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

Pierre Carpentier · Jean-Philippe Chancelier · Michel De Lara · Tristan 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. Their large scale nature makes decomposition methods appealing. We provide a method to decompose multistage stochastic optimization problems by time blocks. Our framework covers both stochastic programming and stochastic dynamic programming. We formulate multistage stochastic optimization problems over a so-called history space, with solutions being history feedbacks. We prove a general dynamic programming equation, with value functions defined on the history space. Then, we consider the question of reducing the history using a compressed "state" variable. This reduction can be done by time blocks, that is, at stages that are not necessarily all the original unit stages. We prove a reduced dynamic programming equation. Then, we apply the reduction method by time blocks to several classes of optimization problems, especially two time-scales stochastic optimization problems and a novel class consisting of decision hazard decision models. Finally, we consider the case of optimization with noise process.

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

MSC: 90C06,90C39,93E20.

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

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.

On the one hand, stochastic programming deals with an underlying random process taking a finite number of values, called scenarios [9]. Solutions are indexed by a scenario tree, the size of which explodes with the number of stages, hence generally few. However, to overcome this obstacle, stochastic programming takes advantage of scenario decomposition methods (Progressive Hedging [8]). On the other hand, stochastic control deals with a state model driven by a white noise, that is, the noise is made of a sequence of independent random variables. Under such assumptions, stochastic dynamic programming is able to handle many stages, as it offers reduction of the search for a solution among state feedbacks (instead of functions of the past noise) [1,6].

In a word, dynamic programming is good at handling multiple stages — but at the price of assuming that noise 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. We formulate multistage stochastic optimization problems over a so-called history space, with solutions being history feedbacks. We prove a general dynamic programming equation, with value functions defined on the history space. In Sect. 3, we consider the question of reducing the history using a compressed "state" variable. This reduction can be done by time blocks, that is, at stages that are not necessarily all the original unit stages. We prove a reduced dynamic programming equation. In Sect. 4, we apply the reduction method by time blocks to several classes of optimization problems, especially two time-scales stochastic optimization problems and a novel class consisting of decision hazard decision models. Finally, we consider the case of optimization with noise process; we show in Sect. 5 that it is a special case of the setting in Sect. 2.

2 Stochastic Dynamic Programming with History Feedbacks

Consider the time span $\{0, 1, 2 \ldots, T-1, T\}$, with *horizon* $T \in \mathbb{N}^*$. 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

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

We present the mathematical formalism to handle such type of problems.

2.1 Histories, Feedbacks and Flows

We first define in §2.1.1 the basic and the composite spaces that we will need to formulate multistage stochastic optimization problems. Then, in §2.1.2, we introduce a class of solutions called history feedbacks; we also define flows.

2.1.1 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

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

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*:

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

We introduce the notations

$$\mathbb{W}_{r:t} = \prod_{s=r}^{t} \mathbb{W}_s \ , \ \ 0 \leqslant r \leqslant t \leqslant T \tag{2b}$$

$$\mathbb{U}_{r:t} = \prod_{s=r}^{t} \mathbb{U}_s \ , \ \ 0 \leqslant r \leqslant t \leqslant T-1 \tag{2c}$$

$$\mathbb{H}_{r:t} = \prod_{s=r-1}^{t-1} (\mathbb{U}_s \times \mathbb{W}_{s+1}) = \mathbb{U}_{r-1} \times \mathbb{W}_r \times \cdots \times \mathbb{U}_{t-1} \times \mathbb{W}_t \ , \ \ 1 \leqslant r \leqslant t \leqslant T \ . \tag{2d}$$

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*

$$[h_t]^{\mathbb{W}}_{r:s} = (w_r, \dots, w_s) = w_{r:s} \in \mathbb{W}_{r:s} \ , \ \ 0 \leqslant r \leqslant s \leqslant t \ , \tag{2e}$$

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

$$[h_t]^{\mathbb{U}}_{r:s} = (u_{r-1}, \dots, u_{s-1}) = u_{r-1:s-1} \in \mathbb{U}_{r-1:s-1} \ , \ \ 1 \leqslant r \leqslant s \leqslant t \ , \tag{2f}$$

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

$$[h_t]_{r:s} = (u_{r-1}, w_r, \dots, u_{s-1}, w_s) = h_{r:s} \in \mathbb{H}_{r:s} \ , \ \ 1 \leqslant r \leqslant s \leqslant t \ , \tag{2g}$$

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

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

2.1.2 Feedbacks and Flows

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

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

$$\gamma_s : \mathbb{H}_s \to \mathbb{U}_s \ . \tag{3}$$

We call $\Gamma_{r:t}$ the set of $(r\!:\!t)$-history feedbacks.

Flows. When $0 \leqslant r < t \leqslant T$, for a $(r\!:\!t-1)$-history feedback $\gamma = \{\gamma_s\}_{s=r,\dots,t-1} \in \Gamma_{r:t-1}$, we define the *flow* $\Phi^{\gamma}_{r:t}$ by

$$\Phi^{\gamma}_{r:t} : \mathbb{H}_r \times \mathbb{W}_{r+1:t} \to \mathbb{H}_t \tag{4a}$$

$$(h_r, w_{r+1:t}) \mapsto (h_r, \gamma_r(h_r), w_{r+1}, \gamma_{r+1}(h_r, \gamma_r(h_r), w_{r+1}), w_{r+2}, \cdots, u_{t-1}, w_t) \ , \tag{4b}$$

that is,

$$\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t}) = (h_r, u_r, w_{r+1}, u_{r+1}, w_{r+2}, \dots, u_{t-1}, w_t) \ , \tag{4c}$$

$$\text{with } \ h_s = (h_r, u_r, w_{r+1}, \dots, u_{s-1}, w_s) \ , \ \ r < s \leqslant t \ , \tag{4d}$$

$$\text{and } \ u_s = \gamma_s(h_s) \ , \ \ r < s \leqslant t-1 \ . \tag{4e}$$

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

$$\Phi^{\gamma}_{r:r} : \mathbb{H}_r \to \mathbb{H}_r \ , \ \ h_r \mapsto h_r \ . \tag{4f}$$

With this convention, the expression $\Phi^{\gamma}_{r:t}$ makes sense when $0 \leqslant r \leqslant t \leqslant T$ for a $(r : t-1)$-history feedback $\gamma = \{\gamma_s\}_{s=r,\dots,t-1} \in \Gamma_{r:t-1}$ (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 $\{\gamma_s\}_{s=r,\dots,t-1} \in \Gamma_{r:t-1}$. An immediate consequence of this definition are the *flow properties*:

$$\Phi^{\gamma}_{r:t+1}(h_r, w_{r+1:t+1}) = \Big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t}), \gamma_t\big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t})\big), w_{t+1}\Big) \ , \ \ 0 \leqslant r \leqslant t \leqslant T-1 \ , \tag{5a}$$

$$\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t}) = \Phi^{\gamma}_{r+1:t}\big((h_r, \gamma_r(h_r), w_{r+1}), w_{r+2:t}\big) \ , \ \ 0 \leqslant r < t \leqslant T \ . \tag{5b}$$

2.2 Optimization with Stochastic Kernels

In §2.2.1, given a history feedback and a sequence of stochastic kernels from partial histories to uncertainties, we will build a new sequence of stochastic kernels, but from partial histories to sequences of uncertainties. With this construction, we introduce a family of optimization problems with stochastic kernels in §2.2.2. Then, in §2.2.3, 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) [5].

2.2.1 Stochastic Kernels

Definition of stochastic kernels. 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 (on $\mathbb{Y}$ knowing $\mathbb{X}$) 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 $\mathcal{Y}$, with the property that the function $x \in \mathbb{X} \mapsto \int_Y \rho(x, \mathrm{d}y)$ is measurable for any $Y \in \mathcal{Y}$.

Building new stochastic kernels from history feedbacks and stochastic kernels.

Definition 1 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

  $$\rho_{s-1:s} : \mathbb{H}_{s-1} \to \Delta(\mathbb{W}_s) \ , \ \ s = r+1, \dots, t \ , \tag{6}$$

  we define a stochastic kernel

  $$\rho^{\gamma}_{r:t} : \mathbb{H}_r \to \Delta(\mathbb{H}_t) \tag{7a}$$

  by, for any $\varphi : \mathbb{H}_t \to [0, +\infty]$, measurable nonnegative numerical function,[1]

  $$\int_{\mathbb{H}_t} \varphi(h'_r, h'_{r+1:t}) \rho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_t) = \int_{\mathbb{W}_{r+1:t}} \varphi\big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t})\big) \prod_{s=r+1}^{t} \rho_{s-1:s}\big(\Phi^{\gamma}_{r:s-1}(h_r, w_{r+1:s-1}), \mathrm{d}w_s\big) \ . \tag{7b}$$

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

  $$\rho^{\gamma}_{r:r} : \mathbb{H}_r \to \Delta(\mathbb{H}_r) \ , \ \ \rho^{\gamma}_{r:r}(h_r, \mathrm{d}h'_r) = \delta_{h_r}(\mathrm{d}h'_r) \ . \tag{7c}$$

[1] We could also consider any $\varphi : \mathbb{H}_t \to \mathbb{R}$, 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.

We detail Equation (7b) in Appendix A. The stochastic kernels $\rho^{\gamma}_{r:t}$ on $\mathbb{H}_t$, given by (7), are of the form

$$\rho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_t) = \rho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_r \mathrm{d}h'_{r+1:t}) = \delta_{h_r}(\mathrm{d}h'_r) \otimes \varrho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_{r+1:t}) \,, \tag{8}$$

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.

Proposition 1 *Following Definition 1, 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 $s < t$,*

$$\rho^{\gamma}_{s:t}(h_s, \mathrm{d}h'_t) = \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}\big(h_s, \mathrm{d}w_{s+1}\big) \rho^{\gamma}_{s+1:t}\Big(\big(h_s, \gamma_s(h_s), w_{s+1}\big), \mathrm{d}h'_t\Big) \,. \tag{9}$$

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

$$\begin{aligned} &\int_{\mathbb{H}_t} \varphi(h'_s, h'_{s+1:t}) \rho^{\gamma}_{s:t}(h_s, \mathrm{d}h'_t) \qquad (10a) \\ &\quad = \int_{\mathbb{W}_{s+1:t}} \varphi\big(\Phi^{\gamma}_{s:t}(h_s, w_{s+1:t})\big) \prod_{s'=s+1}^{t} \rho_{s'-1:s'}\big(\Phi^{\gamma}_{s:s'-1}(h_s, w_{s+1:s'-1}), \mathrm{d}w_{s'}\big) \end{aligned}$$

by the definition (7b) 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 (4f) of the flow $\Phi^{\gamma}_{s:s}$

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

by the flow property (5b)

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

by Fubini Theorem [5, 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 (7b) of $\rho^{\gamma}_{s+1:t}$

$$= \int_{\mathbb{H}_t} \varphi\big((h'_s, \gamma_s(h'_s), w'_{s+1}), h'_{s+2:t}\big) \int_{\mathbb{W}_{s+1}} \rho_{s:s+1}\big(h_s, \mathrm{d}w_{s+1}\big) \rho^{\gamma}_{s+1:t}\big((h_s, \gamma_s(h_s), w_{s+1}), \mathrm{d}h'_t\big) \tag{10b}$$

by Fubini Theorem and by definition (7b) of $\rho^{\gamma}_{s:t}$. As the two expressions (10a) and (10b) are equal for any $\varphi : \mathbb{H}_t \to [0, +\infty]$, we deduce the flow property (9). This ends the proof.

2.2.2 Family of Optimization Problems with Stochastic Kernels

To build a family of optimization problems over the time span $\{0, \dots, T-1\}$, we need two ingredients:

– a family $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant T}$ of stochastic kernels

$$\rho_{s-1:s} : \mathbb{H}_{s-1} \to \Delta(\mathbb{W}_s) \ , \quad s = 1, \dots, T \ , \tag{11}$$

– a numerical function, playing the role of a cost to be minimized,

$$j : \mathbb{H}_T \to [0, +\infty] \ , \tag{12}$$

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

We define, for any $\{\gamma_s\}_{s=t,\dots,T-1} \in \Gamma_{t:T-1}$,

$$V_t^\gamma(h_t) = \int_{\mathbb{H}_T} j(h'_T) \rho_{t:T}^\gamma(h_t, \mathrm{d}h'_T) \ , \quad \forall h_t \in \mathbb{H}_t \ . \tag{13}$$

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

$$\inf_{\gamma_{t:T-1} \in \Gamma_{t:T-1}} \int_{\mathbb{H}_T} j(h'_T) \rho_{t:T}^\gamma(h_t, \mathrm{d}h'_T) \ . \tag{14}$$

For all $t = 0, \dots, T-1$, we define the minimum value of Problem (14) by

$$V_t(h_t) = \inf_{\gamma_{t:T-1} \in \Gamma_{t:T-1}} \int_{\mathbb{H}_T} j(h'_T) \rho_{t:T}^\gamma(h_t, \mathrm{d}h'_T) \tag{15a}$$

$$= \inf_{\gamma_{t:T-1} \in \Gamma_{t:T-1}} V_t^\gamma(h_t) \ , \quad \forall h_t \in \mathbb{H}_t \ , \tag{15b}$$

and we also define

$$V_T(h_T) = j(h_T) \ , \quad \forall h_T \in \mathbb{H}_T \ . \tag{15c}$$

The last notation is consistent with (14) by the definition (7c) of the stochastic kernel $\rho_{T:T}^\gamma$. The numerical function $V_t : \mathbb{H}_t \to [0, +\infty]$ is called *value function*.

