Time Blocks Decomposition of Multistage Stochastic Optimization Problems

Pierre Carpentier, Jean-Philippe Chancelier, Michel de Lara, Thomas Martin, Tristan Rigaut

Time Blocks Decomposition of Multistage Stochastic Optimization Problems

Pierre Carpentier[*], Jean-Philippe Chancelier[†], Michel De Lara[†], Thomas Martin[†], Tristan Rigaut[‡]

December 22, 2022

Abstract

Multistage stochastic optimization problems are, by essence, complex as their solutions are functions of both stages and uncertainties. Their large scale nature makes decomposition methods appealing, like dynamic programming which is a sequential decomposition using a state variable defined at all stages. By contrast, in this paper we introduce the notion of state reduction by time blocks, that is, at stages that are not necessarily all the original stages. Then, we prove a dynamic programming equation with value functions that are functions of a state only at some stages. This equation crosses over time blocks, but involves an optimization inside each block. We illustrate our contribution by showing its potential in three applications in multistage stochastic optimization: mixing dynamic programming and stochastic programming, two-time-scale optimization problems, decision-hazard-decision optimization problems.

Keywords: multistage stochastic optimization, time blocks decomposition, time scales, dynamic programming

1 Introduction

Solutions of multistage stochastic optimization problems are functions of both time and uncertainties. This makes such problems complex. However, their structure makes decomposition methods appealing to solve them [18]. One of the most common approaches are time decomposition (state-based resolution methods), like stochastic dynamic programming, in stochastic optimal control, and scenario decomposition, like progressive hedging, in stochastic programming. On the one hand, stochastic programming deals with an underlying random process taking a finite number of values, called scenarios [19]. Solutions are indexed by a

[*]UMA, ENSTA Paris, Institut Polytechnique de Paris, Palaiseau, France
[†]CERMICS, Ecole des Ponts ParisTech, Marne-la-Vallée, France
[‡]Efficacity, Marne-la-Vallée, France

scenario tree, the size of which increases exponentially with the number of stages (hence generally a few stages in practice). However, to overcome this obstacle, stochastic programming takes advantage of scenario decomposition methods (progressive hedging [17]). On the other hand, stochastic control deals with a state model driven by a white noise, that is, the noise is made of a sequence of independent random variables. Under such assumptions, stochastic dynamic programming is able to handle many stages, as it offers reduction of the search for a solution among state feedbacks (instead of functions of the past noise) [2, 15].

In a word, dynamic programming is good at handling multiple stages — but at the price of assuming that noises are stagewise independent — whereas stochastic programming does not require such assumption, but can only handle a few stages. Could we take advantage of both methods? Is there a way to apply stochastic dynamic programming at a slow time scale — a scale at which noises could be considered statistically independent — crossing over fast time scale optimization problems where independence would not hold? This question is one of the motivations of this paper, and we indeed provide a method to decompose multistage stochastic optimization problems by time blocks. This decomposition method and the main result are, mathematically speaking, quite natural, but the main difficulty is notational. Indeed, the rigorous formulation of multistage stochastic optimization problems on so-called history spaces requires heavy notation.

Although specialists in stochastic optimal control and dynamic programming will find the results as natural and non surprising, or as part of folklore, the fact is that we have not been able to find references that treat the case of *a state defined only at a subset of stages*. This is why we set out to write this paper, without any real theoretical ambition, but with the objective that this result be established and can be used for applications using several forms of decomposition[1]. This is also why we present three (theoretical) applications in multistage stochastic optimization: mixing dynamic programming and stochastic programming, two-time-scale optimization problems, decision-hazard-decision optimization problems.

As there are several ways to tackle the difficulties of dealing with a large number of time steps, we compare our approach with other ones. In this paper, we propose an *exact decomposition* of a multistage stochastic optimization problem *by time blocks* using *a state defined only at a subset of stages*, to be distinguished from either *time aggregation* or *approximate decomposition by timescales*, which both yield *approximate* problems. We discuss both now.

Time aggregation consists in grouping the time steps, that is, in considering a partition of the time steps in time blocks and "aggregating" variables and constraints in each time block. To our knowledge, this approach was initiated in [4] for stochastic linear programs. For such linear programs, it is indeed easily conceived that, by summing ("aggregating") linear constraints, one obtains lower bounds for minimization problems. This approach was generalized in [27] who puts forward a measure-theoretic framework with coarser and finer filtrations, and used linear duality. Then, this was extended in [11] for stochastic convex programs, using filtrations and convex duality. The main idea can be sketched as follows: the

[1]The starting point of our reflections on this subject were conversations that three of us held with Roger Wets in Bogota in 2013. We discussed the interest and the way of mixing the techniques of scenario trees (to be able to take into account correlated noises) with the techniques of dynamic programming (to have a vision of the optimal future costs).

coarser filtration is used to reduce the measurability of the decision variables, whereas the finer filtration is used to enlarge the measurability of the dual variables associated with the constraints, so that the optimal value of the problem obtained by using these two filtrations is an upper bound of the true optimal value; exchanging the role of the filtrations leads to a lower bound. Thus, with time aggregation, one obtains simpler problems that are lower and upper bounds for the original minimization stochastic problems, hence are approximations.

In *approximate decomposition by timescales*, one identifies several timescales in the original multistage stochastic optimization problem and then sets up an optimization problem for each timescale. It is approximate in that the connexion between the problems formulated for each timescale and the whole multistage problem is not explicit.

Approximate decomposition by timescales can be done in the context of dynamic programming, with the value functions obtained for a given timescale entering the final cost of the problem at the finer timescale. This approach gives a cascade of easier to solve optimization problems, and again corresponds to approximate the original problem. An example of this approach can be found in [6] where — for a problem involving both the control of the storage of a battery (5 minutes time steps) and the frequency regulation (2 seconds time steps) — is introduced a first hourly resource model whose resolution by dynamic programming leads to value functions used in a five minute storage model as final costs. The value functions, obtained by solving by dynamic programming this second model, are themselves used in a 2 second frequency model. Another possibility arises when the considered optimization problem displays a *periodical behavior*. In that case, a natural time blocks decomposition is given by the period of the system. In [20], by taking into account such a periodical pattern in the dynamic programming equations, one significantly reduces the computational effort to solve the problem using a fixed point approach. Finally, [14] present a preliminary work on extending the Stochastic Dual Dynamic Programming approach to two-time-scale problems, such as those encountered in energy systems involving both long-term hydro storages and short-term battery storages.

Approximate decomposition by timescales can also be done in the framework of stochastic programming. In [10], the authors introduce a slow scenario tree, that is, a tree involving only the time stages of the slow time scale; but at each node of this slow scenario tree are attached fast time scale scenarios, which do not interfere with the other nodes of the slow scenario tree. The special situation where decisions are taken only at the slow time scale (whereas uncertainties occur at each time stage) is considered in [8]. The authors propose to build a scenario tree branching at the slow time stages, and designed using the theory of bridge processes between two consecutive nodes in order to represent the noise at the fast time scale.

The paper is organized as follows. In Sect. 2, we present stochastic dynamic programming with histories as a way to solve a stochastic optimal control problem formulated in discrete time. In Sect. 3, we revisit the notion of "state" by defining state reduction by time blocks — that is, at stages that are not necessarily all the original stages — and then we prove a reduced dynamic programming equation. This is the central contribution. Then, we illustrate our contribution by showing its potential in three cases. In Sect. 4, we show how to

mix dynamic programming and stochastic programming. In Sect. 5, we detail how to handle problems with two time scales, and illustrate this with the crude oil procurement problem. In Sect. 6, we introduce what we call the *decision-hazard-decision* framework, and we provide a dynamic programming equation. In Appendix A, we survey several frameworks and approaches to solve, by dynamic programming, a stochastic optimal control problem formulated in discrete time. We relegate technical results and proofs in Appendix B, Appendix C and in Appendix D.

2 Stochastic dynamic programming with histories

In this section, we formulate multistage stochastic optimization problems over the so-called history space, with history feedbacks. Then, to prepare the main result in Sect. 3, we establish a dynamic programming equation when the state is the history, that is, the uncertainties and the controls prior to the current stage (see the "canonical construction" in [9, p. 15]). Although quite natural, this equation is generally not written in the literature, as most frameworks in dynamic programming assume the *a priori* existence of a state (see the survey in Appendix A).

We use the notation $[\![r, s]\!] = \{r, r+1, \ldots, s-1, s\}$ for any two natural numbers r, s such that $r \leq s$. We will also use the shorter notation $r{:}s = [\![r, s]\!]$, for example in subscripts as in $h_{r:s}$. From now on, time is discrete and runs among the natural numbers $t \in [\![0, T]\!]$, where $T \in \mathbb{N}^*$ is a positive natural number (and where, for the sake of simplicity, we have taken $t_0 = 0$ regarding the notation in Appendix A). We first define the basic and the composite spaces that we need to formulate multistage stochastic optimization problems. Then, we introduce a class of solutions called history feedbacks.

Histories and history spaces. For each time $t \in [\![0, T-1]\!]$, the control u_t takes its values in a measurable set $\mathbb{U}_t$ equipped with a σ-field $\mathcal{U}_t$. For each time $t \in [\![0, T]\!]$, the uncertainty w_t takes its values in a measurable set $\mathbb{W}_t$ equipped with a σ-field $\mathcal{W}_t$. For $t \in [\![0, T]\!]$, we define the *history space* $\mathbb{H}_t$ equipped with the *history field* $\mathcal{H}_t$

$$\mathbb{H}_t = \mathbb{W}_0 \times \prod_{s=1}^{t} (\mathbb{U}_{s-1} \times \mathbb{W}_s) \ ,$$

$$\mathcal{H}_t = \mathcal{W}_0 \otimes \bigotimes_{s=1}^{t} (\mathcal{U}_{s-1} \otimes \mathcal{W}_s) \ ,$$

with the particular case $\mathbb{H}_0 = \mathbb{W}_0$, $\mathcal{H}_0 = \mathcal{W}_0$. A generic element $h_t = \big(w_0, (u_{s-1}, w_s)_{s=1,\ldots,t}\big) = (w_0,\ u_0,\ w_1,\ u_1,\ w_2,\ \ldots,\ u_{t-2},\ w_{t-1},\ u_{t-1},\ w_t) \in \mathbb{H}_t$ is called a *history* at time t. For $1 \leq r \leq s \leq t$, we introduce the $(r{:}s)$-*history subpart* $h_{r:s} = (u_{r-1}, w_r, \ldots, u_{s-1}, w_s) \in \mathbb{H}_{r:s} = \prod_{\tau=r}^{s} (\mathbb{U}_{\tau-1} \times \mathbb{W}_\tau)$, so that we have $h_t = (h_{r-1}, h_{r:t})$.

History feedbacks. For $0 \leq r \leq t \leq T-1$, we define a $(r{:}t)$-*history feedback* as a sequence $\big\{\gamma_s\big\}_{s=r,\ldots,t}$ of measurable mappings $\gamma_s : (\mathbb{H}_s, \mathcal{H}_s) \to (\mathbb{U}_s, \mathcal{U}_s)$. We call $\Gamma_{r:t}$ the set

of $(r:t)$-history feedbacks. The history feedbacks reflect the following information structure. At the end of the time interval $[t-1, t[$, an uncertainty variable w_t is revealed. Then, at the beginning of the time interval $[t, t+1[$, a decision-maker chooses a control u_t contingent on no more than the past, giving the chronology

$$w_0 \rightsquigarrow u_0 \rightsquigarrow w_1 \rightsquigarrow u_1 \rightsquigarrow \cdots \rightsquigarrow w_t \rightsquigarrow u_t \rightsquigarrow \cdots \rightsquigarrow w_{T-1} \rightsquigarrow u_{T-1} \rightsquigarrow w_T \ . \tag{1}$$

Stochastic kernels. In what follows, given $(\mathbb{Y}, \mathcal{Y})$ a measurable space, $\Delta(\mathbb{Y})$ denotes the space of probability measures over $(\mathbb{Y}, \mathcal{Y})$. Uncertainty is represented by a family $\left\{\rho_{t-1:t}\right\}_{t\in[\![1,T]\!]}$ of stochastic kernels

$$\rho_{t-1:t} : (\mathbb{H}_{t-1}, \mathcal{H}_{t-1}) \to \Delta(\mathbb{W}_t) \ , \ \ \forall t \in [\![1, T]\!] \ , \tag{2}$$

where each $\rho_{t-1:t}(h_{t-1}, \cdot)$ represents the distribution of the next uncertainty w_t parameterized by past history h_{t-1}.

We define, for any feedback $\{\gamma_s\}_{s=t,\dots,T-1} \in \Gamma_{t:T-1}$, a new family of stochastic kernels $\rho^{\gamma}_{t:T} : (\mathbb{H}_t, \mathcal{H}_t) \to \Delta(\mathbb{H}_T)$, that capture the transitions between histories when the dynamics $h_{s+1} = \big(h_s, u_s, w_{s+1}\big)$ is driven by $u_s = \gamma_s(h_s)$ for all s in $[\![t, T-1]\!]$ (see Definition 15 in Appendix B for the detailed construction of $\rho^{\gamma}_{r:t}$). Note that $\rho^{\gamma}_{t:T}$ generates a probability distribution on the space $\mathbb{H}_T$ of histories over the whole timespan $[\![0, T]\!]$.

Cost function. In what follows, we say that a function is *numerical* if it takes its values in $\overline{\mathbb{R}} = [-\infty, +\infty]$ (also called *extended* or *extended real-valued* function). The cost to be minimized is a numerical function

$$j : (\mathbb{H}_T, \mathcal{H}_T) \to [0, +\infty] \ , \tag{3}$$

assumed to be nonnegative[2] and measurable with respect to the history field $\mathcal{H}_T$.

Notice that (3) does not represent a cost at final stage, but a cost function of the whole history $h_T = (w_0, u_0, w_1, u_1, \dots, w_{T-1}, u_{T-1}, w_T) \in \mathbb{H}_T$. As h_T contains *all* past controls and uncertainties, a function $j : \mathbb{H}_T \to [0, +\infty]$ covers the most general case. For instance, the function j can have the special form of a sum of time block costs, like in Equation (20).

Optimization data set and family of optimization problems. We call *optimization data set* the family

$$\mathcal{D} = \Big(\{(\mathbb{U}_t, \mathcal{U}_t)\}_{t\in[\![0,T-1]\!]}, \{(\mathbb{W}_t, \mathcal{W}_t)\}_{t\in[\![0,T]\!]}, \{\rho_{t-1:t}\}_{t\in[\![1,T]\!]}, j\Big) \tag{4}$$

[2]We could also consider a criterion $j : \mathbb{H}_t \to \overline{\mathbb{R}}$, either measurable bounded function, or measurable and uniformly bounded below function. However, for the sake of simplicity, we will deal in the sequel with measurable nonnegative numerical functions. The case $j(h_T) = +\infty$ materializes joint constraints between uncertainties and controls in h_T.

made of the elements presented above. We consider the following family of optimization problems, indexed by t in $[\![0, T-1]\!]$ and parameterized by the history $h_t \in \mathbb{H}_t$: for all t in $[\![0, T-1]\!]$, we define the minimum value

$$V_t(h_t) = \inf_{\gamma_{t:T-1} \in \Gamma_{t:T-1}} \int_{\mathbb{H}_T} j(h'_T) \rho^{\gamma}_{t:T}(\, \mathrm{d}h'_T \,|\, h_t) \;, \quad \forall h_t \in \mathbb{H}_t \;, \tag{5a}$$

and we also define

$$V_T(h_T) = j(h_T) \;, \quad \forall h_T \in \mathbb{H}_T \;. \tag{5b}$$

The numerical function $V_t : \mathbb{H}_t \to [0, +\infty]$ is called the *value function* at time t.

In the next paragraph, we show how the family $\{V_t\}_{t \in [\![0,T]\!]}$ of value functions can be used to solve, via dynamic programming, the optimization problem of interest, that is, the one starting at $t = 0$, whose value is (recall that $h_0 = w_0$)

$$V_0(w_0) = \inf_{\gamma_{0:T-1} \in \Gamma_{0:T-1}} \int_{\mathbb{H}_T} j(h'_T) \rho^{\gamma}_{0:T}(\, \mathrm{d}h'_T \,|\, w_0) \;. \tag{6}$$

Bellman operators and dynamic programming. We show that the value functions in (5) are *Bellman functions*, in that they are solution of a Bellman or dynamic programming equation. However, the proof requires technical assumptions. Indeed, as Bertsekas and Shreve notice at the beginning of [3, §7.6]: "The dynamic programming algorithm is centered around infimization of functions, and this is intimately connected with projections of sets"; "Unfortunately, the projection of a Borel-measurable set need not be Borel-measurable. In Borel spaces, however, the projection of a Borel-measurable set is an analytic set". They devote [3, §7.6] to the definition and study of analytic sets, and in [3, §7.7] define universally measurable functions, as well as lower semianalytic functions.

For any Borel space $\mathbb{X}$, subset $X \subset \mathbb{X}$ and function $\varphi : X \to \overline{\mathbb{R}}$, the function φ is said to be *lower semianalytic* [3, Definition 7.21] if the subset X is analytic and if the subset $\big\{x \in X \,\big|\, \varphi(x) < c\big\}$ is analytic for all $c \in \mathbb{R}$. We denote by $\mathbb{L}^0_+(\mathbb{X})$ the space of lower semianalytic nonnegative numerical functions over $\mathbb{X}$.

Definition 1 *The optimization data set $\mathcal{D}$ in* (4) *is said to be* Borel-analytic *if the spaces $\{(\mathbb{U}_t, \mathcal{U}_t)\}_{t \in [\![0,T-1]\!]}$ and $\{(\mathbb{W}_t, \mathcal{W}_t)\}_{t \in [\![0,T]\!]}$ are Borel spaces, the stochastic kernels $\{\rho_{t-1:t}\}_{t \in [\![1,T]\!]}$ in* (2) *are Borel-measurable, and the criterion j in* (3) *is a (nonnegative) lower semianalytic numerical function.*

The proof of the following theorem is given in Appendix B.

Theorem 2 *Suppose that the optimization data set $\mathcal{D}$ in* (4) *is Borel-analytic. For t in $[\![0, T-1]\!]$, we define the* Bellman operator *$\mathcal{B}_{t+1:t}$ by, for all $\varphi \in \mathbb{L}^0_+(\mathbb{H}_{t+1})$ and for all $h_t \in \mathbb{H}_t$,*

$$\big(\mathcal{B}_{t+1:t}\varphi\big)(h_t) = \inf_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \varphi(h_t, u_t, w_{t+1}) \rho_{t:t+1}(dw_{t+1} \,|\, h_t) \;. \tag{7a}$$

Then, the Bellman operators are such that

$$\mathcal{B}_{t+1:t} : \mathbb{L}^0_+(\mathbb{H}_{t+1}) \to \mathbb{L}^0_+(\mathbb{H}_t) \ , \tag{7b}$$

and the value functions V_t defined in (5) *are universally measurable and satisfy the* Bellman equation*, or* (stochastic) dynamic programming equation*,*

$$V_T = j \ , \ \ V_t = \mathcal{B}_{t+1:t} V_{t+1} \ , \quad \textit{for } t = T-1, \dots, 1, 0 \ . \tag{8}$$

This theorem is mainly inspired by [3, Chap. 8], with the feature that the state x_t is, in our case, the canonical history h_t, with the canonical dynamics $h_{t+1} = \big(h_t, u_t, w_{t+1}\big)$. This very general dynamic programming result is the basis of all future developments in this paper.

The recalls and statements presented in this Sect. 2 are mostly straightforward consequences of results already established in the literature. However, the developements are indispensable to tackle time blocks decomposition in the forthcoming Sect. 3.

3 State reduction by time blocks and dynamic programming

In standard approaches to solve, by dynamic programming, a stochastic optimal control problem formulated in discrete time, either a state is given for all times or no state is given (see Appendix A). In this paper, our approach is intermediate, in that a state is possibly obtained, but only at certain times. Thus, in this section, we consider the question of reducing the history using a compressed "state" variable. Differing with traditional practice, such a variable may not be available at any time $t \in [\![0, T]\!]$, but at some specified stages $0 = t_0 < \cdots < t_N = T$. We have seen in Sect. 2 that the history h_t is itself a state variable with associated canonical dynamics $h_{t+1} = \big(h_t, u_t, w_{t+1}\big)$. However, the size of this canonical state increases with t, which is an unpleasant feature for dynamic programming quickly leading to the well-known curse of dimensionality, hence the practical need to introduce a (ideally low dimensional) state space, at least at some specified stages, as done in this paper. As already said in the introduction, the main difficulty in achieving this goal is notational.

In §3.1, we start by introducing the notion of state reduction on a single time block. In §3.2, we move to state reduction on multiple consecutive time blocks and we give the corresponding dynamic programming equations across time blocks. In §3.3, we conclude on how we obtain reduced optimal feedbacks.

3.1 State reduction on a single time block

We first present the case where the reduction only occurs at two instants denoted by r and t, and such that $0 \leq r < t \leq T$.

