Time Blocks Decomposition of Multistage Stochastic Optimization Problems

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

Noname manuscript No.
(will be inserted by the editor)

Time Blocks Decomposition of Multistage Stochastic Optimization Problems

P. Carpentier · J.-P. Chancelier · M. De Lara · T. Rigaut

the date of receipt and acceptance should be inserted later

Abstract Multistage stochastic optimization problems are, by essence, complex because their solutions are indexed both by stages (time) and by uncertainties (scenarios). Their large scale nature makes decomposition methods appealing. 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. We present a method to decompose multistage stochastic optimization problems by time blocks, which covers both stochastic programming and stochastic dynamic programming. Once established a dynamic programming equation with value functions defined on the history space (a history is a sequence of uncertainties and controls), we provide conditions to reduce the history using a compressed "state" variable. This reduction is done by time blocks, that is, at stages that are not necessarily all the original unit stages, and we prove a reduced dynamic programming equation. Then, we apply the reduction method by time blocks to *two time-scales* stochastic optimization problems and to a novel class of so-called *decision-hazard-decision* problems, arising in many practical situations, like in stock management. The *time blocks decomposition* scheme is as follows: we use dynamic programming at slow time scale where the slow time scale noises are supposed to be stagewise independent, and we produce slow time scale Bellman functions; then, we use stochastic programming at short time scale, within two consecutive slow time steps, with the final short time scale cost given by the slow time scale Bellman functions, and without assuming stagewise independence for the short time scale noises.

Keywords: multistage stochastic optimization, dynamic programming, decomposition, time blocks, two time-scales, decision-hazard-decision.

MSC: 90C06,90C39,93E20.

1 Introduction

Multistage stochastic optimization problems are, by essence, complex because their solutions are indexed both by stages (time) and by uncertainties. Their large scale nature makes decomposition methods appealing. 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.

On the one hand, stochastic programming deals with an underlying random process taking a finite number of values, called scenarios [10]. Solutions are indexed by a scenario tree, the size of which explodes with the number of stages, hence generally few in practice. However, to overcome this obstacle, stochastic programming takes advantage of scenario decomposition methods (progressive hedging [9]). On the other hand, stochastic control deals with a state model driven by a white noise, that is, the noise is made of

UMA, ENSTA ParisTech, Université Paris-Saclay E-mail: pierre.carpentier@ensta-paristech.fr · Université Paris-Est, CERMICS (ENPC) E-mail: chancelier@cermics.enpc.fr · Université Paris-Est, CERMICS (ENPC) E-mail: delara@cermics.enpc.fr · Efficacity E-mail: t.rigaut@efficacity.com

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) [2,8].

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 noise would be statistically independent — crossing over short time scale optimization problems where independence would not hold? This question is one of the motivations of this paper.

We will provide a method to decompose multistage stochastic optimization problems by time blocks. In Sect. 2, we present a mathematical framework that covers both stochastic programming and stochastic dynamic programming. First, in §2.1, we sketch the literature in stochastic dynamic programming, in order to locate our contribution. Second, in §2.2, we formulate multistage stochastic optimization problems over a so-called history space, and we obtain a general dynamic programming equation. Then, we lay out the basic brick of time blocks decomposition, by revisiting the notion of "state" in Sect. 3. We lay out conditions under which we can reduce the history using a compressed "state" variable, but with a reduction done by time blocks, that is, at stages that are not necessarily all the original unit stages. We prove a reduced dynamic programming equation, and apply it to two classes of problems in Sect. 4. In §4.1, we detail the case of two time-scales stochastic optimization problems. In §4.2, we apply the reduction method by time blocks to a novel class consisting of decision-hazard-decision models. In the appendix, we relegate technical results, as well as the specific case of optimization with noise process.

2 Stochastic Dynamic Programming with Histories

We recall the standard approaches used to deal with a stochastic optimal control problem formulated in discrete time, and we highlight the differences with the framework used in this paper.

2.1 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 mapping $\rho : \mathbb{X} \times \mathcal{Y} \to [0, 1]$ such that

- for any $Y \in \mathcal{Y}$, $\rho(\cdot, Y)$ is $\mathcal{X}$-measurable;
- for any $x \in \mathbb{X}$, $\rho(x, \cdot)$ is a probability measure on $\mathcal{Y}$.

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 $x \in \mathbb{X} \mapsto \int_Y \rho(x, \mathrm{d}y)$ is measurable for any $Y \in \mathcal{Y}$.

We now sketch the most classical frameworks for stochastic dynamic programming.

Witsenhausen Approach. The most general stochastic dynamic programming principle is sketched by Witsenhausen in [12]. However, we do not detail it as its formalism is too far from the following ones. We present here what Witsenhausen calls an optimal stochastic control problem in *standard form* (see [11]). The ingredients are the following:

1. time $t = t_0, t_0 + 1, \ldots, T-1, T$ is discrete, with integers $t_0 < T$;
2. $(\mathbb{X}_{t_0}, \mathcal{X}_{t_0})$, $\ldots$, $(\mathbb{X}_T, \mathcal{X}_T)$ are measurable spaces ("state" spaces);
3. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0})$,$\ldots$, $(\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (decision spaces);
4. $\mathcal{I}_t$ is a subfield of $\mathcal{X}_t$, for $t = t_0, \ldots, T-1$ (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 $t = t_0, \ldots, T-1$ (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 $t = t_0, \ldots, T-1$. 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 [6] are the following:

1. time $t = t_0, t_0+1, \ldots, T-1$ is discrete, with integers $t_0 < T$;
2. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0}), \ldots, (\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (decision spaces);
3. $(\Omega, \mathcal{F})$ is a measurable space (Nature);
4. $\{\mathcal{F}_t\}_{t_0,\ldots,T-1}$ is a filtration of $\mathcal{F}$ (information);
5. $\mathbb{P}$ is a probability on $(\Omega, \mathcal{F})$;
6. $j : (\prod_{t=t_0,\ldots,T-1} \mathbb{U}_t \times \Omega, \bigotimes_{t=t_0,\ldots,T-1} \mathcal{U}_t \otimes \mathcal{F}) \to \mathbb{R}$ 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 $\prod_{t=t_0,\ldots,T-1} \mathbb{U}_t$ and adapted to the filtration $\{\mathcal{F}_t\}_{t_0,\ldots,T-1}$. Then, he establishes a dynamic programming equation, where the Bellman function at time t is an $\mathcal{F}_t$-integrand depending on decisions 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 function j, that we do not detail here).

Bertsekas and Shreve Approach. The ingredients of the approach developed in [3] are the following:

1. time $t = t_0, t_0+1, \ldots, T-1, T$ is discrete, with integers $t_0 < T$;
2. $(\mathbb{X}_{t_0}, \mathcal{X}_{t_0}), \ldots, (\mathbb{X}_T, \mathcal{X}_T)$ are measurable spaces (state spaces);
3. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0}), \ldots, (\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (decision spaces);
4. $(\mathbb{W}_{t_0}, \mathcal{W}_{t_0}), \ldots, (\mathbb{W}_T, \mathcal{W}_T)$ are measurable spaces (Nature);
5. $f_t : (\mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_t, \mathcal{X}_t \otimes \mathcal{U}_t \otimes \mathcal{W}_t) \to (\mathbb{X}_{t+1}, \mathcal{X}_{t+1})$ is a measurable mapping, for $t = t_0, \ldots, T-1$ (dynamics);
6. $\rho_{t-1:t} : \mathbb{X}_{t-1} \times \mathbb{U}_{t-1} \to \Delta(\mathbb{W}_t)$ is a stochastic kernel, for $t = t_0, \ldots, T-1$;
7. $L_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \to \mathbb{R}$, for $t = t_0, \ldots, T-1$ and $K : \mathbb{X}_T \to \mathbb{R}$, measurable functions (instantaneous and final costs).

With these ingredients, Bertsekas and Shreve formulate a stochastic optimization problem with time additive additive cost function over given state spaces, action spaces and uncertainty spaces (note that state and action spaces are assumed to be of fixed sizes when time varies, thus a "state" is a priori given). They introduce the notion of history at time t which consists in the states and the actions 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 $\mathbb{X}_{t_0} \times \prod_{s=t_0}^{t-1} (\mathbb{U}_s \times \mathbb{X}_{s+1}) \to \mathbb{U}_t$. 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].

Puterman Approach. The ingredients of the approach developed in [8] are the following:

1. time $t = t_0, t_0+1, \ldots, T-1, T$ is discrete, with integers $t_0 < T$;
2. $(\mathbb{X}_{t_0}, \mathcal{X}_{t_0}), \ldots, (\mathbb{X}_T, \mathcal{X}_T)$ are measurable spaces (state spaces);
3. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0}), \ldots, (\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (decision spaces);
4. $\rho_{t-1:t} : \mathbb{X}_{t-1} \times \mathbb{U}_{t-1} \to \Delta(\mathbb{X}_t)$ is a stochastic kernel, for $t = t_0, \ldots, T-1$;
5. $L_t : \mathbb{X}_t \times \mathbb{U}_t \to \mathbb{R}$, for $t = t_0, \ldots, T-1$ and $K : \mathbb{X}_T \to \mathbb{R}$, measurable functions (instantaneous and final costs).

Puterman shares most of his ingredients with Bertsekas and Shreve, but he does not require uncertainty sets and dynamics, as he directly considers state transition stochastic kernels. With these ingredients, Puterman formulates a stochastic optimization problem, 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.

Approach in this Paper. The ingredients that we will use are the following:

1. time $t = t_0, t_0 + 1, \ldots, T-1, T$ is discrete, with integers $t_0 < T$;
2. $(\mathbb{U}_{t_0}, \mathcal{U}_{t_0})$,..., $(\mathbb{U}_{T-1}, \mathcal{U}_{T-1})$ are measurable spaces (decision spaces);
3. $(\mathbb{W}_{t_0}, \mathcal{W}_{t_0})$,..., $(\mathbb{W}_T, \mathcal{W}_T)$ are measurable spaces (Nature);
4. $\rho_{t-1:t} : \mathbb{W}_0 \times \prod_{s=0}^{t-1}(\mathbb{U}_s \times \mathbb{W}_{s+1}) \to \Delta(\mathbb{W}_t)$ is a stochastic kernel, for $t = t_0, \ldots, T-1$,
5. $j : (\mathbb{W}_0 \times \prod_{s=0}^{T-1}(\mathbb{U}_s \times \mathbb{W}_{s+1}), \mathcal{W}_0 \otimes \bigotimes_{s=0}^{T-1}(\mathcal{U}_s \otimes \mathcal{W}_{s+1})) \to \mathbb{R}$ is a measurable function (criterion).

The main features of the framework developed in this paper are the following: the history at time t consists of all uncertainties and actions prior to time t (rather than states and actions); the cost is a unique function depending on the whole history, from initial time t_0 to the horizon T; the probability distribution of uncertainty at time t depends on the history up to time $t-1$. We will state a dynamic programming equation, where the Bellman functions are function of the history $h_t \in \mathbb{W}_0 \times \prod_{s=0}^{t}(\mathbb{U}_s \times \mathbb{W}_{s+1})$ and where the minimization is done over controls $u_t \in \mathbb{U}_t$.

Our main objective is to establish a dynamic programming equation with a state, not at any time $t \in \{0, \ldots, T\}$, but at some specified instants $0 = t_0 < t_1 < \cdots < t_N = T$. The state spaces are not given a priori, but introduced a posteriori as image sets of history reduction mappings. With this, we can mix dynamic programming and stochastic programming.

Our framework is rather distant with the one of Evstigneev in [6]. It falls in the general framework developed by Witsenhausen (see [11] and [4, § 4.5.4]), *except* for the stochastic kernels (we are more general) and for the information structure (we are less general). Finally, our framework is closest to the one found in Bertsekas and Shreve [3] and Puterman [8], *except* for the state spaces, not given a priori, and for the criterion, function of the whole history.

2.2 Stochastic Dynamic Programming with History Feedbacks

We now present a framework that is adapted to both stochastic programming and stochastic dynamic programming. Time is discrete and runs among the integers $t = 0, 1, 2 \ldots, T-1, T$, where $T \in \mathbb{N}^*$. For $0 \leqslant r \leqslant s \leqslant T$, we introduce the interval $(r:s) = \{t \in \mathbb{N} \mid r \leqslant t \leqslant s\}$.

2.2.1 Histories and Feedbacks

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 $t = 0, 1, 2 \ldots, T-1$, the decision u_t takes its values in a measurable set $\mathbb{U}_t$ equipped with a σ-field $\mathcal{U}_t$. For each time $t = 0, 1, 2 \ldots, T$, the uncertainty w_t takes its values in a measurable set $\mathbb{W}_t$ equipped with a σ-field $\mathcal{W}_t$.

For $t = 0, 1, 2 \ldots, T$, we define the *history space* $\mathbb{H}_t$ equipped with the *history field* $\mathcal{H}_t$ by

$$\mathbb{H}_t = \mathbb{W}_0 \times \prod_{s=0}^{t-1}(\mathbb{U}_s \times \mathbb{W}_{s+1}) \text{ and } \mathcal{H}_t = \mathcal{W}_0 \otimes \bigotimes_{s=0}^{t-1}(\mathcal{U}_s \otimes \mathcal{W}_{s+1}) , \quad t = 0, 1, 2 \ldots, T , \tag{1}$$

with the particular case $\mathbb{H}_0 = \mathbb{W}_0$, $\mathcal{H}_0 = \mathcal{W}_0$. A generic element $h_t \in \mathbb{H}_t$ is called a *history*:

$$h_t = (w_0, (u_s, w_{s+1})_{s=0,\ldots,t-1}) = (w_0, u_0, w_1, u_1, w_2, \ldots, u_{t-2}, w_{t-1}, u_{t-1}, w_t) \in \mathbb{H}_t .$$

For $1 \leqslant r \leqslant s \leqslant t$, we introduce the $(r:s)$-*history subpart*

$$h_{r:s} = (u_{r-1}, w_r, \dots, u_{s-1}, w_s) ,$$

so that we have $h_t = (h_{r-1}, h_{r:t})$.

History Feedbacks. When $0 \leqslant r \leqslant t \leqslant T-1$, we define a $(r:t)$-*history feedback* as a sequence $\{\gamma_s\}_{s=r,\dots,t}$ 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 produced. Then, at the beginning of the time interval $[t, t+1[$, a decision-maker takes a decision u_t, as follows

$$w_0 \rightsquigarrow u_0 \rightsquigarrow w_1 \rightsquigarrow u_1 \rightsquigarrow \quad \dots \quad \rightsquigarrow w_{T-1} \rightsquigarrow u_{T-1} \rightsquigarrow w_T . \tag{2}$$

2.2.2 *Optimization with Stochastic Kernels*

We introduce a family of optimization problems with stochastic kernels. Then, we show how such problems can be solved by stochastic dynamic programming.

