. \end{cases} \tag{2.5}$$

A triple $(Y, M, A) \in \big(\mathcal{L}^{\infty}_{\mathbb{G},T}\big)^3$ is said to be a solution to the system of RBSΔEs (2.5) if M is a $\mathbb{G}$-adapted $\rho_{s,t}$-martingale (applying the definition in Section A.3 of the appendix), A is non-decreasing and $\mathbb{G}$-predictable (with $M_0 = A_0 = 0$), and (Y, M, A) satisfies (2.5). A solution (Y, M, A) is called unique if any other solution (Y', M', A') is indistinguishable as a process from (Y, M, A).

Remark 2.2. The martingale characterisation of the optimal switching value process (see for example [16] under the linear expectation) may be derived from the associated RBSΔE. Under a risk mapping ρ, however, the 'driver' $\rho_t\big(g_i(t) + \Delta M^i_{t+1}\big)$ in (2.5) depends on the $\{\rho_{s,t}\}$-martingale difference ΔM^i_{t+1}, which is natural for general (infinite-state) backward stochastic difference equations—see [6]. Note also that the driver is a function of the mappings $\omega \mapsto \Delta M^i_{t+1}(\omega)$ and $\omega \mapsto g_i(\omega, t)$ and not the realised values of these random variables. Also, we refer to the last line in Equation (2.5) as the *Skorokhod condition*.

The optimal switching problem (2.1) is related to this system of RBSΔEs through the following result.

Theorem 2.2. *The system of RBSΔEs* (2.5) *has a unique solution* (Y, M, A)*. Furthermore, we have* $Y = V$*.*

Proof. We divide the proof into two parts:

Existence: We aim to find a family of $\rho_{s,t}$-martingales $M = \{M^i\}_{i\in\mathcal{I}}$ and non-decreasing $\mathbb{G}$-predictable processes $A = \{A^i\}_{i\in\mathcal{I}}$ such that (V, M, A) solves (2.5). For every $i \in \mathcal{I}$, define the sequence $\{A^i_t\}_{t=0}^{T}$ by

$$\begin{cases} A^i_0 = 0, \\ A^i_t = A^i_{t-1} + \rho_{t-1}\big(g_i(t-1) + V^i_t\big) - V^i_{t-1}, \quad t = 1, \ldots, T. \end{cases}$$

We note that A^i is $\mathbb{G}$-predictable and non-decreasing since, by Theorem 2.1 and the backward induction formula (2.2), $V^i_t \le \rho_t\big(g_i(t) + V^i_{t+1}\big)$. Furthermore, for $t < T$ we have $\Delta A^i_{t+1} = \rho_t\big(g_i(t) + V^i_{t+1}\big) - V^i_t = 0$ on $\big\{V^i_t = \rho_t\big(g_i(t) + V^i_{t+1}\big)\big\} \supset \big\{V^i_t < \min_{j\in\mathcal{I}^{-i}} \rho_t\big(\tilde{g}_{i,j}(t) + V^j_{t+1}\big)\big\}$.

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Let M^i be the martingale in the Doob decomposition (see Lemma A.5 in Appendix A.3) for V^i; that is, $M_0^i = 0$ and $\Delta M_{t+1}^i = V_{t+1}^i - \rho_t\big(V_{t+1}^i\big)$. We have

$$V_t^i = \min_{j\in\mathcal{I}} \rho_T\big(\tilde{g}_{i,j}(T)\big) + \sum_{s=t}^{T-1} \big(V_s^i - V_{s+1}^i\big).$$

Now, as

$$\begin{aligned}\Delta M_{s+1}^i + \Delta A_{s+1}^i &= V_{s+1}^i - \rho_s\big(V_{s+1}^i\big) + \rho_s\big(g_i(t) + V_{s+1}^i\big) - V_s^i \\ &= V_{s+1}^i + \rho_s\big(g_i(t) + V_{s+1}^i - \rho_s\big(V_{s+1}^i\big)\big) - V_s^i \\ &= V_{s+1}^i - V_s^i + \rho_s\big(g_i(t) + \Delta M_{s+1}^i\big),\end{aligned}$$

we get $V_s^i - V_{s+1}^i = \rho_s\big(g_i(s) + \Delta M_{s+1}^i\big) - \Delta M_{s+1}^i + \Delta A_{s+1}^i$, and thus

$$\begin{aligned}V_t^i = {} & \min_{j\in\mathcal{I}} \rho_T\big(\tilde{g}_{i,j}(T)\big) + \sum_{s=t}^{T-1} \rho_s\big(g_i(s) + \Delta M_{s+1}^i\big) - \big(M_T^i - M_t^i\big) \\ & - \big(A_T^i - A_t^i\big).\end{aligned}$$

We conclude that (V, M, A) is a solution to the RBSΔE (2.5).

Uniqueness: Suppose that (Y, N, B) is another solution. Then

$$\Delta Y_{t+1}^i = -\rho_t\big(g_i(t) + \Delta N_{t+1}^i\big) + \Delta N_{t+1}^i + \Delta B_{t+1}^i. \tag{2.6}$$

Applying ρ_t on both sides gives

$$\begin{aligned}\rho_t\big(\Delta Y_{t+1}^i\big) &= -\rho_t\big(g_i(t) + \Delta N_{t+1}^i\big) + \rho_t\big(\Delta N_{t+1}^i + \Delta B_{t+1}^i\big) \\ &= -\rho_t\big(g_i(t) + \Delta N_{t+1}^i\big) + \Delta B_{t+1}^i,\end{aligned}$$

since, by our assumption on solutions to the RBSΔE, ΔB_{t+1}^i is $\mathcal{G}_t$-measurable and N^i is a martingale. Inserted into Equation (2.6), this gives

$$\begin{aligned}\Delta N_{t+1}^i &= \Delta Y_{t+1}^i + \rho_t\big(g_i(t) + \Delta N_{t+1}^i\big) - \Delta B_{t+1}^i \\ &= \Delta Y_{t+1}^i - \rho_t\big(\Delta Y_{t+1}^i\big) \\ &= Y_{t+1}^i - \rho_t\big(Y_{t+1}^i\big)\end{aligned}$$

and

$$\begin{aligned}\Delta B_{t+1}^i &= \rho_t\big(\Delta Y_{t+1}^i\big) + \rho_t\big(g_i(t) + \Delta N_{t+1}^i\big) \\ &= \rho_t\big(\Delta Y_{t+1}^i\big) + \rho_t\big(g_i(t) + Y_{t+1}^i - \rho_t\big(Y_{t+1}^i\big)\big) \\ &= \rho_t\big(g_i(t) + Y_{t+1}^i\big) - Y_t^i.\end{aligned}$$

We conclude that

$$\begin{cases}\Delta N_{t+1}^i = Y_{t+1}^i - \rho_t\big(Y_{t+1}^i\big), \\ \Delta B_{t+1}^i = \rho_t\big(g_i(t) + Y_{t+1}^i\big) - Y_t^i,\end{cases} \tag{2.7}$$

and in particular we have that, given $Y \in \mathcal{L}^{\infty}_{\mathbb{G},T}$, there is at most (up to indistinguishability of processes) one pair (N, B) such that (Y, N, B) solves the RBSΔE (2.5).

Since (Y, N, B) solves the RBSΔE (2.5) we have that

$$Y_t^i \leq \min_{j \in \mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + Y_{t+1}^j\Big),$$

and

$$\begin{aligned} Y_t^i &= Y_{t+1}^i + \rho_t\big(g_i(t) + \Delta N_{t+1}^i\big) - \big(N_{t+1}^i - N_t^i\big) - \big(B_{t+1}^i - B_t^i\big) \\ &\leq Y_{t+1}^i + \rho_t\big(g_i(t) + \Delta N_{t+1}^i\big) - \big(N_{t+1}^i - N_t^i\big) \\ &= Y_{t+1}^i + \rho_t\big(g_i(t) + Y_{t+1}^i - \rho_t\big(Y_{t+1}^i\big)\big) - \big(Y_{t+1}^i - \rho_t\big(Y_{t+1}^i\big)\big) \\ &= \rho_t\big(g_i(t) + Y_{t+1}^i\big). \end{aligned}$$

We conclude that $Y_t^i \leq \min_{j \in \mathcal{I}} \rho_t\big(\tilde{g}_{i,j}(t) + Y_{t+1}^j\big)$ for all $t \leq T$ and $i \in \mathcal{I}$. For $t = T$ this implies that, for all $i \in \mathcal{I}$, $Y_T^i \leq \min_{j \in \mathcal{I}} \rho_T\big(\tilde{g}_{i,j}(t)\big) = V_T^i$. Assume that $t < T$ and $Y_{t+1}^i \leq V_{t+1}^i$ for all $i \in \mathcal{I}$; then

$$\begin{aligned} Y_t^i &\leq \min_{j \in \mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + Y_{t+1}^j\Big) \\ &\leq \min_{j \in \mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + V_{t+1}^j\Big) \\ &\leq V_t^i. \end{aligned}$$

Applying an induction argument, we thus find that if (Y, N, B) solves the RBSΔE (2.5), then $Y_t^i \leq V_t^i$ for all $t \leq T$ and $i \in \mathcal{I}$. To arrive at uniqueness we show that the value Y_t^i is attained by a strategy in which case the reverse inequality follows by optimality of V_t^i.

Define the stopping time $\bar{\tau}_1^{t,i} := \inf\big\{s \geq t : \Delta B_{s+1}^i > 0\big\} \wedge T$ and the $\mathcal{G}_{\bar{\tau}_1^{t,i}}$-measurable $\mathcal{I}$-valued random variable $\bar{\beta}_1^{t,i}$ as a measurable selection of

$$\begin{cases} \arg\min\limits_{j \in \mathcal{I}^{-i}} \rho_{\bar{\tau}_1^{t,i}}\left(\tilde{g}_{i,j}\left(\bar{\tau}_1^{t,i}\right) + Y^j_{\bar{\tau}_1^{t,i}+1}\right), & \bar{\tau}_1^{t,i} < T, \\ \arg\min\limits_{j \in \mathcal{I}} \rho_T\big(\tilde{g}_{i,j}(T)\big), & \bar{\tau}_1^{t,i} = T. \end{cases}$$

Now, as $B^i_{\bar{\tau}_1^{t,i}} - B_t^i = 0$, we have for $t \leq s < \bar{\tau}_1^{t,i}$ the recursion

$$\begin{aligned} Y_s^i &= Y_{s+1}^i + \rho_s\big(g_i(s) + \Delta N_{s+1}^i\big) - \big(\Delta N_{s+1}^i\big) - \big(\Delta B_{s+1}^i\big) \\ &= Y_{s+1}^i + \rho_s\big(g_i(s) + Y_{s+1}^i - \rho_s\big(Y_{s+1}^i\big)\big) - \big(Y_{s+1}^i - \rho_s\big(Y_{s+1}^i\big)\big) \\ &= \rho_s\big(g_i(s) + Y_{s+1}^i\big). \end{aligned}$$

Furthermore, by the Skorokhod condition, on $\big\{\bar{\tau}_1^{t,i} < T\big\}$ we have that

$$\begin{aligned} Y^i_{\bar{\tau}_1^{t,i}} &= \min_{j \in \mathcal{I}^{-i}} \rho_{\bar{\tau}_1^{t,i}}\Big(\tilde{g}_{i,j}\big(\bar{\tau}_1^{t,i}\big) + Y^j_{\bar{\tau}_1^{t,i}+1}\Big) \\ &= \rho_{\bar{\tau}_1^{t,i}}\Big(\tilde{g}_{i,\bar{\beta}_1^{t,i}}\big(\bar{\tau}_1^{t,i}\big) + Y^{\bar{\beta}_1^{t,i}}_{\bar{\tau}_1^{t,i}+1}\Big), \end{aligned}$$

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

and since $Y_T^i = \arg\min_{j\in\mathcal{I}} \rho_T\big(\tilde{g}_{i,j}(T)\big)$, we conclude that

$$Y_t^i = \rho_{t,\tau_1^{t,i}}\left(g_i(t), \ldots, g_i\left(\bar{\tau}_1^{t,i}-1\right), \tilde{g}_{i,\bar{\beta}_1^{t,i}}\left(\bar{\tau}_1^{t,i}\right) + Y^{\bar{\beta}_1^{t,i}}_{\bar{\tau}_1^{t,i}+1}\right),$$

with $Y_{T+1}^j = 0$ for all $j \in \mathcal{I}$.

