



HAL
open science

A Min-plus-SDDP Algorithm for Deterministic Multistage Convex Programming

Marianne Akian, Jean-Philippe Chancelier, Benoît Tran

► **To cite this version:**

Marianne Akian, Jean-Philippe Chancelier, Benoît Tran. A Min-plus-SDDP Algorithm for Deterministic Multistage Convex Programming. CDC 2019 - 58th IEEE Conference on Decision and Control, Dec 2019, Nice, France. hal-02436343

HAL Id: hal-02436343

<https://inria.hal.science/hal-02436343>

Submitted on 13 Jan 2020

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.

A Min-plus-SDDP Algorithm for Deterministic Multistage Convex Programming

Marianne Akian, Jean-Philippe Chancelier and Benoît Tran

Abstract—We consider discrete time optimal control problems with finite horizon involving continuous states and possibly both continuous and discrete controls, subject to non-stationary linear dynamics and convex costs. In this general framework, we present a stochastic algorithm which generates monotone approximations of the value functions as a pointwise supremum or infimum of basic functions (for example affine or quadratic) which are randomly selected. We give sufficient conditions on the way basic functions are selected in order to ensure almost sure convergence of the approximations to the value function on a set of interest. Then we study a linear-quadratic optimal control problem with one control constraint. On this toy example we show how to use our algorithm in order to build lower approximations, like the SDDP algorithm, as supremum of affine cuts and upper approximations, by min-plus techniques, as infimum of quadratic functions.

Index Terms—Deterministic multistage optimization problems, min-plus algebra, tropical algebra, Stochastic Dual Dynamic Programming, Dynamic Programming.

I. INTRODUCTION

Throughout this paper, we aim to solve a deterministic optimal control problem with discrete time. Informally, given a time t and a state x_t , one can apply a control u_t and the next state is given by the dynamic f_t , that is $x_{t+1} = f_t(x_t, u_t)$. Then, one wants to minimize the sum of costs $c_t(x_t, u_t)$ induced by the controls starting from a given state x_0 and during a given time horizon T . Furthermore, one can add some final restrictions on the states at time T which will be modeled by an additional cost function ψ depending only on the final state x_T . We will call such optimal control problems, *multistage optimization problems* and *switched multistage optimization problems* if the controls are both continuous and discrete:

$$\begin{aligned} \min_{\substack{x=(x_0, \dots, x_T) \in \mathbb{X}^{T+1} \\ u=(u_0, \dots, u_{T-1}) \in \mathbb{U}^T}} & \sum_{t=0}^{T-1} c_t(x_t, u_t) + \psi(x_T) \\ \text{s.t. } & x_0 \in \mathbb{X} \text{ given and } \forall t \in \llbracket 0, T-1 \rrbracket, x_{t+1} = f_t(x_t, u_t). \end{aligned} \quad (1)$$

A. Dynamic Programming and value functions

One can solve the multistage optimization problem (1) by Dynamic Programming as introduced by Richard Bellman

M. Akian is with INRIA and CMAP, École polytechnique. Address: CMAP, Ecole polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France. Email: marianne.akian@inria.fr.

J.-P. Chancelier is with CERMICS, École des Ponts ParisTech. Address: École des Ponts ParisTech, CERMICS, 6 et 8 av. B. Pascal, 77455 Marne-la-Vallée, France. Email: jpc@cermics.enpc.fr.

Corresponding author B. Tran is with CERMICS, École des Ponts ParisTech, INRIA and CMAP, École polytechnique. Address: École des Ponts ParisTech, CERMICS, 6 et 8 av. B. Pascal, 77455 Marne-la-Vallée, France. Email: benoit.tran@enpc.fr.

around 1950 [1]. This method breaks the multistage optimization problem (1) into T sub-problems that one can solve by backward recursion over time. More precisely, denote by \mathbb{X} the set of states, by \mathbb{U} the control space, and by \mathcal{B}_t the operator from the set $\overline{\mathbb{R}}^{\mathbb{X}}$ of functions over \mathbb{X} that take values in the extended real line, to itself, defined by

$$\mathcal{B}_t(\phi) : x \mapsto \min_{u \in \mathbb{U}} (c_t(x, u) + \phi(f_t(x, u))) \quad , \quad (2)$$

for all $\phi \in \overline{\mathbb{R}}^{\mathbb{X}}$. One can show that solving problem (1) amounts to solve the following sequence of sub-problems:

$$V_T = \psi \text{ and } \forall t \in \llbracket 0, T-1 \rrbracket, V_t = \mathcal{B}_t(V_{t+1}) \quad . \quad (3)$$

We will call each operator \mathcal{B}_t the *Bellman operator* at time t and each equation in (3) will be called the *Bellman equation* at time t . Lastly, the function V_t defined in (3) will be called the (Bellman) *value function* at time t . Note that the value of (1) is equal to the value function V_0 at point x_0 , that is $V_0(x_0)$, whereas solving Equation (3) means to compute the value functions V_t at each point $x \in \mathbb{X}$ and time $t \in \llbracket 0, T-1 \rrbracket$. Under mild technical assumptions, which are detailed in the Appendix, we devise an algorithm to approximate the solution of the system of Bellman equations (3) (also called the Dynamic Programming formulation of the multistage optimization problem). One issue of using Dynamic Programming to solve multistage optimization problems is the so-called *curse of dimensionality* [1], that is, grid-based methods to compute the value functions have a complexity exponential in the dimension of the state space.