In what follows, we say that a function is *numerical* if it takes its values in $[-\infty, +\infty]$ (also called *extended* or *extended real-valued* function).

Family of Optimization Problems with Stochastic Kernels. To build a family of optimization problems over the time span $\{0, \dots, T-1\}$, we require two ingredients:

- a family $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant T}$ of stochastic kernels
$$\rho_{s-1:s} : (\mathbb{H}_{s-1}, \mathcal{H}_{s-1}) \to \Delta(\mathbb{W}_s) , \quad s = 1, \dots, T , \tag{3}$$
that represents the distribution of the next uncertainty w_s parameterized by past history h_{s-1} (see the chronology in (2)),
- a numerical function, playing the role of a cost to be minimized,
$$j : (\mathbb{H}_T, \mathcal{H}_T) \to [0, +\infty] , \tag{4}$$
assumed to be nonnegative[1] and measurable with respect to the field $\mathcal{H}_T$.

We define, for any $\{\gamma_s\}_{s=t,\dots,T-1} \in \Gamma_{t:T-1}$, 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 $s = t, \dots, T-1$ (see Definition 5 in §A.2 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 horizon $\{0, \dots, T\}$).

We consider the family of optimization problems, indexed by $t = 0, \dots, T-1$ and parameterized by the history $h_t \in \mathbb{H}_t$:

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

the integral in the right-hand side of the above equation corresponding to the cost induced by the feedback $\gamma_{t:T-1}$ when starting at time t with a given history h_t. For all $t = 0, \dots, T-1$, we define the minimum value of Problem (5) by

$$V_t(h_t) = \inf_{\gamma_{t:T-1} \in \Gamma_{t:T-1}} \int_{\mathbb{H}_T} j(h'_T) \rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T) , \quad \forall h_t \in \mathbb{H}_t , \tag{6a}$$

and we also define

$$V_T(h_T) = j(h_T) , \quad \forall h_T \in \mathbb{H}_T . \tag{6b}$$

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

[1] We could also consider any $j : \mathbb{H}_t \to \mathbb{R}$, 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. When $j(h_T) = +\infty$, this materializes joint constraints between uncertainties and controls.

Bellman Operators and Dynamic Programming. We show that the value functions in (6) are *Bellman functions*, in that they are solution of the Bellman or dynamic programming equation.

For $t = 0, \dots, T$, let $\mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t)$ be the space of universally measurable nonnegative numerical functions over $\mathbb{H}_t$ (see [3] for further details). For $t = 0, \dots, T-1$, we define the *Bellman operator* by, for all $\varphi \in \mathbb{L}^0_+(\mathbb{H}_{t+1}, \mathcal{H}_{t+1})$ and for all $h_t \in \mathbb{H}_t$,

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

Since $\varphi \in \mathbb{L}^0_+(\mathbb{H}_{t+1}, \mathcal{H}_{t+1})$, we have that $\mathcal{B}_{t+1:t}\varphi$ is a well defined nonnegative numerical function.

The proof of the following theorem is inspired by [3], and given in §A.3.1.

Theorem 1 *Assume that all the spaces introduced in §2.2.1 are Borel spaces, the stochastic kernels in* (3) *are Borel-measurable, and that the criterion j in* (4) *is a nonnegative lower semianalytic function.*

Then, the Bellman operators in (7) *map $\mathbb{L}^0_+(\mathbb{H}_{t+1}, \mathcal{H}_{t+1})$ into $\mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t)$*

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

and the value functions V_t defined in (6) *are universally measurable and satisfy the* Bellman equation*, or* (stochastic) dynamic programming equation*,*

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

This theorem is mainly inspired by [3], with the feature that the state x_t is in our case the history h_t, with the dynamics:

$$h_{t+1} = \big(h_t, u_t, w_{t+1}\big) \; . \tag{9}$$

This very general dynamic programming result will be the basis of all future developments in this paper. In the sequel, we assume that all the assumptions of Theorem 1 are fulfilled, that is,

- all the spaces (like the ones introduced in §2.2.1) will be supposed to be Borel spaces,
- all the stochastic kernels (like the ones introduced in (3)) will be supposed to be Borel-measurable,
- all the criteria (like the one introduced in (4)) will be supposed to be nonnegative lower semianalytic functions.

3 State Reduction by Time Blocks and Dynamic Programming

In this section, we consider the question of reducing the history using a compressed "state" variable. Differing with traditional practice, such a variable may be not available at any time $t \in \{0, \dots, T\}$, but at some specified instants $0 = t_0 < t_1 < \cdots < t_N = T$. We have see in the previous section that the history h_t is itself a canonical state variable in our framework with associated dynamics (9). However the size of this canonical state increases with t, which is a nasty feature for dynamic programming.

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:

$$0 \leqslant r < t \leqslant T \; .$$

Definition 1 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*

$$\theta_r : \mathbb{H}_r \to \mathbb{X}_r \; , \;\; \theta_t : \mathbb{H}_t \to \mathbb{X}_t \; , \tag{10a}$$

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

$$f_{r:t} : \mathbb{X}_r \times \mathbb{H}_{r+1:t} \to \mathbb{X}_t \; . \tag{10b}$$

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

$$\theta_t\big((h_r, h_{r+1:t})\big) = f_{r:t}\big(\theta_r(h_r), h_{r+1:t}\big) \ , \quad \forall h_t \in \mathbb{H}_t \ . \tag{10c}$$

The state reduction $(\theta_r, \theta_t, f_{r:t})$ is said to be *compatible* with the family $\{\rho_{s-1:s}\}_{r+1 \leqslant s \leqslant t}$ of stochastic kernels (3) if

- there exists a *reduced stochastic kernel*

$$\widetilde{\rho}_{r:r+1} : \mathbb{X}_r \to \Delta(\mathbb{W}_{r+1}) \ , \tag{11a}$$

  such that the stochastic kernel $\rho_{r:r+1}$ in (3) can be factored as

$$\rho_{r:r+1}(h_r, \mathrm{d}w_{r+1}) = \widetilde{\rho}_{r:r+1}\big(\theta_r(h_r), \mathrm{d}w_{r+1}\big) \ , \quad \forall h_r \in \mathbb{H}_r \ , \tag{11b}$$

- for all $s = r+2, \ldots, t$, there exists a *reduced stochastic kernel*

$$\widetilde{\rho}_{s-1:s} : \mathbb{X}_r \times \mathbb{H}_{r+1:s-1} \to \Delta(\mathbb{W}_s) \ , \tag{11c}$$

  such that the stochastic kernel $\rho_{s-1:s}$ can be factored as

$$\rho_{s-1:s}\big((h_r, h_{r+1:s-1}), \mathrm{d}w_s\big) = \widetilde{\rho}_{s-1:s}\Big(\big(\theta_r(h_r), h_{r+1:s-1}\big), \mathrm{d}w_s\Big) \ , \quad \forall h_{s-1} \in \mathbb{H}_{s-1} \ . \tag{11d}$$

According to this definition, the triplet $(\theta_r, \theta_t, f_{r:t})$ is a state reduction across $(r : t)$ if and only if the diagram in Figure 1 is commutative; it is compatible if and only if the diagram in Figure 2 is commutative.

$$\begin{tikzcd}
\mathbb{H}_r \arrow[d, "\theta_r"] & \times & \mathbb{H}_{r+1:t} \arrow[d, "I_d"] \arrow[r, "I_d"] & \mathbb{H}_t \arrow[d, "\theta_t"] \\
\mathbb{X}_r & \times & \mathbb{H}_{r+1:t} \arrow[r, "f_{r:t}"] & \mathbb{X}_t
\end{tikzcd}$$

Fig. 1 Commutative diagram in case of state reduction $(\theta_r, \theta_t, f_{r:t})$

$$\begin{tikzcd}
\mathbb{H}_r \arrow[d, "\theta_r"] & \times & \mathbb{H}_{r+1:s-1} \arrow[d, "I_d"] \arrow[r, "\rho_{s-1:s}"] & \Delta(\mathbb{W}_s) \\
\mathbb{X}_r & \times & \mathbb{H}_{r+1:s-1} \arrow[ur, "\widetilde{\rho}_{s-1:s}"'] &
\end{tikzcd}$$

Fig. 2 Commutative diagram in case of state reduction $(\theta_r, \theta_t, f_{r:t})$ compatible with the family $\{\rho_{s-1:s}\}_{r+1 \leqslant s \leqslant t}$

We define the *Bellman operator across* $(t:r)$ $\mathcal{B}_{t:r} : \mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t) \to \mathbb{L}^0_+(\mathbb{H}_r, \mathcal{H}_r)$ by

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

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

The following proposition, whose proof is given in §A.3.2, is the key ingredient to formulate dynamic programming equations with a reduced state.

Proposition 1 *Suppose that there exists a state reduction* $(\theta_r, \theta_t, f_{r:t})$ *that is compatible with the family* $\{\rho_{s-1:s}\}_{r+1 \leqslant s \leqslant t}$ *of stochastic kernels* (3) *(see Definition 1). Then, there exists a* reduced Bellman operator across $(t\!:\!r)$

$$\widetilde{\mathcal{B}}_{t:r} : \mathbb{L}^0_+(\mathbb{X}_t, \mathcal{X}_t) \to \mathbb{L}^0_+(\mathbb{X}_r, \mathcal{X}_r) \;, \tag{13}$$

such that, for all $\widetilde{\varphi}_t \in \mathbb{L}^0_+(\mathbb{X}_t, \mathcal{X}_t)$*, we have that*

$$\big(\widetilde{\mathcal{B}}_{t:r}\widetilde{\varphi}_t\big) \circ \theta_r = \mathcal{B}_{t:r}(\widetilde{\varphi}_t \circ \theta_t) \;. \tag{14}$$

For all measurable nonnegative numerical function $\widetilde{\varphi}_t : \mathbb{X}_t \to [0, +\infty]$ *and for all* $x_r \in \mathbb{X}_r$*, we have that*

$$\begin{aligned}\big(\widetilde{\mathcal{B}}_{t:r}\widetilde{\varphi}_t\big)(x_r) = \inf_{u_r \in \mathbb{U}_r} \int_{\mathbb{W}_{r+1}} & \widetilde{\rho}_{r:r+1}(x_r, \mathrm{d}w_{r+1}) \\ & \inf_{u_{r+1} \in \mathbb{U}_{r+1}} \int_{\mathbb{W}_{r+2}} \widetilde{\rho}_{r+1:r+2}(x_r, u_r, w_{r+1}, \mathrm{d}w_{r+2}) \ldots \\ & \qquad \inf_{u_{t-1} \in \mathbb{U}_{t-1}} \int_{\mathbb{W}_t} \widetilde{\varphi}_t\big(f_{r:t}(x_r, u_r, w_{r+1}, \ldots, u_{t-1}, w_t)\big) \\ & \qquad\qquad \widetilde{\rho}_{t-1:t}(x_r, u_r, w_{r+1}, \ldots, u_{t-2}, w_{t-1}, \mathrm{d}w_t) \;. \end{aligned} \tag{15}$$

Proposition 1 can be interpreted as follows. Denoting by $\theta_t^\star : \mathbb{L}^0_+(\mathbb{X}_t, \mathcal{X}_t) \to \mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t)$ the operator defined by

$$\theta_t^\star(\widetilde{\varphi}_t) = \widetilde{\varphi}_t \circ \theta_t \;, \quad \forall \widetilde{\varphi}_t \in \mathbb{L}^0_+(\mathbb{X}_t, \mathcal{X}_t) \;,$$

the relation (14) rewrites

$$\theta_r^\star \circ \widetilde{\mathcal{B}}_{t:r} = \mathcal{B}_{t:r} \circ \theta_t^\star \;,$$

that is, Proposition 1 states that the diagram in Figure 3 is commutative.

$$\begin{tikzcd} \mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t) \arrow[r, "\mathcal{B}_{t:r}"] & \mathbb{L}^0_+(\mathbb{H}_r, \mathcal{H}_r) \\ \mathbb{L}^0_+(\mathbb{X}_t, \mathcal{X}_t) \arrow[u, "\theta_t^\star"] \arrow[r, "\widetilde{\mathcal{B}}_{t:r}"] & \mathbb{L}^0_+(\mathbb{X}_r, \mathcal{X}_r) \arrow[u, "\theta_r^\star"'] \end{tikzcd}$$

Fig. 3 Commutative diagram for Bellman operators in case of a compatible state reduction $(\theta_r, \theta_t, f_{r:t})$

3.2 State Reduction on Multiple Consecutive Time Blocks and Dynamic Programming Equations

Proposition 1 can easily be extended to the case of multiple consecutive time blocks $[t_i, t_{i+1}]$, $i = 0, \ldots, N-1$, where

$$0 = t_0 < t_1 < \cdots < t_N = T \;. \tag{16}$$

Definition 2 Let $\{(\mathbb{X}_{t_i}, \mathcal{X}_{t_i})\}_{i=0,\ldots,N}$ be a family of measurable state spaces, $\{\theta_{t_i}\}_{i=0,\ldots,N}$ be a family of measurable reduction mappings $\theta_{t_i} : \mathbb{H}_{t_i} \to \mathbb{X}_{t_i}$, and $\{f_{t_i:t_{i+1}}\}_{i=0,\ldots,N-1}$ 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 triplet $(\{\mathbb{X}_{t_i}\}_{i=0,\ldots,N}, \{\theta_{t_i}\}_{i=0,\ldots,N}, \{f_{t_i:t_{i+1}}\}_{i=0,\ldots,N-1})$ is called a *state reduction across the consecutive time blocks* $[t_i, t_{i+1}]$*,* $i = 0, \ldots, N-1$ if every triplet $(\theta_{t_i}, \theta_{t_{i+1}}, f_{t_i:t_{i+1}})$ is a state reduction, for $i = 0, \ldots, N-1$.

The state reduction across the consecutive time blocks $[t_i, t_{i+1}]$ is said to be *compatible* with the family $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant T}$ of stochastic kernels given in (3) if every triplet $(\theta_{t_i}, \theta_{t_{i+1}}, f_{t_i:t_{i+1}})$ is compatible with the family $\{\rho_{s-1:s}\}_{t_i+1 \leqslant s \leqslant t_{i+1}}$, for $i = 0, \ldots, N-1$.