This process can be repeated to define

$$\bar{\tau}_{k+1}^{t,i} := \inf\left\{s > \bar{\tau}_k^{t,i} \colon \Delta B_{s+1}^{\beta_k^{t,i}} > 0\right\} \wedge T$$

and the $\mathcal{G}_{\bar{\tau}_{k+1}^{t,i}}$-measurable $\mathcal{I}$-valued random variable $\bar{\beta}_{k+1}^{t,i}$ as a measurable selection of

$$\begin{cases} \arg\min_{j\in\mathcal{I}^{-\beta_k^{t,i}}} \rho_{\bar{\tau}_{k+1}^{t,i}}\left(\tilde{g}_{\beta_k^{t,i},j}(\bar{\tau}_{k+1}^{t,i}) + Y^j_{\bar{\tau}_{k+1}^{t,i}+1}\right), & \tau_{k+1}^{t,i} < T, \\ \arg\min_{j\in\mathcal{I}} \rho_T\left(\tilde{g}_{\beta_k^{t,i},j}(T)\right), & \tau_{k+1}^{t,i} = T. \end{cases}$$

Letting $\mathcal{N} := \min\left\{k \geq 1 \colon \bar{\tau}_k^{t,i} \geq T\right\}$,

$$\bar{\xi}_s^{t,i} := i\mathbf{1}_{\left[-1,\tau_1^{t,i}\right)}(s) + \sum_{j=1}^{\mathcal{N}} \beta_j \mathbf{1}_{\left[\bar{\tau}_j^{t,i},\bar{\tau}_{j+1}^{t,i}\right)}(s) + \beta_{\mathcal{N}}\mathbf{1}_{\{s=T\}},$$

and arguing as above, we get that

$$\begin{aligned} Y_t^i &= \rho_{t,T}\left(\tilde{g}_{i,\bar{\xi}_t^{t,i}}(t), \ldots, \tilde{g}_{\bar{\xi}_{T-1}^{t,i},\bar{\xi}_T^{t,i}}(T)\right) \\ &\geq V_t^i. \end{aligned}$$

□

Given a strategy $\xi \in \mathcal{U}_t^i$, we can define its pairs of jump times $\tau_j \geq t$ and positions $\beta_j \in \mathcal{I}$ as follows:

$$\begin{aligned} \tau_1 &= \inf\left\{s \geq t \colon \xi_s \neq i\right\} \wedge T, \\ \beta_1 &= \xi_{\tau_1}, \\ &\vdots \\ \tau_{j+1} &= \inf\left\{s > \tau_j \colon \xi_s \neq \beta_j\right\} \wedge T, \\ \beta_{j+1} &= \xi_{\tau_{j+1}}. \end{aligned} \tag{2.8}$$

(Note that constant strategies $\xi_t \equiv i$ satisfy $\tau_j = T$ and $\beta_j = i$ for all j.)

We have the following characterisation of an optimal strategy.

Corollary 2.1. *A strategy $\xi \in \mathcal{U}_t^i$ is optimal for (2.1) if*

$$\begin{cases} A_{\tau_j}^{\beta_{j-1}} - A_{\tau_{j-1}}^{\beta_{j-1}} = 0, \\ Y_{\tau_j}^{\beta_{j-1}} = \rho_{\tau_j}\left(\tilde{g}_{\beta_{j-1},\beta_j} + Y_{\tau_j+1}^{\beta_j}\right), \end{cases} \tag{2.9}$$

where $\left\{\left(\tau_j, \beta_j\right)\right\}$ are the pairs of jump times and positions of ξ. If ρ has the strong sensitivity property (cf. Appendix A.1), then the condition (2.9) is also necessary for optimality.

Proof. Sufficiency: From the proof of Theorem 2.2 we have that

$$Y_t^i = \rho_{t,T}\big(\tilde{g}_{i,\xi_t}(t), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}\big), \tag{2.10}$$

and optimality follows by the fact that $Y_t^i = V_t^i$.

Necessity: Suppose $\xi \in \mathcal{U}_t^i$ is optimal for (2.1) and ρ is strongly sensitive. Let $\{(\tau_j, \beta_j)\}$ be the pairs of jump times and positions of ξ as defined in (2.8). Then, using (2.10) above, Lemma A.6, the RBSΔEs (2.5), and monotonicity of ρ, we have

$$\begin{aligned}
Y_t^i &= \rho_{t,T}\big(\tilde{g}_{i,\xi_t}(t), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}\big)\\
&= \rho_{t,\tau_1}\Big(g_i(t), \ldots, g_i(\tau_1 - 1), \rho_{\tau_1,T}\big(\tilde{g}_{i,\beta_1}(\tau_1), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}\big)\Big)\\
&\geq \rho_{t,\tau_1}\Big(g_i(t), \ldots, g_i(\tau_1 - 1), \rho_{\tau_1}\Big(\tilde{g}_{i,\beta_1}(\tau_1) + Y_{\tau_1+1}^{\beta_1}\Big)\Big)\\
&\geq \rho_{t,\tau_1}\Big(g_i(t), \ldots, g_i(\tau_1 - 1), Y_{\tau_1}^i\Big)\\
&\;\;\vdots\\
&\geq Y_t^i,
\end{aligned}$$

where we set $Y_{T+1}^j := 0$ for all $j \in \mathcal{I}$. We therefore have

$$\rho_{t,\tau_1}\Big(g_i(t), \ldots, g_i(\tau_1 - 1), \rho_{\tau_1}\Big(\tilde{g}_{i,\beta_1}(\tau_1) + Y_{\tau_1+1}^{\beta_1}\Big)\Big) = \rho_{t,\tau_1}\Big(g_i(t), \ldots, g_i(\tau_1 - 1), Y_{\tau_1}^i\Big),$$

and by strong sensitivity of ρ and the definition of A^i from Theorem 2.2, (2.9) is true for $j = 1$. The general case is proved by induction in a similar manner. □

2.3. The special case of optimal stopping

We now consider the problem of finding

$$F_t := \operatorname*{ess\,inf}_{\tau \in \mathscr{T}_{[t,T]}} \rho_{t,\tau}\big(f(t), \ldots, f(\tau - 1), h(\tau)\big), \tag{2.11}$$

for given sequences $\{f(t)\}_{t=0}^T$ and $\{h(t)\}_{t=0}^T$ in $\left(L_{\mathcal{F}}^\infty\right)^{T+1}$. This problem can be related to optimal switching with two modes $\mathcal{I} := \{1, 2\}$. The optimal stopping problem (2.11) is equivalent to (2.1) if we do the following:

- Set $g_1(t) = f(t)$ for $0 \leq t \leq T - 1$, $g_1(T) = h(T)$, $c_{1,2} \equiv h$, and $g_2 \equiv c_{2,1} \equiv 0$.
- *Mutatis mutandis* let $\mathcal{I}$ depend on the present mode. We then set $\mathcal{I}(1) := \{1, 2\}$ when we are in mode 1 and $\mathcal{I}(2) := \{2\}$ when we are in mode 2. In particular this gives $\mathcal{I}(2)^{-2} = \emptyset$ in (2.5). We additionally use the conventions $\min \emptyset = \infty$ and $-\infty \cdot 0 = \infty \cdot 0 = 0$.
- Optimise over strategies satisfying $\xi_{t-1} = 1$.

We note that in this setting the recursion (2.2) gives $V^2 \equiv 0$. The following result is then a direct consequence of Theorem 2.2:

Theorem 2.3. *The value process F for the optimal stopping problem satisfies*

$$\begin{cases} F_T = \rho_T(h(T)), \\ F_t = \rho_t(f(t) + F_{t+1}) \wedge \rho_t(h(t)), \quad 0 \leq t < T, \end{cases} \tag{2.12}$$

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

and the stopping time $\tau_t \in \mathscr{T}_{[t,T]}$ defined by

$$\tau_t = \inf\{t \le s \le T\colon F_s = \rho_s(h(s))\} \tag{2.13}$$

is optimal for (2.11)*. Furthermore, there exist a $\rho_{s,t}$-martingale M and a non-decreasing $\mathbb{G}$-predictable process A such that (F, M, A) is the unique solution to the following RBSΔE:*

$$\begin{cases} F_t = \rho_T(h(T)) + \sum_{s=t}^{T-1} \rho_s(f(s) + \Delta M_{s+1}) - (M_T - M_t) \\ \qquad -(A_T - A_t), \quad \forall\, t \in \mathbb{T}, \\ F_t \le \rho_t(h(t)), \quad \forall\, t \in \mathbb{T}, \\ \sum_{t=0}^{T-1} (F_t - \rho_t(h(t)))\Delta A_{t+1} = 0. \end{cases} \tag{2.14}$$

Remark 2.3. As is done in [8], Theorem 2.3 can be used to identify the optimal switching problem with a family of optimal stopping problems. If we set

$$h^i(t) = \begin{cases} \min_{j\in\mathcal{I}} \rho_T\big(\tilde{g}_{i,j}(T)\big), & t = T, \\ \min_{j\in\mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + \hat{V}^j_{t+1}\Big), & t < T, \end{cases}$$

and then substitute h^i for h in (2.12) and recall (2.2), it follows by Theorem 3 that for each $i \in \mathcal{I}$ and $t \in \mathbb{T}$ we have

$$\hat{V}^i_t = \operatorname*{ess\,inf}_{\tau\in\mathscr{T}_{[t,T]}} \rho_{t,\tau}\big(\tilde{g}_{i,i}(t), \ldots, \tilde{g}_{i,i}(\tau-1), h^i(\tau)\big).$$

3. Infinite-horizon risk-aware optimal switching under general filtration

In many problems the horizon T is so long that it can be considered infinite, and this motivates us to extend the results obtained in Section 2 to the infinite horizon. We thus let $\mathbb{T} := \mathbb{N}_0$ and define the infinite-horizon aggregated risk mapping $\varrho_s\colon \left(L^\infty_{\mathcal{F}}\right)^{\mathbb{T}} \to m\mathcal{G}_s$ (with $m\mathcal{G}_s$ the set of $\mathcal{G}_s$-measurable random variables) by

$$\varrho_s\big(W_s, W_{s+1}, \ldots\big) = \limsup_{t\to\infty} \rho_{s,t}\big(W_s, W_{s+1}, \ldots, W_t\big). \tag{3.1}$$

We define the value process for the switching problem on an infinite horizon as

$$V^i_t := \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_t} \varrho_t\big(\tilde{g}_{\xi_{t-1},\xi_t}(t), \tilde{g}_{\xi_t,\xi_{t+1}}(t+1), \ldots\big). \tag{3.2}$$

Definition 3.1. Let $\mathcal{L}^\infty_{\mathbb{G}} := \otimes_{t\in\mathbb{T}} L^\infty_{\mathcal{G}_t}$ and

$$\mathcal{L}^{\infty,d}_{\mathbb{G}} := \Big\{W \in \mathcal{L}^\infty_{\mathbb{G}}\colon \lim_{s\to\infty} \operatorname*{ess\,sup}_{\omega} |W_s(\omega)| = 0\Big\}.$$

Also, let $\mathcal{K}^+_d$ denote the set of all non-negative deterministic sequences $\{k_t\}_{t\in\mathbb{T}}$ such that the series $\sum_{t\in\mathbb{T}} k_t$ converges, and define

$$H_{\mathcal{F}} := \Big\{W \in \big(L^\infty_{\mathcal{F}}\big)^{\mathbb{T}}\colon \exists\{k_t\}_{t\in\mathbb{T}} \in \mathcal{K}^+_d \text{ such that } |W_t| \le k_t\ \forall t \in \mathbb{T}\Big\}.$$

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Remark 3.1. If $W \in H_{\mathcal{F}}$ then for every $s \in \mathbb{T}$ the limit

$$\varrho_s\big(W_s, W_{s+1}, \ldots\big) = \lim_{t\to\infty} \rho_{s,t}\big(W_s, \ldots, W_t\big)$$

exists almost surely and belongs to $L^\infty_{\mathcal{G}_s}$ (see Lemma A.2 in the appendix). An example $W \in H_{\mathcal{F}}$ is a discounted sequence $W_t = \alpha^t Z_t$ for some $\alpha \in (0, 1)$ and $\{Z_t\}_{t\in\mathbb{T}} \subset L^\infty_{\mathcal{F}}$ with $\sup_t |Z_t| < C$ for some $C \in (0, \infty)$.

Assumption 3.1. *There exists a sequence* $\{\bar{g}(t)\}_{t\in\mathbb{T}} \in H_{\mathcal{F}}$ *such that* $|\tilde{g}_{i,j}(t)| \le \bar{g}(t)$ *for all* $(t, i, j) \in \mathbb{T} \times \mathcal{I}^2$.

3.1. Dynamic programming equations

For $(t, r) \in \mathbb{T}^2$ we set $\hat{V}^i_{t,r} := \varrho_t\big(g_i(t), g_i(t+1), \ldots\big)$ whenever $t > r$ and define

$$\hat{V}^i_{t,r} = \min_{j\in\mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + \hat{V}^j_{t+1,r}\Big) \tag{3.3}$$

recursively for $t \le r$. By a simple induction argument we note that for each $i \in \mathcal{I}$ and $r \in \mathbb{T}$ the sequence $\big\{\hat{V}^i_{t,r}\big\}_{t\in\mathbb{T}}$ exists as a member of $\mathcal{L}^{\infty,d}_{\mathbb{G}}$. We have the following lemma.