2.2.3 Resolution by Stochastic Dynamic Programming

Now, we show that the value functions in (15) are *Bellman functions*, in that they are solution of the Bellman or Dynamic Programming equation.

The following two assumptions will be made throughout the whole paper.

Assumption 1 (Measurable function) *For all $t = 0, \dots, T-1$ and for all nonnegative measurable numerical function $\varphi : \mathbb{H}_{t+1} \to [0, +\infty]$, the numerical function*

$$h_t \mapsto \inf_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \varphi(h_t, u_t, w_{t+1}) \rho_{t:t+1}(h_t, dw_{t+1}) \tag{16}$$

is measurable[3] from $(\mathbb{H}_t, \mathcal{H}_t)$ to $[0, +\infty]$.

Assumption 2 (Measurable selection) *For all $t = 0, \dots, T-1$, there exists a* measurable selection,[4] *that is, a measurable mapping*

$$\gamma_t^\star : (\mathbb{H}_t, \mathcal{H}_t) \to (\mathbb{U}_t, \mathcal{U}_t) \tag{17a}$$

such that

$$\gamma_t^\star(h_t) \in \arg\min_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} V_{t+1}(h_t, u_t, w_{t+1}) \rho_{t:t+1}(h_t, \mathrm{d}w_{t+1}) \ , \tag{17b}$$

where the numerical function V_{t+1} is given by (15).

[2] See Footnote 1. When $j(h_T) = +\infty$, this materializes joint constraints between uncertainties and controls.

[3] This is a delicate issue, treated in [2].

[4] See [2] and [7] for a precise definition of a measurable selection.

Bellman Operators. For $t = 0, \dots, T$, let $\mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t)$ be the space of nonnegative measurable numerical functions over $\mathbb{H}_t$.

Definition 2 For $t = 0, \dots, T-1$, we define the *Bellman operator*

$$\mathcal{B}_{t+1:t} : \mathbb{L}^0_+(\mathbb{H}_{t+1}, \mathcal{H}_{t+1}) \to \mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t) \tag{18a}$$

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

$$\big(\mathcal{B}_{t+1:t}\varphi\big)(h_t) = \inf_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \varphi(h_t, u_t, w_{t+1}) \rho_{t:t+1}(h_t, dw_{t+1}) \; . \tag{18b}$$

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 and, by Assumption 1, we know that $\mathcal{B}_{t+1:t}\varphi$ is a measurable numerical function, hence belongs to $\mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t)$.

Bellman equation and optimal history feedbacks.

Theorem 1 *The value functions in (15) satisfy the* Bellman equation*, or* (Stochastic) Dynamic Programming equation

$$V_T = j \; , \tag{19a}$$

$$V_t = \mathcal{B}_{t+1:t} V_{t+1} \;\; \textit{for} \;\; t = T-1, \dots, 0 \; . \tag{19b}$$

Moreover, a solution to any Problem (14) — that is, whatever the index $t = 0, \dots, T-1$ and the parameter $h_t \in \mathbb{H}_t$ — is any history feedback $\gamma^\star = \{\gamma^\star_s\}_{s=t,\dots,T-1}$ defined by the collection of mappings $\gamma^\star_s$ in (17).

Notice that, although Problem (14) is parameterized by $h_t \in \mathbb{H}_t$, the optimal history feedback $\gamma^\star = \{\gamma^\star_s\}_{s=t,\dots,T-1}$ is not.

Proof From the definition (13), we have for any $\{\gamma_s\}_{s=t,\dots,T-1} \in \Gamma_{t:T-1}$,

$$V^\gamma_t(h_t) = \int_{\mathbb{H}_T} j(h'_T) \rho^\gamma_{t:T}(h_t, \mathrm{d}h'_T)$$

that only depends on $\{\gamma_s\}_{s=t,\dots,T-1}$

$$= \int_{\mathbb{H}_T} j(h'_T) \int_{\mathbb{W}_{t+1}} \rho_{t:t+1}\big(h_t, \mathrm{d}w_{t+1}\big) \rho^\gamma_{t+1:T}\Big(\big(h_t, \gamma_t(h_t), w_{t+1}\big), \mathrm{d}h'_T\Big)$$

by the flow property (9) for stochastic kernels

$$= \int_{\mathbb{W}_{t+1}} \rho_{t:t+1}\big(h_t, \mathrm{d}w_{t+1}\big) \int_{\mathbb{H}_T} j(h'_T) \rho^\gamma_{t+1:T}\Big(\big(h_t, \gamma_t(h_t), w_{t+1}\big), \mathrm{d}h'_T\Big)$$

by Fubini Theorem [5, p.137]

$$= \int_{\mathbb{W}_{t+1}} \rho_{t:t+1}\big(h_t, \mathrm{d}w_{t+1}\big) V^\gamma_{t+1}\big(h_t, \gamma_t(h_t), w_{t+1}\big)$$

by definition (13) of V^γ_{t+1}

$$\geqslant \int_{\mathbb{W}_{t+1}} \rho_{t:t+1}\big(h_t, \mathrm{d}w_{t+1}\big) V_{t+1}\big(h_t, \gamma_t(h_t), w_{t+1}\big)$$

by definition (15) of the value function V_{t+1}, and as V^γ_{t+1} only depends on $\{\gamma_s\}_{s=t+1,\dots,T-1}$. We deduce that

$$V_t(h_t) \geqslant \inf_{u_t} \int_{\mathbb{W}_{t+1}} \rho_{t:t+1}\big(h_t, \mathrm{d}w_{t+1}\big) V_{t+1}\big(h_t, u_t, w_{t+1}\big) \; . \tag{20a}$$

The inequality (20a) above is in fact an equality, as seen by using any measurable history feedback $\gamma^\star = \{\gamma^\star_s\}_{s=t,\dots,T-1}$ defined by the collection of functions $\gamma^\star_s$ in (17).

This ends the proof.

3 State Reduction by Time Blocks

In this section, we consider the question of reducing the history using a compressed "state" variable. Such a variable may be not available at any time $t \in \{0, \dots, T\}$, but at some specified instants. We have to note that the history h_t is itself a canonical state variable in our framework, the associated dynamics being $h_{t+1} = \big(h_t, u_t, w_{t+1}\big)$.

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:

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

Let $\{\rho_{s-1:s}\}_{r+1 \leqslant s \leqslant t}$ be a family of stochastic kernels

$$\rho_{s-1:s} : \mathbb{H}_{s-1} \to \Delta(\mathbb{W}_s) \ , \ \ s = r+1, \dots, t \ . \tag{21}$$

We define the *Bellman operator across* $(t\!:\!r)$ by

$$\mathcal{B}_{t:r} : \mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t) \to \mathbb{L}^0_+(\mathbb{H}_r, \mathcal{H}_r) \ , \ \ \mathcal{B}_{t:r} = \mathcal{B}_{t:t-1} \circ \cdots \circ \mathcal{B}_{r+1:r} \ , \tag{22}$$

where the one time step operators $\mathcal{B}_{s:s-1}$, for $r+1 \leqslant s \leqslant t$ have been defined in (18).

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

$$\theta_r : \mathbb{H}_r \to \mathbb{X}_r \ , \ \ \theta_t : \mathbb{H}_t \to \mathbb{X}_t \ , \tag{23}$$

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

$$f_{r:t} : \mathbb{X}_r \times \mathbb{H}_{r+1:t} \to \mathbb{X}_t \ . \tag{24}$$

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

$$\theta_t\big((h_r, h_{r+1:t})\big) = f_{r:t}\big(\theta_r(h_r), h_{r+1:t}\big) \ , \ \ \forall h_t \in \mathbb{H}_t \ . \tag{25}$$

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 defined in (21) if

– there exists a *reduced stochastic kernel*

$$\widetilde{\rho}_{r:r+1} : \mathbb{X}_r \to \Delta(\mathbb{W}_{r+1}) \ , \tag{26a}$$

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

$$\rho_{r:r+1}(h_r, \mathrm{d}w_{r+1}) = \widetilde{\rho}_{r:r+1}\big(\theta_r(h_r), \mathrm{d}w_{r+1}\big) \ , \ \ \forall h_r \in \mathbb{H}_r \ , \tag{26b}$$

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

$$\widetilde{\rho}_{s-1:s} : \mathbb{X}_r \times \mathbb{H}_{r+1:s-1} \to \Delta(\mathbb{W}_s) \tag{26c}$$

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

$$\rho_{s-1:s}\big((h_r, h_{r+1:s-1}), \mathrm{d}w_s\big) = \widetilde{\rho}_{s-1:s}\Big(\big(\theta_r(h_r), h_{r+1:s-1}\big), \mathrm{d}w_s\Big) \ , \ \ \forall h_{s-1} \in \mathbb{H}_{s-1} \ . \tag{26d}$$

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

The following theorem is the key ingredient to formulate Dynamic Programming equations with a reduced state.

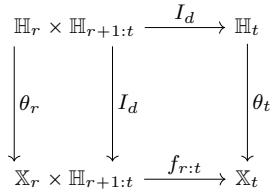

Fig. 1 Commutative diagram in case of state reduction $(\theta_r, \theta_t, f_{r:t})$

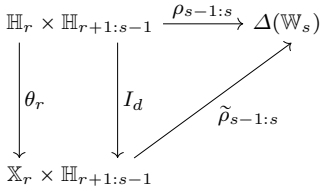

Fig. 2 Commutative diagram in case of state reduction $(\theta_r, \theta_t, f_{r:t})$ compatible with the family $\{\rho_{s-1:s}\}_{r+1 \leqslant s \leqslant t}$

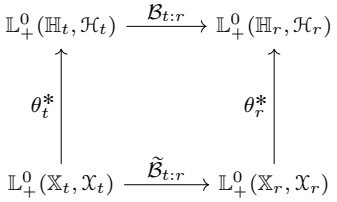

Fig. 3 Commutative diagram for Bellman operators in case of state reduction $(\theta_r, \theta_t, f_{r:t})$ compatible with the family $\{\rho_{s-1:s}\}_{r+1 \leqslant s \leqslant t}$

Theorem 2 *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 (21) (see Definition 3). Then, there exists a* reduced Bellman operator across $(t\!:\!r)$

$$\widetilde{\mathcal{B}}_{t:r} : \mathbb{L}^0_+(\mathbb{X}_t, \mathcal{X}_t) \to \mathbb{L}^0_+(\mathbb{X}_r, \mathcal{X}_r) \,, \tag{27}$$

such that, for any measurable nonnegative numerical function $\widetilde{\varphi}_t : \mathbb{X}_t \to [0, +\infty]$, we have that

$$\big(\widetilde{\mathcal{B}}_{t:r}\widetilde{\varphi}_t\big) \circ \theta_r = \mathcal{B}_{t:r}(\widetilde{\varphi}_t \circ \theta_t) \,. \tag{28}$$

Denoting by $\theta_t^* : \mathbb{L}^0_+(\mathbb{X}_t, \mathcal{X}_t) \to \mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t)$ the operator such that

$$\theta_t^*(\widetilde{\varphi}_t) = \widetilde{\varphi}_t \circ \theta_t \,, \tag{29}$$

the relation (28) rewrites:

$$\theta_r^*\big(\widetilde{\mathcal{B}}_{t:r}\widetilde{\varphi}_t\big) = \mathcal{B}_{t:r}\big(\theta_t^*(\widetilde{\varphi}_t)\big) \,. \tag{30}$$

Equivalently, Theorem 2 states that the diagram in Figure 3 is commutative.

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

$$\varphi_t = \widetilde{\varphi}_t \circ \theta_t \,. \tag{31}$$

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 (22)) to the measurable nonnegative numerical function φ_t:

$$\varphi_r = \mathcal{B}_{t:r}\varphi_t = \mathcal{B}_{r+1:r} \circ \cdots \circ \mathcal{B}_{t:t-1}\varphi_t \,. \tag{32}$$

We will show that there exists a measurable nonnegative numerical function

$$\widetilde{\varphi}_r : \mathbb{X}_r \to [0, +\infty] \tag{33}$$

such that

$$\varphi_r = \widetilde{\varphi}_r \circ \theta_r \ . \tag{34}$$

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 (34).