B. Stochastic Dual Dynamic Programming

One popular algorithm (see [2], [3], [4]) that aims to dampen the curse of dimensionality is the Stochastic Dual Dynamic Programming algorithm (or SDDP for short) introduced by Pereira and Pinto [3]. Assuming that the cost functions c_t are convex and the dynamics f_t are linear, the value functions defined in the Dynamic Programming formulation (3) are convex [2]. Under these assumptions, the SDDP algorithm aims to build lower (or outer) approximations of the value functions as a supremum of affine functions and thus, doesn't rely on a discretisation of the state space in order to compute (approximations of) the value functions. In the aforementioned references, this approach is used to solve stochastic multistage convex optimization problems, however in this article we will restrict our study to deterministic multistage convex optimization problems, that is, the above formulation (1). Still, the SDDP algorithm can be applied to our framework. One of the main drawback of the SDDP algorithm (in the stochastic case) is the lack of an

efficient stopping criterion: it builds lower approximations of the value functions but upper (or inner) approximations are built through a Monte-Carlo scheme that is costly and the associated stopping criteria is not deterministic. We follow another path to provide upper approximations as explained now.

C. The min-plus approach of McEneaney

In [5], Qu devised an algorithm which builds upper approximations of the value functions in an infinite horizon and continuous time framework where the set of controls is both discrete and continuous. This work was inspired by the work of McEneaney [6] using techniques coming from tropical algebra, also called min-plus techniques. Assume that for each fixed discrete control the cost functions are convex quadratic and the dynamics are linear. If the set of discrete controls is finite, then exploiting the min-plus linearity of the Bellman operators, one can show that the value functions can be computed as a finite pointwise infimum of convex quadratic functions: $V_t = \inf_{\phi_t \in F_t} \phi_t$, where F_t is a finite set of convex quadratic forms. Moreover, in this framework, the elements of F_t can be explicitly computed through the Discrete Algebraic Riccati Equation ([7]). Thus, an approximation scheme that computes a non-decreasing sequence of subsets $(F_t^k)_{k \in \mathbb{N}}$ of F_t yields an algorithm that converges after a finite number of improvements

$$V_t^k := \inf_{\phi_t \in F_t^k} \phi_t \approx \inf_{\phi_t \in F_t} \phi_t = V_t. \quad (4)$$

However, the size of the set of functions F_t that need to be computed is growing exponentially with $T - t$. In [5], in order to address the exponential growth of F_t , Qu introduced a probabilistic scheme that adds to F_t^k the best (given the current approximations) element of F_t at some point drawn on the unit sphere (assuming the space of states \mathbb{X} to be Euclidean).

D. Contributions and structure of this article

Our work aims to build a general algorithm that encompasses both a deterministic version of the SDDP algorithm and an adaptation of Qu's work to a discrete time and finite horizon framework. For the sake of presentation, most technical statements and proofs can be found in the companion article [8]. In the following section we define an algorithm which builds approximations of the value functions as a pointwise optimum (i.e. either a pointwise infimum or a pointwise supremum) of basic functions in order to solve the associated Dynamic Programming formulation (3) of the multistage optimization problem (1). At each iteration, the so-called basic function that is added to the current approximation will have to satisfy two key properties at a randomly drawn point, namely, *tightness* and *validity*. A key feature of our algorithm is that it can yield either upper or lower approximations, for example:

- If the basic functions are affine, then approximating the value functions by a pointwise supremum of affine functions will yield the SDDP algorithm.

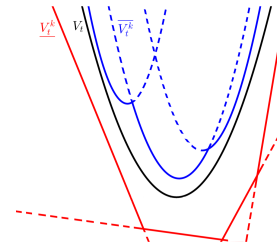


Fig. 1: The lower approximations V_t^k will be built as a supremum of basic functions (here affine functions) that remains below V_t . Likewise, the upper approximations \bar{V}_t^k will be built as an infimum of some other basic functions (here quadratic functions) that remains above V_t .

- If the basic functions are quadratic convex, then approximating the value functions by a pointwise infimum of convex quadratic functions will yield an adaptation of Qu's algorithm.

The convergence of the approximations of the value functions generated by our algorithm for every time $t \in \llbracket 0, T \rrbracket$ is stated and the main ideas of the proof are given. Under mild assumptions our approximating sequence converges almost surely (over the draws) to the value function on a set of interest (that will be specified). In this article we present numerical experiments of our algorithm on a toy example. Moreover, we show how a multistage linear-quadratic problem with constraints can be formulated as a switched linear-quadratic multistage optimization problem with homogeneous costs by adding a dimension to the state variable. We apply two different variants of our algorithm which can be computed in parallel in order to get simultaneously upper and lower approximations of the value functions. Then, we discuss on the effects of each parameters of such framework on the behavior, computational time and precision of the generated approximations.

