HAL open science

Differentiability and Regularization of Parametric Convex Value Functions in Stochastic Multistage Optimization

Adrien Le Franc, Jean-Philippe Chancelier, Pierre Carpentier, Michel de Lara

To cite this version:

Adrien Le Franc, Jean-Philippe Chancelier, Pierre Carpentier, Michel de Lara. Differentiability and Regularization of Parametric Convex Value Functions in Stochastic Multistage Optimization. ENPC - École des Ponts ParisTech. 2022. hal-03903218v1

HAL Id: hal-03903218

https://hal.science/hal-03903218v1

Submitted on 16 Dec 2022 (v1), last revised 11 Jan 2023 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Differentiability and Regularization of Parametric Convex Value Functions in Stochastic Multistage Optimization

Adrien Le Franc* Jean-Philippe Chancelier* Michel De Lara*
Pierre Carpentier†

December 16, 2022

Abstract

In multistage decision problems, it is often the case that an initial strategic decision (such as investment) is followed by many operational ones (operating the investment). Such initial strategic decision can be seen as a parameter affecting a multistage decision problem. More generally, we study in this paper a standard multistage stochastic optimization problem depending on a parameter. When the parameter is fixed, Stochastic Dynamic Programming provides a way to compute the optimal value of the problem. Thus, the value function depends both on the state (as usual) and on the parameter. Our aim is to investigate on the possibility to efficiently compute gradients of the value function with respect to the parameter, when these objects exist. When nondifferentiable, we propose a regularization method based on the Moreau-Yosida envelope. We present a numerical test case from day-ahead power scheduling.

Keywords Stochastic multistage optimization; Dynamic Programming; Marginal function; Differentiability; Moreau-Yosida regulatization

1 Introduction

We consider optimization problems where an *upstream* decision is made in the first place, which stands for a parameter for a *downstream* multistage stochastic optimization problem. Our work is motivated by applications in the field of energy planning, where such decision structure arises naturally. As a typical example, the regulatory rules considered in [32, 22, 25] impose renewable power plants to commit a day-ahead power production

*CERMICS, École des Ponts Marne-la-Vallée, France
†UMA, ENSTA Paris, IP Paris, Palaiseau, France

profile, upstream to the intraday management phase where costs are subject to uncertainties arising from power production. Another important application arises when dealing with large-scale stochastic multistage optimization problems. In many cases, dualizing some coupling constraint allows for decomposition in subproblems, each corresponding to a "small" stochastic multistage optimization problem. The Lagrange multiplier associated with the coupling constraint has then to be optimized, and can be interpreted as a parameter for the multistage stochastic optimization subproblems (see [10, 12] for details).

In this article, we propose a standard formulation for *parametric multistage stochastic optimization problems* (PMSOP). In the formulation we outline, the parameter does not affect the dynamics but affects all instantaneous (and final) costs at all periods, and also possibly the admissible control sets; due to this structure, the parameter cannot be identified with the initial decision of a multistage stochastic optimization problem. When the value of the parameter is fixed, Stochastic Dynamic Programming (see e.g. [7, 27]) is a way to obtain the value of the downstream problem by computing the value functions given by the Bellman equation. Thus, the value functions now depend both on the state (as usual) and on the parameter. On top of that, we investigate on the possibility to efficiently compute additional first-order information, e.g. gradients of the value functions with respect to the parameter, when these objects exist. Our end goal is to formulate first-order oracles which let us enter the world of (primal) first-order optimization methods (see [3] for a recent survey) to solve PMSOPs, that is, to perform optimization with respect to the parameter.

Of course, the interest in such kinds of problems is not new. The reference textbook of Bonnans and Shapiro [9] gathers numerous results on the value functions of a parameterized optimization problem. In the context of multistage stochastic programming, the sensitivity analysis of the value of a downstream problem with respect to some model parameters has been already studied in [13, 17, 33]. These works mainly focus on the computation of directional derivatives of the value function, in the case where the stage cost functions of the problem are affine. In [13], the authors further argue that — since Danskin's Theorem tells us that the value function is locally Lipschitz continuous — the value function is differentiable almost everywhere by Rademacher's Theorem. A similar conclusion is drawn in [33], and formulas to compute the gradient at points where the value function is differentiable are given in both references. However, employing smooth optimization methods to minimize nondifferentiable functions can yield suboptimal solutions, even in the convex case, as illustrated by the example discussed in [3, §8.1.2].

Thus, our work differs from the above references in at least two points: (*i*) we study the existence and provide formulas for the gradient of a parametric value function, whereas previous works concentrate on directional derivatives; (*ii*) we consider convex nonlinear stage costs and constraints, going beyond the usual linear multistage stochastic programming framework. Also, to cover cases where the parametric value functions are convex but nondifferentiable, we propose a regularization method based on the Moreau-Yosida envelope [21, 34]. We study the convergence properties of both the regularized value functions that we introduce, and of the parameter solutions of a PMSOP, as our regularization coefficient tends to zero. Although Moreau-Yosida regularization has been previously employed

in optimal control problems (see e.g. [5]), few anterior studies concentrate on stochastic problems, except recently in [23], where the authors examine the case of a discounted infinite horizon Markov decision process. We share common interests with the latter reference, but our finite horizon context leads us to follow different paths. Finally, we propose a concurrent method and an assessment technique to evaluate the quality of a parameter as a solution to a PMSOP. For both purposes, we rely on the stochastic dual dynamic programming algorithm (SDDP), introduced by Pereira and Pinto [24], and analyzed further by Shapiro [30] and Philpott [26, 16].

The paper is organized as follows. First, in Sect. 2, we introduce the definition of a PMSOP and of parametric value functions. Second, in Sect. 3, we introduce Bellman-like equations to compute the gradient of differentiable convex value functions with respect to their parameter argument. To extend the method to convex nondifferentiable parametric value functions, we propose a regularization scheme, and study the convergence properties of resulting regularized parametric value functions. Finally, in Sect. 4, we present a numerical test case inspired from day-ahead power scheduling. Our assessment method aims at evaluating the efficiency of first-order oracles based on the gradient computation technique introduced in this work.

Background notions and notations

Natural numbers. We use the notation $[\![i, j]\!] = \{i, i+1, \ldots, j-1, j\}$ for any pair of natural numbers such that $i \leq j$.
Probability. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We use bold capital letters, e.g. $\mathbf{Z}$, to denote random variables, and denote by $\sigma(\mathbf{Z})$ the σ-algebra on Ω ($\sigma(\mathbf{Z}) \subset \mathcal{F}$) generated by the random variable $\mathbf{Z}$ and by $\mathrm{supp}(\mathbf{Z})$ the support of $\mathbf{Z}$. Besides, for a topological space $\mathbb{X}$, we denote by $\mathcal{B}(\mathbb{X})$ its Borel σ-field.
Functional analysis and topology. We introduce the extended real line $\overline{\mathbb{R}} = [-\infty, +\infty]$, and we denote $\mathbb{R}_+ = [0, +\infty[$ and $\mathbb{R}_+^* =]0, +\infty[$. Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ be a function. The effective domain $\mathrm{dom} f$ is the set $\left\{x \in \mathbb{X} \,\middle|\, f(x) < +\infty\right\}$, and the function f is said to be proper if $f > -\infty$ and $\mathrm{dom} f \neq \emptyset$. For any subset $X \subseteq \mathbb{X}$, $\delta_X : \mathbb{X} \to \overline{\mathbb{R}}$ denotes the indicator function of the set X: $\delta_X(x) = 0$ if $x \in X$, and $\delta_X(x) = +\infty$ if $x \notin X$. For a topological space $\mathbb{X}$ and $X \subseteq \mathbb{X}$, we recall that the interior $\mathrm{int}(X)$ is defined as the largest open set contained in X.

2 Parametric multistage stochastic optimization problems

In §2.1, we introduce a standard formulation for PMSOPs. Then, in §2.2, we introduce parametric value functions defined by the Bellman equations, and we discuss the specific role of the parameter.

2.1 Problem formulation

We are interested in solving problems of the form

$$\min_{p \in \mathcal{P}_{\mathrm{ad}}} \Phi(p) \ , \tag{1a}$$

in the case where the objective function Φ in (1a) is the value of the following *parametric multistage stochastic optimization problem* (PMSOP):

$$\Phi(p) = \inf_{\mathbf{U}_0, \ldots, \mathbf{U}_{T-1}} \mathbb{E}\Big[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}, p) + K(\mathbf{X}_T, p)\Big] \ , \tag{1b}$$

$$\mathbf{X}_0 = x_0 \ , \tag{1c}$$

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

$$\sigma(\mathbf{U}_t) \subseteq \sigma(\mathbf{W}_1, \ldots, \mathbf{W}_t) \ , \ \ \forall t \in [\![0, T-1]\!] \ . \tag{1e}$$

We now comment on all terms in Problem (1), and we provide measurability[1] and integrability assumptions so that the expected value in (1b) is well defined. We consider a *discrete time span*

$$[\![0, T]\!] = \{0, 1, \ldots, T-1, T\} \ , \tag{2}$$

with *horizon* a natural number $T \in \mathbb{N}^*$. Concerning the *upstream problem* (1a), the variable

$$p \in \mathbb{R}^{n_p} \ , \tag{3}$$

where $n_p \in \mathbb{N}^*$, is a *parameter* which may be chosen in the *parameter set*

$$\mathcal{P}_{\mathrm{ad}} \subseteq \mathbb{R}^{n_p} \ . \tag{4}$$

The parameter p in (3) affects the multistage Problem (1b)–(1e). Concerning the *downstream problem* (1b)–(1e), we introduce the random variables

$$\mathbf{X}_t : (\Omega, \mathcal{F}, \mathbb{P}) \to \big(\mathbb{R}^{n_x}, \mathcal{B}(\mathbb{R}^{n_x})\big) \ , \ \ \forall t \in [\![0, T]\!] \ , \tag{5a}$$

$$\mathbf{U}_t : (\Omega, \mathcal{F}, \mathbb{P}) \to \big(\mathbb{R}^{n_u}, \mathcal{B}(\mathbb{R}^{n_u})\big) \ , \ \ \forall t \in [\![0, T-1]\!] \ , \tag{5b}$$

$$\mathbf{W}_t : (\Omega, \mathcal{F}, \mathbb{P}) \to \big(\mathbb{R}^{n_w}, \mathcal{B}(\mathbb{R}^{n_w})\big) \ , \ \ \forall t \in [\![1, T]\!] \ , \tag{5c}$$

which denote respectively the *state*, *control* and *noise* variables of Problem (1b)–(1e), taking values in real Euclidean spaces of respective finite dimensions $(n_x, n_u, n_w) \in \mathbb{N}^{*3}$. The state variables are initialized by $x_0 \in \mathbb{R}^{n_x}$ and evolve in (1d) according to the measurable *dynamics*

$$f_t : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \to \mathbb{R}^{n_x} \ , \ \ \forall t \in [\![0, T-1]\!] \ . \tag{6}$$

[1] When not explicit, measurability is to be understood as with respect to the underlying Borel σ-fields.

Note that the parameter p does not affect the dynamics f_t, and that the constraints (1c)–(1d) are almost sure (a.s.) constraints. The control variables are constrained by the *nonanticipativity* constraints (1e). Last, the criterion to be minimized in (1b) is the expected value of the sum of the measurable *parametric stage costs*

$$L_t : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \times \mathbb{R}^{n_p} \to \left]-\infty, +\infty\right] , \quad \forall t \in [\![0, T-1]\!] , \tag{7a}$$

with a measurable *parametric final cost*

$$K : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \to \left]-\infty, +\infty\right] . \tag{7b}$$

Remark 2.1 *The cost functions can take values in* $\left]-\infty, +\infty\right]$ *to offer the possibility to implicitly encode constraints through effective domains. In practice, explicit constraints of the form* $\mathbf{U}_t \in \mathcal{U}_t(\mathbf{X}_t, p)$ *a.s. — with set-valued mappings* $\mathcal{U}_t : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \rightrightarrows \mathbb{R}^{n_u}$*, for all* $t \in [\![0, T-1]\!]$ *— can be added in Problem* (1) *(see the example in Sect. 4).*

The mathematical expectation (1b) is well defined. Indeed, as we focus our attention on the differentiability properties of the function Φ, and on the computation of the gradient $\nabla\Phi$ — when it exists — in what follows, we simplify integrability questions. For this purpose, we consider proper cost functions — that never take the value $-\infty$ — and discrete noise variables (5c) — so that all expectations reduce to finite sums of numbers that belong to $\left]-\infty, +\infty\right]$. Questioning whether our results extend to continuous noise variables could be the subject of a following research work.

Remark 2.2 *The formulation of the nonanticipativity constraint in* (1e) *corresponds to problems which formulate naturally in the* decision-hazard *information structure. In particular, the first decision is deterministic, with* $\sigma(\mathbf{U}_0) = \{\emptyset, \Omega\}$.

2.2 Parametric value functions

First, we introduce parametric value functions, defined by the Bellman equations. Second, we discuss the role of the parameter in a PMSOP.

The Bellman equations. As mentioned in §2.1, we consider discrete random variables (5c) in Problem (1). Besides, we make the following (discrete) white noise assumption.

