Time Consistency for Multistage Stochastic Optimization Problems under Constraints in Expectation

Pierre Carpentier, Jean-Philippe Chancelier, Michel de Lara

Time Consistency for Multistage Stochastic Optimization Problems under Constraints in Expectation

Pierre Carpentier[*], Jean-Philippe Chancelier[†], Michel De Lara[†]

August 26, 2022

Abstract

We consider sequences — indexed by time (discrete stages) — of families of multistage stochastic optimization problems. At each time, the optimization problems in a family are parameterized by some quantities (initial states, constraint levels...). In this framework, we introduce an adapted notion of time consistent optimal solutions, that is, solutions that remain optimal after truncation of the past and that are optimal for any values of the parameters. We link this time consistency notion with the concept of state variable in Markov Decision Processes for a class of multistage stochastic optimization problems incorporating state constraints at the final time, either formulated in expectation or in probability. For such problems, when the primitive noise random process is stagewise independent and takes a finite number of values, we show that time consistent solutions can be obtained by considering a finite dimensional state variable. We illustrate our results on a simple dam management problem.

Keywords: Multistage Stochastic Optimization; Time Consistency; Constraints in Expectation; Dynamic Programming

1 Introduction and motivation

The notion of time consistency has been introduced in the field of Economics [11], and developed in the context of risk measures [1, 14, 7, 6]. It has been studied in stochastic optimization, both from the stochastic programming [16, 13] and from the Markov Decision Process [15] points of view. Loosely speaking, time consistency means that strategies obtained by solving the problem at the very first stage do not have to be questioned later on. This definition has been used in [5] to establish links between the concept of state

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

variable and the notion of time consistency. The aim in [5] was to highlight the role of information in time consistency. For example, considering a standard multistage stochastic optimization problem solvable by dynamic programming, it was shown that adding a probabilistic constraint involving the state at the final instant of the time span invalidates the time inconsistency property, in the sense that optimal strategies based on the usual state variable have to be reconsidered at each time stage. It was also shown that it was possible to devise an appropriate state variable, namely the probabilistic distribution of the state variable rather than the state variable itself, to formulate an equivalent problem enjoying the time consistency property. But this state is an infinite dimensional one, so that dynamic programming is usually not implementable. The aim of this article is to give deeper insights into the results established in [5] and to show that it is possible to regain time consistency on such problems by using an extended finite dimensional state variable.

The paper is organized as follows. In Sect. 2, we introduce the notion of universal solution for a family of optimization problems, and we define the notion of time consistency for a sequence of families of optimization problems. Then, we revisit the setting of a discrete time multistage stochastic optimization problem in the standard formulation, and we show that our definition of time consistency applies in this case. In Sect. 3, we add an expectation constraint on the final state to the standard multistage stochastic optimization problem, and we define families of optimization problems parameterized by both the initial state and the level of constraint. We prove that the feedback strategies — obtained by dynamic programming on an extended problem formulation with additional state and control variables — are time consistent. In Sect. 4, we present a toy problem for managing a dam subject to a final constraint in probability, and we give results obtained using the extended formulation. Finally, we draw some conclusions in Sect. 5.

2 Time consistency and multistage stochastic optimization

In §2.1, we introduce the notion of universal solution of a family of optimization problems and the notion of time consistency of a sequence of controls for an optimization data set. In §2.2, we show how these two notions apply in multistage stochastic optimization.

2.1 Universal solutions and time consistency

We start with general considerations on universal solutions and time consistency, before moving to more formal statements.

In optimization, the most natural notion of *universal solution* is the following. Let $\mathbb{A}$ be a set of parameters and $\mathbb{U}$ be a set of (decision) variables. Let $\{J^a\}_{a\in\mathbb{A}}$ be a family of functions[1] $J^a : \mathbb{U} \to \mathbb{R} \cup \{+\infty\}$ indexed by the parameter a. An element $u^\sharp \in \mathbb{U}$ is a

[1]Adopting usage in mathematics, we follow Serge Lang and use "function" only to refer to mappings in which the codomain is numerical — that is, a set of numbers (i.e. a subset of $\mathbb{R}$ or $\mathbb{C}$, or their possible

universal solution for the family of functions $\{J^a\}_{a\in\mathbb{A}}$ when $u^\sharp \in \bigcap_{a\in\mathbb{A}} \arg\min_{u\in\mathbb{U}} J^a(u)$. In a different way, the most natural notion of *time consistency* in (mutistage) optimization is the following. Let $t_i < t_f$ be two integers, $[\![t_i, t_f]\!] = \{t_i, t_i{+}1, \ldots, t_f{-}1, t_f\}$ be the corresponding finite time span, and $\{\mathbb{U}_t\}_{t\in[\![t_i,t_f]\!]}$ be a sequence of control sets. We introduce the *truncation mapping* (projection) at time $t \in [\![t_i, t_f]\!]$, that is, $\mathfrak{T}_t : \mathbb{U}_{t_i} \times \cdots \times \mathbb{U}_{t_f} \to \mathbb{U}_t \times \cdots \times \mathbb{U}_{t_f}$ and $\mathfrak{T}_t(u_{t_i}, \ldots, u_{t_f}) = (u_t, \ldots, u_{t_f})$. Then, considering a sequence of functions $\{J_t\}_{t\in[\![t_i,t_f]\!]}$, with $J_t : \mathbb{U}_t \times \cdots \times \mathbb{U}_{t_f} \to \mathbb{R}\cup\{+\infty\}$, we say that *time consistency* holds when, for all $t \in [\![t_i, t_f]\!]$,

$$\mathfrak{T}_t\Big(\operatorname*{arg\,min}_{(u_{t_i},\ldots,u_{t_f})\in\mathbb{U}_{t_i}\times\cdots\times\mathbb{U}_{t_f}} J_{t_i}(u_{t_i}, \ldots, u_{t_f})\Big) \subset \operatorname*{arg\,min}_{(u_t,\ldots,u_{t_f})\in\mathbb{U}_t\times\cdots\times\mathbb{U}_{t_f}} J_t(u_t, \ldots, u_{t_f}) \; .$$

We now extend and mix these two notions in the case where the (cost) functions depend on both parameters and time.

Definition 1 *We call* optimization data set *a family* $\mathcal{D} = \big(\mathcal{T},\, \{\mathbb{A}_t\}_{t\in\mathcal{T}},\, \{\mathbb{U}_t\}_{t\in\mathcal{T}},\, \{J_t\}_{t\in\mathcal{T}}\big)$, *where* $\mathcal{T} = [\![t_i, t_f]\!]$ *(with* $t_i < t_f$ *two integers), a sequence* $\{\mathbb{A}_t\}_{t\in\mathcal{T}}$ *of parameter sets, a sequence* $\{\mathbb{U}_t\}_{t\in\mathcal{T}}$ *of control sets, a sequence* $\{J_t\}_{t\in\mathcal{T}}$ *of cost functions, with* $J_t : \mathbb{A}_t \times \mathbb{U}_t \times \cdots \times \mathbb{U}_{t_f} \to \mathbb{R}\cup\{+\infty\}$.

For any $t \in \mathcal{T}$, *we call* truncated optimization data set at time t *the optimization data set* $\mathcal{D}_t = \big([\![t, t_f]\!], \{\mathbb{A}_s\}_{s\in[\![t,t_f]\!]}, \{\mathbb{U}_s\}_{s\in[\![t,t_f]\!]}, \{J_s\}_{s\in[\![t,t_f]\!]}\big)$.

The notion of universal solution for an optimization data set at time $t \in \mathcal{T}$ is the following.

Definition 2 *Let* $\mathcal{D}$ *be an optimization data set and let* $t \in \mathcal{T}$ *be given. We say that* $(u_t^\sharp, \ldots, u_{t_f}^\sharp) \in \mathbb{U}_t \times \cdots \times \mathbb{U}_{t_f}$ *is a* universal solution for the data set $\mathcal{D}$ at time t *if it satisfies*

$$(u_t^\sharp, \ldots, u_{t_f}^\sharp) \in \bigcap_{a_t\in\mathbb{A}_t} \arg\min\big\{\mathcal{P}_t^{\mathcal{D}}(a_t)\big\} \; , \tag{1}$$

where, for any $a_t \in \mathbb{A}_t$, *the optimization problem* $\mathcal{P}_t^{\mathcal{D}}(a_t)$ *is defined by*

$$\min_{(u_t,\ldots,u_{t_f})\in\mathbb{U}_t\times\cdots\times\mathbb{U}_{t_f}} J_t(a_t, u_t, \ldots, u_{t_f}) \; . \tag{2}$$

The property of *time consistency* of a sequence $(u_{t_i}^\sharp, \ldots, u_{t_f}^\sharp)$ of controls for an optimization data set $\mathcal{D}$ is defined as follows.

Definition 3 *Let* $\mathcal{D}$ *be an optimization data set and let* $(u_{t_i}^\sharp, \ldots, u_{t_f}^\sharp)$ *be a sequence of controls in* $\mathbb{U}_{t_i} \times \cdots \times \mathbb{U}_{t_f}$. *We say that the sequence* $(u_{t_i}^\sharp, \ldots, u_{t_f}^\sharp)$ *of controls is* time consistent for the optimization data set $\mathcal{D}$ *if, for any* $t \in \mathcal{T}$, *the truncated subsequence* $(u_t^\sharp, \ldots, u_{t_f}^\sharp) = \mathfrak{T}_t(u_{t_i}^\sharp, \ldots, u_{t_f}^\sharp)$ *of controls is a universal solution for the optimization data set* $\mathcal{D}$ *at time* t.

Otherwise stated, given a universal solution $(u_{t_i}^\sharp, \ldots, u_{t_f}^\sharp)$ for the data set $\mathcal{D}$ at initial time t_i, time consistency means that the subsequence $(u_t^\sharp, \ldots, u_{t_f}^\sharp)$ is a universal solution for the data set $\mathcal{D}$ at time t, for any time $t > t_i$.

extensions with $\pm\infty$) — and reserve the term "mapping" for more general codomains.

2.2 Multistage stochastic optimization in the classical case

We study the standard case of a controlled dynamical system influenced by exogenous disturbances. The decision maker has to find strategies to drive the system so as to minimize some objective function over a certain time span.

Let be given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. All random variables and all random processes are defined on $(\Omega, \mathcal{F}, \mathbb{P})$, and we denote them using bold letters. We denote by $\sigma(\boldsymbol{Z}) \subset \mathcal{F}$ the σ-field generated by a random variable $\boldsymbol{Z}$.

We consider a positive integer $T > 0$ and the finite time span $[\![0, T]\!] = \{0, 1, \ldots, T-1, T\}$. We denote by $\boldsymbol{W} = \{\boldsymbol{W}_t\}_{t=1,\ldots,T}$ the primitive (or exogenous) *noise random process*, where each random variable $\boldsymbol{W}_t$ takes values in a measurable space $\mathbb{W}_t$. We denote by $\boldsymbol{U} = \{\boldsymbol{U}_t\}_{t=0,\ldots,T-1}$ the *control random process*, where each random variable $\boldsymbol{U}_t$ takes values in a measurable space $\mathbb{U}_t$, and by $\boldsymbol{X} = \{\boldsymbol{X}_t\}_{t=0,\ldots,T}$ the *state random process*, where each random variable $\boldsymbol{X}_t$ takes values in a measurable space $\mathbb{X}_t$. We consider a sequence $\{f_t\}_{t=0,\ldots,T-1}$ of measurable mappings $f_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \to \mathbb{X}_{t+1}$ (dynamics), a sequence $\{L_t\}_{t=0,\ldots,T-1}$ of measurable functions $L_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \to \mathbb{R}_+ \cup \{+\infty\}$ (instantaneous cost), and a measurable function $K : \mathbb{X}_{T+1} \to \mathbb{R}_+ \cup \{+\infty\}$ (final cost).

The optimization problem we consider below consists in minimizing the expectation of a sum of costs depending on the state, the control and the noise variables over the finite time span $[\![0, T]\!]$. The state variable evolves with respect to the dynamics f_t that depends on the current state, noise and control values. The problem starting at time $t = 0$ is

$$\min_{\boldsymbol{U}, \boldsymbol{X}} \quad \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + K(\boldsymbol{X}_T)\right], \tag{3a}$$

$$\text{s.t.} \quad \boldsymbol{X}_0 = x_0 \,, \tag{3b}$$

$$\boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) \,, \qquad \forall t = 0, \ldots, T-1 \,, \tag{3c}$$

$$\sigma(\boldsymbol{U}_t) \subset \sigma(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_t) \,, \qquad \forall t = 0, \ldots, T-1 \,. \tag{3d}$$

By convention, for $t = 0$, the σ-field $\sigma(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_t)$ is the trivial σ-field $\{\emptyset, \Omega\}$. Under the measurability assumptions made, Problem (3) is well-defined as all functions take extended nonnegative values.

We make the following assumption.

Assumption 1 (Markovian setting) *The noise random variables $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_T$ are independent.*

Using Assumption 1, it is well known [2] that there is no loss of optimality in looking for the optimal control $\boldsymbol{U}_t$ at time t of Problem (3) as a feedback strategy depending on the state variable $\boldsymbol{X}_t$, that is, as a measurable mapping $\phi_t : \mathbb{X}_t \to \mathbb{U}_t$ (state feedback).