– For $s = t$, we have, by (31) and by (25), 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,

$$\varphi_{s+1}(h_{s+1}) = \overline{\varphi}_{s+1}(\theta_r(h_r), h_{r+1:s+1}) \ . \tag{35}$$

Then,

$$\begin{aligned}
\varphi_s(h_s) &= \big(\mathcal{B}_{s+1:s}\varphi_{s+1}\big)(h_s) && (\text{ by (32)}) \\
&= \inf_{u_s \in \mathbb{U}_s} \int_{\mathbb{W}_{s+1}} \varphi_{s+1}\big((h_s, u_s, w_{s+1})\big)\rho_{s:s+1}(h_s, \mathrm{d}w_{s+1})
\end{aligned}$$

by definition (18) 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 (35)

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

by compatibility (26) 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.

Corollary 1 *Under the assumptions of Theorem 2, the expression of the reduced Bellman operator $\widetilde{\mathcal{B}}_{t:r}$ in (27) is available: 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*

$$\begin{aligned}
\big(\widetilde{\mathcal{B}}_{t:r}\widetilde{\varphi}_t\big)(x_r) = \inf_{u_r \in \mathbb{U}_r} \int_{\mathbb{W}_{r+1}} \widetilde{\rho}_{r:r+1}(x_r, \mathrm{d}w_{r+1}) \inf_{u_{r+1} \in \mathbb{U}_{r+1}} \int_{\mathbb{W}_{r+2}} \widetilde{\rho}_{r+1:r+2}(x_r, u_r, w_{r+1}, \mathrm{d}w_{r+2}) \\
\ldots \inf_{u_{t-1} \in \mathbb{U}_{t-1}} \int_{\mathbb{W}_t} \widetilde{\varphi}_t\big(f_{r:t}(x_r, u_r, w_{r+1}, \ldots, u_{t-1}, w_t, u_{t-1}, w_t)\big) \\
\widetilde{\rho}_{t-1:t}(x_r, u_r, w_{r+1}, \ldots, u_{t-2}, w_{t-1}, \mathrm{d}w_t) \ .
\end{aligned} \tag{37}$$

Proof Equation (37) follows from the induction developed in the proof of Theorem 2.

The optimal feedbacks yielded by (37) are mappings $\widetilde{\gamma}_s : \mathbb{X}_r \times \mathbb{H}_{r+1:s} \to \mathbb{U}_s$, for $s = r, \ldots, t-1$. These are no longer history feedbacks, by partially truncated history feedbacks where history h_r has been replaced by state x_r.

3.2 State Reduction on Multiple Consecutive Time Blocks

Theorem 2 can easily be extended to the case of multiple consecutive time blocks $[t_i, t_{i+1}]$, $i = 0, \ldots, N-1$ where

$$0 \leqslant t_0 < t_1 < \cdots < t_N \leqslant T \, . \tag{38}$$

Let $\{\rho_{s-1:s}\}_{t_0+1 \leqslant s \leqslant t_N}$ be a family of stochastic kernels

$$\rho_{s-1:s} : \mathbb{H}_{s-1} \to \Delta(\mathbb{W}_s) \, , \;\; s = t_0 + 1, \ldots, t_N \, . \tag{39}$$

Definition 4 Let $\{\mathbb{X}_{t_i}\}_{i=0,\ldots,N}$ be a family of 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 $\left\{f_{t_i:t_{i+1}}\right\}_{i=0,\ldots,N-1}$ be a family of 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 $(\{\mathbb{X}_{t_i}\}_{i=0,\ldots,N}, \{\theta_{t_i}\}_{i=0,\ldots,N}, \left\{f_{t_i:t_{i+1}}\right\}_{i=0,\ldots,N-1})$ 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 triplet is said to be *compatible* with the family $\{\rho_{s-1:s}\}_{t_0+1 \leqslant s \leqslant t_N}$ of stochastic kernels given in (39) 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 $i = 0, \ldots, N-1$.

Theorem 3 *Suppose that a state reduction* $(\{\mathbb{X}_{t_i}\}_{i=0,\ldots,N}, \{\theta_{t_i}\}_{i=0,\ldots,N}, \left\{f_{t_i:t_{i+1}}\right\}_{i=0,\ldots,N-1})$ *exists across the consecutive time blocks* $[t_i, t_{i+1}]$, $i = 0, \ldots, N-1$, *that is compatible with the family* $\{\rho_{s-1:s}\}_{t_0+1 \leqslant s \leqslant t_N}$ *of stochastic kernels given in (39).*

Then, there exists a family of reduced Bellman operators *across the consecutive* $(t_{i+1} : t_i)$, $i = 0, \ldots, N-1$,

$$\widetilde{\mathcal{B}}_{t_{i+1}:t_i} : \mathbb{L}^0_+(\mathbb{X}_{t_{i+1}}, \mathcal{X}_{t_{i+1}}) \to \mathbb{L}^0_+(\mathbb{X}_{t_i}, \mathcal{X}_{t_i}) \, , \;\; i = 0, \ldots, N-1 \, , \tag{40}$$

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*

$$\big(\widetilde{\mathcal{B}}_{t_{i+1}:t_i} \widetilde{\varphi}_{t_{i+1}}\big) \circ \theta_{t_i} = \mathcal{B}_{t_{i+1}:t_i}(\widetilde{\varphi}_{t_{i+1}} \circ \theta_{t_{i+1}}) \, . \tag{41}$$

Proof This is an immediate consequence of multiple applications of Theorem 2.

4 Stochastic Dynamic Programming by Time Blocks

We apply the reduction by time blocks to several classes of optimization problems: dynamic programming with unit time blocks in §4.1, two time-scales dynamic programming in §4.2; decision hazard decision dynamic programming in §4.3.

4.1 Dynamic Programming with Unit Time Blocks

We now consider the case where a state reduction exists at each time $t = 0, \ldots, T-1$, with associated dynamics. We recover the classical Dynamic Programming equations.

Following the setting in §2.2, we consider a family $\{\rho_{t-1:t}\}_{1 \leqslant t \leqslant T}$ of stochastic kernels as in (11) and a measurable nonnegative numerical cost function j as in (12).

4.1.1 The General Case of Unit Time Blocks

First, we treat the general criterion case. We assume the existence of a family of state spaces $\{\mathbb{X}_t\}_{t=0,\ldots,T}$ and the existence of a family of 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 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

$$\theta_{t+1}\big((h_t, u_t, w_{t+1})\big) = f_{t:t+1}\big(\theta_t(h_t), u_t, w_{t+1}\big) \, , \;\; \forall (h_t, u_t, w_{t+1}) \in \mathbb{H}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \, . \tag{42}$$

The following Proposition is a direct application of Theorem 3.

Proposition 2 *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 (11) (see Definition 4).*

Suppose that there exists a measurable nonnegative numerical function

$$\widetilde{j} : \mathbb{X}_T \to [0, +\infty] \,, \tag{43a}$$

such that the cost function j *in (12) can be factored as*

$$j = \widetilde{j} \circ \theta_T \,. \tag{43b}$$

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

$$\widetilde{V}_T(x_T) = \widetilde{j}(x_T) \,, \quad \forall x_T \in \mathbb{X}_T \,, \tag{44a}$$

$$\widetilde{V}_t(x_t) = \inf_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \widetilde{V}_{t+1}\big(f_{t:t+1}(x_t, u_t, w_{t+1})\big) \widetilde{\rho}_{t:t+1}(x_t, \mathrm{d}w_{t+1}) \,, \quad \forall x_t \in \mathbb{X}_t \,, \tag{44b}$$

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

Then, the family $\{V_t\}_{t=0,\dots,T}$ *of value functions defined by the family of optimization problems (15) satisfies*

$$V_t = \widetilde{V}_t \circ \theta_t \,, \quad t = 0, \dots, T \,. \tag{45}$$

Proof The existence of the family $\{\widetilde{\mathcal{B}}_{t+1:t}\}_{t=0,\dots,T-1}$ of reduced Bellman operators, as well as the relation (45), are a direct consequence of Theorem 3. The specific expression (44b) is induced by Corollary 1 in case of a unit time block.

The expression of the optimal state feedbacks is given by the next Corollary.

Corollary 2 *Suppose that, for* $t = 0, \dots, T-1$*, there exist measurable selections*

$$\widetilde{\gamma}_t^\star : (\mathbb{X}_t, \mathcal{X}_t) \to (\mathbb{U}_t, \mathcal{U}_t) \tag{46a}$$

such that, for all $x_t \in \mathbb{X}_t$*,*

$$\widetilde{\gamma}_t^\star(x_t) \in \operatorname*{arg\,min}_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \widetilde{V}_{t+1}\big(f_{t:t+1}(x_t, u_t, w_{t+1})\big) \widetilde{\rho}_{t:t+1}(x_t, \mathrm{d}w_{t+1}) \,, \tag{46b}$$

where the family $\{\widetilde{V}_t\}_{t=0,\dots,T}$ *of functions is given by (44). Then, the family of history feedbacks* $\{\gamma_s^\star\}_{s=t,\dots,T-1}$ *given by*

$$\gamma_s^\star = \widetilde{\gamma}_s^\star \circ \theta_s \,, \quad s = t, \dots, T-1 \tag{47}$$

is a solution to any Problem (14), that is, whatever the index $t = 0, \dots, T-1$ *and the parameter* $h_t \in \mathbb{H}_t$*.*

Proof The proof is an immediate consequence of Theorem 1 and Theorem 2.

4.1.2 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 §4.1.1, we assume the existence of a family of state spaces $\{\mathbb{X}_t\}_{t=0,\dots,T}$, the existence of a family of mappings $\{\theta_t\}_{t=0,\dots,T}$, and the existence of a family of dynamics such that Equation (42) is fulfilled.

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

$$L_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \to [0, +\infty] \,, \tag{48a}$$

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

$$K : \mathbb{X}_T \to [0, +\infty] \,, \tag{48b}$$

such that the cost function j in (12) writes

$$j(h_T) = \sum_{t=0}^{T-1} L_t\big(\theta_t(h_t), u_t, w_{t+1}\big) + K\big(\theta_T(h_T)\big) \,. \tag{48c}$$

Proposition 3 *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 (11) (see Definition 4).*

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*

$$\widehat{V}_T(x_T) = K(x_T) \ , \ \ \forall x_T \in \mathbb{X}_T \tag{49a}$$

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

$$\widehat{V}_t(x_t) = \min_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \Big(L_t(x_t, u_t, w_{t+1}) + \widehat{V}_{t+1}\big(f_{t:t+1}(x_t, u_t, w_{t+1})\big)\Big) \widetilde{\rho}_{t:t+1}(x_t, \mathrm{d}w_{t+1}) \ . \tag{49b}$$

Then, the family $\{V_t\}_{t=0,\dots,T}$ *of value functions defined by the family of optimization problems (15) satisfies*

$$V_t(h_t) = \sum_{s=0}^{t-1} L_s\big(\theta_s(h_s), u_s, w_{s+1}\big) + \widehat{V}_t\big(\theta_t(h_t)\big) \ , \ \ t = 1, \dots, T \ , \tag{50a}$$

$$V_0(h_0) = \widehat{V}_0\big(\theta_0(h_0)\big) \ . \tag{50b}$$

Proof The proof is an immediate consequence of Theorem 2, of the specific form of the cost function j and of the fact that the additive term $\sum_{s=0}^{t-1} L_s\big(\theta_s(h_s), u_s, w_{s+1}\big)$ only depends on h_t.

Corollary 3 *Suppose that, for* $t = 0, \dots, T-1$*, there exists measurable selections*

$$\widehat{\gamma}_t^\star : (\mathbb{X}_t, \mathfrak{X}_t) \to (\mathbb{U}_t, \mathfrak{U}_t) \ , \tag{51a}$$

such that, for all $x_t \in \mathbb{X}_t$*,*

$$\widehat{\gamma}_t^\star(x_t) \in \operatorname*{arg\,min}_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \Big(L_t(x_t, u_t, w_{t+1}) + \widehat{V}_{t+1}\big(F_t(x_t, u_t, w_{t+1})\big)\Big) \widetilde{\rho}_{t:t+1}(x_t, dw_{t+1}) \ , \tag{51b}$$

where the family $\{\widehat{V}_t\}_{t=0,\dots,T}$*, of functions is given by (49). Then, the family of history feedbacks* $\{\gamma_s^\star\}_{s=t,\dots,T-1}$ *given by*

$$\gamma_s^\star = \widehat{\gamma}_s^\star \circ \theta_s \ , \ \ s = t, \dots, T-1 \tag{52}$$

is a solution to any Problem (14), that is, whatever the index $t = 0, \dots, T-1$ *and the parameter* $h_t \in \mathbb{H}_t$*.*

4.2 Two Time-Scales Dynamic Programming

Let $(D, M) \in \mathbb{N}^{*2}$. We put

$$\mathbb{T} = \{0, \dots, D\} \times \{0, \dots, M\} \cup \{(D+1, 0)\} \ . \tag{53}$$

We can think of the index $d \in \{0, \dots, D+1\}$ as an index of days (slow scale), and $m \in \{0, \dots, M\}$ as an index of minutes (fast scale).