Lemma 3.1. *For* $0 \le t \le r$ *and* $i \in \mathcal{I}$, *let* $\mathcal{U}^i_{t,r} := \big\{\xi \in \mathcal{U}^i_t : \xi_s = \xi_r,\ \forall s > r\big\}$. *Then*

$$\hat{V}^i_{t,r} = \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_{t,r}} \varrho_t\big(\tilde{g}_{\xi_{t-1},\xi_t}(t), \tilde{g}_{\xi_t,\xi_{t+1}}(t+1), \ldots\big).$$

Proof. This follows immediately from Lemma A.3 by applying Theorem 2.1 with cost sequence

$$\Big(\tilde{g}_{i,j}(t), \tilde{g}_{i,j}(t+1), \ldots, \tilde{g}_{i,j}(r-1), \tilde{g}_{i,j}(r) + \varrho_{r+1}\big(g_j(r+1), g_j(r+2), \ldots\big)\Big)_{i,j\in\mathcal{I}}, \tag{3.4}$$

noting that $\varrho_{r+1}\big(g_j(r+1), g_j(r+2), \ldots\big) \in L^\infty_{\mathcal{G}_{r+1}}$. □

We arrive at the following verification theorem.

Theorem 3.1. *The pointwise limits* $\big\{\tilde{V}^i_t\big\}_{t\in\mathbb{T},i\in\mathcal{I}} := \lim_{r\to\infty}\big\{\hat{V}^i_{t,r}\big\}_{t\in\mathbb{T},i\in\mathcal{I}}$ *exist and satisfy*

$$\tilde{V}^i_t = V^i_t, \quad \forall\, t \in \mathbb{T}.$$

Furthermore, starting from any $t \in \mathbb{T}$ *and* $i \in \mathcal{I}$, *the limit family defines an optimal strategy* $\xi^* \in \mathcal{U}^i_t$ *as follows:*

$$\begin{cases} \xi^*_r = i, & r < t, \\ \xi^*_r \in \arg\min_{j\in\mathcal{I}} \rho_r\Big(\tilde{g}_{\xi^*_{r-1},j}(r) + \tilde{V}^j_{r+1}\Big), & r \ge t. \end{cases}$$

Proof. From Lemma 3.1 and as $\mathcal{U}^i_{t,r} \subset \mathcal{U}^i_{t,r+1} \subset \mathcal{U}^i_t$ for all $0 \le t \le r$, the sequence $\big\{\hat{V}^i_{t,r}\big\}_{r\ge 0}$ is non-increasing and $\hat{V}^i_{t,r} \ge V^i_t$ for all $r \ge 0$. Furthermore, as it is bounded from below by the sequence $\big\{\varrho_t\big(-\bar{g}(t), -\bar{g}(t+1), \ldots\big)\big\}_{t\in\mathbb{T}}$ (owing to monotonicity) and $\big\{-\bar{g}(t)\big\}_{t\in\mathbb{T}} \in H_{\mathcal{F}}$, we conclude that the sequence $\Big\{\big\{\hat{V}^i_{t,r}\big\}_{t\in\mathbb{T},i\in\mathcal{I}}\Big\}_{r\ge 0}$ converges pointwise.

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Now, by Assumption 3.1 there is a non-negative decreasing deterministic sequence $\{K_s\}_{s\in\mathbb{T}}$, with $\lim_{s\to\infty} K_s = 0$, such that, for all $\xi \in \mathcal{U}^i_{t,r}$,

$$\begin{aligned}&\big|\varrho_{r+1}\big(\tilde{g}_{\xi_r,\xi_{r+1}}(r+1),\, \tilde{g}_{\xi_{r+1},\xi_{r+2}}(r+2),\, \ldots\big)\big| \\ &= \sum_{j\in\mathcal{I}} \mathbf{1}_{\{\xi_r=j\}}\big|\varrho_{r+1}\big(g_j(r+1),\, g_j(r+2),\, \ldots\big)\big| \le K_{r+1},\end{aligned} \tag{3.5}$$

and

$$\big|\varrho_{r+1}\big(-\bar{g}(r+1),\, -\bar{g}(r+2),\, \ldots\big)\big| \le K_{r+1}. \tag{3.6}$$

Then, by Lemma A.3, (3.5) gives

$$\begin{aligned}\hat{V}^i_{t,r} &= \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_{t,r}} \varrho_t\big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \tilde{g}_{\xi_t,\xi_{t+1}}(t+1),\, \ldots\big) \\ &\le \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_{t,r}} \rho_{t,r+1}\big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \ldots,\, \tilde{g}_{\xi_{r-1},\xi_r}(r),\, K_{r+1}\big) \\ &= \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_{t,r}} \rho_{t,r}\big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \ldots,\, \tilde{g}_{\xi_{r-1},\xi_r}(r)\big) + K_{r+1},\end{aligned}$$

and (3.6) implies that

$$\begin{aligned}V^i_t &= \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_t} \varrho_t\big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \tilde{g}_{\xi_t,\xi_{t+1}}(t+1),\, \ldots\big) \\ &\ge \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_t} \rho_{t,r+1}\big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \ldots,\, \tilde{g}_{\xi_{r-1},\xi_r}(r),\, -K_{r+1}\big) \\ &= \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_{t,r}} \rho_{t,r}\big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \ldots,\, \tilde{g}_{\xi_{r-1},\xi_r}(r)\big) - K_{r+1}.\end{aligned}$$

We conclude that $\hat{V}^i_{t,r} - V^i_t \le 2K_{r+1}$. Letting $r\to\infty$ gives the first statement.

For the second part, first note that the following inequality holds:

$$V^i_t \ge \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_{t,r}} \rho_{t,r}\Big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \ldots,\, \tilde{g}_{\xi_{r-1},\xi_r}(r) + V^{\xi_r}_{r+1}\Big), \qquad 0\le t\le r. \tag{3.7}$$

Indeed, for every $0\le t\le r$ and $\xi\in\mathcal{U}^i_t$ we can use Lemma A.3 to argue that

$$\begin{aligned}\varrho_t\Big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \tilde{g}_{\xi_t,\xi_{t+1}}(t+1),\, \ldots\Big) &= \rho_{t,r}\Big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \ldots,\, \tilde{g}_{\xi_{r-1},\xi_r}(r) + \hat{V}^{\xi_r}_{r+1,r}\Big) \\ &\ge \rho_{t,r}\Big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \ldots,\, \tilde{g}_{\xi_{r-1},\xi_r}(r) + V^{\xi_r}_{r+1}\Big) \\ &\ge \operatorname*{ess\,inf}_{\xi'\in\mathcal{U}^i_{t,r}} \rho_{t,r}\Big(\tilde{g}_{\xi'_{t-1},\xi'_t}(t),\, \ldots,\, \tilde{g}_{\xi'_{r-1},\xi'_r}(r) + V^{\xi'_r}_{r+1}\Big),\end{aligned}$$

and since this is true for every $\xi\in\mathcal{U}^i_t$ we get (3.7). Next, momentarily fix $0\le t\le r$ and replace $g_j(r)$ with $g_j(r) + V^j_{r+1}$. Then, using Theorem 2.1 with $T = r$, we have

$$\begin{aligned}V^i_t &\ge \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_{t,r}} \rho_{t,r}\Big(\tilde{g}_{\xi_{t-1},\xi_t}(t),\, \ldots,\, \tilde{g}_{\xi_{r-1},\xi_r}(r) + V^{\xi_r}_{r+1}\Big) \\ &= \rho_{t,r}\Big(\tilde{g}_{i,\xi^*_t}(t),\, \tilde{g}_{\xi^*_t,\xi^*_{t+1}}(t+1),\, \ldots,\, \tilde{g}_{\xi^*_{r-1},\xi^*_r}(r) + V^{\xi^*_r}_{r+1}\Big) \\ &\ge \rho_{t,r}\Big(\tilde{g}_{i,\xi^*_t}(t),\, \tilde{g}_{\xi^*_t,\xi^*_{t+1}}(t+1),\, \ldots,\, \tilde{g}_{\xi^*_{r-1},\xi^*_r}(r)\Big) - K_{r+1}.\end{aligned}$$

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Letting $r \to \infty$ we conclude that

$$V_t^i \geq \varrho_t\big(\tilde{g}_{i,\xi_t^*}(t), \tilde{g}_{\xi_t^*,\xi_{t+1}^*}(t+1), \ldots\big),$$

from which it follows that ξ^* is an optimal strategy. □

We also record the following corollary, which will be used in the proof of Theorem 3.2.

Corollary 3.1. *The value process for the infinite-horizon optimal switching problem (3.2) satisfies the following dynamic programming principle:*

$$V_t^i = \operatorname*{ess\,inf}_{\xi \in \mathcal{U}_{t,r}^i} \rho_{t,r}\Big(\tilde{g}_{\xi_{t-1},\xi_t}(t), \ldots, \tilde{g}_{\xi_{r-1},\xi_r}(r) + V_{r+1}^{\xi_r}\Big), \quad 0 \leq t \leq r.$$

Proof. We only need to prove that the following recursion holds:

$$V_t^i = \min_{j \in \mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + V_{t+1}^j\Big). \tag{3.8}$$

The general result then follows from Theorem 2.1 with $T = r$ and replacing $g_j(r)$ with $g_j(r) + V_{r+1}^j$. Taking limits on both sides in (3.3) gives

$$\begin{aligned} V_t^i &= \lim_{r\to\infty} \min_{j \in \mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + \hat{V}_{t+1,r}^j\Big) \\ &\leq \lim_{r\to\infty} \min_{j \in \mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + V_{t+1}^j + 2K_{r+1}\Big) \\ &= \lim_{r\to\infty} \Big\{\min_{j \in \mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + V_{t+1}^j\Big) + 2K_{r+1}\Big\} \\ &= \min_{j \in \mathcal{I}} \rho_t\Big(\tilde{g}_{i,j}(t) + V_{t+1}^j\Big). \end{aligned}$$

Since the reverse inequality follows as special case of (3.7), the proof is complete. □

3.2. Relation to systems of RBSΔEs

Definition 3.2 *(Infinite-horizon RBSΔEs).* The infinite-horizon extension of Definition 2.1 (with $\mathbb{T} = \mathbb{N}_0$) is given by

$$\begin{cases} Y_t^i = Y_T^i + \sum_{s=t}^{T-1} \rho_s\big(g_i(s) + \Delta M_{s+1}^i\big) - \big(M_T^i - M_t^i\big) \\ \qquad -\big(A_T^i - A_t^i\big), \quad \forall\, t, T \in \mathbb{T} \text{ with } t \leq T, \\ Y_t^i \leq \min_{j \in \mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + Y_{t+1}^j\Big), \quad \forall\, t \in \mathbb{T}, \\ \sum_{t \in \mathbb{T}} \Big(Y_t^i - \min_{j \in \mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + Y_{t+1}^j\Big)\Big)\Delta A_{t+1}^i = 0. \end{cases} \tag{3.9}$$

A solution to the system of RBSΔEs (3.9) is a triple $(Y, M, A) \in \mathcal{L}_{\mathbb{G}}^{\infty,d} \times \big(\mathcal{L}_{\mathbb{G}}^{\infty}\big)^2$ with M a $\rho_{s,t}$-martingale and A a $\mathbb{G}$-predictable non-decreasing process.

Remark 3.2. In the special case when the limits $M_\infty = \lim_{t\to\infty} M_t$ and $A_\infty = \lim_{t\to\infty} A_t$ exist $\mathbb{P}$-almost surely as members of $L^\infty_{\mathcal{G}}$, the infinite-horizon RBSΔE (3.9) can be written

$$\begin{cases} Y^i_t = \sum_{s=t}^{\infty} \rho_s\Big(g_i(s) + \Delta M^i_{s+1}\Big) - \big(M^i_\infty - M^i_t\big) \\ \qquad -\big(A^i_\infty - A^i_t\big), \quad \forall\, t \in \mathbb{T}, \\ Y^i_t \le \min_{j\in\mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + Y^j_{t+1}\Big), \quad \forall\, t \in \mathbb{T}, \\ \sum_{t\in\mathbb{T}} \Big(Y^i_t - \min_{j\in\mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + Y^j_{t+1}\Big)\Big)\Delta A^i_{t+1} = 0. \end{cases} \tag{3.10}$$

We also emphasise that $Y \in \mathcal{L}^{\infty,d}_{\mathbb{G}}$ implies the boundary condition $\lim_{T\to\infty} Y^i_T = 0$ for all $i \in \mathcal{I}$.

We have the following extension of Theorem 2.2.