Assuming the existence of a state reduction across the consecutive time blocks $[t_i, t_{i+1}]$ compatible with the family of stochastic kernels (3), we obtain the existence of a *family of reduced Bellman operators* across the consecutive $(t_{i+1} : t_i)$ as an immediate consequence of multiple applications of Proposition 1, that is,

$$\widetilde{\mathcal{B}}_{t_{i+1}:t_i} : \mathbb{L}^0_+(\mathbb{X}_{t_{i+1}}, \mathcal{X}_{t_{i+1}}) \to \mathbb{L}^0_+(\mathbb{X}_{t_i}, \mathcal{X}_{t_i}) \ , \ \ i = 0, \dots, N-1 \ ,$$

such that, for any function $\widetilde{\varphi}_{t_{i+1}} \in \mathbb{L}^0_+(\mathbb{X}_{t_{i+1}}, \mathcal{X}_{t_{i+1}})$, we have that

$$\big(\widetilde{\mathcal{B}}_{t_{i+1}:t_i} \widetilde{\varphi}_{t_{i+1}}\big) \circ \theta_{t_i} = \mathcal{B}_{t_{i+1}:t_i}\big(\widetilde{\varphi}_{t_{i+1}} \circ \theta_{t_{i+1}}\big) \ .$$

We now consider the family of optimization problems (5) and the associated value functions (6). Thanks to the state reductions, we are able to state the following theorem which establishes dynamic programming equations *across* consecutive time blocks. Its proof is an immediate consequence of multiple applications of Theorem 1 and Proposition 1.

Theorem 2 *Suppose that a state reduction* $\big(\{\mathbb{X}_{t_i}\}_{i=0,\dots,N}, \{\theta_{t_i}\}_{i=0,\dots,N}, \{f_{t_i:t_{i+1}}\}_{i=0,\dots,N-1}\big)$ *exists across the consecutive time blocks* $[t_i, t_{i+1}]$, $i = 0, \dots, N-1$ *as in* (16)*, that is compatible with the family* $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant T}$ *of stochastic kernels given in* (3)*.*

Assume that there exists a reduced criterion

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

such that the cost function j *in* (4) *can be factored as*

$$j = \widetilde{j} \circ \theta_{t_N} \ .$$

We define the family of reduced value functions $\{\widetilde{V}_{t_i}\}_{i=0,\dots,N}$ *by*

$$\widetilde{V}_{t_N} = \widetilde{j} \ , \tag{18a}$$

$$\widetilde{V}_{t_i} = \widetilde{\mathcal{B}}_{t_{i+1}:t_i} \widetilde{V}_{t_{i+1}} \ , \quad \textit{for } i = N-1, \dots, 0 \ . \tag{18b}$$

Then, the family $\{V_{t_i}\}_{i=0,\dots,N}$ *in* (6) *satisfies*

$$V_{t_i} = \widetilde{V}_{t_i} \circ \theta_{t_i} \ , \ \ i = 0, \dots, N \ . \tag{18c}$$

To obtain such a dynamic programming equation across time blocks, we needed the detour of Sect. 2, with a dynamic programming equation over the history space. Thus equipped, it is now possible to propose a decomposition scheme for optimization problems with multiple time scales, using both stochastic programming and stochastic dynamic programming. We detail applications of this scheme in Sect. 4.

4 Applications of Time Blocks Dynamic Programming

We present in this section two applications of the state reduction result stated in Theorem 2.

The first one corresponds to a *two time-scales* optimization problem. A typical instance of such a problem is to optimize long-term investment decisions (slow time-scale) — for example the renewal of batteries in an energy system — but the optimal long-term decisions highly depend on short-term operating decisions (fast time-scale) — for example the way the battery is operated in real-time.

The second application corresponds to a class of stochastic multistage optimization problems arising often in practice, especially when managing stocks (dams for instance). The decision-maker takes two decisions at each time step t: at the beginning of the time interval $[t, t+1[$, the first decision (quantity of water to be turbinated to produce electricity for instance) is taken without knowing the uncertainty that will occur during the time step (decision-hazard framework); at the end of the time interval $[t, t+1[$, an uncertainty variable w_{t+1} is produced and the second decision (quantity of water to be released to avoid dam overflow for instance) is taken once the uncertainty at time step t is revealed (hazard-decision framework). This new class of problems is called *decison-hazard-decision* optimization problems.

4.1 Two Time-Scales Multistage Optimization Problems

In this class of problems, each time index t is represented by a couple (d, m) of indices, with $d \in \{0, \ldots, D+1\}$ and $m \in \{0, \ldots, M\}$: we can think of the index d as an index of days (slow time-scale), and m as an index of minutes (fast time-scale). The corresponding set of time indices is thus

$$\mathbb{T} = \{0, \ldots, D\} \times \{0, \ldots, M\} \cup \{(D+1, 0)\} . \tag{19}$$

At the end of every minute $m-1$ of every day d, that is, at the end of the time interval $\big[(d, m-1), (d, m)\big)$, $0 \leqslant d \leqslant D$ and $1 \leqslant m \leqslant M$, an uncertainty variable $w_{d,m}$ becomes available. Then, at the beginning of the minute m, a decision-maker takes a decision $u_{d,m}$. Moreover, at the beginning of every day d, an uncertainty variable $w_{d,0}$ is produced, followed by a decision $u_{d,0}$. The interplay between uncertainties and decision is thus as follows (compare the chronology with the one in (2)):

$$\begin{aligned} &w_{0,0} \rightsquigarrow u_{0,0} \rightsquigarrow w_{0,1} \rightsquigarrow u_{0,1} \rightsquigarrow \cdots \\ &\qquad \cdots \rightsquigarrow w_{0,M-1} \rightsquigarrow u_{0,M-1} \rightsquigarrow w_{0,M} \rightsquigarrow u_{0,M} \rightsquigarrow w_{1,0} \rightsquigarrow u_{1,0} \rightsquigarrow w_{1,1} \cdots \\ &\qquad\qquad \cdots \rightsquigarrow w_{D,M} \rightsquigarrow u_{D,M} \rightsquigarrow w_{D+1,0} . \end{aligned}$$

We assume that a state reduction (as in Definition 2) is available at the beginning of each day d, so that it becomes possible to write dynamic programming equations by time blocks as stated by Theorem 2. Such state reductions will be for example available when the noises of the different days are stochastically independent.

We present the mathematical formalism to handle such type of problems. In this application, the difficulty to apply Theorem 2 is mainly notational.

Time Span. We consider the set $\mathbb{T}$ equipped with the *lexicographical order*

$$(0,0) < (0,1) < \cdots < (d, M) < (d+1, 0) < \cdots < (D, M-1) < (D, M) < (D+1, 0) . \tag{20a}$$

The set $\mathbb{T}$ of couples in (19) is in one to one correspondence with the (linear) *time span* $\{0, \ldots, T\}$, where

$$T = (D+1) \times (M+1) + 1 , \tag{20b}$$

by the *lexicographic mapping* τ

$$\tau : \{0, \ldots, T\} \to \mathbb{T} \tag{20c}$$

$$t \mapsto \tau(t) = (d, m) . \tag{20d}$$

In the sequel, we will denote by $(d, m) \in \mathbb{T}$ the element of $\{0, \ldots, T\}$ given by $\tau^{-1}(d, m) = d \times (M+1) + m$:

$$\mathbb{T} \ni (d, m) \leftrightarrow \tau^{-1}(d, m) = d \times (M+1) + m \in \{0, \ldots, T\} . \tag{20e}$$

For $(d, m) \leqslant (d', m')$, as ordered by the lexicographical order (20a), we introduce the time interval $((d, m) {:} (d', m')) = \{(d'', m'') \in \mathbb{T} \mid (d, m) \leqslant (d'', m'') \leqslant (d', m')\}$.

History Spaces. For all $(d, m) \in \{0, \ldots, D\} \times \{0, \ldots, M\}$, the decision $u_{d,m}$ takes its values in a measurable set $\mathbb{U}_{d,m}$ equipped with a σ-field $\mathcal{U}_{d,m}$. For all $(d, m) \in \{0, \ldots, D\} \times \{0, \ldots, M\} \cup \{(D+1, 0)\}$, the uncertainty $w_{d,m}$ takes its values in a measurable set $\mathbb{W}_{d,m}$ equipped with a σ-field $\mathcal{W}_{d,m}$.

With the identification (20e), for all $(d, m) \in \mathbb{T}$, we define the *history space* $\mathbb{H}_{(d,m)}$

$$\mathbb{H}_{(d,m)} = \mathbb{W}_{0,0} \times \mathbb{U}_{0,0} \times \mathbb{W}_{0,1} \times \cdots \times \mathbb{U}_{d,m-1} \times \mathbb{W}_{d,m} , \tag{21a}$$

equipped with the *history field* $\mathcal{H}_{(d,m)}$ as in (1). For all $d \in \{0, \ldots, D+1\}$, we define the *slow scale history* h_d element of the *slow scale history space* $\mathbb{H}_d$

$$h_d = h_{(d,0)} \in \mathbb{H}_d = \mathbb{H}_{(d,0)} , \tag{21b}$$

equipped with the *slow scale history field* $\mathcal{H}_d = \mathcal{H}_{(d,0)}$. For all $d \in \{1, \ldots, D\}$, we define the *slow scale partial history space* $\mathbb{H}_{d:d+1}$

$$\mathbb{H}_{d:d+1} = \mathbb{H}_{(d,1):(d+1,0)} = \mathbb{U}_{d,0} \times \mathbb{W}_{d,1} \times \cdots \times \mathbb{U}_{d,M-1} \times \mathbb{W}_{d,M} \times \mathbb{U}_{d,M} \times \mathbb{W}_{d+1,0} , \tag{21c}$$

equipped with the associated *slow scale partial history field* $\mathcal{H}_{d:d+1}$, the case $d = 0$ being

$$\mathbb{H}_{0:1} = \mathbb{H}_{(1,0)} = \mathbb{W}_{0,0} \times \mathbb{U}_{0,0} \times \mathbb{W}_{0,1} \times \cdots \times \mathbb{U}_{0,M-1} \times \mathbb{W}_{0,M} \times \mathbb{U}_{0,M} \times \mathbb{W}_{1,0} . \tag{21d}$$

Stochastic Kernels. Because of the jump from one day to the next, we introduce two families of stochastic kernels[2]:

– a family $\{\rho_{(d,M):(d+1,0)}\}_{0 \leqslant d \leqslant D}$ of stochastic kernels *across* consecutive slow scale steps

$$\rho_{(d,M):(d+1,0)} : \mathbb{H}_{(d,M)} \to \Delta(\mathbb{W}_{d+1,0}) \ , \ \ d = 0, \ldots, D \ , \tag{22a}$$

– a family $\{\rho_{(d,m-1):(d,m)}\}_{0 \leqslant d \leqslant D, 1 \leqslant m \leqslant M}$ of stochastic kernels *within* consecutive slow scale steps

$$\rho_{(d,m-1):(d,m)} : \mathbb{H}_{(d,m-1)} \to \Delta(\mathbb{W}_{d,m}) \ , \ \ d = 0, \ldots, D \ , \ \ m = 1, \ldots, M \ . \tag{22b}$$

History Feedbacks. A history feedback at index $(d, m) \in \mathbb{T}$ is a measurable mapping

$$\gamma_{(d,m)} : \mathbb{H}_{(d,m)} \to \mathbb{U}_{(d,m)} \ .$$

For $(d, m) \leqslant (d', m')$, as ordered by the lexicographical order (20a), we denote by $\Gamma_{(d,m):(d',m')}$ the set of $((d, m) \colon (d', m'))$-history feedbacks.

Slow Scale Value Functions. We suppose given a nonnegative numerical function

$$j : \mathbb{H}_{D+1} \to [0, +\infty] \ , \tag{23}$$

assumed to be measurable with respect to the field $\mathcal{H}_{D+1}$ associated to $\mathbb{H}_{D+1}$.

For $d = 0, \ldots, D$, we build the new stochastic kernels $\rho^{\gamma}_{(d,0):(D+1,0)} : \mathbb{H}_d \to \Delta(\mathbb{H}_{D+1})$ (see Definition 5 in §A.2 for their construction), and we define the *slow scale value functions*

$$V_d(h_d) = \inf_{\gamma \in \Gamma_{(d,0):(D,M)}} \int_{\mathbb{H}_{D+1}} j(h'_{D+1}) \rho^{\gamma}_{(d,0):(D+1,0)}(h_d, \mathrm{d}h'_{D+1}) \ , \ \ \forall h_d \in \mathbb{H}_d \ , \tag{24a}$$

$$V_{D+1} = j \ . \tag{24b}$$

For $d = 0, \ldots, D$, we define a *family of slow scale Bellman operators across* $(d + 1 \colon d)$

$$\mathcal{B}_{d+1:d} : \mathbb{L}^0_+(\mathbb{H}_{d+1}, \mathcal{H}_{d+1}) \to \mathbb{L}^0_+(\mathbb{H}_d, \mathcal{H}_d) \ , \ \ d = 0, \ldots, D \ , \tag{25a}$$

by

$$\mathcal{B}_{d+1:d} = \mathcal{B}_{(d+1,0):(d,0)} = \mathcal{B}_{(d+1,0):(d,M)} \circ \mathcal{B}_{(d,M):(d,M-1)} \circ \ldots \circ \mathcal{B}_{(d,1):(d,0)} \ . \tag{25b}$$

Then, applying repeatedly Theorem 1 leads to the fact that the family $\{V_d\}_{d=0,\ldots,D+1}$ of slow scale value functions (24) satisfies

$$V_{D+1} = j \ , \tag{26a}$$

$$V_d = \mathcal{B}_{d+1:d} V_{d+1} \ , \quad \text{for } d = D, D-1, \ldots, 0 \ . \tag{26b}$$

[2] These families are defined over the time span $\{0, \ldots, T\} \equiv \mathbb{T}$ by the identification (20e) in such a way that the notation is consistent with the notation (3).

Compatible State Reductions. We now rewrite Definition 2 in the context of the two time-scales problem.