II. A STOCHASTIC ALGORITHM FOR SOLVING MULTISTAGE OPTIMIZATION PROBLEMS

A. Notations

Denote by $\mathbb{X} := \mathbb{R}^n$ the continuous set of states endowed with its euclidean structure and its Borel σ -algebra, $\mathbb{U} = \mathbb{R}^m$ a continuous control space, T a finite integer that we will call the *horizon*. Denote opt an operation that is either the pointwise infimum or the pointwise supremum of functions which we will call the *pointwise optimum*. Denote $\bar{\mathbb{R}}$ the extended real line endowed with the operations $+\infty + (-\infty) = -\infty + \infty = +\infty$. For every $t \in \llbracket 0, T \rrbracket$, fix \mathbf{F}_t and \mathbb{F}_t two subsets of $\bar{\mathbb{R}}^{\mathbb{X}}$, the set of functions over \mathbb{X} , such that $\mathbf{F}_t \subset \mathbb{F}_t$. We will say that a function ϕ is a *basic function* if it is an element of \mathbf{F}_t for some $t \in \llbracket 0, T \rrbracket$. For every set $X \subset \mathbb{X}$, denote by δ_X the function equal to 0 on X and $+\infty$ elsewhere. For every $t \in \llbracket 0, T \rrbracket$ and every set of basic functions $F_t \subset \mathbf{F}_t$, we denote by \mathcal{V}_{F_t} its pointwise optimum, $\mathcal{V}_{F_t} := \text{opt}_{\phi \in F_t} \phi$, i.e. $\mathcal{V}_{F_t} = (x \mapsto \text{opt} \{ \phi(x) \mid \phi \in F_t \})$. Fix a function $\psi : \mathbb{X} \rightarrow \mathbb{R} \cup \{+\infty\}$, the *final cost function* and fix for every $t \in \llbracket 0, T - 1 \rrbracket$ the *cost functions* $c_t : \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{R} \cup \{+\infty\}$.

For every $t \in \llbracket 0, T-1 \rrbracket$, denote by \mathcal{B}_t the Bellman operator defined in Equation (2). For every $t \in \llbracket 0, T \rrbracket$, denote by V_t the value functions solutions of (3). Denote by \mathcal{S}_n^+ (resp. \mathcal{S}_n^{++}) the set of real $n \times n$ symmetric semidefinite (resp. definite) positive matrices.

B. Compatible selection functions

From a set of basic functions $F_t \subset \mathbf{F}_t$, one can build its pointwise optimum $\mathcal{V}_{F_t} = \text{opt}_{\phi \in F_t} \phi$. We build a monotone sequence of approximations of the value functions as optima of basic functions which will be computed through *compatible selection functions* as defined below.

Definition 1 (Compatible selection functions): Let a time step $t \in \llbracket 0, T-1 \rrbracket$ be fixed. A *compatible selection function* is a function ϕ_t^\sharp from $2^{\mathbf{F}_{t+1}} \times \mathbb{X}$ to \mathbf{F}_t which is

- **Valid:** for every set of basic functions $F_{t+1} \subset \mathbf{F}_{t+1}$ and every $x \in \mathbb{X}$, we have $\phi_t^\sharp(F_{t+1}, x) \leq \mathcal{B}_t(V_{F_{t+1}})$ (resp. $\phi_t^\sharp(F_{t+1}, x) \geq \mathcal{B}_t(V_{F_{t+1}})$) when $\text{opt} = \text{sup}$ (resp. $\text{opt} = \text{inf}$).
- **Tight:** for every set of basic functions $F_{t+1} \subset \mathbf{F}_{t+1}$ and every $x \in \mathbb{X}$ the functions $\phi_t^\sharp(F_{t+1}, x)$ and $\mathcal{B}_t(V_{F_{t+1}})$ coincide at point x , that is $\phi_t^\sharp(F_{t+1}, x)(x) = \mathcal{B}_t(V_{F_{t+1}})(x)$.

For $t = T$, ϕ_T^\sharp is *valid* if, for every $F_T \subset \mathbf{F}_T$ and $x \in \mathbb{X}$, the function $\phi_T^\sharp(F_T, x)$ remains below (resp. above) the value function at time T when $\text{opt} = \text{sup}$ (resp. $\text{opt} = \text{inf}$). Moreover, the function ϕ_T^\sharp is *tight* if for every $F_T \subset \mathbf{F}_T$ and $x \in \mathbb{X}$ we have $\phi_T^\sharp(F_T, x)(x) = V_T(x)$.