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

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

We present the mathematical formalism to handle such type of problems.

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

$$(0,0) < (0,1) < \cdots < (d,M) < (d+1,0) < \cdots < (D,M-1) < (D,M) < (D+1,0)\ . \tag{54a}$$

This set is in one to one correspondence with the *time span* $\{0,\dots,T\}$, where

$$T = (D+1)\times(M+1)+1 \tag{54b}$$

by the *lexicographic mapping* τ

$$\tau : \{0,\dots,T\} \to \mathbb{T} \tag{54c}$$
$$t \mapsto \tau(t) = (d,m)\ . \tag{54d}$$

By abuse of notation, we will simply denote by $(d,m) \in \mathbb{T}$ the element of $\{0,\dots,T\}$ given by $\tau^{-1}(d,m) = d\times(M+1)+m$

$$\mathbb{T} \ni (d,m) \equiv \tau^{-1}(d,m) \in \{0,\dots,T\}\ . \tag{54e}$$

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

History spaces. With the identification (54e), for all $(d,m) \in \mathbb{T}$, we define the *history space* $\mathbb{H}_{(d,m)}$ equipped with the *history field* $\mathcal{H}_{(d,m)}$ as in (1). For all $d \in \{0,\dots,D+1\}$, we define the *slow scale history* h_d element of the *slow scale history space* $\mathbb{H}_d$ equipped with the *slow scale history field* $\mathcal{H}_d$ as in (1) by:

$$h_d = h_{(d,0)} \in \mathbb{H}_d = \mathbb{H}_{(d,0)}\ ,\ \ \mathcal{H}_d = \mathcal{H}_{(d,0)}\ . \tag{55a}$$

For all $d \in \{0,\dots,D\}$, we define the *slow scale partial history space* $\mathbb{H}_{d:d+1}$ equipped with the *slow scale partial history field* $\mathcal{H}_{d:d+1}$ as in (2d) by:

$$\mathbb{H}_{d:d+1} = \mathbb{H}_{(d,1):(d+1,0)} = \mathbb{U}_{d,0}\times\mathbb{W}_{d,1}\times\cdots\times\mathbb{U}_{d,M-1}\times\mathbb{W}_{d,M}\times\mathbb{U}_{d,M}\times\mathbb{W}_{d+1,0}\ , \tag{55b}$$
$$\mathcal{H}_{d:d+1} = \mathcal{H}_{(d,1):(d+1,0)} = \mathcal{U}_{d,0}\otimes\mathcal{W}_{d,1}\otimes\cdots\otimes\mathcal{U}_{d,M-1}\otimes\mathcal{W}_{d,M}\otimes\mathcal{U}_{d,M}\otimes\mathcal{W}_{d+1,0}\ . \tag{55c}$$

Stochastic kernels. Because of the jump from one day to the next, we introduce two families of stochastic kernels[5]:

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

$$\rho_{(d,M):(d+1,0)} : \mathbb{H}_{(d,M)} \to \Delta(\mathbb{W}_{d+1,0})\ ,\ \ d = 0,\dots,D\ , \tag{56a}$$

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

$$\rho_{(d,m-1):(d,m)} : \mathbb{H}_{(d,m-1)} \to \Delta(\mathbb{W}_{d,m})\ ,\ \ d = 0,\dots,D\ ,\ \ m = 1,\dots,M\ . \tag{56b}$$

History feedbacks. Following the notation in §2.1.2, a history feedback at index $(d,m) \in \mathbb{T}$ is a measurable mapping

$$\gamma_{(d,m)} : \mathbb{H}_{(d,m)} \to \mathbb{U}_{(d,m)}\ . \tag{57}$$

For $(d,m) \leqslant (d',m')$, we denote by $\Gamma_{(d,m):(d',m')}$ the set of $((d,m):(d',m'))$-history feedbacks.

[5] These families are defined over the time span $\{0,\dots,T\} \equiv \mathbb{T}$ by the identification (54e) in such a way that the notation is consistent with the notation (11).

Slow scale value functions. We suppose given a nonnegative numerical function

$$j : \mathbb{H}_{D+1} \to [0, +\infty] \,, \tag{58}$$

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 $\rho^{\gamma}_{(d,0):(D+1,0)}(h_d, \mathrm{d}h'_{D+1}) : \mathbb{H}_d \to \Delta(\mathbb{H}_{D+1})$ thanks to Definition 1, and we are then to define the *slow scale value functions*

$$V_d(h_d) = \min_{\gamma \in \Gamma_{(d,0):(D,M)}} \int_{\mathbb{H}_{D+1}} j(h'_{D+1}) \rho^{\gamma}_{(d,0):(D+1,0)}(h_d, \mathrm{d}h'_{D+1}) \,, \quad \forall h_d \in \mathbb{H}_d \,, \tag{59}$$

and $V_{D+1} = j$.

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

$$\mathcal{B}_{d+1:d} : \mathbb{L}^0_+(\mathbb{H}_{d+1}, \mathcal{H}_{d+1}) \to \mathbb{L}^0_+(\mathbb{H}_d, \mathcal{H}_d) \,, \quad d = 0, \ldots, D \,, \tag{60a}$$

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

$$\begin{aligned} \big(\mathcal{B}_{d+1:d}\varphi\big)(h_d) = \inf_{u_{d,0} \in \mathbb{U}_{d,0}} \int_{\mathbb{W}_{d,1}} & \rho_{(d,0):(d,1)}(h_d, \mathrm{d}w_{d,1}) \ldots \\ & \inf_{u_{d,M-1} \in \mathbb{U}_{d,M-1}} \int_{\mathbb{W}_{d,M}} \rho_{(d,M-1):(d,M)}(h_d, u_{d,0}, w_{d,1}, \cdots, w_{d,M-1}, \mathrm{d}w_{d,M}) \\ & \inf_{u_{d,M} \in \mathbb{U}_{d,M}} \int_{\mathbb{W}_{d+1,0}} \varphi\big(h_d, u_{d,0}, w_{d,1}, \cdots, u_{d,M-1}, w_{d,M}, u_{d,M}, w_{d+1,0}\big) \\ & \qquad \rho_{(d,M):(d+1,0)}(h_d, u_{d,0}, w_{d,1}, \cdots, w_{d,M}, \mathrm{d}w_{d+1,0}) \,. \end{aligned} \tag{60b}$$

Proposition 4 *The family* $\{V_d\}_{d=0,\ldots,D+1}$ *of slow scale value functions (59) satisfies*

$$V_{D+1} = j \,, \tag{61a}$$

$$V_d = \mathcal{B}_{d+1:d} V_{d+1} \,, \quad \textit{for } d = D, \ldots, 0 \,. \tag{61b}$$

Proof With the identification (54e), a general two-time scales stochastic dynamic optimization problem as (59) takes the usual form (14). Since we have

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

we can apply Theorem 1 repeatedly, which leads to the result.

Definition 5 (Compatible slow scale reduction) Let $\{\mathbb{X}_d\}_{d=0,\ldots,D+1}$ be a family of state spaces, $\{\theta_d\}_{d=0,\ldots,D+1}$ be family of measurable reduction mappings such that

$$\theta_d : \mathbb{H}_d \to \mathbb{X}_d \,, \tag{62a}$$

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

$$f_{d:d+1} : \mathbb{X}_d \times \mathbb{H}_{d:d+1} \to \mathbb{X}_{d+1} \,. \tag{62b}$$

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

$$\theta_{d+1}\big((h_d, h_{d:d+1})\big) = f_{d:d+1}\big(\theta_d(h_d), h_{d:d+1}\big) \,, \quad \forall (h_d, h_{d:d+1}) \in \mathbb{H}_{d+1} \,. \tag{62c}$$

The slow scale state reduction $\big(\{\mathbb{X}_d\}_{d=0,\ldots,D+1}, \{\theta_d\}_{d=0,\ldots,D+1}, \{f_{d:d+1}\}_{d=0,\ldots,D}\big)$ 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 (56a)–(56b) if for any $d = 0, \ldots, D$, we have that

– there exists a *reduced stochastic kernel*

$$\widetilde{\rho}_{(d,M):(d+1,0)} : \mathbb{X}_d \times \mathbb{H}_{(d,0):(d,M)} \to \Delta(\mathbb{W}_{d+1,0}) \,, \tag{63a}$$

such that the stochastic kernel $\rho_{(d,M):(d+1,0)}$ in (56a) can be factored as

$$\rho_{(d,M):(d+1,0)}(h_{d,M}, \mathrm{d}w_{d+1,0}) = \widetilde{\rho}_{(d,M):(d+1,0)}\big(\theta_d(h_d), h_{(d,0):(d,M)}, \mathrm{d}w_{d+1,0}\big) \,, \quad \forall h_{d,M} \in \mathbb{H}_{(d,M)} \,, \tag{63b}$$

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

$$\widetilde{\rho}_{(d,m-1):(d,m)} : \mathbb{X}_d \times \mathbb{H}_{(d,0):(d,m-1)} \to \Delta(\mathbb{W}_{d,m}) \,, \tag{63c}$$

such that the stochastic kernel $\rho_{(d,m-1):(d,m)}$ in (56b) can be factored as

$$\rho_{(d,m-1):(d,m)}(h_{d,m-1}, \mathrm{d}w_{d,m}) = \widetilde{\rho}_{(d,m-1):(d,m)}\big(\theta_d(h_d), h_{(d,0):(d,m-1)}, \mathrm{d}w_{d,m}\big) \,, \quad \forall h_{d,m-1} \in \mathbb{H}_{(d,m-1)} \,. \tag{63d}$$

Theorem 4 *Assume that there exists a slow scale state reduction* $\big(\{\mathbb{X}_d\}_{d=0,\ldots,D+1}, \{\theta_d\}_{d=0,\ldots,D+1}, \{f_{d:d+1}\}_{d=0,\ldots,D}\big)$ *and that there exists a* reduced criterion

$$\widetilde{j} : \mathbb{X}_{D+1} \to [0, +\infty] \,, \tag{64a}$$

such that the cost function j in (58) can be factored as

$$j = \widetilde{j} \circ \theta_{D+1} \,. \tag{64b}$$

Using the reduced stochastic kernels of Definition 5, we define a family of slow scale reduced Bellman operators across $(d+1:d)$

$$\widetilde{\mathcal{B}}_{d+1:d} : \mathbb{L}^0_+(\mathbb{X}_{d+1}, \mathcal{X}_{d+1}) \to \mathbb{L}^0_+(\mathbb{X}_d, \mathcal{X}_d) \,, \;\; d = 0, \ldots, D \,, \tag{65a}$$

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

$$\begin{aligned}\big(\widetilde{\mathcal{B}}_{d+1:d}\widetilde{\varphi}\big)(x_d) = \inf_{u_{d,0} \in \mathbb{U}_{d,0}} \int_{\mathbb{W}_{d,1}} \widetilde{\rho}_{(d,0):(d,1)}(x_d, \mathrm{d}w_{d,1}) \ldots \\ \inf_{u_{d,M-1} \in \mathbb{U}_{d,M-1}} \int_{\mathbb{W}_{d,M}} \widetilde{\rho}_{(d,M-1):(d,M)}(x_d, u_{d,0}, w_{d,1}, \cdots, w_{d,M-1}, \mathrm{d}w_{d,M}) \\ \inf_{u_{d,M} \in \mathbb{U}_{d,M}} \int_{\mathbb{W}_{d+1,0}} \widetilde{\varphi}\big(\widetilde{f}_{d:d+1}(x_d, u_{d,0}, w_{d,1}, \cdots, u_{d,M-1}, w_{d,M}, u_{d,M}, w_{d+1,0})\big) \\ \widetilde{\rho}_{(d,M):(d+1,0)}(x_d, u_{d,0}, w_{d,1}, \cdots, w_{d,M}, \mathrm{d}w_{d+1,0}) \,.\end{aligned} \tag{65b}$$

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

$$\widetilde{V}_{D+1} = \widetilde{j} \,, \tag{66a}$$

$$\widetilde{V}_d = \widetilde{\mathcal{B}}_{d+1:d}\widetilde{V}_{d+1} \,, \quad \textit{for } d = D, \ldots, 0 \,. \tag{66b}$$

Then, the family $\{V_d\}_{d=0,\ldots,D+1}$ of slow scale value functions (59) satisfies

$$V_d = \widetilde{V}_d \circ \theta_d \,, \;\; d = 0, \ldots, D \,. \tag{66c}$$

Proof The triplet $(\{\mathbb{X}_d\}_{d=0,\ldots,D+1}, \{\theta_d\}_{d=0,\ldots,D+1}, \{f_{d:d+1}\}_{d=0,\ldots,D})$ is a state reduction accross the time blocks $[(d,0),(d+1,0)]$, which is compatible with the family $\big\{\rho_{(d,0):(d+1,0)}\big\}_{0 \leqslant d \leqslant D}$ of stochastic kernels. Hence, we can apply Theorem 3, which leads to the expressions (66c). The expression (60) of the reduced Bellman operators is a consequence of Corollary 1.