Assumption 2.3 (discrete white noise) *The sequence* $\{\mathbf{W}_t\}_{t \in [\![1,T]\!]}$ *of noise variables in* (5c) *is stagewise independent, and each noise variable* $\mathbf{W}_t$ *has a finite support.*

The above assumption has a direct consequence on the solutions of the multistage Problem (1b)-(1e). Indeed, if we consider a fixed value of the parameter $p \in \mathbb{R}^{n_p}$ in (3), we retrieve a standard *multistage stochastic optimization problem.* Therefore, Stochastic

Dynamic Programming gives us a method for computing the optimal solution of the multistage Problem (1b)-(1e), and thus to evaluate $\Phi(p)$ in (1a). We introduce the sequence $\{V_t\}_{t\in[\![0,T]\!]}$ of *parametric value functions*, defined by backward induction by the *Bellman equations*

$$V_T(x,p) = K(x,p)\ ,\ \ \forall (x,p) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_p}\ , \tag{8a}$$

$$V_t(x,p) = \inf_{u\in\mathbb{R}^{n_u}} \mathbb{E}\Big[L_t(x,u,\mathbf{W}_{t+1},p) + V_{t+1}\big(f_t(x,u,\mathbf{W}_{t+1}),p\big)\Big]\ , \tag{8b}$$
$$\forall (x,p) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_p}\ ,\ \ \forall t \in [\![0,T-1]\!]\ .$$

Under Assumption 2.3, the value functions $\{V_t\}_{t\in[\![0,T]\!]}$ in (8) give the optimal value of the multistage Problem (1b)-(1e), in the sense that

$$\Phi(p) = V_0(x_0,p)\ ,\ \ \forall p \in \mathbb{R}^{n_p}\ . \tag{9}$$

We refer to Bertsekas [7] and Puterman [27] for a comprehensive presentation of the Stochastic Dynamic Programming method.

Since parametric value functions are defined as the infimum of a certain criterion in the Bellman equations, they are "marginal functions", a class of function of rich properties [9]. To ease applications to our context, we introduce the (parametric) *Q-functions*

$$Q_t(x,u,p) = \mathbb{E}\Big[L_t(x,u,\mathbf{W}_{t+1},p) + V_{t+1}\big(f_t(x,u,\mathbf{W}_{t+1}),p\big)\Big]\ , \tag{10a}$$
$$\forall (x,u,p) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p}\ ,\ \ \forall t \in [\![0,T-1]\!]\ ,$$

so that the parametric value functions $\{V_t\}_{t\in[\![0,T-1]\!]}$ formulate explicitly as marginal functions:

$$V_t(x,p) = \inf_{u\in\mathbb{R}^{n_u}} Q_t(x,u,p)\ ,\ \ \forall (x,p) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_p}\ ,\ \ \forall t \in [\![0,T-1]\!]\ . \tag{10b}$$

For the same reason, we also introduce the (possibly empty) parametric *solution sets*

$$\mathcal{U}_t^*(x,p) = \arg\min_{u\in\mathbb{R}^{n_u}} Q_t(x,u,p)\ ,\ \ \forall (x,p) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_p}\ ,\ \ \forall t \in [\![0,T-1]\!]\ . \tag{11}$$

The role of the parameter. The parameter p, as well as the control u_0, can be seen as an initial decision as they are both constant. However, whereas u_0 only appears at time $t = 0$, p appears in the cost functions at all times. This is why the parameter p can not be seen as an initial control variable in a stochastic control formulation, amenable to dynamic programming.

By contrast, it is clear from the definitions of Problem (1) and of the parametric value functions $\{V_t\}_{t\in[\![0,T-1]\!]}$ in (8) that the parameter p in (3) could be treated as a state, by reformulating the state variable as

$$\tilde{\mathbf{X}}_t = (\mathbf{X}_t, p)\ ,\ \ \forall t \in [\![0,T]\!]\ . \tag{12}$$

However, we wish to avoid this strategy, due to the exponential growth of Stochastic Dynamic Programming's complexity with respect to the dimension of the state space (termed curse of dimensionality: see [4]). Therefore, rather than computing a global approximation of the objective function Φ in Problem (1), we intend to build an efficient first-order oracle, defined as a mapping

$$p \mapsto (\Phi(p), \nabla\Phi(p)) \ . \tag{13}$$

This mapping returns the value of the objective function Φ together with the value of the gradient of the objective (when it exists) to apply iterative optimization steps with respect to the parameter p. For this reason, we intentionally keep the parameter apart from the state variables.[2]

3 Gradient of a convex parametric value function

We recall that the properties of the objective function Φ in (1) are inherited from those of the parametric value functions $\{V_t\}_{t\in[\![0,T]\!]}$ defined in (8) through $\Phi = V_0(x_0, \cdot)$ in (9), under the discrete white noise Assumption 2.3.

First, in §3.1, we provide sufficient conditions for the gradients $\{\nabla_p V_t\}_{t\in[\![0,T]\!]}$ to be well-defined, and introduce a method to compute such gradients by backward induction. Second, in §3.2, we propose a regularization method to handle cases where the parametric value functions $\{V_t\}_{t\in[\![0,T]\!]}$ are nondifferentiable with respect to the parameter p. Third, in §3.3, we study the convergence properties of such regularized value functions.

In this section, we consider two different situations.

- In §3.1, the effective domain of the function Φ, named $\mathcal{P}$, is derived from the data of the parametric multistage stochastic optimization problem (1b)–(1e) and the function Φ is shown to be differentiable on $\operatorname{int}\mathcal{P}$ when assuming differentiability properties on the data of the problem. Thus, for a closed convex $\mathcal{P}_{\text{ad}} \subset \operatorname{int}\mathcal{P}$, the upstream optimization problem (1a) is a convex differentiable optimization problem.

- In §3.2, we drop the differentiability properties made on the data of the parametric multistage stochastic optimization problem (1b)–(1e) By a regularization procedure using the Moreau envelope, we build a convex differentiable parametric multistage stochastic optimization problem over $\mathbb{R}^{n_p}$ which approximates Problem (1b–1e). Thus, using the same steps as in §3.1, we obtain that, for any closed convex $\mathcal{P}_{\text{ad}} \subset \mathbb{R}^{n_p}$, the regularized upstream optimization problem is a convex differentiable optimization problem.

[2]See however §4.1 where a first-order oracle is computed thanks to SDDP with the extended state variable (x, p).

3.1 Differentiability and gradient computation by backward induction

We provide assumptions on the data of the optimization problem (1) to enforce the differentiability of the corresponding value function Φ, from the differentiability of the parametric value functions $\{V_t\}_{t\in[\![0,T]\!]}$ in (8). In our main result, we also introduce a Bellman-like backward recursion to compute the gradients $\{\nabla_p V_t\}_{t\in[\![0,T]\!]}$, and thus the gradient $\nabla\Phi$ of the objective function Φ defined in Equation (1a).

Parametric differentiability is preserved by the Bellman equation. We start with a lemma that gives sufficient conditions to propagate convexity and parametric differentiability properties through the Bellman equations. We introduce three definitions.

Definition 3.1 (Γ) *For an Euclidean space $\mathbb{Y}$, we denote by $\Gamma[\mathbb{Y},\mathbb{R}^{n_p}]$ the set of lsc convex functions $\gamma : \mathbb{Y}\times\mathbb{R}^{n_p} \to]-\infty,+\infty]$.*

Definition 3.2 (Θ) *For an Euclidean space $\mathbb{Y}$ and a given set $\mathcal{P}\subset\mathbb{R}^{n_p}$, we denote by $\Theta[\mathbb{Y},\mathcal{P}]$ the subset of functions θ in $\Gamma[\mathbb{Y},\mathbb{R}^{n_p}]$ satisfying*

1. *the effective domain of θ is a (possibly empty) product: $\mathrm{dom}\theta = Y_\theta\times\mathcal{P}\subset\mathbb{Y}\times\mathbb{R}^{n_p}$,*
2. *for all $y\in Y_\theta$, the function $\theta(y,\cdot)$ is differentiable on $\mathrm{int}\,\mathcal{P}$.*

We recall that in the context of Definition 3.2, the notions of Fréchet and Gâteaux differentiability coincide [2, Corollary 17.44]. To handle marginal functions, we are interested in compacity properties when minimizing with respect to controls $u\in\mathbb{R}^{n_u}$.

Definition 3.3 (compacity) *We denote by $\Gamma_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$ the subset of functions γ in $\Gamma[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$ for which there exists a compact set $\mathcal{K}_\gamma\subset\mathbb{R}^{n_u}$ such that*[3]

$$\forall (x,p)\in\mathbb{R}^{n_x}\times\mathbb{R}^{n_p}\ ,\ \ \mathrm{dom}\gamma(x,\cdot,p)\subset\mathcal{K}_\gamma\ . \tag{14}$$

We also introduce the set $\Theta_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathcal{P}] = \Theta[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathcal{P}]\cap\Gamma_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$, for a given set $\mathcal{P}\subset\mathbb{R}^{n_p}$.

The next theorem shows that the set $\Theta[\mathbb{R}^{n_x},\mathcal{P}]$ of functions in Definition 3.1 is stable by the Bellman operator, and we give a formula to propagate gradients.

[3] Indeed we also have

$$\left\{u\in\mathbb{R}^{n_u}\,\middle|\,\exists (x,p)\in\mathbb{R}^{n_x}\times\mathbb{R}^{n_p}\text{ s.t. }(x,u,p)\in\mathrm{dom}\gamma\right\} = \bigcup_{(x,p)\in\mathbb{R}^{n_x}\times\mathbb{R}^{n_p}}\mathrm{dom}\gamma(x,\cdot,p)\subset\mathcal{K}_\gamma\ .$$

Theorem 3.4 *Given a function $L : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \times \mathbb{R}^{n_p} \to]-\infty, +\infty]$ and a mapping $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \to \mathbb{R}^{n_x}$, we consider the Bellman operator $\mathcal{B}$ defined, for any function $\varphi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \to]-\infty, +\infty]$, by*

$$\mathcal{B}(\varphi)(x,p) = \inf_{u \in \mathbb{R}^{n_u}} \mathbb{E}\Big[L(x,u,\mathbf{W},p) + \varphi\big(f(x,u,\mathbf{W}),p\big)\Big] \ , \ \ \forall (x,p) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \ . \tag{15}$$

Assume that the mapping f is linear in (x,u). For any w in the (finite) support of $\mathbf{W}$, we denote $L_w = L(\cdot,\cdot,w,\cdot)$. We have that,

1. *if $\{L_w\}_{w \in \mathrm{supp}(\mathbf{W})} \subset \Gamma_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathbb{R}^{n_p}]$, then the Bellman operator $\mathcal{B}$ is an endomorphism on $\Gamma[\mathbb{R}^{n_x}, \mathbb{R}^{n_p}]$,*

2. *if $\{L_w\}_{w \in \mathrm{supp}(\mathbf{W})} \subset \Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}]$ for a given set $\mathcal{P} \subset \mathbb{R}^{n_p}$, then the Bellman operator $\mathcal{B}$ is an endomorphism on $\Theta[\mathbb{R}^{n_x}, \mathcal{P}]$. Moreover, for all $(x,p) \in \mathrm{dom}\big(\mathcal{B}(\varphi)\big)$ with $p \in \mathrm{int}\,\mathcal{P}$, the gradient of the function $\mathcal{B}(\varphi)(x,\cdot)$ at p is given by*

$$\nabla_p \mathcal{B}(\varphi)(x,p) = \mathbb{E}\Big[\nabla_p L(x,u^*,\mathbf{W},p) + \nabla_p \varphi\big(f(x,u^*,\mathbf{W}),p\big)\Big] \ , \tag{16}$$

for any $u^* \in \mathbb{R}^{n_u}$ in the nonempty argmin set of (15).

Proof. We concentrate on the proof of claim (*2.*).

(i) To begin with, for any function $\varphi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \to]-\infty, +\infty]$, we define

$$\widetilde{\mathcal{B}}(\varphi)(x,u,p) = \mathbb{E}\Big[L(x,u,\mathbf{W},p) + \varphi\big(f(x,u,\mathbf{W}),p\big)\Big] \ , \ \ \forall (x,u,p) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \ , \tag{17}$$

— that is, Equation (15) before taking the infimum — and, assuming that $\varphi \in \Theta[\mathbb{R}^{n_x}, \mathcal{P}]$ and $\{L_w\}_{w \in \mathrm{supp}(\mathbf{W})} \subset \Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}]$, we prove that the function $\widetilde{\mathcal{B}}(\varphi)$ belongs to $\Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}]$.

First, we consider the function $Q_L = \mathbb{E}\Big[L(\cdot,\cdot,\mathbf{W},\cdot)\Big] : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to]-\infty, +\infty]$. As, by assumption, $L_w \in \Theta[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}]$ for all w in the support of $\mathbf{W}$, we get that Q_L is convex lsc, with $\mathrm{dom} L_w = Y_{L_w} \times \mathcal{P}$ for some $Y_{L_w} \subset \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}$. We deduce that the effective domain of the function Q_L is given by

$$\mathrm{dom} Q_L = \bigcap_{w \in \mathrm{supp}(\mathbf{W})} \mathrm{dom} Q_{L_w} = \Big(\bigcap_{w \in \mathrm{supp}(\mathbf{W})} Y_{L_w}\Big) \times \mathcal{P} = Y_{Q_L} \times \mathcal{P} \ , \ \text{ with } \ \ Y_{Q_L} = \bigcap_{w \in \mathrm{supp}(\mathbf{W})} Y_{L_w} \ .$$

For any $(x,u) \in Y_{Q_L}$, the differentiability of the function $Q_L(x,u,\cdot)$ is a straightforward consequence of the assumption that $L_w \in \Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}]$ for all $w \in \mathrm{supp}(\mathbf{W})$. Indeed, the function $Q_L(x,u,\cdot)$ is a finite convex combination of the functions $L_w(x,u,\cdot)$, that are differentiable in $\mathrm{int}\,\mathcal{P}$. We thus obtain that $Q_L \in \Theta[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}]$. Moreover, with the notations of Definition 3.3, as $L_w \in \Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}]$ for all $w \in \mathrm{supp}(\mathbf{W})$, we have that

$$\{u \in \mathbb{R}^{n_u} \,|\, \exists (x,u,p) \in Y_{Q_L} \times \mathcal{P}\} \subset \bigcap_{w \in \mathrm{supp}(\mathbf{W})} \{u \in \mathbb{R}^{n_u} \,|\, \exists (x,u,p) \in Y_{L_w} \times \mathcal{P}\} \subset \bigcap_{w \in \mathrm{supp}(\mathbf{W})} \mathcal{K}_{L_w} \ ,$$

which implies that the function Q_L belongs to $\Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}]$.