C. Oracle

In Algorithm 1 we generate for every time t a sequence of random points where the selection functions will be evaluated, given the set F_t^k which characterizes the current approximation. In order to generate those points, we will assume that we have at our disposition an Oracle which, given $T+1$ sets of functions (characterizing the current approximations), computes $T+1$ compact sets and a probability law. The output of this Oracle need to satisfy the following assumption.

Assumption 1: The Oracle takes as input $T+1$ sets of functions. Its output includes $T+1$ compact sets K_0, \dots, K_T , each included in \mathbb{X} and a probability measure μ on \mathbb{X}^{T+1} such that for every $t \in \llbracket 0, T \rrbracket$, $K_t \subset \text{dom}(V_t)$ and there exists a function $g_t : \mathbb{R}_+^* \rightarrow (0, 1)$ such that for every $\eta > 0$ and every $x \in K_t$, $\mu(\mathbb{X}^t \times (\text{Ball}(x, \eta) \cap K_t) \times \mathbb{X}^{T-t}) \geq g_t(\eta)$. Note that this assumption implies that for every time $t \in \llbracket 0, T \rrbracket$, the support of the t -th marginal of μ contains K_t .

D. Tropical Dynamic Programming (TDP)

For every time $t \in \llbracket 0, T \rrbracket$, we construct a sequence of functions $(V_t^k)_{k \in \mathbb{N}}$ belonging to \mathbb{F}_t as follows. At each iteration $k \geq 0$, we build a subset F_t^k of the set \mathbf{F}_t and define the sequence of functions by pointwise optimum $V_t^k := V_{F_t^k} = \text{opt}_{\phi \in F_t^k} \phi$. As described here, the functions are just byproducts of Algorithm 1, which only describes the way the sets F_t^k are computed.

We show below in Theorem 1 that TDP generates approximations $(V_t^k)_k$ of V_t which converge to some limit function

Algorithm 1 Tropical Dynamic Programming (TDP)

Input: For every $t \in \llbracket 0, T \rrbracket$, ϕ_t^\sharp a compatible selection function, an Oracle satisfying Assumption 1, $T+1$ compact sets K_0^0, \dots, K_T^0 and a probability measure μ^0 on \mathbb{X}^{T+1} of support equal to the product of these $T+1$ compact sets.

Output: For every $t \in \llbracket 0, T \rrbracket$, a sequence of sets $(F_t^k)_{k \in \mathbb{N}}$. Define for every $t \in \llbracket 0, T \rrbracket$, $F_t^0 := \emptyset$.

for $k \geq 1$ **do**

Draw some points $(x_t^{k-1})_{t \in \llbracket 0, T \rrbracket}$ over $K_0^{k-1} \times K_1^{k-1} \times \dots \times K_T^{k-1}$ according to μ^{k-1} , independently from previous draws at iterations $k' < k$.

Compute $\phi_T^k := \phi_T^\sharp(F_T^{k-1}, x_T^{k-1})$.

Define $F_T^k := F_T^{k-1} \cup \{\phi_T^k\}$.

for t from $T-1$ to 0 **do**

Compute $\phi_t^k := \phi_t^\sharp(F_{t+1}^k, x_t^{k-1})$.

Define $F_t^k := F_t^{k-1} \cup \{\phi_t^k\}$.

end for

Compute $(K_0^k, \dots, K_T^k, \mu^k) = \text{Oracle}(F_0^k, \dots, F_T^k)$.

end for

V_t^* . One cannot hope for V_t^* to be equal to V_t everywhere, however we introduce the notion of optimal sets $(X_t)_{t \in \llbracket 0, T \rrbracket}$ with respect to a sequence of functions $(\phi_t)_{t \in \llbracket 0, T \rrbracket}$ as a condition on the sets $(X_t)_{t \in \llbracket 0, T \rrbracket}$ ensuring that they are rich enough to have $V_t^* = V_t$ on such sets.

Definition 2 (Optimal sets): Let $(\phi_t)_{t \in \llbracket 0, T \rrbracket}$ be $T+1$ functions on \mathbb{X} . A sequence of sets $(X_t)_{t \in \llbracket 0, T \rrbracket}$ is said to be (ϕ_t) -optimal or optimal with respect to $(\phi_t)_{t \in \llbracket 0, T \rrbracket}$, if for every $t \in \llbracket 0, T-1 \rrbracket$, we have $\mathcal{B}_t(\phi_{t+1} + \delta_{X_{t+1}}) + \delta_{X_t} = \mathcal{B}_t(\phi_{t+1}) + \delta_{X_t}$.

Denote by μ the countable product probability measure of the sequence of probability measures $(\mu^k)_{k \in \mathbb{N}}$ generated by Algorithm 1 through the Oracle.