Definition 3 *Let* $(\mathbb{X}_r, \mathcal{X}_r)$ *and* $(\mathbb{X}_t, \mathcal{X}_t)$ *be two measurable* state spaces, θ_r *and* θ_t *be two measurable* reduction mappings

$$\theta_r : \mathbb{H}_r \to \mathbb{X}_r \ , \ \ \theta_t : \mathbb{H}_t \to \mathbb{X}_t \ , \tag{9a}$$

and $f_{r:t}$ *be a measurable* dynamics

$$f_{r:t} : \mathbb{X}_r \times \mathbb{H}_{r+1:t} \to \mathbb{X}_t \ . \tag{9b}$$

The triplet $(\theta_r, \theta_t, f_{r:t})$ *is called a* state reduction across $[\![r, t]\!]$ *if we have*[3]

$$\theta_t\big((h_r, h_{r+1:t})\big) = f_{r:t}\big(\theta_r(h_r), h_{r+1:t}\big) \ , \ \ \forall h_t \in \mathbb{H}_t \ . \tag{9c}$$

The state reduction $(\theta_r, \theta_t, f_{r:t})$ *is said to be* compatible *with the family* $\{\rho_{s-1:s}\}_{r+1\leq s\leq t}$ *of stochastic kernels* (2) *if*

- *there exists a* reduced stochastic kernel $\widetilde{\rho}_{r:r+1} : \mathbb{X}_r \to \Delta(\mathbb{W}_{r+1})$*, such that the stochastic kernel* $\rho_{r:r+1}$ *in* (2) *can be factored, for all* $h_r \in \mathbb{H}_r$*, as* $\rho_{r:r+1}(\,\mathrm{d}w_{r+1} \,|\, h_r) = \widetilde{\rho}_{r:r+1}\big(\,\mathrm{d}w_{r+1} \,\big|\, \theta_r(h_r)\big)$*,*
- *for all* s *in* $[\![r+2, t]\!]$*, there exists a* reduced stochastic kernel $\widetilde{\rho}_{s-1:s} : \mathbb{X}_r \times \mathbb{H}_{r+1:s-1} \to \Delta(\mathbb{W}_s)$*, such that the stochastic kernel* $\rho_{s-1:s}$ *can be factored, for all* $h_{s-1} \in \mathbb{H}_{s-1}$*, as* $\rho_{s-1:s}\big(\,\mathrm{d}w_s \,\big|\, (h_r, h_{r+1:s-1})\big) = \widetilde{\rho}_{s-1:s}\big(\,\mathrm{d}w_s \,\big|\, \big(\theta_r(h_r), h_{r+1:s-1}\big)\big)$*.*

We call single time block compatible reduction data set *the family*

$$\mathcal{R}_{r,t} = \Big((\mathbb{X}_r, \mathcal{X}_r), (\mathbb{X}_t, \mathcal{X}_t), \theta_r, \theta_t, f_{r:t}, \{\widetilde{\rho}_{s-1:s}\}_{s\in[\![r+1,t]\!]}\Big) \tag{11}$$

made of the elements presented above. We say that the single time block compatible reduction data set $\mathcal{R}_{r,t}$ *is* Borel *if the spaces* $(\mathbb{X}_r, \mathcal{X}_r)$ *and* $(\mathbb{X}_t, \mathcal{X}_t)$ *are Borel spaces, the reduction mappings* θ_r *and* θ_t *are Borel-measurable, the dynamics* $f_{r:t}$ *is Borel-measurable, and the reduced stochastic kernels* $\widetilde{\rho}_{r:r+1}$ *and* $\widetilde{\rho}_{s-1:s}$*, for* $s \in [\![r+2, t]\!]$*, are Borel-measurable.*

The above definition is similar to the sufficient statistics idea in stochastic control: the state variable, which summarizes the history, is sufficient for the controller to design its control policy ([22, p. 19], [3, Definition 10.6], [21]). However, sufficient statistics in the stochastic control literature are defined at the original time stages. By contrast, Definition 3 — and the coming Definition 5 — consider a notion of sufficient statistics *only for a subset of stages*.

According to Definition 3, the triplet $(\theta_r, \theta_t, f_{r:t})$ is a state reduction across $[\![r, t]\!]$ if and only if the diagram in the left part of Figure 1 is commutative; it is compatible if and only if the diagram in the middle part of Figure 1 is commutative.

The following proposition is the key ingredient to formulate dynamic programming equations with a reduced state.

[3]Notice that, if only the couple $(\theta_r, f_{r:t})$ is given, we can define θ_t by (9c), and thus obtain a triplet $(\theta_r, \theta_t, f_{r:t})$ which is a state reduction across $[\![r, t]\!]$.

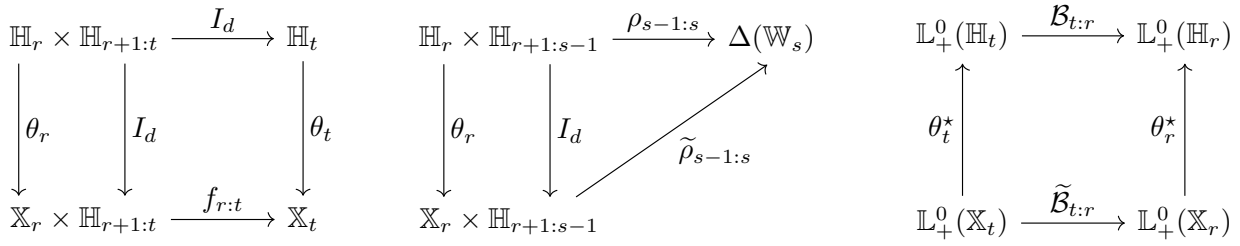

Figure 1: Commutative diagrams in case of state reduction

Proposition 4 *Suppose that the optimization data set $\mathcal{D}$ in* (4) *is Borel-analytic. Then, we can define the* Bellman operator across $[\![t, r]\!]$, $\mathcal{B}_{t:r} : \mathbb{L}^0_+(\mathbb{H}_t) \to \mathbb{L}^0_+(\mathbb{H}_r)$ *by*

$$\mathcal{B}_{t:r} = \mathcal{B}_{r+1:r} \circ \cdots \circ \mathcal{B}_{t:t-1} \ , \tag{12}$$

where the one time step operators $\mathcal{B}_{s:s-1}$, for s in $[\![r+1, t]\!]$ are defined in (7a).

Suppose that there exists a state reduction $(\theta_r, \theta_t, f_{r:t})$ that is compatible with the family $\{\rho_{s-1:s}\}_{r+1\le s\le t}$ of stochastic kernels (2)*, and that the single time block compatible reduction data set $\mathcal{R}_{r,t}$ in* (11) *is Borel.*

Then, there exists a reduced Bellman operator across $[\![t, r]\!]$, $\widetilde{\mathcal{B}}_{t:r} : \mathbb{L}^0_+(\mathbb{X}_t) \to \mathbb{L}^0_+(\mathbb{X}_r)$, *such that*

$$\big(\widetilde{\mathcal{B}}_{t:r}\tilde{\varphi}_t\big) \circ \theta_r = \mathcal{B}_{t:r}(\tilde{\varphi}_t \circ \theta_t) \ , \ \ \forall \tilde{\varphi}_t \in \mathbb{L}^0_+(\mathbb{X}_t) \ . \tag{13}$$

For any $\tilde{\varphi}_t \in \mathbb{L}^0_+(\mathbb{X}_t)$ and for any $x_r \in \mathbb{X}_r$, we have that

$$\begin{aligned}\big(\widetilde{\mathcal{B}}_{t:r}\tilde{\varphi}_t\big) = \inf_{u_r\in\mathbb{U}_r} \int_{\mathbb{W}_{r+1}} \widetilde{\rho}_{r:r+1}(\,\mathrm{d}w_{r+1}\,|\,x_r) \\ \inf_{u_{r+1}\in\mathbb{U}_{r+1}} \int_{\mathbb{W}_{r+2}} \widetilde{\rho}_{r+1:r+2}(\,\mathrm{d}w_{r+2}\,|\,x_r, u_r, w_{r+1}) \quad \cdots \\ \inf_{u_{t-1}\in\mathbb{U}_{t-1}} \int_{\mathbb{W}_t} \widetilde{\rho}_{t-1:t}(\,\mathrm{d}w_t\,|\,x_r, u_r, w_{r+1}, \ldots, u_{t-2}, w_{t-1}) \\ \tilde{\varphi}_t\big(f_{r:t}(x_r, u_r, w_{r+1}, \ldots, u_{t-1}, w_t)\big) \ .\end{aligned} \tag{14}$$

The formula (14) represents a nested sequence of infima of integrals (with respect to different kernels).

The proof of Proposition 4 is given in Appendix B. Proposition 4 can be interpreted as follows. Denoting by $\theta^\star_t : \mathbb{L}^0_+(\mathbb{X}_t) \to \mathbb{L}^0_+(\mathbb{H}_t)$ the operator defined by $\theta^\star_t(\tilde{\varphi}_t) = \tilde{\varphi}_t \circ \theta_t$ for any $\tilde{\varphi}_t \in \mathbb{L}^0_+(\mathbb{X}_t)$, the relation (13) rewrites as $\theta^\star_r \circ \widetilde{\mathcal{B}}_{t:r} = \mathcal{B}_{t:r} \circ \theta^\star_t$, that is, Proposition 4 states that the diagram in the right part of Figure 1 is commutative.

3.2 State reduction on multiple consecutive time blocks and dynamic programming equations

Proposition 4 can easily be extended to the case of multiple consecutive time blocks $[\![t_i, t_{i+1}]\!]$, with $N \in \mathbb{N}^*$, $i \in [\![0, N-1]\!]$ and $0 = t_0 < \cdots < t_N = T$.

Definition 5 *Let $\{(\mathbb{X}_{t_i}, \mathcal{X}_{t_i})\}_{i\in[\![0,N]\!]}$ be a family of measurable state spaces, $\{\theta_{t_i}\}_{i\in[\![0,N]\!]}$ be a family of measurable reduction mappings $\theta_{t_i} : \mathbb{H}_{t_i} \to \mathbb{X}_{t_i}$, and $\{f_{t_i:t_{i+1}}\}_{i\in[\![0,N-1]\!]}$ be a family of measurable dynamics*

$$f_{t_i:t_{i+1}} : \mathbb{X}_{t_i} \times \mathbb{H}_{t_i+1:t_{i+1}} \to \mathbb{X}_{t_{i+1}} \ .$$

The collection $\big(\{(\mathbb{X}_{t_i}, \mathcal{X}_{t_i})\}_{i\in[\![0,N]\!]}, \{\theta_{t_i}\}_{i\in[\![0,N]\!]}, \{f_{t_i:t_{i+1}}\}_{i\in[\![0,N-1]\!]}\big)$ is called a state reduction across the consecutive time blocks $[\![t_i, t_{i+1}]\!]$, $i \in [\![0, N-1]\!]$ *if every triplet $(\theta_{t_i}, \theta_{t_{i+1}}, f_{t_i:t_{i+1}})$ is a state reduction across $[\![t_i, t_{i+1}]\!]$, for i in $[\![0, N-1]\!]$. The state reduction is said to be* Borel *if the state spaces $\{(\mathbb{X}_{t_i}, \mathcal{X}_{t_i})\}_{i\in[\![0,N]\!]}$ are Borel spaces, the reduction mappings $\{\theta_{t_i}\}_{i\in[\![0,N]\!]}$ are Borel-measurable, the dynamics $\{f_{t_i:t_{i+1}}\}_{i\in[\![0,N-1]\!]}$ are Borel-measurable.*

The state reduction across the consecutive time blocks $[\![t_i, t_{i+1}]\!]$ is said to be compatible *with the family $\{\rho_{s-1:s}\}_{s\in[\![1,T]\!]}$ of stochastic kernels given in (2) if every triplet $(\theta_{t_i}, \theta_{t_{i+1}}, f_{t_i:t_{i+1}})$ is compatible with the family $\{\rho_{s-1:s}\}_{s\in[\![t_i+1,t_{i+1}]\!]}$, for i in $[\![0, N-1]\!]$. In this case, we call* multiple time block compatible reduction data set *the family*

$$\mathcal{R}_{\{t_i\}_{i\in[\![0,N]\!]}} = \Big(\{(\mathbb{X}_{t_i}, \mathcal{X}_{t_i})\}_{i\in[\![0,N]\!]}, \{\theta_{t_i}\}_{i\in[\![0,N]\!]}, \{f_{t_i:t_{i+1}}\}_{i\in[\![0,N-1]\!]}, \big\{\{\widetilde{\rho}_{s-1:s}\}_{s\in[\![t_i+1,t_{i+1}]\!]}\big\}_{i\in[\![0,N-1]\!]}\Big) \ , \quad (15)$$

made of the elements presented above. We say that the multiple time block compatible reduction data set $\mathcal{R}_{\{t_i\}_{i\in[\![0,N]\!]}}$ is Borel *if the state reduction is Borel, and the reduced stochastic kernels $\big\{\{\widetilde{\rho}_{s-1:s}\}_{s\in[\![t_i+1,t_{i+1}]\!]}\big\}_{i\in[\![0,N-1]\!]}$ are Borel-measurable.*

Remark 6 (Composed state dynamics as a reduction mapping)

There is a practical case where state reductions can readily be obtained, namely, when the model is given by controlled state dynamics driven by noises. Indeed, in that case,we are given a family of measurable state spaces $\big\{(\mathbb{X}_s, \mathcal{X}_s)\big\}_{s\in[\![0,T]\!]}$ and a family $\{f_{s:s+1}\}_{s\in[\![0,T-1]\!]}$ of measurable dynamics

$$f_{s:s+1} : \mathbb{X}_s \times \mathbb{U}_s \times \mathbb{W}_{s+1} \to \mathbb{X}_{s+1} \ . \quad (16)$$

For any time $s \in [\![0, T-1]\!]$, we define the composition $f_{0:s+1} = f_{s:s+1} \circ f_{s-1:s} \circ \ldots \circ f_{0:1}$ with the abuse of notation that the composition is performed on the state argument. Setting $\mathbb{W}_0 = \mathbb{X}_0$, we obtain that $f_{0:s+1} : \mathbb{H}_{s+1} \to \mathbb{X}_{s+1}$ is a mapping from the history space $\mathbb{H}_{s+1}$ taking values in the state space $\mathbb{X}_{s+1}$.

Now, given a natural number $N > 0$ and an increasing sequence $0 = t_0 < \cdots < t_N = T$ of times, we define the family $\big\{\theta_{t_i}\big\}_{i\in[\![0,N]\!]}$ of measurable reduction mappings by $\theta_{t_i} = f_{0:t_i}$:

$\mathbb{H}_{t_i} \to \mathbb{X}_{t_i}$ *for* $i > 0$, *and by* $\theta_0 = I_d$ *(the identity mapping on* $\mathbb{W}_0$*) for* $i = 0$. *Moreover, given* i *and* $j \in [\![0, N]\!]$, *with* $i < j$ *we obtain, for all* $h_{t_j} \in \mathbb{H}_{t_j}$, *that*

$$\theta_{t_j}(h_{t_j}) = \theta_{t_j}\big((h_{t_i}, h_{t_i+1:t_j})\big) = f_{t_i:t_j}\big(\theta_{t_i}(h_{t_i}), h_{t_i+1:t_j}\big)\ , \tag{17}$$

with $f_{t_i:t_j} = f_{t_j-1:t_j} \circ f_{t_j-2:t_j-1} \circ \ldots \circ f_{t_i:t_i+1}$ *which gives the state reduction Equation* (9c).

Remark 7 (Block independent exogenous noises and stochastic kernels)
There is a practical case where compatible *state reductions can readily be obtained. Assume that the family* $\{\rho_{s-1:s}\}_{s\in[\![1,T]\!]}$ *of stochastic kernels in* (2) *are mappings whose arguments do not include the control part (that is, depend at most on the history uncertainty part (see* (48a)*). If we interpret stochastic kernels as (conditional) distributions of noises (random process), this means that the system dynamics are driven by an exogenous noise process, say* $\{\mathbf{W}_t\}_{t\in[\![1,T]\!]}$.

Assume moreover that the stochastic kernels give rise to noises that are independent block by block, in the sense that the random vectors $\mathbf{W}_0$, $(\mathbf{W}_t)_{t\in[\![1,t_1]\!]}$, $(\mathbf{W}_t)_{t\in[\![t_1+1,t_2]\!]}$, $\ldots$, $(\mathbf{W}_t)_{t\in[\![t_i+1,t_{i+1}]\!]}$, $\ldots$, $(\mathbf{W}_t)_{t\in[\![t_{N-2}+1,t_{N-1}]\!]}$, $(\mathbf{W}_t)_{t\in[\![t_{N-1}+1,t_N]\!]}$ *are stochastically independent. Then, from Definitions 3 and 5, we deduce that any state reduction across the same time blocks is compatible with the stochastic kernels. If the optimization data set* $\mathcal{D}$ *in* (4) *is Borel-analytic and the state reduction is Borel, then the corresponding multiple time block compatible reduction data set is Borel.*

Assuming the existence of a state reduction across the consecutive time blocks $[\![t_i, t_{i+1}]\!]$ compatible with the family of stochastic kernels (2), we obtain the existence of a *family of reduced Bellman operators* across the consecutive $[\![t_i, t_{i+1}]\!]$ as an immediate consequence of multiple applications of Proposition 4, that is, $\widetilde{\mathcal{B}}_{t_{i+1}:t_i} : \mathbb{L}^0_+(\mathbb{X}_{t_{i+1}}) \to \mathbb{L}^0_+(\mathbb{X}_{t_i})$, $i \in [\![0, N-1]\!]$, such that, for any function $\tilde{\varphi}_{t_{i+1}} \in \mathbb{L}^0_+(\mathbb{X}_{t_{i+1}})$, we have that $\big(\widetilde{\mathcal{B}}_{t_{i+1}:t_i}\tilde{\varphi}_{t_{i+1}}\big) \circ \theta_{t_i} = \mathcal{B}_{t_{i+1}:t_i}(\tilde{\varphi}_{t_{i+1}} \circ \theta_{t_{i+1}})$. We now consider the family of optimization problems defined by the associated value functions (5). Thanks to the state reductions, we can enounce the following two theorems which establish dynamic programming equations *across* consecutive time blocks. The first one, Theorem 8, states a dynamic programming equation for an optimization problem in Mayer form (that is, just involving a final cost). The second one, Theorem 9, is more general as it involves both instantaneous costs and a final cost. As it is well known that the second case can be reduced to a Mayer form through a state augmentation, the proof of Theorem 9 easily follows from the proof of Theorem 8.

Theorem 8 (Time block decomposition for the Mayer form) *Suppose that the optimization data set* $\mathcal{D}$ *in* (4) *is Borel-analytic. Suppose that a state reduction* $\big(\{(\mathbb{X}_{t_i}, \mathfrak{X}_{t_i})\}_{i\in[\![0,N]\!]}$, $\{\theta_{t_i}\}_{i\in[\![0,N]\!]}$, $\{f_{t_i:t_{i+1}}\}_{i\in[\![0,N-1]\!]}\big)$ *exists across the consecutive time blocks* $\{[\![t_i, t_{i+1}]\!]\}_{i\in[\![0,N-1]\!]}$, *satisfying* $0 = t_0 < \cdots < t_N = T$, *which is compatible with the family* $\{\rho_{s-1:s}\}_{s\in[\![1,T]\!]}$ *of stochastic kernels given in* (2), *and that the multiple time block compatible reduction data set* $\mathcal{R}_{\{t_i\}_{i\in[\![0,N]\!]}}$ *in* (15) *is Borel.*

Suppose that there exists a reduced criterion $\tilde{\jmath} : \mathbb{X}_T \to [0, +\infty]$, *which is a nonnegative lower semianalytic function and is such that the cost function* j *in* (3) *can be factored as*

$j = \tilde{j} \circ \theta_T$. *We define the family of* reduced value functions $\{\tilde{V}_{t_i}\}_{i\in[\![0,N]\!]}$, *where* $\tilde{V}_{t_i} : \mathbb{X}_{t_i} \to [0,+\infty]$ *for* $i \in [\![0,N]\!]$, *by*

$$\tilde{V}_{t_N} = \tilde{j} \;\; \textit{and} \;\; \tilde{V}_{t_i} = \widetilde{\mathcal{B}}_{t_{i+1}:t_i} \tilde{V}_{t_{i+1}} \;\; , \;\; \forall i \in [\![0, N-1]\!] \; , \tag{19}$$

where the reduced Bellman operators $\{\widetilde{\mathcal{B}}_{t_{i+1}:t_i}\}_{i\in[\![0,N-1]\!]}$ *across the intervals* $\{[\![t_i, t_{i+1}]\!]\}_{i\in[\![0,N-1]\!]}$ *are given in* (14). *Then, the family* $\{V_{t_i}\}_{i\in[\![0,N]\!]}$ *in* (5) *satisfies* $V_{t_i} = \tilde{V}_{t_i} \circ \theta_{t_i}$, *for all* $i \in [\![0,N]\!]$.

The proof is an immediate consequence of multiple applications of Theorem 2 and Proposition 4. It is left to the reader to prove that the following theorem holds true (the proof uses [3, Lemma 7.30(3,4)] on the stability of lower semianalytic functions under addition and under right composition with a Borel mapping).

Theorem 9 (Taking care of instantaneous costs in addition to final cost)
Suppose that the assumptions of Theorem 8 are satisfied, but for the criterion $j : \mathbb{H}_T \to [0,+\infty]$ *which is supposed to be factored as*

$$j(h_T) = \sum_{i=0}^{N-1} \ell_{t_i}\big(\theta_{t_i}(h_{t_i}), h_{t_i+1:t_{i+1}}\big) + \ell_{t_N}\big(\theta_{t_N}(h_{t_N})\big) \; , \tag{20}$$

where the numerical functions $\{\ell_{t_i}\}_{i\in[\![0,N]\!]}$ *are nonnegative lower semianalytic, with* $\ell_{t_i} : \mathbb{X}_{t_i} \times \mathbb{H}_{t_i+1:t_{i+1}} \to [0,+\infty]$ *for* $i \in [\![0,N]\!]$.