Definition 3 (Compatible slow scale reduction) Let $\{(\mathbb{X}_d, \mathcal{X}_d)\}_{d=0,\dots,D+1}$ be a family of measurable state spaces, $\{\theta_d\}_{d=0,\dots,D+1}$ be family of measurable reduction mappings such that

$$\theta_d : \mathbb{H}_d \to \mathbb{X}_d \ ,$$

and $\{f_{d:d+1}\}_{d=0,\dots,D}$ be a family of measurable dynamics such that

$$f_{d:d+1} : \mathbb{X}_d \times \mathbb{H}_{d:d+1} \to \mathbb{X}_{d+1} \ .$$

The triplet $\big(\{\mathbb{X}_d\}_{d=0,\dots,D+1}, \{\theta_d\}_{d=0,\dots,D+1}, \{f_{d:d+1}\}_{d=0,\dots,D}\big)$ is said to be a *slow scale state reduction* if for all $d = 0, \dots, D$

$$\theta_{d+1}\big((h_d, h_{d:d+1})\big) = f_{d:d+1}\big(\theta_d(h_d), h_{d:d+1}\big) \ , \ \ \forall (h_d, h_{d:d+1}) \in \mathbb{H}_{d+1} \ .$$

The slow scale state reduction $\big(\{\mathbb{X}_d\}_{d=0,\dots,D+1}, \{\theta_d\}_{d=0,\dots,D+1}, \{f_{d:d+1}\}_{d=0,\dots,D}\big)$ is said to be *compatible with the two families* $\big\{\rho_{(d,M):(d+1,0)}\big\}_{0\leqslant d\leqslant D}$ *and* $\big\{\rho_{(d,m-1):(d,m)}\big\}_{0\leqslant d\leqslant D, 1\leqslant m\leqslant M}$ *of stochastic kernels* defined in (22a)–(22b) if for any $d = 0, \dots, D$, we have that

- there exists a *reduced stochastic kernel*

  $$\widetilde{\rho}_{(d,M):(d+1,0)} : \mathbb{X}_d \times \mathbb{H}_{(d,0):(d,M)} \to \Delta(\mathbb{W}_{d+1,0}) \ ,$$

  such that the stochastic kernel $\rho_{(d,M):(d+1,0)}$ in (22a) can be factored as

  $$\rho_{(d,M):(d+1,0)}(h_{d,M}, \mathrm{d}w_{d+1,0}) = \widetilde{\rho}_{(d,M):(d+1,0)}\big(\theta_d(h_d), h_{(d,0):(d,M)}, \mathrm{d}w_{d+1,0}\big) \ , \ \ \forall h_{d,M} \in \mathbb{H}_{(d,M)} \ ,$$

- for each $m = 1, \dots, M$, there exists a *reduced stochastic kernel*

  $$\widetilde{\rho}_{(d,m-1):(d,m)} : \mathbb{X}_d \times \mathbb{H}_{(d,0):(d,m-1)} \to \Delta(\mathbb{W}_{d,m}) \ ,$$

  such that the stochastic kernel $\rho_{(d,m-1):(d,m)}$ in (22b) can be factored as

  $$\rho_{(d,m-1):(d,m)}(h_{d,m-1}, \mathrm{d}w_{d,m}) = \widetilde{\rho}_{(d,m-1):(d,m)}\big(\theta_d(h_d), h_{(d,0):(d,m-1)}, \mathrm{d}w_{d,m}\big) \ , \ \ \forall h_{d,m-1} \in \mathbb{H}_{(d,m-1)} \ .$$

Dynamic Programming Equations. Using the reduced stochastic kernels of Definition 3, we apply Proposition 1 and obtain a *family of slow scale reduced Bellman operators across* $(d+1\!:\!d)$

$$\widetilde{\mathcal{B}}_{d+1:d} : \mathbb{L}^0_+(\mathbb{X}_{d+1}, \mathcal{X}_{d+1}) \to \mathbb{L}^0_+(\mathbb{X}_d, \mathcal{X}_d) \ , \ \ d = 0, \dots, D \ . \tag{29}$$

We are now able to state the main result of this section.

Theorem 3 *Assume that there exists a compatible slow scale state reduction* $\big(\{\mathbb{X}_d\}_{d=0,\dots,D+1}, \{\theta_d\}_{d=0,\dots,D+1}, \{f_{d:d+1}\}_{d=0,\dots,D}\big)$ *and that there exists a* reduced criterion

$$\widetilde{j} : \mathbb{X}_{D+1} \to [0, +\infty] \ ,$$

such that the cost function j *in* (23) *can be factored as*

$$j = \widetilde{j} \circ \theta_{D+1} \ .$$

We define the family of reduced value functions $\{\widetilde{V}_d\}_{d=0,\dots,D+1}$ *by*

$$\widetilde{V}_{D+1} = \widetilde{j} \ , \tag{31a}$$

$$\widetilde{V}_d = \widetilde{\mathcal{B}}_{d+1:d}\widetilde{V}_{d+1} \ , \quad \text{for } d = D, \dots, 0 \ . \tag{31b}$$

Then, the family $\{V_d\}_{d=0,\dots,D+1}$ *of slow scale value functions* (24) *satisfies*

$$V_d = \widetilde{V}_d \circ \theta_d \ , \ \ d = 0, \dots, D \ . \tag{31c}$$

Proof Since the triplet $(\{\mathbb{X}_d\}_{d=0,\dots,D+1}, \{\theta_d\}_{d=0,\dots,D+1}, \{f_{d:d+1}\}_{d=0,\dots,D})$ is a state reduction across the time blocks $[(d,0),(d+1,0)]$, which is compatible with the family $\{\rho_{(d,0):(d+1,0)}\}_{0\leqslant d\leqslant D}$ of stochastic kernels, the proof is an immediate consequence of Theorem 2.

Thanks to Theorem 3, we are able to replace the optimization problem formulated on the whole time set $\mathbb{T}$ by a sequence of D optimization subproblems formulated each on a single time block $[(d,0),(d+1,0)]$. Moreover, the numerical burden of the method remains reasonable provided that the dimensions of the spaces $\mathbb{X}_d$ remain small, thus avoiding the curse of dimensionality. This is the benefit induced by *dynamic programming* which makes possible a time decomposition of the problem. However, to make the method operational, we need to compute the functions $\widetilde{V}_d$, whose expression is available thanks to Proposition 1:

$$\begin{aligned}\widetilde{V}_d(x_d) = \inf_{u_{d,0}\in\mathbb{U}_{d,0}} \int_{\mathbb{W}_{d,1}} &\widetilde{\rho}_{(d,0):(d,1)}(x_d, \mathrm{d}w_{d,1}) \dots \\ &\inf_{u_{d,M-1}\in\mathbb{U}_{d,M-1}} \int_{\mathbb{W}_{d,M}} \widetilde{\rho}_{(d,M-1):(d,M)}(x_d, u_{d,0}, w_{d,1}, \cdots, w_{d,M-1}, \mathrm{d}w_{d,M}) \\ &\quad\inf_{u_{d,M}\in\mathbb{U}_{d,M}} \int_{\mathbb{W}_{d+1,0}} \widetilde{V}_{d+1}\big(\widetilde{f}_{d:d+1}(x_d, u_{d,0}, w_{d,1}, \cdots, u_{d,M-1}, w_{d,M}, u_{d,M}, w_{d+1,0})\big) \\ &\qquad\qquad\widetilde{\rho}_{(d,M):(d+1,0)}(x_d, u_{d,0}, w_{d,1}, \cdots, w_{d,M}, \mathrm{d}w_{d+1,0}) \; . \end{aligned} \tag{32}$$

In many practical situations, this computation is tractable by using *stochastic programming*. For example, if the stochastic kernels $\widetilde{\rho}_{(d,m):(d,m+1)}$ do not depend on the past controls $(u_{d,0},\cdots,u_{d,m-1})$, then it is possible to approximate the optimization problem (32) by using scenario tree techniques. Note that these last techniques 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 slow time scale across consecutive slow time steps, when the slow time scale noises are supposed to be stochastically independent; produce slow time scale Bellman functions;
- use stochastic programming at short time scale, within two consecutive slow time steps; the final short time scale cost is given by the slow time scale Bellman functions; no stagewise independence assumption is required for the short time scale noises.

4.2 Decision-Hazard-Decision Optimization Problems

We apply the reduction by time blocks to the so-called *decision-hazard-decision* dynamic programming.

4.2.1 Motivation for the Decision-Hazard-Decision Framework

We illustrate our motivation with a single dam management problem. We can model the dynamics of the water volume in a dam by

$$S_{t+1} = \min\{S^\sharp, S_t - q_t + a_{t+1}\} \; , \tag{33}$$

where $t = t_0, t_0+1, \dots, T-1$ and

- $S^\sharp$ is the maximal dam volume,
- S_t is the volume (stock) of water at the beginning of period $[t, t+1[$,
- a_{t+1} is the inflow water volume (rain, etc.) during $[t, t+1[$,
- q_t is the turbined outflow volume during $[t, t+1[$ (control variable),
  - decided at the *beginning* of period $[t, t+1[$,
  - chosen such that $0 \leqslant q_t \leqslant S_t$,
  - supposed to depend on the stock S_t but not on the inflow water a_{t+1}.

The min operation in Equation (33) ensures that the dam volume always remains below its maximal capacity, but induces a non linearity in the dynamics.

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

$$S_{t+1} = S_t - q_t - a_{t+1} - r_{t+1} \; , \tag{34}$$

where $t = t_0, t_0+1, \dots, T-1$ and

- r_{t+1} is the spilled volume
  - decided at the *end* of period $[t, t+1[$,
  - supposed to depend on the stock S_t and on the inflow water a_{t+1},
  - and chosen such that $0 \leqslant S_t - q_t + a_{t+1} - r_{t+1} \leqslant S^\sharp$.

Thus, with the formulation (34), we pay the price to add one control r_{t+1}, but we obtain a linear model instead of the nonlinear model (33). This is especially interesting when using the stochastic dual dynamic programming (SDDP), for which the linearity of the dynamics is used to obtain the convexity properties required by the algorithm.

4.2.2 Decision-Hazard-Decision Framework

We consider stochastic optimization problems where, during the time interval between two time steps, the decision-maker takes two decisions. At the end of the time interval $[s-1, s[$, an uncertainty variable $w_s^\flat$ is produced, and then, at the beginning of the time interval $[s, s+1[$, the decision-maker takes a *head decision* $u_s^\sharp$. What is new is that, at the end of the time interval $[s, s+1[$, when an uncertainty variable $w_{s+1}^\flat$ is produced, the decision-maker has the possibility to make a *tail decision* $u_{s+1}^\flat$. This latter decision $u_{s+1}^\flat$ can be thought as a *recourse* variable for a two stage stochastic optimization problem that would take place inside the time interval $[s, s+1[$. 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 (2)):

$$w_0^\sharp \rightsquigarrow u_0^\sharp \rightsquigarrow w_1^\flat \rightsquigarrow u_1^\flat \rightsquigarrow u_1^\sharp \rightsquigarrow w_2^\flat \rightsquigarrow \quad \ldots \quad \rightsquigarrow w_{S-1}^\flat \rightsquigarrow u_{S-1}^\flat \rightsquigarrow u_{S-1}^\sharp \rightsquigarrow w_S^\flat \rightsquigarrow u_S^\flat \; .$$

Let $S \in \mathbb{N}^*$. For each time $s = 0, 1, 2 \ldots, S-1$, the *head decision* $u_s^\sharp$ takes values in a measurable set $\mathbb{U}_s^\sharp$, equipped with a σ-field $\mathcal{U}_s^\sharp$. For each time $s = 1, 2 \ldots, S$, the *tail decision* $u_s^\flat$ takes values in measurable set $\mathbb{U}_s^\flat$, equipped with a σ-field $\mathcal{U}_s^\flat$. For each time $s = 1, 2 \ldots, S$, the uncertainty $w_s^\flat$ takes its values in a measurable set $\mathbb{W}_s^\flat$, equipped with a σ-field $\mathcal{W}_s^\flat$. For time $s = 0$, the uncertainty $w_0^\sharp$ takes its values in a measurable set $\mathbb{W}_0^\sharp$, equipped with a σ-field $\mathcal{W}_0^\sharp$.

Again, in this application, the difficulty to apply Theorem 2 is mainly notational.

History Spaces. For $s = 0, 1, 2 \ldots, S$, we define the *head history space*

$$\mathbb{H}_s^\sharp = \mathbb{W}_0^\sharp \times \prod_{s'=0}^{s-1} \left(\mathbb{U}_{s'}^\sharp \times \mathbb{W}_{s'+1}^\flat \times \mathbb{U}_{s'+1}^\flat \right) , \tag{35a}$$

and its associated *head history field* $\mathcal{H}_s^\sharp$. We also define, for $s = 1, 2 \ldots, S$, the *tail history space*

$$\mathbb{H}_s^\flat = \mathbb{H}_{s-1}^\sharp \times \mathbb{U}_{s-1}^\sharp \times \mathbb{W}_s^\flat , \tag{35b}$$

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

Stochastic Kernels. We introduce a family of stochastic kernels $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant S}$, with

$$\rho_{s-1:s} : \mathbb{H}_{s-1}^\sharp \to \Delta(\mathbb{W}_s^\flat) . \tag{36}$$

History Feedbacks. For $s = 0, \ldots, S-1$, a *head history feedback* at time s is a measurable mapping

$$\gamma_s^\sharp : \mathbb{H}_s^\sharp \to \mathbb{U}_s^\sharp .$$

We call $\Gamma_s^\sharp$ the *set of head history feedbacks at time s*, and we define $\Gamma_{s:S}^\sharp = \Gamma_s^\sharp \times \cdots \times \Gamma_S^\sharp$. We also define, for all $s = 1, 2 \ldots, S$, a *tail history feedback* at time s as a measurable mapping

$$\gamma_s^\flat : \mathbb{H}_s^\flat \to \mathbb{U}_s^\flat .$$

We call $\Gamma_s^\flat$ the *set of tail history feedbacks at time s*, and we define $\Gamma_{s:S}^\flat = \Gamma_s^\flat \times \cdots \times \Gamma_S^\flat$.