4.3 Decision Hazard Decision Dynamic Programming

We consider stochastic optimization problems where, during the time interval between two time steps, the decision-maker takes two decisions. As outlined at the beginning of Sect. 2, 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. This gives the following sequence of events:

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

Let $S \in \mathbb{N}^*$. For each time $s = 0, 1, 2 \dots, S-1$, 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 $s = 1, 2 \dots, S$, 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 $s = 1, 2 \dots, S$, 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 $\mathcal{W}_0^\sharp$.

Decision Hazard Decision history spaces and fields. We define,
for $s = 0, 1, 2 \dots, S$, the *head history space*

$$\mathbb{H}_s^\sharp = \mathbb{W}_0^\sharp \times \prod_{s'=0}^{s-1} \left(\mathbb{U}_{s'}^\sharp \times \mathbb{W}_{s'+1}^\flat \times \mathbb{U}_{s'+1}^\flat \right) , \tag{67a}$$

for $s = 0, 1, 2 \dots, S$, the *head history field*

$$\mathcal{H}_s^\sharp = \mathcal{W}_0^\sharp \otimes \bigotimes_{s'=0}^{s-1} \left(\mathcal{U}_{s'}^\sharp \otimes \mathcal{W}_{s'+1}^\flat \otimes \mathcal{U}_{s'+1}^\flat \right) , \tag{67b}$$

for $s = 1, 2 \dots, S$, the *tail history space*

$$\mathbb{H}_s^\flat = \mathbb{H}_{s-1}^\sharp \times \mathbb{U}_{s-1}^\sharp \times \mathbb{W}_s^\flat \ , \tag{67c}$$

for $s = 1, 2 \dots, S$, the *tail history field*

$$\mathcal{H}_s^\flat = \mathcal{H}_{s-1}^\sharp \otimes \mathcal{U}_{s-1}^\sharp \otimes \mathcal{W}_s^\flat \ . \tag{67d}$$

Decision Hazard Decision history feedbacks. For all $s = 0, \dots, S-1$, a *head history feedback* at time s is a measurable mapping

$$\gamma_s^\sharp : \left(\mathbb{H}_s^\sharp, \mathcal{H}_s^\sharp \right) \to \left(\mathbb{U}_s^\sharp, \mathcal{U}_s^\sharp \right) . \tag{68a}$$

We call $\Gamma_s^\sharp$ the *set of head history feedbacks at time s*. In addition, for $0 \leqslant s \leqslant S-1$, we define

$$\Gamma_{s:S}^\sharp = \Gamma_s^\sharp \times \cdots \times \Gamma_S^\sharp \ . \tag{68b}$$

For all $s = 1, 2 \dots, S$, a *tail history feedback* at time s is a measurable mapping

$$\gamma_s^\flat : \left(\mathbb{H}_s^\flat, \mathcal{H}_s^\flat \right) \to \left(\mathbb{U}_s^\flat, \mathcal{U}_s^\flat \right) . \tag{68c}$$

We call $\Gamma_s^\flat$ the *set of tail history feedbacks at time s*. In addition, for $1 \leqslant s \leqslant S$, we define

$$\Gamma_{s:S}^\flat = \Gamma_s^\flat \times \cdots \times \Gamma_S^\flat \ . \tag{68d}$$

Decision Hazard Decision stochastic kernels. For $s = 1, 2 \dots, S$, we define a *DHD stochastic kernel between time $s-1$ and s* as a measurable mapping

$$\rho_{s-1:s} : \left(\mathbb{H}_{s-1}^\sharp, \mathcal{H}_{s-1}^\sharp \right) \to \Delta(\mathbb{W}_s^\flat) \ , \quad s = 1, \dots, S \ . \tag{69}$$

Let $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant S}$ be a family of DHD stochastic kernels.

Decision Hazard Decision value functions. We consider a nonnegative numerical function

$$j : \mathbb{H}_S^\sharp \to [0, +\infty] \,, \tag{70}$$

supposed to be measurable with respect to the σ-field $\mathcal{H}_S^\sharp$ in (67b).

We define *DHD value functions* by, for all $s = 0, \ldots, S$,

$$V_s(h_s^\sharp) = \min_{\gamma^\sharp \in \Gamma_{s:S-1}^\sharp, \gamma^\flat \in \Gamma_{s+1:S}^\sharp} \int_{\mathbb{H}_S^\sharp} j(h'_S) \rho_{s:S}^{\gamma^\sharp, \gamma^\flat}(h_s^\sharp, \mathrm{d}h'_S) \,, \quad \forall h_s^\sharp \in \mathbb{H}_s^\sharp \,, \tag{71}$$

where $\rho_{s:S}^{\gamma^\sharp, \gamma^\flat}$ has to be understood as $\rho_{s:S}^{\gamma}$ as in (7a) with

$$\gamma_s(h_s^\sharp) = \gamma_s^\sharp(h_s^\sharp) \,, \quad \forall h_s^\sharp \in \mathbb{H}_s^\sharp \,, \tag{72a}$$

$$\gamma_{s'}(h_{s'}^\flat) = \Big(\gamma_{s'}^\flat(h_{s'}^\flat), \gamma_{s'}^\sharp\big(h_{s'}^\flat, \gamma_{s'}^\flat(h_{s'}^\flat)\big)\Big) \,, \quad \forall s' = s+1, \ldots, S-1 \,, \quad \forall h_{s'}^\flat \in \mathbb{H}_{s'}^\flat \,, \tag{72b}$$

$$\gamma_S(h_S^\flat) = \gamma_S^\flat(h_S^\flat) \,, \quad \forall h_S^\flat \in \mathbb{H}_S^\flat \,. \tag{72c}$$

Theorem 5 *For $s = 0, \ldots, S-1$, we define the* DHD Bellman operator

$$\mathcal{B}_{s+1:s} : \mathbb{L}_+^0(\mathbb{H}_{s+1}^\sharp, \mathcal{H}_{s+1}^\sharp) \to \mathbb{L}_+^0(\mathbb{H}_s^\sharp, \mathcal{H}_s^\sharp) \tag{73a}$$

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

$$\big(\mathcal{B}_{s+1:s}\varphi\big)(h_s^\sharp) = \inf_{u_s^\sharp \in \mathbb{U}_s^\sharp} \int_{\mathbb{W}_{s+1}^\flat} \inf_{u_{s+1}^\flat \in \mathbb{U}_{s+1}^\flat} \varphi(h_s^\sharp, u_s^\sharp, w_{s+1}^\flat, u_{s+1}^\flat) \rho_{s:s+1}(h_s^\sharp, dw_{s+1}^\flat) \,. \tag{73b}$$

Then the value functions (71) satisfy

$$V_S = j \,, \tag{73c}$$

$$V_s = \mathcal{B}_{s+1:s} V_{s+1}, \forall s = 0, \ldots, S-1 \,. \tag{73d}$$

Proof We will show that the proof follows from Theorem 4. Indeed, we will now show that the setting in §4.3 is a particular kind of two time scales problem as seen in §4.2. 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 $\{\varnothing, \mathbb{W}_s^\sharp\}$, for each time $s = 1, 2 \ldots, S$. Now, we obtain the following sequence of events:

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

which coincides with a two time scales problem:

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

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 §4.2 are in one to one correspondence with the Decision Hazard Decision history spaces and fields in (67a)–(67c) as follows:

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

$$\mathbb{H}_{d,0} = \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \big(\mathbb{U}_{d',0} \times \mathbb{W}_{d'+1,1} \times \mathbb{U}_{d'+1,1} \times \mathbb{W}_{d'+1,0}\big) \tag{74a}$$

$$= \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \big(\mathbb{U}_{d'}^\sharp \times \mathbb{W}_{d'+1}^\flat \times \mathbb{U}_{d'+1}^\flat \times \mathbb{W}_{d'+1}^\sharp\big) \tag{74b}$$

$$\equiv \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \big(\mathbb{U}_{d'}^\sharp \times \mathbb{W}_{d'+1}^\flat \times \mathbb{U}_{d'+1}^\flat\big) = \mathbb{H}_d^\sharp \,, \tag{74c}$$

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

$$\mathcal{H}_{d,0} = \mathcal{W}_0^\sharp \otimes \bigotimes_{d'=0}^{d-1} \big(\mathcal{U}_{d'}^\sharp \otimes \mathcal{W}_{d'+1}^\flat \otimes \mathcal{U}_{d'+1}^\flat \otimes \mathcal{W}_{d'+1}^\sharp\big) \,, \tag{74d}$$

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

$$\mathbb{H}_{d,1} = \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \big(\mathbb{U}_{d',0} \times \mathbb{W}_{d'+1,1} \times \mathbb{U}_{d'+1,1} \times \mathbb{W}_{d'+1,0}\big) \times \mathbb{U}_{d,0} \times \mathbb{W}_{d+1,1} \tag{74e}$$

$$= \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \big(\mathbb{U}_{d'}^\sharp \times \mathbb{W}_{d'+1}^\flat \times \mathbb{U}_{d'+1}^\flat \times \mathbb{W}_{d'+1}^\sharp\big) \times \mathbb{U}_d^\sharp \times \mathbb{W}_{d+1}^\flat \tag{74f}$$

$$\equiv \mathbb{W}_0^\sharp \times \prod_{d'=0}^{d-1} \big(\mathbb{U}_{d'}^\sharp \times \mathbb{W}_{d'+1}^\flat \times \mathbb{U}_{d'+1}^\flat\big) \times \mathbb{U}_d^\sharp \times \mathbb{W}_{d+1}^\flat = \mathbb{H}_{d+1}^\flat \,, \tag{74g}$$

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

$$\mathcal{H}_{d,1} = \mathcal{W}_0^\sharp \otimes \bigotimes_{d'=0}^{d-1} \big(\mathcal{U}_{d'}^\sharp \otimes \mathcal{W}_{d'+1}^\flat \otimes \mathcal{U}_{d'+1}^\flat \otimes \mathcal{W}_{d'+1}^\sharp\big) \otimes \mathcal{U}_d^\sharp \otimes \mathcal{W}_{d+1}^\flat \,. \tag{74h}$$

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 (70) (Decision Hazard Decision setting) corresponds to $j \circ \big[\,\cdot\,\big]^\sharp$ in the two time scales setting in §4.2. The feedbacks in the two time scales setting in §4.2 are in one to one correspondence with the same elements (72) in the Decision Hazard Decision setting, namely

$$\gamma_{d,0} = \gamma_d^\sharp \circ \big[\,\cdot\,\big]^\sharp \,, \quad \gamma_{d,1} = \gamma_{d+1}^\flat \circ \big[\,\cdot\,\big]^\sharp \,. \tag{75}$$

Now we define two famillies 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

$$\rho_{(d,0):(d,1)} : \mathbb{H}_{d,0} \to \Delta(\mathbb{W}_{d,1}) \,, \tag{76a}$$

$$h_{d,0} \mapsto \rho_{d:d+1} \circ \big[\,\cdot\,\big]^\sharp \,. \tag{76b}$$

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

$$\rho_{(d,1):(d+1,0)} : \mathbb{H}_{d,1} \to \Delta(\mathbb{W}_{d+1,0}) \,, \tag{77a}$$

$$h_{d,1} \mapsto \delta_{\overline{w}_{d+1}^\sharp}(\cdot) \,, \tag{77b}$$

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

With these notations, we can apply Theorem 4 to obtain equation (73b), where only one integral appears because of the Dirac stochastic kernels in (77). 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 $\widetilde{\varphi} : \mathbb{H}_{d+1}^\sharp \to [0, +\infty]$ such that $\varphi = \widetilde{\varphi} \circ \big[\ \cdot\ \big]^\sharp$, we obtain that

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

by the Dirac probability in (77)

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

This ends the proof.