Second, we consider the function $Q_\varphi = \mathbb{E}\Big[\varphi\big(f(\cdot,\cdot,\mathbf{W}),\cdot\big)\Big] : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \to]-\infty,+\infty]$. From the assumptions that $\varphi \in \Theta[\mathbb{R}^{n_x},\mathcal{P}]$ and f is linear, we have that the function Q_φ is convex lsc and that its effective domain is a Cartesian product[4]. The differentiability of the function Q_φ with respect to p is straightforward.

Gathering the results obtained for Q_L and Q_φ, since $\widetilde{\mathcal{B}}(\varphi) = Q_L + Q_\varphi$ and using the fact that $\mathrm{dom}(Q_L + Q_\varphi) \subset \mathrm{dom} Q_L$, we obtain that $\widetilde{\mathcal{B}}(\varphi)$ belongs to $\Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u},\mathcal{P}]$.

(ii) We have just proven that the function $\widetilde{\mathcal{B}}(\varphi)$ belongs to the set $\Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u},\mathcal{P}]$ when $\varphi \in \Theta[\mathbb{R}^{n_x},\mathcal{P}]$ and $\{L_w\}_{w\in\mathrm{supp}(\mathbf{W})} \subset \Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u},\mathcal{P}]$. Now, there remains to take the infimum to prove that the function $\mathcal{B}(\varphi)$ belongs to $\Theta[\mathbb{R}^{n_x},\mathcal{P}]$.

First, we prove the effective domain rectangular property in Definition 3.2. Since by definition,

$$\mathcal{B}(\varphi)(x,p) = \inf_{u\in\mathbb{R}^{n_u}} \widetilde{\mathcal{B}}(\varphi)(x,u,p)\ ,\ \ \forall (x,u) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}\ , \tag{18}$$

we deduce that the function $\mathcal{B}(\varphi)$ is convex as the marginal of a convex function. Then, let $K_{\widetilde{\mathcal{B}}(\varphi)} \subset \mathbb{R}^{n_u}$ be the compact set associated to $\widetilde{\mathcal{B}}(\varphi)$ as in Definition 3.3. For all $(x,u) \in \mathbb{R}^{n_x}\times\mathbb{R}^{n_u}$, we have

$$\mathcal{B}(\varphi)(x,p) = \inf_{u\in\mathbb{R}^{n_u}} \widetilde{\mathcal{B}}(\varphi)(x,u,p) = \inf_{u\in\mathrm{dom}\widetilde{\mathcal{B}}(\varphi)(x,\cdot,p)} \widetilde{\mathcal{B}}(\varphi)(x,u,p) = \inf_{u\in K_{\widetilde{\mathcal{B}}(\varphi)}} \widetilde{\mathcal{B}}(\varphi)(x,u,p)\ . \tag{19}$$

Using the last expression in Equation (19), the function $\mathcal{B}(\varphi)$ is lsc and the infimum above — hence in (15) — is attained by [2, Lemma 1.30]. We deduce that

$$\forall (x,p) \in \mathbb{R}^{n_x} \times \mathcal{P}\ ,\ \ \exists u^* \in \mathbb{R}^{n_u}\ ,\ \ \mathcal{B}(\varphi)(x,p) = \widetilde{\mathcal{B}}(\varphi)(x,u^*,p)\ . \tag{20}$$

Therefore, the function $\mathcal{B}(\varphi)$ never takes the value $-\infty$ (as so does $\widetilde{\mathcal{B}}(\varphi)$). Clearly, if $\mathrm{dom}\big(\mathcal{B}(\varphi)\big) = \emptyset$, then $\mathcal{B}(\varphi) \in \Theta[\mathbb{R}^{n_x},\mathcal{P}]$. Therefore, we now consider the case where $\mathrm{dom}\big(\mathcal{B}(\varphi)\big) \neq \emptyset$. Let $(x,p) \in \mathrm{dom}\big(\mathcal{B}(\varphi)\big)$ and $u^* \in \mathbb{R}^{n_u}$ be a minimizer of (15) at (x,p). As $\mathcal{B}(\varphi)(x,p) \leq \widetilde{\mathcal{B}}(\varphi)(x,u^*,p)$ and $\widetilde{\mathcal{B}}(\varphi) \in \Theta[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u},\mathcal{P}]$, we deduce that $\mathcal{P} \subset \mathrm{dom}\big(\mathcal{B}(\varphi)(x,\cdot)\big)$, and conclude from (20) that $\mathrm{dom}\big(\mathcal{B}(\varphi)(x,\cdot)\big) = \mathcal{P}$. This proves that $\mathrm{dom}\big(\mathcal{B}(\varphi)\big) = X \times \mathcal{P}$, for some set $X \subset \mathbb{R}^{n_x}$.

Second, we turn to the parametric differentiability of the function $\mathcal{B}(\varphi) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \to]-\infty,+\infty]$. For a fixed $x \in X$, we introduce the two functions

$$\psi_x = \mathcal{B}(\varphi)(x,\cdot) \quad \text{and} \quad \Psi_x = \widetilde{\mathcal{B}}(\varphi)(x,\cdot,\cdot)\ . \tag{21}$$

The function $\psi_x : \mathbb{R}^{n_p} \to]-\infty,+\infty]$ is proper, convex and lsc, hence subdifferentiable over the interior of its domain [2, Proposition 16.27]. Let $p \in \mathrm{int}\,\mathcal{P}$, $s \in \partial\psi_x(p)$, and $u^* \in \mathbb{R}^{n_u}$ be a minimizer as in (20). From [20, Corollary 2.63], we get that $(0,s) \in \partial\Psi_x(u^*,p)$, and therefore $s \in \partial_p\Psi_x(u^*,p)$, so that $s = \nabla_p\Psi_x(u^*,p)$ (as $\widetilde{\mathcal{B}}(\varphi) \in \Theta[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u},\mathcal{P}]$) hence $\partial\psi_x(p)$ is a singleton. We conclude that the function ψ_x is differentiable at p and that $\nabla\psi_x(p) = \nabla_p\Psi_x(u^*,p)$ [2, Proposition 17.45], that is,

$$\nabla_p\mathcal{B}(\varphi)(x,p) = \nabla_p\mathbb{E}\Big[L(x,u^*,\mathbf{W},p) + \varphi\big(f(x,u^*,\mathbf{W}),p\big)\Big]\ . \tag{22}$$

[4]Had the mapping f depended on the parameter p, we could not have concluded that the effective domain of the function Q_φ be a Cartesian product.

This finally proves that $\mathcal{B}(\varphi) \in \Theta[\mathbb{R}^{n_x}, \mathcal{P}]$, that is, the Bellman operator $\mathcal{B}$ is an endomorphism on $\Theta[\mathbb{R}^{n_x}, \mathcal{P}]$. Moreover, since the support of $\mathbf{W}$ is a finite set, exchanging the derivation and expectation operators in (22) is trivial, so that (16) holds true.

As for claim (*1.*), when $\varphi \in \Gamma[\mathbb{R}^{n_x}, \mathbb{R}^{n_p}]$ and $\{L_w\}_{w \in \mathrm{supp}(\mathbf{W})} \subset \Gamma_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathbb{R}^{n_p}]$, we obtain with analogous arguments that (i) the function $\widetilde{\mathcal{B}}(\varphi)$ belongs to the set $\Gamma_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathbb{R}^{n_p}]$, (ii) the function $\mathcal{B}(\varphi)$ belongs to the set $\Gamma[\mathbb{R}^{n_x}, \mathbb{R}^{n_p}]$. $\square$

Figure 1 illustrates the links induced by the Bellman operator $\mathcal{B}$ between the different subsets of functions, as revealed by Theorem 3.4. The operator $\widetilde{\mathcal{B}}$ is defined in (17) and corresponds to the operator $\mathcal{B}$ before taking the infimum in u.

$$\begin{tikzcd}
\Theta[\mathbb{R}^{n_x}, \mathcal{P}] \arrow[d, hook, "\cap"] & \Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}] \arrow[l, "\min_u"'] \arrow[d, hook, "\cap"] & \Theta[\mathbb{R}^{n_x}, \mathcal{P}] \arrow[l, "\tilde{\mathcal{B}}", "(L_w \in \Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathbb{R}^{n_p}])"'] \arrow[d, hook, "\cap"] \\
\Gamma[\mathbb{R}^{n_x}, \mathbb{R}^{n_p}] & \Gamma_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathcal{P}] \arrow[l, "\min_u"'] & \Gamma[\mathbb{R}^{n_x}, \mathbb{R}^{n_p}] \arrow[l, "\tilde{\mathcal{B}}", "(L_w \in \Gamma_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathbb{R}^{n_p}])"']
\end{tikzcd}$$

Figure 1: Stability of the classes of functions Θ and Γ by the Bellman operator $\mathcal{B}$.

Application to parametric value functions. We introduce assumptions on the data of Problem (1) to enforce the convexity and the parametric differentiability of the parametric value functions $\{V_t\}_{t \in [\![0,T]\!]}$ in (8).

Assumption 3.5 (convex multistage problem) *We assume that*

1. *Problem* (1) *is feasible,*
2. *the dynamics* $\{f_t\}_{t \in [\![0,T-1]\!]}$ *in* (6) *are linear with respect to their arguments* (x, u)*,*
3. *for all* $t \in [\![0, T-1]\!]$*, the functions* $\{L_t(\cdot, \cdot, w, \cdot)\}_{w \in \mathrm{supp}(\mathbf{W}_t)}$ *—defined after the stage cost* L_t *in* (7a)*— belong to* $\Gamma_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathbb{R}^{n_p}]$*,*
4. *the final cost* K *in* (7b) *belongs to* $\Gamma[\mathbb{R}^{n_x}, \mathbb{R}^{n_p}]$*.*

Assumption 3.5 puts us in a standard context for convex stochastic multistage optimization problems (see e.g. [16, assumption H_1]). By assuming that Problem (1) is feasible, we obtain that the function V_0 is proper — as $\Phi = V_0(x_0, \cdot)$ — and therefore that all parametric value functions $\{V_t\}_{t \in [\![0,T]\!]}$ are proper — due to the Bellman equations (8). At a lower level of detail, this can also be enforced by a "relatively complete recourse" assumption (see e.g [16, 19]).

We make a second assumption to handle parametric differentiability.

Assumption 3.6 (parametric differentiability) *Let* $\mathcal{P}$ *be a given subset of* $\mathbb{R}^{n_p}$*. We assume that*

1. *for all* $t \in [\![0, T-1]\!]$, *the functions* $\{L_t(\cdot,\cdot,w,\cdot)\}_{w\in\mathrm{supp}(\mathbf{W}_t)}$ *—defined after the stage cost* L_t *in* (7a)*— belong to* $\Theta_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathcal{P}]$,

2. *the final cost* K *in* (7b) *belongs to* $\Theta[\mathbb{R}^{n_x},\mathcal{P}]$.

We now state our main result regarding the differentiability of the parametric value functions $\{V_t\}_{t\in[\![0,T]\!]}$.

Theorem 3.7 *Under the discrete white noise Assumption 2.3 and the convex multistage problem Assumption 3.5, we have that,*

1. *the parametric value functions* $\{V_t\}_{t\in[\![0,T]\!]}$ *defined in* (8) *are proper and belong to* $\Gamma[\mathbb{R}^{n_x},\mathbb{R}^{n_p}]$,

2. *moreover, under the parametric differentiability Assumption 3.6, the parametric value functions* $\{V_t\}_{t\in[\![0,T]\!]}$ *belong to* $\Theta[\mathbb{R}^{n_x},\mathcal{P}]$ *and, for all* $p\in\operatorname{int}\mathcal{P}$, *their gradients can be computed by backward induction, with, at final stage* T,

$$\nabla_p V_T(x,p) = \nabla_p K(x,p)\ ,\ \ \forall x\in \mathrm{dom}V_T(\cdot,p)\ , \tag{23a}$$

and, at any stage $t\in[\![0,T-1]\!]$ *and for all* $x\in\mathrm{dom}V_t(\cdot,p)$,

$$\nabla_p V_t(x,p) = \mathbb{E}\Big[\nabla_p L_t(x,u^*,\mathbf{W}_{t+1},p) + \nabla_p V_{t+1}\big(f_t(x,u^*,\mathbf{W}_{t+1}),p\big)\Big]\ , \tag{23b}$$

where the controls u^* *is any control in the solution set* $\mathcal{U}_t^*(x,p)$ *defined in* (11),

3. *for any closed convex* $\mathcal{P}_{ad}\subset\mathcal{P}$, *the upstream optimization problem* (1a) *is a convex differentiable optimization problem.*

Proof. For each time $t\in[\![0,T]\!]$, we consider the Bellman operator $\mathcal{B}_t$ defined, for any function $\varphi:\mathbb{R}^{n_x}\times\mathbb{R}^{n_p}\to\,]-\infty,+\infty]$ and any for all $(x,p)\in\mathbb{R}^{n_x}\times\mathbb{R}^{n_p}$, by

$$\mathcal{B}_t(\varphi)(x,p) = \inf_{u\in\mathbb{R}^{n_u}} \mathbb{E}\Big[L_t(x,u,\mathbf{W},p) + \varphi\big(f(x,u,\mathbf{W}),p\big)\Big]\ .$$

Thus, the value functions $\{V_t\}_{t\in[\![0,T]\!]}$ defined in (8) satisfy $V_T = K$ and for all $t\in[\![0,T-1]\!]$, $V_t = \mathcal{B}_t(V_{t+1})$. The proof follows by applying backward in time Theorem 3.4 to the Bellman operator $\mathcal{B}_t$ and the value function V_{t+1}. □

Remark 3.8 *In the case where Problem* (1b)–(1e) *has explicit constraints of the form* $\mathbf{U}_t\in\mathcal{U}_t(\mathbf{X}_t,p)$, *at each time* t, *the parametric stage cost* L_t *in* (7a) *incorporates the indicator function* $\delta_{\mathcal{U}_t(x,p)}$ *(see Remark 2.1). Then, a condition to ensure the differentiability of* L_t *is that the set-valued function* $\mathcal{U}_t$ *does not depend on the parameter* p.