Value Functions. We consider a nonnegative numerical function

$$j : \mathbb{H}^\sharp_S \to [0, +\infty] \ , \tag{38}$$

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

For $s = 0, \dots, S$, we define *value functions* by

$$V_s(h^\sharp_s) = \inf_{\gamma^\sharp \in \Gamma^\sharp_{s:S-1}, \gamma^\flat \in \Gamma^\sharp_{s+1:S}} \int_{\mathbb{H}^\sharp_S} j(h'_S) \rho^{\gamma^\sharp, \gamma^\flat}_{s:S}(h^\sharp_s, \mathrm{d}h'_S) \ , \ \ \forall h^\sharp_s \in \mathbb{H}^\sharp_s \ , \tag{39}$$

where $\rho^{\gamma^\sharp, \gamma^\flat}_{s:S}$ has to be understood as $\rho^{\gamma}_{s:S}$ (see Definition 5), with

$$\gamma_s(h^\sharp_s) = \gamma^\sharp_s(h^\sharp_s) \ , \ \ \forall h^\sharp_s \in \mathbb{H}^\sharp_s \ , \tag{40a}$$

$$\gamma_{s'}(h^\flat_{s'}) = \Big(\gamma^\flat_{s'}(h^\flat_{s'}), \gamma^\sharp_{s'}\big(h^\flat_{s'}, \gamma^\flat_{s'}(h^\flat_{s'})\big) \Big) \ , \ \ \forall s' = s+1, \dots, S-1 \ , \ \ \forall h^\flat_{s'} \in \mathbb{H}^\flat_{s'} \ , \tag{40b}$$

$$\gamma_S(h^\flat_S) = \gamma^\flat_S(h^\flat_S) \ , \ \ \forall h^\flat_S \in \mathbb{H}^\flat_S \ . \tag{40c}$$

The following proposition, whose proof has been relegated in A.3.3, characterizes the dynamic programming equations in the decision-hazard-decision framework.

Proposition 2 *For $s = 0, \dots, S-1$, we define the* Bellman operator

$$\mathcal{B}_{s+1:s} : \mathbb{L}^0_+(\mathbb{H}^\sharp_{s+1}, \mathcal{H}^\sharp_{s+1}) \to \mathbb{L}^0_+(\mathbb{H}^\sharp_s, \mathcal{H}^\sharp_s) \tag{41a}$$

such that, for all $\varphi \in \mathbb{L}^0_+(\mathbb{H}^\sharp_{s+1}, \mathcal{H}^\sharp_{s+1})$ and for all $h^\sharp_s \in \mathbb{H}^\sharp_s$,

$$\big(\mathcal{B}_{s+1:s}\varphi\big)(h^\sharp_s) = \inf_{u^\sharp_s \in \mathbb{U}^\sharp_s} \int_{\mathbb{W}^\flat_{s+1}} \Big(\inf_{u^\flat_{s+1} \in \mathbb{U}^\flat_{s+1}} \varphi(h^\sharp_s, u^\sharp_s, w^\flat_{s+1}, u^\flat_{s+1}) \Big) \rho_{s:s+1}(h^\sharp_s, dw^\flat_{s+1}) \ . \tag{41b}$$

Then the value functions (39) *satisfy*

$$V_S = j \ , \tag{41c}$$

$$V_s = \mathcal{B}_{s+1:s} V_{s+1} \ , \ \ \forall s = 0, \dots, S-1 \ . \tag{41d}$$

Compatible State Reductions. We now rewrite Definition 2 in the context of a decision-hazard-decision problem.

Definition 4 (Compatible state reduction) Let $\{\mathbb{X}_s\}_{s=0,\dots,S}$ be a family of state spaces, $\{\theta_s\}_{s=0,\dots,S}$ be a family of measurable reduction mappings such that

$$\theta_s : \mathbb{H}^\sharp_s \to \mathbb{X}_s \ ,$$

and $\{f_{s:s+1}\}_{s=0,\dots,S-1}$ be a family of measurable dynamics such that

$$f_{s:s+1} : \mathbb{X}_s \times \mathbb{U}^\sharp_s \times \mathbb{W}_{s+1} \times \mathbb{U}^\flat_{s+1} \to \mathbb{X}_{s+1} \ .$$

The triplet $\big(\{\mathbb{X}_s\}_{s=0,\dots,S}, \{\theta_s\}_{s=0,\dots,S}, \{f_{s:s+1}\}_{s=0,\dots,S-1} \big)$ is said to be a *decision-hazard-decision state reduction* if, for all $s = 0, \dots, S-1$, we have that

$$\theta_{s+1}\big((h_s, u^\sharp_s, w_{s+1}, u^\flat_{s+1})\big) = f_{s:s+1}\big(\theta_s(h_s), u^\sharp_s, w_{s+1}, u^\flat_{s+1}\big) \ ,$$
$$\forall (h_s, u^\sharp_s, w_{s+1}, u^\flat_{s+1}) \in \mathbb{H}^\sharp_s \times \mathbb{U}^\sharp_s \times \mathbb{W}_{s+1} \times \mathbb{U}^\flat_{s+1} \ .$$

The decision-hazard-decision state reduction is said to be *compatible with the family $\{\rho_{s:s+1}\}_{0 \leqslant s \leqslant S-1}$ of stochastic kernels* in (36) if there exists a family $\{\widetilde{\rho}_{s:s+1}\}_{0 \leqslant s \leqslant S-1}$ of *reduced stochastic kernels*

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

such that, for each $s = 0, \dots, S-1$, the stochastic kernel $\rho_{s:s+1}$ in (36) can be factored as

$$\rho_{s:s+1}(h^\sharp_s, \mathrm{d}w_{s+1}) = \widetilde{\rho}_{s:s+1}\big(\theta_s(h^\sharp_s), \mathrm{d}w_{s+1}\big) \ , \ \ \forall h^\sharp_s \in \mathbb{H}^\sharp_s \ .$$

Dynamic Programming Equations. We state the main result of this section.

Theorem 4 *Assume that there exists a decision-hazard-decision state reduction* $\big(\{\mathbb{X}_s\}_{s=0,\dots,S}, \{\theta_s\}_{s=0,\dots,S}, \{f_{s:s+1}\}_{s=0,\dots,S-1}\big)$ *and that there exists a* reduced criterion

$$\widetilde{j} : \mathbb{X}_S \to [0, +\infty] \; ,$$

such that the cost function j in (38) *can be factored as*

$$j = \widetilde{j} \circ \theta_S \; .$$

We define a family of reduced Bellman operators across $(s+1 : s)$

$$\widetilde{\mathcal{B}}_{s+1:s} : \mathbb{L}^0_+(\mathbb{X}_{s+1}, \mathcal{X}_{s+1}) \to \mathbb{L}^0_+(\mathbb{X}_s, \mathcal{X}_s) \; , \quad s = 1, \dots, S-1 \; , \tag{45a}$$

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

$$(\widetilde{\mathcal{B}}_{s+1:s} \widetilde{\varphi})(x_s) = \inf_{u^\sharp_s \in \mathbb{U}^\sharp_s} \int_{\mathbb{W}_{s+1}} \Big(\inf_{u^\flat_{s+1} \in \mathbb{U}^\flat_{s+1}} \widetilde{\varphi}\big(f_{s:s+1}(x_s, u^\sharp_s, w_{s+1}, u^\flat_{s+1})\big)\Big) \widetilde{\rho}_{s:s+1}(x_s, dw_{s+1}) \; . \tag{45b}$$

**We define the family of* reduced value functions $\{\widetilde{V}_s\}_{s=0,\dots,S}$ *by*

$$\widetilde{V}_S = \widetilde{j} \tag{46a}$$

$$\widetilde{V}_s = \widetilde{\mathcal{B}}_{s+1:s} \widetilde{V}_{s+1} \quad \text{for } s = S-1, \dots, 0 \; . \tag{46b}$$

Then, the value functions V_s defined by (39) *satisfy*

$$V_s = \widetilde{V}_s \circ \theta_s \; , \quad s = 0, \dots, S \; . \tag{47}$$

Proof It has been shown in the proof of Proposition 2 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 Theorem 3.

Theorem 4 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 §4.2.1.

5 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). One also finds space decomposition methods [1].

This paper is part of a general research program that consists in *mixing* different decomposition bricks. Here, we tackled the issue of mixing time decomposition (stochastic dynamic programming) with scenario decomposition. For this purpose, we have revisited the notion of state, and have provided a way to perform time decomposition but only accross specified time blocks. Inside a time block, one can then use stochastic programming methods, like scenario decomposition. Our time blocks decomposition scheme is especially adapted to multi time-scales stochastic optimization problems. In this vein, we have shown its application to two time-scales and to the novel class of decision-hazard-decision problems.

We are currently working on how to mix time decomposition (stochastic dynamic programming) with space/units decomposition.

Acknowledgements. We thank Roger Wets for the fruitful discussions about the possibility of mixing stochastic dynamic programming with progressive hedging. We thank an anonymous reviewer for challenging our first version of the paper: the current version has been deeply restructured.

A Technical Details and Proofs

In this section, we provide technical details, constructions and proofs of results in the paper.

A.1 Histories, Feedbacks and Flows

We introduce the notations

$$\mathbb{W}_{r:t} = \prod_{s=r}^{t} \mathbb{W}_s \ , \quad 0 \leqslant r \leqslant t \leqslant T \tag{48a}$$

$$\mathbb{U}_{r:t} = \prod_{s=r}^{t} \mathbb{U}_s \ , \quad 0 \leqslant r \leqslant t \leqslant T-1 \tag{48b}$$

$$\mathbb{H}_{r:t} = \prod_{s=r-1}^{t-1} (\mathbb{U}_s \times \mathbb{W}_{s+1}) = \mathbb{U}_{r-1} \times \mathbb{W}_r \times \cdots \times \mathbb{U}_{t-1} \times \mathbb{W}_t \ , \quad 1 \leqslant r \leqslant t \leqslant T \ . \tag{48c}$$

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

$$[h_t]^{\mathbb{W}}_{r:s} = (w_r, \ldots, w_s) = w_{r:s} \in \mathbb{W}_{r:s} \ , \quad 0 \leqslant r \leqslant s \leqslant t \ , \tag{49a}$$

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

$$[h_t]^{\mathbb{U}}_{r:s} = (u_{r-1}, \ldots, u_{s-1}) = u_{r-1:s-1} \in \mathbb{U}_{r-1:s-1} \ , \quad 1 \leqslant r \leqslant s \leqslant t \ , \tag{49b}$$

and the $(r\!:\!s)$-*history subpart*

$$[h_t]_{r:s} = (u_{r-1}, w_r, \ldots, u_{s-1}, w_s) = h_{r:s} \in \mathbb{H}_{r:s} \ , \quad 1 \leqslant r \leqslant s \leqslant t \ , \tag{49c}$$

so that we obtain, for $0 \leqslant r+1 \leqslant s \leqslant t$,

$$h_t = (\underbrace{w_0, u_0, w_1, \ldots, u_{r-1}, w_r}_{h_r}, \underbrace{u_r, w_{r+1}, \ldots, u_{t-2}, w_{t-1}, u_{t-1}, w_t}_{h_{r+1:t}}) = (h_r, h_{r+1:t}) \ . \tag{49d}$$

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

$$\Phi^{\gamma}_{r:t} : \mathbb{H}_r \times \mathbb{W}_{r+1:t} \to \mathbb{H}_t \tag{50a}$$

$$(h_r, w_{r+1:t}) \mapsto (h_r, \gamma_r(h_r), w_{r+1}, \gamma_{r+1}(h_r, \gamma_r(h_r), w_{r+1}), w_{r+2}, \cdots, \gamma_{t-1}(h_{t-1}), w_t) \ , \tag{50b}$$

that is,

$$\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t}) = (h_r, u_r, w_{r+1}, u_{r+1}, w_{r+2}, \ldots, u_{t-1}, w_t) \ , \tag{50c}$$

$$\text{with } \ h_s = (h_r, u_r, w_{r+1}, \ldots, u_{s-1}, w_s) \ , \quad r < s \leqslant t \ , \tag{50d}$$

$$\text{and } \ u_s = \gamma_s(h_s) \ , \quad r < s \leqslant t-1 \ . \tag{50e}$$

When $0 \leqslant r = t \leqslant T$, we put

$$\Phi^{\gamma}_{r:r} : \mathbb{H}_r \to \mathbb{H}_r \ , \quad h_r \mapsto h_r \ . \tag{50f}$$

With this convention, the expression $\Phi^{\gamma}_{r:t}$ makes sense when $0 \leqslant r \leqslant t \leqslant T$: when $r = t$, no $(r:r-1)$-history feedback exists, but none is needed. The mapping $\Phi^{\gamma}_{r:t}$ gives the history at time t as a function of the initial history h_r at time r and of the history feedbacks $\{\gamma_s\}_{s=r,\ldots,t-1} \in \Gamma_{r:t-1}$. An immediate consequence of this definition are the *flow properties*:

$$\Phi^{\gamma}_{r:t+1}(h_r, w_{r+1:t+1}) = \Big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t}), \gamma_t\big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t})\big), w_{t+1}\Big) \ , \quad 0 \leqslant r \leqslant t \leqslant T-1 \ , \tag{51a}$$

$$\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t}) = \Phi^{\gamma}_{r+1:t}\big((h_r, \gamma_r(h_r), w_{r+1}), w_{r+2:t}\big) \ , \quad 0 \leqslant r < t \leqslant T \ . \tag{51b}$$

A.2 Building Stochastic Kernels from History Feedbacks

Definition 5 Let r and t be given such that $0 \leqslant r \leqslant t \leqslant T$.