Theorem 3.2. *The system of RBSΔEs* (3.9) *has a unique solution. Furthermore, the solution satisfies $Y = V$.*

Proof. Existence: By Corollary 3.1, the value process V satisfies the following dynamic programming relation for any $T \in \mathbb{T}$:

$$V^i_t = \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_{t,T}} \rho_{t,T}\Big(\tilde{g}_{\xi_{t-1},\xi_t}(t), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}(T) + V^{\xi_T}_{T+1}\Big), \quad 0 \le t \le T.$$

Using Theorem 2.2, this implies for every $T \in \mathbb{T}$ that (V, M, A) is the unique solution to

$$\begin{cases} V^i_t = V^i_T + \sum_{s=t}^{T-1} \rho_s\big(g_i(s) + \Delta M^i_{s+1}\big) - \big(M^i_T - M^i_t\big) - \big(A^i_T - A^i_t\big), \\ \qquad t = 0, \ldots, T, \\ V^i_t \le \min_{j\in\mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + V^j_{t+1}\Big), \quad t = 0, \ldots, T, \\ \sum_{t=0}^{T} \Big(V^i_t - \min_{j\in\mathcal{I}^{-i}} \rho_t\Big(\tilde{g}_{i,j}(t) + V^j_{t+1}\Big)\Big)\Delta A^i_{t+1} = 0, \end{cases}$$

where $M^i_0 = A^i_0 = 0$ and

$$\begin{cases} \Delta M^i_{t+1} = V^i_{t+1} - \rho_t\big(V^i_{t+1}\big), \\ \Delta A^i_{t+1} = \rho_t\big(g_i(t) + V^i_{t+1}\big) - V^i_t. \end{cases}$$

Furthermore, since this unique definition for the vector-valued processes M and A is independent of T, it follows that (V, M, A) satisfies Equation (3.9).

By the proof of Theorem 3.1, there exists a decreasing deterministic sequence $\{K_t\}_{t\in\mathbb{T}}$ such that $\big|V^i_T\big| \le K_T$ and $\lim_{T\to\infty} K_T = 0$. Therefore $V \in \mathcal{L}^{\infty,d}_{\mathbb{G}}$ and we conclude that (V, M, A) is a solution to (3.9).

Uniqueness: To show uniqueness, we note that if (Y, N, B) is any other solution to (3.9), then by again truncating at time $T \ge t$ and applying Theorem 2.2 we have that

$$Y^i_t = \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_t} \rho_{t,T}\Big(\tilde{g}_{\xi_{t-1},\xi_t}(t), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}(T) + Y^{\xi_T}_{T+1}\Big).$$

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Since $Y, V \in \mathcal{L}^{\infty,d}_{\mathbb{G}}$ and $\mathcal{I}$ is finite, we can define a deterministic sequence $\{K_s\}_{s\in\mathbb{T}}$ with $K_s \to 0$ as $s \to \infty$ such that $\left|V^i_t - Y^i_t\right| \le K_t$ for all $t \in \mathbb{T}$ and $i \in \mathcal{I}$. Appealing once more to the dynamic programming relation, we obtain

$$\begin{aligned} Y^i_t &\le \operatorname*{ess\,inf}_{\xi\in\mathcal{U}^i_t} \rho_{t,T}\Big(\tilde{g}_{\xi_{t-1},\xi_t}(t), \ldots, \tilde{g}_{\xi_{T-1},\xi_T}(T) + V^{\xi_T}_{T+1} + K_{T+1}\Big) \\ &= V^i_t + K_{T+1}, \end{aligned}$$

and similarly we have that $V^i_t \le Y^i_t + K_{T+1}$. Letting $T \to \infty$ we find that $V^i_t = Y^i_t$ for all $i \in \mathcal{I}$, and uniqueness follows. □

3.3. Relation to optimal stopping

As an extension to Section 2.3 above, we specialise to the case of optimal stopping on an infinite horizon:

$$F_t := \operatorname*{ess\,inf}_{\tau\in\mathscr{T}_t} \rho_{t,\tau}(f(t), \ldots, f(\tau-1), h(\tau)). \tag{3.11}$$

The above result for infinite-horizon optimal switching problems naturally extends the results in Section 2.3 on optimal stopping in finite horizon to infinite horizon. We have the following.

Corollary 3.2. *The value process F satisfies the dynamic programming relation*

$$F_t = \rho_t(f(t) + F_{t+1}) \wedge \rho_t(h(t))$$

for all $t \in \mathbb{T}$, and an optimal stopping time τ^*_t is given by

$$\tau^*_t := \inf\{s \ge t : F_s = \rho_s(h(s))\}.$$

Furthermore, there exists a $\rho_{s,t}$-martingale M and a non-decreasing $\mathbb{G}$-predictable process A such that (F, M, A) is the unique solution to the RBSΔE

$$\begin{cases} F_t = F_T + \sum_{s=t}^{T-1} \rho_s\big(f(s) + \Delta M_{s+1}\big) - \big(M_T - M_t\big) \\ \qquad -(A_T - A_t), \quad \forall\, t \in \mathbb{T} \text{ and } T \in \mathbb{T} \text{ with } t \le T, \\ F_t \le \rho_t(h(t)), \quad \forall\, t \in \mathbb{T}, \\ \sum_{t\in\mathbb{T}} \big(F_t - \rho_t(h(t))\big)\Delta A_{t+1} = 0. \end{cases} \tag{3.12}$$

Proof. This follows immediately from Theorem 3.1 through the analogy between optimal switching problems and optimal stopping problems described in Section 2.3. □

4. Example: delayed or missing observations

In this section we aim to add some colour to the above results by illustrating the interplay between delayed or missing observations and risk awareness. We demonstrate that this issue should be treated differently than in the case of linear expectation; otherwise suboptimal actions may result.

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space, and consider either the finite- or the infinite-horizon problem above. Let the process of essentially bounded costs $(\tilde{g}_{i,j}(t), i, j \in \mathcal{I})_{t\in\mathbb{T}}$ be adapted (and, in the infinite-horizon case, also satisfying Assumption 3.1), and let $\rho^{\mathbb{F}}$ be an $\mathbb{F}$-conditional risk mapping. Suppose that the observation at some time s is delayed. To model this, let $\mathbb{G}$ be the filtration given by

$$\mathcal{G}_t = \begin{cases} \mathcal{F}_{s-1}, & t = s, \\ \mathcal{F}_t, & \text{otherwise}, \end{cases}$$

and let ρ be the conditional risk mapping given by

$$\rho_t = \begin{cases} \rho^{\mathbb{F}}_{s-1}, & t = s, \\ \rho^{\mathbb{F}}_t, & \text{otherwise.} \end{cases}$$

Indeed, since we will examine the decision taken at time s rather than at later times, the observation at time s may equivalently be missing rather than delayed.

For any time $t \in \mathbb{T}$ with $t \neq s$, the value processes at time t are given by the dynamic programming equations (2.2) or (3.8) and conditional translation invariance,

$$\hat{V}^i_t = \min_{j \in \mathcal{I}} \left(\tilde{g}_{i,j}(t) + \rho_t\left(\hat{V}^j_{t+1} \right) \right), \tag{4.1}$$

while the missing observation at time s means that

$$\hat{V}^i_s = \min_{j \in \mathcal{I}} \rho_s\left(\tilde{g}_{i,j}(s) + \hat{V}^j_{s+1} \right), \tag{4.2}$$

$$\xi^i_s \in \arg\min_{j \in \mathcal{I}} \rho_s\left(\tilde{g}_{i,j}(s) + \hat{V}^j_{s+1} \right). \tag{4.3}$$

When $\rho^{\mathbb{F}}$ is the linear (conditional) expectation, this is equivalent to the following value and choice of mode:

$$\check{V}^i_s = \min_{j \in \mathcal{I}} \left(\rho_s(\tilde{g}_{i,j}(s)) + \rho_s\left(\hat{V}^j_{s+1} \right) \right), \tag{4.4}$$

$$\check{\xi}^i_s \in \arg\min_{j \in \mathcal{I}} \left(\rho_s(\tilde{g}_{i,j}(s)) + \rho_s\left(\hat{V}^j_{s+1} \right) \right). \tag{4.5}$$

The intuitively obvious fact that the selections (4.3) and (4.5) may differ can be confirmed by suitably modifying the costs at time s, as follows. For $f \in m\mathcal{F}$ define

$$\begin{cases} \check{C}_{i,j}(f) & = \rho_s\big(\tilde{g}_{i,j}(s)\big) + \rho_s\left(\hat{V}^j_{s+1} \right) - f, \\ \hat{C}_{i,j}(f) & = \rho_s\left(\tilde{g}_{i,j}(s) + \hat{V}^j_{s+1} \right) - f. \end{cases} \tag{4.6}$$

We assume that

$$\check{C}_{i,j}(0) > \hat{C}_{i,j}(0) \text{ for each } i, j \in \mathcal{I} \tag{4.7}$$

(which is true for example if the risk mapping $\rho^{\mathbb{F}}$ is subadditive), and that for some $l \in \mathcal{I}$ we have

$$\mathbb{P}\big(\check{C}_{l,1}(0) - \hat{C}_{l,1}(0) = \check{C}_{l,2}(0) - \hat{C}_{l,2}(0) \big) < 1, \tag{4.8}$$

setting $l = 1$ without loss of generality.

Remark 4.1. Clearly, these assumptions fail when $\rho^{\mathbb{F}}$ is linear (and in the finite-horizon case, they require that $s < T$). They can be understood as ensuring that $\rho^{\mathbb{F}}$ is 'sufficiently nonlinear' on the problem data. The inequality (4.7) serves to reduce combinatorial complexity.

We argue as follows:

1. Defining for each $n \in \mathbb{N}$ and $i = 1, 2$ the events
$$A_n^i = \left\{\omega \in \Omega : \check{C}_{1,3-i}(0) - \hat{C}_{1,3-i}(0) > \check{C}_{1,i}(0) - \hat{C}_{1,i}(0) + 2/n\right\}, \tag{4.9}$$
by the assumption (4.8) at least one of these events (A_n^1, say) has positive probability.

2. Setting $f_{1,1} = \check{C}_{1,1}(0) - \check{C}_{1,2}(0) + 1/n$, we have
$$\check{C}_{1,2}(0) = \check{C}_{1,1}(f_{1,1}) + 1/n. \tag{4.10}$$

3. We now further reduce combinatorial complexity by adjusting costs so that under both selections (4.3) and (4.5), when started in state $i = 1$ at time $s - 1$, at time s only either remaining in state 1 or switching to mode 2 can be optimal. That is, we would like the following to hold:
$$\begin{cases} \arg\min_{j \in \mathcal{I}} \left\{\hat{C}_{1,j}(\bar{f}_{1,j})\right\} \subset \{1, 2\}, \\ \arg\min_{j \in \mathcal{I}} \left\{\check{C}_{1,j}(\bar{f}_{1,j})\right\} \subset \{1, 2\}. \end{cases} \tag{4.11}$$
By straightforward linear algebra and (4.7), this can be achieved by taking
$$\begin{aligned} \bar{f} &= 1 + \operatorname{ess\,sup}\left\{\check{C}_{1,1}(f_{1,1}), \hat{C}_{1,1}(f_{1,1}), \check{C}_{1,2}(0), \hat{C}_{1,2}(0)\right\} - \operatorname*{ess\,inf}_{k>2}\left\{\check{C}_{1,k}(0), \hat{C}_{1,k}(0)\right\} \\ &= 1 + \operatorname{ess\,sup}\left\{\check{C}_{1,1}(f_{1,1}), \check{C}_{1,2}(0)\right\} - \operatorname*{ess\,inf}_{k>2}\left\{\hat{C}_{1,k}(0)\right\}. \end{aligned} \tag{4.12}$$

4. Finally, to observe a difference between the selections (4.3) and (4.5), set
$$\bar{f}_{1,j} = \begin{cases} \bar{f} + f_{1,1}, & j = 1, \\ \bar{f}, & j = 2, \\ 0, & j > 2, \end{cases} \tag{4.13}$$
since then on A_n^1 we have
$$\check{C}_{1,2}(\bar{f}_{1,2}) > \check{C}_{1,1}(\bar{f}_{1,1}) > \hat{C}_{1,1}(\bar{f}_{1,1}) > \hat{C}_{1,2}(\bar{f}_{1,2}), \tag{4.14}$$
where the first inequality comes from combining (4.6), (4.10), and (4.13), the second from (4.7), and the third from combining (4.9) with (4.10).

Noting that $\bar{f}_{1,1}$ and $\bar{f}_{1,2}$ are $\mathcal{G}_s$-measurable, we see from (4.6) that the two selections differ if we modify just the two costs $\tilde{g}_{1,1}(s)$ and $\tilde{g}_{1,2}(s)$ by replacing $\tilde{g}_{1,j}(s)$ with $\tilde{g}_{1,j}(s) - \bar{f}_{1,j}$ for $j \in \{1, 2\}$. In particular, from (4.11) and (4.14), if the system is in mode 1 at time $s - 1$, then on A_n^1, (4.5) selects mode 2 at time s while (4.3) selects mode 1 at time s.

5. Example: a hydropower planning problem

In this section we first illustrate the above framework for risk-aware optimal switching under general filtration by formulating a non-Markovian hydropower planning problem (Sections 5.1–5.4). In Sections 5.5–5.8 we provide practical dynamic programming equations, an approximate numerical scheme for the problem, a solution algorithm using neural networks, and a discussion of numerical results.