3.2 Regularization of convex nondifferentiable parametric value functions

We now turn to the non differentiable case, that is, we do not assume anymore that the functions $\{L_t\}_{t\in[\![0,T-1]\!]}$ and K are differentiable with respect to the parameter p. To overcome this drawback and go back to the differentiable situation, we appeal to the Morean-Yosida regularization We recall the definition of the Moreau envelope [21, 34].

Definition 3.9 *Let $n \in \mathbb{N}^*$, $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function and $\mu \in \mathbb{R}_+^*$ be a regularization coefficient. The Moreau envelope of f is the function*

$$f^\mu : \mathbb{R}^n \to \overline{\mathbb{R}} \ , \ \ z \mapsto \inf_{z'\in\mathbb{R}^n} \left(f(z') + \frac{1}{2\mu}||z - z'||_2^2 \right) . \tag{24}$$

We refer to [28, Chapter 1, §G] and [2, Chapter 12, §4] for a review of the properties of the Moreau envelope. Given values of $(x, u, w) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w}$ and a regularization coefficient $\mu \in \mathbb{R}_+^*$, we introduce the *parametric Moreau envelopes* $\{L_t^\mu(x, u, w, \cdot)\}_{t\in[\![0,T-1]\!]}$ and $K^\mu(x, \cdot)$ of the parametric cost functions $\{L_t\}_{t\in[\![0,T-1]\!]}$ in (7a) and K in (7b), with respect to the parameter p in (3), defined as

$$L_t^\mu(x, u, w, p) = \inf_{p'\in\mathbb{R}^{n_p}} \left(L_t(x, u, w, p') + \frac{1}{2\mu}||p - p'||_2^2 \right) , \ \ \forall t \in [\![0, T - 1]\!] \ , \ \ \forall p \in \mathbb{R}^{n_p} \ , \tag{25a}$$

$$K^\mu(x, p) = \inf_{p'\in\mathbb{R}^{n_p}} \left(K(x, p') + \frac{1}{2\mu}||p - p'||_2^2 \right) , \ \ \forall p \in \mathbb{R}^{n_p} \ . \tag{25b}$$

In order to ensure that the regularized parametric cost functions $\{L_t^\mu\}_{t\in[\![0,T-1]\!]}$ and K^μ are lower semi continuous, we introduce the following parametric compacity assumption.

Assumption 3.10 (parametric compacity) *We assume that there exists a compact set $\mathcal{P} \subset \mathbb{R}^{n_p}$ such that for all $t \in [\![0, T-1]\!]$ and for all $w \in \mathrm{supp}(\mathbf{W}_t)$, the stage costs L_t in* (7a) *satisfy*

$$\forall (x, u) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \ , \ \ \mathrm{dom} L_t(x, u, w, \cdot) \subset \mathcal{P} \ , \tag{26a}$$

and the final cost K in (7b) *satisfies*

$$\forall x \in \mathbb{R}^{n_x} \ , \ \ \mathrm{dom} K(x, \cdot) \subset \mathcal{P} \ . \tag{26b}$$

Lemma 3.11 *Let $\mu > 0$. Under the parametric compacity Assumption 3.10, we have that*

1. *for all $t \in [\![0, T-1]\!]$ and all $w \in \mathrm{supp}(\mathbf{W}_t)$, if $L_t(\cdot, \cdot, w, \cdot) \in \Gamma_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathbb{R}^{n_p}]$, then $L_t^\mu(\cdot, \cdot, w, \cdot) \in \Theta_{\mathcal{K}}[\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathbb{R}^{n_p}]$,*
2. *if $K \in \Gamma[\mathbb{R}^{n_x}, \mathbb{R}^{n_p}]$, then $K^\mu \in \Theta[\mathbb{R}^{n_x}, \mathbb{R}^{n_p}]$.*

Proof. Let us consider a function $L_t(\cdot,\cdot,w,\cdot)$ and its partial Moreau envelope $L_t^\mu(\cdot,\cdot,w,\cdot)$ defined by (25).

First, we prove that $L_t^\mu(\cdot,\cdot,w,\cdot) \in \Gamma_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$. As the marginal of a convex function, the function $L_t^\mu(\cdot,\cdot,w,\cdot)$ is convex. The infimum in (25) can be taken equivalently over the fixed compact set $\mathcal{P}\subset\mathbb{R}^{n_p}$ given in Assumption 3.10 . As $L_t(\cdot,\cdot,w,\cdot)\in\Gamma_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$ is lsc, we deduce that the marginal function $L_t^\mu(\cdot,\cdot,w,\cdot)$ is also lsc, and that the infimum in (25) is attained [2, Lemma 1.30]. It follows that $L_t^\mu(\cdot,\cdot,w,\cdot)$ takes values in $]-\infty,+\infty]$ — as so does $L_t(\cdot,\cdot,w,\cdot)$ — and therefore that $L_t^\mu(\cdot,\cdot,w,\cdot)\in\Gamma[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$. Now, let $u\in\mathbb{R}^{n_u}$ be such that there exists $(x,p)\in\mathbb{R}^{n_x}\times\mathbb{R}^{n_p}$ so that $(x,u,p)\in\mathrm{dom}L_t^\mu(\cdot,\cdot,w,\cdot)$. As we have just seen, there exists a minimizer $p^*\in\mathbb{R}^{n_p}$ such that

$$L_t^\mu(x,u,w,p) = L_t(x,u,w,p^*) + \frac{1}{2\mu}||p-p^*||_2^2 \ . \tag{27}$$

Necessary, $(x,u,p^*)\in\mathrm{dom}L_t(\cdot,\cdot,w,\cdot)$, which proves that $u\in\mathcal{K}_{L_t(\cdot,\cdot,w,\cdot)}$, following Definition 3.3. We deduce that $L_t^\mu(\cdot,\cdot,w,\cdot)\in\Gamma_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$.

Second, we prove that $L_t^\mu(\cdot,\cdot,w,\cdot)\in\Theta_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$. Following Definition 3.3, we only need to prove that $L_t^\mu(\cdot,\cdot,w,\cdot)\in\Theta[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$. Let us consider again $(x,u,p)\in\mathrm{dom}L_t^\mu(\cdot,\cdot,w,\cdot)$. The function $L_t(x,u,w,\cdot)$ is convex, lsc, never takes the value $-\infty$ (by assumption) and its domain is nonempty (by (27)). It follows that its Moreau envelope $L_t^\mu(x,u,w,\cdot)$ is finite valued everywhere on $\mathbb{R}^{n_p}$ ([2, Proposition 12.15]) and differentiable on $\mathbb{R}^{n_p}$ ([2, Proposition 12.30]). We deduce that $L_t^\mu(\cdot,\cdot,w,\cdot)\in\Theta[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$, and finally that $L_t^\mu(\cdot,\cdot,w,\cdot)\in\Theta_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$.

The proof that $K^\mu\in\Theta_{\mathcal{K}}[\mathbb{R}^{n_x},\mathbb{R}^{n_p}]$ is analogous. □

We are now ready to introduce the *lower smooth parametric value functions*. For any regularization coefficient $\mu\in\mathbb{R}_+^*$, we define

$$\underset{\sim}{V}_T^\mu(x,p) = K^\mu(x,p)\ ,\ \ \forall (x,p)\in\mathbb{R}^{n_x}\times\mathbb{R}^{n_p}\ , \tag{28a}$$

$$\underset{\sim}{V}_t^\mu(x,p) = \inf_{u\in\mathbb{R}^{n_u}} \mathbb{E}\Big[L_t^\mu(x,u,\mathbf{W}_{t+1},p) + \underset{\sim}{V}_{t+1}^\mu\big(f_t(x,u,\mathbf{W}_{t+1}),p\big)\Big]\ ,$$
$$\forall (x,p)\in\mathbb{R}^{n_x}\times\mathbb{R}^{n_p}\ ,\ \ \forall t\in[\![0,T-1]\!]\ . \tag{28b}$$

The lower smooth parametric value functions $\{\underset{\sim}{V}_t^\mu\}_{t\in[\![0,T]\!]}$ have several interesting properties, which we gather in Theorem 3.12.

Theorem 3.12 *Let $\mu\in\mathbb{R}_+^*$ be a regularization coefficient, and let $\{\underset{\sim}{V}_t^\mu\}_{t\in[\![0,T]\!]}$ be the lower smooth parametric value functions defined in* (28)*. Under the discrete white noise Assumption 2.3, the convex multistage problem Assumption 3.5, and the parametric compacity Assumption 3.10,*

1. *the functions $\{\underset{\sim}{V}_t^\mu\}_{t\in[\![0,T]\!]}$ provide lower bounds on the value functions $\{V_t\}_{t\in[\![0,T]\!]}$ defined in* (8)*, that is,*

$$\underset{\sim}{V}_t^\mu \le V_t\ ,\ \ \forall t\in[\![0,T]\!]\ , \tag{29}$$

2. *the functions* $\{\underset{\sim}{V}_t^\mu\}_{t\in[\![0,T]\!]}$ *are proper and belong to* $\Theta[\mathbb{R}^{n_x},\mathbb{R}^{n_p}]$,[5] *and their gradients can be computed by backward induction, with, at final stage* T,

$$\nabla_p \underset{\sim}{V}_T^\mu(x,p) = \nabla_p K^\mu(x,p)\ , \tag{30a}$$

and, at any stage $t \in [\![0,T-1]\!]$,

$$\nabla_p \underset{\sim}{V}_t^\mu(x,p) = \mathbb{E}\Big[\nabla_p L_t^\mu(x,u^*,\mathbf{W}_{t+1},p) + \nabla_p \underset{\sim}{V}_{t+1}^\mu\big(f_t(x,u^*,\mathbf{W}_{t+1}),p\big)\Big]\ , \tag{30b}$$

for any u^* *in the nonempty solution set* $\mathcal{U}_t^*(x,p)$ *of Problem* (28b)

Proof. First, we prove Inequality (29) proceeding by backward induction. By the properties of the Moreau envelope, we have that $V_T^\mu = K^\mu \leq K = V_T$ [2, Proposition 12.9]. Let $t \in [\![0,T-1]\!]$ and let us assume that (29) holds true at stage $t+1$. Consequently, by application of [2, Proposition 12.9], we have that for any $(x,u,p) \in \mathbb{R}^{n_x}\times\mathbb{R}^{n_u}\times\mathbb{R}^{n_p}$ and $w \in \mathrm{supp}(\mathbf{W}_{t+1})$,

$$L_t^\mu(x,u,w,p) + \underset{\sim}{V}_{t+1}^\mu\big(f_t(x,u,w),p\big) \leq L_t(x,u,w,p) + V_{t+1}\big(f_t(x,u,w),p\big)\ ,$$

from which we deduce that for any $(x,p) \in \mathbb{R}^{n_x}\times\mathbb{R}^{n_p}$,

$$\begin{aligned}\underset{\sim}{V}_t^\mu(x,p) &= \inf_{u\in\mathbb{R}^{n_u}} \mathbb{E}\Big[L_t^\mu(x,u,\mathbf{W}_{t+1},p) + \underset{\sim}{V}_{t+1}^\mu\big(f_t(x,u,\mathbf{W}_{t+1}),p\big)\Big]\ ,\\ &\leq \inf_{u\in\mathbb{R}^{n_u}} \mathbb{E}\Big[L_t(x,u,\mathbf{W}_{t+1},p) + V_{t+1}\big(f_t(x,u,\mathbf{W}_{t+1}),p\big)\Big]\ ,\\ &= V_t(x,p)\ ,\end{aligned}$$

so that Inequality (29) holds true.