Let us embed Problem (3) in the framework developed in §2.1. For that purpose, we build an optimization data set

$$\mathcal{S} = \left(\mathcal{T}, \{\mathbb{X}_t\}_{t\in\mathcal{T}}, \{\mathcal{U}_t\}_{t\in\mathcal{T}}, \{J_t\}_{t\in\mathcal{T}}\right) . \tag{4}$$

The discrete time span $\mathcal{T}$ is $[\![0, T-1]\!]$, the sequence of parameter sets is $\{\mathbb{X}_t\}_{t\in\mathcal{T}}$, and the sequence of control spaces is $\{\mathcal{U}_t\}_{t\in\mathcal{T}}$, $\mathcal{U}_t$ being the space of measurable mappings $\phi_t : \mathbb{X}_t \to \mathbb{U}_t$ (state feedbacks). The sequence $\{J_t\}_{t\in\mathcal{T}}$ of cost functions

$$\begin{aligned} J_t \ : \mathbb{X}_t \times \mathcal{U}_t \times \cdots \times \mathcal{U}_{T-1} \ &\longrightarrow \ \mathbb{R} \cup \{+\infty\} \\ (x_t, \phi_t, \ldots, \phi_{T-1}) \ &\longmapsto \ J_t(x_t, \phi_t, \ldots, \phi_{T-1}) \ , \end{aligned}$$

is defined by

$$\begin{aligned} J_t(x_t, \phi_t, \ldots, \phi_{T-1}) = \mathbb{E}\left[\sum_{\tau=t}^{T-1} L_\tau\big(\boldsymbol{X}_\tau, \phi_\tau(\boldsymbol{X}_\tau), \boldsymbol{W}_{\tau+1}\big) + K(\boldsymbol{X}_T) \right] , \\ \text{with} \quad \boldsymbol{X}_t = x_t \ , \\ \boldsymbol{X}_{\tau+1} = f_\tau\big(\boldsymbol{X}_\tau, \phi_\tau(\boldsymbol{X}_\tau), \boldsymbol{W}_{\tau+1}\big) \ , \ \ \forall \tau \in [\![t, T-1]\!] \ . \end{aligned}$$

Thanks to dynamic programming, we obtain the following result. The sequence $(\phi_0^\sharp, \ldots, \phi_{T-1}^\sharp)$ of optimal strategies of Problem (3), obtained by solving the dynamic programming equation backward in time

$$V_T(x) = K(x) \ , \tag{5a}$$

$$V_t(x) = \min_{u\in\mathbb{U}_t} \mathbb{E}\Big[L_t(x, u, \boldsymbol{W}_{t+1}) + V_{t+1}\big(f_t(x, u, \boldsymbol{W}_{t+1})\big) \Big] \ , \tag{5b}$$

with

$$\phi_t^\sharp(x) \in \operatorname*{arg\,min}_{u\in\mathbb{U}_t} \mathbb{E}\Big[L_t(x, u, \boldsymbol{W}_{t+1}) + V_{t+1}\big(f_t(x, u, \boldsymbol{W}_{t+1})\big) \Big] \ , \tag{5c}$$

is time consistent, in the sense of Definition 3, for the optimization data set $\mathcal{S}$ defined in Equation (4). Indeed, letting time $t \in \mathcal{T}$ be given, we build from the data set $\mathcal{S}$ the family $\mathcal{P}_t^{\mathcal{S}} = \big\{\mathcal{P}_t^{\mathcal{S}}(x_t)\big\}_{x_t\in\mathbb{X}_t}$ of optimization problems as in Equation (2), with Problem $\mathcal{P}_t^{\mathcal{S}}(x_t)$ being

$$\min_{(\phi_t,\ldots,\phi_{T-1})\in\mathcal{U}_t\times\cdots\times\mathcal{U}_{T-1}} \ J_t(x_t, \phi_t, \ldots, \phi_{T-1}) \ . \tag{6}$$

It is clear that Problem $\mathcal{P}_0^{\mathcal{S}}(x_0)$ coincides with Problem (3). From the Bellman theory, we know that the sequence $(\phi_0^\sharp, \ldots, \phi_{T-1}^\sharp)$ of strategies obtained by solving the dynamic programming equation (5) is such that, for any $t \in \mathcal{T}$, the truncated sequence $(\phi_t^\sharp, \ldots, \phi_{T-1}^\sharp)$ is an optimal solution of Problem (6) for any initial state x_t. Thus, according to Definition 3, the sequence $(\phi_0^\sharp, \ldots, \phi_{T-1}^\sharp)$ of controls is time consistent for the optimization data set $\mathcal{S}$.

Remark 4 *The notion of time consistency crucially depends on the nature of the solutions of the family of optimization problem under consideration. As a matter of fact, consider Problem* (3) *and its solution* $(\phi_0^\sharp, \ldots, \phi_{T-1}^\sharp)$ *in terms of feedback strategies: as already explained, there is no difficulty to apply the truncated sequence* $(\phi_t^\sharp, \ldots, \phi_{T-1}^\sharp)$ *to Problem* (6)

since this truncated sequence is admissible for the problem starting at time t. But consider again Problem (3) *and its solution* $\boldsymbol{U}^\sharp = (\boldsymbol{U}_0^\sharp, \ldots, \boldsymbol{U}_{T-1}^\sharp)$ *in terms of random variables. Problem* (6) *is equivalent to*

$$\min_{\boldsymbol{U},\boldsymbol{X}} \quad \mathbb{E}\left[\sum_{\tau=t}^{T-1} L_\tau(\boldsymbol{X}_\tau, \boldsymbol{U}_\tau, \boldsymbol{W}_{\tau+1}) + K(\boldsymbol{X}_T)\right], \tag{7a}$$

$$s.t. \quad \boldsymbol{X}_t = x_t \,, \tag{7b}$$

$$\boldsymbol{X}_{\tau+1} = f_\tau(\boldsymbol{X}_\tau, \boldsymbol{U}_\tau, \boldsymbol{W}_{\tau+1}) \,, \qquad \forall \tau \in [\![t, T-1]\!] \,, \tag{7c}$$

$$\sigma(\boldsymbol{U}_\tau) \subset \sigma(\boldsymbol{W}_{t+1}, \ldots, \boldsymbol{W}_\tau) \,, \qquad \forall \tau \in [\![t, T-1]\!] \,. \tag{7d}$$

We note that the truncated subsequence $(\boldsymbol{U}_t^\sharp, \ldots, \boldsymbol{U}_{T-1}^\sharp)$ *of* $\boldsymbol{U}^\sharp$ *is not even admissible for Problem 7 as it does not satisfy* (7d). *Indeed, each* $\boldsymbol{U}_\tau^\sharp$ *for* $\tau \in [\![t, T-1]\!]$ *is by construction (see Constraint* (3d)*) measurable with respect to the* σ*-field* $\sigma(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_\tau)$ *and thus does not satisfy Constraint* (7d).

3 Multistage stochastic optimization with a final constraint in expectation

In §3.1, we modify the framework studied in §2.2 by adding to Problem (3) a constraint in expectation involving the final state $\boldsymbol{X}_T$, which leads to the optimization problem (8) below. In §3.2, we propose a reformulation of Problem (8) involving a finite dimensional state, and an optimization data set (including the initial state and the level of the expectation constraint) for which time consistency holds. In §3.3, we propose a dual problem formulation of Problem (8), and we illustrate the fact that such a reformulation is not time consistent.

3.1 Standard formulation

We use the notations defined in §2.2. We consider a measurable function $g : \mathbb{X}_T \to \mathbb{R}_+^m$, For convenience, we denote $\mathbb{Z}_T = \mathbb{R}^m$. The stochastic optimization problem starting at time $t_0 \in [\![0, T-1]\!]$ with a final constraint in expectation at time T is

$$\min_{\boldsymbol{U},\boldsymbol{X}} \quad \mathbb{E}\left[\sum_{t=t_0}^{T-1} L_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\big) + K(\boldsymbol{X}_T)\right], \tag{8a}$$

$$\text{s.t.} \quad \boldsymbol{X}_{t_0} = x_{t_0} \,, \tag{8b}$$

$$\boldsymbol{X}_{t+1} = f_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\big) \,, \qquad \forall t \in [\![t_0, T-1]\!] \,, \tag{8c}$$

$$\sigma\big(\boldsymbol{U}_t\big) \subset \sigma\big(\boldsymbol{W}_{t_0+1}, \ldots, \boldsymbol{W}_t\big) \,, \qquad \forall t \in [\![t_0, T-1]\!] \,, \tag{8d}$$

$$\mathbb{E}\big[g(\boldsymbol{X}_T)\big] - b_{t_0} \leq 0 \,, \tag{8e}$$

with $b_{t_0} \in \mathbb{R}^m$. Again, Problem (8) is assumed to be well-defined.

Even under the Markovian Assumption 1, the presence of Constraint (8e) makes it difficult to write a dynamic programming equation for solving Problem (8) starting at time $t_0 = 0$. Indeed Constraint (8e) is not a pointwise constraint at the final stage T, so that we do not know how to incorporate it easily in the dynamic programming equation. In §3.3, using an indirect way of proceeding, we show that it is possible to obtain an optimal solution of Problem (8) starting at time $t_0 = 0$ in terms of feedback strategies depending on the state $\boldsymbol{X}_t$. But these feedbacks are implicitly parameterized by both the initial state x_0 and the constraint level b_0, so that they do not satisfy the time consistency property for a data set in which the parameter at time t is the initial state x_t (see §3.3 for further details).

A partial answer to the question of time consistency of the solution of Problem (8) starting at time $t_0 = 0$ has been given in [5]. Indeed, as detailed in [5], Problem (8) can be written in an equivalent way as a deterministic distributed optimal control problem in which the state variable is the probability distribution of $\boldsymbol{X}_t$, the dynamics of which is given by the Fokker-Planck equation. This deterministic problem can be solved by dynamic programming, which thus produces a sequence $(\Psi_0^\sharp, \ldots, \Psi_{T-1}^\sharp)$ of strategies (with the mapping Ψ_t defined over probability distributions on $\mathbb{X}_t$ and taking values in $\mathbb{U}_t$ [3]), which is time consistent for a data set in which the parameter at time t is the initial state probability distribution. But these optimal strategies depend on the specific value b_0 in the right-hand side of the expectation constraint, and thus have to be recomputed if this value changes. Moreover, the computation of the Bellman functions involves an infinite dimensional state, so that it is generally not tractable.

Our goal is to obtain a solution for Problem (8) starting at time $t_0 = 0$ which, on the one hand is computable in practice (that is, involves a finite dimensional state), and on the other hand is time consistent for a data set (to be specified) in which the parameter at time t consists of both the initial state x_t and the constraint level b_t. More precisely, we want to compute a solution for Problem (8) starting at time $t_0 = 0$ which is optimal for any value of both the initial state x_0 in (8b) and the final constraint level b_0 in (8e). Moreover, for any $t \in [\![0, T-1]\!]$, this solution after truncation has to be a universal solution (Definition 2) for the parameters (x_t, b_t) for Problem (8) starting at $t_0 = t$. As already explained in Remark 4, time consistency is not available for a solution in terms of random variables. We now present a reformulation of Problem (8) involving a finite dimensional state, whose solution in terms of state feedback strategies meets the goal described in this paragraph.

3.2 Formulation with martingale-type constraints

Following the same path as in [4], but in a discrete time context, we show that Problem (8) is equivalent to a multistage stochastic optimization problem subject to an almost sure contraint on the final state (see also [9, 10, 12]). For that purpose, we introduce a new state process $\boldsymbol{Z} = (\boldsymbol{Z}_0, \ldots, \boldsymbol{Z}_T)$ and a new control process $\boldsymbol{V} = (\boldsymbol{V}_0, \ldots, \boldsymbol{V}_{T-1})$. The random variables $\boldsymbol{Z}_t$ and $\boldsymbol{V}_t$ take their values respectively in spaces $\mathbb{V}_t$ and $\mathbb{Z}_t$, all identical to the space $\mathbb{Z}_T = \mathbb{R}^m$ where $\mathbb{R}^m$ is the codomain of the function g introduced at the beginning of §3.1: $\mathbb{V}_t = \mathbb{Z}_t = \mathbb{Z}_T = \mathbb{R}^m$ for all $t \in [\![0, T-1]\!]$. Now, we consider the optimization problem

starting at time t_0

$$\min_{(\boldsymbol{U},\boldsymbol{V},\boldsymbol{X},\boldsymbol{Z})} \mathbb{E}\left[\sum_{t=t_0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + K(\boldsymbol{X}_T)\right] , \tag{9a}$$

subject, for all $t = t_0, \ldots, T-1$, to dynamic constraints

$$\boldsymbol{X}_{t_0} = x_{t_0} , \qquad \boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) , \tag{9b}$$

$$\boldsymbol{Z}_{t_0} = z_{t_0} , \qquad \boldsymbol{Z}_{t+1} = \boldsymbol{Z}_t + \boldsymbol{V}_t , \tag{9c}$$

to measurability constraints

$$\sigma\big(\boldsymbol{U}_t\big) \subset \sigma\big(\boldsymbol{W}_{t_0+1}, \ldots, \boldsymbol{W}_t\big) , \tag{9d}$$

$$\sigma\big(\boldsymbol{V}_t\big) \subset \sigma\big(\boldsymbol{W}_{t_0+1}, \ldots, \boldsymbol{W}_{t+1}\big) , \tag{9e}$$

to martingale-type constraints

$$\boldsymbol{V}_t \text{ is integrable and } \mathbb{E}\big[\boldsymbol{V}_t \,\big|\, \sigma(\boldsymbol{W}_{t_0+1}, \ldots, \boldsymbol{W}_t)\big] = 0 , \tag{9f}$$

and to almost sure final constraint

$$g(\boldsymbol{X}_T) - \boldsymbol{Z}_T \leq 0 . \tag{9g}$$

Note that, in this formulation, the control variable $\boldsymbol{U}_t$ taken at time t does not depend on the noise $\boldsymbol{W}_{t+1}$ (*Decision–Hazard* framework), whereas the control variable $\boldsymbol{V}_t$, also taken at time t, does depend on the noise $\boldsymbol{W}_{t+1}$ (*Hazard–Decision* framework). At time t, the martingale-type constraint (9f) introduces a coupling between all the realizations of the decision random variable $\boldsymbol{V}_t$. In the sequel, for all $t \in [\![t_0{+}1, T]\!]$, we denote by $\mathcal{F}_{t_0:t}$ the σ-field generated by the sequence $(\boldsymbol{W}_{t_0+1}, \ldots, \boldsymbol{W}_t)$ of random variables:

$$\mathcal{F}_{t_0:t} = \sigma\big(\boldsymbol{W}_{t_0+1}, \ldots, \boldsymbol{W}_t\big) . \tag{10}$$

By convention, $\mathcal{F}_{t_0:t_0} = \{\emptyset, \Omega\}$.