Definition 6 (Decision Hazard Decision compatible state reduction) Let $\{\mathbb{X}_s\}_{s=0,\dots,S}$ be a family of state spaces, $\{\theta_s\}_{s=0,\dots,S}$ be family of measurable reduction mappings such that

$$\theta_s : \mathbb{H}_s^\sharp \to \mathbb{X}_s\ , \tag{78a}$$

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

$$f_{s:s+1} : \mathbb{X}_s \times \mathbb{U}_s^\sharp \times \mathbb{W}_{s+1} \times \mathbb{U}_{s+1}^\flat \to \mathbb{X}_{s+1}\ . \tag{78b}$$

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 *DHD state reduction* if, for all $s = 0, \dots, S-1$, we have that

$$\begin{aligned}\theta_{s+1}\big((h_s, u_s^\sharp, w_{s+1}, u_{s+1}^\flat)\big) = f_{s:s+1}\big(\theta_s(h_s), u_s^\sharp, w_{s+1}, u_{s+1}^\flat\big)\ , \\ \forall (h_s, u_s^\sharp, w_{s+1}, u_{s+1}^\flat) \in \mathbb{H}_s^\sharp \times \mathbb{U}_s^\sharp \times \mathbb{W}_{s+1} \times \mathbb{U}_{s+1}^\flat\ .\end{aligned} \tag{78c}$$

The DHD state reduction is said to be *compatible with the family* $\{\rho_{s:s+1}\}_{0\leqslant s\leqslant S-1}$ *of DHD stochastic kernels* in (69) if there exists a family $\{\widetilde{\rho}_{s:s+1}\}_{0\leqslant s\leqslant S-1}$ of *reduced DHD stochastic kernels*

$$\widetilde{\rho}_{s:s+1} : \mathbb{X}_s \to \Delta(\mathbb{W}_{s+1})\ , \tag{79a}$$

such that, for each $s = 0, \dots, S-1$, the stochastic kernel $\rho_{s:s+1}$ in (69) can be factored as

$$\rho_{s:s+1}(h_s^\sharp, \mathrm{d}w_{s+1}) = \widetilde{\rho}_{s:s+1}\big(\theta_s(h_s^\sharp), \mathrm{d}w_{s+1}\big)\ ,\ \ \forall h_s^\sharp \in \mathbb{H}_s^\sharp\ . \tag{79b}$$

Theorem 6 *Assume that there exists a slow scale 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

$$\widetilde{j} : \mathbb{X}_S \to [0, +\infty]\ , \tag{80a}$$

such that the cost function j in (70) can be factored as

$$j = \widetilde{j} \circ \theta_S\ . \tag{80b}$$

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

$$\widetilde{\mathcal{B}}_{s+1:s} : \mathbb{L}^0_+(\mathbb{X}_{s+1}, \mathcal{X}_{s+1}) \to \mathbb{L}^0_+(\mathbb{X}_s, \mathcal{X}_s) \ , \ \ s = 1, \ldots, S-1 \ , \tag{81a}$$

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

$$(\widetilde{\mathcal{B}}_{s+1:s}\widetilde{\varphi})(x_s) = \inf_{u_s^\sharp \in \mathbb{U}_s^\sharp} \int_{\mathbb{W}_{s+1}} \inf_{u_{s+1}^\flat \in \mathbb{U}_{s+1}^\flat} \widetilde{\varphi}(f_{s:s+1}(x_s, u_s^\sharp, w_{s+1}, u_{s+1}^\flat))\widetilde{\rho}_{s:s+1}(x_s, dw_{s+1}) \ . \tag{81b}$$

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

$$\widetilde{V}_S = \widetilde{j} \tag{82a}$$

$$\widetilde{V}_s = \widetilde{\mathcal{B}}_{s+1:s}\widetilde{V}_{s+1} \quad \textit{for } s = S-1, \ldots, 0 \ . \tag{82b}$$

Then, the value functions V_s *defined by (71) satisfy*

$$V_s = \widetilde{V}_s \circ \theta_s \ , \ \ s = 0, \ldots, S \ . \tag{83}$$

Proof See proof of Theorem 5 and apply Theorem 4.

5 The Case of Optimization with Noise Process

In this Section, we suppose the that, for any $s = 0, \ldots, T-1$, the set $\mathbb{U}_s$ is a separable complete metric space. Optimization with noise process now becomes a special case of the setting in Sect. 2, as we will show in §5.1. Therefore, we can apply the results obtained in Sect. 3 and in Sect. 4.

5.1 Optimization with Noise Process

Noise Process. Let $(\Omega, \mathcal{A})$ be a measurable space. For $t = 0, \ldots, T$, 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 $\{\mathbf{W}_t\}_{t=0,\ldots,T}$ called *noise process.*

The following assumption will be made in the sequel.

Assumption 3 *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 3, we can introduce the family $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant T}$ of stochastic kernels

$$\rho_{s-1:s} : \mathbb{H}_{s-1} \to \mathcal{\Delta}(\mathbb{W}_s) \ , \ \ s = 1, \ldots, T \ , \tag{84a}$$

defined by

$$\rho_{s-1:s}(h_{s-1}, \mathrm{d}w_s) = \mathbb{P}_{\mathbf{W}_s}^{\mathbf{W}_{0:s-1}}\big([h_{s-1}]_{0:s-1}^{\mathbb{W}}, \mathrm{d}w_s\big) \ , \ \ s = 1, \ldots, T \ . \tag{84b}$$

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

$$\mathcal{A}_{t:t} = \{\varnothing, \Omega\} \ , \ \ \mathcal{A}_{t:t+1} = \sigma(\mathbf{W}_{t+1}) \ , \ \ \ldots \ , \ \ \mathcal{A}_{t:T-1} = \sigma(\mathbf{W}_{t+1}, \ldots, \mathbf{W}_{T-1}) \ . \tag{85}$$

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

$$\sigma(\mathbf{U}_s) \subset \mathcal{A}_{t:s} \ , \ \ s = t, \ldots, T-1 \ . \tag{86}$$

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, we consider a measurable nonnegative numerical cost function j as in Equation (12).

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

$$\check{V}_t(h_t) = \inf_{(\mathbf{U}_{t:T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \dots, \mathbf{U}_{T-1}, \mathbf{W}_T) \;\Big|\; \mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t} \Big] \,. \tag{87}$$

Theorem 7 *Let $t \in \{0, \dots, T-1\}$ and $h_t \in \mathbb{H}_t$ be given. Problem (14) and Problem (87) coincide, that is,*

$$\check{V}_t(h_t) = \inf_{(\mathbf{U}_{t:T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \dots, \mathbf{U}_{T-1}, \mathbf{W}_T) \;\Big|\; \mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t} \Big] \tag{88a}$$

$$= \inf_{\gamma_{t:T-1} \in \mathit{\Gamma}_{t:T-1}} \int_{\mathbb{H}_T} j(h'_T) \rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T) \tag{88b}$$

$$= V_t(h_t) \,, \tag{88c}$$

where $\rho^{\gamma}_{t:T}$ is given by Definition 1 with the family $\{\rho_{s-1:s}\}_{1 \leqslant s \leqslant T}$ of stochastic kernels defined in (84), and where the value function $\{V_t\}$ is defined by (15).

In addition, any optimal history feedback $\gamma^\star = \{\gamma^\star_s\}_{s=t,\dots,T-1}$ for Problem (14) yields an optimal adapted control process $(\mathbf{U}^\star_t, \dots, \mathbf{U}^\star_{T-1})$ for Problem (87) by

$$(\mathbf{U}^\star_t, \dots, \mathbf{U}^\star_{T-1}) = \big[\mathit{\Phi}^{\gamma^\star}_{t:T}(h_t, \mathbf{W}_{t+1}, \dots, \mathbf{W}_T) \big]^{\mathbb{U}}_{t+1:T} \,, \tag{89a}$$

(where $[\cdot]^{\mathbb{U}}_{t+1:T}$ is defined in (2f)), or, equivalently, by

$$\mathbf{U}^\star_t = \gamma^\star_t(h_t) \,, \tag{89b}$$

$$\mathbf{U}^\star_{t+1} = \gamma^\star_{t+1}(h_t, \mathbf{U}^\star_t, \mathbf{W}_{t+1}) \,, \tag{89c}$$

$$\vdots$$

$$\mathbf{U}^\star_{T-1} = \gamma^\star_{T-1}(h_t, \mathbf{U}^\star_t, \mathbf{W}_{t+1}, \dots, \mathbf{U}^\star_{T-2}, \mathbf{W}_{T-1}) \,. \tag{89d}$$

Proof Let $t \in \{0, \dots, T-1\}$ and $h_t \in \mathbb{H}_t$ be given. We show that Problem (87) and Problem (14) are in one-to-one correspondence.

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

$$(\mathbf{U}_t, \dots, \mathbf{U}_{T-1}) = \big[\mathit{\Phi}^{\gamma}_{t:T}(h_t, \mathbf{W}_{t+1}, \dots, \mathbf{W}_T) \big]^{\mathbb{U}}_{t+1:T} \,, \tag{90}$$

where the flow $\mathit{\Phi}^{\gamma}_{t:T}$ has been defined in (4) and the history control part $[\cdot]^{\mathbb{U}}_{t+1:T}$ in (2f). By the expression (84b) of $\rho_{s:s+1}(h'_s, \mathrm{d}w_{s+1})$ and by Definition 1 of the stochastic kernel $\rho^{\gamma}_{t:T}$, we obtain that (see details for the expression of $\rho^{\gamma}_{t:T}$ in Appendix A)

$$\mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \dots, \mathbf{U}_{T-1}, \mathbf{W}_T) \;\Big|\; \mathbf{W}_{0:t} = [h_t]^{\mathbb{W}} \Big] = \mathbb{E}\Big[j(\mathit{\Phi}^{\gamma}_{t:T}(h_t, \mathbf{W}_{t+1}, \dots, \mathbf{W}_T)) \;\Big|\; \mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t} \Big]$$
$$= \int_{\mathbb{H}_T} j(h'_T) \rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T) \,. \qquad \text{(by (129) in Appendix A)}$$

As a consequence

$$\inf_{(\mathbf{U}_{t:T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \dots, \mathbf{U}_{T-1}, \mathbf{W}_T) \;\Big|\; \mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t} \Big]$$
$$\leqslant \inf_{\gamma_{t:T-1} \in \mathit{\Gamma}_{t:T-1}} \int_{\mathbb{H}_T} j(h'_T) \rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T) \,. \tag{92}$$

– Second, we define a $(t{:}T-1)$-*noise feedback* as a sequence $\lambda = \{\lambda_s\}_{s=t,\dots,T-1}$ 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 $\mathit{\Lambda}_{t:T-1}$ the set of $(t{:}T-1)$-noise feedbacks.
Let $(\mathbf{U}_t, \dots, \mathbf{U}_{T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})$. As each set $\mathbb{U}_s$ is a separable complete metric space, for $s = t, \dots, T-1$, we can invoke Doob Theorem (see [3, Chapter 1, p. 18]). Therefore, there exists a $(t{:}T-1)$-noise feedback $\lambda = \{\lambda_s\}_{s=t,\dots,T-1} \in \mathit{\Lambda}_{t:T-1}$ such that

$$\mathbf{U}_t = \lambda_t \ , \ \ \mathbf{U}_s = \lambda_s(\mathbf{W}_{t+1:s}) \ , \ \ t+1 \leqslant s \leqslant T-1 \ . \tag{93}$$

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

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

By the expression (84b) of $\rho_{s:s+1}(h'_s, \mathrm{d}w_{s+1})$ and by Definition 1 of the stochastic kernel $\rho^{\gamma}_{t:T}$, we obtain that (see Appendix A for details)

$$\int_{\mathbb{H}_T} j(h'_T)\rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T) = \mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \dots, \mathbf{U}_{T-1}, \mathbf{W}_T) \ \Big| \ \mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t}\Big] \ . \tag{94}$$

As a consequence

$$\begin{aligned}
&\inf_{\gamma_{t:T-1} \in \mathit{\Gamma}_{t:T-1}} \int_{\mathbb{H}_T} j(h'_T)\rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T) \\
&\qquad \leqslant \inf_{(\mathbf{U}_t, \dots, \mathbf{U}_{T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\Big[j(h_t, \mathbf{U}_t, \mathbf{W}_{t+1}, \dots, \mathbf{U}_{T-1}, \mathbf{W}_T) \ \Big| \ \mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t}\Big] \ .
\end{aligned} \tag{95}$$

Gathering inequalities (92) and (95) leads to (88).

The relations (89) allowing to build an optimal adapted control process $(\mathbf{U}^{\star}_t, \dots, \mathbf{U}^{\star}_{T-1})$ for Problem (87) when starting from an optimal history feedback $\gamma^{\star} = \{\gamma^{\star}_s\}_{s=t,\dots,T-1}$ for Problem (14) follow easily. This ends the proof.

An immediate consequence of Theorem 7 and Theorem 1 is the following.

Corollary 4 *The family* $\Big\{\breve{V}_t\Big\}_{t=0,\dots,T}$ *of functions in (87) satisfies the backward induction*

$$\breve{V}_T(h_T) = j(h_T) \ , \ \ \forall h_T \in \mathbb{H}_T \ , \tag{96a}$$

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

$$\begin{aligned}
\breve{V}_t(h_t) &= \inf_{u_t} \int_{\mathbb{W}_{t+1}} \breve{V}_{t+1}\big(h_t, u_t, w_{t+1}\big)\mathbb{P}^{\mathbf{W}_{0:t}}_{\mathbf{W}_{t+1}}\big([h_t]^{\mathbb{W}}_{0:t}, \mathrm{d}w_{t+1}\big) &\text{(96b)} \\
&= \inf_{u_t} \mathbb{E}\big[\breve{V}_{t+1}\big(h_t, u_t, \mathbf{W}_{t+1}\big) \ \big| \ \mathbf{W}_{0:t} = [h_t]^{\mathbb{W}}_{0:t}\big] \ , \ \ \forall h_t \in \mathbb{H}_t \ . &\text{(96c)}
\end{aligned}$$

5.2 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 $t = 0, \dots, T-1$.