We define the family of reduced value functions $\{\tilde{V}_{t_i}\}_{i\in[\![0,N]\!]}$, *where* $\tilde{V}_{t_i} : \mathbb{X}_{t_i} \to [0,+\infty]$ *for* $i \in [\![0,N]\!]$, *by*

$$\tilde{V}_{t_N} = \ell_{t_N} \;\; \textit{and} \;\; \tilde{V}_{t_i} = \overline{\mathcal{B}}_{t_{i+1}:t_i} \tilde{V}_{t_{i+1}} \;\; , \;\; \forall i \in [\![0, N-1]\!] \; , \tag{21}$$

where the reduced Bellman operator $\overline{\mathcal{B}}_{t_{i+1}:t_i}$ *across* $[\![t_i, t_{i+1}]\!]$ *are given, for any* $i \in [\![0,N-1]\!]$, *for any* $\tilde{\varphi}_{t_{i+1}} \in \mathbb{L}^0_+(\mathbb{X}_{t_{i+1}})$ *and for any* $x_{t_i} \in \mathbb{X}_{t_i}$, *by*

$$\begin{aligned}
\big(\overline{\mathcal{B}}_{t_{i+1}:t_i}\tilde{\varphi}_{t_{i+1}}\big)(x_{t_i}) = \inf_{u_{t_i}\in\mathbb{U}_{t_i}} \int_{\mathbb{W}_{t_i+1}} & \widetilde{\rho}_{t_i:t_i+1}(\,\mathrm{d}w_{t_i+1} \,|\, x_{t_i}) \\
& \inf_{u_{t_i+1}\in\mathbb{U}_{t_i+1}} \int_{\mathbb{W}_{t_i+2}} \widetilde{\rho}_{t_i+1:t_i+2}(\,\mathrm{d}w_{t_i+2} \,|\, x_{t_i}, u_{t_i}, w_{t_i+1}) \quad \cdots \\
& \inf_{u_{t_{i+1}-1}\in\mathbb{U}_{t_{i+1}-1}} \int_{\mathbb{W}_{t_{i+1}}} \widetilde{\rho}_{t_{i+1}-1:t_{i+1}} \\
& \qquad (\,\mathrm{d}w_{t_{i+1}} \,|\, x_{t_i}, u_{t_i}, w_{t_i+1}, \ldots, u_{t_{i+1}-2}, w_{t_{i+1}-1}) \\
& \Big(\ell_{t_i}(x_{t_i}, u_{t_i}, w_{t_i+1}, \ldots, u_{t_{i+1}-1}, w_{t_{i+1}}) \\
& \qquad + \tilde{\varphi}_{t_{i+1}}\big(f_{t_i:t_{i+1}}(x_{t_i}, u_{t_i}, w_{t_i+1}, \ldots, u_{t_{i+1}-1}, w_{t_{i+1}})\big)\Big) \; .
\end{aligned} \tag{22}$$

Then, the family $\{V_{t_i}\}_{i\in[\![0,N]\!]}$ *in* (21) *satisfies* $V_{t_i} = \tilde{V}_{t_i} \circ \theta_{t_i}$, *for all* $i \in [\![0,N]\!]$.

Here again, Formula (22) represents a nested sequence of infima of integrals (with respect to different kernels).

Of course, solving Equation (19) or Equation (22) can be as difficult as solving the original Bellman equation. However, the interest of such time block decomposition will be illustrated on different applications in Sect. 4, Sect. 5 and Sect. 6.

3.3 State reduction on multiple consecutive time blocks and reduced optimal feedbacks

As in the classical dynamic programming framework [3, p. 190], we recover the property that the search of an optimal policy among all policies (history feedbacks) can be limited to the search of an optimal state feedback. This is the most important result in practice.

Proposition 10 *Under the assumptions of Theorem 8, the reduced value functions* $\{\tilde{V}_{t_i}\}_{i\in[\![0,N]\!]}$ *defined in* (19) *give the minimum value of the following optimization problems, parameterized by the reduced history (state)* $x_{t_i} \in \mathbb{X}_{t_i}$*:*

$$\tilde{V}_{t_i}(x_{t_i}) = \inf_{\gamma_{t_i:T-1}\in\Gamma_{t_i:T-1}^{x_{t_i}}} \int_{\mathbb{H}_{t_{i+1}:T}} \bar{\jmath}(x_{t_i}, h'_{t_i+1:T})\tilde{\varrho}^{\gamma}_{t_i:T}(\,\mathrm{d}h'_{t_i+1:T}\,|\,x_{t_i}, h_{t_{i+1}:t})\ ,\ \ \forall x_{t_i}\in\mathbb{X}_{t_i}\ , \quad \text{(23a)}$$

and

$$\tilde{V}_T(x_T) = \tilde{\jmath}(x_T)\ ,\ \ \forall x_T \in \mathbb{X}_T\ , \quad \text{(23b)}$$

where the mapping $\bar{\jmath}$ *is given by* $\bar{\jmath} = \tilde{\jmath}\circ f_{t_{N-1}:t_N}\circ f_{t_{N-2}:t_{N-1}}\circ\ldots\circ f_{t_i:t_{i+1}}$ *(with, as already noted, the abuse of notation that the composition is performed on the state argument), where* $\tilde{\varrho}^{\gamma}_{t_i:T}$ *is the reduced stochastic kernel associated with the kernel* $\varrho^{\gamma}_{t_i:T}$,[4] *and where* $\Gamma^{x_{t_i}}_{t_i:T-1}$ *is the set of* $(t_i{:}T{-}1)$*-reduced history feedbacks, that is, the set of sequences* $\{\gamma_s\}_{s=t_i,\ldots,T-1}$ *of measurable mappings* $\gamma_s : (\mathbb{X}_{t_i}\times\mathbb{H}_{t_i+1:s}, \mathcal{X}_{t_i}\otimes\mathcal{H}_{t_i+1:s}) \to (\mathbb{U}_s, \mathcal{U}_s)$*.*

Proof. Using Theorem 8, we have that, for all $i \in [\![0, N]\!]$, $V_{t_i} = \tilde{V}_{t_i}\circ\theta_{t_i}$, with $\tilde{V}_{t_i}$ satisfying the Bellman equation (19). For establishing that $\tilde{V}_{t_i}$ is a value function satisfying Equation (23), we now prove that, in the definition of V_{t_i} in Equation (5), we can replace the set of admissible controls $\Gamma_{t_i:T-1}$ by the set $\Gamma^{x_{t_i}}_{t_i:T-1}$. We proceed as follows. Following [3, Chapter 8] we use the Bellman equation (22) to obtain ϵ-minimizers for each problem (5) (note that ϵ-minimizers can be built with the required analytic regularity which would not be the case for $\epsilon = 0$ minimizers without extra assumptions). As ϵ-minimizers are built by recursively solving the equations (22), they are obtained by solving (up to ϵ) parametric optimization problems. Thus, we easily obtain that the ϵ-minimizer at time $t \in [\![t_i, t_{i+1}]\!]$, can be obtained as a function of $(\theta(h_{t_i}), h_{t_{i+1}:t})$. From this last fact, we obtain that — in the value function definition of V_{t_i} given in Equation (5) — the set $\Gamma_{0:T-1}$ can be replaced by the set of feedbacks of the form $(\theta(h_{t_i}), h_{t_{i+1}:t})$ without changing the value function.

[4] The stochastic kernels $\rho^{\gamma}_{r:t}$, given by (51), are of the form $\rho^{\gamma}_{r:t}(\,\mathrm{d}h'_t\,|\,h_r) = \delta_{h_r}(\,\mathrm{d}h'_r)\otimes\varrho^{\gamma}_{r:t}(\,\mathrm{d}h'_{r+1:t}\,|\,h_r)$.

Finally, when considering Equation (5) — with this restricted family of feedbacks, and considered at time t_i for $i \in [\![0, N]\!]$ — we obtain that the cost to be minimized is now parameterized by $\theta(h_{t_i})$ — as it is the case for the cost to be integrated and also for the stochatic kernels induced by the feedbacks. By setting $x_i = \theta(h_{t_i})$, the obtained optimization problem is the right hand side of (23), that we call $\tilde{\tilde{V}}_{t_i}(x_{t_i})$, and we have that, for all $i \in [\![0, N]\!]$, $V_{t_i} = \tilde{\tilde{V}}_{t_i} \circ \theta_{t_i}$. It remains to prove that $\tilde{V}_{t_i} = \tilde{\tilde{V}}_{t_i}$, for all $i \in [\![0, N]\!]$. By a proof similar to the one of Theorem 8, we show that the family $\{\tilde{\tilde{V}}_{t_i}\}_{i \in [\![0,N]\!]}$ satisfies the Bellman equation (19). By uniqueness, the family $\{\tilde{\tilde{V}}_{t_i}\}_{i \in [\![0,N]\!]}$ coincides with the family $\{\tilde{V}_{t_i}\}_{i \in [\![0,N]\!]}$. □

4 Mixing dynamic programming and stochastic programming

As a first application of the formalism developed in §3.2, we show how dynamic programming and stochastic programming can be mixed (which was the original motivation for the paper, see Footnote 1). The proof of the following proposition is a straightforward application of Theorem 9 combined with Remark 7.

Proposition 11 *Suppose that the optimization data set $\mathcal{D}$ in* (4) *is Borel-analytic. We consider multiple consecutive time blocks $[\![t_i, t_{i+1}]\!]$, with $N \in \mathbb{N}^*$, $i \in [\![0, N-1]\!]$ and $0 = t_0 < \cdots < t_N = T$, and we assume that*

- *a Borel state reduction $\big(\{(\mathbb{X}_{t_i}, \mathcal{X}_{t_i})\}_{i \in [\![0,N]\!]}, \{\theta_{t_i}\}_{i \in [\![0,N]\!]}, \{f_{t_i:t_{i+1}}\}_{i \in [\![0,N-1]\!]}\big)$ exists across the consecutive time blocks $\{[\![t_i, t_{i+1}]\!]\}_{i \in [\![0,N-1]\!]}$,*
- *the noises are exogeneous and* time block independent*, that is, the elements of the family $\{\rho_{s-1:s}\}_{s \in [\![1,T]\!]}$ in* (2) *are of the form*

$$\rho_{r:r+1} : \mathbb{W}_{t_i} \times \cdots \times \mathbb{W}_r \to \Delta(\mathbb{W}_{r+1}) \;, \tag{24}$$

  for all $i \in [\![0, N-1]\!]$ and $r \in [t_i, t_{i+1})$, which means that the distribution of the uncertainty w_{r+1} is only function of the past uncertainties $(w_{t_i}, \ldots, wr)$ within the time block,

- *the criterion $j : \mathbb{H}_T \to [0, +\infty]$ can be factored as*

$$j(h_T) = \sum_{i=0}^{N-1} \ell_{t_i}\big(\theta_{t_i}(h_{t_i}), h_{t_i+1:t_{i+1}}\big) + \ell_{t_N}\big(\theta_{t_N}(h_{t_N})\big) \;, \tag{25}$$

  where the numerical functions $\{\ell_{t_i}\}_{i \in [\![0,N]\!]}$ are nonnegative lower semianalytic, with $\ell_{t_i} : \mathbb{X}_{t_i} \times \mathbb{H}_{t_i+1:t_{i+1}} \to [0, +\infty]$ for $i \in [\![0, N]\!]$.

Then, the multistage stochastic optimization problem (6) *can be solved by the following algorithm.*

Initialization. *Define* $\tilde{V}_{t_N} = \ell_{t_N} : \mathbb{X}_T \to [0, +\infty]$.

Backward recursion. *Suppose that the function* $\tilde{V}_{t_{i+1}} : \mathbb{X}_{t_{i+1}} \to [0, +\infty]$ *is known at index* $i+1 \in [\![1, N]\!]$. *Then, for each state* $x_{t_i} \in \mathbb{X}_{t_i}$ *(for instance on a grid approximating the set* $\mathbb{X}_{t_i}$*, or on* $\mathbb{X}_{t_i}$ *itself when finite and small enough), compute the previous Bellman value function* $\tilde{V}_{t_i}$ *at index* i *as*

$$\tilde{V}_{t_i}(x_{t_i}) = \inf_{u_{t_i} \in \mathbb{U}_{t_i}} \int_{\mathbb{W}_{t_i+1}} \rho_{t_i:t_i+1}(\,\mathrm{d}w_{t_i+1} \,|\, w_{t_i}) \tag{26}$$

$$\inf_{u_{t_i+1} \in \mathbb{U}_{t_i+1}} \int_{\mathbb{W}_{t_i+2}} \rho_{t_i+1:t_i+2}(\,\mathrm{d}w_{t_i+2} \,|\, w_{t_i}, w_{t_i+1}) \quad \cdots$$

$$\inf_{u_{t_{i+1}-1} \in \mathbb{U}_{t_{i+1}-1}} \int_{\mathbb{W}_{t_{i+1}}} \rho_{t_{i+1}-1:t_{i+1}}(\,\mathrm{d}w_{t_{i+1}} \,|\, w_{t_i}, w_{t_i+1}, \ldots, w_{t_{i+1}-1})$$

$$\Big(\ell_{t_i}(x_{t_i}, u_{t_i}, w_{t_i+1}, \ldots, u_{t_{i+1}-1}, w_{t_{i+1}})$$

$$+ \tilde{V}_{t_{i+1}}\big(f_{t_i:t_{i+1}}(x_{t_i}, u_{t_i}, w_{t_i+1}, \ldots, u_{t_{i+1}-1}, w_{t_{i+1}})\big)\Big)\,. \tag{27}$$

Final step. *Compute* $V_0(w_0) = \tilde{V}_{t_0}\big(\theta_{t_0}(w_{t_0})\big)$.

In many practical situations, all the uncertainty sets $\mathbb{W}_0$, …, $\mathbb{W}_T$ are finite and the computation in (26) is tractable by using *stochastic programming* and *scenario tree* techniques, which do not require stagewise independence of the noises. We are thus able to take advantage of both the dynamic programming world and the stochastic programming world:

- use dynamic programming at a selection of stages (for instance, at stages corresponding to the slow time scale) and across the corresponding time blocks (for instance, across consecutive slow time stages), when noises are stochastically independent block by block; that produces Bellman value functions only for the chosen selection of stages (for instance, at the slow time scale);
- use stochastic programming inside time blocks (for instance, at fast time scale, within two consecutive slow time stages); the fast time scale final cost function of a block is given by the Bellman value function computed at the slow time scale which corresponds to the terminal stage of the block; no stagewise independence assumption is required within time blocks (for instance, for the short time scale noises).

Remark 12 *As a special case, it is straightforward to check that the triplet* $\big(\{\mathbb{W}_{t_i}\}_{i \in [\![0,N]\!]}, \{\theta_{t_i}\}_{i \in [\![0,N]\!]}, \{f_{t_i:t_{i+1}}\}_{i \in [\![0,N-1]\!]}\big)$*, with*

- *the reduction mapping* θ_{t_i} *given by* $\theta_{t_i}(h_{t_i}) = w_{t_i}$ *for all* $i \in [\![0, N]\!]$,
- *the dynamics* $f_{t_i:t_{i+1}}$ *given by* $f_{t_i:t_{i+1}}(w_{t_i}, h_{t_i+1:t_{i+1}}) = w_{t_{i+1}}$*, for all* $i \in [\![0, N-1]\!]$.

is a state reduction across the consecutive time blocks $[\![t_i, t_{i+1}]\!]$, $i \in [\![0, N-1]\!]$ *which is compatible with the family of stochastic kernels given by Equation* (24)*. Thus, Proposition 11 applies.*

But, in this special case, the optimal controls can be computed in parallel with respect to time blocks, as the term $\tilde{V}_{t_{i+1}}\big(w_{t_{i+1}}\big)$ *is a constant in* (26)*. What is interesting in* (26) *is the added fact that the optimal strategy which was, a priori, searched as feedbacks depending on the whole history is in fact made up of independent strategies, each defined on a single time block and made up of feedbacks depending only on the block history (the history within the block).*

Numerical illustration. To numerically illustrate the mixing between dynamic programming and stochastic programming, we consider a toy optimization problem over a time span $[\![0, T]\!]$, where T is an even natural number (for instance $T = 24$ for an hourly period problem during a day). The problem involves a storage the state x_t of which is driven by a dynamics involving a control variable u_t and a noise variable w_t. We assume that the noises during the first half time span, that is, for $t \in [\![1, T/2]\!]$, are independent of the noises during the second half time span, that is, for $t \in [\![T/2, T]\!]$. We also assume that each noise variable w_t can only take 2 possibles values, so that the whole uncertainty process can be represented by a binary tree.

In this problem, we consider the two consecutive time blocks $[\![0, T/2]\!]$ and $[\![T/2, T]\!]$, and the state reduction is given in a straightforward manner by the variable x_t (as explained in Remark 6). Thus, we are able to compute Bellman values functions by the algorithm given in Proposition 11. We illustrate the algorithm for the horizon $T = 24$.

- The Bellman value function $\tilde{V}_{24}$ is given by the final cost function of the problem.
- The Bellman value function $\tilde{V}_{12}$ is approximated by discretization and it is computed on a grid involving n points $(x^1, \dots, x^n)$. For $i \in [\![1, n]\!]$, each value $\tilde{V}_{12}(x^i)$ is obtained by solving a stochastic programming problem on the time span $[\![12, 24]\!]$, that is, on a tree involving 2^{12} leaves.
- The optimal cost of the optimization problem is $\tilde{V}_0(x_0)$, obtained again by stochastic programming on the time span $[\![0, 12]\!]$, that is, by solving a stochastic optimization problem on a tree involving 2^{12} leaves, the final cost being given by the function $\tilde{V}_{12}$.

Gathering the calculations performed by this algorithm, we obtain that solving the global problem by mixing dynamic programming and stochastic programming is done by solving $(n + 1)$ stochastic optimization programs on scenario trees, each involving 2^{12} leaves. The total number of leaves to explore when solving the problem by this mixing method is

$$(n + 1)2^{12} \approx 4(n + 1)\, 10^3 \ ,$$

which gives an estimation of the algorithm computational effort.

This mixing method is to be compared with a pure scenario tree method, that is, when the problem is solved by a stochastic optimization program on a scenario tree over the whole time horizon on 24 hours, the total number of leaves to explore being

$$2^{24} \approx 1.6\,10^7 \ .$$

Even using a fairly fine state discretization grid, for example a grid containing 100 points, the resolution by mixing dynamic programming and stochastic programming compared to the pure stochastic programming approach leads to a quite significant gain, namely a factor $1.6\ 10^7/4(100+1)\ 10^3 \approx 40$ in our case.

We performed numerical experiments with a single computer equipped with 12 Intel Core i5-10500 CPU and 16 GB of RAM. We used the LP package of the solver Gurobi 9.51. Apart from the solver, all our code has been implemented with the Julia language and the JuMP modeler. As we failed to obtain a solution for the original problem on a tree for the horizon $T = 24$, we performed numerical tests for shorter horizons, hence for smaller numbers of time steps. For every $T \in \{12, 14, 16, 18, 20, 24\}$, we have considered that a state reduction existed at time $T/2$. The results are gathered in Table 1, and show that the computational time[5] needed by the pure scenario tree method is very rapidly increasing with the number of time steps, whereas the computational time needed by the mixing method grows very slowly with the number of time steps, at least for the different horizons under consideration. Finally, note that the mixing method can be easily parallelized since the computation of the n values $\left\{\tilde{V}_{T/2}(x^i)\right\}_{i\in[\![1,n]\!]}$ of the Bellman function $\tilde{V}_{T/2}$ can be performed in parallel, reducing the CPU time by a factor $(n+1)/2$, that is, approximately 50 in our case.

Horizon T	Mixing method	Pure scenario tree method
12	6.5 s	0.5 s
14	6.5 s	5.0 s
16	6.6 s	71.2 s
18	6.8 s	1009.5 s
20	6.8 s	137,296.0 s
24	7.2 s	$\times$

Table 1: Computational time (in seconds) needed for solving the problem by both the mixing method and the pure scenario tree method, for different horizons; the symbol $\times$ means that the optimization failed

5 Two-time-scale optimization problems

As a second application of the formalism developed in §3.2, we show how to tackle two-time-scale optimization problems. Indeed, some decisions problems naturally involve two different

[5]that is, the CPU time needed to create the LP model by JuMP and to solve it by Gurobi

time scales, because of the timing of decisions — as for example long term investment decision and short term monitoring of physical devices.

In §5.1 and §5.2 we detail the structure and we formulate the two-time-scale optimization problems that we consider. In §5.3, we show how to decompose such problems by time blocks. In §5.4, we illustrate the approach on a crude oil procurement problem.

5.1 Structure of a two-time-scale optimization problem

We provide the data for a two-time-scale optimization problem.

Two time scales. We consider a multistage decision problem, with two time scales. The slow time scale is represented by a finite totally ordered set $(\mathsf{S}, \preceq)$ as follows — where s^+ denotes the *successor* of $s \in \mathsf{S}$ and s^- its *predecessor*, and where we use the notation $t \prec t'$ for $t \preceq t'$ and $t \neq t'$ —

$$\min \mathsf{S} = \underline{s} \prec \cdots \prec s^- \prec s \prec s^+ \prec \cdots \prec \bar{s} = \max \mathsf{S} \,, \tag{28a}$$

and the fast time scale by a finite totally ordered set $(\mathsf{F}, \preceq)$:

$$\min \mathsf{F} = \underline{f} \prec \cdots \prec f^- \prec f \prec f^+ \prec \cdots \prec \bar{f} = \max \mathsf{F} \,. \tag{28b}$$

In a sense to be made more rigorous later (once a unified timeline will have been defined), each slow time interval $[s, s^+[$ is made up of $|\mathsf{F}|$ (cardinality of F) fast time steps, hence the denomination "two-time-scale". For instance, $\mathsf{S} = \{Mo, Tu, We, Th, Fr, Sa, Su\}$ may represent days, whereas $\mathsf{F} = [\![1, 24]\!]$ may represent hours within a day. In some problems, we might even take $\mathsf{F} = [\![0, 24]\!]$ to handle the fact that two decisions (one slow and one fast) are taken at midnight, hence an additional fast time step 0.

Unified timeline. We define the unified timeline of the decision problem in two steps. First, we equip the product set $\mathsf{S} \times \mathsf{F}$ with the following lexicographic order:

$$\begin{aligned} &(\underline{s}, \underline{f}) \prec \cdots \prec (s^-, \bar{f}) \prec (s, \underline{f}) \prec (s, \underline{f}^+) \prec \cdots \\ &\qquad \cdots \prec (s, \bar{f}^-) \prec (s, \bar{f}) \prec (s^+, \underline{f}) \prec \cdots \prec (\bar{s}, \bar{f}) \,. \end{aligned} \tag{29}$$

More formally, we denote by $(s, f)^+$ the successor of (s, f) in $\mathsf{S} \times \mathsf{F} \setminus \{(\bar{s}, \bar{f})\}$, with

$$(s, f)^+ = \begin{cases} (s, f^+) & \text{if } f \neq \bar{f} \,, \\ (s^+, \underline{f}) & \text{if } f = \bar{f} \,. \end{cases} \tag{30a}$$

Similarly, we denote by $(s, f)^-$ the predecessor of (s, f) in $\mathsf{S} \times \mathsf{F} \setminus \{(\underline{s}, \underline{f})\}$, with

$$(s, f)^- = \begin{cases} (s, f^-) & \text{if } f \neq \underline{f} \,, \\ (s^-, \bar{f}) & \text{if } f = \underline{f} \,. \end{cases} \tag{30b}$$

We adopt the convention that the slow time $\mathsf{s} \in \mathsf{S}$ is identified with the two scale time $(s, \bar{f})$, as illustrated in Figure 2. For instance Monday is identified with $(Mo, 24)$.