3.2.1 Equivalence with the standard formulation

The link between the multistage stochastic optimization control problem (9) incorporating a martingale-type constraint and the initial problem (8) is given by the following proposition.

Proposition 5 *We suppose that the data of the Problems* (8) *and* (9) *are linked by*

$$b_{t_0} = z_{t_0} . \tag{11}$$

Then, Problem (8) *and Problem* (9) *are equivalent, in the sense that*

- *a solution* $(\boldsymbol{U}^\flat, \boldsymbol{V}^\flat, \boldsymbol{X}^\flat, \boldsymbol{Z}^\flat)$ *of Problem* (9) *can be deduced from a solution* $(\boldsymbol{U}^\sharp, \boldsymbol{X}^\sharp)$ *of Problem* (8)*,*
- *the two first components* $(\boldsymbol{U}^\flat, \boldsymbol{X}^\flat)$ *of a solution* $(\boldsymbol{U}^\flat, \boldsymbol{V}^\flat, \boldsymbol{X}^\flat, \boldsymbol{Z}^\flat)$ *of Problem* (9) *is a solution of Problem* (8)*.*

Proof. Let $(\boldsymbol{U}^\sharp, \boldsymbol{X}^\sharp)$ be a solution of Problem (8). We define the random processes $\boldsymbol{V}^\sharp$ and $\boldsymbol{Z}^\sharp$ by

$$\boldsymbol{V}_t^\sharp = \mathbb{E}\big[g(\boldsymbol{X}_T^\sharp) \,\big|\, \mathcal{F}_{t_0:t+1}\big] - \mathbb{E}\big[g(\boldsymbol{X}_T^\sharp) \,\big|\, \mathcal{F}_{t_0:t}\big] \;, \qquad \forall t \in [\![t_0, T-1]\!] \;, \tag{12a}$$

$$\boldsymbol{Z}_{t_0}^\sharp = z_{t_0} \;, \qquad \boldsymbol{Z}_{t+1}^\sharp = \boldsymbol{Z}_t^\sharp + \boldsymbol{V}_t^\sharp \;, \qquad \forall t \in [\![t_0, T-1]\!] \;. \tag{12b}$$

The random vector $\boldsymbol{V}_t^\sharp$ is well defined (hence so is $\boldsymbol{Z}_t^\sharp$) and is integrable. Indeed, as the function $g : \mathbb{X}_T \to \mathbb{R}_+^m$ is assumed nonnegative in §3.1 and using Inequality (8e), we get $0 \leq \mathbb{E}[g(\boldsymbol{X}_T^\sharp)] \leq b_{t_0}$. As $b_{t_0} \in \mathbb{R}^m$, the random vector $g(\boldsymbol{X}_T^\sharp)$ is integrable, hence $\boldsymbol{V}_t^\sharp$ in (12a) is the difference between two random vectors in $\mathbb{R}^m$, hence is well defined, and is integrable.

By construction, the two processes $\boldsymbol{V}^\sharp$ and $\boldsymbol{Z}^\sharp$ satisfy the constraints (9c)–(9e)–(9f). Moreover, we have that

$$\begin{aligned} \boldsymbol{Z}_T^\sharp &= z_{t_0} + \sum_{t=t_0}^{T-1} \boldsymbol{V}_t^\sharp && \text{(by (12b))} \\ &= z_{t_0} + g(\boldsymbol{X}_T^\sharp) - \mathbb{E}[g(\boldsymbol{X}_T^\sharp)] \;, && \text{(by telescoping sum using (12a))} \end{aligned}$$

and hence, by (8e) and (11), we get that

$$g(\boldsymbol{X}_T^\sharp) - \boldsymbol{Z}_T^\sharp = \mathbb{E}\big[g(\boldsymbol{X}_T^\sharp)\big] - z_{t_0} \leq b_{t_0} - z_{t_0} = 0 \;,$$

so that Constraint (9g) is also fulfilled. We deduce that $(\boldsymbol{U}^\sharp, \boldsymbol{V}^\sharp, \boldsymbol{X}^\sharp, \boldsymbol{Z}^\sharp)$ is admissible for Problem (9). Suppose that there would exist a solution $(\boldsymbol{U}^\flat, \boldsymbol{V}^\flat, \boldsymbol{X}^\flat, \boldsymbol{Z}^\flat)$ of Problem (9) with a strictly lower cost value than $(\boldsymbol{U}^\sharp, \boldsymbol{V}^\sharp, \boldsymbol{X}^\sharp, \boldsymbol{Z}^\sharp)$. From the dynamics (9c), we would have

$$\boldsymbol{Z}_T^\flat = z_{t_0} + \sum_{t=t_0}^{T-1} \boldsymbol{V}_t^\flat \;,$$

so that $\mathbb{E}[\boldsymbol{Z}_T^\flat] = z_{t_0}$ by repeated uses of (9f). Taking the expectation in (9g) would lead thus to

$$\mathbb{E}\big[g(\boldsymbol{X}_T^\flat)\big] \leq \mathbb{E}[\boldsymbol{Z}_T^\flat] = z_{t_0} = b_{t_0} \;.$$

Then, $(\boldsymbol{U}^\flat, \boldsymbol{X}^\flat)$ would be admissible for Problem (8) with a strictly lower cost value than the optimal solution $(\boldsymbol{U}^\sharp, \boldsymbol{X}^\sharp)$, which contradicts the assumed optimality of $(\boldsymbol{U}^\sharp, \boldsymbol{X}^\sharp)$. We conclude that $(\boldsymbol{U}^\sharp, \boldsymbol{V}^\sharp, \boldsymbol{X}^\sharp, \boldsymbol{Z}^\sharp)$ is an optimal solution of Problem (9).

Conversely, let $(\boldsymbol{U}^\flat, \boldsymbol{V}^\flat, \boldsymbol{X}^\flat, \boldsymbol{Z}^\flat)$ be an optimal solution of Problem (9). As shown in the first part of the proof, we have $\mathbb{E}\big[g(\boldsymbol{X}_T^\flat)\big] \leq b_{t_0}$, so that $(\boldsymbol{U}^\flat, \boldsymbol{X}^\flat)$ is admissible for Problem (8). Suppose that there would exist a solution $(\boldsymbol{U}^\sharp, \boldsymbol{X}^\sharp)$ of Problem (8) with a strictly lower cost value than $(\boldsymbol{U}^\flat, \boldsymbol{X}^\flat)$. Then, the quadruplet $(\boldsymbol{U}^\sharp, \boldsymbol{V}^\sharp, \boldsymbol{X}^\sharp, \boldsymbol{Z}^\sharp)$ obtained by constructing $\boldsymbol{V}^\sharp$ and $\boldsymbol{Z}^\sharp$ by (12)

would give a strictly lower cost value for Problem (9) than $(\boldsymbol{U}^\flat, \boldsymbol{V}^\flat, \boldsymbol{X}^\flat, \boldsymbol{Z}^\flat)$, which would be absurd. We conclude that $(\boldsymbol{U}^\flat, \boldsymbol{X}^\flat)$ is an optimal solution of Problem (8). □

Let us make a few comments about Problem (9).

- The nice features of the equivalent formulation (9) of Problem (8) are double. On the one hand, the initial constraint (8e) in expectation is replaced by an almost sure constraint (9g) on the final state, hence paving the way to use dynamic programming to solve Problem (9). On the other hand, the parameter defining the right-hand side of the constraint (8e) in expectation in formulation (8) becomes a component (9c) of the initial state in Problem (9), thus leading to time consistency as a consequence of dynamic programming.

- The conditional expectation $\mathbb{E}\big[g(\boldsymbol{X}_T) \,\big|\, \mathcal{F}_{t_0:t}\big]$ can be interpreted as the "perception of the risk constraint (8e)" at time t. From the very definition of $\boldsymbol{V}_t$, we have that

$$\boldsymbol{V}_t = \mathbb{E}\big[g(\boldsymbol{X}_T) \,\big|\, \mathcal{F}_{t_0:t+1}\big] - \mathbb{E}\big[g(\boldsymbol{X}_T) \,\big|\, \mathcal{F}_{t_0:t}\big] \;,$$

  from which we deduce that the control $\boldsymbol{V}_t$ corresponds to the variation of this perception between time t and time $t+1$. The additional state $\boldsymbol{Z}_{t+1}$ in (9c) is thus the cumulative variation of the risk contraint perception up to time $t+1$. Therefore, this new added state seems to be the minimal information which has to be added to the standard state in order to recover a dynamic programming principle.

- Nevertheless, Problem (9) is rather more intricate that Problem (8):

  1. there are additional state and control processes $\boldsymbol{Z}$ and $\boldsymbol{V}$,
  2. the new control variables $\boldsymbol{V}_t$ has to be searched in the *Hazard–Decision* framework as in (9e),
  3. a new expectation constraint (9f) on the controls $\boldsymbol{V}_t$ appears at each time step.

3.2.2 Extended dynamic programming equation and time consistency

The interest of Problem (9) is highlighted by the following theorem.

Theorem 6 *Suppose that the primitive noise random process $\boldsymbol{W} = (\boldsymbol{W}_1, \ldots, \boldsymbol{W}_T)$ takes a finite number of values, and that the following induction (Bellman equation)*

$$\mathcal{V}_T(x,z) = K(x) + \chi_{\{g(x)-z\leq 0\}}(x,z) \;, \tag{13a}$$

$$\mathcal{V}_t(x,z) = \min_u \min_{\substack{\sigma(\boldsymbol{V})\subset\sigma(\boldsymbol{W}_{t+1})\\ \mathbb{E}[\boldsymbol{V}]=0}} \mathbb{E}\Big[L_t(x,u,\boldsymbol{W}_{t+1}) + \mathcal{V}_{t+1}\big(f_t(x,u,\boldsymbol{W}_{t+1}), z+\boldsymbol{V}\big)\Big] \tag{13b}$$

is well-defined in the sense that all the functions $\mathcal{V}_t : \mathbb{X}_t \times \mathbb{Z}_t \to \mathbb{R}_+ \cup \{+\infty\}$ are measurable, for $t \in [\![0,T]\!]$.

Then, under Assumption 1, Problem (9) *starting at time $t_0 = 0$ can be solved by dynamic programming with associated Bellman equation* (13)*, and its optimal value is $\mathcal{V}_0(x_0, z_0)$.*

Proof. The proof of Theorem 6 in given in Appendix B. □

We deduce from Equation (13) that there is no loss of optimality in looking for the optimal control $\boldsymbol{U}_t$ at time t of Problem (9) as induced by a measurable mapping $\phi_t : \mathbb{X}_t \times \mathbb{Z}_t \to \mathbb{U}_t$, and for the optimal control $\boldsymbol{V}_{t+1}$ at time t as induced by a measurable mapping $\varphi_t : \mathbb{X}_t \times \mathbb{Z}_t \times \mathbb{W}_{t+1} \to \mathbb{V}_t$.