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

5.1. Decision space and market

Consider a hydropower producer whose interventions take the form of bidding into a market. The producer sells electricity in a daily spot market at noon on the day before delivery. Let $T = 9$, $\mathbb{T} := \{0, \ldots, T\}$, and $\mathbb{T}^+ := \{0, \ldots, T + 1\}$. Here, $t \in \mathbb{T} \cup \{-1\}$ represents a decision epoch at hour 12 of day t, where day -1 is the last day of the previous planning period. We assume one-hour planning periods, so that at decision epoch $t \in \mathbb{T}$, the producer hands in a list of bids $B_t := \big(B^E_{t+1,1}, \ldots, B^E_{t+1,24}; B^P_{t+1,1}, \ldots, B^P_{t+1,24}\big)$, where $B^E_{t+1,l}$ specifies the quantity of electrical energy offered and $B^P_{t+1,l}$ the acceptable price for hour l of day $t + 1$. Just after decision epoch t, the market clears and the prices of electricity are published. If the market price $R_{t+1,l}$ of electricity for hour l exceeds the producer's bid price $B^P_{t+1,l}$, the producer is obligated to deliver the bidden volume $B^E_{t+1,l}$ of electrical energy during hour l of day $t + 1$. For this the producer receives a payment $R_{t+1,l}B^E_{t+1,l}$. The total income arising from the bid vector B_t made at decision epoch t is thus given by

$$\sum_{l=1}^{24} \mathbf{1}_{\left\{B^P_{t+1,l} \le R_{t+1,l}\right\}} R_{t+1,l} B^E_{t+1,l}. \tag{5.1}$$

If, on the other hand, a bid is accepted and the reservoir contains insufficient water to deliver the bidden volume, the producer has to purchase the undelivered energy from the balancing power market at a price R^F, which is usually higher than the spot price. This induces the cost

$$\sum_{l=1}^{24} \mathbf{1}_{\left\{B^P_{t+1,l} \le R_{t+1,l}\right\}} R^F_{t+1,l} \big(B^E_{t+1,l} - E_{t+1,l}\big)^+ \tag{5.2}$$

of undelivered energy, where $E_{t,l}$ is the electrical energy produced during hour $l \in \{1, \ldots, 24\}$ of day t.

5.2. Probability space, inflow and price processes

We take $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}^+}, \mathbb{P})$ to be a filtered probability space, with $\mathcal{F}_t$ representing the information available at noon on day $t \in \mathbb{T}^+$. This space will be rich enough to support a Markovian price process $(\tilde{R}_t)_{t \in \mathbb{T}^+}$ and a non-Markovian inflow process $(\tilde{I}_t)_{t \in \mathbb{T}^+}$, as follows.

As is common in electricity planning problems, we assume that the electricity price vector $(\tilde{R}_t)_{t \in \mathbb{T}^+} := (R_{t,1}, \ldots, R_{t,24})_{t \in \mathbb{T}^+}$ is a bounded Markov process adapted to $\mathbb{F}$. Regarding the inflow process, even under normal conditions, heavy rainfall only leads to increased inflows to a reservoir after a time delay, as the water is filtered through the catchment area surrounding the reservoir. Moreover, the hydropower station may be located in a mountainous area where river flows depend heavily on the melting of snow masses in a spring flood. To model the discrete-time process of inflows $\{I_{t,j}\}_{t \in \mathbb{T}}^{1 \le j \le 24}$, where $I_{t,l}$ is the inflow of water from the surroundings during hour l of day t, let $(H_s)_{s \ge 0}$ be a continuous-time Markov process representing relevant environmental conditions. To account for the dependence of inflows on environmental conditions, set

$$I_{t,j} = \int_{t+(j-1)/24}^{t+j/24} \int_0^{\delta} h(s) H_{(r-s)^+} \, ds \, dr, \tag{5.3}$$

where δ is a constant time lag and h a deterministic function. Then $(\tilde{I}_t)_{t \in \mathbb{T}^+} := (I_{t-1,13}, \ldots, I_{t,12})_{t \in \mathbb{T}^+}$ is adapted to $\mathbb{F}$ and non-Markovian.

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

5.3. Dynamics of the hydropower system

We assume that the hydropower system consists of one reservoir containing the volume $M_{t,l}$ at the beginning of hour l of day t and a plant that produces electricity

$$E_{t,l} := \eta\big(M_{t,l}, F_{t,l}\big), \tag{5.4}$$

where $F_{t,l}$ is the flow of water directed through the turbines and $\eta : \mathbb{R}^2_+ \to [0, C]$ is a deterministic function describing the efficiency of the plant, with $C > 0$ the installed capacity. We assume that the function $y \mapsto \eta(m, y)$ is strictly increasing for each fixed m lying between the reservoir minimum level $M_{\min}$ and maximum $M_{\max}$. The process $M = (M_{t,l})_{t,l}$ of reservoir levels follows the dynamics

$$\begin{aligned} M_{t,l} = \min \big\{ &\mathbf{1}_{[l>1]}\big(M_{t,l-1} - F_{t,l-1} + I_{t,l-1}\big) \\ &+ \mathbf{1}_{[l=0]}\big(M_{t-1,24} - F_{t-1,24} + I_{t-1,24}\big), M_{\max} \big\}, \end{aligned} \tag{5.5}$$

where $M_{0,13}$ is the volume in the reservoir at the first decision epoch.

Also, as explained in [12], changing the production level by altering the flow $F_{t,l}$ may necessitate the startup or shutdown of turbines, resulting in both wear and tear and temporarily decreased efficiency. This feature motivates the inclusion of switching costs in the optimisation problem.

5.4. The optimisation problem

The controllable parameters in the problem are the bid vectors $\{B_t\}_{t\in\mathbb{T}}$. With the reasonable assumption that these bids take values in a finite set $\mathcal{I} \subset \mathbb{R}^{48}$ we have a switching problem. Let $\xi := (\xi_t)_{t\in\mathbb{T}}$ denote the switching control, so that $\xi_t = B_t$ for each $t \in \mathbb{T}$.

By inverting η, from the production plan and the reservoir level we obtain the flow

$$F_{t+1,l} = \min \Big(f\Big(B^E_{t+1,l}, M_{t+1,l}\Big)\mathbf{1}_{\big\{B^P_{t+1,l}\le R_{t+1,l}\big\}}, M_{t+1,l} - M_{\min}\Big). \tag{5.6}$$

Substituting (5.6) into (5.5) we see that M_t depends both on ω and on the entire history of ξ up to time t. It follows that the switching costs are also dependent on this history. Therefore, recalling (5.1)–(5.6) and letting $\mathcal{I}_t := (\mathcal{I})^{t+1}$, for $(i_{-1}, \dots, i_{t-1}, i_t) \in \mathcal{I}_{t+1}$ we may define the rewards for the planning problem as

$$\begin{aligned} \tilde{g}_{i_{-1:t-1},i_{-1:t}}(t) := \; & -c_{i_{-1:t}}\big(\tilde{R}_{t+1}\big) + \sum_{l=1}^{24} \mathbf{1}_{\big\{i_{t,24+l}\le R_{t+1,l}\big\}}\big(R_{t+1,l}i_{t,l} \\ & - R^F_{t+1,l}\big(i_{t,l} - \eta\big(M^{i_{-1:t}}_{t+1,l}, \min\big(f(i_{t,l}, M^{i_{-1:t}}_{t+1,l}), M^{i_{-1:t}}_{t+1,l} - M_{\min}\big)\big)\big)^+\big) \\ & + \mathbf{1}_{\{t=T\}}R^M M^{i_{-1:T}}_{T+2,1}, \end{aligned} \tag{5.7}$$

where

- $i_{-1:t} := (i_{-1}, \dots, i_{t-1}, i_t)$;
- $M^{i_{-1:t}}_{t+1,l}$ is the reservoir level at hour l on day $t+1$ corresponding to the bid history $i_{-1:t} \in \mathcal{I}_{t+1}$;
- $i_{t,m}$ is the mth component of i_t;

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

- R^M is the value of water stored at the end of the planning period;
- and for each $r \in \mathbb{R}_+^{24}$, $c_{i_{-1:t}}(r)$ is the cost rendered by switching from bid i_{t-1} to i_t when the price vector is r and the bid history is $i_{-1:t}$.

If the producer has risk mapping ρ, then for each $t \in \mathbb{T}$, given a bid history $i_{-1:t-1} \in \mathcal{I}_t$, the objective is to find

$$V_t^{i_{-1:t-1}} := \operatorname*{ess\,sup}_{\xi \in \mathcal{U}_t^{i_{-1:t-1}}} \rho_{t,T}\left(\tilde{g}_{\xi_{-1:t-1},\xi_{-1:t}}(t), \tilde{g}_{\xi_{-1:t},\xi_{-1:t+1}}(t+1), \ldots, \tilde{g}_{\xi_{-1:T-1},\xi_{-1:T}}(T)\right), \tag{5.8}$$

where $\mathcal{U}_t^{i_{-1:t-1}}$ is the set of $\mathbb{F}$-adapted, $\mathcal{I}$-valued processes $(\xi_s)_{s\in\mathbb{T}}$ such that $\xi_{-1:t-1} = i_{-1:t-1}$. Note that the reward $\tilde{g}_{i_{-1:t-1},i_{-1:t}}(t)$ is $\mathcal{F}_{t+1}$-measurable but not $\mathcal{F}_t$-measurable. The producer's problem is thus one of non-adapted (in this case, delayed) information.

5.5. Dynamic programming equations

By modifying the proof of Theorem 2.1 accordingly we can show that the value processes $(V_t^{i_{-1:t-1}} : i_{-1:t-1} \in \mathcal{I}_t)_{t\in\mathbb{T}}$ corresponding to (5.8) satisfy the following analogue of (2.2):

$$\begin{cases} V_T^{i_{-1:T-1}} = \max_{j\in\mathcal{I}} \rho_T\left(\tilde{g}_{i_{-1:T-1},(i_{-1:T-1},j)}(T)\right), \\ V_t^{i_{-1:t-1}} = \max_{j\in\mathcal{I}} \rho_t\left(\tilde{g}_{i_{-1:t-1},(i_{-1:t-1},j)}(t) + V_{t+1}^{(i_{-1:t-1},j)}\right), \quad \text{for } 0 \le t < T, \end{cases} \tag{5.9}$$

where for $i_{-1:t-1} \in \mathcal{I}_t$ and $j \in \mathcal{I}$ we define $(i_{-1:t-1}, j) = (i_{-1}, \ldots, i_{t-1}, j)$. In order to obtain a practical solution algorithm we observe that the same optimal control can be obtained by dynamic programming without requiring the entire bid history. Recalling (5.7), given $\omega \in \Omega$, for $(i_{-1}, \ldots, i_{t-1}, i_t) \in \mathcal{I}_{t+1}$ the cost $\tilde{g}_{i_{-1:t-1},i_{-1:t}}(t)$ depends on $i_{-1:t-1}$ only through its final bid vector i_{t-1} and the reservoir level $M_{t+1,1}^{i_{-1:t}}$. Moreover, by (5.5) and (5.6), $M_{t+1,1}^{i_{-1:t}}$ depends on $i_{-1:t-1}$ only through $M_{t,13}^{i_{-1:t-1}}$ and the final bid vector i_{t-1}. Thus for $i_{-1:t} \in \mathcal{I}_{t+1}$ and $m \in [M_{\min}, M_{\max}]$ we may define new (random) rewards $\tilde{g}_{i_{t-1},i_t}(t, m)$ such that

$$\begin{aligned} \tilde{g}_{i_{t-1},i_t}(t,m) := & -c_{i_{-1:t}}(R_{t+1}) + \sum_{l=1}^{24} \mathbf{1}_{\left\{i_{t,24+l} \le R_{t+1,l}\right\}} \Big(R_{t+1,l} i_{t,l} \\ & - R_{t+1,l}^F \left(i_{t,l} - \eta\left(M_{t+1,l}^{m,i_t}, \min\left(f\left(i_{t,l}, M_{t+1,l}^{m,i_t}\right), M_{t+1,l}^{m,i_t} - M_{\min}\right)\right)\right)^+\Big) \\ & + \mathbf{1}_{\{t=T\}} R^M M_{T+2,1}^{i_{-1:T}}, \end{aligned} \tag{5.10}$$

where M_{t+1}^{m,i_t} is the vector of reservoir levels on day $t+1$ given that on day t the reservoir was at level m at the beginning of hour 13 (i.e. at noon) and the bid vector was i_t. That is, $\tilde{g}_{i_{t-1},i_t}(t, m)$ and $\tilde{g}_{i_{-1:t-1},i_{-1:t}}(t)$ coincide when $M_{t,13}^{i_{-1:t-1}} = m$. Then define auxiliary value processes by

$$V_t^i(m) = \begin{cases} \max_{j\in\mathcal{I}} \rho_T\left(\tilde{g}_{i,j}(T, m)\right), \\ \max_{j\in\mathcal{I}} \rho_t\left(\tilde{g}_{i,j}(t,m) + V_{t+1}^j\left(M_{t+1,13}^{m,j}\right)\right), \quad \text{for } 0 \le t < T. \end{cases} \tag{5.11}$$

By construction we have $V_t^{i_{t-1}}\left(M_{t,13}^{i_{-1:t-1}}\right) = V_t^{i_{-1:t-1}}$; this can be confirmed by backward induction. Therefore, if the auxiliary value function $V_{t+1}^j(m)$ can be computed for each $j \in \mathcal{I}$

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

and $m \in [M_{\min}, M_{\max}]$, then (5.9) and (5.11) provide equivalent dynamic programming equations over the set of modes $\mathcal{I}$. The benefit of (5.11) is that we do not need to remember the switching control's entire history. Note that this reformulation is non-Markovian since $(M_t)_{t\in\mathbb{T}}$ is not a Markov process. In the next section we present a numerical approximation to this scheme using neural networks.