Second, we apply Theorem 3.7. To do this, we observe that the parametric value functions $\{\underset{\sim}{V}_t^\mu\}_{t\in[\![0,T]\!]}$ are the value functions of the new PMSOP $\min_{p\in\mathcal{P}_{\mathrm{ad}}}\Phi^\mu(p)$, whose definition follows the one of the original PMSOP (1), except that the data of the problem $\big(\{f_t\}_{t\in[\![0,T-1]\!]}, \{L_t\}_{t\in[\![0,T-1]\!]}, K\big)$ is replaced by $\big(\{f_t\}_{t\in[\![0,T-1]\!]}, \{L_t^\mu\}_{t\in[\![0,T-1]\!]}, K^\mu\big)$. Lemma 3.11 tells us that the new components of this data triplet satisfies the conditions of Theorem 3.7, so that we only need to prove that the new PMSOP $\min_{p\in\mathcal{P}_{\mathrm{ad}}}\Phi^\mu(p)$ is feasible (to fulfill Assumption 3.5). Again, under the white noise Assumption 2.3, we have from the dynamic programming principle that $\Phi^\mu = \underset{\sim}{V}_0^\mu(x_0,\cdot)$. By definition of the functions $\{\underset{\sim}{V}_t^\mu\}_{t\in[\![0,T]\!]}$ in (28), and by the properties of the cost functions $\{L_t^\mu\}_{t\in[\![0,T-1]\!]}$ and K^μ in Lemma 3.11, we obtain that all functions in $\{\underset{\sim}{V}_t^\mu\}_{t\in[\![0,T]\!]}$ belong to $\Gamma[\mathbb{R}^{n_x},\mathbb{R}^{n_p}]$ (applying Theorem 3.4 backward in time). It follows that $\underset{\sim}{V}_0^\mu$ takes values in $]-\infty,+\infty]$. Now, as the original PMSOP (1) is feasible, taking $p \in \mathrm{dom}\Phi \cap \mathcal{P}_{\mathrm{ad}}$, we have that $\Phi^\mu(p) = \underset{\sim}{V}_0^\mu(x_0,p) \leq V_0(x_0,p) = \Phi(p) < +\infty$, which proves that the problem $\min_{p\in\mathcal{P}_{\mathrm{ad}}}\Phi^\mu(p)$ is feasible too. Therefore, we can apply Theorem 3.7. This concludes the proof. □

To summarize, in the non differentiable case, we have obtained differentiable value functions by Theorem 3.12 that are lower bounds of the original value functions. The gradients of these differentiable value functions can be computed by backward induction as stated by Theorem 3.7.

[5] and therefore to $\Gamma[\mathbb{R}^{n_x},\mathbb{R}^{n_p}]$

3.3 Convergence properties of regularized convex parametric value functions

Finally, we prove some convergence properties of the lower smooth parametric value functions $\{\underset{\sim}{V}_t^{\mu}\}_{t\in[\![0,T]\!]}$ defined in (28), which show that they are suitable candidates to approximate the original value functions $\{V_t\}_{t\in[\![0,T]\!]}$ in (8) for solving Problem (1). We refer the reader to the definition of pointwise convergence in [28, §7.A], denoted by "$\xrightarrow{\mathrm{P}}$", and to the definition of epiconvergence in [28, §7.B] denoted by "$\xrightarrow{\mathrm{e}}$".

Proposition 3.13 *Let $\{\mu_n\}_{n\in\mathbb{N}} \in (\mathbb{R}_+^*)^{\mathbb{N}}$ be a nonincreasing sequence of positive real numbers such that $\lim_{n\to+\infty}\mu_n = 0$, let $\{\underset{\sim}{V}_t^{\mu_n}\}_{t\in[\![0,T]\!],n\in\mathbb{N}}$ be lower smooth parametric value functions as defined in (28), and let $\{V_t\}_{t\in[\![0,T]\!]}$ be the parametric value functions defined in (8).*

Under the discrete white noise Assumption 2.3, the convex multistage problem Assumption 3.5 and the parameteric compacity Assumption 3.10, we have the following convergence property for all $t \in [\![0,T]\!]$:

$$\underset{\sim}{V}_t^{\mu_n} \xrightarrow[n\to+\infty]{\mathrm{e}} V_t \ . \tag{31}$$

Proof. We know by Theorem 3.12 that, for any $n \in \mathbb{N}$, the functions $\{\underset{\sim}{V}_t^{\mu_n}\}_{t\in[\![0,T]\!]}$ are lsc. Moreover, for each $t \in [\![0,T]\!]$, $\{\underset{\sim}{V}_t^{\mu_n}\}_{n\in\mathbb{N}}$ is a nondecreasing sequence of functions since $\{\mu_n\}_{n\in\mathbb{N}}$ is a nonincreasing sequence of positive real numbers. This ensures the equivalence between pointwise convergence and epiconvergence for the sequence $\{\underset{\sim}{V}_t^{\mu_n}\}_{n\in\mathbb{N}}$ by [28, Proposition 7.4(d)], so that it is sufficient to prove the pointwise convergence of the sequence $\{\underset{\sim}{V}_t^{\mu_n}\}_{n\in\mathbb{N}}$ to obtain (31):

$$\underset{\sim}{V}_t^{\mu_n} \xrightarrow[n\to+\infty]{\mathrm{P}} V_t \ , \tag{32}$$

We proceed by backward induction.

• We start by proving (32) at stage T. Let $n \in \mathbb{N}$, from (28a), $\underset{\sim}{V}_T^{\mu_n} = K^{\mu_n}$, where for $x \in \mathbb{R}^{n_x}$, $K^{\mu_n}(x,\cdot)$ is defined in (25b) as the Moreau envelope of $K(x,\cdot)$. By Assumption 3.5, the function K is proper and belongs to $\Gamma[\mathbb{R}^{n_x},\mathbb{R}^{n_p}]$, so that either $K(x,\cdot) = +\infty$, in which case $\underset{\sim}{V}_T^{\mu_n}(x,\cdot) = K^{\mu_n}(x,\cdot) = +\infty$, or $K(x,\cdot)$ is proper, in which case $\underset{\sim}{V}_T^{\mu_n}(x,\cdot) = K^{\mu_n}(x,\cdot)$ converges pointwise to $K(x,\cdot)$ and is a nondecreasing sequence, from the properties of the Moreau envelope (see [2, Proposition 12.33]). This proves (32) at final stage T.

• Now, fix $t \in [\![0,T-1]\!]$, and assume that the statement (32) holds at stage $t+1$. Let $n \in \mathbb{N}$ and, for any $(x,u,p) \in \mathbb{R}^{n_x}\times\mathbb{R}^{n_u}\times\mathbb{R}^{n_p}$, we define the lower smooth Q-function $\underset{\sim}{Q}_t^{\mu_n}$ by

$$\underset{\sim}{Q}_t^{\mu_n}(x,u,p) = \mathbb{E}\Big[L_t^{\mu_n}(x,u,\mathbf{W}_{t+1},p) + \underset{\sim}{V}_{t+1}^{\mu_n}\big(f_t(x,u,\mathbf{W}_{t+1}),p\big)\Big] \ .$$

As a first step, we show that the sequence of functions $\{\underset{\sim}{Q}_t^{\mu_n}(x,\cdot,p)\}_{n\in\mathbb{N}}$ epiconverges to $Q_t(x,\cdot,p)$, with Q_t defined in (10a). From the properties of the Moreau envelope, the sequence $\{L_t^{\mu_n}\}_{n\in\mathbb{N}}$ in (25a) is nondecreasing and, from Lemma 3.11, for any $w \in \mathrm{supp}(\mathbf{W}_{t+1})$, each function $L_t^{\mu_n}(\cdot,\cdot,w,\cdot)$ is lsc. Similarly, by assumption, the sequence $\{\underset{\sim}{V}_{t+1}^{\mu_n}\}_{n\in\mathbb{N}}$ is nondecreasing and, from Theorem 3.12, each function $\underset{\sim}{V}_{t+1}^{\mu_n}$ is lsc. It follows that (i) $\{\underset{\sim}{Q}_t^{\mu_n}(x,\cdot,p)\}_{n\in\mathbb{N}}$ is nondecreasing, and (ii) each function $\underset{\sim}{Q}_t^{\mu_n}$ is lsc, since the expectation above is a finite sum. Therefore, from [28, Proposition 7.4(d)], we obtain the epiconvergence

$$\underset{\sim}{Q}_t^{\mu_n}(x,\cdot,p) \xrightarrow[n\to+\infty]{\mathrm{e}} \sup_{n\in\mathbb{N}}\big(\mathrm{lsc}\big(\underset{\sim}{Q}_t^{\mu_n}\big)(x,\cdot,p)\big) = \sup_{n\in\mathbb{N}}\big(\underset{\sim}{Q}_t^{\mu_n}(x,\cdot,p)\big) \ .$$

Moreover, the nondecreasing sequences $\{L_t^{\mu_n}\}_{n\in\mathbb{N}}$ and $\{\underset{\sim}{V}_{t+1}^{\mu_n}\}_{n\in\mathbb{N}}$ of functions converge pointwise respectively to the function L_t, using the properties of the Moreau envelope (see [2, Proposition 12.33(ii)]), and to V_{t+1}, using the backward induction assumption on $\{\underset{\sim}{V}_{t+1}^{\mu_n}\}_{n\in\mathbb{N}}$. It follows that $\sup_{n\in\mathbb{N}}\left(\underset{\sim}{Q}_t^{\mu_n}(x,\cdot,p)\right) = Q_t(x,\cdot,p)$.

As a second step, we show that (32) holds at stage t. For that purpose we fix $(x,p) \in \mathbb{R}^{n_x}\times\mathbb{R}^{n_p}$ and consider the sequence of mappings $\{\underset{\sim}{Q}_t^{\mu_n}(x,\cdot,p)\}_{n\in\mathbb{N}}$. Since $\mathrm{dom}(\underset{\sim}{Q}_t^{\mu_n}(x,\cdot,p)) \subseteq \cap_{w\in\mathrm{supp}(\mathbf{W}_{t+1})}\mathrm{dom}L_t^{\mu_n}(x,\cdot,w,p)$ — where $\mathrm{dom}L_t^{\mu_n}(x,\cdot,w,p)$ is bounded as $L_t^{\mu_n}(\cdot,\cdot,w,\cdot) \in \Gamma_{\mathcal{K}}[\mathbb{R}^{n_x}\times\mathbb{R}^{n_u},\mathbb{R}^{n_p}]$ — the sequence $\{\underset{\sim}{Q}_t^{\mu_n}(x,\cdot,p)\}_{n\in\mathbb{N}}$ of functions is *eventually level-bounded* (see [28, Exercice 7.32(a)]). Therefore, by application of [28, Theorem 7.33], we obtain the convergence of the infimums, that is

$$\inf_{u\in\mathbb{R}^{n_u}} \underset{\sim}{Q}_t^{\mu_n}(x,u,p) \xrightarrow[n\to+\infty]{} \inf_{u\in\mathbb{R}^{n_u}} Q_t(x,u,p)\ .$$

Thus, we obtain the pointwise convergence of the sequence $\{\underset{\sim}{V}_t^{\mu_n}\}_{n\in\mathbb{N}}$ as for any fixed value $(x,p) \in \mathbb{R}^{n_x}\times\mathbb{R}^{n_p}$ we have that

$$\underset{\sim}{V}_t^{\mu_n}(x,p) = \inf_{u\in\mathbb{R}^{n_u}} \underset{\sim}{Q}_t^{\mu_n}(x,u,p) \xrightarrow[n\to+\infty]{} \inf_{u\in\mathbb{R}^{n_u}} Q_t(x,u,p) = V_t(x,p)\ ,$$

Moreover, for all $n \in \mathbb{N}$, we have that $\underset{\sim}{V}_t^{\mu_n} \leq \underset{\sim}{V}_t^{\mu_{n+1}}$ since the sequence $\{\underset{\sim}{Q}_t^{\mu_n}(x,\cdot,p)\}_{n\in\mathbb{N}}$ is nondecreasing. We conclude that the statement (32) holds at time t. This ends the proof. □

As a consequence, we obtain the following corollary.

Corollary 3.14 *Under the assumptions of Proposition 3.13, let x_0 be the initial state in* (1c)*, let the set $\mathcal{P}_{ad}$ be compact, and let $\Phi^* = \inf_{p\in\mathcal{P}_{ad}}\Phi(p)$ be the optimal value of Problem* (1)*. Then, we have that*

$$\inf_{p\in\mathcal{P}_{ad}} \underset{\sim}{V}_0^{\mu_n}(x_0,p) \leq \Phi^*\ ,\ \ \forall n\in\mathbb{N}\ ,\ \ \textit{and}\ \ \inf_{p\in\mathcal{P}_{ad}} \underset{\sim}{V}_0^{\mu_n}(x_0,p) \xrightarrow[n\to+\infty]{} \Phi^*\ . \tag{33}$$

Proof. From Proposition 3.13, the sequence $\{\underset{\sim}{V}_0^{\mu_n}\}_{n\in\mathbb{N}}$ of functions converges pointwise and epiconverges to V_0 in (8). Moreover, the function $\delta_{\{x_0\}\times\mathcal{P}_{\mathrm{ad}}}$ is lsc. It follows from [28, Proposition 7.46] that the sequence $\{\underset{\sim}{V}_0^{\mu_n} + \delta_{\{x_0\}\times\mathcal{P}_{\mathrm{ad}}}\}_{n\in\mathbb{N}}$ of functions epiconverges to $V_0 + \delta_{\{x_0\}\times\mathcal{P}_{\mathrm{ad}}}$. Then, as the effective domain $\mathrm{dom}(\underset{\sim}{V}_0^{\mu_n} + \delta_{\{x_0\}\times\mathcal{P}_{\mathrm{ad}}}) \subseteq \{x_0\}\times\mathcal{P}_{\mathrm{ad}}$ is bounded, the sequence $\{\underset{\sim}{V}_0^{\mu_n} + \delta_{\{x_0\}\times\mathcal{P}_{\mathrm{ad}}}\}_{n\in\mathbb{N}}$ is *eventually level-bounded* (see [28, Exercice 7.32(a)]). Therefore, by application of [28, Theorem 7.33], we obtain the convergence of the infimum

$$\inf_{p\in\mathcal{P}_{\mathrm{ad}}} \underset{\sim}{V}_0^{\mu_n}(x_0,p) \xrightarrow[n\to+\infty]{} \inf_{p\in\mathcal{P}_{\mathrm{ad}}} V_0(x_0,p) = \Phi^*\ .$$

Then, for any $n\in\mathbb{N}$, the inequality $\inf_{p\in\mathcal{P}_{\mathrm{ad}}} \underset{\sim}{V}_0^{\mu_n}(x_0,p) \leq \Phi^*$ follows from $\underset{\sim}{V}_0^{\mu_n} \leq \underset{\sim}{V}_0^{\mu_{n+1}}$ as shown in the proof of Proposition 3.13. □

Remark 3.15 *In Corollary 3.14, by application of [28, Theorem 7.33], we also obtain that all accumulation points of a sequence $\{p_n\}_{n\in\mathbb{N}}$ satisfying $p_n \in \arg\min_{p\in\mathcal{P}} \underset{\sim}{V}_0^{\mu_n}(x_0,p)$ for $n\in\mathbb{N}$ is a solution of $\inf_{p\in\mathcal{P}_{ad}}\Phi(p)$ in* (1)*. In particular, if Problem* (1) *has a unique solution $p^*\in\mathcal{P}_{ad}$, then $p_n \xrightarrow[n\to+\infty]{} p^*$.*

4 A numerical example in day-ahead power scheduling

We now consider a numerical example for the purpose of testing the resolution of Problem (1) with the differentiability and regularization methods introduced in §3.