Let us embed Problem (9) starting at time $t_0 = 0$ in the framework developed in §2.1. The finite time span is $\mathcal{T} = [\![0, T-1]\!]$, the sequence of parameter sets is $\{\mathbb{X}_t \times \mathbb{Z}_t\}_{t\in\mathcal{T}}$, the sequences of control spaces are made of two sequences, $\{\mathcal{U}_t\}_{t\in\mathcal{T}}$ with $\mathcal{U}_t$ the space of measurable mappings defined on $\mathbb{X}_t \times \mathbb{Z}_t$ and taking values in $\mathbb{U}_t$, and $\{\mathcal{V}_t\}_{t\in\mathcal{T}}$ with $\mathcal{V}_t$ the space of measurable mappings defined on $\mathbb{X}_t \times \mathbb{Z}_t \times \mathbb{W}_{t+1}$ and taking values in $\mathbb{V}_t$. Notice that there exists an additional set $\mathbb{X}_T \times \mathbb{Z}_T$, where the final state of the system takes values, but that this set is not part of the sequence of parameter sets. The sequence of cost functions $\{J_t\}_{t\in\mathcal{T}}$, with

$$J_t \ : \mathbb{X}_t \times \mathbb{Z}_t \times \mathcal{U}_t \times \mathcal{V}_t \times \cdots \times \mathcal{U}_{T-1} \times \mathcal{V}_{T-1} \longrightarrow \mathbb{R} \,,$$

is defined by

$$\begin{aligned} J_t(x_t, z_t, \phi_t, \varphi_t, \ldots, \phi_{T-1}, \varphi_{T-1}) = \ \mathbb{E}\Bigg(& \sum_{\tau=t}^{T-1} L_\tau\big(\boldsymbol{X}_\tau, \phi_\tau(\boldsymbol{X}_\tau, \boldsymbol{Z}_\tau), \boldsymbol{W}_{\tau+1}\big) \\ & + K(\boldsymbol{X}_T) + \chi_{\{g(x)-z\leq 0\}}(\boldsymbol{X}_T, \boldsymbol{Z}_T)\Bigg) \,, \\ \text{with} \ \ \boldsymbol{X}_t = x_t \,, \ \ \boldsymbol{X}_{\tau+1} &= f_\tau\big(\boldsymbol{X}_\tau, \phi_\tau(\boldsymbol{X}_\tau, \boldsymbol{Z}_\tau), \boldsymbol{W}_{\tau+1}\big) \,, \quad \text{for } \tau \in [\![t, T-1]\!] \,, \\ \boldsymbol{Z}_t = z_t \,, \ \ \boldsymbol{Z}_{\tau+1} &= \boldsymbol{Z}_\tau + \varphi_\tau(\boldsymbol{X}_\tau, \boldsymbol{Z}_\tau, \boldsymbol{W}_{\tau+1}) \,, \quad \text{for } \tau \in [\![t, T-1]\!] \,, \end{aligned}$$

if $\mathbb{E}\Big[\varphi_\tau(\boldsymbol{X}_\tau, \boldsymbol{Z}_\tau, \boldsymbol{W}_{\tau+1}) \,\Big|\, \boldsymbol{X}_\tau, \boldsymbol{Z}_\tau\Big] = 0 \,, \ \forall \tau \in [\![t, T-1]\!]$, and by

$$J_t(x_t, z_t, \phi_t, \varphi_t, \ldots, \phi_{T-1}, \varphi_{T-1}) = +\infty \qquad \text{otherwise.}$$

From the optimization data set $\mathcal{E} = \big(\mathcal{T}, \{\mathbb{X}_t \times \mathbb{Z}_t\}_{t\in\mathcal{T}}, \{\mathcal{U}_t \times \mathcal{V}_t\}_{t\in\mathcal{T}}, \{J_t\}_{t\in\mathcal{T}}\big)$ and for a given $t \in \mathcal{T}$, we build, as in Definition 2, the family of optimization problems $\mathcal{P}_t^{\mathcal{E}} = \big\{\mathcal{P}_t^{\mathcal{E}}(x_t, z_t)\big\}_{(x_t,z_t)\in\mathbb{X}_t\times\mathbb{Z}_t}$, with Problem $\mathcal{P}_t^{\mathcal{E}}(x_t, z_t)$ being

$$\min_{\substack{(\phi_t,\ldots,\phi_{T-1})\in\mathcal{U}_t\times\cdots\times\mathcal{U}_{T-1} \\ (\varphi_t,\ldots,\varphi_{T-1})\in\mathcal{V}_t\times\cdots\times\mathcal{V}_{T-1}}} J_t(x_t, z_t, \phi_t, \varphi_t, \ldots, \phi_{T-1}, \varphi_{T-1}) \,. \tag{14}$$

Optimal strategies $(\phi_0^\sharp, \ldots, \phi_{T-1}^\sharp)$ and $(\varphi_0^\sharp, \ldots, \varphi_{T-1}^\sharp)$ obtained by solving for $t = 0$ Problem (14) using the dynamic programming equation (13) are such that, for any $t \in \mathcal{T}$, $(\phi_t^\sharp, \ldots, \phi_{T-1}^\sharp)$ and $(\varphi_t^\sharp, \ldots, \varphi_{T-1}^\sharp)$ is an optimal solution of Problem (14) for any initial state (x_t, z_t).

We deduce that solving Problem (9) dy dynamic programming fully answers the goal of time consistency enounced at the end of §3.1. Indeed, for all $t \in \mathcal{T}$, the subsequence of

optimal strategies $(\phi_t^\sharp, \dots, \phi_{T-1}^\sharp)$ is a universal solution for the family $\{\mathcal{P}_t^\mathcal{E}(x_t, z_t)\}_{(x_t,z_t)\in\mathbb{X}_t\times\mathbb{Z}_t}$ of optimization problems (14). By Proposition 5, a solution of Problem $\mathcal{P}_t^\mathcal{E}(x_t, z_t)$ induces a solution of Problem (8) starting at time $t_0 = t$ with initial state x_t and final constraint level $b_t = z_t$.

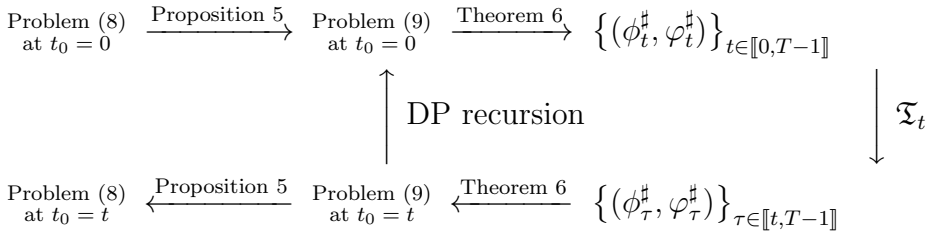

Figure 1: Links between the original problem (8) and the extended problem (9)

Moreover, the Bellman functions defined by (13) involve a finite dimensional state, so that their computation becomes tractable. We will provide a numerical illustration of this last point in Sect. 4.

3.3 Formulation with dualized constraint

We finish this section by presenting a way to solve Problem (8) using Lagrangian duality, and we show that the dualized problem does *not* display time consistency.

3.3.1 Dualized formulation

By dualizing the expectation constraint (8e) in Problem (8) with a given (fixed) multiplier $\lambda \in \mathbb{R}^m$, we obtain the following problem[2]

$$\min_{\boldsymbol{U},\boldsymbol{X}} \quad \mathbb{E}\left[\sum_{t=0}^{T-1} L_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\big) + K(\boldsymbol{X}_T) + \lambda \cdot g(\boldsymbol{X}_T)\right], \tag{15a}$$

$$\text{s.t.} \quad \boldsymbol{X}_0 = x_0\,, \tag{15b}$$

$$\boldsymbol{X}_{t+1} = f_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\big)\,,\ \ \forall t = 0, \dots, T-1\,, \tag{15c}$$

$$\sigma\big(\boldsymbol{U}_t\big) \subset \sigma\big(\boldsymbol{W}_1, \dots, \boldsymbol{W}_t\big)\,,\ \ \forall t = 0, \dots, T-1\,. \tag{15d}$$

By weak duality, we have that the optimal value of this problem is a lower bound of the optimal value of Problem (8) for any value λ. In some cases, Problems (15) and (8) are equivalent as specified by the following theorem.

[2]For the sake of simplicity, we suppose that the function $g : \mathbb{X}_T \to \mathbb{R}_+^m$, introduced at the beginning of §3.1, is bounded to ensure integrability in (15a)

Theorem 7 *Assume that there exists $\lambda_0^\sharp \in \mathbb{R}^m$ such that Problem* (15) *with $\lambda = \lambda_0^\sharp$ admits a solution $(\boldsymbol{U}^\sharp, \boldsymbol{X}^\sharp)$. We denote by $\bar{b}_0^{\lambda_0^\sharp}$ the expectation of $g(\boldsymbol{X}_T^\sharp)$ (which exists by Footnote 2)*

$$\bar{b}_0^{\lambda_0^\sharp} = \mathbb{E}[g(\boldsymbol{X}_T^\sharp)] \ ,$$

and we assume that the constraint level b_0 in Problem (8) *is such that $b_0 \in B_0^{\lambda_0^\sharp}$, with*

$$B_0^{\lambda_0^\sharp} = \big\{ b \in \mathbb{R}^m \ \ s.t. \ \ b_i = (\bar{b}_0^{\lambda_0^\sharp})_i \ \textit{if} \ (\lambda_0^\sharp)_i > 0 \\ \textit{and} \ \ b_i \geq (\bar{b}_0^{\lambda_0^\sharp})_i \ \textit{if} \ (\lambda_0^\sharp)_i = 0 \ , \ \ \forall i = 1, \ldots, m \big\} \ .$$

Then, the solution $(\boldsymbol{U}^\sharp, \boldsymbol{X}^\sharp)$ of Problem (15) *is also a solution of Problem* (8).

Proof. The result is a direct consequence of the extension of Everett's Theorem given in Appendix A. □

Remark 8 *The Everett argument that has been used here can be replaced by a more binding duality argument, namely, Problem* (8) *admits a saddle point $(\boldsymbol{U}^\sharp, \boldsymbol{X}^\sharp, \lambda_0^\sharp)$ when dualizing Constraint* (8e).

3.3.2 Discussion about time consistency

Since Problem (15) falls within the standard dynamic programming framework, there is no loss of optimality to look for the optimal controls $\boldsymbol{U}_t^\sharp$ of Problem (15), and hence of Problem (8), as feedback strategies $\phi_t^\sharp : \mathbb{X}_t \to \mathbb{U}_t$ depending on the state variable $\boldsymbol{X}_t$.

However, we do not claim that the optimal feedbacks $\phi_t^\sharp$ obtained by this argument have any specific properties in terms of time consistency. Indeed, assume as in Theorem 7 that there exists a $\lambda_0^\sharp$ such that $b_0 \in B_0^{\lambda_0^\sharp}$. The parameter $\lambda_0^\sharp$ implicitly depends on both the initial condition x_0 *and* the constraint level b_0, so that the optimal feedbacks $\phi_t^\sharp$ of Problem (15), which are parameterized by $\lambda_0^\sharp$, are accordingly implicitly parameterized by the pair (x_0, b_0) and therefore do not satisfy the property of time consistency. Moreover, if we write an optimization problem similar to Problem (15) starting at an initial time $t > 0$ with this value $\lambda_0^\sharp$

$$\min_{\boldsymbol{U}, \boldsymbol{X}} \quad \mathbb{E}\left[\sum_{\tau=t}^{T-1} L_\tau\big(\boldsymbol{X}_\tau, \boldsymbol{U}_\tau, \boldsymbol{W}_{\tau+1}\big) + K(\boldsymbol{X}_T) + \lambda_0^\sharp \cdot g(\boldsymbol{X}_T) \right] , \tag{16a}$$

$$\text{s.t.} \quad \boldsymbol{X}_t = x_t \ , \tag{16b}$$

$$\boldsymbol{X}_{\tau+1} = f_\tau\big(\boldsymbol{X}_\tau, \boldsymbol{U}_\tau, \boldsymbol{W}_{\tau+1}\big) \ , \ \ \forall \tau \in [\![t, T-1]\!] \ , \tag{16c}$$

$$\sigma\big(\boldsymbol{U}_\tau\big) \subset \sigma\big(\boldsymbol{W}_{t+1}, \ldots, \boldsymbol{W}_\tau\big) \ , \ \ \forall \tau = t, \ldots, T-1 \ , \tag{16d}$$

there is no reason that the optimal solution $(\boldsymbol{X}_t^\sharp, \ldots, \boldsymbol{X}_T^\sharp)$ of Problem (16) satisfies the constraint $\mathbb{E}[g(\boldsymbol{X}_T^\sharp)] \leq b_0$, that is, there is no reason to satisfy the relation $b_0 \in B_t^{\lambda_0^\sharp}$. Of course, it may exists some $\lambda_t^\sharp$ such that $b_0 \in B_t^{\lambda_t^\sharp}$, but usually $\lambda_t^\sharp \neq \lambda_0^\sharp$.

Thus, the sequence $(\phi_0^\sharp, \ldots, \phi_{T-1}^\sharp)$ of controls obtained by the dual approach, that is, by solving Problem (15), is not time consistent (in the sense of Definition 3).

4 Numerical experiments

We illustrate numerically whether time consistency holds true or not on a simple dam management problem developed in §4.1. In §4.2, we provide a numerical resolution of the problem with dualized expectation constraint (as seen in §3.3). In §4.3, we provide a numerical resolution by extended dynamic programming (as seen in §3.2).

4.1 A dam management problem

We consider here a basic dam model for a management problem. Let $T > 0$ denote a positive integer (horizon) and $[\![0, T]\!]$ be the optimization time span, and let $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ be a probability space. For any time t in $[\![0, T-1]\!]$, we consider the following real valued random variables:

- $\boldsymbol{X}_t$, the *water storage volume* in the dam at the beginning of time interval $[t, t+1)$,
- $\boldsymbol{U}_t$, the *decided amount of water to be turbinated* during time interval $[t, t+1)$, set at the beginning of the time interval $[t, t+1)$, and constrained to belong to an interval $[\underline{u}, \overline{u}]$,
- $\boldsymbol{W}_{t+1}$, the *amount of water inflow* in the dam during time interval $[t, t+1)$.