Theorem 1 ([8]): Fix $t \in \llbracket 0, T \rrbracket$. Under mild technical assumptions given in Appendix I, the sequence of functions $(V_t^k)_{k \in \mathbb{N}}$ defined by $V_t^k := \mathcal{V}_{F_t^k}$, where the sets F_t^k are generated by Algorithm 1, μ -a.s. converges uniformly on every compact set included in the domain of V_t to a (random) function V_t^* . Moreover, define $K_t^* := \limsup_k K_t^k$, for every time $t \in \llbracket 0, T \rrbracket$. Assume that, μ -a.s. the sets $(K_t^*)_{t \in \llbracket 0, T \rrbracket}$ are (V_t) -optimal when $\text{opt} = \text{inf}$ (resp. (V_t^*) -optimal when $\text{opt} = \text{sup}$). Then, μ -a.s. the function V_t^* is equal to the value function V_t on K_t^* .

Proof: We only give a sketch of the proof, details can be found in [8]. First we show the existence of limit candidates. Under the structural assumptions given in Appendix I, for every time $t \in \llbracket 0, T \rrbracket$, the approximation functions $(V_t^k)_{k \geq 1}$ generated by Algorithm 1 are equicontinuous and bounded on every compact subset of the domain of V_t . Hence by Arzelà-Ascoli theorem and monotonicity, the sequence $(V_t^k)_{k \geq 1}$ converges uniformly to a continuous function V_t^* on every compact subset of the domain of V_t .

Fix a time $t \in \llbracket 0, T-1 \rrbracket$. Using again Arzelà-Ascoli theorem, one can show, under the same assumptions, that on every compact subset of the domain of V_t we have $\mathcal{B}_t(V_{t+1}^k) \xrightarrow[k \rightarrow \infty]{} \mathcal{B}_t(V_{t+1}^*)$, where the convergence is in the

sup-norm restricted to the given compact set.

We claim that our candidates $(V_t^*)_{t \in \llbracket 0, T \rrbracket}$ satisfy a system of modified Bellman equations. Indeed, as Algorithm 1 used compatible selection functions and an Oracle satisfying Assumption 1, we can show that the approximations (V_t^*) satisfy the system of equations:

$$V_T^* + \delta_{K_T^*} = \psi + \delta_{K_T^*} \text{ and } \forall t, \mathcal{B}_t(V_{t+1}^*) + \delta_{K_t^*} = V_t^* + \delta_{K_t^*}.$$

Under the assumption that the sets $(K_t^*)_{t \in \llbracket 0, T \rrbracket}$ are (V_t) -optimal when $\text{opt} = \inf$ (resp. (V_t^*) -optimal when $\text{opt} = \sup$), this allows us to conclude that for every $t \in \llbracket 0, T \rrbracket$ and every $x \in K_t^*$, we have $V_t^*(x) = V_t(x)$. ■

III. A TOY EXAMPLE: LINEAR-QUADRATIC MULTISTAGE OPTIMIZATION PROBLEM WITH ONE CONTROL CONSTRAINT

A. Constrained linear-quadratic framework

Let a, b be two given reals such that $a < b$, we study the following multistage linear quadratic problem involving a constraint on one of the controls:

$$\min_{\substack{x \in \mathbb{X}^{T+1} \\ (u, v) \in (\mathbb{U} \times \mathbb{V})^T}} \sum_{t=0}^{T-1} c_t(x_t, u_t, v_t) + \psi(x_T) \quad (5)$$

such that $x_0 \in X$ is given, $x_{t+1} = f_t(x_t, u_t, v_t)$ and $(u_t, v_t) \in \mathbb{U} \times [\beta, \gamma]$ where $\mathbb{X} = \mathbb{R}^n$ and $\mathbb{U} = \mathbb{R}^m$, with quadratic convex costs functions of the form $c_t(x, u, v) = x^T C_t x + u^T D_t u + v^2 d_t$, where $C_t \in \mathcal{S}_n^+$, $D_t \in \mathcal{S}_m^{++}$ and $d_t > 0$, linear dynamics $f_t(x, u, v) = A_t x + B_t u + v b_t$, where A_t (resp. B_t) is a $n \times n$ (resp. $n \times m$) matrix, $b_t \in \mathbb{X}$, and final cost function $\psi := x^T M x$ with $M \in \mathcal{S}_n^{++}$.

B. SDDP lower approximations of the value functions

The following result validates that this framework is well suited for lower approximations of the value functions by suprema of affine cuts. We refer to [8] for a proof in a more general framework.

Proposition 1 ([8]): For every $t \in \llbracket 0, T \rrbracket$, the value function V_t is L_t -Lipschitz continuous and convex. Moreover the Lipschitz constant $L_t > 0$ can be explicitly computed.