In the product set $\mathsf{S}\times\mathsf{F}$, the first time $(\underline{s}, \underline{f})$ does not coincide with a slow time (the couple $(Mo, 0)$ does not correspond to Monday in our running example). Thus, we add to the product set $\mathsf{S}\times\mathsf{F}$ an extra time denoted by $(\underline{s}^-, \bar{f})$, corresponding to the extra slow time $\underline{s}^-$, which is such that $(\underline{s}, \underline{f})^- = (\underline{s}^-, \bar{f})$. We denote by $\overline{\mathsf{S}}$ the set $\{\underline{s}^-\} \cup \mathsf{S}$ and by $\overline{\mathsf{S}\times\mathsf{F}}$ the set $(\underline{s}^-, \bar{f}) \cup (\mathsf{S}\times\mathsf{F})$, also called the *extended timeline* when equipped with an order $\preceq$ as follows — where we use the notation $(s, f) \prec (s', f')$ for $(s, f) \preceq (s', f')$ and $(s, f) \neq (s', f')$ —

$$\begin{aligned}(\underline{s}^-, \bar{f}) \prec (\underline{s}, \underline{f}) \prec \cdots \prec (s^-, \bar{f}) \prec (s, \underline{f}) \prec (s, \underline{f}^+) \prec \cdots \\ \cdots \prec (s, \bar{f}^-) \prec (s, \bar{f}) \prec (s^+, \underline{f}) \prec \cdots \prec (\bar{s}, \bar{f}) \,. \end{aligned} \tag{31}$$

The two-time-scale optimization problem will be formulated on the extended timeline $\overline{\mathsf{S}\times\mathsf{F}}$, which we trivially identify with the time set $[\![0, T]\!]$, where $T = |\mathsf{S}| \times |\mathsf{F}|$.

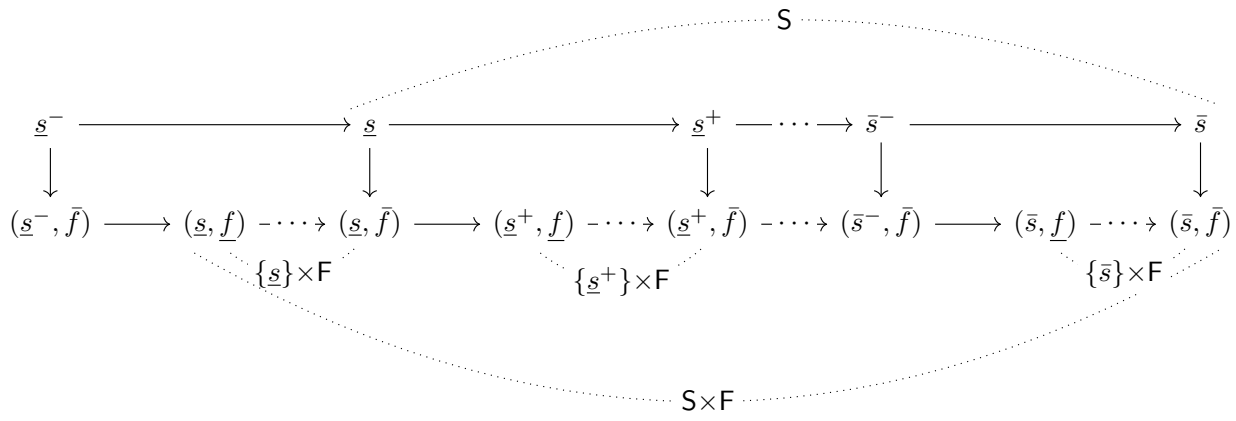

Figure 2: The product timeline with an extra starting point $(\underline{s}^-, \bar{f})$

Decisions. We suppose given

- a family $\{\mathbb{U}^{\mathsf{s}}_s\}_{s\in\overline{\mathsf{S}}\setminus\{\bar{s}\}}$ of *slow time scale decision Borel sets*, and a family $\{\mathbb{W}^{\mathsf{s}}_s\}_{s\in\mathsf{S}}$ of *slow time scale uncertainty Borel sets*,
- a family $\{\mathbb{U}^{\mathsf{sf}}_{(s,f)}\}_{(s,f)\in\mathsf{S}\times(\mathsf{F}\setminus\{\bar{f}\})}$ of *fast time scale decision Borel sets*, and a family $\{\mathbb{W}^{\mathsf{sf}}_{(s,f)}\}_{(s,f)\in\mathsf{S}\times(\mathsf{F}\setminus\{\underline{f}\})}$ of *fast time scale uncertainty Borel sets*.

Dynamics. We suppose given a family $\{\mathbb{X}^{\mathsf{sf}}_{(s,f)}\}_{(s,f)\in\mathsf{S}\times(\mathsf{F}\setminus\{\bar{f}\})}$ of *fast time scale state Borel sets*. For the sake of simplicity, we set $\mathbb{X}^{\mathsf{s}}_s = \mathbb{X}^{\mathsf{sf}}_{(s,\bar{f})}$ for all $s \in \mathsf{S}$ (thus, the slow time $\mathsf{s} \in \mathsf{S}$ is

identified with the two scale time $(s, \bar{f})$). We also suppose given a family $\{\mathcal{F}^{\mathsf{s}}_s\}_{s\in\overline{\mathsf{S}}\setminus\{\bar{\mathsf{s}}\}}$ of *slow time scale dynamics Borel-measurable mappings*, that represent the evolution "driven at the slow time scale" given, for $s \in \overline{\mathsf{S}} \setminus \{\bar{\mathsf{s}}\}$, by[6]

$$\begin{aligned}\mathcal{F}^{\mathsf{s}}_s : \mathbb{X}^{\mathsf{s}}_s \times \mathbb{U}^{\mathsf{s}}_s \times \mathbb{W}^{\mathsf{s}}_{s^+} &\to \mathbb{X}^{\mathsf{sf}}_{(s^+,\underline{f})} \ , \\ \big(x^{\mathsf{s}}_s, u^{\mathsf{s}}_s, w^{\mathsf{s}}_{s^+}\big) &\mapsto x^{\mathsf{sf}}_{(s^+,\underline{f})} = \mathcal{F}^{\mathsf{s}}_s\big(x^{\mathsf{s}}_s, u^{\mathsf{s}}_s, w^{\mathsf{s}}_{s^+}\big) \ . \end{aligned} \tag{32a}$$

We suppose given a family $\{\mathcal{F}^{\mathsf{sf}}_{(s,f)}\}_{(s,f)\in\mathsf{S}\times(\mathsf{F}\setminus\{\bar{f}\})}$ of *fast time scale dynamics Borel-measurable mappings*, that represent the evolution "driven at the fast time scale" given, for all $s \in \mathsf{S}$ and $f \in \mathsf{F} \setminus \{\bar{f}\}$, by

$$\begin{aligned}\mathcal{F}^{\mathsf{sf}}_{(s,f)} : \mathbb{X}^{\mathsf{sf}}_{(s,f)} \times \mathbb{U}^{\mathsf{sf}}_{(s,f)} \times \mathbb{W}^{\mathsf{sf}}_{(s,f)^+} &\to \mathbb{X}^{\mathsf{sf}}_{(s,f)^+} \ , \\ \big(x^{\mathsf{sf}}_{(s,f)}, u^{\mathsf{sf}}_{(s,f)}, w^{\mathsf{sf}}_{(s,f)^+}\big) &\mapsto x^{\mathsf{sf}}_{(s,f)^+} \\ &= \mathcal{F}^{\mathsf{sf}}_{(s,f)}\big(x^{\mathsf{sf}}_{(s,f)}, u^{\mathsf{sf}}_{(s,f)}, w^{\mathsf{sf}}_{(s,f)^+}\big) \ . \end{aligned} \tag{32b}$$

Criterion. We suppose given a family $\{\Lambda_s\}_{s\in\overline{\mathsf{S}}\setminus\{\bar{\mathsf{s}}\}}$ of *slow time scale nonnegative lower semianalytic cost functions*, with

$$\Lambda_{s^-} : \mathbb{X}^{\mathsf{s}}_{s^-} \times \mathbb{U}^{\mathsf{s}}_{s^-} \times \mathbb{W}^{\mathsf{s}}_s \times \underbrace{\prod_{f\in\mathsf{F}\setminus\{\bar{f}\}} \big(\mathbb{X}^{\mathsf{sf}}_{(s,f)} \times \mathbb{U}^{\mathsf{sf}}_{(s,f)} \times \mathbb{W}^{\mathsf{sf}}_{(s,f)^+}\big)}_{\text{interval } [s^-,s[} \to [0,+\infty] \ ,$$

for $s \in \mathsf{S}$, and we suppose given a nonnegative lower semianalytic function $\Lambda_{\bar{s}}$ representing a final cost, with

$$\Lambda_{\bar{s}} : \mathbb{X}^{\mathsf{s}}_{\bar{s}} \to [0,+\infty] \ ,$$

that make up, by summation, an intertemporal criterion

$$\sum_{s\in\mathsf{S}} \Lambda_{s^-}\big(x^{\mathsf{s}}_{s^-}, u^{\mathsf{s}}_{s^-}, w^{\mathsf{s}}_s, \{x^{\mathsf{sf}}_{(s,f)}, u^{\mathsf{sf}}_{(s,f)}, w^{\mathsf{sf}}_{(s,f)^+}\}_{f\in\mathsf{F}\setminus\{\bar{f}\}}\big) + \Lambda_{\bar{s}}\big(x^{\mathsf{s}}_{\bar{s}}\big) \ . \tag{34}$$

Stochastic kernels. Finally, we suppose given a family of *constant slow time scale stochastic kernels* $\{\rho^{\mathsf{s}}_{s:s^+}\}_{s\in\overline{\mathsf{S}}\setminus\{\bar{s}\}}$

$$\rho^{\mathsf{s}}_{s:s^+} \in \Delta(\mathbb{W}^{\mathsf{s}}_{s^+}) \ , \quad \forall s \in \overline{\mathsf{S}} \setminus \{\bar{s}\} \ , \tag{35a}$$

[6]We stress that the slow time scale dynamics (32a) yields as output the first fast state of the slow period (and not the next slow state). Thus, the slow time scale dynamics (32a) is *not* a dynamics from one slow state to the next slow state.

and, for each $s \in \mathsf{S}$, a family $\{\rho^{\mathsf{sf}}_{(s,f):(s,f)^+}\}_{f \in \mathsf{F}\backslash\{\bar{f}\}}$ of *fast time scale stochastic kernels*

$$\rho^{\mathsf{sf}}_{(s,f):(s,f)^+} : \mathbb{W}^{\mathsf{s}}_s \times \underbrace{\prod_{f'=\underline{f}^+}^{f} \mathbb{W}^{\mathsf{sf}}_{(s,f')}}_{\text{interval } [s^-,s[} \longrightarrow \Delta(\mathbb{W}^{\mathsf{sf}}_{(s,f)^+}) \ ,$$
$$\forall s \in \mathsf{S} \ , \ \forall f \in \mathsf{F}\backslash\{\bar{f}\} \ , \tag{35b}$$

with the convention that the Cartesian products of spaces in Equations (35a) and (35b) reduce to the empty set when the upper index of the Cartesian product is strictly lower that the corresponding lower index. Note that, for a given $s \in \mathsf{S}$, each fast time scale stochastic kernel $\rho^{\mathsf{sf}}_{(s,f):(s,f)^+}$, only depends on the noises of the slow time block s.[7]

5.2 Formulation of a two-time-scale optimization problem on the product timeline

To apply Theorem 8, we introduce sets associated with the extended timeline (31) by

$$\mathbb{X}_{(s,f)} = \begin{cases} \mathbb{X}^{\mathsf{s}}_s & \text{if } f = \bar{f} \\ \mathbb{X}^{\mathsf{sf}}_{(s,f)} & \text{if } f \neq \bar{f} \end{cases} \ , \ \forall (s,f) \in \overline{\mathsf{S}\times\mathsf{F}} \ , \tag{36a}$$

$$\mathbb{U}_{(s,f)} = \begin{cases} \mathbb{U}^{\mathsf{s}}_s & \text{if } f = \bar{f} \\ \mathbb{U}^{\mathsf{sf}}_{(s,f)} & \text{if } f \neq \bar{f} \end{cases} \ , \ \forall (s,f) \in \overline{\mathsf{S}\times\mathsf{F}} \setminus \{(\bar{s},\bar{f})\} \ , \tag{36b}$$

$$\mathbb{W}_{(s,f)} = \begin{cases} \mathbb{W}^{\mathsf{s}}_s & \text{if } f = \underline{f} \\ \mathbb{W}^{\mathsf{sf}}_{(s,f)} & \text{if } f \neq \underline{f} \end{cases} \ , \ \forall (s,f) \in \mathsf{S}\times\mathsf{F} \ , \tag{36c}$$

with the particular case of the extra initial slow time

$$\mathbb{W}_{(\underline{s}^-,\bar{f})} = \mathbb{X}^{\mathsf{s}}_{\underline{s}^-} \ , \tag{36d}$$

and a family of state dynamics $\mathcal{F}_{(s,f)} : \mathbb{X}_{(s,f)} \times \mathbb{U}_{(s,f)} \times \mathbb{W}_{(s,f)^+} \to \mathbb{X}_{(s,f)^+}$ defined by

$$\mathcal{F}_{(s,f)} = \begin{cases} \mathcal{F}^{\mathsf{s}}_s & \text{if } f = \bar{f} \\ \mathcal{F}^{\mathsf{sf}}_{(s,f)} & \text{if } f \neq \bar{f} \end{cases} \ , \ \forall (s,f) \in \overline{\mathsf{S}\times\mathsf{F}} \setminus \{(\bar{s},\bar{f})\} \ . \tag{37}$$

From these sets, we deduce the history sets and the histories for all $(s,f) \in \overline{\mathsf{S}\times\mathsf{F}}$

$$\mathbb{H}_{(s,f)} = \mathbb{W}_{(\underline{s}^-,\bar{f})} \times \prod_{(\underline{s},\underline{f}) \preceq (s',f') \preceq (s,f)} \left(\mathbb{U}_{(s',f')^-} \times \mathbb{W}_{(s',f')} \right) , \tag{38a}$$

$$h_{(s,f)} = \left(w_{(\underline{s}^-,\bar{f})}, \left(u_{(s',f')^-}, w_{(s',f')} \right)_{(\underline{s},\underline{f}) \preceq (s',f') \preceq (s,f)} \right) , \tag{38b}$$

[7] The (constant) assumption (35a) and the (single block) assumption (35b) correspond to stochastic independence between time blocks, and will be useful in the proof of Proposition 13.

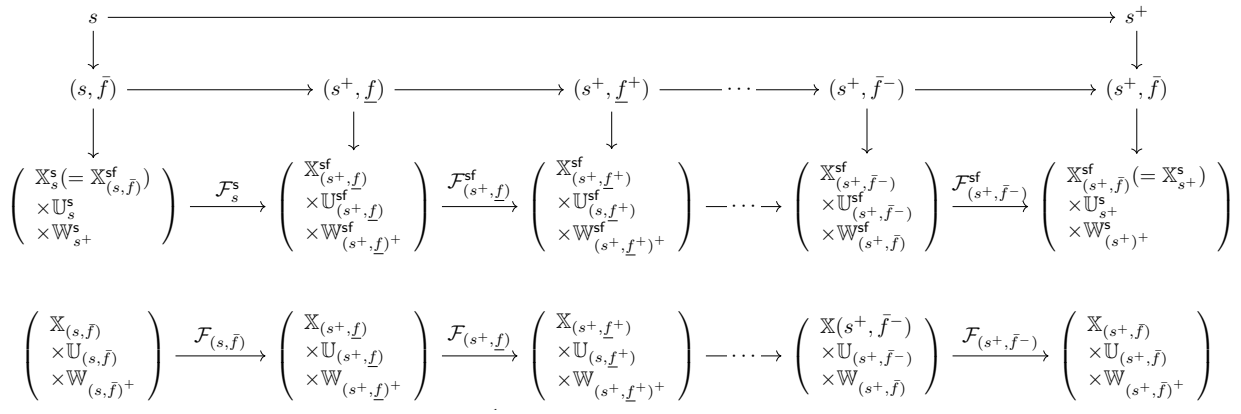

Figure 3: Original dynamics and their reformulation on the product timeline on the slow time interval $[s, s^+[$

and, for suitable indices, the partial history sets and the partial histories

$$\mathbb{H}_{(s,f):(s',f')} = \prod_{(s,f)\preceq(s'',f'')\preceq(s',f')} \left(\mathbb{U}_{(s'',f'')^-} \times \mathbb{W}_{(s'',f'')}\right) , \tag{39a}$$

$$h_{(s,f):(s',f')} = \left((u_{(s'',f'')^-}, w_{(s'',f'')}\right)_{(s,f)\preceq(s'',f'')\preceq(s',f')}\Big) . \tag{39b}$$

The criterion formulated in Equation (34) combined with state dynamics leads to a (lower semianalytic) criterion $j : \mathbb{H}_{(\bar{s},\bar{f})} \to [0, +\infty]$.

Based on the stochastic kernels (35a) and (35b), we introduce stochastic kernels $\rho_{(s,f):(s,f)^+}$ associated with the extended timeline (31), for each $(s, f) \in \overline{\mathsf{S}\times\mathsf{F}} \setminus \{\bar{s}, \bar{f}\}$, by $\rho_{(s,f):(s,f)^+} : \mathbb{H}_{(s,f)} \longrightarrow \Delta(\mathbb{W}_{(s,f)^+})$ with

$$\begin{aligned}
&\rho_{(s,f):(s,f)^+}\big(\,\mathrm{d}w_{(s,f)^+} \,\big|\, h_{(s,f)}\big) \\
&= \begin{cases} \rho^{\mathsf{s}}_{s:s^+}(\,\mathrm{d}w^{\mathsf{s}}_{s^+}) & \text{if } f = \bar{f}\,, \\ \rho^{\mathsf{sf}}_{(s,f):(s,f)^+}\big(\,\mathrm{d}w^{\mathsf{sf}}_{(s,f^+)} \,\big|\, w^{\mathsf{s}}_{s}, w^{\mathsf{sf}}_{(s,\underline{f}^+)}, \cdots, w^{\mathsf{sf}}_{(s,f)}\big) & \text{if } f \neq \bar{f}\,. \end{cases}
\end{aligned}$$

Note that, for $f \neq \bar{f}$, the kernels $\rho_{(s,f):(s,f)^+} : \mathbb{H}_{(s,\underline{f}):(s,f)} \to \Delta(\mathbb{W}_{(s,f)^+})$, only depend on the partial history uncertainty part from $(s, \underline{f})$ to (s, f), and not on the (past) controls.

The components of the problem are now formulated on the extended timeline $\overline{\mathsf{S}\times\mathsf{F}}$, already identified with the time set $[\![0, T]\!]$. Thus, we are in the framework of Sect. 2 and we aim at solving an optimization problem as formulated in Equation (6).

5.3 Two-time-scale decomposition

The existence of Bellman equations for a two-time-scale optimization problem is given by the following proposition.