First, in §4.1, we present our experimental assessment method that is based on the SDDP algorithm. Second, in §4.2, we introduce a parametric multistage stochastic optimization problem that represents the minimization of the expected intraday management cost of a solar plant. Third, in §4.3, we showcase the numerical results obtained.

4.1 Experimental assessment method based on SDDP

The principle of the Stochastic Dual Dynamic Programming (SDDP) algorithm [24] is to run a sequence of $k \in \mathbb{N}$ *forward* and *backward* passes in order to compute lower polyhedral value functions $\{\underline{V}_t^k\}_{t\in[\![0,T]\!]}$ which approximate the original value functions of a multistage convex stochastic optimization problem. At each pass $i \in [\![1, k-1]\!]$ of SDDP, the current approximation $\underline{V}_t^i$ is refined with a new cut (affine minorant), giving

$$\underline{V}_t^{i+1} = \sup\left\{\underline{V}_t^i\ ,\ \ \theta_t^i + \left\langle \cdot - \tilde{x}_t^i\,, \lambda_t^i \right\rangle\right\}\ , \tag{34}$$

where, denoting by $\tilde{x}$ the state variable, $\tilde{x}_t^i$ is a state value computed during the ith forward pass, and $(\theta_t^i, \lambda_t^i)$ are the cut parameters computed during the ith backward pass.

In our experiment, we use SDDP for two different purposes.

Approximate first-order oracle. In order to solve Problem (1), we propose an approximate first-order oracle based on SDDP. We compute a lower polyhedral approximation $\underline{V}_0^k$ of the parametric value function V_0 in (8) by running $k \in \mathbb{N}$ forward-backward passes of SDDP. In that case, the state variable is extended to $\tilde{x} = (x, p)$ in (34) together with a trivial stationary dynamics for its parameter component p. Then, we can build a (k)-oracle

$$p \mapsto \big(\underline{V}_0^k\big((x_0, p)\big), q \in \partial_p \underline{V}_0^k\big((x_0, p)\big)\big)\ , \tag{35}$$

computing an approximate value and an approximate subgradient of the value function V_0 by linear programming. Due to the convergence properties of SDDP [16], this approach lets us approximate the original value function V_0 as closely as desired, provided we run SDDP long enough. We refer to [18] for an analysis of the convergence rate of the SDDP algorithm.

Such convergence properties make this SDDP-based oracle a competitive alternative to the first-order oracle $p \mapsto \big(V_0(x_0, p), \nabla_p V_0(x_0, p)\big)$ which is based on the backward induction introduced in Theorem 3.7 — and to its regularized counterpart with $\underset{\sim}{V}_0^\mu$, following naturally from Theorem 3.12.

Estimation of $\Phi(p^*)$. Given a candidate optimal solution $p^* \in \mathcal{P}_{\text{ad}}$ for Problem (1a), we compute an estimation of $\Phi(p^*)$ with the SDDP algorithm. At this evaluation stage, the parameter p is not a component of the state variable but appears as a fixed parameter. We highlight this point by using the notation $(\cdot\,; p^*)$ in the arguments of ad hoc value functions, given by

$$\underline{V}_T(x; p^*) = K(x)\ ,\ \ \forall x \in \mathbb{R}^{n_x}\ , \tag{36a}$$

$$\underline{V}_t(x; p^*) = \inf_{u \in \mathcal{U}_t(x)} \mathbb{E}\Big[L_t(x, u, \mathbf{W}_{t+1}, p^*) + \underline{V}_{t+1}\big(f_t(x, u, \mathbf{W}_{t+1}); p^*\big)\Big]\ , \tag{36b}$$

$$\forall x \in \mathbb{R}^{n_x}\ ,\ \ \forall t \in [\![0, T-1]\!]\ ,$$

where the state variable is reduced to $\tilde{x} = x$ in (34). Once obtained the value functions $\underline{V}_t$ by SDDP, we use the resulting policy

$$\pi_t(x; p^*) \in \arg\min_{u \in \mathcal{U}_t(x)} \mathbb{E}\Big[L_t(x, u, \mathbf{W}_{t+1}, p^*) + \underline{V}_{t+1}\big(f_t(x, u, \mathbf{W}_{t+1}); p^*\big)\Big]\ , \tag{37}$$

$$\forall x \in \mathbb{R}^{n_x}\ ,\ \ \forall t \in [\![0, T-1]\!]\ ,$$

to compute the expected simulation cost

$$\overline{V}_0(x_0) = \mathbb{E}\Big[\sum_{t=0}^{T-1} L_t\big(\mathbf{X}_t, \pi_t(\mathbf{X}_t; p^*), \mathbf{W}_{t+1}, p^*\big) + K(\mathbf{X}_T)\Big]\ . \tag{38}$$

Since SDDP provides polyhedral lower estimate functions $\{\underline{V}_t\}_{t \in [\![0,T]\!]}$ of the true value functions $\{V_t(\cdot, p^*)\}_{t \in [\![0,T]\!]}$ defined by (8), and since $\{\pi_t\}_{t \in [\![0,T-1]\!]}$ in (37) is a suboptimal policy for Problem (1), we have the inequality

$$\underline{V}_0(x_0; p^*) \leq \Phi(p^*) \leq \overline{V}_0(x_0; p^*)\ . \tag{39}$$

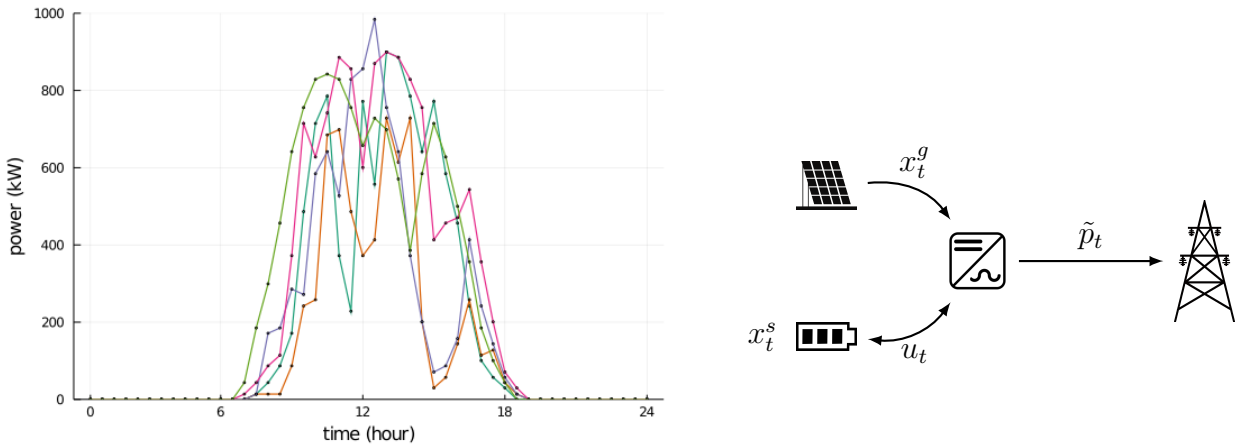
In practice, we compute $\{\underline{V}_t\}_{t \in [\![0,T]\!]}$ by running $k \in \mathbb{N}$ forward-backward passes of the SDDP algorithm, and the expectation in (38) is computed by Monte-Carlo simulation, generating scenarios with the discrete probability distributions of the noise variables $\{\mathbf{W}_t\}_{t \in [\![1,T]\!]}$ in (5c).

4.2 Problem statement

We introduce a PMSOP as defined in (1) for the daily management of a solar plant. Our example is inspired by the French regulation for day-ahead power scheduling of renewable units in islanded and overseas territories [14].

The schematic organization of the solar plant is given in Figure 2b. The main components of the power unit are a DC/AC inverter, a solar panel of installed peak power $\overline{p} \in \mathbb{R}_+^*$ (MW) and a lithium-ion battery characterized by the coefficients $(\kappa, \underline{u}, \overline{u}, \rho_c, \rho_d)$ referring respectively to the battery's capacity (MWh), minimum load (MW), maximum load (MW), charge and discharge efficiency coefficients.

We consider the time span of one operating day, with time intervals of length $\Delta_t = 30$ minutes, hence a problem horizon of $T = 48$. We now introduce all components of a dynamical system to formalize our management problem.



(a) Example of daily generated power scenarios

(b) Schematic power plant

Figure 2: Description of the solar power plant: (a) examples of daily generated power scenarios (data from [1]) and (b) schematic organization

Variables and parameter. We introduce the state variable

$$x_t = \begin{pmatrix} x_t^s \\ x_t^g \end{pmatrix} \in \mathbb{R}^2 \ , \ \ \forall t \in [\![0, T]\!] \ , \tag{40}$$

where $x_t^s \in [0, 1]$ is the state of charge of the battery and $x_t^g \in [0, \overline{p}]$ is the generated power of the solar panels, both observed at stage $t \in [\![0, T]\!]$. The control

$$u_t \in [\underline{u}, \overline{u}] \ , \ \ \forall t \in [\![0, T-1]\!] \ , \tag{41}$$

taken at the beginning of every time interval $[t, t+1[$, accounts for the charging power ($u_t \geq 0$) or discharging power ($u_t \leq 0$) applied to the battery during $[t, t+1[$. Lastly, we introduce the noise variable

$$w_t \in \mathbb{R} \ , \ \ \forall t \in [\![1, T]\!] \ , \tag{42}$$

to represent uncertainties in the evolution of the generated power x_t^g. All in all, the two dimensional state x_t evolves according to the dynamics[6]

$$x_0 = \begin{pmatrix} x_0^s \\ 0 \end{pmatrix} \ , \ \ x_{t+1} = f_t(x_t, u_t, w_{t+1}) \ , \ \ \forall t \in [\![0, T-1]\!] \ , \tag{43a}$$

whose components are given by

$$f_t(x, u, w) = \begin{pmatrix} f(x^s, u) \\ \alpha_t x^g + \beta_t + w \end{pmatrix} \ , \ \ \forall (x, u, w) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R} \ , \ \ \forall t \in [\![0, T-1]\!] \ , \tag{43b}$$

$$f(x^s, u) = x^s + \left(\rho_c u^+ - \frac{1}{\rho_d} u^-\right) \frac{\Delta_t}{\kappa} \ , \ \ \forall (x^s, u) \in [0, 1] \times \mathbb{R} \ , \tag{43c}$$

[6]The initial value $x_0^g = 0$ reflects the absence of sun at midnight in our use case.

where $u^+ = \max(0, u)$ and $u^- = \max(0, -u)$; f in (43c) is the dynamics of the state of charge of the battery; (α_t, β_t) in (43b) are the weights of the linear dynamics (AR(1) model) of the generated power at stage $t \in [\![0, T-1]\!]$.

We model the uncertainty of the noise variable w_t in (42) with a stochastic noise process $\mathbf{W} = \{\mathbf{W}_t\}_{t \in [\![1,T]\!]}$ as in (5c), that we assume to be stagewise independent with finitely supported random variables $\mathbf{W}_t$, $t \in [\![1, T]\!]$ — in line with the discrete white noise Assumption 2.3— and also introduce a state process $\mathbf{X} = \{\mathbf{X}_t\}_{t \in [\![0,T]\!]}$ as in (5a) and a control process $\mathbf{U} = \{\mathbf{U}_t\}_{t \in [\![0,T-1]\!]}$ as in (5b)[7].