Let $\underline{x}$ (resp. $\overline{x}$) denotes the minimum (resp. maximum) water volume of the dam, and let x_0 be the dam volume at time 0. The decision $\boldsymbol{U}_t$ can be implemented only if there is enough water in the dam, that is, the turbinated water during a time interval cannot exceed the quantity of water present in the dam. Then, the *real amount of turbinated water* during the time interval $[t, t+1)$ is

$$\widetilde{\boldsymbol{U}}_{t+1} = \min\left\{\boldsymbol{U}_t, \boldsymbol{X}_t + \boldsymbol{W}_{t+1} - \underline{x}\right\} .$$

The maximal dam volume $\overline{x}$ is taken into account by accepting reservoir overflow: if the forthcoming water volume $\boldsymbol{X}_t + \boldsymbol{W}_{t+1} - \widetilde{\boldsymbol{U}}_t$ is greater than $\overline{x}$, then the dam water surplus $\boldsymbol{X}_t + \boldsymbol{W}_{t+1} - \widetilde{\boldsymbol{U}}_t - \overline{x}$ spills out. The dam dynamics is written accordingly:

$$\begin{aligned}
\boldsymbol{X}_0 &= x_0 \ , \\
\boldsymbol{X}_{t+1} &= \min\left\{\overline{x}, \boldsymbol{X}_t - \widetilde{\boldsymbol{U}}_{t+1} + \boldsymbol{W}_{t+1}\right\} = \min\left\{\overline{x}, \max\{\underline{x}, \boldsymbol{X}_t - \boldsymbol{U}_t + \boldsymbol{W}_{t+1}\}\right\} .
\end{aligned}$$

The turbinated water during the time interval $[t, t+1)$ produces electricity which is sold at a given price p_t. We assume that the sequence $(p_0, \dots, p_{T-1})$ of prices is deterministic. The dam revenue to be maximized is thus

$$\mathbb{E}\left[\sum_{t=0}^{T-1} p_t \widetilde{\boldsymbol{U}}_t\right] .$$

Compared to classical dam management model, we do not add an explicit final cost, but we rather constrain the final level of water in the dam at the end of the time span by a risk contraint. Indeed, we consider a probability constraint on the dam water volume at final time T, namely

$$\mathbb{P}\big[\boldsymbol{X}_T \geq \ell\big] \geq \pi ,$$

where the water level ℓ and the probability level π are given real numbers. Ultimately, the problem we have to solve is

$$\min_{\boldsymbol{U}, \boldsymbol{X}} \ \mathbb{E}\left[\sum_{t=0}^{T-1} -p_t \cdot \min\big\{\boldsymbol{U}_t, \boldsymbol{X}_t + \boldsymbol{W}_{t+1} - \underline{x}\big\}\right] , \tag{17a}$$

$$\text{s.t.} \quad \boldsymbol{X}_0 = x_0 , \tag{17b}$$

$$\boldsymbol{X}_{t+1} = \min\big\{\overline{x}, \max\{\underline{x}, \boldsymbol{X}_t - \boldsymbol{U}_t + \boldsymbol{W}_{t+1}\}\big\} , \tag{17c}$$

$$\underline{u} \leq \boldsymbol{U}_t \leq \overline{u} , \tag{17d}$$

$$\sigma\big(\boldsymbol{U}_t\big) \subset \sigma\big(\boldsymbol{W}_1, \dots, \boldsymbol{W}_t\big) , \tag{17e}$$

$$\mathbb{P}\big[\boldsymbol{X}_T \geq \ell\big] \geq \pi . \tag{17f}$$

We recall that the probability constraint (17f) can be rewritten as an expectation constraint

$$\pi - \mathbb{E}\big[\mathbf{1}_{\mathbb{R}^+}(\boldsymbol{X}_T - \ell)\big] \leq 0 , \tag{18}$$

where $\mathbf{1}_{\mathbb{R}^+} : \mathbb{R} \to \mathbb{R}$ is the Heaviside step function:

$$\mathbf{1}_{\mathbb{R}^+}(y) = \begin{cases} 0 & \text{if } y < 0 \\ 1 & \text{if } y \geq 0 \end{cases} .$$

We assume that the water inflows $\boldsymbol{W}_1, \dots, \boldsymbol{W}_T$ are independent random variables with a known probability distribution on the interval $[\underline{w}, \overline{w}]$.

To numerically solve Problem (17), we use the following parameter values: final time $T = 12$; initial state $x_0 = 10$; state bounds $[\underline{x}, \overline{x}] = [0, 20]$; control bounds $[\underline{u}, \overline{u}] = [0, 3]$; noise bounds $[\underline{w}, \overline{w}] = [0, 4]$; price sequence $p = (10, 10, 10, 8, 6, 4, 4, 4, 4, 6, 8, 10)$; final dam water level $\ell = 10$; required probability level $\pi = 0.9$. Moreover, we assume that the variables $\boldsymbol{X}_t$, $\boldsymbol{U}_t$ and $\boldsymbol{W}_t$ take discrete values within their respective bounds, with respective discretization steps equal to 0.1, 0.3 and 0.2. The optimization problem thus corresponds to the control of a discrete state space Markov chain. The discrete probability distribution of each random variable $\boldsymbol{W}_t$ is uniform. We represent on Figure 2 some selected trajectories of the noise process $(\boldsymbol{W}_1, \dots, \boldsymbol{W}_T)$ and the sequence of the deterministic prices p. These noise trajectories are used in the sequel to illustrate the behavior of the different optimization algorithms.

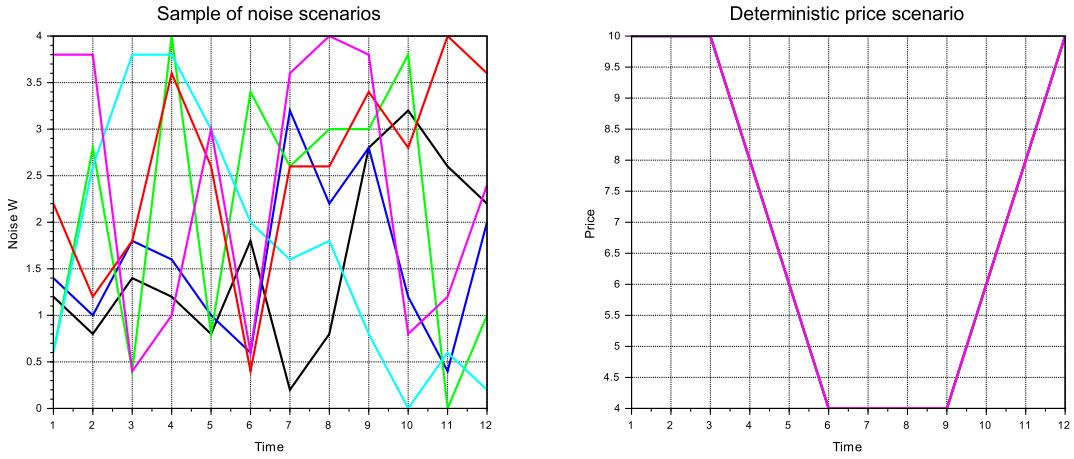

Figure 2: Samples of the noise process $\mathbf{W}$ (left) and price scenario p (right)

4.2 Resolution of the problem with dualized expectation constraint

The method used here to solve Problem (17) is based on the duality argument described in §3.3. The probability constraint (17f) is dualized with associated multiplier λ, which leads, in the cost function (17a), to an added term of the form:

$$\lambda\, \mathbb{E}\big[\pi - \mathbf{1}_{\mathbb{R}^+}(\boldsymbol{X}_T - \ell)\big] \ .$$

Instead of guessing the parameter value λ leading to the constraint level ℓ, we solve the dual problem

$$\max_{\lambda \geq 0} \ \ \varphi(\lambda) \ , \tag{19}$$

with

$$\varphi(\lambda) \ = \min_{\boldsymbol{U},\boldsymbol{X}} \mathbb{E}\left[\sum_{t=0}^{T-1} -p_t \cdot \widetilde{\boldsymbol{U}}_t \ - \ \lambda \cdot \mathbf{1}_{\mathbb{R}^+}(\boldsymbol{X}_T - \ell)\right] , \tag{20a}$$

$$\text{under constraints (17b)–(17c)–(17d)–(17e).} \tag{20b}$$

The Uzawa algorithm consists in maximizing the dual function φ using a projected gradient algorithm. At iteration k of the algorithm, knowing the value $\lambda^{(k)}$ of the multiplier, we perform the three following steps.

- Compute $\varphi(\lambda^{(k)})$, that is, solve the minimization problem (20) with $\lambda = \lambda^{(k)}$; this minimization is performed using dynamic programming (1-dimensional state variable), hence furnishing optimal feedbacks $\phi_t^{(k+1)}$ and optimal state variables $\boldsymbol{X}_t^{(k+1)}$.

- Compute the probability distribution of the random variable $\boldsymbol{X}_T^{(k+1)}$ by integrating the dynamics (17b)-(17c) using the optimal feedbacks $\phi_t^{(k+1)}$, and then compute the value of the constraint $\mathbb{P}[\boldsymbol{X}_T^{(k+1)} \geq \ell]$.

- Update the multiplier λ by a projected gradient step:

$$\lambda^{(k+1)} = \mathrm{proj}_{\mathbb{R}^+}\left(\lambda^{(k)} + \rho\big(\pi - \mathbb{P}[\boldsymbol{X}_T^{(k+1)} \geq \ell]\big)\right) .$$

For the problem under consideration, and despite the potential nonconvexity induced by the final probability constraint (17f), the Uzawa algorithm converges in about 10 iterations, leading to an optimal multiplier value $\lambda^\sharp = 62.405$. Once the algorithm has converged, we obtain the optimal feedback sequence $\{\phi_t^\sharp\}_{t=0,\dots,T-1}$ by solving Problem (20) by dynamic programming with $\lambda = \lambda^\sharp$. Then, we simulate the dynamics of the dam along some noise trajectories using these optimal feedbacks $\phi_t^\sharp$. The results given in Table 1 are obtained by simulating 10,000 noise trajectories, and illustrates the adequacy between optimization and simulation.

Uzawa optimization	**Monte Carlo simulation**
Bellman value at $t = 0$: -188.90	Monte Carlo cost: -188.94
Required probability: 0.900	Estimated probability: 0.903

Table 1: Optimization and simulation for the duality method

On Figure 3, we represent the dam water level and control trajectories over $[0, T]$ obtained by simulating with the optimal feedbacks $\phi_t^\sharp$ along the noise trajectories depicted on Figure 2. We observe that the optimization "gives up" for certain trajectories (the lowest one to the left of Figure 3) to reach the final level ℓ appearing in the constraint in probability: we turbine as much water as possible, leaving the state evolve towards the minimum level $\underline{x}$. This observation is conform to the expected behavior of an optimization problem with a probability constraint.

Finally, we can use the optimal feedbacks $\phi_t^\sharp$ to simulate the dam starting at any initial time $t_i > 0$ from any given initial state x_{t_i}. For example starting at time $t_i = 3$ from the initial state $x_{t_i} = 5$ and simulating the optimal feedbacks $\phi_t^\sharp$ along 10,000 scenarios leads to the results given in Table 2. As expected, the final constraint level reached in this last simulation is not equal to the required level $\pi = 0.9$, which illustrates that time consistency does not hold true for the problem formulation with dualized constraint.

Monte Carlo simulation
Monte Carlo cost: -92.04
Estimated probability: 0.833

Table 2: Simulation for the duality method starting from $t_i = 3$ and $x_{t_i} = 5$

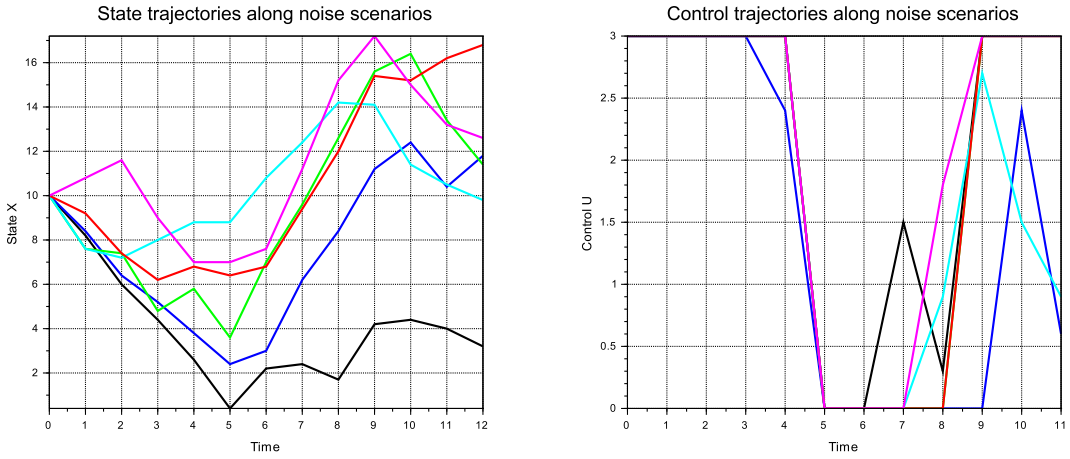

Figure 3: Optimal state $\mathbf{X}$ and control $\mathbf{U}$ trajectories over $[0, T]$ obtained by the duality method

4.3 Resolution by extended dynamic programming

We now use the equivalent formulation of Problem (17) incorporating an additional state process $\boldsymbol{Z} = (\boldsymbol{Z}_0, \ldots, \boldsymbol{Z}_T)$, an additional control process $\boldsymbol{V} = (\boldsymbol{V}_0, \ldots, \boldsymbol{V}_{T-1})$ and an almost sure contraint on the final state. As it has been explained in §3.2, the expression of the new dynamics, here 1-dimensional, is

$$\boldsymbol{Z}_0 = z_0 \ , \qquad \boldsymbol{Z}_{t+1} = \boldsymbol{Z}_t + \boldsymbol{V}_t \ ,$$

and the form of the final constraint is

$$-\mathbf{1}_{\mathbb{R}^+}(\boldsymbol{X}_T - \ell) - \boldsymbol{Z}_T \leq 0 \ .$$