5.2.1 The Case of Final Cost Function

We first treat the case of a general criterion, as in §4.1.1.

Proposition 5 *Suppose that there exists a family* $\{\mathbb{X}_t\}_{t=0,\dots,T}$ *of state spaces, with* $\mathbb{X}_0 = \mathbb{W}_0$*, and a family* $\{f_{t:t+1}\}_{t=0,\dots,T-1}$ *of dynamics*

$$f_{t:t+1} : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \to \mathbb{X}_{t+1} \ . \tag{97}$$

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

$$\widetilde{j} : \mathbb{X}_T \to [0, +\infty] \ , \tag{98}$$

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

$$\widetilde{V}_T(x_T) = \widetilde{j}(x_T) \ , \ \ \forall x_T \in \mathbb{X}_T \ , \tag{99a}$$

$$\widetilde{V}_t(x_t) = \inf_{u_t \in \mathbb{U}_t} \mathbb{E}\big[\widetilde{V}_{t+1}\big(x_t, u_t, \mathbf{W}_{t+1}\big)\big] \ , \ \ \forall x_t \in \mathbb{X}_t \ , \tag{99b}$$

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* $x_t \in \mathbb{X}_t$*,*

$$\widetilde{V}_t(x_t) = \inf_{\mathbf{U}_{t:T-1} \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\big[\widetilde{j}(\mathbf{X}_T)\big] \ , \tag{100a}$$

where

$$\mathbf{X}_s = x_t \ , \ \ \mathbf{X}_{s+1} = f_{s:s+1}\big(\mathbf{X}_s, \mathbf{U}_s, \mathbf{W}_{s+1}\big) \ , \ \ \forall s = t, \dots, T-1 \ . \tag{100b}$$

Proof We define a family $\{\theta_t\}_{t=0,\dots,T}$ of reduction mappings $\theta_t : \mathbb{H}_t \to \mathbb{X}_t$ as in (23) 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 (42) to define the mappings $\theta_1, \dots, \theta_T$. As a consequence, the triplet $(\{\mathbb{X}_t\}_{t=0,\dots,T}, \{\theta_t\}_{t=0,\dots,T}, \{f_{t:t+1}\}_{t=0,\dots,T-1})$ is a state reduction across the consecutive time blocks $[t, t+1]_{t=0,\dots,T-1}$ of the time span.

Since the noise process $\{\mathbf{W}_t\}_{t=0,\dots,T}$ 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 (84) is given by

$$\rho_{s-1:s} : \mathbb{H}_{s-1} \to \mathit{\Delta}(\mathbb{W}_s) \ , \ \ s = 1, \dots, T \ , \tag{101a}$$

$$h_{s-1} \mapsto \mathbb{P}_{\mathbf{W}_s}(\mathrm{d}w_s) \ . \tag{101b}$$

As a consequence, we have by (26) that the triplet $(\{\mathbb{X}_t\}_{t=0,\dots,T}, \{\theta_t\}_{t=0,\dots,T}, \{f_{t:t+1}\}_{t=0,\dots,T-1})$ is compatible (see Definition 4) with the family $\{\rho_{t-1:t}\}_{t=1,\dots,T}$ of stochastic kernels in (101). In addition, the reduced stochastic kernels in (26) coincide with the original stochastic kernels in (101).

Define the cost function j as

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

Corollary 2 applies, so that the family $\{V_t\}_{t=0,\dots,T}$ of value functions defined for the family of optimization problems (14) satisfies

$$V_t = \widetilde{V}_t \circ \theta_t \ , \ \ t = 0, \dots, T \ . \tag{102}$$

By means of Theorem 7, we deduce that

$$\check{V}_t(h_t) = \widetilde{V}_t \circ \theta_t(h_t) \ , \tag{103}$$

for all $t = 0, \dots, T$ and for any $h_t \in \mathbb{H}_t$. From the definition (87) of the family of functions $\check{V}_t$, we obtain the expression (100) of functions $\widetilde{V}_t$.

The expression of the optimal state feedbacks is given by the next corollary.

Corollary 5 *Suppose that, for $t = 0, \dots, T-1$, there exist measurable selections*

$$\widetilde{\gamma}_t^\star : (\mathbb{X}_t, \mathcal{X}_t) \to (\mathbb{U}_t, \mathcal{U}_t) \tag{104a}$$

such that

$$\widetilde{\gamma}_t^\star(x_t) \in \operatorname*{arg\,min}_{u_t \in \mathbb{U}_t} \mathbb{E}\big[\widetilde{V}_{t+1}\big(x_t, u_t, \mathbf{W}_{t+1}\big)\big] \ , \ \ \forall x_t \in \mathbb{X}_t \ , \ \ \forall t = T-1, \dots, 0 \ , \tag{104b}$$

where the family $\{\widetilde{V}_t\}_{t=0,\dots,T}$ of functions is given by (99). Then, the family of random variables $\{\mathbf{U}_s^\star\}_{s=t,\dots,T-1}$ defined by

$$\mathbf{U}_s^\star = \widetilde{\gamma}_s^\star \circ \mathbf{X}_s^\star \ , \ \ s = t, \dots, T-1 \ , \tag{105a}$$

where

$$\mathbf{X}_s^\star = x_t \ , \ \ \mathbf{X}_{s+1}^\star = f_{s:s+1}\big(\mathbf{X}_s^\star, \mathbf{U}_s^\star, \mathbf{W}_{s+1}\big) \ , \ \ \forall s = t, \dots, T-1 \ , \tag{105b}$$

is a solution to Problem (100).

Proof The result directly follows from Corollary 2.

5.2.2 The Case of Time Additive Cost Functions

We make the same assumptions than in §4.1.2. We leave the proofs to the reader.

Proposition 6 *Suppose that there exists a family $\{\mathbb{X}_t\}_{t=0,\dots,T}$ of state spaces, with $\mathbb{X}_0 = \mathbb{W}_0$, and a family $\{f_{t:t+1}\}_{t=0,\dots,T-1}$ of dynamics*

$$f_{t:t+1} : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \to \mathbb{X}_{t+1} \ . \tag{106}$$

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

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

$$\widehat{V}_T(x_T) = K(x_T) \ , \ \ \forall x_T \in \mathbb{X}_T \ , \tag{107a}$$

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

$$\widehat{V}_t(x_t) = \inf_{u_t \in \mathbb{U}_t} \mathbb{E}\big[L_t(x_t, u_t, \mathbf{W}_{t+1}) + \widehat{V}_{t+1}\big(f_{t:t+1}(x_t, u_t, \mathbf{W}_{t+1})\big)\big] \ . \tag{107b}$$

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

$$\widehat{V}_t(x_t) = \inf_{(\mathbf{U}_t, \dots, \mathbf{U}_{T-1}) \in \mathbb{L}^0(\Omega, \mathcal{A}_{t:T-1}, \mathbb{U}_{t:T-1})} \mathbb{E}\Big[\sum_{s=t}^{T-1} L_s\big(\mathbf{X}_s, \mathbf{U}_s, \mathbf{W}_{s+1}\big) + K\big(\mathbf{X}_T\big)\Big] \ , \tag{108a}$$

where

$$\mathbf{X}_s = x_t \ , \ \ \mathbf{X}_{s+1} = f_{s:s+1}\big(\mathbf{X}_s, \mathbf{U}_s, \mathbf{W}_{s+1}\big) \ , \ \ \forall s = t, \dots, T-1 \ . \tag{108b}$$

Corollary 6 *Suppose that, for $t = 0, \dots, T-1$, there exists measurable selections*

$$\widehat{\gamma}_t^\star : (\mathbb{X}_t, \mathcal{X}_t) \to (\mathbb{U}_t, \mathcal{U}_t) \ , \tag{109a}$$

such that, for all $x_t \in \mathbb{X}_t$,

$$\widehat{\gamma}_t^\star(x_t) \in \operatorname*{arg\,min}_{u_t \in \mathbb{U}_t} \mathbb{E}\big[L_t(x_t, u_t, \mathbf{W}_{t+1}) + \widehat{V}_{t+1}\big(f_{t:t+1}(x_t, u_t, \mathbf{W}_{t+1})\big)\big] \ . \tag{109b}$$

where the family $\{\widehat{V}_t\}_{t=0,\dots,T}$, of functions is given by (107).

Then, the family of random variables $\{\mathbf{U}_s^\star\}_{s=t,\dots,T-1}$ defined by

$$\mathbf{U}_s^\star = \widehat{\gamma}_s^\star \circ \mathbf{X}_s^\star \ , \ \ s = t, \dots, T-1 \ , \tag{110a}$$

where

$$\mathbf{X}_s^\star = x_t \ , \ \ \mathbf{X}_{s+1}^\star = f_{s:s+1}\big(\mathbf{X}_s^\star, \mathbf{U}_s^\star, \mathbf{W}_{s+1}\big) \ , \ \ \forall s = t, \dots, T-1 \ , \tag{110b}$$

is a solution to Problem (108).

5.3 Two Time-Scales Dynamic Programming

We adopt the notation of § 4.2. We suppose given a two time-scales noise process

$$\mathbf{W}_{(0,0):(D+1,0)} = \big(\mathbf{W}_{0,0}, \mathbf{W}_{0,1}, \dots, \mathbf{W}_{0,M}, \mathbf{W}_{1,0}, \dots, \mathbf{W}_{D,M}, \mathbf{W}_{D+1,0}\big) \ . \tag{111}$$

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

$$\mathcal{A}_{d,0} = \{\varnothing, \Omega\} \ , \ \ \mathcal{A}_{d,m} = \sigma(\mathbf{W}_{(d,1):(d,m)}) \ , \ \ m = 1, \dots, M \ . \tag{112}$$

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

Proposition 7 *Suppose that there exists a family $\{\mathbb{X}_d\}_{d=0,\dots,D+1}$ of state spaces, with $\mathbb{X}_0 = \mathbb{W}_{0,0}$, and a family $\{f_{d:d+1}\}_{d=0,\dots,D}$ of dynamics*

$$f_{d:d+1} : \mathbb{X}_d \times \mathbb{H}_{d:d+1} \to \mathbb{X}_{d+1} \ . \tag{113}$$

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

For a measurable nonnegative numerical cost function

$$\widetilde{j} : \mathbb{X}_{D+1} \to [0, +\infty] \ , \tag{114}$$

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

$$\widetilde{V}_{D+1}(x_{D+1}) = \widetilde{j}(x_{D+1}) \ , \ \ \forall x_{D+1} \in \mathbb{X}_{D+1} \ , \tag{115a}$$

$$\widetilde{V}_d(x_d) = \inf_{\mathbf{U}_{(d,0):(d,M)} \in \mathbb{L}^0(\Omega, \mathcal{A}_{(d,0):(d,M)}, \mathbb{U}_{(d,0):(d,M)})} \mathbb{E}\Big[\widetilde{V}_{d+1}\big(f_{d:d+1}(x_d, \mathbf{U}_{d,0}, \mathbf{W}_{d,1}, \cdots, \mathbf{U}_{d,M}, \mathbf{W}_{d+1,0})\big)\Big] \ , \ \ \forall x_d \in \mathbb{X}_d \ , \tag{115b}$$

for $d = D, \dots, 0$.