Proposition 13 *Consider a two-time-scale optimization problem as formulated in §5.1 and §5.2. The optimization problem* (6) *has a solution given by a dynamic programming equation at the slow scale. More precisely, let* $\{V_s\}_{s\in\overline{\mathsf{S}}}$ *be given by* $V_{\bar{s}} = \Lambda_{\bar{s}}$ *and, for* $s \in \overline{\mathsf{S}} \setminus \{\bar{s}\}$*, by the backward induction*[8]

$$
\begin{aligned}
V_s(x_s^{\mathsf{s}}) = \inf_{u_s^{\mathsf{s}}\in\mathbb{U}_s^{\mathsf{s}}} \int_{\mathbb{W}_{s^+}^{\mathsf{s}}} \rho_{s:s^+}^{\mathsf{s}}(\,\mathrm{d}w_{s^+}^{\mathsf{s}}) \\
\inf_{u_{(s^+,\underline{f})}^{\mathsf{sf}}\in\mathbb{U}_{(s^+,\underline{f})}^{\mathsf{sf}}} \int_{\mathbb{W}_{(s^+,\underline{f}^+)}^{\mathsf{sf}}} \rho_{(s^+,\underline{f}):(s^+,\underline{f}^+)}^{\mathsf{sf}}(\,\mathrm{d}w_{(s^+,\underline{f}^+)}^{\mathsf{sf}} \,|\, w_{s^+}^{\mathsf{s}}) \;\cdots \\
\inf_{u_{(s^+,\bar{f}^-)}^{\mathsf{sf}}\in\mathbb{U}_{(s^+,\bar{f}^-)}^{\mathsf{sf}}} \int_{\mathbb{W}_{(s^+,\bar{f})}^{\mathsf{sf}}} \rho_{(s^+,\bar{f}^-):(s^+,\bar{f})}^{\mathsf{sf}}(\,\mathrm{d}w_{(s^+,\bar{f})}^{\mathsf{sf}} \,|\, w_{s^+}^{\mathsf{s}}, w_{(s^+,\underline{f}^+)}^{\mathsf{sf}}, \cdots, w_{(s^+,\bar{f}^-)}^{\mathsf{sf}}) \\
\Big(\Lambda_s(x_s^{\mathsf{s}}, u_s^{\mathsf{s}}, w_{s^+}^{\mathsf{s}}, \ldots, u_{(s^+,\bar{f}^-)}^{\mathsf{sf}}, w_{(s^+,\bar{f})}^{\mathsf{sf}}) \\
+ V_{s^+}\big(\mathcal{F}_{s:s^+}(x_s^{\mathsf{s}}, u_s^{\mathsf{s}}, w_{s^+}^{\mathsf{s}}, \ldots, u_{(s^+,\bar{f}^-)}^{\mathsf{sf}}, w_{(s^+,\bar{f})}^{\mathsf{sf}})\big)\Big)\,,
\end{aligned}
\tag{40}
$$

where $\mathcal{F}_{s:s^+}$ *is the composition* $\mathcal{F}_{s:s^+} = \mathcal{F}_{(s^+,\bar{f}^-)}^{\mathsf{sf}} \circ \cdots \circ \mathcal{F}_{(s^+,\underline{f})}^{\mathsf{sf}} \circ \mathcal{F}_s^{\mathsf{s}}$ *associated with the state dynamics defined in* (32)*. Then, the value of the optimization problem* (6) *is given by* $V_{\underline{s}^-}(x_{\underline{s}^-}^{\mathsf{s}})$*.*

Proof. The proof is an application of Theorem 9 with the help of Remarks 6 and 7. First, we have re-framed in §5.2 the two-time-scale optimization problems described in §5.1 in the formalism of Sect. 2 with the help of the extended timeline (31). Second, as we are given state dynamics (37) on the extended timeline and thanks to Remark 6, we obtain a Borel state reduction at times $\{(s,\bar{f})\}_{s\in\mathsf{S}}$ by composition of the state dynamics. Moreover, as the slow time scale kernels given by Equation (35a) are constant, the state reduction across the slow time scale is compatible with the stochastic kernels (see Remark 7), and is Borel. We are thus able to apply Theorem 9 and obtain the slow time scale Bellman recursion (40) as a special case of Equation (22). □

The slow time scale Bellman equation (40) is as difficult to solve as the Bellman equation on the extended timeline. However, the interest of (40) lies elsewhere. Imagine that one is able to obtain, in a relatively easy way, lower $\underline{V_s}$ and upper $\overline{V_s}$ approximations of V_s in (40) (see [16] in which such approximations are obtained on a battery management problem). Then, in order to obtain optimal controls for the optimization problem (6), one can proceed as follows. By replacing the last term V_{s^+} of (40) by either $\underline{V_{s^+}}$ or $\overline{V_{s^+}}$, one can now solve a (lower or upper) surrogate of Equation (40) and thus obtain the optimal controls on the time block $[\![s, s^+]\!]$. For instance, one could use scenario decomposition methods, like progressive hedging [17], that do not require statistical independence of noises within the slow time interval $[s, s^+[$. Thus, the two-time-scale optimization problem as formulated in §5.1 and §5.2 can be approximatively solved, from below and from above, by a mix of slow time scale dynamic programming and of (for example) progressive hedging (or any other method, including dynamic programming).

[8] Here again, the formula (40) represents a nested sequence of infima of integrals (with respect to different kernels).

5.4 Illustration with the crude oil procurement problem

Crude oil procurement is the part of the oil supply chain that sits between the production of crude oil and its processing in a refinery. The goal of procurement is to purchase crude oil from various suppliers around the world and having it delivered in time to the refinery to be processed. As illustrated in Figure 4, every month (on the bottom line) a refinery receives crudes that have been bought during the 8 previous weeks (on the upper line).

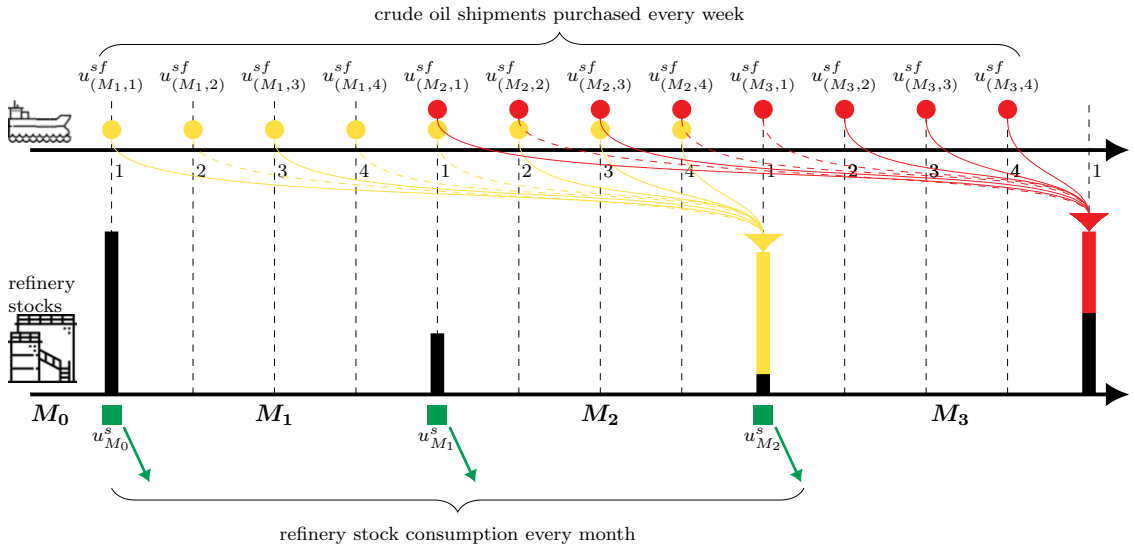

Figure 4: Procurement of crude oil over 3 months M_1, M_2 and M_3, where a circle ○ denotes purchase decisions and a square □ denotes consumption decisions

The problem naturally displays two time scales. On the one hand, deliveries to the refinery are made at the beginning of each month, and crude consumption is set once a month. On the other hand, crude oil shipments can be purchased at the frequency of the week; every week, a selection of shipments is presented to the decision-maker who must decide which shipments to purchase. Following the construction of the extended timeline in (31), we represent by the sequence

$$\begin{aligned}(M_0,5) \prec (M_1,1) \prec (M_1,2) \prec (M_1,3) \prec (M_1,4) \prec (M_1,5)\\ \prec (M_2,1) \prec (M_2,2) \prec (M_2,3) \prec (M_2,4) \prec (M_2,5)\\ \prec (M_3,1) \prec (M_3,2) \prec (M_3,3) \prec (M_3,4) \prec (M_3,5)\end{aligned} \tag{41}$$

the timeline associated with Figure 4 (notice that we consider that a month is made of 4 weeks). The initial stage $(M_0,5)$ corresponds to the additional stage $(\underline{s}^-,\bar{f})$ in (31). The stages $(M_1,5)$ and $(M_2,5)$ both represent the "end of the month" when a consumption decision (slow scale decision u_s^{s} on the bottom line of Figure 4) is taken.

We now illustrate how the crude oil procurement problem can be put in the form of a

Notations from §5.1	Crude oil procurement
S	set of months during which we manage the refinery; in Figure 4, $\mathsf{S} = \{M_1, M_2, M_3\}$
F	set of weeks in each month; in Figure 4, $\mathsf{F} = \{1, 2, 3, 4, 5\}$
$\mathbb{U}^{\mathsf{s}}_s$	set of crude oil consumptions during the month s^+
$\mathbb{W}^{\mathsf{s}}_{s^+}$	set of product prices for the month s^+
$\mathbb{U}^{\mathsf{sf}}_{(s,f)}$	set of crude shipments purchased in week (s, f)
$\mathbb{W}^{\mathsf{sf}}_{(s,f)^+}$	set of crude oil prices in week (s, f)
$\mathcal{F}^{\mathsf{sf}}_{(s,f)}$	accumulation of shipments purchased in (s, f)
$\mathcal{F}^{\mathsf{s}}_s$	delivery of orders and consumption of crude oil for the month s^+
Λ_s	operational costs during the month s (crude oil purchases during s - earnings from production)
Λ_{M_4}	end cost associated with the state $x^s_{M_3} = x^{sf}_{(M_3,5)}$ valuation of the buffers and stocks in the refinery before the beginning of the month M_4

Table 2: Identification of the elements introduced in §5.1 with elements of the crude oil procurement problem

two-time-scale optimization problem such as presented in §5.1. For this purpose, we proceed to the identifications in Table 2.

We call $s-$buffer (resp. s^--buffer), the temporary stock that is created at the beginning of the month s (resp. s^-) and that will be delivered two months after. For instance, in Figure 4, the yellow disks represent the M_1-buffer and the red disks represent the M_2-buffer. We introduce the state variable $x^{\mathsf{sf}}_{(s,f)} = \Big(s^- - \text{buffer}, s - \text{buffer}, \text{refinery stocks}\Big)$, together with the accumulation dynamics $\mathcal{F}^{\mathsf{sf}}_{(s,f)}$ for the buffers, and the accumulation dynamics $\mathcal{F}^{\mathsf{s}}_s$ for the stocks. Regarding the criterion to minimize, it is an intertemporal criterion like in (34) with slow time scale cost functions Λ_{s^-} made of minus the purchases of crude oil plus the selling of finished products inside a week. Supposing that the products prices are independent month by month, we represent this assumption by a family of constant kernels $\{\rho^{\mathsf{s}}_{s:s^+}\}_{s\in\overline{\mathsf{S}}\setminus\{\bar{s}\}}$. By contrast, we do not assume that the crude prices are independent week by week, and the possible dependency is modeled by stochastic kernels $\rho^{\mathsf{sf}}_{(s,f):(s,f)^+}$ for $f \in \mathsf{F}\setminus\{\bar{f}\}$.

Now that all the elements from §5.1 have been identified, Proposition 13 enables us to write a dynamic programming equation such as (40) at the scale of the month, without losing the time-dependency of crude prices inside the month. This illustration stems from a research work done in partnership with TotalEnergies, in the context of a PhD thesis [13].

6 Decision-Hazard-Decision Optimization Problems

In multistage stochastic optimization, the *decision-hazard-decision* framework corresponds to the case where, at the beginning of each time interval, a decision is taken without knowing the uncertainty that will materialize at the end of the time interval, and, at the end of the time interval, a *recourse* decision is possible.

In §6.1, we provide motivation for the decision-hazard-decision framework. In §6.2, we formalize the decision-hazard-decision framework and we provide a dynamic programming equation. Proofs and details are relegated in Appendix D.

6.1 Motivation for the Decision-Hazard-Decision Framework

We illustrate our motivation to develop a formalism for the Decision-Hazard-Decision framework with a single dam management problem. We suppose given a stochastic process $\{\mathbf{A}_t\}_{t\in[\![1,T]\!]}$ on a probability space. We can model the dynamics of the water volume in a dam by

$$\mathbf{S}_{t+1} = \min\left\{S^\sharp, \mathbf{S}_t - \mathbf{Q}_t + \mathbf{A}_{t+1}\right\}\ ,\ \ \forall t \in [\![0, T-1]\!]\ , \tag{42}$$

where $S^\sharp$ is the maximal dam volume, $\mathbf{S}_t$ is the volume (stock) of water at the beginning of period $[t, t+1[$, $\mathbf{A}_{t+1}$ is the inflow water volume (rain, etc.) during $[t, t+1[$, $\mathbf{Q}_t$ is the turbined outflow volume during $[t, t+1[$. The control variable $\mathbf{Q}_t$ is decided at the *beginning* of period $[t, t+1[$, chosen such that $0 \leq \mathbf{Q}_t \leq \mathbf{S}_t$, supposed to depend on the stock $\mathbf{S}_t$ but not on the inflow water $\mathbf{A}_{t+1}$ (as $\mathbf{A}_{t+1}$ takes place during $[t, t+1[$, hence materializes at $t+1$, hence the time index $t+1$). The min operation in Equation (42) ensures that the dam volume always remains below its maximal capacity $S^\sharp$, but induces a nonlinearity in the dynamics. This nonlinear dynamics is an obstacle to apply stochastic dual dynamic programming (SDDP).

Alternatively, we can model the dynamics of the water volume in a dam by

$$\mathbf{S}_{t+1} = \mathbf{S}_t - \mathbf{Q}_t - \mathbf{A}_{t+1} - \mathbf{R}_{t+1}\ ,\ \ \forall t \in [\![0, T-1]\!]\ , \tag{43}$$

where the new control variable $\mathbf{R}_{t+1}$ is the spilled volume, decided at the *end* of period $[t, t+1[$, supposed to depend on the stock $\mathbf{S}_t$ and on the inflow water $\mathbf{A}_{t+1}$, and chosen such that $0 \leq \mathbf{S}_t - \mathbf{Q}_t + \mathbf{A}_{t+1} - \mathbf{R}_{t+1} \leq S^\sharp$. Thus, with the formulation (43), we "pay the price" to add one control $\mathbf{R}_{t+1}$, but we obtain a linear model instead of the nonlinear model (42). This is especially interesting when using the stochastic dual dynamic programming (SDDP) algorithm, for which the linearity of the dynamics is used to obtain the convexity properties required by the algorithm.

6.2 Decision-Hazard-Decision framework and dynamic programming equation

Let $\{\mathbb{U}_t^\flat\}_{t\in[\![0,T-1]\!]}$, $\{\mathbb{U}_t^\sharp\}_{t\in[\![1,T]\!]}$, $\{\mathbb{W}_t\}_{t\in[\![1,T]\!]}$ and $\{\mathbb{X}_t\}_{t\in[\![0,T]\!]}$ be sequences of measurable sets, equipped with corresponding σ-fields $\{\mathcal{U}_t^\sharp\}_{t\in[\![0,T-1]\!]}$, $\{\mathcal{U}_t^\flat\}_{t\in[\![0,T-1]\!]}$, $\{\mathcal{W}_t\}_{t\in[\![1,T]\!]}$ and $\{\mathcal{X}_t\}_{t\in[\![0,T]\!]}$.

Let also be given measurable dynamics mappings

$$f_t : \mathbb{X}_t \times \mathbb{U}_t^\flat \times \mathbb{U}_{t+1}^\sharp \times \mathbb{W}_{t+1} \to \mathbb{X}_{t+1} \ , \ \ \forall t \in [\![0, T-1]\!] \ , \tag{44a}$$

measurable instantaneous cost functions

$$L_t : \mathbb{X}_t \times \mathbb{U}_t^\flat \times \mathbb{U}_{t+1}^\sharp \times \mathbb{W}_{t+1} \to [0, +\infty] \ , \ \ \forall t \in [\![0, T-1]\!] \ , \tag{44b}$$

and measurable final cost function

$$K : \mathbb{X}_T \to [0, +\infty] \ . \tag{44c}$$

Finally, we suppose given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a random variable $\mathbf{X}_0 : \Omega \to \mathbb{X}_0$, and a stochastic process $\{\mathbf{W}_t\}_{t \in [\![1,T]\!]}$ (*noise process*), where $\mathbf{W}_t : \Omega \to \mathbb{W}_t$ for $t \in [\![1, T]\!]$.

Thus equipped, we consider the following multistage stochastic optimization problem

$$\inf_{\{(\mathbf{U}_t^\flat, \mathbf{U}_{t+1}^\sharp)\}_{t \in [\![0,T-1]\!]}} \mathbb{E}\big[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t^\flat, \mathbf{U}_{t+1}^\sharp, \mathbf{W}_{t+1}) + K(\mathbf{X}_T)\big] \ , \tag{45a}$$

$$\sigma(\mathbf{U}_t^\flat) \subset \sigma(\mathbf{X}_0, \mathbf{W}_1, \ldots, \mathbf{W}_t) \ , \ \ \forall t \in [\![0, T-1]\!] \ , \tag{45b}$$

$$\sigma(\mathbf{U}_t^\sharp) \subset \sigma(\mathbf{X}_0, \mathbf{W}_1, \ldots, \mathbf{W}_t) \ , \ \ \forall t \in [\![1, T]\!] \ , \tag{45c}$$

$$\mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t^\flat, \mathbf{U}_{t+1}^\sharp, \mathbf{W}_{t+1}) \ , \ \ \forall t \in [\![0, T-1]\!] \ . \tag{45d}$$

Thus, in the above setting, during the time interval between two time steps, the decision-maker makes two decisions. At the end of the time interval $[t-1, t[$, a random variable $\mathbf{W}_t$ is revealed, and then, at the beginning of the time interval $[t, t+1[$, the decision-maker makes a *head decision* $\mathbf{U}_t^\sharp$. What is new — in comparison with the classical decision-hazard framework — is that, at the end of the time interval $[t, t+1[$, when a next random variable $\mathbf{W}_{t+1}$ is revealed, the decision-maker has the possibility to make a *tail decision* $\mathbf{U}_{t+1}^\flat$. This latter decision $\mathbf{U}_{t+1}^\flat$ can be thought as a *recourse* variable for a two stage stochastic optimization problem that would take place inside the time interval $[t, t+1[$.

Proposition 14 *Assume that the spaces* $\{(\mathbb{U}_t^\flat, \mathcal{U}_t^\flat)\}_{t \in [\![0,T-1]\!]}$, $\{(\mathbb{U}_t^\sharp, \mathcal{U}_t^\sharp)\}_{t \in [\![1,T]\!]}$, $\{(\mathbb{W}_t, \mathcal{W}_t)\}_{t \in [\![1,T]\!]}$ *and* $\{(\mathbb{X}_t, \mathcal{X}_t)\}_{t \in [\![0,T]\!]}$ *are Borel spaces, that all mappings in* (44) *are Borel-measurable, and that all functions in* (44) *are lower semianalytic.*

If the a random variable $\mathbf{X}_0$ *is deterministic with value* x_0*, and if the noise process* $\{\mathbf{W}_t\}_{t \in [\![1,T]\!]}$ *is white, that is, is made of independent random variables, then the value* $V_t(x_0)$ *of the multistage stochastic optimization problem* (45) *is given by the dynamic programming backward induction*

$$V_T(x) = K(x) \ , \tag{46a}$$

$$V_t(x) = \inf_{u_t^\sharp \in \mathbb{U}_t^\sharp} \mathbb{E}\Big[\inf_{u_{t+1}^\flat \in \mathbb{U}_{t+1}^\flat} \Big\{ L_t(x_t, u_t^\flat, u_{t+1}^\sharp, \mathbf{W}_{t+1}) + V_{t+1}\big(f_t(x_t, u_t^\flat, u_{t+1}^\sharp, \mathbf{W}_{t+1})\big) \Big\}\Big] \ . \tag{46b}$$

Theorem 19 allows to develop dynamic programming equations in the decision-hazard-decision framework. Such equations can be solved using the stochastic dual dynamic programming (SDDP) algorithm provided that convexity of the value functions is preserved. This requires linearity in the dynamics, a feature that may be recovered by modeling the problem in the decision-hazard-decision framework as illustrated in §6.1.

7 Conclusion and perspectives

As said in the introduction, decomposition methods are appealing to tackle multistage stochastic optimization problems, as they are naturally large scale. The most common approaches are time decomposition (and state-based resolution methods, like stochastic dynamic programming, in stochastic optimal control), and scenario decomposition (like progressive hedging in stochastic programming).

This paper is part of a general research program that consists in *mixing* different decomposition bricks. Space decomposition methods have been investigated in [1] and [5]. Here, we have tackled the issue of using time blocks decomposition in such a way that stochastic dynamic programming is used at the slow time scale with an appropriate white noise assumption, whereas stochastic programming methods such as progressive hedging can be used at the fast time scale where such an independence assumption does not hold. This approach paves the way of mixing time decomposition with scenario decomposition. For this purpose, we have revisited the notion of state, and have provided a way to perform time decomposition but only across specified time blocks.

Acknowledgements. We thank Roger Wets for fruitful discussions about the possibility of mixing stochastic dynamic programming with Progressive Hedging.

A Background on stochastic dynamic programming

We first recall the notion of stochastic kernel, used in the modeling of stochastic control problems. Let $(\mathbb{X}, \mathcal{X})$ and $(\mathbb{Y}, \mathcal{Y})$ be two measurable spaces. A *stochastic kernel* from $(\mathbb{X}, \mathcal{X})$ to $(\mathbb{Y}, \mathcal{Y})$ is a function $\rho : \mathbb{X} \times \mathcal{Y} \to [0, 1]$ such that, for any $Y \in \mathcal{Y}$, the function $\rho(\cdot, Y) : \mathbb{X} \to [0, 1]$ is $\mathcal{X}$-measurable and, for any $x \in \mathbb{X}$, the function $\rho(x, \cdot) : \mathcal{Y} \to [0, 1]$ is a probability measure. By a slight abuse of notation, a stochastic kernel is also denoted as a mapping $\rho : \mathbb{X} \to \Delta(\mathbb{Y})$ from the measurable space $(\mathbb{X}, \mathcal{X})$ towards the space $\Delta(\mathbb{Y})$ of probability measures over $(\mathbb{Y}, \mathcal{Y})$, with the property that the function $x \in \mathbb{X} \mapsto \int_Y \rho(\,\mathrm{d}y \,|\, x)$ is measurable for any $Y \in \mathcal{Y}$.

We now sketch the most classical frameworks for stochastic dynamic programming in discrete time. In what follows, $t_0 \in \mathbb{N}$ and $T \in \mathbb{N}^*$ are two natural numbers such that $t_0 < T$.

Witsenhausen approach. The most general stochastic dynamic programming principle is sketched by Witsenhausen at the end of [25]. However, we do not detail it as its formalism

is too far from the following ones, though we will touch the subject when we discuss Yüksel's approach below. We present here what Witsenhausen calls an optimal stochastic control problem in *standard form* (see [23]). The ingredients are the following:

1. time $t \in [\![t_0, T]\!]$ is discrete and runs among a finite set of consecutive natural numbers;
2. $(\mathbb{X}_{t_0}, \mathcal{X}_{t_0})$ (nature), $(\mathbb{X}_{t_0+1}, \mathcal{X}_{t_0+1})$, ..., $(\mathbb{X}_T, \mathcal{X}_T)$ (state spaces) are measurable spaces;
3. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0})$,..., $(\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (control spaces);
4. $\mathcal{I}_t$ is a subfield of $\mathcal{X}_t$, for $t \in [\![t_0, T-1]\!]$ (information);
5. $f_t : (\mathbb{X}_t \times \mathbb{U}_t, \mathcal{X}_t \otimes \mathcal{U}_t) \to (\mathbb{X}_{t+1}, \mathcal{X}_{t+1})$ is measurable, for $t \in [\![t_0, T-1]\!]$ (dynamics);
6. π_{t_0} is a probability on $(\mathbb{X}_{t_0}, \mathcal{X}_{t_0})$;
7. $j : (\mathbb{X}_T, \mathcal{X}_T) \to \mathbb{R}$ is a measurable function (criterion).