Lastly, for each time interval $[t, t+1[$, the solar plant is engaged to deliver a certain value of committed power $p_t \in \mathbb{R}$, composing all together the parameter

$$p = \{p_t\}_{t \in [\![0,T-1]\!]} \in \mathbb{R}^T \ . \tag{44}$$

Constraints and costs. Controls are constrained by the admissibility sets

$$\mathcal{U}_t(x) = \left\{ u \in \mathbb{R} \,\middle|\, \underline{u} \leq u \leq \overline{u} \text{ and } 0 \leq f(x^s, u) \leq 1 \right\} \ , \ \ \forall x \in \mathbb{R}^2 \ , \ \ \forall t \in [\![0, T-1]\!] \ . \tag{45}$$

Besides, given the stagewise independence assumption on noises and the expression of the state variable x_t in (40), the nonanticipativity constraint in (1e) can be reformulated without loss of optimality [11, §4.4] as

$$\sigma(\mathbf{U}_t) \subseteq \sigma(\mathbf{X}_t) \ , \ \ \forall t \in [\![0, T-1]\!] \ . \tag{46}$$

Stage costs depend on the delivered power $\tilde{p}_{t+1}$ over the interval $[t, t+1[$, given by

$$\tilde{p}_{t+1} = x^g_{t+1} - u_t = \alpha_t x^g_t + \beta_t + w_{t+1} - u_t \in \mathbb{R} \ , \ \ \forall t \in [\![0, T-1]\!] \ . \tag{47}$$

Thus, for $t \in [\![0, T-1]\!]$ and $(x, u, w, p_t) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, we define stage costs as

$$L_t(x, u, w, p) = L^{\mathrm{e}}_t(\alpha_t x^g + \beta_t + w - u) + L^{\mathrm{p}}_t(\alpha_t x^g + \beta_t + w - u, p_t) \ , \tag{48a}$$

with the energy cost L^{e}_t and the penalty cost L^{p}_t given by

$$L^{\mathrm{e}}_t(\tilde{p}_{t+1}) = -c_t \Delta_t \tilde{p}_{t+1} \ , \tag{48b}$$

$$L^{\mathrm{p}}_t(\tilde{p}_{t+1}, p_t) = \lambda c_t \Delta_t |\tilde{p}_{t+1} - p_t| \ , \tag{48c}$$

where c_t is the (deterministic) energy price, expressed in €/MWh, for $t \in [\![0, T]\!]$ and $\lambda \geq 1$ is a penalty coefficient. The final cost is then defined as

$$K(x) = -c_T x^s_T \kappa \ , \ \ \forall x \in \mathbb{R}^2 \ . \tag{49}$$

The cost structure defined by (48)-(49) reflects the original formulation in [14], except for the penalty cost L^{p}_t in (48c) that we have simplified for this illustrative example.

[7] With the dynamics f_t in (43b), we can only guarantee $\mathbb{P}(\mathbf{X}^g_t \in [0, \overline{p}]) \geq 1 - \epsilon$ for some $\epsilon > 0$. In practice, we round values $x^g \notin [0, \overline{p}]$ by projecting them to $[0, \overline{p}]$.

Optimization problem. Gathering all components introduced above, we define the *intraday value function* Φ as the value of the PMSOP defined by (1). Finally, we introduce the *set of admissible parameter profiles*

$$\mathcal{P}_{\mathrm{ad}} = [0, \overline{p}]^T \ , \tag{50}$$

and we consider the day-ahead optimization of the expected management cost in (48a) with respect to the commitment profile $p \in \mathcal{P}_{\mathrm{ad}}$, that is, we want to solve

$$\min_{p \in \mathcal{P}_{\mathrm{ad}}} \Phi(p) \ . \tag{51}$$

4.3 Numerical experiments

We perform numerical experiments with a single computer equipped with 4 Intel Core i7-7700K CPU and 15 GB of RAM. We use the package `SDDP.jl` [15] for SDDP together with the LP solver of CPLEX 12.9. Apart from the solver, all our code is implemented with the Julia language [8]. Further implementation details are provided in Appendix A.1.

First, we describe our experimental protocol, where we introduce two candidate methods to address Problem (51). Second, we comment on the results of the two methods separately, before finally confronting the results of both methods.

Experimental protocol. Our goal is to evaluate optimization methods for solving Problem (51). The objective function $\Phi = V_0(x_0, \cdot)$ is convex (see Theorem 3.7) and polyhedral (due to the polyhedral penalty L_t^{p} in (48c) and other affine constraints, see e.g. arguments in [31, §3.2.1]), hence nondifferentiable. We consider two methods.

- ***μ*SDP**: we address Problem (51) by solving

$$\min_{p \in \mathcal{P}_{\mathrm{ad}}} \ \underset{\sim}{V}_0^{\mu}(x_0, p) \ , \tag{52}$$

  where the lower smooth value function $\underset{\sim}{V}_0^{\mu}$ is defined by (28). In this method, we embed the first order oracle $p \mapsto \big(\underset{\sim}{V}_0^{\mu}(x_0, p), \nabla_p \underset{\sim}{V}_0^{\mu}(x_0, p)\big)$ within a projected gradient descent algorithm [6]. We evaluate the function $\underset{\sim}{V}_0^{\mu}(x_0, \cdot)$ in (28) by Stochastic Dynamic Programming (SDP) and the gradient $\nabla_p \underset{\sim}{V}_0^{\mu}(x_0, \cdot)$ using the Bellman-like induction (23) of Theorem 3.12. We use a constant value for the regularization coefficient, namely, $\mu = 0.1$.

- ***k*SDDP**: we address Problem (51) by solving

$$\min_{p \in \mathcal{P}_{\mathrm{ad}}} \ \underline{V}_0^{k}(x_0, p) \ , \tag{53}$$

  where the lower polyhedral value function $\underline{V}_0^k$ is defined by (34) with $\tilde{x} = (x, p)$. In this method, we embed the first order oracle $p \mapsto \big(\underline{V}_0^k\big((x_0, p)\big), q \in \partial_p \underline{V}_0^k\big((x_0, p)\big)\big)$ within a projected subgradient algorithm [3, §8.2]. We evaluate the function $\underline{V}_0^k\big((x_0, \cdot)\big)$ and an element of the subdifferential $\partial_p \underline{V}_0^k\big((x_0, \cdot)\big)$ after running $k \in \mathbb{N}$ forward-backward passes of the SDDP algorithm, following the method developed in §4.1.

For both methods, we consider several implementations, playing on the coefficients that affect the precision in the approximation of Φ and the computing time. For μSDP, we have 12 implementations, characterized by the size of the discrete grids for state and control variables.[8] For kSDDP, we have 9 implementations, characterized by the number of forward-backward passes $k \in \mathbb{N}$ run by the SDDP algorithm. Each implementation of a method is processed as follows.

(i) First, we compute a commitment profile $p^* \in \mathcal{P}_{\text{ad}}$ as a solution of (52) or (53). We initialize $p_0 = 0 \in \mathbb{R}^T$ and use a dynamic step size of $\eta_i = \frac{10^3}{i}$ for both methods μSDP and kSDDP. Note that (52) is a smooth problem, and we could expect better performance with line search. Yet, we keep the same step size for both methods to focus the experiment on the performance of the oracles. As for our stopping rule, we stop the computation if we exceed 100 iterations, or if the progress of the objective value is not larger than $\pm 0.5\%$ for 5 consecutive iterations. We report time performance for computing p^*, both in term of overall computing time (Figure 3, X-axis in log scale, the lower the better) and in term of average time per oracle call (Figure 4, idem).

(ii) Second, we compute an estimation of $\Phi(p^*)$ with SDDP, following the methodology introduced in §4.1. In Figure 3 and Figure 4, we report cost performance (Y-axis, the lower the better), where the height of a marker spans over the interval $[\underline{V}_0(x_0; p^*), \overline{V}_0(x_0; p^*)]$ on the Y-axis, representing an estimation of the intraday value $\Phi(p^*)$, from the inequality (39). We report that, for each implementation, the gap between the lower bound $\underline{V}_0(x_0; p^*)$ in (36) — computed by running 2000 forward-backward passes of SDDP— and the expected simulation cost $\overline{V}_0(x_0; p^*)$ in (38) — computed by Monte-Carlo simulation, sampling 25000 scenarios — is lower than 1.7%.

Further detailed results are available in Appendix A.2.

Results of μSDP. We comment on the results of implementations of μSDP, represented by blue markers on Figure 3 and Figure 4. The performance for this method is related to the size of the discrete grid introduced for the state and control variables (x, u) in (40) and (41). In this experiment, we use 12 grid sizes ranging from (5×5, 11) points to (101×101, 201) points.

We expect that the finer the grid discretization, the more accurate the computation of the value function $\underset{\sim}{V}_0^\mu$ and of the gradient $\nabla_p \underset{\sim}{V}_0^\mu$, but also the longer the computing time per oracle call. Indeed, the worst cost performance (highest estimated value of $\Phi(p^*)$ on the Y-axis of Figure 3 and Figure 4) is obtained with the coarsest grid discretization. Although this implementation has the shortest average oracle time (lowest value on the X-axis of Figure 4), it requires more iterations of the projected gradient method to stabilize, and is not the fastest implementation in term of overall computing time (second lowest value on the X-axis of Figure 3). More surprisingly, we find that the cost performance is greatly improved by only adding one grid point to the state space of the worst cost-performing implementation (marker with the second lowest value on the X-axis of Figure 4), and that cost performance stabilizes for further refinements of the grid despite longer average oracle

[8] We recall that the regularization coefficient μ is fixed and equal to 0.1 for all 12 implementations.

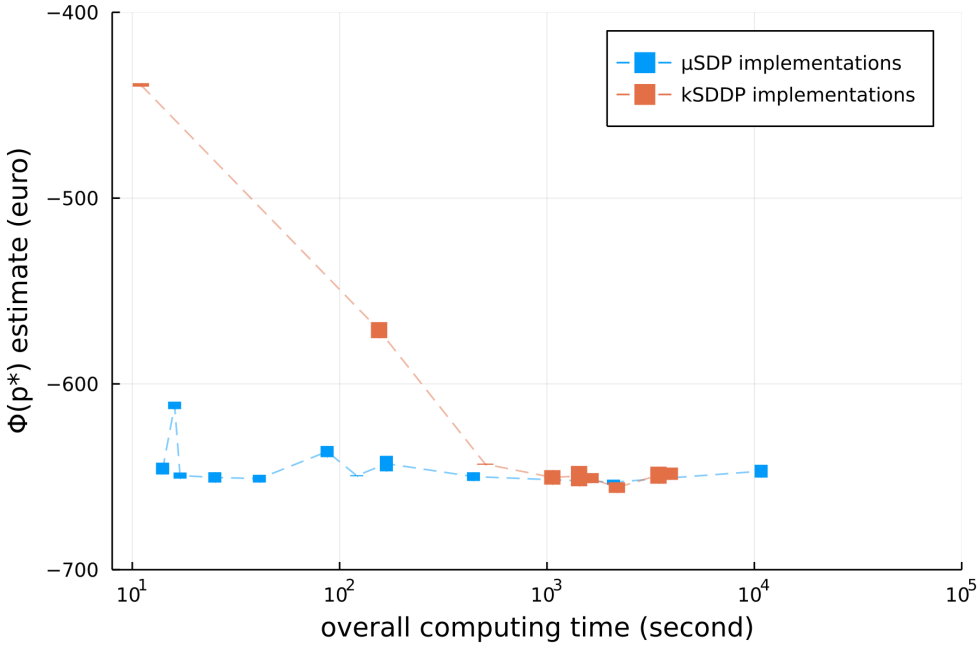


Figure 3: Estimate of $\Phi(p^*) \in [\underline{V}_0(x_0), \overline{V}_0(x_0)]$ in (39) (marker span on the Y-axis) for implementations of μSDP (in blue) and kSDDP (in orange); and overall computing time for the computation of p^* (X-axis in log scale). For both axis, the lower the values the better

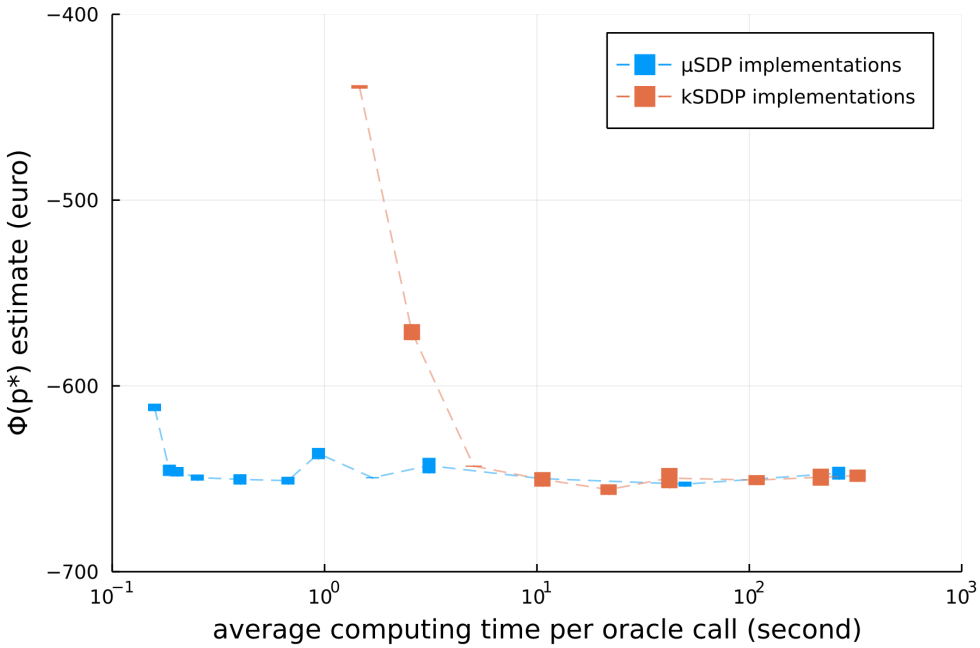


Figure 4: Estimate of $\Phi(p^*) \in [\underline{V}_0(x_0), \overline{V}_0(x_0)]$ in (39) (marker span on the Y-axis) for implementations of μSDP (in blue) and kSDDP (in orange); and average computing time per oracle call for the computation of p^* (X-axis in log scale). For both axis, the lower the values the better

times (third to twelfth markers on the same axis).