As (lower semicontinuous proper) convex functions can be approximated by a supremum of affine functions, for every $t \in \llbracket 0, T \rrbracket$ we define $\mathbf{F}_t^{\text{SDDP}}$ to be the set of affine functions $\phi : x \in \mathbb{X} \mapsto \langle a, x \rangle + b \in \mathbb{R}$, $a \in \mathbb{X}$, $b \in \mathbb{R}$ with $\|a\|_2 \leq L_t$. Moreover we will denote by $\mathbb{F}_t^{\text{SDDP}}$ the set of convex functions $\phi : \mathbb{X} \mapsto \overline{\mathbb{R}}$ which are L_t -Lipschitz continuous. The structural assumptions given in Appendix I can then be checked. We now construct a compatible selection function. Let $t \in \llbracket 0, T-1 \rrbracket$, for any $F \subset \mathbf{F}_t^{\text{SDDP}}$ and $x \in \mathbb{X}$, define the following quadratic program

$$\min_{\substack{x' \in \mathbb{X} \\ (u, v) \in \mathbb{U} \times [\beta, \gamma] \\ \lambda \in \mathbb{R}}} c_t(x', u, v) + \lambda \text{ s.t. } \begin{cases} x' = x \text{ and } \forall \phi \in F \\ \phi(f_t(x', u, v)) \leq \lambda. \end{cases} \quad (6)$$

If we denote by b its optimal value and by a a Lagrange multiplier of the constraint $x' - x = 0$ at the optimum, that is such that (x', u, v, λ, a) is a stationary point of the Lagrangian $c_t(x', u, v) + \lambda - \langle a, x' - x \rangle$, then we define $\phi_t^{\text{SDDP}}(F, x) :=$

$\langle a, \cdot - x \rangle + b$. At time $t = T$, for any $F \subset \mathbf{F}_T^{\text{SDDP}}$ and $x \in \mathbb{X}$, fix $a \in \partial V_T(x)$ and define $\phi_T^{\text{SDDP}}(F, x) := \langle a, \cdot - x \rangle + V_T(x)$.

As done in the literature of the Stochastic Dual Dynamic Programming algorithm (see for example [2] or [3]), one can study the case when the draws are made along the optimal trajectories of the current approximations. This will satisfy Assumption 1.

More precisely, fix $k \in \mathbb{N}$, we say that a sequence of states $(x_0^k, x_1^k, \dots, x_T^k)$ is an *optimal trajectory for the k -th approximations starting from x_0* if $x_0^k := x_0$ and for every $t \in \llbracket 0, T-1 \rrbracket$, $x_{t+1}^k := f_t(x_t^k, u_t^k, v_t^k)$ where (u_t^k, v_t^k) is a minimizer of $\mathcal{B}_t(V_{t+1}^k)(x_t^k) = \min_{\substack{u \in \mathbb{U} \\ v \in [\beta, \gamma]}} c_t(x, u, v) + V_{t+1}^k(f_t(x, u, v))$.

Proposition 2 ([8]): For every $k \in \mathbb{N}$, define the sequence of singletons, $K_0^k := \{x_0\}$ and for $t \in \llbracket 0, T-1 \rrbracket$, $K_{t+1}^k := \{x_{t+1}^k\}$, where $(x_0^k, x_1^k, \dots, x_T^k)$ is an optimal trajectory for the k -th approximations starting from x_0 . Lastly, for every $t \in \llbracket 0, T \rrbracket$, define $K_t^* := \limsup_k K_t^k$. Then, the sequence of sets $(K_t^*)_{t \in \llbracket 0, T \rrbracket}$ is optimal with respect to $(V_t^*)_{t \in \llbracket 0, T \rrbracket}$ in the sense of Definition 2.

C. Discretization

We will approximate (5) by an unconstrained switched multistage linear quadratic problem by discretizing the constrained control. Fix an integer $N \geq 2$, set $v_i = \beta + i \frac{\gamma - \beta}{N-1}$ for every $0 \leq i \leq N-1$ and set $\mathbb{V} := \{v_0, v_1, \dots, v_{N-1}\}$. We define the following unconstrained switched multistage linear quadratic problem:

$$\begin{aligned} \min_{\substack{x \in \mathbb{X}^T \\ (u, v) \in (\mathbb{U} \times \mathbb{V})^{T-1}}} \sum_{t=0}^{T-1} c_t^{v_t}(x_t, u_t) + \psi(x_T) \\ \text{s.t. } \begin{cases} x_0 \in \mathbb{X} \text{ is given, } \forall t \in \llbracket 0, T-1 \rrbracket, v_t \in \mathbb{V} \\ \forall t \in \llbracket 0, T-1 \rrbracket, x_{t+1} = f_t^{v_t}(x_t, u_t), \end{cases} \end{aligned} \quad (7)$$

where for every $v \in \mathbb{V}$, $f_t^v = f_t(\cdot, \cdot, v)$ and $c_t^v = c_t(\cdot, \cdot, v)$. As the set of controls of problem (5) contains the set of controls of problem (7), upper approximations of the value functions of (7) are also upper approximations of the value functions of (5). Thus we will construct upper approximations for (7).