5.6. Numerical scheme

Let $\eta(M, F) = \eta_0 MF$ with $\eta_0 = 0.1$ and $R_{t,l} = \big(1 + |\sin(l\pi/12)|\big)\big(0 \vee \tilde{R}_{t+(l-1)/24} \wedge C_R\big)$, where the multiplicative coefficient models the daily trend, $C_R = 4$ is a price ceiling, and $\tilde{R}$ solves the stochastic difference equation

$$\tilde{R}_{t+1} - \tilde{R}_t = 0.02\big(1 - \tilde{R}_t\big) + 0.05N_t,$$

where $(N_t)_{t\in\mathbb{T}}$ are standard normal random variables.

For the processes I and H of (5.3) we take $\delta = 2/24$, $h(s) := \sin(s\pi/\delta)$, and H to be a pure jump Markov process taking values in $\{0, 0.5, 1\}$ with transition intensity matrix

$$Q_H := \begin{bmatrix} -1 & 0.5 & 0.5 \\ 1 & -2 & 1 \\ 2 & 0.5 & -2.5 \end{bmatrix},$$

representing no, medium, and heavy rainfall respectively. For numerical purposes we approximate H by a discrete-time Markov chain updating k times per hour, with transition matrix $\exp\left(\frac{1}{24k}Q_H\right)$.

Moreover, let $\tilde{I}$ be a discretisation of the set $[0, 2]^{24} \times [0, 4]^{24}$ (representing the fact that market bids have limited precision, for example 1 MWh and 0.01 euro), and let the hydropower producer's risk aversion be modelled by an entropic risk measure, i.e.

$$\rho_t(X) = -\frac{1}{\theta}\log\left(\mathbb{E}\big[e^{-\theta X}\big|\mathcal{F}_t\big]\right),$$

with parameter $\theta > 0$. Finally, we assume that changes in production level cost 0.1 Euro per MW and set $R^F = 10$, $R^M := 4$, $M_{\min} = 10$, $M_{\max} = 50$, and $k = 2$.

5.6.1. *State-space description.* To obtain a state-space description of our problem we introduce the state $(x_t)_{t\in\mathbb{T}}$, where x_t is the non-redundant information available at hand at noon on day t; that is,

$$x_t := \begin{bmatrix} M_{t,13} \\ R_{t,24} \\ \{P_{t,j}\}_{13\le j\le 24} \\ \left\{H^k_{t+1/2-l/24k}\right\}_{l\in\{0,\ldots,24\delta k\}} \end{bmatrix}, \tag{5.12}$$

where $\{P_{t,j}\}_{13\le j\le 24}$ is the $\mathcal{F}_t$-measurable production schedule for the hours between noon and midnight of day t. In particular, the state contains the discretised weather trajectory for the past two hours (10 a.m. to noon), since, according to (5.3), the impact of precipitation is only fully revealed after this delay. Recalling the notation ξ^* of Theorem 2.1 for an optimal strategy, from

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

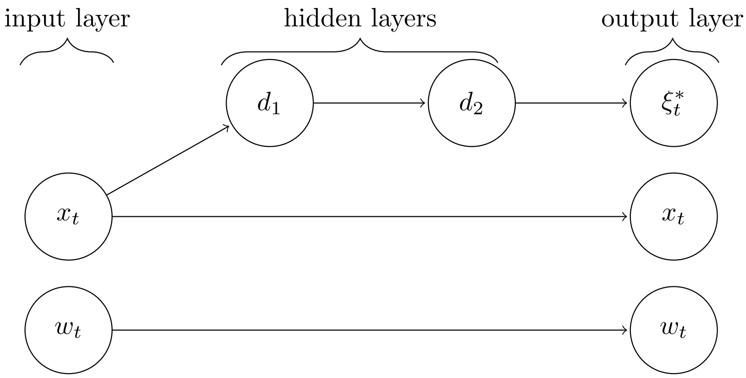

FIGURE 1: Architecture of the bid neural network. Nodes d_1 and d_2 represent dense layers with sigmoid activation function.

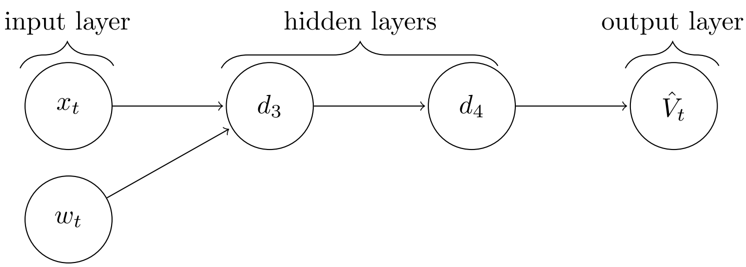

FIGURE 2: Architecture of the value neural network. To reduce dimension, in the state vector x_t the production schedule $\{P_{t,j}\}_{13\le j\le 24}$ is replaced by the sum of its entries. Nodes d_3 and d_4 represent dense layers with sigmoid activation function.

Section 5.5 the optimal mode (bid vector) ξ_t^* depends on its previous value ξ_{t-1}^* only through the production schedule $\{P_{t,j}\}_{13\le j\le 24}$, so we may write $\xi_t^* = \xi_t^*(x_t)$.

It follows from Equations (5.5)–(5.7) and (5.3) that given the system state x_t and bid vector $j = B_t$ at noon on day t, both the reward $\tilde{g}_{i,j}(t)$ and the new state x_{t+1} are measurable with respect to the noise vector w_t, where

$$w_t := \begin{bmatrix} \left\{R_{t+1,j}\right\}_{1\le j\le 24} \\ \left\{H^k_{t+1/2-l/24k}\right\}_{l\in\left\{0,\ldots,24\delta k\right\}} \end{bmatrix},$$

which is not $\mathcal{F}_t$-measurable.

5.7. Algorithm

In this section we describe an implementation of the numerical scheme of Section 5.6. Code implementing this scheme, and also a risk-neutral scheme, is available at https://github.com/moriartyjm/optimalswitching/tree/main/hydro and is described in Algorithm 1. For practicality it employs the neural networks shown in Figures 1 and 2.

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Algorithm 1. Hydropower planning over T days

Input : M independently sampled price and weather trajectories and initial reservoir levels

for t **in** $\{T, T-1, \ldots, 0\}$ **do**

  Train bid NN for day t with states $(x_t^i)_{i=1,\ldots,M}$ and noise $(w_t^i)_{i=1,\ldots,M}$, store as `model_bids[t]`, and predict optimal bid vectors $((\xi_t^*)^i)_{i=1,\ldots,M}$

  Train value NN for day t with states $(x_t^i)_{i=1,\ldots,M}$, bid vectors $((\xi_t^*)^i)_{i=1,\ldots,M}$, and noise $(w_t^i)_{i=1,\ldots,M}$, and store as `mdl_E_exp[t]`

end

Output: Neural networks `model_bids`, `mdl_E_exp` approximating optimal bid vector and value for each day t, state x_t

The bid neural network, whose architecture is given in Figure 1, aims to solve the following optimisation problem:

$$\begin{cases} \xi_T^*(x_T) \in \arg\max_{j\in\mathcal{I}} \left\{-\frac{1}{\theta}\log\left(E_T^{x_T}\left[e^{-\theta\tilde{g}_{i,j}(T)}\right]\right)\right\}, \\ \xi_t^*(x_t) \in \arg\max_{j\in\mathcal{I}} \left\{-\frac{1}{\theta}\log\left(E_t^{x_t}\left[e^{-\theta\left(\tilde{g}_{i,j}(t)+\hat{V}^j_{t+1}(x_{t+1})\right)}\right]\right)\right\}, \quad \text{for } 0 \le t < T, \end{cases} \tag{5.13}$$

where $x_t \mapsto E_t^{x_t}$ approximates the conditional expectation with respect to $\mathcal{F}_t$ using the state vector, and $\hat{V}^j_{t+1}(x_{t+1})$ approximates the continuation value using the value neural network, whose architecture is given in Figure 2. Continuation values $\hat{V}^j_{T+1}(x_{T+1})$ are set equal to zero. Note that these equations do not simplify further since the rewards $\tilde{g}_{ij}(t)$ are non-adapted.

The optimisation is performed by first training the bid neural network on M independent noise realisations with target values equal to zero and loss function equal to

$$-\frac{1}{\theta}\log\left(\frac{1}{M}\sum_{\ell=1}^{M}\left[e^{-\theta\left(\tilde{g}_{i,\xi_t^*\left(x_t^\ell\right)}(t)+\hat{V}^j_{t+1}\left(x^\ell_{t+1}\right)\right)}\right]\right),$$

where x_t^ℓ denotes the state vector x_t under the ℓth noise realisation. (Note that since the state x_t^ℓ contains the production schedule $\left\{P^\ell_{t,j}\right\}_{13\le j\le 24}$, it also depends on the bid vector submitted at time $t-1$; we omit this dependency in order to lighten the notation.) After the bid neural network has been trained, the value neural network is trained on the M independent noise realisations with target values equal to $\exp\left(-\theta\left(\tilde{g}_{i,\xi_t^*\left(x_t^\ell\right)}(t)+\hat{V}^j_{t+1}\left(x^\ell_{t+1}\right)\right)\right)$ and the mean squared error as the loss function. Initial reservoir levels $M_{0,13}$ are drawn uniformly at random between $M_{\min}$ and $M_{\max}$, while initial market prices $R_{0,24}$ and weather values $\left\{H^k_{t+1/2-l/24k}\right\}_{l\in\{0,\ldots,24\delta k\}}$ are drawn from the corresponding stationary distribution.

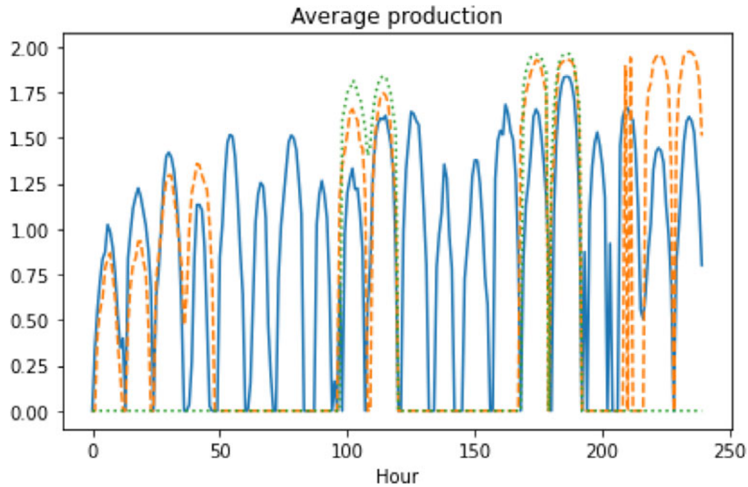

FIGURE 3: Average production curves by hour for risk sensitivity $\theta = 0,\ 0.01,\ 0.02$ (blue solid, orange dashed, and green dotted lines, respectively).

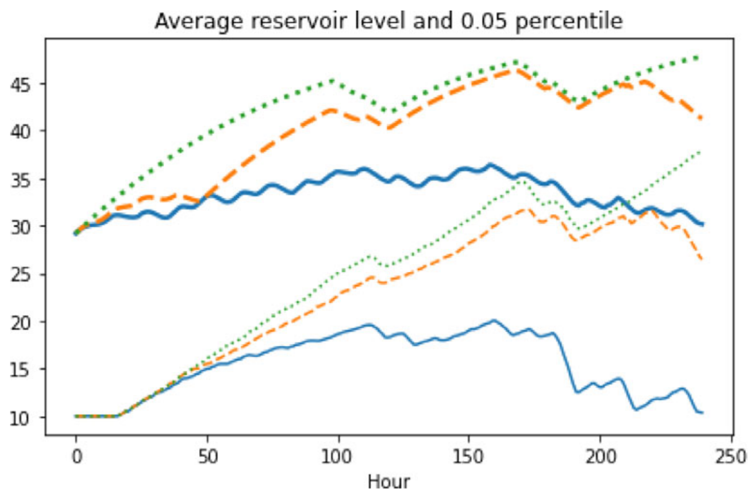

FIGURE 4: Plots for risk sensitivity $\theta = 0$, 0.01, 0.02 of the reservoir level M_t by hour: mean (thick blue solid, orange dashed, and green dotted lines, respectively) and 0.05 percentile (thinner lines).

5.8. Numerical results and discussion

In this section we present and discuss numerical results obtained using Algorithm 1 over an optimisation horizon of 10 days and with 50,000 independent noise realisations. Identifying the risk-neutral case with $\theta = 0$, we plot results for θ equal to 0, 0.01, and 0.02 in blue (solid), orange (dashed), and green (dotted), respectively.