– When $0 \leqslant r < t \leqslant T$, for
  1. a $(r:t-1)$-history feedback $\gamma = \{\gamma_s\}_{s=r,\dots,t-1} \in \Gamma_{r:t-1}$,
  2. a family $\{\rho_{s-1:s}\}_{r+1 \leqslant s \leqslant t}$ of stochastic kernels

$$\rho_{s-1:s} : \mathbb{H}_{s-1} \to \Delta(\mathbb{W}_s) \ , \quad s = r+1, \dots, t \ ,$$

we define a stochastic kernel

$$\rho^{\gamma}_{r:t} : \mathbb{H}_r \to \Delta(\mathbb{H}_t) \tag{52a}$$

by, for any $\varphi : \mathbb{H}_t \to [0, +\infty]$, measurable nonnegative numerical function, that is, $\varphi \in \mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t)$, [3]

$$\int_{\mathbb{H}_t} \varphi(h'_r, h'_{r+1:t}) \rho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_t) = \int_{\mathbb{W}_{r+1:t}} \varphi\big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t})\big) \prod_{s=r+1}^{t} \rho_{s-1:s}\big(\Phi^{\gamma}_{r:s-1}(h_r, w_{r+1:s-1}), \mathrm{d}w_s\big) \ . \tag{52b}$$

– When $0 \leqslant r = t \leqslant T$, we define

$$\rho^{\gamma}_{r:r} : \mathbb{H}_r \to \Delta(\mathbb{H}_r) \ , \quad \rho^{\gamma}_{r:r}(h_r, \mathrm{d}h'_r) = \delta_{h_r}(\mathrm{d}h'_r) \ . \tag{52c}$$

The stochastic kernels $\rho^{\gamma}_{r:t}$ on $\mathbb{H}_t$, given by (52), are of the form

$$\rho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_t) = \rho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_r \mathrm{d}h'_{r+1:t}) = \delta_{h_r}(\mathrm{d}h'_r) \otimes \varrho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_{r+1:t}) \ , \tag{53}$$

where, for each $h_r \in \mathbb{H}_r$, the probability distribution $\varrho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_{r+1:t})$ 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 3 *Following Definition 5, we can define a family $\{\rho^{\gamma}_{s:t}\}_{r \leqslant s \leqslant t}$ of stochastic kernels. This family has the flow property, that is, for $s < t$,*

$$\rho^{\gamma}_{s:t}(h_s, \mathrm{d}h'_t) = \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}\big(h_s, \mathrm{d}w_{s+1}\big) \rho^{\gamma}_{s+1:t}\Big(\big(h_s, \gamma_s(h_s), w_{s+1}\big), \mathrm{d}h'_t\Big) \ . \tag{54}$$

Proof Let $s < t$. For any $\varphi : \mathbb{H}_t \to [0, +\infty]$, we have that

$$\int_{\mathbb{H}_t} \varphi(h'_s, h'_{s+1:t}) \rho^{\gamma}_{s:t}(h_s, \mathrm{d}h'_t) \tag{55a}$$

$$= \int_{\mathbb{W}_{s+1:t}} \varphi\big(\Phi^{\gamma}_{s:t}(h_s, w_{s+1:t})\big) \prod_{s'=s+1}^{t} \rho_{s'-1:s'}\big(\Phi^{\gamma}_{s:s'-1}(h_s, w_{s+1:s'-1}), \mathrm{d}w_{s'}\big)$$

by the definition (52b) of the stochastic kernel $\rho^{\gamma}_{s:t}$,

$$= \int_{\mathbb{W}_{s+1:t}} \varphi\big(\Phi^{\gamma}_{s:t}(h_s, w_{s+1:t})\big) \rho_{s:s+1}\big(h_s, \mathrm{d}w_{s+1}\big) \prod_{s'=s+2}^{t} \rho_{s'-1:s'}\big(\Phi^{\gamma}_{s:s'-1}(h_s, w_{s+1:s'-1}), \mathrm{d}w_{s'}\big)$$

by the property (50f) of the flow $\Phi^{\gamma}_{s:s}$,

$$= \int_{\mathbb{W}_{s+1:t}} \varphi\big(\Phi^{\gamma}_{s+1:t}\big((h_s, \gamma_s(h_s), w_{s+1}), w_{s+2:t}\big)\big) \rho_{s:s+1}\big(h_s, \mathrm{d}w_{s+1}\big) \prod_{s'=s+2}^{t} \rho_{s'-1:s'}\big(\Phi^{\gamma}_{s+1:s'-1}\big((h_s, \gamma_s(h_s), w_{s+1}), w_{s+2:s'-1}\big), \mathrm{d}w_{s'}\big)$$

by the flow property (51b),

$$= \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}\big(h_s, \mathrm{d}w_{s+1}\big) \int_{\mathbb{W}_{s+2:t}} \varphi\big(\Phi^{\gamma}_{s+1:t}\big((h_s, \gamma_s(h_s), w_{s+1}), w_{s+2:t}\big)\big) \prod_{s'=s+2}^{t} \rho_{s'-1:s'}\big(\Phi^{\gamma}_{s+1:s'-1}\big((h_s, \gamma_s(h_s), w_{s+1}), w_{s+2:s'-1}\big), \mathrm{d}w_{s'}\big)$$

[3] See Footnote 1.

by Fubini Theorem [7, p.137],

$$
= \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}\big(h_s, \mathrm{d}w_{s+1}\big) \int_{\mathbb{H}_t} \varphi\big((h'_s, \gamma_s(h'_s), w'_{s+1}), h'_{s+2:t}\big) \rho^{\gamma}_{s+1:t}\big((h_s, \gamma_s(h_s), w_{s+1}), \mathrm{d}h'_t\big)
$$

by definition (52b) of $\rho^{\gamma}_{s+1:t}$,

$$
= \int_{\mathbb{H}_t} \varphi\big((h'_s, \gamma_s(h'_s), w'_{s+1}), h'_{s+2:t}\big) \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}\big(h_s, \mathrm{d}w_{s+1}\big) \rho^{\gamma}_{s+1:t}\big((h_s, \gamma_s(h_s), w_{s+1}), \mathrm{d}h'_t\big) \tag{55b}
$$

by Fubini Theorem and by definition (52b) of $\rho^{\gamma}_{s:t}$. As the two expressions (55a) and (55b) are equal for any $\varphi : \mathbb{H}_t \to [0, +\infty]$, we deduce the flow property (54). This ends the proof.

A.3 Proofs

A.3.1 Proof of Theorem 1

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

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

$$
f\big(h_t, u_t, w_{t+1}\big) = \big(h_t, u_t, w_{t+1}\big) = h_{t+1} \in \mathbb{H}_{t+1} = \mathbb{H}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \ . \tag{56}
$$

Then, the family $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant T}$ of stochastic kernels (3) gives a family of disturbance kernels that do not depend on the current control. The criterion to be minimized (4) is a function of the history at time T, thus of the state at time T. Problem (5) 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 1 follows from the results developed in Chap. 7, 8 and 10 of [3] in a Borel setting. 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 (56) 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 (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 1 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 (4) is nonnegative.

A.3.2 Proof of Proposition 1

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

$$
\varphi_t = \widetilde{\varphi}_t \circ \theta_t \ . \tag{57}
$$

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

$$
\varphi_r = \mathcal{B}_{t:r}\varphi_t = \mathcal{B}_{r+1:r} \circ \cdots \circ \mathcal{B}_{t:t-1}\varphi_t \ . \tag{58}
$$

We will show that there exists a measurable nonnegative numerical function

$$
\widetilde{\varphi}_r : \mathbb{X}_r \to [0, +\infty]
$$

such that

$$
\varphi_r = \widetilde{\varphi}_r \circ \theta_r \ . \tag{59}
$$

First, we show by backward induction that, for all $s \in \{r, \ldots, t\}$, 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 $\widetilde{\varphi}_r = \overline{\varphi}_r$ satisfies (59).

- For $s = t$, we have, by (57) and by (10c), that

  $$
  \varphi_t(h_t) = \widetilde{\varphi}_t\big(\theta_t(h_t)\big) = \widetilde{\varphi}_t\big(f_{r:t}(\theta_r(h_r), h_{r+1:t})\big) \ ,
  $$

  so that the measurable nonnegative numerical function $\overline{\varphi}_t$ is given by $\widetilde{\varphi}_t \circ f_{r:t}$.

– Assume that, at $s+1$, the result holds true, that is,

$$\varphi_{s+1}(h_{s+1}) = \overline{\varphi}_{s+1}(\theta_r(h_r), h_{r+1:s+1}) \; . \tag{60}$$

Then, by (58),

$$\begin{aligned}\varphi_s(h_s) &= \big(\mathcal{B}_{s+1:s}\varphi_{s+1}\big)(h_s) \\ &= \inf_{u_s \in \mathbb{U}_s} \int_{\mathbb{W}_{s+1}} \varphi_{s+1}\big((h_s, u_s, w_{s+1})\big)\rho_{s:s+1}(h_s, \mathrm{d}w_{s+1})\end{aligned}$$

by definition (7) of the Bellman operator

$$= \inf_{u_s \in \mathbb{U}_s} \int_{\mathbb{W}_{s+1}} \overline{\varphi}_{s+1}\big((\theta_r(h_r), (h_{r+1:s}, u_s, w_{s+1}))\big)\rho_{s:s+1}(h_s, \mathrm{d}w_{s+1})$$

by induction assumption (60)

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

by compatibility (11) of the stochastic kernel

$$= \overline{\varphi}_s\big(\theta_r(h_r), h_{r+1:s}\big) \; ,$$

where

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

The result thus 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 $\widetilde{\varphi}_r = \overline{\varphi}_r$ gives the expected result.

A.3.3 Proof of Proposition 2

Proof We now show that the setting in §4.2 is a particular kind of two time scales problem as seen in §4.1. For this purpose, we introduce a *spurious uncertainty variable* $w_s^\sharp$ taking values in a singleton set $\mathbb{W}_s^\sharp = \{\overline{w}_s^\sharp\}$, equipped with the trivial σ-field $\{\emptyset, \mathbb{W}_s^\sharp\}$, for each time $s = 1, 2 \ldots, S$. Now, we obtain the following sequence of events:

$$w_0^\sharp \rightsquigarrow u_0^\sharp \rightsquigarrow w_1^\flat \rightsquigarrow u_1^\flat \rightsquigarrow w_1^\sharp \rightsquigarrow u_1^\sharp \rightsquigarrow w_2^\flat \rightsquigarrow u_2^\flat \rightsquigarrow w_2^\sharp \rightsquigarrow u_2^\sharp \rightsquigarrow \quad \ldots$$
$$\rightsquigarrow w_{S-1}^\flat \rightsquigarrow u_{S-1}^\flat \rightsquigarrow w_{S-1}^\sharp \rightsquigarrow u_{S-1}^\sharp \rightsquigarrow w_S^\flat \rightsquigarrow u_S^\flat \rightsquigarrow w_S^\sharp \; ,$$

which coincides with a two time scales problem:

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

We introduce the sets

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

As a consequence, we observe that the two time scales history spaces in §4.1 are in one to one correspondence with the decision-hazard-decision history spaces and fields in (35a)–(35b) as follows:

for $d = 0, 1, 2 \dots, S$,

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

for $d = 0, 1, 2 \dots, S$,

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

for $d = 0, 1, 2 \dots, S-1$,

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

for $d = 0, 1, 2 \dots, S-1$,

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

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 (38) (decision-hazard-decision setting) corresponds to $j \circ \big[\,\cdot\,\big]^\sharp$ in the two time scales setting in §4.1. The feedbacks in the two time scales setting in §4.1 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 $\big\{\rho_{(d,0):(d,1)}\big\}_{0 \leqslant d \leqslant D}$ of stochastic kernels within two consecutive slow scale indexes

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

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

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

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

With these notations, we obtain Equation (41b), 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

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

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

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

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

$$
= \inf_{u^{\sharp}_{d} \in \mathbb{U}^{\sharp}_{d}} \int_{\mathbb{W}^{\flat}_{d+1}} \rho_{(d,0):(d,1)}\Big(h^{\sharp}_{d}, \mathrm{d}w^{\flat}_{d+1}\Big) \inf_{u^{\flat}_{d+1} \in \mathbb{U}^{\flat}_{d+1}} \widetilde{\varphi}(h^{\sharp}_{d}, u^{\sharp}_{d}, w^{\flat}_{d+1}, u^{\flat}_{d+1})
$$

This ends the proof.

B Dynamic Programming with Unit Time Blocks

Here, we recover the classical dynamic programming equations when a state reduction exists at each time $t = 0, \ldots, T-1$, with associated dynamics. Following the setting in §2.2.2, we consider a family $\{\rho_{t-1:t}\}_{1 \leqslant t \leqslant T}$ of stochastic kernels as in (3) and a measurable nonnegative numerical cost function j as in (4).

B.1 The General Case of Unit Time Blocks

First, we treat the general criterion case. We assume the existence of a family of measurable state spaces $\{\mathbb{X}_t\}_{t=0,\ldots,T}$ and the existence of a family of measurable mappings $\{\theta_t\}_{t=0,\ldots,T}$ with $\theta_t : \mathbb{H}_t \to \mathbb{X}_t$. We suppose that there exists a family of measurable dynamics $\{f_{t:t+1}\}_{t=0,\ldots,T-1}$ with $f_{t:t+1} : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \to \mathbb{X}_{t+1}$, such that

$$
\theta_{t+1}\big((h_t, u_t, w_{t+1})\big) = f_{t:t+1}\big(\theta_t(h_t), u_t, w_{t+1}\big) , \quad \forall (h_t, u_t, w_{t+1}) \in \mathbb{H}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} . \tag{65}
$$

The following proposition is a immediate application of Theorem 2 and Proposition 1.

Proposition 4 *Suppose that the triplet* $(\{\mathbb{X}_t\}_{t=0,\ldots,T}, \{\theta_t\}_{t=0,\ldots,T}, \{f_{t:t+1}\}_{t=0,\ldots,T-1})$*, which is a state reduction across the consecutive time blocks* $[t, t+1]_{t=0,\ldots,T-1}$ *of the time span, is compatible with the family* $\{\rho_{t-1:t}\}_{t=1,\ldots,T}$ *of stochastic kernels in* (3) *(see Definition 2).*

Suppose that there exists a measurable nonnegative numerical function

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

such that the cost function j *in* (4) *can be factored as*

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

Define the family $\Big\{\widetilde{V}_t\Big\}_{t=0,\ldots,T}$ *of functions by the backward induction*

$$
\widetilde{V}_T(x_T) = \widetilde{j}(x_T) , \quad \forall x_T \in \mathbb{X}_T , \tag{67a}
$$

$$
\widetilde{V}_t(x_t) = \inf_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \widetilde{V}_{t+1}\big(f_{t:t+1}(x_t, u_t, w_{t+1})\big) \widetilde{\rho}_{t:t+1}(x_t, \mathrm{d}w_{t+1}) , \quad \forall x_t \in \mathbb{X}_t , \tag{67b}
$$

for $t = T-1, \ldots, 0$.

Then, the family $\{V_t\}_{t=0,\ldots,T}$ *of value functions defined by the family of optimization problems* (6) *satisfies*

$$
V_t = \widetilde{V}_t \circ \theta_t , \quad t = 0, \ldots, T . \tag{68}
$$

B.2 The Case of Time Additive Cost Functions

A time additive stochastic optimal control problem is a particular form of the stochastic optimization problem presented previously. As in §B.1, we assume the existence of a family of measurable state spaces $\{\mathbb{X}_t\}_{t=0,\dots,T}$, the existence of a family of measurable mappings $\{\theta_t\}_{t=0,\dots,T}$, and the existence of a family of measurable dynamics such that Equation (65) is fulfilled.