D. Homogenization

We add a dimension to the state space in order to homogenize the costs and dynamics, when a sequence of switching controls is fixed. Define the following homogenized costs and dynamics for every $t \in \llbracket 0, T-1 \rrbracket$ by:

$$\begin{aligned} \tilde{f}_t^v(x, y, u) &= \begin{pmatrix} A_t & v b_t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} B_t \\ 0 \end{pmatrix} u, \\ \tilde{c}_t^v(x, y, u) &= \begin{pmatrix} x \\ y \end{pmatrix}^T \begin{pmatrix} C_t & 0 \\ 0 & v^2 d_t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + u^T D_t u, \end{aligned}$$

And as the final cost function is already homogeneous, $\tilde{\psi}(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}^T \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. From these homogenized functions we define a multistage optimization problem with

one more (compared to (7)) dimension on the state variable:

$$\begin{aligned} & \min_{\substack{(x,y) \in (\mathbb{X} \times \mathbb{R})^T \\ (u,v) \in (\mathbb{U} \times \mathbb{V})^{T-1}}} \sum_{t=0}^{T-1} \tilde{c}_t^{v_t}(x_t, y_t, u_t) + \tilde{\Psi}(x_T, y_T) \\ \text{s.t. } & (x_0, y_0) \in \mathbb{X} \times \mathbb{R} \text{ given, and } (x_{t+1}, y_{t+1}) = \tilde{f}_t^{v_t}(x_t, y_t, u_t). \end{aligned} \quad (8)$$

One can deduce the value functions $V_{t,N}$ of the multistage optimization problem (7) (with non-homogeneous costs and dynamics) from the value functions $\tilde{V}_{t,N}$ of (8) (with homogeneous costs and dynamics).

Proposition 3 ([8]): Fix $t \in \llbracket 0, T \rrbracket$. For every $x \in \mathbb{X}$, we have that $V_{t,N}(x) = \tilde{V}_{t,N}(x, 1)$.

The following property allows us to restrict the study of the value functions to the unit sphere, which is compact.

Proposition 4: For every time step $t \in \llbracket 0, T \rrbracket$, the value function $\tilde{V}_{t,N}$ solution of (8) is 2-homogeneous, *i.e.* for every $(x, y) \in \mathbb{X} \times \mathbb{R}$ and every $\lambda \in \mathbb{R}$, we have

$$\tilde{V}_{t,N}(\lambda x, \lambda y) = \lambda^2 \tilde{V}_{t,N}(x, y).$$

E. Qu min-plus upper approximations of the value functions

Let $v \in \mathbb{V}$ be a given switching control, we define the operator \mathcal{B}_t^v by, for every function $\phi : \mathbb{X} \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ and every point $(x, y) \in \mathbb{X} \times \mathbb{R}$, $\mathcal{B}_t^v(\phi)(x, y) = \inf_{u \in \mathbb{U}} \tilde{c}_t^v(x, y, u) + \phi(\tilde{f}_t^v(x, y, u))$. Then, for every time $t \in \llbracket 0, T-1 \rrbracket$, the dynamic programming operator \mathcal{B}_t associated to (8) satisfies $\mathcal{B}_t(\phi) := \inf_{v \in \mathbb{V}} \mathcal{B}_t^v(\phi)$. A key property of the operators \mathcal{B}_t^v and \mathcal{B}_t is that they are min-additive, meaning that for every functions $\phi_1, \phi_2 : \mathbb{X} \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$, $\mathcal{B}_t^v(\inf(\phi_1, \phi_2)) = \inf(\mathcal{B}_t^v(\phi_1), \mathcal{B}_t^v(\phi_2))$ and a similar equation for \mathcal{B}_t . Moreover, by Riccati formula ([7]), the image of a convex quadratic function by \mathcal{B}_t^v is also a convex quadratic function.

Proposition 5 (Similar to [5], [6], [9]): For every time $t \in \llbracket 0, T \rrbracket$, there exists a finite set F_t of convex quadratic functions such that $\tilde{V}_{t,N} = \inf_{\phi \in F_t} \phi$.

As done in [8], using Proposition 5 one can consider as basic functions, the quadratic functions bounded in the Loewner sense between 0 and $\alpha_t \mathbf{I}$, where $\alpha_t > 0$ are known real numbers such that, if ϕ is a quadratic form bounded between 0 and $\alpha_{t+1} \mathbf{I}$, then $\mathcal{B}_t^v(\phi)$ is bounded between 0 and $\alpha_t \mathbf{I}$. Moreover, using the aforementioned properties, one will be able to compute $\mathcal{B}_t^v(\mathcal{V}_F)$ for a given switching control v and $\mathcal{B}_t(\mathcal{V}_F)$, for any finite set F of convex quadratic functions. Therefore, given a time $t \in \llbracket 0, T-1 \rrbracket$, we define the selection function ϕ_t^{Qu} as follows. For any given $F \subset \mathbf{F}_{t+1}^{\text{Qu}}$ and $(x, y) \in \mathbb{X} \times \mathbb{R}$, $\phi_t^{\text{Qu}}(F, x, y) = \mathcal{B}_t^v(\phi)$ for some $(v, \phi) \in \arg \min_{(v, \phi) \in \mathbb{V} \times F} \mathcal{B}_t^v(\phi)(x, y)$. Lastly, at time $t = T$, for any $F \subset \mathbf{F}_T^{\text{Qu}}$ and $(x, y) \in \mathbb{X} \times \mathbb{R}$, we set $\phi_T^{\text{Qu}}(F, x, y) = \tilde{\Psi}(x, y) = \psi(x)$. As done in [8] one can check that the structural assumptions given in Appendix I are satisfied. By 2-homogeneity, it is enough to know the value functions of (8) on the sphere to know them on the whole state space. Thus the random draws of TDP for the basic functions \mathbf{F}_t^{Qu} , $1 \leq t \leq T$ and the selection functions ϕ_t^{Qu} will be made uniformly on the unit euclidean sphere.