Then, the value functions $\widetilde{V}_d$ in (115) are the solution of the following family of optimization problems, indexed by $d = 0, \dots, D$ and by $x_d \in \mathbb{X}_d$,

$$\widetilde{V}_d(x_d) = \inf_{U_{(d,0):(D,M)} \in \mathbb{L}^0(\Omega, \mathcal{A}_{(d,0):(D,M)}, \mathbb{U}_{(d,0):(D,M)})} \mathbb{E}\big[\widetilde{j}(\mathbf{X}_{D+1})\big] \ , \tag{116a}$$

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

$$\mathbf{X}_d = x_d \ , \ \ \mathbf{X}_{d'+1} = f_{d':d'+1}\big(\mathbf{X}_{d'}, \mathbf{U}_{d',0}, \mathbf{W}_{d',1}, \cdots, \mathbf{U}_{d',M}, \mathbf{W}_{d'+1,0}\big) \ . \tag{116b}$$

5.4 Decision Hazard Decision Dynamic Programming

We adopt the notation of § 4.3. We suppose given a noise process

$$\mathbf{W}_{0:S} = \big(\mathbf{W}_0^\sharp, \mathbf{W}_1^\flat, \dots, \mathbf{W}_S^\flat\big) \ . \tag{117}$$

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

$$\mathcal{A}_s = \{\varnothing, \Omega\} \ , \ \ \mathcal{A}_{s'} = \sigma(\mathbf{W}^\flat_{s+1:s'}) \ , \ \ s' = s+1, \dots, S \ . \tag{118}$$

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

Proposition 8 *Suppose that there exists a family* $\{\mathbb{X}_s\}_{s=0,\dots,S}$ *of state spaces, with* $\mathbb{X}_0 = \mathbb{W}_0^\sharp$*, and a family* $\{f_{s:s+1}\}_{s=0,\dots,S-1}$ *of dynamics*

$$f_{s:s+1} : \mathbb{X}_s \times \mathbb{U}_s^\sharp \times \mathbb{W}_{s+1}^\flat \times \mathbb{U}_{s+1}^\flat \to \mathbb{X}_{s+1} \; . \tag{119}$$

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

For a measurable nonnegative numerical cost function

$$\widetilde{j} : \mathbb{X}_S \to [0, +\infty] \; , \tag{120}$$

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

$$\widetilde{V}_S(x_S) = \widetilde{j}(x_S) \; , \quad \forall x_S \in \mathbb{X}_S \; , \tag{121a}$$

$$\widetilde{V}_s(x_s) = \inf_{u_s^\sharp \in \mathbb{U}_s^\sharp} \mathbb{E}\Big[\inf_{u_{s+1}^\flat \in \mathbb{U}_{s+1}^\flat} \widetilde{V}_{s+1}\Big(f_{s':s'+1}\big(x_s, u_s^\sharp, \mathbf{W}_{s+1}^\flat, u_{s+1}^\flat\big)\Big)\Big] \tag{121b}$$

$$\forall x_s \in \mathbb{X}_s \; , \quad \forall s = S-1, \dots, 0 \; . \tag{121c}$$

Then, the value functions $\widetilde{V}_s$ *in (121) are the solution of the following family of optimization problems, indexed by* $s = 0, \dots, S-1$ *and by* $x_s \in \mathbb{X}_s$*,*

$$\widetilde{V}_s(x_s) = \inf_{\mathbf{U}_{s:S-1}^\sharp \in \mathbb{L}^0(\Omega, \mathcal{A}_{s:S-1}, \mathbb{U}_{s:S-1}^\sharp)} \; \inf_{\mathbf{U}_{s+1:S}^\flat \in \mathbb{L}^0(\Omega, \mathcal{A}_{s+1:S}, \mathbb{U}_{s+1:S}^\flat)} \mathbb{E}\big[\widetilde{j}(\mathbf{X}_S)\big] \; , \tag{122a}$$

where

$$\mathbf{X}_{s'} = x_s \; , \quad \mathbf{X}_{s'+1} = f_{s':s'+1}\big(\mathbf{X}_{s'}, \mathbf{U}_{s'}^\sharp, \mathbf{W}_{s'+1}^\flat, \mathbf{U}_{s'+1}^\flat\big) \; , \quad \forall s' = s, \dots, S-1 \; . \tag{122b}$$

6 Conclusion

As said in the Introduction Sect. 1, the large scale nature of multistage stochastic optimization problems makes decomposition methods appealing. We have provided a method to decompose multistage stochastic optimization problems by time blocks.

In the case of optimization with noise process, we do not require noise independence within the time blocks. This opens the possibility to apply stochastic dynamic programming between the extremities of the time blocks — at a slow time scale for which noise would be statistically independent — and to apply stochastic programming within the time blocks. Therefore, our time block decomposition paves the way for mixing and reconciliating stochastic dynamic programming and stochastic programming methods.

Such an approach is part of a larger research program, where we aim at mixing various decomposition-coordination methods in multistage stochastic optimization, be they spatial, temporal or by scenarios [4].

A Construction of the stochastic kernels $\boldsymbol{\rho_{r:t}^\gamma}$

We detail here the construction of the stochastic kernels $\rho_{r:t}^\gamma$ in (7a) when $0 \leqslant r < t \leqslant T$. We assume that the $(\mathbb{W}_s)_{s=0,\dots,T}$ are measurable spaces and we denote by $(\mathcal{W}_s)_{s=0,\dots,T}$ the associated σ-fields.

1. In the first step, we build a family of stochastic kernels $(\nu_{r,s:s+1}^\gamma)_{s=r,\cdots,t-1}$ using composition and then we follow [5, p.138] (see also [2, Proposition 7.28]) to define a stochastic kernel product $\mu_{r:t+1}^\gamma = \bigotimes_{s=r}^{t-1} \nu_{r,s:s+1}^\gamma$. More precisely, let r and t be fixed (such that $0 \leqslant r < t \leqslant T$). First, for $s = r$, we simply define $\nu_{r,r:r+1}^\gamma = \rho_{r:r+1}$. Second, for each s such that $0 \leqslant r < s < t$, we define a new stochastic kernel $\nu_{r,s:s+1}^\gamma$ by the composition $\nu_{r,s:s+1}^\gamma = \rho_{s:s+1} \circ \Phi_{r:s}^\gamma$:

$$\begin{array}{ccccc} & & \overset{\nu_{r,s:s+1}^\gamma}{\frown} & & \\ \mathbb{H}_r \times \mathbb{W}_{r+1:s} & \xrightarrow{\Phi_{r:s}^\gamma} & \mathbb{H}_s & \xrightarrow{\rho_{s:s+1}} & \Delta(\mathbb{W}_{s+1}) \; . \end{array} \tag{123}$$

Now, proceeding as in [5, p.138], we construct a stochastic kernel

$$\mu^{\gamma}_{r:t} : \mathbb{H}_r \to \Delta(\mathbb{W}_{r+1:t}) \,,$$

obtained as a product of the stochastic kernels $(\nu^{\gamma}_{r,s:s+1})_{s=r,\dots,t-1}$. The construction if as follows: for a fixed $h_r \in \mathbb{H}_r$ and a fixed sequence of measurable sets $B_{r+1:t} \in \mathcal{W}_{r+1:t}$, we put

$$\begin{aligned}\mu^{\gamma}_{r:t}(h_r, B_{r+1:t}) = \int_{B_{r+1}} \cdots \int_{B_{s+1}} \cdots \Big(\int_{B_t} \nu^{\gamma}_{r,t-1:t}(h_r, w_{r+1:t-1}, \mathrm{d}w_t) \Big) \\ \cdots \nu^{\gamma}_{r,s:s+1}(h_r, w_{r+1:s}, \mathrm{d}w_{s+1}) \cdots \nu^{\gamma}_{r,r:r+1}(h_r, \mathrm{d}w_{r+1}) \,. \end{aligned} \tag{124}$$

2. The second step is to define the stochastic kernel $\rho^{\gamma}_{r:t} : \mathbb{H}_r \to \Delta(\mathbb{H}_t)$ from the stochastic kernel $\mu^{\gamma}_{r:t}$ using transport with the flow $\Phi^{\gamma}_{r:t} : \mathbb{H}_r \times \mathbb{W}_{r+1:t} \to \mathbb{H}_t$. More precisely, for any measurable nonnegative function $\varphi : \mathbb{H}_t \to [0, +\infty]$, we define the integral with respect to the stochastic kernel $\rho^{\gamma}_{r:t}$ as the integral of the function $\varphi \circ \Phi^{\gamma}_{r:t}$ with respect to the kernel $\mu^{\gamma}_{r:t}$:

$$\int_{\mathbb{H}_t} \varphi(h'_t) \rho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_t) = \int_{\mathbb{W}_{r+1:t}} \varphi\Big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t})\Big) \mu^{\gamma}_{r:t}(h_r, \mathrm{d}w_{r+1:t}) \,. \tag{125}$$

$$\begin{array}{ccc} \mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t) \ni \varphi & \xrightarrow{\quad \rho^{\gamma}_{r:t} \quad} & \langle \varphi \,, \rho^{\gamma}_{r:t} \rangle \in \mathbb{L}^0_+(\mathbb{H}_r, \mathcal{H}_r) \\ \Big\downarrow & & \Big\uparrow = \\ \mathbb{L}^0_+\Big(\mathbb{H}_r \times \mathbb{W}_{r+1:t}, \mathcal{H}_r \otimes \mathcal{W}_{r+1:t}\Big) \ni \psi = \varphi \circ \Phi^{\gamma}_{r:t} & \xrightarrow{\quad \mu^{\gamma}_{r:t} \quad} & \langle \psi \,, \mu^{\gamma}_{r:t} \rangle \,. \end{array}$$

This ends the construction.

B Specialization to the noise case

We turn now to the special case where, for any $s = 0, \dots, T-1$, the stochastic kernel $\rho_{s:s+1}$ is the regular conditional distribution $\mathbb{P}^{\mathbf{W}_{0:s}}_{\mathbf{W}_{s+1}}$ of the random variable $\mathbf{W}_{s+1}$ knowing the random process $\mathbf{W}_{0:s}$, that is,

$$\rho_{s:s+1}(h_s, \mathrm{d}w_{s+1}) = \mathbb{P}^{\mathbf{W}_{0:s}}_{\mathbf{W}_{s+1}}\big([h_s]^{\mathbb{W}}_{0:s}, \mathrm{d}w_{s+1}\big) \,. \tag{126}$$

For any s such that $0 \leqslant r < s < t$ and $B_{s+1} \in \mathcal{W}_{s+1}$, we have that

$$\begin{aligned} \nu^{\gamma}_{r,s:s+1}\big((h_r, w_{r+1:s}), B_{s+1}\big) &= \rho_{s:s+1}\big(\Phi^{\gamma}_{r:s}(h_r, w_{r+1:s}), B_{s+1}\big) && \text{(by (123))} \\ &= \mathbb{P}^{\mathbf{W}_{0:s}}_{\mathbf{W}_{s+1}}\big(\big[\Phi^{\gamma}_{r:s}(h_r, w_{r+1:s})\big]^{\mathbb{W}}_{0:s}, B_{s+1}\big) \,, && \text{(by (126))} \end{aligned}$$

which, using Equations (2e) and (4b), gives

$$= \mathbb{P}^{\mathbf{W}_{0:s}}_{\mathbf{W}_{s+1}}\big(([h_r]^{\mathbb{W}}_{0:r}, w_{r+1:s}), B_{s+1}\big) \,. \tag{127}$$

We observe that the stochastic kernel $\nu^{\gamma}_{r,s:s+1}$ does not depend on the history feedback γ. As a consequence, the stochastic kernel $\mu^{\gamma}_{r:t} : \mathbb{H}_r \to \Delta(\mathbb{W}_{r+1:t})$ obtained by product in (124), does not depend on the history feedback γ either, and can be expressed using the regular conditional distribution of $\mathbf{W}_{r+1:t}$ knowing the random process $\mathbf{W}_{0:r}$. By (127) and (124), for a fixed sequence $B_{r+1:t} \in \mathcal{B}(\mathbb{W}_{r+1:t})$ of Borel sets, we have

$$\mu^{\gamma}_{r:t}(h_r, B_{r+1:t}) = \mathbb{P}^{\mathbf{W}_{0:r}}_{\mathbf{W}_{r+1:t}}\big([h_r]^{\mathbb{W}}_{0:r}, B_{r+1:t}\big) \,. \tag{128}$$

Now, for any measurable nonnegative function $\varphi : \mathbb{H}_t \to [0, +\infty]$, the integral with respect to the stochastic kernel $\rho^{\gamma}_{r:t}$ is defined by (125) as the integral of the function $\varphi \circ \Phi^{\gamma}_{r:t-1}$ with respect to the kernel $\mu^{\gamma}_{r:t}$. Using Equation (128), this gives

$$\begin{aligned} \int_{\mathbb{H}_t} \varphi(h'_t) \rho^{\gamma}_{r:t}(h_r, \mathrm{d}h'_t) &= \int_{\mathbb{W}_{r+1:t}} \varphi\Big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t})\Big) \mu^{\gamma}_{r:t}(h_r, \mathrm{d}w_{r+1:t}) \\ &= \int_{\mathbb{W}_{r+1:t}} \varphi\Big(\Phi^{\gamma}_{r:t}(h_r, w_{r+1:t})\Big) \mathbb{P}^{\mathbf{W}_{0:r}}_{\mathbf{W}_{r+1:t}}\big([h_r]^{\mathbb{W}}_{0:r}, \mathrm{d}w_{r+1:t}\big) \,. \\ &= \mathbb{E}\Big[\varphi(\Phi^{\gamma}_{r:t}(h_r, \mathbf{W}_{r+1:t})) \;\Big|\; \mathbf{W}_{0:r} = [h_r]^{\mathbb{W}}_{0:r}\Big] \,. \end{aligned} \tag{129}$$

References

1. R. E. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, N.J., 1957.
2. D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific, Belmont, Massachusets, 1996.
3. C. Dellacherie and P.A. Meyer. *Probabilités et potentiel*. Hermann, Paris, 1975.
4. M. De Lara, P. Carpentier, J.-P. Chancelier, and V. Leclère. Optimization methods for the smart grid. Report commissioned by the Conseil Français de l'Énergie, École des Ponts ParisTech, October 2014.
5. M. Loève. *Probability Theory I*. Springer-Verlag, New York, fourth edition, 1977.
6. M. L. Puterman. *Markov Decision Processes*. Wiley, New York, 1994.
7. R. T. Rockafellar and R. J-B. Wets. *Variational Analysis*. Springer-Verlag, Berlin, 1998.
8. 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.
9. 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.