We then assume that, for $t = 0, \dots, T-1$, there exist measurable nonnegative numerical functions (*instantaneous cost*)

$$L_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \to [0, +\infty] ,$$

and that there exists a measurable nonnegative numerical function (*final cost*)

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

such that the cost function j in (4) writes

$$j(h_T) = \sum_{t=0}^{T-1} L_t\big(\theta_t(h_t), u_t, w_{t+1}\big) + K\big(\theta_T(h_T)\big) .$$

The following proposition is an immediate consequence of the specific form of the cost function j when applying Proposition 4.

Proposition 5 *Suppose that the triplet* $(\{\mathbb{X}_t\}_{t=0,\dots,T}, \{\theta_t\}_{t=0,\dots,T}, \{f_{t:t+1}\}_{t=0,\dots,T-1})$*, which is a state reduction across the consecutive time blocks* $[t, t+1]_{t=0,\dots,T-1}$ *of the time span, is compatible with the family* $\{\rho_{t-1:t}\}_{t=1,\dots,T}$ *of stochastic kernels in* (3) *(see Definition 2).*

We inductively define the family of functions $\{\widehat{V}_t\}_{t=0,\dots,T}$*, with* $\widehat{V}_t : \mathbb{X}_t \to [0, +\infty]$*, by the relations*

$$\widehat{V}_T(x_T) = K(x_T) , \quad \forall x_T \in \mathbb{X}_T \tag{70a}$$

and, for $t = T-1, \dots, 0$ *and for all* $x_t \in \mathbb{X}_t$*,*

$$\widehat{V}_t(x_t) = \inf_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \Big(L_t(x_t, u_t, w_{t+1}) + \widehat{V}_{t+1}\big(f_{t:t+1}(x_t, u_t, w_{t+1})\big)\Big) \widetilde{\rho}_{t:t+1}(x_t, \mathrm{d}w_{t+1}) . \tag{70b}$$

Then, the family $\{V_t\}_{t=0,\dots,T}$ *of value functions defined by the family of optimization problems* (6) *satisfies*

$$V_t(h_t) = \sum_{s=0}^{t-1} L_s\big(\theta_s(h_s), u_s, w_{s+1}\big) + \widehat{V}_t\big(\theta_t(h_t)\big) , \quad t = 1, \dots, T , \tag{71a}$$

$$V_0(h_0) = \widehat{V}_0\big(\theta_0(h_0)\big) . \tag{71b}$$

C The Case of Optimization with Noise Process

In this section, the *noise* at time t is modeled as a random variable $\mathbf{W}_t$. We suppose given a stochastic process $\{\mathbf{W}_t\}_{t=0,\dots,T}$ called *noise process*. Then, optimization with noise process becomes a special case of the setting in §2.2. Therefore, we can apply the results obtained in Sect. 3.

We moreover assume that, for any $s = 0, \dots, T-1$, the set $\mathbb{U}_s$ in §2.2.1 is a separable complete metric space.

C.1 Optimization with Noise Process

Noise Process and Stochastic Kernels. Let $(\Omega, \mathcal{A})$ be a measurable space. For $t = 0, \dots, T$, the *noise* at time t is modeled as a random variable $\mathbf{W}_t$ defined on Ω and taking values in $\mathbb{W}_t$. Therefore, we suppose given a stochastic process $\{\mathbf{W}_t\}_{t=0,\dots,T}$ called *noise process*. The following assumption is made in the sequel.

Assumption 1 *For any* $1 \leqslant s \leqslant T$*, there exists a regular conditional distribution of the random variable* $\mathbf{W}_s$ *knowing the random process* $\mathbf{W}_{0:s-1}$*, denoted by* $\mathbb{P}_{\mathbf{W}_s}^{\mathbf{W}_{0:s-1}}(w_{0:s-1}, \mathrm{d}w_s)$.

Under Assumption 1, we can introduce the family $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant T}$ of stochastic kernels

$$\rho_{s-1:s} : \mathbb{H}_{s-1} \to \Delta(\mathbb{W}_s) , \quad s = 1, \dots, T , \tag{72a}$$

defined by

$$\rho_{s-1:s}(h_{s-1}, \mathrm{d}w_s) = \mathbb{P}_{\mathbf{W}_s}^{\mathbf{W}_{0:s-1}}\big([h_{s-1}]_{0:s-1}^{\mathbb{W}}, \mathrm{d}w_s\big) , \quad s = 1, \dots, T , \tag{72b}$$

where $[h_{s-1}]_{0:s-1}^{\mathbb{W}} = (w_0, w_1, \dots, w_{s-1})$ is the uncertainty part of the history h_{s-1} (see Equation (49a)).

Then, using Definition 5, the stochastic kernels $\rho^{\gamma}_{r:t} : \mathbb{H}_r \to \Delta(\mathbb{H}_t)$ are defined, for any measurable nonnegative numerical function $\varphi : \mathbb{H}_t \to [0, +\infty]$, by

$$\begin{aligned}\int_{\mathbb{H}_t} \varphi(h'_t)\rho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_t) &= \int_{\mathbb{W}_{r+1:t}} \varphi\Big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t})\Big)\mathbb{P}^{\mathbf{W}_{0:r}}_{\mathbf{W}_{r+1:t}}\big([h_r]^{\mathbb{W}}_{0:r}, \mathrm{d}w_{r+1:t}\big) \ . \\ &= \mathbb{E}\Big[\varphi(\Phi^{\gamma}_{r:t}(h_r, \mathbf{W}_{r+1:t}))\Big]\mathbf{W}_{0:r} = [h_r]^{\mathbb{W}}_{0:r} \ , \end{aligned} \tag{73}$$

where $\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t}) = (h_r, \gamma_r(h_r), w_{r+1}, \gamma_{r+1}(h_r, \gamma_r(h_r), w_{r+1}), w_{r+2}, \cdots, \gamma_{t-1}(h_{t-1}), w_t)$ is the flow induced by the feedback γ (see §A.1).

Adapted Control Processes. Let t be given such that $0 \leqslant t \leqslant T-1$. We introduce

$$\mathcal{A}_{t:t} = \{\emptyset, \Omega\} \ , \ \mathcal{A}_{t:t+1} = \sigma(\mathbf{W}_{t+1}) \ , \ \ldots, \ , \ \mathcal{A}_{t:T-1} = \sigma(\mathbf{W}_{t+1}, \ldots, \mathbf{W}_{T-1}) \ .$$

Let $\mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})$ be the space of $\mathcal{A}$-*adapted control processes* $(\mathbf{U}_t, \ldots, \mathbf{U}_{T-1})$ with values in $\mathbb{U}_{t:T-1}$, that is, such that

$$\sigma(\mathbf{U}_s) \subset \mathcal{A}_{t:s} \ , \ \ s = t, \ldots, T-1 \ .$$

Family of Optimization Problems over Adapted Control Processes. We suppose here that the measurable space $(\Omega, \mathcal{A})$ is equipped with a probability $\mathbb{P}$, so that $(\Omega, \mathcal{A}, \mathbb{P})$ is a probability space. Following the setting given in §2.2.2, we consider a measurable nonnegative numerical cost function j as in Equation (4).

We consider the following family of optimization problems, indexed by $t = 0, \ldots, T-1$ and by $h_t \in \mathbb{H}_t$,

$$\check{V}_t(h_t) = \inf_{(\mathbf{U}_{t:T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \ldots, \mathbf{U}_{T-1}, \mathbf{W}_T)\Big]\mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t} \ . \tag{74}$$

Proposition 6 *Let $t \in \{0, \ldots, T-1\}$ and $h_t \in \mathbb{H}_t$ be given. Problem* (5) *and Problem* (74) *coincide, that is,*

$$\check{V}_t(h_t) = V_t(h_t) \ , \tag{75}$$

where the family of value functions $\{V_t\}_{t=0,\ldots,T}$ is defined by (6).

Proof Let $t \in \{0, \ldots, T-1\}$ and $h_t \in \mathbb{H}_t$ be given. We show that Problem (74) and Problem (5) are in one-to-one correspondence.

- First, for any history feedback $\gamma_{t:T-1} = \{\gamma_s\}_{s=t,\ldots,T-1} \in \Gamma_{t:T-1}$, we define $(\mathbf{U}_{t:T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})$ by

$$(\mathbf{U}_t, \ldots, \mathbf{U}_{T-1}) = \big[\Phi^{\gamma}_{t:T}(h_t, \mathbf{W}_{t+1}, \ldots, \mathbf{W}_T)\big]^{\mathbb{U}}_{t+1:T} \ , \tag{76}$$

  where the flow $\Phi^{\gamma}_{t:T}$ has been defined in (50) and the history control part $[\cdot]^{\mathbb{U}}_{t+1:T}$ in (49b). By the expression (72b) of $\rho_{s:s+1}(h'_s, \mathrm{d}w_{s+1})$ and by Definition 5 of the stochastic kernel $\rho^{\gamma}_{t:T}$, we obtain that

$$\begin{aligned}\mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \ldots, \mathbf{U}_{T-1}, \mathbf{W}_T)\Big]\mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t} &= \mathbb{E}\Big[j(\Phi^{\gamma}_{t:T}(h_t, \mathbf{W}_{t+1}, \ldots, \mathbf{W}_T))\Big]\mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t} \\ &= \int_{\mathbb{H}_T} j(h'_T)\rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T) \ . \end{aligned} \tag{77}$$

  As a consequence

$$\begin{aligned}&\inf_{(\mathbf{U}_{t:T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \ldots, \mathbf{U}_{T-1}, \mathbf{W}_T)\Big]\mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t} \\ &\qquad\qquad \leqslant \inf_{\gamma_{t:T-1} \in \Gamma_{t:T-1}} \int_{\mathbb{H}_T} j(h'_T)\rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T) \ . \end{aligned} \tag{78}$$

- Second, we define a $(t{:}T-1)$-*noise feedback* as a sequence $\lambda = \{\lambda_s\}_{s=t,\ldots,T-1}$ of measurable mappings (the mapping λ_t is constant)

$$\lambda_t \in \mathbb{U}_t \ , \ \ \lambda_s : \mathbb{W}_{t+1:s} \to \mathbb{U}_s \ , \ \ t+1 \leqslant s \leqslant T-1 \ .$$

  We denote by $\Lambda_{t:T-1}$ the set of $(t{:}T-1)$-noise feedbacks. Let $(\mathbf{U}_t, \ldots, \mathbf{U}_{T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})$. As each set $\mathbb{U}_s$ is a separable complete metric space, for $s = t, \ldots, T-1$, we can invoke Doob Theorem (see [5, Chap. 1, p. 18]). Therefore, there exists a $(t{:}T-1)$-noise feedback $\lambda = \{\lambda_s\}_{s=t,\ldots,T-1} \in \Lambda_{t:T-1}$ such that

$$\mathbf{U}_t = \lambda_t \ , \ \ \mathbf{U}_s = \lambda_s(\mathbf{W}_{t+1:s}) \ , \ \ t+1 \leqslant s \leqslant T-1 \ .$$

  Then, we define the history feedback $\gamma_{t:T-1} = \{\gamma_s\}_{s=t,\ldots,T-1} \in \Gamma_{t:T-1}$ by, for any history $h'_r \in \mathbb{H}_r$, $r = t, \ldots, T-1$:

$$\begin{aligned}\gamma_t(h'_t) &= \lambda_t \ , \\ \gamma_{t+1}(h'_{t+1}) &= \lambda_{t+1}\Big(\big[h'_{t+1}\big]^{\mathbb{W}}_{t+1:t+1}\Big) = \lambda_{t+1}(w'_{t+1}) \ , \\ &\vdots \\ \gamma_{T-1}(h'_{T-1}) &= \lambda_{T-1}\Big(\big[h'_{T-1}\big]^{\mathbb{W}}_{t+1:T-1}\Big) = \lambda_{T-1}(w'_{t+1}, \cdots, w'_{T-1}) \ . \end{aligned}$$

By the expression (72b) of $\rho_{s:s+1}(h'_s, \mathrm{d}w_{s+1})$ and by Definition 5 of the stochastic kernel $\rho^{\gamma}_{t:T}$, we obtain that

$$\int_{\mathbb{H}_T} j(h'_T)\rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T) = \mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \ldots, \mathbf{U}_{T-1}, \mathbf{W}_T)\Big|\mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t}\ .$$

As a consequence

$$\inf_{\gamma_{t:T-1}\in\Gamma_{t:T-1}} \int_{\mathbb{H}_T} j(h'_T)\rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T)$$
$$\leqslant \inf_{(\mathbf{U}_t,\ldots,\mathbf{U}_{T-1})\in\mathbb{L}^0(\Omega,\mathcal{A}_{t:T-1},\mathbb{U}_{t:T-1})} \mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \ldots, \mathbf{U}_{T-1}, \mathbf{W}_T)\Big|\mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t}\ . \quad (79)$$

Gathering inequalities (78) and (79) leads to (75). This ends the proof.

The following proposition is an immediate consequence of Theorem 1 and Proposition 6.

Proposition 7 *The family* $\left\{\check{V}_t\right\}_{t=0,\ldots,T}$ *of functions in* (74) *satisfies the backward induction*

$$\check{V}_T(h_T) = j(h_T)\ ,\ \ \forall h_T \in \mathbb{H}_T\ , \quad (80a)$$

and, for $t = T-1, \ldots, 0$,

$$\check{V}_t(h_t) = \inf_{u_t} \int_{\mathbb{W}_{t+1}} \check{V}_{t+1}\big(h_t, u_t, w_{t+1}\big)\mathbb{P}^{\mathbf{W}_{0:t}}_{\mathbf{W}_{t+1}}\big([h_t]^{\mathbb{W}}_{0:t}, \mathrm{d}w_{t+1}\big) \quad (80b)$$
$$= \inf_{u_t}\ \mathbb{E}\big[\check{V}_{t+1}\big(h_t, u_t, \mathbf{W}_{t+1}\big)\big]\mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t}\ ,\ \ \forall h_t \in \mathbb{H}_t\ . \quad (80c)$$

C.2 Two Time-Scales Dynamic Programming

We adopt the notation of §4.1. We suppose given a two time-scales noise process

$$\mathbf{W}_{(0,0):(D+1,0)} = \big(\mathbf{W}_{0,0}, \mathbf{W}_{0,1}, \ldots, \mathbf{W}_{0,M}, \mathbf{W}_{1,0}, \ldots, \mathbf{W}_{D,M}, \mathbf{W}_{D+1,0}\big)\ .$$

For any $d \in \{0, 1, \ldots, D\}$, we introduce the σ-fields

$$\mathcal{A}_{d,0} = \{\emptyset, \Omega\}\ ,\ \ \mathcal{A}_{d,m} = \sigma\big(\mathbf{W}_{(d,1):(d,m)}\big)\ ,\ \ m = 1, \ldots, M\ .$$

The proof of the following proposition is left to the reader.