With these ingredients, Witsenhausen formulates a stochastic optimization problem, whose solutions are to be searched among adapted feedbacks, namely $\lambda_t : (\mathbb{X}_t, \mathcal{X}_t) \to (\mathbb{U}_t, \mathcal{U}_t)$ with the property that $\lambda_t^{-1}(\mathcal{U}_t) \subset \mathcal{I}_t$ for all $t \in [\![t_0, T-1]\!]$. Then, he establishes a dynamic programming equation, where the Bellman functions are function of the (unconditional) distribution of the original state $x_t \in \mathbb{X}_t$, and where the minimization is done over adapted feedbacks. The main objective of Witsenhausen is to establish a dynamic programming equation for nonclassical information patterns.

Evstigneev approach. The ingredients of the approach developed in [7] are the following:

1. time $t \in [\![t_0, T]\!]$ is discrete and runs among a finite set of consecutive natural numbers;
2. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0})$,..., $(\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (control spaces);
3. $(\Omega, \mathcal{F})$ is a measurable space (nature);
4. $\left\{\mathcal{F}_t\right\}_{t \in [\![t_0, T-1]\!]}$ is a filtration of $\mathcal{F}$ (information);
5. $\mathbb{P}$ is a probability on $(\Omega, \mathcal{F})$;
6. $j : (\Omega \times \prod_{t \in [\![t_0, T-1]\!]} \mathbb{U}_t, \mathcal{F} \otimes \bigotimes_{t \in [\![t_0, T-1]\!]} \mathcal{U}_t) \to \mathbb{R}$ is a measurable function (criterion).

With these ingredients, Evstigneev formulates a stochastic optimization problem, whose solutions are to be searched among adapted processes, namely random processes with values in $\prod_{t \in [\![t_0, T-1]\!]} \mathbb{U}_t$ and adapted to the filtration $\left\{\mathcal{F}_t\right\}_{t \in [\![t_0, T-1]\!]}$. Then, he establishes a dynamic programming equation, where the Bellman function at time t is an $\mathcal{F}_t$-integrand depending on controls up to time t (random variables) and where the minimization is done over $\mathcal{F}_t$-measurable random variables at time t. The main objective of Evstigneev is to establish an existence theorem for an optimal adapted process (under proper technical assumptions, especially on the objective function j, that we do not detail here). Notice that there is no notion of state variable.

Puterman approach. The ingredients of the approach developed in [15, Sect. 2.1] are the following:

1. time $t \in [\![t_0, T]\!]$ is discrete and runs among a finite set of consecutive natural numbers;
2. $(\mathbb{X}_{t_0}, \mathcal{X}_{t_0})$, ..., $(\mathbb{X}_T, \mathcal{X}_T)$ are measurable spaces (state spaces);
3. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0})$,..., $(\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (control spaces);
4. $\nu_{t:t+1} : \mathbb{X}_t \times \mathbb{U}_t \to \Delta(\mathbb{X}_{t+1})$ is a stochastic kernel, for $t \in [\![t_0, T-1]\!]$ (transitions);
5. $L_t : \mathbb{X}_t \times \mathbb{U}_t \to \mathbb{R}$, for $t \in [\![t_0, T-1]\!]$, and $K : \mathbb{X}_T \to \mathbb{R}$, are measurable functions (instantaneous and final costs).

With these ingredients, Puterman formulates a stochastic optimization problem with a time additive cost function over given state and control spaces, whose solutions are to be searched among history feedbacks, namely sequences of mappings $\mathbb{X}_{t_0} \times \prod_{s=t_0}^{t-1}(\mathbb{U}_s \times \mathbb{X}_{s+1}) \to \mathbb{U}_t$. Then, he establishes a dynamic programming equation, where the Bellman functions are function of the history $h_t \in \mathbb{X}_{t_0} \times \prod_{s=t_0}^{t-1}(\mathbb{U}_s \times \mathbb{X}_{s+1})$. He identifies cases where no loss of optimality results from reducing the search to Markovian feedbacks $\mathbb{X}_t \to \mathbb{U}_t$. In such cases, the Bellman functions are function of the state $x_t \in \mathbb{X}_t$, and the minimization in the dynamic programming equation is done over controls $u_t \in \mathbb{U}_t$. The main objective of Puterman is to explore infinite horizon criteria, average reward criteria, the continuous time case, and to present many examples.

Hernández-Lerma and Lasserre approach. The ingredients of the approach developed in [9, §2.2, §3.2, §3.3] are the following:

1. time $t \in [\![t_0, T]\!]$ is discrete and runs among a finite set of consecutive natural numbers;
2. $(\mathbb{X}_{t_0}, \mathcal{X}_{t_0})$, ..., $(\mathbb{X}_T, \mathcal{X}_T)$ are Borel spaces (state spaces);
3. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0})$,..., $(\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are Borel spaces (control spaces); there are also feasible state-dependent control constraints that we do not present here;
4. $\nu_{t:t+1} : \mathbb{X}_t \times \mathbb{U}_t \to \Delta(\mathbb{X}_{t+1})$, for $t \in [\![t_0, T-1]\!]$, are Borel-measurable stochastic kernels (transitions);
5. $L_t : \mathbb{X}_t \times \mathbb{U}_t \to \mathbb{R}$, for $t \in [\![t_0, T-1]\!]$, and $K : \mathbb{X}_T \to \mathbb{R}$ are Borel-measurable functions (instantaneous and final costs).

With these ingredients, Hernández-Lerma and Lasserre formulate a stochastic optimization problem with a time additive cost function over given state and control spaces. They introduce the "canonical construction" where the history at time t consists in the states and the controls prior to t. Then, they study optimization problems whose solutions (policies) are to be searched among history feedbacks (or randomized history feedbacks), namely sequences of mappings $\mathbb{X}_{t_0} \times \prod_{s=t_0}^{t-1}(\mathbb{U}_s \times \mathbb{X}_{s+1}) \to \mathbb{U}_t$. They identify cases where no loss of optimality

results from reducing the search to (relaxed) Markovian feedbacks $\mathbb{X}_t \to \mathbb{U}_t$. Then, they establish a dynamic programming equation, where the Bellman functions are function of the state $x_t \in \mathbb{X}_t$, and where the minimization is done over controls $u_t \in \mathbb{U}_t$. For finite horizon problems, the mathematical challenge is to set up a mathematical framework — the Borel assumptions plus additional topological ones presented in [9, §3.3] — for which optimal policies exists. The main objective of [9] is to offer a unified and comprehensive treatment of discrete-time Markov control processes, with emphasis on the case of Borel state and control spaces, and possibly unbounded costs and noncompact control constraint sets.

Bertsekas and Shreve approach. The ingredients of the approach developed in [3] (more precisely in [3, Definition 10.1]) are the following:

1. time $t \in [\![t_0, T]\!]$ is discrete and runs among a finite set of consecutive natural numbers;
2. $(\mathbb{X}_{t_0}, \mathcal{X}_{t_0})$, ..., $(\mathbb{X}_T, \mathcal{X}_T)$ are Borel spaces (state spaces);
3. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0})$,..., $(\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are Borel spaces (control spaces); there are also feasible state-dependent control constraints that we do not present here;
4. $(\mathbb{W}_{t_0}, \mathcal{W}_{t_0})$,..., $(\mathbb{W}_T, \mathcal{W}_T)$ are Borel spaces (noise);
5. $f_t : (\mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_t, \mathcal{X}_t \otimes \mathcal{U}_t \otimes \mathcal{W}_t) \to (\mathbb{X}_{t+1}, \mathcal{X}_{t+1})$, for $t \in [\![t_0, T-1]\!]$, are Borel-measurable mappings (dynamics);
6. $\rho_{t:t+1} : \mathbb{X}_t \times \mathbb{U}_t \to \Delta(\mathbb{W}_{t+1})$, for $t \in [\![t_0, T-1]\!]$, are Borel-measurable stochastic kernels (noise distributions);
7. $L_t : \mathbb{X}_t \times \mathbb{U}_t \to \mathbb{R}$, for $t \in [\![t_0, T-1]\!]$, and $K : \mathbb{X}_T \to \mathbb{R}$ are lower semianalytic functions (instantaneous and final costs).

With these ingredients, Bertsekas and Shreve formulate a stochastic optimization problem with a time additive cost function over given state spaces, control spaces and uncertainty spaces. They introduce the notion of history at time t which consists in the states and the controls prior to t and study optimization problems whose solutions (policies) are to be searched among history feedbacks (or relaxed history feedbacks), namely sequences of mappings from history space $\mathbb{X}_{t_0} \times \prod_{s=t_0}^{t-1} (\mathbb{U}_s \times \mathbb{X}_{s+1}) \to \mathbb{U}_t$. They identify cases where no loss of optimality results from reducing the search to (relaxed) Markovian feedbacks $\mathbb{X}_t \to \mathbb{U}_t$. Then, they establish a dynamic programming equation, where the Bellman functions are function of the state $x_t \in \mathbb{X}_t$, and where the minimization is done over controls $u_t \in \mathbb{U}_t$. For finite horizon problems, the mathematical challenge is to set up a mathematical framework (the Borel assumptions) for which optimal policies exists. The main objective of Bertsekas and Shreve is to state conditions under which the dynamic programming equation is mathematically sound, namely with universally measurable Bellman functions and with universally measurable relaxed control strategies in the context of Borel spaces. The interested reader will find all the subtleties about Borel spaces and universally measurable concepts in [3, Chapter 7].

Yüksel approach. As said at the beginning, the most general stochastic dynamic programming principle is sketched by Witsenhausen at the end of [25]. This approach builds upon the so-called Witsenhausen intrinsic model [24] which does not consider state, but information under the form of σ-fields (see [26] for the functional form). In [23], Witsenhausen provides conditions to express stochastic control optimization problems — with information constraints, but without state — in standard form with a state (the first approach that we have considered above).

Although Witsenhausen established a dynamic programming equation in [23], Yüksel notes in [28] that "Witsenhausen's construction [...] does not address the well-posedness of such a dynamic program" and that "the existence problem was not considered". In the spirit of [23], Yüksel entails in [28] "a general approach establishing that any sequential team optimization may admit a formulation appropriate for a dynamic programming analysis". One of the contributions of [28] is to propose a construction of standard Borel controlled state and action spaces and to establish a universal dynamic program for stochastic control optimization problems — with information constraints, but without state — thus addressing some of the issues raised and left open by Witsenhausen. The ingredients are the following:

1. time $t \in [\![t_0, T]\!]$ is discrete and runs among a finite set of consecutive natural numbers;
2. $(\Omega, \mathcal{F})$ is a measurable space (nature);
3. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0})$,..., $(\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (control spaces);
4. $(\mathbb{Y}_{t_0}, \mathcal{Y}_{t_0})$, ..., $(\mathbb{Y}_{T-1}, \mathcal{Y}_{T-1})$ are measurable spaces ("observation" spaces);
5. $\big\{\eta_t : (\Omega \times \prod_{s\in[\![t_0,t]\!]} \mathbb{U}_s \ , \mathcal{F} \otimes \bigotimes_{s\in[\![t_0,t]\!]} \mathcal{U}_s) \to (\mathbb{U}_t, \mathcal{U}_t)\big\}_{t\in[\![t_0,T-1]\!]}$ are measurable mappings ("measurement constraints");
6. $\mathbb{P}$ is a probability on $(\Omega, \mathcal{F})$;
7. $j : (\Omega \times \prod_{t\in[\![t_0,T-1]\!]} \mathbb{U}_t, \mathcal{F} \otimes \bigotimes_{t\in[\![t_0,T-1]\!]} \mathcal{U}_t) \to \mathbb{R}_+$ is a measurable function (criterion).

With these ingredients, Yüksel formulates a stochastic team optimization problem whose solutions (policies) are to be searched among sequences of measurable mappings ("design constraints") $\mathbb{Y}_{t-1} \to \mathbb{U}_t$, and their "randomized" versions (so-called strategic measures). He establishes a dynamic programming equation, where the Bellman functions are function of probability distributions and where the minimization is done over proper design mappings. One objective of Yüksel is to set up a mathematical framework under which the dynamic programming equation is mathematically sound [28, Theorem 3.6].

Our approach. The ingredients that we use (in Sect. 2 and in Sect. 3) are the following:

1. time $t \in [\![t_0, T]\!]$ is discrete and runs among a finite set of consecutive natural numbers;
2. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0})$,..., $(\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (control spaces);

3. $(\mathbb{W}_{t_0}, \mathcal{W}_{t_0}),\ldots, (\mathbb{W}_T, \mathcal{W}_T)$ are measurable spaces (noise);

4. $\rho_{t:t+1} : \mathbb{W}_{t_0} \times \prod_{s=t_0}^{t-1}(\mathbb{U}_s \times \mathbb{W}_{s+1}) \to \Delta(\mathbb{W}_{t+1})$, for $t \in [\![t_0, T-1]\!]$, are stochastic kernels (noise distributions);

5. $j : (\mathbb{W}_{t_0} \times \prod_{s=t_0}^{T-1}(\mathbb{U}_s \times \mathbb{W}_{s+1}), \mathcal{W}_{t_0} \otimes \bigotimes_{s=t_0}^{T-1}(\mathcal{U}_s \otimes \mathcal{W}_{s+1})) \to [0, +\infty]$ is a measurable function (criterion);

6. $t_0 < \cdots < t_N = T$ are the indices of multiple consecutive time blocks $[\![t_0, t_1]\!]$, $\ldots$, $[\![t_{N-1}, t_N]\!]$, with $N \geq 1$ a natural number;

7. $\big\{(\mathbb{X}_{t_j}, \mathcal{X}_{t_j})\big\}_{j\in[\![0,N]\!]}$ are measurable spaces (time block state spaces);

8. $\Big\{\theta_{t_j} : \mathbb{W}_{t_0} \times \prod_{s=t_0}^{t_j-1}(\mathbb{U}_s \times \mathbb{W}_{s+1}) \to \mathbb{X}_{t_j}\Big\}_{j\in[\![1,N]\!]}$ and $\theta_{t_0} : \mathbb{W}_{t_0} \to \mathbb{X}_{t_0}$ are measurable mappings (time block reduction of history towards state);

9. $\Big\{f_{t_j:t_{j+1}} : \mathbb{X}_{t_j} \times \prod_{s=t_j}^{t_{j+1}-1}(\mathbb{U}_s \times \mathbb{W}_{s+1}) \to \mathbb{X}_{t_{j+1}}\Big\}_{j\in[\![0,N-1]\!]}$ are measurable mappings (time block dynamics).

The framework developed in this paper is intermediate between the ones of Evstigneev in [7] and of Yüksel in [28] — notable by the absence of a state space — and the ones of Witsenhausen [23], Hernández-Lerma and Lasserre [9], Bertsekas and Shreve [3] and Puterman [15] — where the state spaces are given for all times.

This said, our preoccupation could be adapted to any of the above frameworks. Indeed, our objective is to establish a dynamic programming equation with a state, not at any time $t \in [\![t_0, T]\!]$, but at some specified instants $t_0 < t_1 < \cdots < t_N = T$. In §3.2, the state spaces are introduced as image sets (codomains) of what we call *(time block) history reduction mappings* (where history at time t consists of all uncertainties and controls prior to time t).

B Technical details and proofs (Sect. 2)

We introduce the notations

$$\mathbb{W}_{r:t} = \prod_{s=r}^{t} \mathbb{W}_s \ , \ \ 0 \leq r \leq t \leq T \ ,$$
$$\mathbb{U}_{r:t} = \prod_{s=r}^{t} \mathbb{U}_s \ , \ \ 0 \leq r \leq t \leq T-1 \ . \tag{47}$$

Let $0 \leq r \leq s \leq t \leq T$. From a history $h_t \in \mathbb{H}_t$, we can extract the $(r\,{:}\,s)$-*history uncertainty part*

$$[h_t]^{\mathbb{W}}_{r:s} = (w_r, \ldots, w_s) = w_{r:s} \in \mathbb{W}_{r:s} \ , \ \ 0 \leq r \leq s \leq t \ , \tag{48a}$$

the $(r\!:\!s)$-*history control part* (notice that the indices are special)

$$\begin{aligned}[h_t]_{r:s}^{\mathbb{U}} &= (u_{r-1}, \dots, u_{s-1}) \\ &= u_{r-1:s-1} \in \mathbb{U}_{r-1:s-1} \ , \ \ 1 \leq r \leq s \leq t \ . \end{aligned} \tag{48b}$$

Flows. Let r and t be given such that $0 \leq r < t \leq T$. For a $(r\!:\!t-1)$-history feedback $\gamma = \{\gamma_s\}_{s=r,\dots,t-1} \in \Gamma_{r:t-1}$, we define the *flow* $\Phi_{r:t}^{\gamma}$ by

$$\begin{aligned} &\Phi_{r:t}^{\gamma} : \mathbb{H}_r \times \mathbb{W}_{r+1:t} \to \mathbb{H}_t \\ &\qquad (h_r, w_{r+1:t}) \mapsto \Big(h_r, \gamma_r(h_r), w_{r+1}, \gamma_{r+1}\big(h_r, \gamma_r(h_r), w_{r+1}\big), w_{r+2}, \cdots, \gamma_{t-1}(h_{t-1}), w_t\Big) \ . \end{aligned} \tag{49a}$$

Otherwise stated, the flow is given by

$$\Phi_{r:t}^{\gamma}(h_r, w_{r+1:t}) = (h_r, u_r, w_{r+1}, u_{r+1}, w_{r+2}, \dots, u_{t-1}, w_t) \ , \tag{49b}$$

$$\text{with} \ \ h_s = (h_r, u_r, w_{r+1}, \dots, u_{s-1}, w_s) \ , \ \ r < s \leq t \ , \tag{49c}$$

$$\text{and} \ \ u_s = \gamma_s(h_s) \ , \ \ r \leq s \leq t-1 \ . \tag{49d}$$

When $0 \leq r = t \leq T$, we put $\Phi_{r:r}^{\gamma} : \mathbb{H}_r \to \mathbb{H}_r$, $h_r \mapsto h_r$. With this convention, the expression $\Phi_{r:t}^{\gamma}$ makes sense when $0 \leq r \leq t \leq T$. The mapping $\Phi_{r:t}^{\gamma}$ gives the history at time t as a function of the initial history h_r at time r and of the history feedbacks $\{\gamma_s\}_{s=r,\dots,t-1} \in \Gamma_{r:t-1}$. An immediate consequence of this definition are the *flow properties*:

$$\Phi_{r:t+1}^{\gamma}(h_r, w_{r+1:t+1}) = \Big(\Phi_{r:t}^{\gamma}(h_r, w_{r+1:t}), \gamma_t\big(\Phi_{r:t}^{\gamma}(h_r, w_{r+1:t})\big), w_{t+1}\Big) \ , \ \ 0 \leq r \leq t \leq T-1 \ , \tag{50a}$$

$$\Phi_{r:t}^{\gamma}(h_r, w_{r+1:t}) = \Phi_{r+1:t}^{\gamma}\big((h_r, \gamma_r(h_r), w_{r+1}), w_{r+2:t}\big) \ , \ \ 0 \leq r < t \leq T \ . \tag{50b}$$

Definition 15 *Let r and t be given such that $0 \leq r \leq t \leq T$.*

- *When $0 \leq r < t \leq T$, for a $(r\!:\!t-1)$-history feedback $\gamma = \{\gamma_s\}_{s \in [\![r,t-1]\!]} \in \Gamma_{r:t-1}$, and for a family $\{\rho_{s-1:s}\}_{r+1 \leq s \leq t}$ of stochastic kernels $\rho_{s-1:s} : \mathbb{H}_{s-1} \to \Delta(\mathbb{W}_s)$, $s \in [\![r+1, t]\!]$, we define a stochastic kernel $\rho_{r:t}^{\gamma} : \mathbb{H}_r \to \Delta(\mathbb{H}_t)$ such that, for any numerical function $\varphi \in \mathbb{L}_+^0(\mathbb{H}_t)$[9], we have that*

$$\begin{aligned} &\int_{\mathbb{H}_t} \varphi(h'_r, h'_{r+1:t}) \rho_{r:t}^{\gamma}(\,\mathrm{d}h'_t \,|\, h_r) \\ &\qquad = \int_{\mathbb{W}_{r+1:t}} \varphi\big(\Phi_{r:t}^{\gamma}(h_r, w_{r+1:t})\big) \prod_{s=r+1}^{t} \rho_{s-1:s}\big(\,\mathrm{d}w_s \,\big|\, \Phi_{r:s-1}^{\gamma}(h_r, w_{r+1:s-1})\big) \ . \end{aligned} \tag{51}$$

- *When $0 \leq r = t \leq T$, we define $\rho_{r:r}^{\gamma} : \mathbb{H}_r \to \Delta(\mathbb{H}_r)$ by $\rho_{r:r}^{\gamma}(\,\mathrm{d}h'_r \,|\, h_r) = \delta_{h_r}(\,\mathrm{d}h'_r)$.*

[9] space of universally measurable nonnegative numerical functions over $\mathbb{H}_t$: see Footnote 2

The stochastic kernels $\rho^{\gamma}_{r:t}$ on $\mathbb{H}_t$, given by (51), are of the form $\rho^{\gamma}_{r:t}(\,\mathrm{d}h'_t \,|\, h_r) = \rho^{\gamma}_{r:t}(\,\mathrm{d}h'_r\,\mathrm{d}h'_{r+1:t} \,|\, h_r) = \delta_{h_r}(\,\mathrm{d}h'_r) \otimes \varrho^{\gamma}_{r:t}(\,\mathrm{d}h'_{r+1:t} \,|\, h_r)$, where, for each $h_r \in \mathbb{H}_r$, the probability distribution $\varrho^{\gamma}_{r:t}(\,\mathrm{d}h'_{r+1:t} \,|\, h_r)$ only charges the histories visited by the flow from $r+1$ to t. The construction of the stochastic kernels $\rho^{\gamma}_{r:t}$ is developed in [3, p. 190] for relaxed history feedbacks and obtained by using [3, Proposition 7.45].