For each hour in the optimisation, Figure 3 shows the production level under the respective optimal strategies, averaged across all noise realisations. Similarly, Figure 4 plots the mean water level under the optimal strategies, together with the 0.05 percentiles (dashed lines). In order to represent the value processes, Figure 5 plots the prediction $\hat{V}_0(x_0)$ made by the value

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

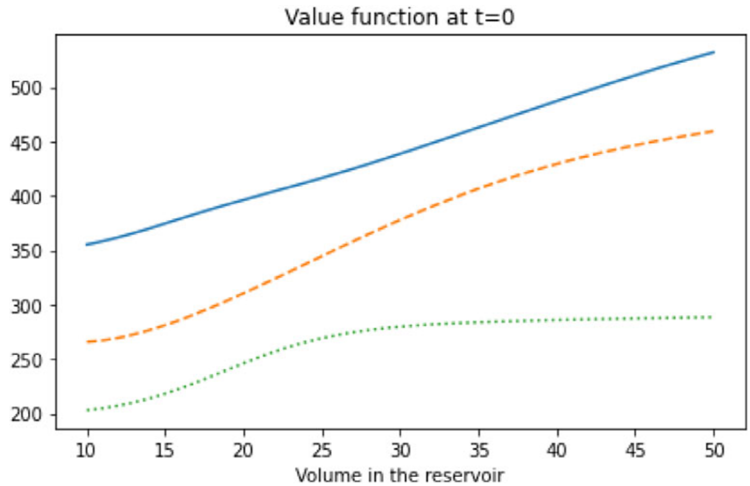

FIGURE 5: Predictions made by the $t = 0$ value neural network for $\theta = 0, 0.01, 0.02$ (blue solid, orange dashed, and green dotted lines, respectively) for the input $x_0 = [m, 0, \mathbf{0}, \mathbf{0}]^\top$, for different initial reservoir levels m.

neural network when the input is $x_0 = [m, 0, \mathbf{0}, \mathbf{0}]^\top$, for different values of m (recall (5.12); $\top$ denotes transpose).

The reservoir's physical constraints $M_{\min}$ and $M_{\max}$ create risks for the hydropower producer. When the reservoir level is near $M_{\min}$ the producer risks being unable to fulfil the bid volume and receiving a penalty for under-production. Conversely, if the reservoir reaches its maximum level $M_{\max}$ then she risks spilling the water inflow, which would otherwise be stored and used profitably later.

From Figure 4, the risk-neutral producer maintains the reservoir at an intermediate water level on average. Furthermore, in at least 5% of cases she allows the water level to fall rather close to the minimum level. In contrast, in at least 95% of cases the optimal strategy of the risk-averse producer first drives the initial water level up by trading less, and production increases only once the reservoir is at least approximately half filled. Indeed, for $\theta = 0.02$ the average water level is seen to increase towards $M_{\max}$ over the time horizon. Thus increases in θ incentivise the producer to avoid the risk of under-production penalties. (The risk of spilling water at level $M_{\max}$ appears to have comparatively less influence on the optimal strategies.)

These observations are also borne out in Figure 5. In the risk-neutral case, the marginal value of water is approximately constant as the water level varies. However, locally around $M_{\min}$, where the risk of penalties has more influence, the marginal value of water becomes lower as the risk sensitivity parameter θ increases.

Figure 3 confirms that the risk-neutral strategy involves producing every day, and also involves following the daily price trend within each day. As the risk aversion parameter θ increases, the number of production days, and also the total produced volume, decrease.

Appendix A. Properties of conditional risk mappings

Here we review definitions and preliminary results on conditional risk mappings that are used in the main text. References for this material include [9, 7, 13, 19, 18, 5, 17, 8], among many others. Proofs are provided for results if they are not readily available in these references.

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

We are given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\mathbb{G} = \{\mathcal{G}_t\}_{t\in\mathbb{T}}$ of sub-σ-algebras of $\mathcal{F}$. All random variables below are defined with respect to this probability space, and (in-)equalities between random variables are in the $\mathbb{P}$-almost-sure sense.

A.1. Conditional risk mappings

A $\mathbb{G}$-conditional risk mapping is a family of mappings $\{\rho_t\}_{t\in\mathbb{T}}$, $\rho_t\colon L^\infty_{\mathcal{F}} \to L^\infty_{\mathcal{G}_t}$, satisfying the following for all $t \in \mathbb{T}$:

Normalisation: $\rho_t(0) = 0$.

Conditional translation invariance: for all $W \in L^\infty_{\mathcal{F}}$ and $Z \in L^\infty_{\mathcal{G}_t}$,

$$\rho_t(Z + W) = Z + \rho_t(W).$$

Monotonicity: for all $W, Z \in L^\infty_{\mathcal{F}}$,

$$W \le Z \implies \rho_t(W) \le \rho_t(Z).$$

For each $t \in \mathbb{T}$ we refer to ρ_t as a conditional risk mapping. Note that in contrast to the one-step conditional risk mappings ρ_t of [17], whose respective domains would be $L^\infty_{\mathcal{G}_{t+1}}$ in this context, here the domain of each ρ_t is $L^\infty_{\mathcal{F}}$. Conditional risk mappings and the *monetary conditional risk mappings* of [8] are interchangeable via the mapping $Z \mapsto \rho_t(-Z)$. Each $\mathbb{G}$-conditional risk mapping satisfies the following property (cf. [4, Proposition 3.3], [8, Exercise 11.1.2]):

Conditional locality: for every W and Z in $L^\infty_{\mathcal{F}}$, $t \in \mathbb{T}$, and $A \in \mathcal{G}_t$,

$$\rho_t(\mathbf{1}_A W + \mathbf{1}_{A^c} Z) = \mathbf{1}_A \rho_t(W) + \mathbf{1}_{A^c} \rho_t(Z).$$

A $\mathbb{G}$-conditional risk mapping is said to be strongly sensitive if it satisfies the following:

Strong sensitivity: for all $W, Z \in L^\infty_{\mathcal{F}}$ and $t \in \mathbb{T}$,

$$W \le Z \text{ and } \rho_t(W) = \rho_t(Z) \iff W = Z.$$

The strong sensitivity and monotonicity properties are sometimes jointly called the strict (or strong) monotonicity property.

A.2. Aggregated conditional risk mappings

A.2.1. *Finite horizon.* Where it simplifies notation we will write $W_{s:t} = (W_s, \dots, W_t)$ for tuples of length $t - s + 1$, with $W_{s:s} = W_s$, and use the componentwise partial order $W_{s:t} \le W'_{s:t} \iff W_r \le W'_r$, $r = s, \dots, t$. If α and β are real-valued random variables then we write $\alpha W_{s:t} + \beta Z_{s:t} = (\alpha W_s + \beta Z_s, \dots, \alpha W_t + \beta Z_t)$.

Lemma A.1. *The aggregated risk mapping $\{\rho_{s,t}\}$ has the following properties: for all $s, t \in \mathbb{T}$ with $s \le t$,*

Normalisation: $\rho_{s,t}(0, \dots, 0) = 0$.

Conditional translation invariance: for all $\{W_r\}_{r=s}^t \in \otimes^{t-s+1} L^\infty_{\mathcal{F}}$ with $W_s \in \mathcal{G}_s$,

$$\rho_{s,t}(W_s, \dots, W_t) = W_s + \rho_{s,t}(0, W_{s+1}, \dots, W_t).$$

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Monotonicity: for all $\{W_r\}_{r=s}^t, \{Z_r\}_{r=s}^t \in \otimes^{t-s+1} L^\infty_{\mathcal{F}}$,

$$W_{s:t} \le Z_{s:t} \implies \rho_{s,t}(W_{s:t}) \le \rho_{s,t}(Z_{s:t}).$$

Conditional locality: for all $\{W_r\}_{r=s}^t$ and $\{Z_r\}_{r=s}^t$ in $\otimes^{t-s+1} L^\infty_{\mathcal{F}}$,

$$\rho_{s,t}(\mathbf{1}_A W_{s:t} + \mathbf{1}_{A^c} Z_{s:t}) = \mathbf{1}_A \rho_{s,t}(W_{s:t}) + \mathbf{1}_{A^c} \rho_{s,t}(Z_{s:t}), \quad \forall A \in \mathcal{G}_s.$$

Recursivity: for all $s, r, t \in \mathbb{T}$ with $0 \le s < r \le t$,

$$\rho_{s,t}(W_{s:t}) = \rho_{s,r}(W_{s:r-1}, \rho_{r,t}(W_{r:t})).$$

Proof. The proof follows from expanding the recursive definition of $\rho_{s,t}$ and using the properties of its generator. □

A.2.2. *Infinite horizon.*

Lemma A.2. *Recalling Definition 3.1, for all $W \in H_{\mathcal{F}}$ we have*

$$\varrho_s\big(W_s, W_{s+1}, \dots\big) = \lim_{t\to\infty} \rho_{s,t}\big(W_s, \dots, W_t\big) \quad \forall s \in \mathbb{T}.$$

Proof. Let $W \in H_{\mathcal{F}}$ and $\{k_t\}_{t\in\mathbb{T}}$ be as in the definition of $H_{\mathcal{F}}$. Set $K_t := \sum_{n\ge 0} k_{t+n}$. Note that $\{K_t\}_{t\in\mathbb{T}}$ is a non-negative, non-increasing deterministic sequence such that $\lim_{t\to\infty} K_t = 0$. For every $0 \le s \le t$ and $n \ge 1$,

$$\begin{aligned}
\rho_{s,t+n}\big(W_s, \dots, W_{t+n}\big) &= \rho_{s,t+1}\big(W_s, \dots, W_t, \rho_{t+1,t+n}\big(W_{t+1}, \dots, W_{t+n}\big)\big) \\
&\le \rho_{s,t+1}\left(W_s, \dots, W_t, \sum_{m=1}^n k_{t+m}\right) \\
&= \rho_{s,t}(W_s, \dots, W_t) + \sum_{m=1}^n k_{t+m}.
\end{aligned}$$

Similarly we have

$$\rho_{s,t+n}(W_s, \dots, W_{t+n}) \ge \rho_{s,t}(W_s, \dots, W_t) - \sum_{m=1}^n k_{t+m},$$

and we conclude that $\mathbb{P}$-almost surely, the sequence $\{\rho_{s,t}(W_s, \dots, W_t)\}_{t\in\mathbb{T}}$ is Cauchy. □

Lemma A.3. *For all $W \in H_{\mathcal{F}}$ we have*

$$\varrho_s(W_s, W_{s+1}, \dots) = \rho_{s,s+1}(W_s, \varrho_{s+1}(W_{s+1}, W_{s+2}, \dots)).$$

Proof. Arguing as in the proof of Lemma A.2, there is a deterministic positive sequence $\{K_t\}_{t\in\mathbb{T}}$, with $\lim_{t\to\infty} K_t = 0$, such that for every $0 \le s \le t$ we have

$$|\varrho_{s+1}(W_{s+1}, W_{s+2}, \dots) - \rho_{s+1,t}(W_{s+1}, \dots, W_t)| \le K_t \quad \text{almost surely.}$$

The monotonicity and conditional translation invariance of $\rho_{s+1,t}$ imply that

$$\begin{aligned}
\rho_{s,s+1}(W_s, \varrho_{s+1}(W_{s+1}, W_{s+2}, \dots)) &\le \rho_{s,s+1}(W_s, \rho_{s+1,t}(W_{s+1}, \dots, W_t) + K_t) \\
&= \rho_{s,t}(W_s, W_{s+1}, \dots, W_t) + K_t.
\end{aligned}$$

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Taking the limit as $t \to \infty$ we find that

$$\rho_{s,s+1}(W_s, \varrho_{s+1}(W_{s+1}, W_{s+2}, \ldots)) \le \varrho_s(W_s, W_{s+1}, \ldots).$$

A similar argument can be applied to find the reverse inequality. □

All of the properties in Lemma A.1 for finite sequences extend to infinite sequences in $H_{\mathcal{F}}$ with ϱ_s playing the role of $\rho_{s,\infty}$.

A.3. Martingales for aggregated conditional risk mappings

We close by presenting elementary martingale theory for aggregated conditional risk mappings (see also [8, 11]).

Let $f = \{f_t\}_{t\in\mathbb{T}}$ be a sequence in $L^\infty_{\mathcal{F}}$. We say that $W \in \mathcal{L}^\infty_{\mathbb{G}}$ is an f-extended $\{\rho_{s,t}\}$-submartingale (-supermartingale) if

$$W_s \le (\ge)\, \rho_{s,t}\big(f_s, \ldots, f_{t-1}, W_t\big), \quad 0 \le s \le t,$$

and an f-extended $\{\rho_{s,t}\}$ martingale if it has both these properties. Note that we use the convention

$$\rho_{s,t}\big(f_s, \ldots, f_{t-1}, W_t\big) = \rho_{t,t}(W_t) \text{ if } s = t.$$

If $f \equiv 0$ then the qualifier 'f-extended' is omitted.