Results of *k*SDDP. We comment on the results of implementations of kSDDP, represented by orange markers on Figure 3 and Figure 4. The performance of these implementations is related to the number $k \in \mathbb{N}^*$ of forward-backward passes performed by the SDDP algorithm at each iteration of the subgradient method. We report results for a total of 9 implementations, with $k \in \{10, 20, 40, 80, 150, 250, 500, 750, 1000\}$.

As expected, we observe that the more forward-backward passes of the SDDP algorithm we run, the more accurate the approximations of the value function V_0 by $\underline{V}_0^k$ and of the subdifferential $\partial_p V_0$ by $\partial_p \underline{V}_0^k$. This phenomenon is well illustrated by Figure 4, where better cost performances, corresponding to lower values on the Y-axis, are obtained with longer computing time per oracle call, characterized by higher values on the X-axis (where the higher the value of k, the higher the position of the marker for implementations of kSDDP). In general, we also report that low values of k increase the noise of the oracle. Indeed, for implementations of kSDDP with $k \in \{40, 80, 150\}$, the values of $\underline{V}_0^k(x_0, \cdot)$ do not stabilize after 100 iterations of the subgradient method, whereas implementations with $k \in \{250, 500, 750, 1000\}$ reach convergence in respectively $\{34, 15, 16, 12\}$ iterations. As for $k \in \{10, 20\}$, the SDDP algorithm only samples a few scenarios in these cases, and seems to fail to obtain accurate representations of the value function V_0. Observations on oracle stability are backed by more detailed results available in Appendix A.2.

Cross-method comparison. We compare the results obtained with μSDP versus the ones obtained with kSDDP. We observe in Figures 4 and 3 that μSDP (blue markers) almost attains its best cost performance with a value of the expected simulation cost $\overline{V}_0(x_0)$ in (38) of -648 € (upper value of the span of the markers on the Y-axis) with only 0.25 seconds spent per oracle call (Figure 4, X-axis) and only 17 seconds of overall computing time (Figure 3, X-axis). Comparatively, for kSDDP implementations (orange markers), we need to perform at least $k = 80$ forward-backward passes of the SDDP algorithm to attain $\overline{V}_0(x_0)$= -646 € (upper value of the span of the markers on the Y-axis). For this value of k, we spend on average 10.2 seconds per oracle call (Figure 4, X-axis), and the overall computing time is of 1,061 seconds (Figure 3, X-axis). We conclude that the μSDP method performs better than kSDDP in our experiments, both in term of time performance and in term of cost performance. In particular, we argue that this result illustrates the pertinence of treating the parameter p in (3) apart from the state variables to address Problem (1), especially when the parameter space $\mathbb{R}^{n_p}$ is of much larger dimension than the state space $\mathbb{R}^{n_x}$.

5 Conclusion

We have studied the differentiability properties of a class of parametric multistage stochastic optimization problems. Our main finding is that, under differentiability and convexity

assumptions, we manage to compute the gradient of the value function of the problem with respect to the parameter by backward induction.

In the case where the differentiability assumption is not fulfilled, we have proposed a method for obtaining lower approximations of the original parametric value functions by regularization. We have also shown that such approximations let us approach the value and a solution of the original problem as closely as desired — provided the regularization coefficient is sufficiently small.

A numerical test case inspired from power scheduling reveals that our gradient computation technique is efficient to formulate first-order oracles in convex parametric multistage stochastic optimization.

In particular, we report that treating the parameter apart from the state variables — as proposed in our backward induction for gradients — can be advantageous in term of computing time performance, especially when the problem is structured with a large parameter space.

A Appendix to numerical experiments

A.1 Data and implementation details

First, we detail the implementation of the probabilistic model introduced in §4.2. We use one year of historical photovoltaic power data from the public platform of the Australian transmission system operator Ausgrid [1]. We scale the generated power data to simulate the operating of a solar power plant with an installed peak power $\overline{p} = 1$ MW. Then, we use a standard linear regression to calibrate the weights of the linear model in (43b), and we perform a quantization of the support of the error noise process $\{\mathbf{W}_t\}_{t\in[\![1,T]\!]}$ with the K-means algorithm. This latter technique lets us compute discrete probability distributions for each random variable in $\{\mathbf{W}_t\}_{t\in[\![1,T]\!]}$. We refer the reader to [29] for the theoretical motivations of this quantization scheme.

Second, we provide implementation details for the oracles of the two methods introduced in §4.3. For the oracle of the μSDP method, we implement the backward recursions in (28) (for value functions) and in (23) (for gradients) by parallelizing computations across a discrete grid of states. We also use a discrete grid for controls. As for the oracle of the kSDDP method, we use the built-in parallelization scheme of `SDDP.jl` to run forward-backward passes in asynchronous mode.

Lastly, concerning other parameters, we take $\kappa = 1$ MWh, $\overline{u} = -\underline{u} = 1$ MW, and $\rho_c = \rho_d = 0.95$ for the battery parameters; $c_t = 0.4$ € for the off-peak energy price and $c_t = 0.6$ € for the on-peak energy price, with a 2-hours peak spanning over [19:00, 21:00]; and $\lambda = 2$ for the penalty cost in (48c).

A.2 Detailed results for all implementations

We provide additional details on the numerical results of each implementations of the μSDP and kSDDP methods considered in the experiments of §4.3.

(x, u) grid size	Iterative steps	Overall time (seconds)	Avg. time / oracle call (seconds)	$\underline{V}_0(x_0)$ in (36) (€)	$\overline{V}_0(x_0)$ in (38) (€)	Gap (%)
5×5, 11	97	16	0.16	-613.6	-609.6	0.7
5×6, 11	78	14	0.19	-648.5	-642.4	0.9
6×6, 11	69	14	0.20	-648.8	-643.6	0.8
6×6, 21	66	17	0.25	-651.0	-647.7	0.5
6×11, 21	63	25	0.40	-653.1	-647.5	0.9
11×11, 21	54	41	0.67	-653.1	-648.9	0.6
11×11, 41	93	87	0.94	-639.4	-633.4	0.9
11×21, 41	72	121	1.7	-649.7	-649.1	0.1
21×21, 41	54	168	3.1	-647.1	-638.7	1.3
21×21, 201	42	442	10.5	-652.3	-638.7	1.3
21×101, 201	42	2092	49.8	-654.2	-651.5	0.4
101×101, 201	41	10781	263.4	-650.4	-643.5	1.0

Table 1: Detailed numerical performances for implementations of the μSDP method, characterized by the size of the discrete grids for state and control variables in the first column. Other columns report the number of iterations performed (second column), time performances (third and fourth columns), together with the lower bound $\underline{V}_0(x_0)$ in (36) (fifth column), the expected simulation cost $\overline{V}_0(x_0)$ in (38) (sixth column), and the estimation gap (seventh column) expressed as a percentage of $\underline{V}_0(x_0)$. For columns 2-6, the lower the values the better the performance of the instance

k	**Iterative steps**	**Overall time** (seconds)	**Avg. time / oracle call** (seconds)	$\underline{V}_0(x_0)$ **in** (36) (€)	$\overline{V}_0(x_0)$ **in** (38) (€)	**Gap** (%)
10	7	11	1.5	-440.0	-438.1	0.4
20	60	155	2.6	-575.4	-566.7	1.5
40	100	504	5.0	-643.5	-642.9	0.1
80	100	1061	10.6	-654.2	-646.4	1.2
150	100	2173	21.7	-658.7	-652.9	0.8
250	34	1428	42.0	-655.1	-644.1	1.7
500	15	1622	108.2	-653.4	-648.0	0.8
750	16	3448	216.9	-653.7	-644.5	1.4
1000	12	3912	323.4	-651.6	-645.0	1.0

Table 2: Detailed numerical performances for implementations of the kSDDP method, characterized by k in the first column. Other columns report the number of iterations performed (second column), time performances (third and fourth columns), together with the lower bound $\underline{V}_0(x_0)$ in (36) (fifth column), the expected simulation cost $\overline{V}_0(x_0)$ in (38) (sixth column), and the estimation gap (seventh column) expressed as a percentage of $\underline{V}_0(x_0)$. For columns 2-6, the lower the values the better the performance of the instance

References

[1] Ausgrid. Solar home electricity data. URL `https://www.ausgrid.com.au/Industry/Our-Research/Data-to-share/Solar-home-electricity-data`.

[2] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer-Verlag, second edition, 2017.

[3] A. Beck. *First-order methods in optimization*. SIAM, 2017.

[4] R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.

[5] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch. A comparison of interior point methods and a Moreau-Yosida based active set strategy for constrained optimal control problems. In *SIAM Journal on Optimization*. Citeseer, 1998.

[6] D. P. Bertsekas. On the Goldstein-Levitin-Polyak gradient projection method. *IEEE Transactions on automatic control*, 21(2):174–184, 1976.

[7] D. P. Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena scientific Belmont, MA, 1995.

[8] J. Bezanson, S. Karpinski, V. B. Shah, and A. Edelman. Julia: A fast dynamic language for technical computing. *arXiv preprint arXiv:1209.5145*, 2012.

[9] J. F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer, 2013.

[10] D. B. Brown and J. Zhang. On the strength of relaxations of weakly coupled stochastic dynamic programs. *Operations Research*, 2022. doi: 10.1287/opre.2022.2287.

[11] P. Carpentier, J.-P. Chancelier, G. Cohen, and M. De Lara. Stochastic multi-stage optimization. Springer, 2015.

[12] P. Carpentier, J.-P. Chancelier, M. De Lara, and F. Pacaud. Mixed spatial and temporal decompositions for large-scale multistage stochastic optimization problems. *Journal of Optimization Theory and Applications*, 186(3):985–1005, 2020.

[13] Z. Cen, J. F. Bonnans, and T. Christel. Sensitivity analysis of energy contracts management problem by stochastic programming techniques. In R. Carmona, P. D. Moral, P. Hu, and N. Oudjane, editors, *Numerical Methods in Finance*, volume 12 (2012) of *Springer Proceeding in Mathematics*, pages 447–471. Springer, 2012.

[14] Commission de régulation de l'énergie (CRE). Cahier des charges des appels d'offres portant sur la réalisation et l'exploitation d'installations de production d'électricité à partir de l'énergie solaire et situées dans les zones non interconnectées, 2019.

[15] O. Dowson and L. Kapelevich. SDDP.jl: a Julia package for stochastic dual dynamic programming. *INFORMS Journal on Computing*, 2020. doi: https://doi.org/10.1287/ijoc.2020.0987. Articles in Advance.

[16] P. Girardeau, V. Leclère, and A. B. Philpott. On the convergence of decomposition methods for multistage stochastic convex programs. *Mathematics of Operations Research*, 40(1):130–145, 2015.

[17] V. Guigues, A. Shapiro, and Y. Cheng. Duality and sensitivity analysis of multistage linear stochastic programs. *arXiv preprint arXiv:1911.07080*, 2019.

[18] G. Lan. Complexity of stochastic dual dynamic programming. *Mathematical Programming*, pages 1–38, 2020.

[19] V. Leclère, P. Carpentier, J.-P. Chancelier, A. Lenoir, and F. Pacaud. Exact converging bounds for stochastic dual dynamic programming via Fenchel duality. *SIAM Journal on Optimization*, 30(2):1223–1250, 2020.

[20] B. S. Mordukhovich and N. M. Nam. *An Easy Path to Convex Analysis and Applications*. Morgan & Claypool Publishers, 1st edition, 2013. ISBN 1627052372.

[21] J.-J. Moreau. Proximité et dualité dans un espace Hilbertien. *Bulletin de la Société mathématique de France*, 93:273–299, 1965.

[22] A. N'Goran, B. Daugrois, M. Lotteau, and S. Demassey. Optimal engagement and operation of a grid-connected PV/battery system. In *2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pages 1–5. IEEE, 2019.

[23] R. I. Ortega-Gutiérrez and H. Cruz-Suárez. A Moreau-Yosida regularization for Markov decision processes. *Proyecciones (Antofagasta)*, 40(1):117–137, 2021.

[24] M. V. Pereira and L. M. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical programming*, 52(1):359–375, 1991.

[25] P. Pflaum, M. Alamir, and M. Y. Lamoudi. Battery sizing for PV power plants under regulations using randomized algorithms. *Renewable Energy*, 113:596–607, 2017.

[26] A. B. Philpott and Z. Guan. On the convergence of stochastic dual dynamic programming and related methods. *Operations Research Letters*, 36(4):450–455, 2008.

[27] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., 1rst edition, 1994. ISBN 0471619779.

[28] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317. Springer, 2009.

[29] N. Rujeerapaiboon, K. Schindler, D. Kuhn, and W. Wiesemann. Scenario reduction revisited: Fundamental limits and guarantees. *Mathematical Programming*, pages 1–36, 2018.

[30] A. Shapiro. Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209(1):63–72, 2011.

[31] A. Shapiro, D. Dentcheva, and A. Ruszczynski. *Lectures on Stochastic Programming: Modeling and Theory, Second Edition*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2014. ISBN 1611973422, 9781611973426.

[32] D. Team-Drives, A. Energy, A. Cervone, E. Santini, S. Teodori, D. Z. Romito, et al. Impact of regulatory rules on economic performance of pv power plants. *Renewable Energy*, 74:78–86, 2015.

[33] G. Terça and D. Wozabal. Envelope theorems for multistage linear stochastic optimization. *Operations Research*, 2020.

[34] K. Yosida. *Functional analysis*. Springer Berlin Heidelberg, 1971.