Proposition 16 *The family $\{\rho^{\gamma}_{s:t}\}_{s=r,\dots,t}$ of stochastic kernels of Definition 15 has the flow property:*

$$\rho^{\gamma}_{s:t}(\,\mathrm{d}h'_t \,|\, h_s) = \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}(\,\mathrm{d}w_{s+1} \,|\, h_s)\rho^{\gamma}_{s+1:t}\Big(\,\mathrm{d}h'_t \,\Big|\, \big(h_s, \gamma_s(h_s), w_{s+1}\big)\Big)\ ,\ \ \forall s < t\ . \tag{52}$$

Proof. Let $s < t$. For any $\varphi \in \mathbb{L}^0_+(\mathbb{H}_t)$, we have that

$$\begin{aligned}
&\int_{\mathbb{H}_t} \varphi(h'_s, h'_{s+1:t})\rho^{\gamma}_{s:t}(\,\mathrm{d}h'_t \,|\, h_s) && \text{(53a)}\\
&\quad= \int_{\mathbb{W}_{s+1:t}} \varphi\big(\Phi^{\gamma}_{s:t}(h_s, w_{s+1:t})\big) \\
&\qquad\qquad \prod_{s'=s+1}^{t} \rho_{s'-1:s'}\big(\,\mathrm{d}w_{s'} \,\big|\, \Phi^{\gamma}_{s:s'-1}(h_s, w_{s+1:s'-1})\big) && \text{(by Definition (51))}\\
&\quad= \int_{\mathbb{W}_{s+1:t}} \varphi\big(\Phi^{\gamma}_{s:t}(h_s, w_{s+1:t})\big)\rho_{s:s+1}\big(\,\mathrm{d}w_{s+1} \,\big|\, h_s\big) \\
&\qquad\qquad \prod_{s'=s+2}^{t} \rho_{s'-1:s'}\big(\,\mathrm{d}w_{s'} \,\big|\, \Phi^{\gamma}_{s:s'-1}(h_s, w_{s+1:s'-1})\big)\\
&\quad= \int_{\mathbb{W}_{s+1:t}} \varphi\big(\Phi^{\gamma}_{s+1:t}\big((h_s, \gamma_s(h_s), w_{s+1}), w_{s+2:t}\big)\big) \\
&\qquad \rho_{s:s+1}\big(\,\mathrm{d}w_{s+1} \,\big|\, h_s\big) \\
&\qquad\quad \prod_{s'=s+2}^{t} \rho_{s'-1:s'}\big(\,\mathrm{d}w_{s'} \,\big|\, \Phi^{\gamma}_{s+1:s'-1}\big((h_s, \gamma_s(h_s), w_{s+1}), w_{s+2:s'-1}\big)\big) && \text{(by the flow property (50b))}\\
&\quad= \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}\big(\,\mathrm{d}w_{s+1} \,\big|\, h_s\big) \\
&\qquad\quad \int_{\mathbb{W}_{s+2:t}} \varphi\big(\Phi^{\gamma}_{s+1:t}\big((h_s, \gamma_s(h_s), w_{s+1}), w_{s+2:t}\big)\big) \\
&\qquad\qquad \prod_{s'=s+2}^{t} \rho_{s'-1:s'}\big(\,\mathrm{d}w_{s'} \,\big|\, \Phi^{\gamma}_{s+1:s'-1}\big((h_s, \gamma_s(h_s), w_{s+1}), w_{s+2:s'-1}\big)\big)
\end{aligned}$$

by Fubini Theorem [12, p.137]

$$
\begin{aligned}
&= \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}\big(\,\mathrm{d}w_{s+1} \,\big|\, h_s\big) \int_{\mathbb{H}_t} \varphi\big((h'_s, \gamma_s(h'_s), w'_{s+1}), h'_{s+2:t}\big) \\
&\qquad \rho^{\gamma}_{s+1:t}\big(\,\mathrm{d}h'_t \,\big|\, (h_s, \gamma_s(h_s), w_{s+1})\big) \qquad \text{(by Definition (51))} \\
&= \int_{\mathbb{H}_t} \varphi\big((h'_s, \gamma_s(h'_s), w'_{s+1}), h'_{s+2:t}\big) \\
&\qquad \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}\big(\,\mathrm{d}w_{s+1} \,\big|\, h_s\big) \rho^{\gamma}_{s+1:t}\big(\,\mathrm{d}h'_t \,\big|\, (h_s, \gamma_s(h_s), w_{s+1})\big) \qquad \text{(53b)}
\end{aligned}
$$

by Fubini Theorem. As the two expressions (53a) and (53b) are equal for any $\varphi \in \mathbb{L}^0_+(\mathbb{H}_t)$, we deduce the flow property (52). □

Proof of Theorem 2. We only give a sketch of the proof, as it is a variation on different results of [3], the framework of which we follow.

Proof. By assumption, the optimization data set $\mathcal{D}$ in (4) is Borel-analytic, so that we are in the setting of [3, Chap. 7, Chap. 8].

We take the history space $\mathbb{H}_t$ for state space, and the state dynamics

$$
f\big(h_t, u_t, w_{t+1}\big) = \big(h_t, u_t, w_{t+1}\big) = h_{t+1} \in \mathbb{H}_{t+1} = \mathbb{H}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \,. \tag{54}
$$

Then, the family $\{\rho_{s-1:s}\}_{s\in[\![1,T]\!]}$ of stochastic kernels (2) gives a family of disturbance kernels that do not depend on the current control. The criterion to be minimized (3) is a function of the history at time T, thus of the state at time T. The optimization problem defined by the associated value function (6)

$$
\begin{aligned}
V_0(w_0) &= \inf_{\gamma_{0:T-1}\in\Gamma_{0:T-1}} \int_{\mathbb{H}_T} j(h'_T) \rho^{\gamma}_{0:T}(\,\mathrm{d}h'_T \,|\, w_0) \\
&= \inf_{\gamma_{0:T-1}\in\Gamma_{0:T-1}} \int_{\mathbb{W}_{1:T}} j\big(\Phi^{\gamma}_{0:T}(w_{0:T})\big) \prod_{s=1}^{T} \rho_{s-1:s}\big(\,\mathrm{d}w_s \,\big|\, \Phi^{\gamma}_{0:s-1}(w_{0:s-1})\big) \,,
\end{aligned}
$$

by (51) — where the flows $\Phi^{\gamma}_{0:s}$ for $s \in [\![0, T-1]\!]$ are defined by Equation (49a) — is thus a finite horizon model with a final cost, and we are minimizing over the so-called state-feedbacks. Then, the proof of Theorem 2 follows from the results developed in Chap. 7, 8 and 10 of [3] in a Borel setting.

The Bellman operators in (7a) satisfy (7b) because — as the optimization data set $\mathcal{D}$ in (4) is Borel-analytic, and by Lemma 7.30(4) and Propositions 7.47 and 7.48 as stated in [3, p. 196] — we have that $\varphi \in \mathbb{L}^0_+(\mathbb{H}_{t+1}) \implies \mathcal{B}_{t+1:t}\varphi \in \mathbb{L}^0_+(\mathbb{H}_t)$, for t in $[\![0, T-1]\!]$.

Since we are considering a finite horizon model with a final cost, we detail the steps needed to use the results of [3, Chap. 8]. The final cost at time T can be turned into an instantaneous cost at time $T-1$ by inserting the state dynamics (54) in the final cost. Getting rid of the disturbance in the expected cost by using the disturbance kernel is standard practice. Then, we can turn this non-homogeneous finite horizon model into a finite horizon model with homogeneous dynamics and costs by following the steps of [3, Chap. 10]. Using [3, Proposition 8.2], we obtain that the family

of optimization problems defined by the associated value functions (5), when minimizing over the relaxed state feedbacks, satisfies the Bellman equation (8); we conclude with [3, Proposition 8.4] which covers the minimization over state feedbacks. □

To summarize, Theorem 2 is valid under the general Borel assumptions of [3, Chap. 8] and with the specific (F^-) assumption needed for [3, Proposition 8.4]; this last assumption is fulfilled here since we have assumed that the criterion (3) is nonnegative.

C Technical details and proofs (Sect. 3)

Proof of Proposition 4.

Proof. By assumption, the optimization data set $\mathcal{D}$ in (4) is Borel-analytic and the single time block compatible reduction data set $\mathcal{R}_{r,t}$ in (11) is Borel. Thus, we are in the setting of [3, Chap. 7, Chap. 8].

Let $\tilde{\varphi}_t : \mathbb{X}_t \to [0, +\infty]$ be a given lower semianalytic nonnegative numerical function, and let $\varphi_t : \mathbb{H}_t \to [0, +\infty]$ be

$$\varphi_t = \tilde{\varphi}_t \circ \theta_t \ . \tag{55}$$

Let $\varphi_r : \mathbb{H}_r \to [0, +\infty]$ be the lower semianalytic nonnegative numerical function obtained by applying the Bellman operator $\mathcal{B}_{t:r}$ across $[\![t, r]\!]$ (see (12)) to the lower semianalytic nonnegative numerical function φ_t:

$$\varphi_r = \mathcal{B}_{t:r}\varphi_t = \mathcal{B}_{r+1:r} \circ \cdots \circ \mathcal{B}_{t:t-1}\varphi_t \ . \tag{56}$$

By [3, Lemma 7.30(3)] — on the stability of lower semianalytic functions under right composition with a Borel mapping — we have that the nonnegative numerical function φ_r is lower semianalytic. We show that there exists a lower semianalytic nonnegative numerical function $\tilde{\varphi}_r : \mathbb{X}_r \to [0, +\infty]$ such that

$$\varphi_r = \tilde{\varphi}_r \circ \theta_r \ . \tag{57}$$

First, we show by backward induction that, for all $s \in [\![r, t]\!]$, there exists a measurable nonnegative numerical function $\overline{\varphi}_s$ such that $\varphi_s(h_s) = \overline{\varphi}_s(\theta_r(h_r), h_{r+1:s})$. Second, we prove that the function $\tilde{\varphi}_r = \overline{\varphi}_r$ satisfies (57) and is lower semianalytic.

- For $s = t$, we have, by (55) and by (9c), that $\varphi_t(h_t) = \tilde{\varphi}_t\big(\theta_t(h_t)\big) = \tilde{\varphi}_t\big(f_{r:t}(\theta_r(h_r), h_{r+1:t})\big)$, so that the nonnegative numerical function $\overline{\varphi}_t$ is given by $\tilde{\varphi}_t \circ f_{r:t}$. By [3, Lemma 7.30(3)], $\overline{\varphi}_t$ is a lower semianalytic numerical function.

- Assume that, at $s + 1$, the result holds true, that is, $\varphi_{s+1}(h_{s+1}) = \overline{\varphi}_{s+1}(\theta_r(h_r), h_{r+1:s+1})$,

where the numerical function $\overline{\varphi}_{s+1}$ is nonnegative lower semianalytic. Then, by (56),

$$\begin{aligned}
\varphi_s(h_s) &= \big(\mathcal{B}_{s+1:s}\varphi_{s+1}\big)(h_s) \\
&= \inf_{u_s\in\mathbb{U}_s} \int_{\mathbb{W}_{s+1}} \varphi_{s+1}\big((h_s,u_s,w_{s+1})\big)\rho_{s:s+1}(\,\mathrm{d}w_{s+1}\,|\,h_s) \\
&\qquad\qquad \text{(by definition (7a) of the Bellman operator)} \\
&= \inf_{u_s\in\mathbb{U}_s} \int_{\mathbb{W}_{s+1}} \overline{\varphi}_{s+1}\big((\theta_r(h_r),(h_{r+1:s},u_s,w_{s+1}))\big) \\
&\qquad\qquad \rho_{s:s+1}(\,\mathrm{d}w_{s+1}\,|\,h_s) \qquad \text{(by the induction assumption)} \\
&= \inf_{u_s\in\mathbb{U}_s} \int_{\mathbb{W}_{s+1}} \overline{\varphi}_{s+1}\big((\theta_r(h_r),(h_{r+1:s},u_s,w_{s+1}))\big) \\
&\qquad\qquad \widetilde{\rho}_{s:s+1}\big(\,\mathrm{d}w_{s+1}\,\big|\,(\theta_r(h_r),h_{r+1:s})\big) \\
&\qquad\qquad \text{(by compatibility (10) of the stochastic kernel)} \\
&= \overline{\varphi}_s\big(\theta_r(h_r),h_{r+1:s}\big)\ ,
\end{aligned}$$

where

$$\overline{\varphi}_s\big(x_r,h_{r+1:s}\big) = \inf_{u_s\in\mathbb{U}_s} \int_{\mathbb{W}_{s+1}} \overline{\varphi}_{s+1}\big((x_r,(h_{r+1:s},u_s,w_{s+1}))\big) \widetilde{\rho}_{s:s+1}\big(\,\mathrm{d}w_{s+1}\,\big|\,(x_r,h_{r+1:s})\big)\ .$$

By [3, p. 196] (right before Lemma 8.2), we get that the numerical function $\overline{\varphi}_s$ is nonnegative lower semianalytic. Thus, we have shown that the result holds true at time s.

The induction implies that, at time r, the expression of $\varphi_r(h_r)$ is $\varphi_r(h_r) = \overline{\varphi}_r\big(\theta_r(h_r)\big)$, since the term $h_{r+1:r}$ vanishes. Choosing $\tilde{\varphi}_r = \overline{\varphi}_r$ gives the expected result. □

D Technical details and proofs (Sect.6)

In §D.1, we formally define the decision-hazard-decision framework within the history formalism of Sect. 2. Then, in §D.2, we prove a dynamic programming equation. Finally, in §D.3, we express the formalism of state reduction, developed in Sect. 3, for the decision-hazard-decision framework, and then we prove a reduced dynamic programming equation. Again, in this application, the difficulty is mainly notational.

D.1 Multistage stochastic optimization problem with decision-hazard-decision information structure

We flesh out the ingredients necessary to formulate a multistage stochastic optimization problem with decision-hazard-decision information structure.

History. Let $T \in \mathbb{N}^*$. For each time $t \in [\![0, T-1]\!]$, the *head decision* $u_t^\sharp$ takes values in a measurable set $\mathbb{U}_t^\sharp$, equipped with a σ-field $\mathcal{U}_t^\sharp$. For each time $t \in [\![1, T]\!]$, the *tail decision* $u_t^\flat$ takes values in measurable set $\mathbb{U}_t^\flat$, equipped with a σ-field $\mathcal{U}_t^\flat$. For each time $t \in [\![1, T]\!]$, the uncertainty $w_t^\flat$ takes its values in a measurable set $\mathbb{W}_t^\flat$, equipped with a σ-field $\mathcal{W}_t^\flat$. For time $t = 0$, the uncertainty $w_0^\sharp$ takes its values in a measurable set $\mathbb{W}_0^\sharp$, equipped with a σ-field $\mathcal{W}_0^\sharp$.

At the beginning of the time interval $[t, t+1[$, the decision-maker makes a *head decision* $u_t^\sharp$. What is new — in comparison with the classical decision-hazard framework — is that, at the end of the time interval $[t, t+1[$, when an uncertainty variable $w_{t+1}^\flat$ is revealed, the decision-maker has the possibility to make a *tail decision* $u_{t+1}^\flat$. This latter decision $u_{t+1}^\flat$ can be thought as a *recourse* variable for a two stage stochastic optimization problem that would take place inside the time interval $[t, t+1[$. We call $w_0^\sharp$ the uncertainty happening right before the first decision. The interplay between uncertainties and decisions is thus as follows (compare the chronology with the one in (1)):

$$w_0^\sharp \rightsquigarrow u_0^\sharp \rightsquigarrow w_1^\flat \rightsquigarrow u_1^\flat \rightsquigarrow u_1^\sharp \rightsquigarrow w_2^\flat \rightsquigarrow \quad \dots \quad \rightsquigarrow w_{T-1}^\flat \rightsquigarrow u_{T-1}^\flat \rightsquigarrow u_{T-1}^\sharp \rightsquigarrow w_T^\flat \rightsquigarrow u_T^\flat \ .$$

History Spaces. For $t \in [\![0, T]\!]$, we define the *head history space*

$$\mathbb{H}_t^\sharp = \mathbb{W}_0^\sharp \times \prod_{t'=0}^{t-1} \left(\mathbb{U}_{t'}^\sharp \times \mathbb{W}_{t'+1}^\flat \times \mathbb{U}_{t'+1}^\flat \right) , \tag{59a}$$

and its associated *head history field* $\mathcal{H}_t^\sharp$. We also define, for $t \in [\![1, T]\!]$, the *tail history space*

$$\mathbb{H}_t^\flat = \mathbb{H}_{t-1}^\sharp \times \mathbb{U}_{t-1}^\sharp \times \mathbb{W}_t^\flat , \tag{59b}$$

and its associated *tail history field* $\mathcal{H}_t^\flat$.

Stochastic kernels. We introduce a family of stochastic kernels $\{\rho_{t-1:t}\}_{1 \leq t \leq T}$, with

$$\rho_{t-1:t} : \mathbb{H}_{t-1}^\sharp \to \Delta(\mathbb{W}_t^\flat) \ . \tag{60}$$

History feedbacks. For $t \in [\![0, T-1]\!]$, a *head history feedback* at time t is a measurable mapping

$$\gamma_t^\sharp : \mathbb{H}_t^\sharp \to \mathbb{U}_t^\sharp \ .$$

We call $\Gamma_t^\sharp$ the *set of head history feedbacks at time* t, and we define $\Gamma_{t:T}^\sharp = \Gamma_t^\sharp \times \cdots \times \Gamma_T^\sharp$. We also define, for all $t \in [\![1, T]\!]$, a *tail history feedback* at time t as a measurable mapping

$$\gamma_t^\flat : \mathbb{H}_t^\flat \to \mathbb{U}_t^\flat \ .$$

We call $\Gamma_t^\flat$ the *set of tail history feedbacks at time* t, and we define $\Gamma_{t:T}^\flat = \Gamma_t^\flat \times \cdots \times \Gamma_T^\flat$.

Value functions. We consider a nonnegative numerical function

$$j : \mathbb{H}_T^\sharp \to [0, +\infty] \ , \tag{62}$$

assumed to be measurable with respect to the head history field $\mathcal{H}_T^\sharp$.

For $s \in [\![0, T]\!]$, we define *value functions* by

$$V_t(h_t^\sharp) = \inf_{\gamma^\sharp \in \Gamma_{t:T-1}^\sharp, \gamma^\flat \in \Gamma_{t+1:T}^\sharp} \int_{\mathbb{H}_T^\sharp} j(h'_T) \rho_{t:T}^{\gamma^\sharp, \gamma^\flat}(\, \mathrm{d}h'_T \,|\, h_t^\sharp) \ , \ \ \forall h_t^\sharp \in \mathbb{H}_t^\sharp \ , \tag{63}$$

where $\rho_{t:T}^{\gamma^\sharp, \gamma^\flat}$ has to be understood as $\rho_{t:T}^{\gamma}$ (see Definition 15), with

$$\gamma_t(h_t^\sharp) = \gamma_t^\sharp(h_t^\sharp) \ , \ \ \forall h_t^\sharp \in \mathbb{H}_t^\sharp \ , \tag{64a}$$

$$\gamma_{t'}(h_{t'}^\flat) = \Big(\gamma_{t'}^\flat(h_{t'}^\flat), \gamma_{t'}^\sharp\big(h_{t'}^\flat, \gamma_{t'}^\flat(h_{t'}^\flat)\big)\Big) \ , \ \ \forall t' \in [\![t+1, T-1]\!] \ , \ \ \forall h_{t'}^\flat \in \mathbb{H}_{t'}^\flat \ , \tag{64b}$$

$$\gamma_T(h_T^\flat) = \gamma_T^\flat(h_T^\flat) \ , \ \ \forall h_T^\flat \in \mathbb{H}_T^\flat \ . \tag{64c}$$

D.2 Dynamic programming equations in the decision-hazard-decision framework

The following proposition characterizes the dynamic programming equations in the decision-hazard-decision framework.