Lemma A.4. *The definitive property for an f-extended $\{\rho_{s,t}\}$-submartingale (-supermartingale) W is equivalent to the one-step property,*

$$W_t \le (\ge)\, \rho_{t,t+1}(f_t, W_{t+1}), \quad t \in \mathbb{T}.$$

Proof. If $\{W_t\}_{t\in\mathbb{T}}$ is a one-step f-extended $\{\rho_{s,t}\}$-submartingale, then for all $s, t \in \mathbb{T}$ such that $s < t$ we have

$$\begin{aligned}\rho_{s,t}\big(f_s, \ldots, f_{t-1}, W_t\big) &= \rho_{s,t-1}\big(f_s, \ldots, f_{t-2}, \rho_{t-1,t}(f_{t-1}, W_t)\big)\\ &\ge \rho_{s,t-1}\big(f_s, \ldots, f_{t-2}, W_{t-1}\big) \ldots \ge W_s.\end{aligned}$$

The case $s = t$ and the converse implication that an f-extended $\{\rho_{s,t}\}$-submartingale satisfies the one-step property are both trivial and thus omitted. □

Lemma A.5. (Doob decomposition) *Let $W \in \mathcal{L}^\infty_{\mathbb{G}}$. There exists an almost surely unique $\{\rho_{s,t}\}$-martingale M and $\mathbb{G}$-predictable process A such that $M_0 = A_0$ and*

$$W_t = W_0 + M_t + A_t. \tag{A.1}$$

The processes A and M are defined recursively as follows:

$$\begin{cases} A_0 = 0, \\ A_{t+1} = A_t + \big(\rho_t\big(W_{t+1}\big) - W_t\big), \quad t \in \mathbb{T}, \end{cases}$$

$$\begin{cases} M_0 = 0, \\ M_{t+1} = M_t + \big(W_{t+1} - \rho_t\big(W_{t+1}\big)\big), \quad t \in \mathbb{T}. \end{cases}$$

If W is a $\{\rho_{s,t}\}$-submartingale (-supermartingale) then A is increasing (decreasing).

Proof. This is proved in the same way as Lemma 5.1 of [11]. □

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

A.3.1. *Optional stopping properties.* First let $\tau \in \mathscr{T}$ be a stopping time. For sequences $\{f_t\}_{t\in\mathbb{T}}$ and $\{W_t\}_{t\in\mathbb{T}}$ in $H_{\mathcal{F}}$, define the aggregated cost $\rho_{t,\tau}(f_t, \ldots, f_{\tau-1}, W_\tau)$ as

$$\rho_{t,\tau}(f_t, \ldots, f_{\tau-1}, W_\tau) = \begin{cases} 0, & \text{on } \{\tau < t\}, \\ \rho_t(W_t), & \text{on } \{\tau = t\}, \\ \rho_t\big(f_t + \rho_{t+1,\tau}(f_{t+1}), \ldots, f_{\tau-1}, W_\tau)\big), & \text{on } \{\tau > t\}. \end{cases} \tag{A.2}$$

Given another stopping time $\varsigma \in \mathscr{T}$, define the aggregated cost $\rho_{\varsigma,\tau}(f_\varsigma, \ldots, f_{\tau-1}, W_\tau)$ as

$$\begin{aligned} \rho_{\varsigma,\tau}(f_\varsigma, \ldots, f_{\tau-1}, W_\tau) &= \sum_{t\in\mathbb{T}} \mathbf{1}_{\{\varsigma=t\}} \rho_{t,\tau}(f_t, \ldots, f_{\tau-1}, W_\tau) \\ &= \begin{cases} 0, & \text{on } \{\tau < \varsigma\}, \\ \rho_\varsigma(W_\varsigma), & \text{on } \{\tau = \varsigma\}, \\ \rho_\varsigma\big(f_\varsigma + \rho_{\varsigma+1,\tau}(f_{\varsigma+1}, \ldots, f_{\tau-1}, W_\tau)\big), & \text{on } \{\tau > \varsigma\}. \end{cases} \end{aligned} \tag{A.3}$$

Without loss of generality we can assume $\tau \geq t$ and $\tau \geq \varsigma$ in (A.2) and (A.3) respectively. The following lemma shows that the recursive property of aggregated conditional risk mappings extends to stopping times.

Lemma A.6. *If ς, $\tilde{\varsigma}$, and τ are bounded stopping times in $\mathscr{T}$ such that $\varsigma \leq \tilde{\varsigma} \leq \tau$, then for all sequences $\{f_t\}_{t\in\mathbb{T}}$ and $\{W_t\}_{t\in\mathbb{T}}$ in $\mathcal{L}^\infty_{\mathcal{F}}$ we have*

$$\rho_{\varsigma,\tau}\big(f_\varsigma, \ldots, f_{\tau-1}, W_\tau\big) = \rho_{\varsigma,\tilde{\varsigma}}\big(f_\varsigma, \ldots, f_{\tilde{\varsigma}-1}, \rho_{\tilde{\varsigma},\tau}\big(f_{\tilde{\varsigma}}, \ldots, f_{\tau-1}, W_\tau\big)\big).$$

Proof. Since τ is bounded it follows that $\tau \in \mathscr{T}_{[0,T]}$ for some integer $0 < T < \infty$. Furthermore, by (A.3) it suffices to prove for all $0 \leq t \leq T$ that

$$\mathbf{1}_{\{\tilde{\varsigma}\geq t\}} \rho_{t,\tilde{\varsigma}}\big(f_t, \ldots, f_{\tilde{\varsigma}-1}, \rho_{\tilde{\varsigma},\tau}\big(f_{\tilde{\varsigma}}, \ldots, f_{\tau-1}, W_\tau\big)\big) = \mathbf{1}_{\{\tilde{\varsigma}\geq t\}} \rho_{t,\tau}\big(f_t, \ldots, f_{\tau-1}, W_\tau\big). \tag{A.4}$$

By decomposing $\{\tilde{\varsigma} \geq t\}$ into the disjoint events $\{\tilde{\varsigma} = t\}$ and $\{\tilde{\varsigma} \geq t+1\}$ we have

$$\begin{aligned} &\mathbf{1}_{\{\tilde{\varsigma}\geq t\}} \rho_{t,\tilde{\varsigma}}\big(f_t, \ldots, f_{\tilde{\varsigma}-1}, \rho_{\tilde{\varsigma},\tau}\big(f_{\tilde{\varsigma}}, \ldots, f_{\tau-1}, W_\tau\big)\big) = \mathbf{1}_{\{\tilde{\varsigma}=t\}} \rho_{t,\tau}\big(f_t, \ldots, f_{\tau-1}, W_\tau\big) \\ &+ \mathbf{1}_{\{\tilde{\varsigma}\geq t+1\}} \rho_{t,t+1}\big(f_t, \rho_{t+1,\tilde{\varsigma}}\big(f_{t+1}, \ldots, f_{\tilde{\varsigma}-1}, \rho_{\tilde{\varsigma},\tau}\big(f_{\tilde{\varsigma}}, \ldots, f_{\tau-1}, W_\tau\big)\big)\big). \end{aligned}$$

If $t < T$ and if (A.4) holds with $t+1$ in place of t, then using conditional translation invariance we get

$$\begin{aligned} &\mathbf{1}_{\{\tilde{\varsigma}\geq t\}} \rho_{t,\tilde{\varsigma}}\big(f_t, \ldots, f_{\tilde{\varsigma}-1}, \rho_{\tilde{\varsigma},\tau}\big(f_{\tilde{\varsigma}}, \ldots, f_{\tau-1}, W_\tau\big)\big) \\ &\quad = \mathbf{1}_{\{\tilde{\varsigma}=t\}} \rho_{t,\tau}\big(f_t, \ldots, f_{\tau-1}, W_\tau\big) \\ &\qquad + \mathbf{1}_{\{\tilde{\varsigma}\geq t+1\}} \rho_{t,t+1}\big(f_t, \rho_{t+1,\tilde{\varsigma}}\big(f_{t+1}, \ldots, f_{\tilde{\varsigma}-1}, \rho_{\tilde{\varsigma},\tau}\big(f_{\tilde{\varsigma}}, \ldots, f_{\tau-1}, W_\tau\big)\big)\big) \\ &\quad = \mathbf{1}_{\{\tilde{\varsigma}=t\}} \rho_{t,\tau}\big(f_t, \ldots, f_{\tau-1}, W_\tau\big) \\ &\qquad + \mathbf{1}_{\{\tilde{\varsigma}\geq t+1\}} \rho_{t,t+1}\big(f_t, \rho_{t+1,\tau}\big(f_{t+1}, \ldots, f_{\tau-1}, W_\tau\big)\big) \\ &\quad = \mathbf{1}_{\{\tilde{\varsigma}\geq t\}} \rho_{t,\tau}\big(f_t, \ldots, f_{\tau-1}, W_\tau\big), \end{aligned}$$

and we conclude using backward induction. □

https://doi.org/10.1017/apr.2021.44 Published online by Cambridge University Press

Acknowledgements

The authors would like to thank the editor and two anonymous referees for their suggestions, which helped improve the presentation of the paper.

Funding information

This work was partially supported by the EPSRC through grants no. EP/N013492/1 and EP/P002625/1, by the Lloyd's Register Foundation–Alan Turing Institute programme on data-centric engineering under the LRF grant G0095, and by the Swedish Energy Agency through grants no. 42982-1 and 48405-1.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process for this article.

Data

The code used in Section 5 can be found at https://github.com/moriartyjm/optimalswitching/tree/main/hydro.

References

[1] AN, L., COHEN, S. N. AND JI, S. (2013). Reflected backward stochastic difference equations and optimal stopping problems under g-expectation. Preprint. Available at https://arxiv.org/abs/1305.0887.

[2] BÄUERLE, N. AND JAŚKIEWICZ, A. (2018). Stochastic optimal growth model with risk sensitive preferences. *J. Econom. Theory* **173**, 181–200.

[3] CARMONA, R. AND LUDKOVSKI, M. (2010). Valuation of energy storage: an optimal switching approach. *Quant. Finance* **10**, 359–374.

[4] CHERIDITO, P., DELBAEN, F. AND KUPPER, M. (2006). Dynamic monetary risk measures for bounded discrete-time processes. *Electron. J. Prob.* **11**, 57–106.

[5] CHERIDITO, P. AND KUPPER, M. (2011). Composition of time-consistent dynamic monetary risk measures in discrete time. *Internat. J. Theoret. Appl. Finance* **14**, 137–162.

[6] COHEN, S. N. AND ELLIOTT, R. J. (2011). Backward stochastic difference equations and nearly time-consistent nonlinear expectations. *SIAM J. Control Optimization* **49**, 125–139.

[7] DETLEFSEN, K. AND SCANDOLO, G. (2005). Conditional and dynamic convex risk measures. *Finance Stoch.* **9**, 539–561.

[8] FOLLMER, H. AND SCHIED, A. (2016). *Stochastic Finance: an Introduction in Discrete Time*, 4th edn. De Gruyter, Berlin.

[9] FRITTELLI, M. AND ROSAZZA GIANIN, E. (2002). Putting order in risk measures. *J. Banking Finance* **26**, 1473–1486.

[10] KOSE, U. AND RUSZCZYNSKI, A. (2020). Risk-averse learning by temporal difference methods. Preprint. Available at https://arxiv.org/abs/2003.00780v1.

[11] KRÄTSCHMER, V. AND SCHOENMAKERS, J. (2010). Representations for optimal stopping under dynamic monetary utility functionals. *SIAM J. Financial Math.* **1**, 811–832.

[12] LUNDSTRÖM, N. L. P., OLOFSSON, M. AND ÖNSKOG, T. (2020). Management strategies for run-of-river hydropower plants—an optimal switching approach. Preprint. Available at https://arxiv.org/abs/2009.10554.

[13] PFLUG, G. C. AND PICHLER, A. (2014). *Multistage Stochastic Optimization*. Springer, Cham.

[14] PICHLER, A. AND SCHLOTTER, R. (2020). Martingale characterizations of risk-averse stochastic optimization problems. *Math. Program.* **181**, 377–403.

[15] PICHLER, A. AND SHAPIRO, A. (2019). Risk averse stochastic programming: time consistency and optimal stopping. Preprint. Available at https://arxiv.org/abs/1808.10807.

[16] RIEDER, U. (1976). On optimal policies and martingales in dynamic programming. *J. Appl. Prob.* **13**, 507–518.

[17] RUSZCZYŃSKI, A. (2010). Risk-averse dynamic programming for Markov decision processes. *Math. Program.* **125**, 235–261.

[18] RUSZCZYŃSKI, A. AND SHAPIRO, A. (2006). Conditional risk mappings. *Math. Operat. Res.* **31**, 544–561.

[19] SHAPIRO, A., DENTCHEVA, D. AND RUSZCZYNSKI, A. (2014). *Lectures on Stochastic Programming*, 2nd edn. Society for Industrial and Applied Mathematics/Mathematical Optimization Society, Philadelphia.
[20] SHEN, Y., STANNAT, W. AND OBERMAYER, K. (2013). Risk-sensitive Markov control processes. *SIAM J. Control Optimization* **51**, 3652–3672.
[21] UGURLU, K. (2018). Robust optimal control using conditional risk mappings in infinite horizon. *J. Comput. Appl. Math.* **344**, 275–287.