The equivalent problem for the case study under consideration is

$$\min_{\boldsymbol{U},\boldsymbol{V},\boldsymbol{X},\boldsymbol{Z}} \quad \mathbb{E}\left[\sum_{t=0}^{T-1} -p_t \cdot \min\left\{\boldsymbol{U}_t, \boldsymbol{X}_t + \boldsymbol{W}_{t+1} - \underline{x}\right\}\right] , \tag{21a}$$

$$\text{s.t.} \quad \boldsymbol{X}_0 = x_0 \ , \quad \boldsymbol{X}_{t+1} = \min\left\{\overline{x}, \max\{\underline{x}, \boldsymbol{X}_t - \boldsymbol{U}_t + \boldsymbol{W}_{t+1}\}\right\} , \tag{21b}$$

$$\boldsymbol{Z}_0 = z_0 \ , \quad \boldsymbol{Z}_{t+1} = \boldsymbol{Z}_t + \boldsymbol{V}_t \ , \tag{21c}$$

$$\underline{u} \leq \boldsymbol{U}_t \leq \overline{u} \ , \tag{21d}$$

$$\sigma(\boldsymbol{U}_t) \subset \sigma(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_t) \ , \tag{21e}$$

$$\sigma(\boldsymbol{V}_t) \subset \sigma(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_{t+1}) \ , \tag{21f}$$

$$\mathbb{E}\left[\boldsymbol{V}_t \,\middle|\, \sigma(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_t)\right] = 0 \ , \tag{21g}$$

$$\mathbf{1}_{\mathbb{R}^+}(\boldsymbol{X}_T - \ell) + \boldsymbol{Z}_T \geq 0 \ . \tag{21h}$$

From Proposition 5, we have that Problems (17) and (21) are equivalent under the condition

$$z_0 = -\pi \ .$$

Moreover, the special form (21h) of the final constraint makes it possible to bound the variables $\boldsymbol{V}_t$ and $\boldsymbol{Z}_t$. Indeed, from the proof of Proposition 5, we deduce from the expression (12a) of the optimal control $\boldsymbol{V}_t^\sharp$ with $g = \mathbf{1}_{\mathbb{R}^+}$ that it is sufficient to search for the control $\boldsymbol{V}_t$ in $[-1, 1]$. Moreover, the optimal state $\boldsymbol{Z}_t^\sharp$ being obtained by a telescoping sum, it is sufficient to search for the state $\boldsymbol{Z}_t$ in $[z_0 - 1, z_0 + 1]$. Problem (21) can be solved by dynamic programming with the extended state variable $(\boldsymbol{X}_t, \boldsymbol{Z}_t)$, which corresponds to a dynamic programming equation with a 2-dimensional state variable. Then, we simulate the dynamics of the dam — using the same 10,000 noise trajectories as those previously used to obtain Table 1 — with the optimal feedbacks given by dynamic programming with the extended state variable $(\boldsymbol{X}_t, \boldsymbol{Z}_t)$. The associated results are given in Table 3. We observe a good adequacy between optimization and simulation, and we observe that the costs are pretty much identical between Table 1 and Table 3.

Extended Dynamic Programming	**Monte Carlo Simulation**
Bellman value at $t = 0$: -188.67	Monte Carlo cost: -187.47
Initial state z_0: -0.900	Estimated probability: 0.896

Table 3: Optimization and simulation for the extended dynamic programming method

Some simulation trajectories are represented on Figure 4 and Figure 5. Figure 4 gives the same information (dam water level and control trajectories over $[0, T]$) as the one presented for the duality based algorithm in §4.2, whereas Figure 5 depicts trajectories of the optimal state process $\boldsymbol{Z}$ and the optimal control process $\boldsymbol{V}$. We observe that the results depicted on Figure 3 (duality method) and on Figure 4 (extended dynamic programming method) are very close (with tiny differences induced by numerical resolution), which illustrates the equivalence between Problem (17) and Problem (21).

Finally, we can use the optimal feedbacks obtained when solving Problem (21) by dynamic programming to simulate the dam starting at any initial time $t_i > 0$ from any given initial state (x_{t_i}, z_{t_i}). For example starting at time $t_i = 3$ from the initial state $(x_{t_i}, z_{t_i}) = (5, -0.9)$ — using the same values than those used in §4.2 — and simulating the optimal feedbacks along 10,000 scenarios leads to the results given in Table 4: the final probability level to be reached is by construction equal to $\pi = 0.9$, and the Monte Carlo simulation induces a very similar level of probability, which numerically illustrates that the time consistency property is fulfilled. The associated simulation trajectories are represented on Figure 6, using the same noises trajectories as those that had been used to obtain Figure 3.

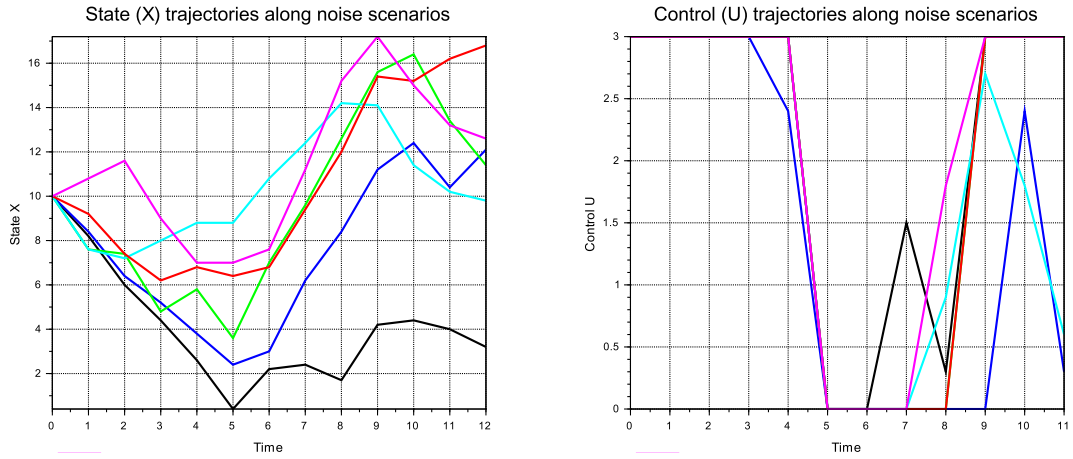

Figure 4: Optimal state X and control U trajectories over $[0, T]$ obtained by the extended dynamic programming method

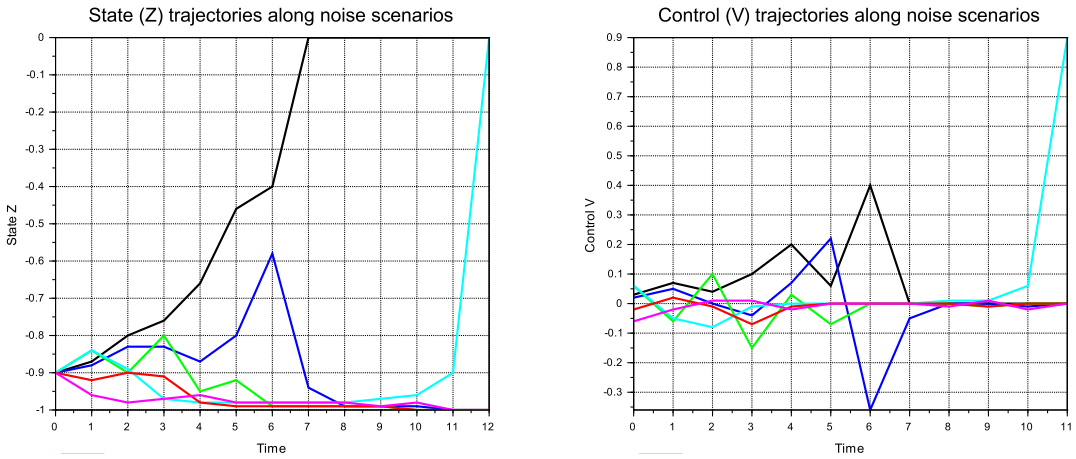

Figure 5: Optimal state Z and control V trajectories over $[0, T]$ obtained by the extended dynamic programming method

Extended Dynamic Programming	**Monte Carlo Simulation**
Bellman value at $t = 3$: -87.78	Monte Carlo cost: -87.71
Initial state (x_3, z_3): $(5, -0.900)$	Estimated probability: 0.902

Table 4: Optimization and simulation for the extended dynamic programming method

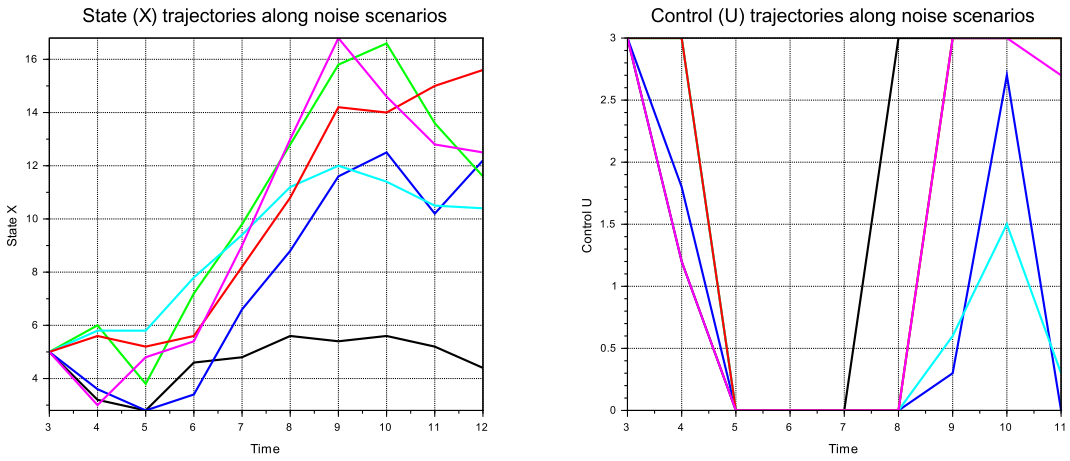

Figure 6: Optimal state X and control U trajectories over $[3, T]$ obtained by the extended dynamic programming method

5 Conclusion

In this paper, we have proposed a formal definition of time consistency for families of optimization problems, by introducing the notion of universal solution. With this, we have shown that — for the class of problems where risk is modeled in the form of constraints in probability or in expectation — the property of time consistency depends on the notion of state that one chooses, which must be suited to the problem studied. In particular, we have shown that, even if the "right" notion of state for the class of multistage stochastic optimization problems with a final expectation state constraint was of infinite dimension (the conditional probability distribution of the state), it is possible to display a state of finite dimension, so that solving the problem by dynamic programming becomes conceivable again.

Acknowledgements: This paper builds upon results obtained by Pierre Girardeau during his PhD thesis [9], supervised by the three authors.

A An extension of Everett's theorem

A result due to Everett (see [8]) links the solution of an optimization problem under constraint and the one of the related dualized optimization problem. We give here a slight extension, which relaxes an assumption of Everett's theorem.

Let $\mathbb{U}$ be a set and let U^{ad} be a subset of $\mathbb{U}$. Let $\Theta : \mathbb{U} \to \mathbb{R}^m$ be a (mulivalued) function. We deal with the following optimization problem

$$J^\sharp = \min_{u \in U^{\mathrm{ad}}} J(u) \quad \text{s.t.} \quad \Theta(u) - b \leq 0 \in \mathbb{R}^m \,, \tag{22}$$

hence subject to a finite number m of inequality constraints.

Theorem 9 *Let $\lambda \in \mathbb{R}^m_+$ be given. We consider $\overline{u}^\lambda$, solution of the optimization problem*

$$\min_{u \in U^{\mathrm{ad}}} J(u) + \big\langle \lambda,\, \Theta(u) \big\rangle \,, \tag{23}$$

and we set $\overline{b}^\lambda = \Theta(\overline{u}^\lambda)$. We introduce the set $B^\lambda \subset \mathbb{R}^m$ defined by

$$B^\lambda = \Big\{ b \in \mathbb{R}^m \,\Big|\, b_i = \overline{b}^\lambda_i \ \textit{if } \lambda_i > 0 \ \ \textit{and} \ \ b_i \geq \overline{b}^\lambda_i \ \textit{if } \lambda_i = 0 \,, \ \ \forall i = 1, \ldots, m \Big\} \,. \tag{24}$$

Then, a solution $\overline{u}^\lambda$ of Problem (23) is a solution of Problem (22) for any $b \in B^\lambda$.

Proof. Let $\overline{u}^\lambda$ be a solution of Problem (23) and let $b \in B^\lambda$. We have

$$\begin{aligned}
J(\overline{u}^\lambda) &= \min_{u \in U^{\mathrm{ad}}} J(u) + \big\langle \lambda,\, \Theta(u) - \Theta(\overline{u}^\lambda) \big\rangle \,, && \text{(by definition of } \overline{u}^\lambda) \\
&= \min_{u \in U^{\mathrm{ad}}} J(u) + \big\langle \lambda,\, \Theta(u) - \overline{b}^\lambda \big\rangle \,, && \text{(by definition of } \overline{b}^\lambda = \Theta(\overline{u}^\lambda)) \\
&= \min_{u \in U^{\mathrm{ad}}} J(u) + \big\langle \lambda,\, \Theta(u) - b \big\rangle \,, && \text{(by definition of } B^\lambda \text{ in (24))} \\
&\leq \sup_{\mu \geq 0} \min_{u \in U^{\mathrm{ad}}} J(u) + \big\langle \mu,\, \Theta(u) - b \big\rangle \,, && \text{(as } \lambda \geq 0) \\
&\leq \inf_{u \in U^{\mathrm{ad}}} \sup_{\mu \geq 0} J(u) + \big\langle \mu,\, \Theta(u) - b \big\rangle \,, && \text{(by weak duality)} \\
&= J^\sharp \,. && \text{(by (22))}
\end{aligned}$$

Since $\overline{u}^\lambda$ is such that $\Theta(\overline{u}^\lambda) = \overline{b}^\lambda \leq b$ by definition of the set B^λ in (24), we deduce that $\overline{u}^\lambda$ is admissible for Problem (22), and hence is an optimal solution of this problem. □

B Dynamic programming for the optimization problem involving martingale-type constraints

We prove in §B.2 that Problem (9) can be solved by dynamic programming under the additional assumption that, for any time t in $[\![0, T]\!]$, the random variable $\boldsymbol{W}_t$ can take only a finite number of values. The proof is based on a so-called interchange (between minimization and integration) lemma given in §B.1.