Proposition 17 *Assume that all the spaces introduced in §D.1 are Borel spaces, the stochastic kernels in* (60) *are Borel-measurable, and that the criterion j in* (62) *is a nonnegative lower semianalytic numerical function.*

For $t \in [\![0, T-1]\!]$, we define the Bellman operator

$$\mathcal{B}_{t+1:t} : \mathbb{L}_+^0(\mathbb{H}_{t+1}^\sharp) \to \mathbb{L}_+^0(\mathbb{H}_t^\sharp) \tag{65a}$$

such that, for all $\varphi \in \mathbb{L}_+^0(\mathbb{H}_{t+1}^\sharp)$ and for all $h_t^\sharp \in \mathbb{H}_t^\sharp$,

$$\big(\mathcal{B}_{t+1:t}\varphi\big)(h_t^\sharp) = \inf_{u_t^\sharp \in \mathbb{U}_t^\sharp} \int_{\mathbb{W}_{t+1}^\flat} \Big(\inf_{u_{t+1}^\flat \in \mathbb{U}_{t+1}^\flat} \varphi(h_t^\sharp, u_t^\sharp, w_{t+1}^\flat, u_{t+1}^\flat) \Big) \rho_{t:t+1}(h_t^\sharp, dw_{t+1}^\flat) \ . \tag{65b}$$

Then the value functions (63) *satisfy*

$$V_T = j \ , \tag{65c}$$

$$V_t = \mathcal{B}_{t+1:t} V_{t+1} \ , \ \ \forall t \in [\![0, T-1]\!] \ . \tag{65d}$$

Proof. We now show that the setting in §D.1 is a particular kind of two-time-scale problem as formalized in Sect. 5. For this purpose, we introduce a *spurious uncertainty variable* $w_t^\sharp$ taking values in a singleton set $\mathbb{W}_t^\sharp = \{\bar{w}_t^\sharp\}$, equipped with the trivial σ-field $\{\emptyset, \mathbb{W}_t^\sharp\}$, for each time $t \in [\![1, T]\!]$. Now, we obtain the following sequence of events

$$w_0^\sharp \rightsquigarrow u_0^\sharp \rightsquigarrow w_1^\flat \rightsquigarrow u_1^\flat \rightsquigarrow w_1^\sharp \rightsquigarrow u_1^\sharp \rightsquigarrow w_2^\flat \rightsquigarrow u_2^\flat \rightsquigarrow w_2^\sharp \rightsquigarrow u_2^\sharp \rightsquigarrow \quad \dots \\ \rightsquigarrow w_{T-1}^\flat \rightsquigarrow u_{T-1}^\flat \rightsquigarrow w_{T-1}^\sharp \rightsquigarrow u_{T-1}^\sharp \rightsquigarrow w_T^\flat \rightsquigarrow u_T^\flat \rightsquigarrow w_T^\sharp \ ,$$

which coincides with a two-time-scale problem:

$$\underbrace{w_{0,0} = w_0^\sharp \rightsquigarrow u_{0,0} = u_0^\sharp \rightsquigarrow w_{0,1} = w_1^\flat \rightsquigarrow u_{0,1} = u_1^\flat}_{\text{slow time cycle}} \rightsquigarrow \\ \underbrace{w_{1,0} = w_1^\sharp \rightsquigarrow u_{1,0} = u_1^\sharp \rightsquigarrow w_{1,1} = w_2^\flat \rightsquigarrow u_{1,1} = u_2^\flat}_{\text{slow time cycle}} \rightsquigarrow \\ \cdots \rightsquigarrow \underbrace{w_{T-1,0} = w_{T-1}^\sharp \rightsquigarrow u_{T-1,0} = u_{T-1}^\sharp \rightsquigarrow w_{T-1,1} = w_T^\flat \rightsquigarrow u_{T-1,1} = u_T^\flat}_{\text{slow time cycle}} \rightsquigarrow w_{T,0} = w_T^\sharp \ .$$

We introduce the sets

$$\begin{aligned} \mathbb{W}_{d,0} &= \mathbb{W}_d^\sharp, \text{ for } \ d \in [\![0, T]\!], \\ \mathbb{W}_{d,1} &= \mathbb{W}_{d+1}^\flat, \text{ for } \ d \in [\![0, T-1]\!], \\ \mathbb{U}_{d,0} &= \mathbb{U}_d^\sharp, \text{ for } \ d \in [\![0, T-1]\!], \\ \mathbb{U}_{d,1} &= \mathbb{U}_{d+1}^\flat, \text{ for } \ d \in [\![0, T-1]\!]. \end{aligned}$$

As a consequence, we observe that the two-time-scale history spaces in §D.1 are in one to one correspondence with the decision-hazard-decision history spaces and fields in (59a)–(59b) as follows: for $d \in [\![0, T]\!]$,

$$\begin{aligned} \mathbb{H}_{d,0} &= \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \left(\mathbb{U}_{d',0} \times \mathbb{W}_{d',1} \times \mathbb{U}_{d',1} \times \mathbb{W}_{d'+1,0} \right) \\ &= \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \left(\mathbb{U}_{d'}^\sharp \times \mathbb{W}_{d'+1}^\flat \times \mathbb{U}_{d'+1}^\flat \times \mathbb{W}_{d'+1}^\sharp \right) \\ &\equiv \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \left(\mathbb{U}_{d'}^\sharp \times \mathbb{W}_{d'+1}^\flat \times \mathbb{U}_{d'+1}^\flat \right) = \mathbb{H}_d^\sharp \ , \end{aligned}$$

for $d \in [\![0, T]\!]$,

$$\mathcal{H}_{d,0} = \mathcal{W}_0^\sharp \otimes \bigotimes_{d'=0}^{d-1} \left(\mathcal{U}_{d'}^\sharp \otimes \mathcal{W}_{d'+1}^\flat \otimes \mathcal{U}_{d'+1}^\flat \otimes \mathcal{W}_{d'+1}^\sharp \right) \ ,$$

for $d \in [\![0, T-1]\!]$,

$$\begin{aligned}\mathbb{H}_{d,1} &= \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \big(\mathbb{U}_{d',0} \times \mathbb{W}_{d',1} \times \mathbb{U}_{d',1} \times \mathbb{W}_{d'+1,0}\big) \times \mathbb{U}_{d,0} \times \mathbb{W}_{d,1} \\ &= \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \big(\mathbb{U}_{d'}^\sharp \times \mathbb{W}_{d'+1}^\flat \times \mathbb{U}_{d'+1}^\flat \times \mathbb{W}_{d'+1}^\sharp\big) \times \mathbb{U}_d^\sharp \times \mathbb{W}_{d+1}^\flat \\ &\equiv \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \big(\mathbb{U}_{d'}^\sharp \times \mathbb{W}_{d'+1}^\flat \times \mathbb{U}_{d'+1}^\flat\big) \times \mathbb{U}_d^\sharp \times \mathbb{W}_{d+1}^\flat = \mathbb{H}_{d+1}^\flat \ ,\end{aligned}$$

for $d \in [\![0, T-1]\!]$,

$$\mathcal{H}_{d,1} = \mathcal{W}_0^\sharp \otimes \bigotimes_{d'=0}^{d-1} \big(\mathcal{U}_{d'}^\sharp \otimes \mathcal{W}_{d'+1}^\flat \otimes \mathcal{U}_{d'+1}^\flat \otimes \mathcal{W}_{d'+1}^\sharp\big) \otimes \mathcal{U}_d^\sharp \otimes \mathcal{W}_{d+1}^\flat \ .$$

For any element h of $\mathbb{H}_{d,0}$ or $\mathbb{H}_{d,1}$ we call $\big[h\big]^\sharp$ the element of $\mathbb{H}_d^\sharp$ or $\mathbb{H}_d^\flat$ corresponding to h with all the spurious uncertainties removed. By a slight abuse of notation, the criterion j in (62) (decision-hazard-decision setting) corresponds to $j \circ \big[\cdot\big]^\sharp$ in the two-time-scale setting in Sect. 5. The feedbacks in the two-time-scale setting in Sect. 5 are in one to one correspondence with the same elements in the decision-hazard-decision setting, namely

$$\gamma_{d,0} = \gamma_d^\sharp \circ \big[\cdot\big]^\sharp \ , \ \ \gamma_{d,1} = \gamma_{d+1}^\flat \circ \big[\cdot\big]^\sharp \ .$$

Now we define two families of stochastic kernels

- a family $\{\rho_{(d,0):(d,1)}\}_{0 \le d \le D}$ of stochastic kernels within two consecutive slow scale indexes

$$\begin{aligned}\rho_{(d,0):(d,1)} : \mathbb{H}_{d,0} &\to \Delta(\mathbb{W}_{d,1}) \ , \\ h_{d,0} &\mapsto \rho_{d:d+1} \circ \big[\cdot\big]^\sharp \ .\end{aligned}$$

- a family $\{\rho_{(d,1):(d+1,0)}\}_{0 \le d \le D-1}$ of stochastic kernels across two consecutive slow scale indexes

$$\begin{aligned}\rho_{(d,1):(d+1,0)} : \mathbb{H}_{d,1} &\to \Delta(\mathbb{W}_{d+1,0}) \ , \\ h_{d,1} &\mapsto \delta_{\bar{w}_{d+1}^\sharp}(\cdot) \ ,\end{aligned}$$

where we recall that $\mathbb{W}_{d+1,0} = \mathbb{W}_{d+1}^\sharp = \{\bar{w}_{d+1}^\sharp\}$.

With these notations, we obtain Equation (65b), where only one integral appears because of the Dirac in the stochastic kernels $\rho_{(d,1):(d+1,0)}$. Indeed, for any measurable function $\varphi : \mathbb{H}_{d+1,0} \to [0, +\infty]$, we have that

$$\begin{aligned}\big(\mathcal{B}_{d+1:d}\varphi\big)(h_{d,0}) = \inf_{u_{d,0} \in \mathbb{U}_{d,0}} \int_{\mathbb{W}_{d,1}} \rho_{(d,0):(d,1)}\Big(\mathrm{d}w_{d,1} \,\Big|\, h_{d,0}\Big) \\ \inf_{u_{d,1} \in \mathbb{U}_{d,1}} \int_{\mathbb{W}_{d+1,0}} \varphi\big(h_{d,0}, u_{d,0}, w_{d,1}, u_{d,1}, w_{d+1,0}\big) \rho_{(d,1):(d+1,0)}\Big(\mathrm{d}w_{d+1,0} \,\Big|\, h_{d,0}, h_{d:d+1}\Big) \ .\end{aligned}$$

Now, if there exists $\tilde{\varphi} : \mathbb{H}_{d+1}^\sharp \to [0, +\infty]$ such that $\varphi = \tilde{\varphi} \circ \big[\cdot\big]^\sharp$, we obtain that

$$\begin{aligned}\big(\mathcal{B}_{d+1:d}\varphi\big)(h_{d,0}) &= \inf_{u_{d,0}\in\mathbb{U}_{d,0}} \int_{\mathbb{W}_{d,1}} \rho_{(d,0):(d,1)}\Big(\mathrm{d}w_{d,1} \,\Big|\, h_{d,0}\Big) \inf_{u_{d,1}\in\mathbb{U}_{d,1}} \tilde{\varphi}\big(\big[h_{d,0}\big]^\sharp, u_{d,0}, w_{d,1}, u_{d,1}\big) \\ &\qquad \int_{\mathbb{W}_{d+1,0}} \rho_{(d,1):(d+1,0)}\Big(\mathrm{d}w_{d+1,0} \,\Big|\, h_{d,0}, h_{d:d+1}\Big) \\ &= \inf_{u_{d,0}\in\mathbb{U}_{d,0}} \int_{\mathbb{W}_{d,1}} \rho_{(d,0):(d,1)}\Big(\mathrm{d}w_{d,1} \,\Big|\, h_{d,0}\Big) \inf_{u_{d,1}\in\mathbb{U}_{d,1}} \tilde{\varphi}\big(\big[h_{d,0}\big]^\sharp, u_{d,0}, w_{d,1}, u_{d,1}\big)\end{aligned}$$

by the Dirac probability of the stochastic kernels $\rho_{(d,1):(d+1,0)}$,

$$= \inf_{u_d^\sharp\in\mathbb{U}_d^\sharp} \int_{\mathbb{W}_{d+1}^\flat} \rho_{(d,0):(d,1)}\Big(\mathrm{d}w_{d+1}^\flat \,\Big|\, h_d^\sharp\Big) \inf_{u_{d+1}^\flat\in\mathbb{U}_{d+1}^\flat} \tilde{\varphi}(h_d^\sharp, u_d^\sharp, w_{d+1}^\flat, u_{d+1}^\flat)$$

This ends the proof. □

D.3 State reduction in the decision-hazard-decision framework

We express the formalism of state reduction, developed in Sect. 3, for the decision-hazard-decision framework, and then we prove a reduced dynamic programming equation.

Compatible state reductions. We now rewrite Definition 5 in the context of a decision-hazard-decision problem.

Definition 18 (Compatible state reduction) *Let $\{(\mathbb{X}_t, \mathcal{X}_t)\}_{t\in[\![0,T]\!]}$ be a family of state spaces, $\{\theta_t\}_{t\in[\![0,T]\!]}$ be a family of measurable reduction mappings such that*

$$\theta_t : \mathbb{H}_t^\sharp \to \mathbb{X}_t \ ,$$

and $\{f_{t:t+1}\}_{t\in[\![0,T-1]\!]}$ be a family of measurable dynamics such that

$$f_{t:t+1} : \mathbb{X}_t \times \mathbb{U}_t^\sharp \times \mathbb{W}_{t+1} \times \mathbb{U}_{t+1}^\flat \to \mathbb{X}_{t+1} \ .$$

The triplet $\big(\{\mathbb{X}_t\}_{t\in[\![0,T]\!]}, \{\theta_t\}_{t\in[\![0,T]\!]}, \{f_{t:t+1}\}_{t\in[\![0,T-1]\!]}\big)$ is said to be a decision-hazard-decision state reduction *if, for all $t \in [\![0, T-1]\!]$, we have that*

$$\theta_{t+1}\big((h_t, u_t^\sharp, w_{t+1}, u_{t+1}^\flat)\big) = f_{t:t+1}\big(\theta_t(h_t), u_t^\sharp, w_{t+1}, u_{t+1}^\flat\big) \ , \qquad \forall (h_t, u_t^\sharp, w_{t+1}, u_{t+1}^\flat) \in \mathbb{H}_t^\sharp \times \mathbb{U}_t^\sharp \times \mathbb{W}_{t+1} \times \mathbb{U}_{t+1}^\flat \ .$$

The decision-hazard-decision state reduction is said to be compatible with the family $\{\rho_{t:t+1}\}_{t\in[\![0,T-1]\!]}$ of stochastic kernels *in* (60) *if there exists a family $\{\widetilde{\rho}_{t:t+1}\}_{t\in[\![0,T-1]\!]}$ of* reduced stochastic kernels

$$\widetilde{\rho}_{t:t+1} : \mathbb{X}_t \to \Delta(\mathbb{W}_{t+1}) \ ,$$

such that, for each $t \in [\![0, T-1]\!]$*, the stochastic kernel* $\rho_{t:t+1}$ *in* (60) *can be factored as*

$$\rho_{t:t+1}(\, \mathrm{d}w_{t+1} \,|\, h_t^\sharp) = \widetilde{\rho}_{t:t+1}\big(\, \mathrm{d}w_{t+1} \,\big|\, \theta_t(h_t^\sharp)\big) \ , \ \ \forall h_t^\sharp \in \mathbb{H}_t^\sharp \ .$$

As in Definition 5, we call decision-hazard-decision compatible reduction data set *the family*

$$\mathcal{R}_{dhd} = \big(\{(\mathbb{X}_t, \mathcal{X}_t)\}_{t\in[\![0,T]\!]}, \{\theta_t\}_{t\in[\![0,T]\!]}, \{f_{t:t+1}\}_{t\in[\![0,T-1]\!]}, \{\widetilde{\rho}_{t:t+1}\}_{t\in[\![0,T-1]\!]}\big) \tag{71}$$

made of the elements presented above. We say that the decision-hazard-decision compatible reduction data set $\mathcal{R}_{dhd}$ *is* Borel *if the state spaces* $\{(\mathbb{X}_t, \mathcal{X}_t)\}_{t\in[\![0,T]\!]}$ *are Borel spaces, the reduction mappings* $\{\theta_t\}_{t\in[\![0,T]\!]}$ *are Borel-measurable, the dynamics* $\{f_{t:t+1}\}_{t\in[\![0,T-1]\!]}$ *are Borel-measurable, the reduced stochastic kernels* $\{\widetilde{\rho}_{t:t+1}\}_{t\in[\![0,T-1]\!]}$ *are Borel-measurable.*

Dynamic programming equations. We state the main result of this section.

Theorem 19 *Suppose that all the assumptions of Proposition 17 hold true. Suppose that there exists a compatible decision-hazard-decision state reduction with Borel decision-hazard-decision compatible reduction data set* $\mathcal{R}_{dhd}$ *as in* (71)*, and that there exists a* reduced criterion*, which is a nonnegative lower semianalytic numerical function,*

$$\tilde{j} : \mathbb{X}_T \to [0, +\infty] \ ,$$

such that the cost function j *in* (62) *can be factored as*

$$j = \tilde{j} \circ \theta_T \ .$$

We can define a family of reduced Bellman operators across $[\![t+1, t]\!]$

$$\widetilde{\mathcal{B}}_{t+1:t} : \mathbb{L}^0_+(\mathbb{X}_{t+1}) \to \mathbb{L}^0_+(\mathbb{X}_t, \mathcal{X}_t) \ , \ \ t \in [\![1, T-1]\!] \ , \tag{73a}$$

by, for any measurable function $\tilde{\varphi} : \mathbb{X}_{t+1} \to [0, +\infty]$*,*

$$(\widetilde{\mathcal{B}}_{t+1:t}\tilde{\varphi})(x_t) = \inf_{u_t^\sharp \in \mathbb{U}_t^\sharp} \int_{\mathbb{W}_{t+1}} \Big(\inf_{u_{t+1}^\flat \in \mathbb{U}_{t+1}^\flat} \tilde{\varphi}\big(f_{t:t+1}(x_t, u_t^\sharp, w_{t+1}, u_{t+1}^\flat)\big)\Big) \widetilde{\rho}_{t:t+1}(x_t, dw_{t+1}) \ . \tag{73b}$$

We define the family of reduced value functions $\{\tilde{V}_t\}_{t\in[\![0,T]\!]}$ *by*

$$\tilde{V}_T = \tilde{j} \tag{74a}$$

$$\tilde{V}_t = \widetilde{\mathcal{B}}_{t+1:t}\tilde{V}_{t+1} \quad \textit{for } t = T-1, \ldots, 0 \ . \tag{74b}$$

Then, the value functions V_t *defined by* (63) *satisfy*

$$V_t = \tilde{V}_t \circ \theta_t \ , \ \ t \in [\![0, T]\!] \ . \tag{75}$$

Proof. It has been shown in the proof of Proposition 17 that the setting of a decision-hazard-decision problem was a particular kind of two time-scales problem. The proof of the theorem is then a direct application of Proposition 13. □

References

References

[1] Kengy Barty, Pierre Carpentier, and Pierre Girardeau. Decomposition of large-scale stochastic optimal control problems. *RAIRO Operations Research*, 44(3):167–183, 2010.

[2] R. E. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, N.J., 1957.

[3] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific, Belmont, Massachusetts, 1996.

[4] John R. Birge. Aggregation bounds in stochastic linear programming. *Mathematical Programming*, 31:25–41, 1985.

[5] Pierre Carpentier, Jean-Philippe Chancelier, Michel De Lara, and François Pacaud. Mixed spatial and temporal decompositions for large-scale multistage stochastic optimization problems. *Journal of Optimization Theory and Applications*, 186(3):985–1005, 2020.

[6] Bolong Cheng and Warren B. Powell. Co-optimizing battery storage for the frequency regulation and energy arbitrage using multi-scale dynamic programming. *IEEE Transactions on Smart Grid*, 9(3):1997–2005, 2016.

[7] I. V. Evstigneev. Measurable selection and dynamic programming. *Mathematics of Operations Research*, 1(3):267–272, 1976.

[8] Martin Glanzer and Georg Ch. Pflug. Multiscale stochastic optimization: modeling aspects and scenario generation. *Computational Optimization and Applications*, 75(1):1–34, 2020.

[9] Onésimo Hernández-Lerma and Jean Bernard Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Applications of Mathematics Stochastic Modelling and Applied Probability. Springer, New York, 1996.

[10] Michal Kaut, Kjetil T. Midthun, Adrian S. Werner, Asgeir Tomasgard, Lars Hellemo, and Marte Fodstad. Multi-horizon stochastic programming. *Computational Management Science*, 11(1):179–193, 2014.

[11] Daniel Kuhn. Aggregation and discretization in multistage stochastic programming. *Mathematical Programming*, 113:61–94, 2008.

[12] M. Loève. *Probability Theory I*. Springer-Verlag, New York, fourth edition, 1977.

[13] Thomas Martin. *Stochastic Optimization for the Procurement of Crude Oil in Refineries*. PhD thesis, École des Ponts ParisTech, 2021.

[14] Rodrigo Porteiro, Andres Ferragut, and Fernando Paganini. Towards multi-timescale energy provisioning using stochastic dual dynamic programming. In *2018 IEEE 9th Power, Instrumentation and Measurement Meeting (EPIM)*, pages 1–6, 2018.

[15] M. L. Puterman. *Markov Decision Processes*. Wiley, New York, 1994.

[16] Tristan Rigaut. *Time Decomposition Methods For Optimal Management Of Energy Storage Under Stochasticity*. PhD thesis, Université Paris-Est, 2019.

[17] R.T. Rockafellar and R. J-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of operations research*, 16(1):119–147, 1991.

[18] Andrzej Ruszczyński. Decomposition methods. In *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 141–211. Elsevier, 2003.

[19] A. Shapiro, D. Dentcheva, and A. Ruszczynski. *Lectures on stochastic programming: modeling and theory*. The society for industrial and applied mathematics and the mathematical programming society, Philadelphia, USA, 2009.

[20] Alexander Shapiro and Lingquan Ding. Periodical multistage stochastic programs. *SIAM Journal of Optimization*, 30(3):2083–2102, 2020.

[21] Jayakumar Subramanian and Aditya Mahajan. Approximate information state for partially observed systems. *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 1629–1636, 2019.

[22] P. Whittle. *Optimization over Time: Dynamic Programming and Stochastic Control*, volume 1. John Wiley & Sons, New York, 1982.

[23] H. S. Witsenhausen. A standard form for sequential stochastic control. *Mathematical Systems Theory*, 7(1):5–11, 1973.

[24] H. S. Witsenhausen. The intrinsic model for discrete stochastic control: Some open problems. In A. Bensoussan and J. L. Lions, editors, *Control Theory, Numerical Methods and Computer Systems Modelling*, volume 107 of *Lecture Notes in Economics and Mathematical Systems*, pages 322–335. Springer-Verlag, 1975.

[25] H. S. Witsenhausen. On policy independence of conditional expectations. *Information and Control*, 28(1):65–75, 1975.

[26] H. S. Witsenhausen. Equivalent stochastic control problems. *Mathematics of Control, Signals, and Systems*, 1(1):3–7, 1988.

[27] S. E. Wright. Primal-dual aggregation and disaggregation for stochastic linear programs. *Mathematics of Operations Research*, 19(4):893–908, 1994.

[28] Serdar Yüksel. A universal dynamic program and refined existence results for decentralized stochastic control. *SIAM Journal on Control and Optimization*, 58(5):2711–2739, 2020.