Proposition 8 *Suppose that there exists a family* $\{\mathbb{X}_d\}_{d=0,\ldots,D+1}$ *of measurable state spaces, with* $\mathbb{X}_0 = \mathbb{W}_{0,0}$, *and a family* $\{f_{d:d+1}\}_{d=0,\ldots,D}$ *of measurable dynamics*

$$f_{d:d+1} : \mathbb{X}_d \times \mathbb{H}_{d:d+1} \to \mathbb{X}_{d+1}\ .$$

Suppose that the slow scale subprocesses $\mathbf{W}_{(d,1):(d+1,0)} = \big(\mathbf{W}_{d,1}, \cdots, \mathbf{W}_{d+1,0}\big)$, $d = 0, \ldots, D$, *are independent (under the probability law* $\mathbb{P}$*).*

For a measurable nonnegative numerical cost function

$$\widetilde{j} : \mathbb{X}_{D+1} \to [0, +\infty]\ ,$$

we define the family $\left\{\widetilde{V}_d\right\}_{d=0,\ldots,D+1}$ *of functions by the backward induction*

$$\widetilde{V}_{D+1}(x_{D+1}) = \widetilde{j}(x_{D+1})\ , \quad (81a)$$
$$\widetilde{V}_d(x_d) = \inf_{\mathbf{U}_{(d,0):(d,M)}\in\mathbb{L}^0(\Omega,\mathcal{A}_{(d,0):(d,M)},\mathbb{U}_{(d,0):(d,M)})} \mathbb{E}\Big[\widetilde{V}_{d+1}\big(f_{d:d+1}(x_d, \mathbf{U}_{d,0}, \mathbf{W}_{d,1}, \cdots, \mathbf{U}_{d,M}, \mathbf{W}_{d+1,0})\big)\Big]\ . \quad (81b)$$

Then, the value functions $\widetilde{V}_d$ *are the solution of the following family of optimization problems, indexed by* $d = 0, \ldots, D$ *and by* $x_d \in \mathbb{X}_d$,

$$\widetilde{V}_d(x_d) = \inf_{U_{(d,0):(D,M)}\in\mathbb{L}^0(\Omega,\mathcal{A}_{(d,0):(D,M)},\mathbb{U}_{(d,0):(D,M)})} \mathbb{E}\big[\widetilde{j}(\mathbf{X}_{D+1})\big]\ , \quad (82a)$$

where, for all $d' = d, \ldots, D$,

$$\mathbf{X}_d = x_d\ ,\ \ \mathbf{X}_{d'+1} = f_{d':d'+1}\big(\mathbf{X}_{d'}, \mathbf{U}_{d',0}, \mathbf{W}_{d',1}, \cdots, \mathbf{U}_{d',M}, \mathbf{W}_{d'+1,0}\big)\ . \quad (82b)$$

C.3 Decision-Hazard-Decision Dynamic Programming

We adopt the notation of §4.2. We suppose given a noise process

$$\mathbf{W}_{0:S} = \big(\mathbf{W}^{\sharp}_0, \mathbf{W}^{\flat}_1, \dots, \mathbf{W}^{\flat}_S\big) \; . \tag{83}$$

For any $s \in \{0, 1, \dots, S-1\}$, we introduce the σ-fields

$$\mathcal{A}_s = \{\varnothing, \Omega\} \; , \; \mathcal{A}_{s'} = \sigma(\mathbf{W}^{\flat}_{s+1:s'}) \; , \; s' = s+1, \dots, S \; . \tag{84}$$

The proof of the following proposition is left to the reader.

Proposition 9 *Suppose that there exists a family $\{\mathbb{X}_s\}_{s=0,\dots,S}$ of measurable state spaces, with $\mathbb{X}_0 = \mathbb{W}^{\sharp}_0$, and a family $\{f_{s:s+1}\}_{s=0,\dots,S-1}$ of measurable dynamics*

$$f_{s:s+1} : \mathbb{X}_s \times \mathbb{U}^{\sharp}_s \times \mathbb{W}^{\flat}_{s+1} \times \mathbb{U}^{\flat}_{s+1} \to \mathbb{X}_{s+1} \; .$$

Suppose that the noise process $\{\mathbf{W}^{\flat}_s\}_{s=0,\dots,S}$ is made of independent random variables (under the probability law $\mathbb{P}$).

For a measurable nonnegative numerical cost function

$$\widetilde{j} : \mathbb{X}_S \to [0, +\infty] \; , \tag{85}$$

we define the family of functions $\big\{\widetilde{V}_s\big\}_{s=0,\dots,S}$ by the backward induction

$$\widetilde{V}_S(x_S) = \widetilde{j}(x_S) \; , \tag{86a}$$

$$\widetilde{V}_s(x_s) = \inf_{u^{\sharp}_s \in \mathbb{U}^{\sharp}_s} \mathbb{E}\Big[\inf_{u^{\flat}_{s+1} \in \mathbb{U}^{\flat}_{s+1}} \widetilde{V}_{s+1}\Big(f_{s':s'+1}\big(x_s, u^{\sharp}_s, \mathbf{W}^{\flat}_{s+1}, u^{\flat}_{s+1}\big)\Big)\Big] \; . \tag{86b}$$

Then, the value functions $\widetilde{V}_s$ in (86) *are the solution of the following family of optimization problems, indexed by $s = 0, \dots, S-1$ and by $x_s \in \mathbb{X}_s$,*

$$\widetilde{V}_s(x_s) = \inf_{\mathbf{U}^{\sharp}_{s:S-1} \in \mathbb{L}^0(\Omega, \mathcal{A}_{s:S-1}, \mathbb{U}^{\sharp}_{s:S-1})} \; \inf_{\mathbf{U}^{\flat}_{s+1:S} \in \mathbb{L}^0(\Omega, \mathcal{A}_{s+1:S}, \mathbb{U}^{\flat}_{s+1:S})} \mathbb{E}\big[\widetilde{j}(\mathbf{X}_S)\big] \; , \tag{87a}$$

where

$$\mathbf{X}_{s'} = x_s \; , \; \mathbf{X}_{s'+1} = f_{s':s'+1}\big(\mathbf{X}_{s'}, \mathbf{U}^{\sharp}_{s'}, \mathbf{W}^{\flat}_{s'+1}, \mathbf{U}^{\flat}_{s'+1}\big) \; , \; \forall s' = s, \dots, S-1 \; . \tag{87b}$$

C.4 Dynamic Programming with Unit Time Blocks

In the setting of optimization with noise process, we now consider the case where a state reduction exists at each time $t = 0, \dots, T-1$. We will use a standard assumption in Dynamic Programming, that is, $\{\mathbf{W}_t\}_{t=0,\dots,T}$ is a white noise process.

C.4.1 The Case of Final Cost Function

We first treat the case of a general criterion, as in §B.1.

Proposition 10 *Suppose that there exists a family $\{\mathbb{X}_t\}_{t=0,\dots,T}$ of measurable state spaces, with $\mathbb{X}_0 = \mathbb{W}_0$, and a family $\{f_{t:t+1}\}_{t=0,\dots,T-1}$ of measurable dynamics*

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

Suppose that the noise process $\{\mathbf{W}_t\}_{t=0,\dots,T}$ is made of independent random variables (under the probability law $\mathbb{P}$).

For a measurable nonnegative numerical cost function

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

we define the family $\big\{\widetilde{V}_t\big\}_{t=0,\dots,T}$ of functions by the backward induction

$$\widetilde{V}_T(x_T) = \widetilde{j}(x_T) \; , \; \forall x_T \in \mathbb{X}_T \; , \tag{88a}$$

$$\widetilde{V}_t(x_t) = \inf_{u_t \in \mathbb{U}_t} \mathbb{E}\big[\widetilde{V}_{t+1}\big(x_t, u_t, \mathbf{W}_{t+1}\big)\big] \; , \; \forall x_t \in \mathbb{X}_t \; , \tag{88b}$$

for $t = T-1, \dots, 0$. Then, the value functions $\widetilde{V}_t$ are the solution of the following family of optimization problems, indexed by $t = 0, \dots, T-1$ and by $x_t \in \mathbb{X}_t$,

$$\widetilde{V}_t(x_t) = \inf_{\mathbf{U}_{t:T-1} \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\big[\widetilde{j}(\mathbf{X}_T)\big] \; , \tag{89a}$$

where

$$\mathbf{X}_s = x_t \; , \; \mathbf{X}_{s+1} = f_{s:s+1}\big(\mathbf{X}_s, \mathbf{U}_s, \mathbf{W}_{s+1}\big) \; , \; \forall s = t, \dots, T-1 \; . \tag{89b}$$

Proof We define a family $\{\theta_t\}_{t=0,\ldots,T}$ of reduction mappings $\theta_t : \mathbb{H}_t \to \mathbb{X}_t$ as in (10a) by induction. First, as $\mathbb{X}_0 = \mathbb{W}_0 = \mathbb{H}_0$ by assumption, we put $\theta_0 = \mathrm{I_d} : \mathbb{H}_0 \to \mathbb{X}_0$. Then, we use (65) to define the mappings $\theta_1, \ldots, \theta_T$. As a consequence, the triplet $(\{\mathbb{X}_t\}_{t=0,\ldots,T}, \{\theta_t\}_{t=0,\ldots,T}, \{f_{t:t+1}\}_{t=0,\ldots,T-1})$ is a state reduction across the consecutive time blocks $[t, t+1]_{t=0,\ldots,T-1}$ of the time span.

Since the noise process $\{\mathbf{W}_t\}_{t=0,\ldots,T}$ is made of independent random variables (under $\mathbb{P}$), the family $\{\rho_{s-1:s}\}_{1\leqslant s\leqslant T}$ of stochastic kernels defined in (72) is given by

$$\rho_{s-1:s} : \mathbb{H}_{s-1} \to \Delta(\mathbb{W}_s) , \quad s = 1, \ldots, T , \tag{90a}$$
$$h_{s-1} \mapsto \mathbb{P}_{\mathbf{W}_s}(\mathrm{d}w_s) . \tag{90b}$$

As a consequence, we have by (11) that the triplet $(\{\mathbb{X}_t\}_{t=0,\ldots,T}, \{\theta_t\}_{t=0,\ldots,T}, \{f_{t:t+1}\}_{t=0,\ldots,T-1})$ is compatible (see Definition 2) with the family $\{\rho_{t-1:t}\}_{t=1,\ldots,T}$ of stochastic kernels in (90). In addition, the reduced stochastic kernels in (11) coincide with the original stochastic kernels in (90).

Define the cost function j as

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

Proposition 4 applies, so that the family $\{V_t\}_{t=0,\ldots,T}$ of value functions defined for the family of optimization problems (5) satisfies

$$V_t = \widetilde{V}_t \circ \theta_t , \quad t = 0, \ldots, T .$$

By means of Proposition 6, we deduce that

$$\check{V}_t(h_t) = \widetilde{V}_t \circ \theta_t(h_t) ,$$

for all $t = 0, \ldots, T$ and for any $h_t \in \mathbb{H}_t$. From the definition (74) of the family of functions $\check{V}_t$, we obtain the expression (89) of functions $\widetilde{V}_t$.

C.4.2 The Case of Time Additive Cost Functions

We make the same assumptions than in §B.2. The proof is left to the reader.

Proposition 11 *Suppose that there exists a family $\{\mathbb{X}_t\}_{t=0,\ldots,T}$ of measurable state spaces, with $\mathbb{X}_0 = \mathbb{W}_0$, and a family $\{f_{t:t+1}\}_{t=0,\ldots,T-1}$ of measurable dynamics*

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

Suppose that the noise process $\{\mathbf{W}_t\}_{t=0,\ldots,T}$ is made of independent random variables (under the probability law $\mathbb{P}$).

We define the family $\left\{\widehat{V}_t\right\}_{t=0,\ldots,T}$ of functions by the backward induction

$$\widehat{V}_T(x_T) = K(x_T) , \quad \forall x_T \in \mathbb{X}_T , \tag{91a}$$

and, for $t = T-1, \ldots, 0$ and for all $x_t \in \mathbb{X}_t$

$$\widehat{V}_t(x_t) = \inf_{u_t \in \mathbb{U}_t} \mathbb{E}\big[L_t(x_t, u_t, \mathbf{W}_{t+1}) + \widehat{V}_{t+1}\big(f_{t:t+1}(x_t, u_t, \mathbf{W}_{t+1})\big)\big] . \tag{91b}$$

Then, the value functions $\widehat{V}_t$ are the solution of the following family of optimization problems, indexed by $t = 0, \ldots, T-1$ and by $x_t \in \mathbb{X}_t$,

$$\widehat{V}_t(x_t) = \inf_{(\mathbf{U}_t,\ldots,\mathbf{U}_{T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\Big[\sum_{s=t}^{T-1} L_s\big(\mathbf{X}_s, \mathbf{U}_s, \mathbf{W}_{s+1}\big) + K\big(\mathbf{X}_T\big)\Big] , \tag{92a}$$

where

$$\mathbf{X}_s = x_t , \quad \mathbf{X}_{s+1} = f_{s:s+1}\big(\mathbf{X}_s, \mathbf{U}_s, \mathbf{W}_{s+1}\big) , \quad \forall s = t, \ldots, T-1 . \tag{92b}$$

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. P. Carpentier, J.-P. Chancelier, G. Cohen, and M. De Lara. *Stochastic Multi-Stage Optimization. At the Crossroads between Discrete Time Stochastic Control and Stochastic Programming.* Springer-Verlag, Berlin, 2015.
5. C. Dellacherie and P.A. Meyer. *Probabilités et potentiel.* Hermann, Paris, 1975.
6. I. V. Evstigneev. Measurable selection and dynamic programming. *Mathematics of Operations Research*, 1(3):267–272, 1976.
7. M. Loève. *Probability Theory I.* Springer-Verlag, New York, fourth edition, 1977.

8. M. L. Puterman. *Markov Decision Processes.* Wiley, New York, 1994.
9. 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.
10. 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.
11. H. S. Witsenhausen. A standard form for sequential stochastic control. *Mathematical Systems Theory*, 7(1):5–11, 1973.
12. H. S. Witsenhausen. On policy independence of conditional expectations. *Information and Control*, 28(1):65–75, 1975.