B.1 An interchange Lemma

Lemma 10 *Let $\mathbb{Y}$, $\mathbb{U}$, $\mathbb{V}$, $\mathbb{W}'$ and $\mathbb{W}''$ be measurable spaces, and let $\varphi : \mathbb{Y} \times \mathbb{U} \times \mathbb{V} \times \mathbb{W}'' \to \mathbb{R}_+ \cup \{+\infty\}$ be a measurable extended real function. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space[3].*

[3] All random variables are defined on $(\Omega, \mathcal{F}, \mathbb{P})$, and we denote them using bold letters.

Given three random variables $\boldsymbol{Y}$, $\boldsymbol{W}'$ and $\boldsymbol{W}''$ taking values in $\mathbb{Y}$, $\mathbb{W}'$ and $\mathbb{W}''$ respectively, we consider the optimization problem $\mathcal{P}$ defined by

$$\mathcal{V}_{\mathcal{P}}[\boldsymbol{Y}] = \inf_{(\boldsymbol{U},\boldsymbol{V})} \mathbb{E}\big[\varphi(\boldsymbol{Y},\boldsymbol{U},\boldsymbol{V},\boldsymbol{W}'')\big] \tag{25a}$$

$$s.t. \quad \sigma(\boldsymbol{U}) \subset \sigma(\boldsymbol{W}') \ , \ \ \sigma(\boldsymbol{V}) \subset \sigma(\boldsymbol{W}',\boldsymbol{W}'') \ , \tag{25b}$$

$$\boldsymbol{V} \text{ is integrable and } \mathbb{E}\big[\boldsymbol{V} \,\big|\, \sigma(\boldsymbol{W}')\big] = 0 \ , \tag{25c}$$

where the minimization is done over couples of random variables $\boldsymbol{U} : \Omega \to \mathbb{U}$ and $\boldsymbol{V} : \Omega \to \mathbb{V}$.

We define the function $\psi : \mathbb{Y} \to \mathbb{R}_+ \cup \{+\infty\}$ by

$$\psi : \mathbb{Y} \ni y \mapsto \inf_{(u,\boldsymbol{V})} \mathbb{E}\big[\varphi(y,u,\boldsymbol{V},\boldsymbol{W}'')\big] \tag{26}$$

$$s.t. \quad \sigma(\boldsymbol{V}) \subset \sigma(\boldsymbol{W}'') \ ,$$

$$\boldsymbol{V} \text{ is integrable and } \mathbb{E}\big[\boldsymbol{V}\big] = 0 \ .$$

where the minimization is done over variables $u \in \mathbb{U}$ and random variables $\boldsymbol{V} : \Omega \to \mathbb{V}$.

We suppose that the two random variables $\boldsymbol{W}'$ and $\boldsymbol{W}''$ are independent, that they each take a finite number of values, and that the random variable $\boldsymbol{Y}$ is $\sigma(\boldsymbol{W}')$-measurable, that is, $\sigma(\boldsymbol{Y}) \subset \sigma(\boldsymbol{W}')$. Then, the optimal value $\mathcal{V}_{\mathcal{P}}[\boldsymbol{Y}]$ of Problem $\mathcal{P}$ satisfies the following interchange formula

$$\mathcal{V}_{\mathcal{P}}[\boldsymbol{Y}] = \mathbb{E}\big[\psi(\boldsymbol{Y})\big] \ . \tag{27}$$

Proof. Letting $\{w'_i\}_{i\in[\![0,N']\!]}$ and $\{w''_i\}_{i\in[\![0,N'']\!]}$ be the sets of values taken by, respectively, the random variables $\boldsymbol{W}'$ and $\boldsymbol{W}''$, we denote

$$\boldsymbol{W}' : \Omega \to \{w'_i\}_{i\in[\![0,N']\!]} \quad \text{with} \quad \mathbb{P}\big[\{\boldsymbol{W} = w'_i\}\big] = \pi'_i \ , \ \ \forall i \in [\![0,N']\!] \ , \tag{28}$$

$$\boldsymbol{W}'' : \Omega \to \{w''_i\}_{i\in[\![0,N'']\!]} \quad \text{with} \quad \mathbb{P}\big[\{\boldsymbol{W} = w''_i\}\big] = \pi''_i \ , \ \ \forall i \in [\![0,N'']\!] \ . \tag{29}$$

Now, since $\boldsymbol{Y}$ is a $\sigma(\boldsymbol{W}')$-measurable random variable and from the measurability constraints on the random variables $\boldsymbol{U}$ and $\boldsymbol{V}$ in Equation (25b), we can represent these random variables as follows:

$$\boldsymbol{Y} = \sum_{i=0}^{N'} y_i \mathbf{1}_{w'_i}(\boldsymbol{W}') \ , \ \ \boldsymbol{U} = \sum_{i=0}^{N'} u_i \mathbf{1}_{w'_i}(\boldsymbol{W}') \ , \ \ \boldsymbol{V} = \sum_{i=0}^{N'} \sum_{j=0}^{N''} v_{i,j} \mathbf{1}_{w'_i}(\boldsymbol{W}') \mathbf{1}_{w''_j}(\boldsymbol{W}'') \ . \tag{30}$$

We have just expressed the fact that the set of $\sigma(\boldsymbol{W})$-measurable random variables taking values in a set $\mathbb{F}$ is in bijection with the product space $\mathbb{F}^N$ if the random variable $\boldsymbol{W}$ takes N different values.

We start the proof by using Equation (30) to establish the following equalities

$$\mathbb{E}\big[\varphi(\boldsymbol{Y},\boldsymbol{U},\boldsymbol{V},\boldsymbol{W}'') \,\big|\, \boldsymbol{W}'\big] = \sum_{i=0}^{N'} \Big(\sum_{j=0}^{N''} \pi''_j \varphi\Big(y_i, u_i, v_{i,j}, w''_j\Big)\Big) \mathbf{1}_{w'_i}(\boldsymbol{W}') \ , \tag{31a}$$

$$\mathbb{E}\big[\varphi(\boldsymbol{Y},\boldsymbol{U},\boldsymbol{V},\boldsymbol{W}'')\big] = \sum_{i=0}^{N'} \pi'_i \Big(\sum_{j=0}^{N''} \pi''_j \varphi\Big(y_i, u_i, v_{i,j}, w''_j\Big)\Big) \ , \tag{31b}$$

$$\mathbb{E}\big[\boldsymbol{V} \,\big|\, \boldsymbol{W}'\big] = \sum_{i=0}^{N'} \Big(\sum_{j=0}^{N''} \pi''_j v_{i,j} \Big) \mathbf{1}_{w'_i}(\boldsymbol{W}') \ . \tag{31c}$$

All the manipulations below are easy to check, and are justified because all quantities take extended nonnegative values.

- For Equation (31a):

$$\begin{aligned}
&\mathbb{E}\big(\varphi(\boldsymbol{Y},\boldsymbol{U},\boldsymbol{V},\boldsymbol{W''}) \,\big|\, \boldsymbol{W'}\big) \\
&\quad= \mathbb{E}\Big[\varphi\Big(\sum_{i=0}^{N'} y_i \mathbf{1}_{w_i'}(\boldsymbol{W'}), \sum_{i=0}^{N'} u_i \mathbf{1}_{w_i'}(\boldsymbol{W'}), \sum_{i=0}^{N'}\sum_{j=0}^{N''} v_{i,j}\mathbf{1}_{w_i'}(\boldsymbol{W'})\mathbf{1}_{w_j''}(\boldsymbol{W''}), \boldsymbol{W''}\Big) \,\Big|\, \boldsymbol{W'}\Big] \qquad \text{(by (30))} \\
&\quad= \mathbb{E}\Big[\sum_{i=0}^{N'} \varphi\Big(y_i, u_i, \sum_{j=0}^{N''} v_{i,j}\mathbf{1}_{w_j''}(\boldsymbol{W''}), \boldsymbol{W''}\Big)\mathbf{1}_{w_i'}(\boldsymbol{W'}) \,\Big|\, \boldsymbol{W'}\Big] \\
&\quad= \sum_{i=0}^{N'} \mathbb{E}\Big[\varphi\Big(y_i, u_i, \sum_{j=0}^{N''} v_{i,j}\mathbf{1}_{w_j''}(\boldsymbol{W''}), \boldsymbol{W''}\Big) \,\Big|\, \boldsymbol{W'}\Big]\mathbf{1}_{w_i'}(\boldsymbol{W'}) \\
&\quad= \sum_{i=0}^{N'} \mathbb{E}\Big[\varphi\Big(y_i, u_i, \sum_{j=0}^{N''} v_{i,j}\mathbf{1}_{w_j''}(\boldsymbol{W''}), \boldsymbol{W''}\Big)\Big]\mathbf{1}_{w_i'}(\boldsymbol{W'}) \qquad \text{(as } \boldsymbol{W'} \text{ and } \boldsymbol{W''} \text{ are independent)} \\
&\quad= \sum_{i=0}^{N'} \mathbb{E}\Big[\sum_{j=0}^{N''} \varphi\Big(y_i, u_i, v_{i,j}, w_j''\Big)\mathbf{1}_{w_j''}(\boldsymbol{W''})\Big]\mathbf{1}_{w_i'}(\boldsymbol{W'}) \\
&\quad= \sum_{i=0}^{N'} \Big(\sum_{j=0}^{N''} \pi_j'' \varphi\Big(y_i, u_i, v_{i,j}, w_j''\Big)\Big)\mathbf{1}_{w_i'}(\boldsymbol{W'})\,. \qquad \text{(by (29))}
\end{aligned}$$

- For Equation (31b):

$$\begin{aligned}
\mathbb{E}\big[\varphi(\boldsymbol{Y},\boldsymbol{U},\boldsymbol{V},\boldsymbol{W''})\big] &= \mathbb{E}\Big[\mathbb{E}\big[\varphi(\boldsymbol{Y},\boldsymbol{U},\boldsymbol{V},\boldsymbol{W''}) \,\big|\, \boldsymbol{W'}\big]\Big] \\
&= \mathbb{E}\Big[\sum_{i=0}^{N'} \Big(\sum_{j=0}^{N''} \pi_j'' \varphi\Big(y_i, u_i, v_{i,j}, w_j''\Big)\Big)\mathbf{1}_{w_i'}(\boldsymbol{W'})\Big] \qquad \text{(by (31a))} \\
&= \sum_{i=0}^{N'} \pi_i' \Big(\sum_{j=0}^{N''} \pi_j'' \varphi\Big(y_i, u_i, v_{i,j}, w_j''\Big)\Big)\,. \qquad \text{(by (28))}
\end{aligned}$$

- For Equation (31c):

$$\begin{aligned}
\mathbb{E}\big[\boldsymbol{V} \,\big|\, \boldsymbol{W'}\big] &= \mathbb{E}\Big[\sum_{i=0}^{N'}\sum_{j=0}^{N''} v_{i,j}\mathbf{1}_{w_i'}(\boldsymbol{W'})\mathbf{1}_{w_j''}(\boldsymbol{W''}) \,\Big|\, \boldsymbol{W'}\Big] \qquad \text{(by (30))} \\
&= \sum_{i=0}^{N'} \mathbb{E}\Big[\sum_{j=0}^{N''} v_{i,j}\mathbf{1}_{w_j''}(\boldsymbol{W''}) \,\Big|\, \boldsymbol{W'}\Big]\mathbf{1}_{w_i'}(\boldsymbol{W'}) \\
&= \sum_{i=0}^{N'} \mathbb{E}\Big[\sum_{j=0}^{N''} v_{i,j}\mathbf{1}_{w_j''}(\boldsymbol{W''})\Big]\mathbf{1}_{w_i'}(\boldsymbol{W'}) \qquad \text{(as } \boldsymbol{W'} \text{ and } \boldsymbol{W''} \text{ are independent)} \\
&= \sum_{i=0}^{N'} \Big(\sum_{j=0}^{N''} \pi_j'' v_{i,j}\Big)\mathbf{1}_{w_i'}(\boldsymbol{W'})\,. \qquad \text{(by (29))}
\end{aligned}$$

Using Equations (30) and (31c), for any $\sigma(\boldsymbol{W}'')$-measurable random variable $\boldsymbol{V}$, we have the equivalence

$$\mathbb{E}\big[\boldsymbol{V} \,\big|\, \boldsymbol{W}'\big] = 0 \iff \sum_{j=0}^{N''} \pi''_j v_{i,j} = 0 \;, \quad \forall i \in [\![0, N']\!] \;. \tag{35}$$

Using again Equations (30) and (31), we obtain that the optimization Problem (25) is equivalent to the following optimization problem

$$\inf_{\substack{\{u_i\}_{i\in[\![0,N']\!]} \\ \{v_{i,j}\}_{i\in[\![0,N']\!],j\in[\![0,N'']\!]}}} \quad \sum_{i=0}^{N'} \pi'_i \Big(\sum_{j=0}^{N''} \pi''_j \varphi(y_i, u_i, v_{i,j}, w''_i) \Big) \tag{36a}$$

$$\text{s.t.} \quad \sum_{j=0}^{N''} \pi''_j v_{i,j} = 0 \;, \quad \forall i \in [\![0, N']\!] \;. \tag{36b}$$

The optimization problem (36) trivially splits into a family $\{\mathcal{P}_i\}_{i\in[\![0,N']\!]}$ of N' independent optimization problems, Problem $\mathcal{P}_i$ being defined by

$$\mathcal{V}_{\mathcal{P}_i}[y_i] = \inf_{(u_i, \{v_{i,j}\}_{j\in[\![0,N'']\!])} \sum_{j=0}^{N''} \pi''_j \varphi(y_i, u_i, v_{i,j}, w''_i) \tag{37a}$$

$$\text{s.t.} \quad \sum_{j=0}^{N''} \pi''_j v_{i,j} = 0 \;, \tag{37b}$$

and the value of Problem (36) is the weighted sum of the values of the family of problems $\{\mathcal{P}_i\}_{i\in[\![0,N']\!]}$:

$$\mathcal{V}_{\mathcal{P}}[\boldsymbol{Y}] = \sum_{i=0}^{N'} \pi'_i \mathcal{V}_{\mathcal{P}_i}[y_i] \;.$$

We notice that $\mathcal{V}_{\mathcal{P}_i}[y_i]$ in (37) is exactly $\psi(y_i)$ in (26), so that the above equation gives (27). □

B.2 Proof of Theorem 6

Proof. For any $\tau \in [\![0, T]\!]$ we consider the minimization Problem $\mathcal{P}_\tau$ defined by[4]

$$\min_{(\boldsymbol{U},\boldsymbol{V},\boldsymbol{X},\boldsymbol{Z})} \mathbb{E}\left[\sum_{t=0}^{\tau-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + \mathcal{V}_\tau(\boldsymbol{X}_\tau, \boldsymbol{Z}_\tau) \right] , \tag{38a}$$

$$\text{s.t. } \boldsymbol{X}_0 = x_0 \;, \;\; \boldsymbol{Z}_0 = z_0 \;, \tag{38b}$$

$$\text{and, for all } t \in [\![0, \tau-1]\!],$$

$$\boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) \;, \;\; \boldsymbol{Z}_{t+1} = \boldsymbol{Z}_t + \boldsymbol{V}_t \;, \tag{38c}$$

$$\sigma\big(\boldsymbol{U}_t\big) \subset \sigma\big(\boldsymbol{W}_1, \dots, \boldsymbol{W}_t\big) \;, \;\; \sigma\big(\boldsymbol{V}_t\big) \subset \sigma\big(\boldsymbol{W}_1, \dots, \boldsymbol{W}_{t+1}\big) \;, \tag{38d}$$

$$\boldsymbol{V}_t \text{ is integrable and } \mathbb{E}\big[\boldsymbol{V}_t \,\big|\, \boldsymbol{W}_1, \dots, \boldsymbol{W}_t\big] = 0 \;, \tag{38e}$$

[4]For $\tau = 0$, the value of Problem $\mathcal{P}_0$ is simply $\mathcal{V}_0(x_0, z_0)$.

where the sequence $\{\mathcal{V}_\tau\}_{\tau\in[\![0,T]\!]}$ of value functions, with $\mathcal{V}_\tau : \mathbb{X}_\tau \times \mathbb{Z}_\tau \to \mathbb{R}_+ \cup \{+\infty\}$, appearing in the cost function (38a) is given by the Bellman recursion (13). To simplify the notation, we denote by Λ_τ the set of random variables $(\boldsymbol{U}_t, \boldsymbol{V}_t)_{t\in[\![0,\tau]\!]}$ and $(\boldsymbol{X}_t, \boldsymbol{Z}_t)_{t\in[\![0,\tau+1]\!]}$ satisfying the constraints (38b) – (38c) – (38d) – (38e). We recall that, by Equation (10), $\mathcal{F}_\tau$ represents the σ-field generated by $(\boldsymbol{W}_1, \dots, \boldsymbol{W}_\tau)$ for all $\tau \in [\![1, T]\!]$.

We are now going to prove, by backward induction, that the value of Problem (9) with $t_0 = 0$ is equal to the value of Problem $\mathcal{P}_\tau$ in (38) for any $\tau \in [\![0, T]\!]$.

First, the value of Problem (9), with $t_0 = 0$, is equal to the value of Problem $\mathcal{P}_\tau$ in (38) for $\tau = T$. Indeed, the criterion (38a) in Problem (38), satisfies, for $\tau = T$,

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + \mathcal{V}_T(\boldsymbol{X}_T, \boldsymbol{Z}_T)\right] \\
&= \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + \big(K(x) + \chi_{\{g(x)-z\le 0\}}\big)(\boldsymbol{X}_T, \boldsymbol{Z}_T)\right] \qquad \text{(as } \mathcal{V}_T \text{ is defined by (13a))} \\
&= \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + K(\boldsymbol{X}_T)\right] \text{ under the almost sure final constraint} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad g(\boldsymbol{X}_T) - \boldsymbol{Z}_T \le 0\,.
\end{aligned}
$$

Thus, we obtain that Problem (38) for $\tau = T$ is the same as Problem (9) with $t_0 = 0$, the only difference being that the almost sure final constraint (9g) has been moved in the final cost in (38a).

Second, we prove by backward induction that the value of Problem (9) is equal to the value of Problem $\mathcal{P}_\tau$ for any $\tau \in [\![0, T]\!]$. For this purpose, assuming that the value of Problem (9) is equal to the value of Problem $\mathcal{P}_{\tau+1}$, we prove that it is also equal to the value of Problem $\mathcal{P}_\tau$. We immediately get that

$$
\begin{aligned}
&\min_{\big((\boldsymbol{U}_t,\boldsymbol{V}_t)_{t\in[\![0,\tau]\!]},(\boldsymbol{X}_t,\boldsymbol{Z}_t)_{t\in[\![0,\tau+1]\!]}\big)\in\Lambda_\tau} \mathbb{E}\left[\sum_{t=0}^{\tau} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + \mathcal{V}_{\tau+1}(\boldsymbol{X}_{\tau+1}, \boldsymbol{Z}_{\tau+1})\right] \\
&= \min_{\big((\boldsymbol{U}_t,\boldsymbol{V}_t)_{t\in[\![0,\tau-1]\!]},(\boldsymbol{X}_t,\boldsymbol{Z}_t)_{t\in[\![0,\tau]\!]}\big)\in\Lambda_{\tau-1}} \left(\mathbb{E}\left[\sum_{t=0}^{\tau-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1})\right]\right. \qquad (39) \\
&\qquad \left. + \min_{\substack{(\boldsymbol{U}_\tau,\boldsymbol{V}_\tau)\\ \sigma(\boldsymbol{U}_\tau)\subset\mathcal{F}_\tau\\ \sigma(\boldsymbol{V}_\tau)\subset\mathcal{F}_{\tau+1}\\ \mathbb{E}[\boldsymbol{V}_\tau\,|\,\mathcal{F}_\tau]=0}} \mathbb{E}\Big[L_\tau(\boldsymbol{X}_\tau, \boldsymbol{U}_\tau, \boldsymbol{W}_{\tau+1}) + \mathcal{V}_{\tau+1}\big(f_\tau(\boldsymbol{X}_\tau, \boldsymbol{U}_\tau, \boldsymbol{W}_{\tau+1}), \boldsymbol{Z}_\tau + \boldsymbol{V}_\tau\big)\Big]\right)\,,
\end{aligned}
$$

because all quantities take extended nonnegative values.

Now, we apply Lemma 10 to the inner minimization, with $\boldsymbol{Y} = (\boldsymbol{X}_\tau, \boldsymbol{Z}_\tau)$, $\boldsymbol{W}' = (\boldsymbol{W}_1, \dots, \boldsymbol{W}_\tau)$, $\boldsymbol{W}'' = \boldsymbol{W}_{\tau+1}$ and with the function $\varphi\big((x,z),u,v,w''\big) = L_\tau(x,u,w'') + \mathcal{V}_{\tau+1}\big(f_\tau(x,u,w''),\, z+v\big)$ and deduce that

$$
\min_{\substack{(\boldsymbol{U}_\tau,\boldsymbol{V}_\tau)\\ \sigma(\boldsymbol{U}_\tau)\subset\mathcal{F}_\tau\\ \sigma(\boldsymbol{V}_\tau)\subset\mathcal{F}_{\tau+1}\\ \mathbb{E}[\boldsymbol{V}_\tau\,|\,\mathcal{F}_\tau]=0}} \mathbb{E}\Big[L_\tau(\boldsymbol{X}_\tau, \boldsymbol{U}_\tau, \boldsymbol{W}_{\tau+1}) + \mathcal{V}_{\tau+1}\big(f_\tau(\boldsymbol{X}_\tau, \boldsymbol{U}_\tau, \boldsymbol{W}_{\tau+1}), \boldsymbol{Z}_\tau + \boldsymbol{V}_\tau\big)\Big] = \mathbb{E}\big[\mathcal{V}_\tau(\boldsymbol{X}_\tau, \boldsymbol{Z}_\tau)\big]\,,
$$

because $\psi(y)$ in (26) is exactly $\mathcal{V}_\tau(x,z)$ in (13b). Combined with Equation (39), this leads to

$$\min_{\left((\boldsymbol{U}_t,\boldsymbol{V}_t)_{t\in[\![0,\tau]\!]},(\boldsymbol{X}_t,\boldsymbol{Z}_t)_{t\in[\![0,\tau+1]\!]}\right)\in\Lambda_\tau} \mathbb{E}\left[\sum_{t=0}^{\tau} L_t(\boldsymbol{X}_t,\boldsymbol{U}_t,\boldsymbol{W}_{t+1}) + \mathcal{V}_{\tau+1}(\boldsymbol{X}_{\tau+1},\boldsymbol{Z}_{\tau+1})\right]$$

$$= \min_{\left((\boldsymbol{U}_t,\boldsymbol{V}_t)_{t\in[\![0,\tau-1]\!]},(\boldsymbol{X}_t,\boldsymbol{Z}_t)_{t\in[\![0,\tau]\!]}\right)\in\Lambda_{\tau-1}} \mathbb{E}\left[\sum_{t=0}^{\tau-1} L_t(\boldsymbol{X}_t,\boldsymbol{U}_t,\boldsymbol{W}_{t+1}) + \mathcal{V}_\tau\big(\boldsymbol{X}_\tau,\boldsymbol{Z}_\tau\big)\right].$$

We conclude that the value of Problem (9) is equal to the value of Problem $\mathcal{P}_\tau$, so that we have by induction that the value of Problem (9) is equal to the value of Problem $\mathcal{P}_\tau$ in (38) for any $\tau \in [\![0,T]\!]$.

The value of Problem (9) is thus equal to the value of problem $\mathcal{P}_0$, namely $\mathcal{V}_0(x_0,z_0)$, and can therefore be computed by using the dynamic programming equation (13). □

References

[1] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku. Coherent multiperiod risk-adjusted values and Bellman's principle. *Annals of Operations Research*, 152(1):5–22, July 2007.

[2] D. P. Bertsekas. *Dynamic Programming and Optimal Control, Vol. I.* Athena Scientific, Belmont, Massachusets, second edition, 2005.

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

[4] B. Bouchard, R. Elie, and N. Touzi. Stochastic target problems with controlled loss. *SIAM Journal on Control and Optimization*, 48(5):3123–3150, 2009.

[5] P. Carpentier, J.-P. Chancelier, G. Cohen, M. De Lara, and P. Girardeau. Dynamic consistency for stochastic optimal control problems. *Annals of Operation Research*, 200(1):247–263, 2012.

[6] P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, 11(3):57–106, 2006.

[7] K. Detlefsen and G. Scandolo. Conditional and dynamic convex risk measures. *Finance and Stochastics*, 9(4):539–561, October 2005.

[8] H. Everett. Generalized Lagrange multiplier method for solving problems of optimum allocation of resources. *Operations research*, 11:399–417, 1963.

[9] P. Girardeau. *Résolution de grands problèmes en optimisation stochastique dynamique et synthèse de lois de commande.* PhD thesis, Université Paris-Est, december 2010. `http://www.theses.fr/2010PEST1026`.

[10] G. Granato. *Optimisation de lois de gestion énergétiques des véhicules hybrides*. PhD thesis, École Polytechnique, december 2012.

[11] P. J. Hammond. Consistent plans, consequentialism, and expected utility. *Econometrica*, 57(6):1445–1449, 1989.

[12] L. Pfeiffer. Two approaches to stochastic optimal control problems with a final-time expectation constraint. *Appl Math Optim*, 77:377–404, 2018.

[13] G. C. Pflug and A. Pichler. Time-inconsistent multistage stochastic programs: Martingale bounds. *European Journal of Operational Research*, 249(1):155–163, 2016.

[14] F. Riedel. Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112(2):185 – 200, 2004.

[15] A. Ruszczynski. Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming*, 125:235–261, 2010.

[16] A. Shapiro. On a time consistency concept in risk averse multistage stochastic programming. *Operations Research Letters*, 37(3):143 – 147, 2009.