F. Upper and lower approximations of the value functions

For a large number of discretization points N (defined in Section III-C), one would expect that the value functions $V_{t,N}$ of (7) approximate the value functions V_t of (5). Indeed, one can show that for every time step $t \in \llbracket 0, T \rrbracket$, the approximation error is bounded by $C_t T/N^2$ in \mathbb{X} , for some constant $C_t > 0$. Thus, for large N we have $V_{t,N} \approx V_t$ and by Proposition 3, for every $N \geq 2$, we have $\tilde{V}_{t,N}(\cdot, 1) = V_{t,N} \geq V_t$. In the following Proposition we approximate $\tilde{V}_{t,N}$ from above by a min-plus algorithm and V_t from below by SDDP and using the convergence result of Theorem 1, we obtain the following one.

Theorem 2 ([8]): For every time step $t \in \llbracket 0, T \rrbracket$, denote by $(\overline{V}_t^k)_{k \in \mathbb{N}}$ (resp. $(\underline{V}_t^k)_{k \in \mathbb{N}}$) the sequence of functions generated by TDP with the selection function ϕ_t^{Qu} (resp. ϕ_t^{SDDP}) and the draws made uniformly over the euclidean sphere of $\mathbb{X} \times \mathbb{R}$ (resp. made as described in Proposition 2).

Then the sequence $(\overline{V}_t^k)_{k \in \mathbb{N}}$ (resp. $(\underline{V}_t^k)_{k \in \mathbb{N}}$) is non-increasing (resp. non-decreasing), bounded from below (resp. above) by $\tilde{V}_{t,N}$ (resp. V_t) and converges uniformly to $\tilde{V}_{t,N}$ (resp. V_t) on any compact subset of $\mathbb{X} \times \mathbb{R}$ (resp. K_t^* defined in Proposition 2).

IV. NUMERICAL EXPERIMENTS

The following data was used as a specific case of (5). For every time $t \in \llbracket 0, T-1 \rrbracket$, $A_t = (1 - 0.1) \text{Id}$, $b_t = (1 \ \dots \ 1)^T$, $B_t = b_t b_t^T$, $C_t = 0.1 \text{Id}$, $D_t = 0.1 \text{Id}$ and $d_t = 0.1$. The time horizon is $T = 15$, the states are in $\mathbb{X} = \mathbb{R}^n$ with $n = 25$, the unconstrained continuous controls are in $\mathbb{U} = \mathbb{R}^m$ with $m = 3$, the constrained continuous control is in $[\beta, \gamma]$, with $[\beta, \gamma] = [1, 5]$ in the first example and $[\beta, \gamma] = [-3, 5]$ in the second one. Moreover, we start from the initial point $x_0 = 0.2(1, \dots, 1)^T$ when TDP is applied with the selection function ϕ_t^{SDDP} and the number of discretization points N is varying from 5 to 200, for TDP with the selection function ϕ_t^{Qu} . In Figure 2, we give graphs representing the values $\overline{V}_t^k(x_t^k)$ and $\overline{V}_t^k(x_t^k, 1)$ at each time step t where the sequence of states $(x_t^k)_{k \in \mathbb{N}}$ is the optimal trajectory for the current lower approximations $(\underline{V}_t^k)_{k \in \mathbb{N}}$. From Theorem 2, we know that for every $t \in \llbracket 0, T-1 \rrbracket$ the gap $\overline{V}_t^k(x_t^k, 1) - \underline{V}_t^k(x_t^k)$ should be close to 0 as k increases assuming that N is large enough to have $V_t \approx V_{t,N}$.

On those two examples, we exhibit two convergence behaviors. In the first example, the constrained control has to be greater than 1, thus avoiding 0 which would have been (or almost) the optimal control if there were no constraint. The optimal constrained control is the projection on $\mathbb{U} \times [\beta, \gamma]$ of the optimal unconstrained control, thus the switching control is most of the time equal to the lower bound $\beta = 1$. From this observation we deduce two properties. First, the upper approximation given by Qu algorithm is good, even for a small N , as the optimal switch is (most of the time) equal to β . Second, this implies that at iteration k , the set F_t^k is of small cardinality. Moreover, in this example the number of switches is $N = 5$ thus few computations of $\mathcal{B}_t^v(\phi)(x)$ need

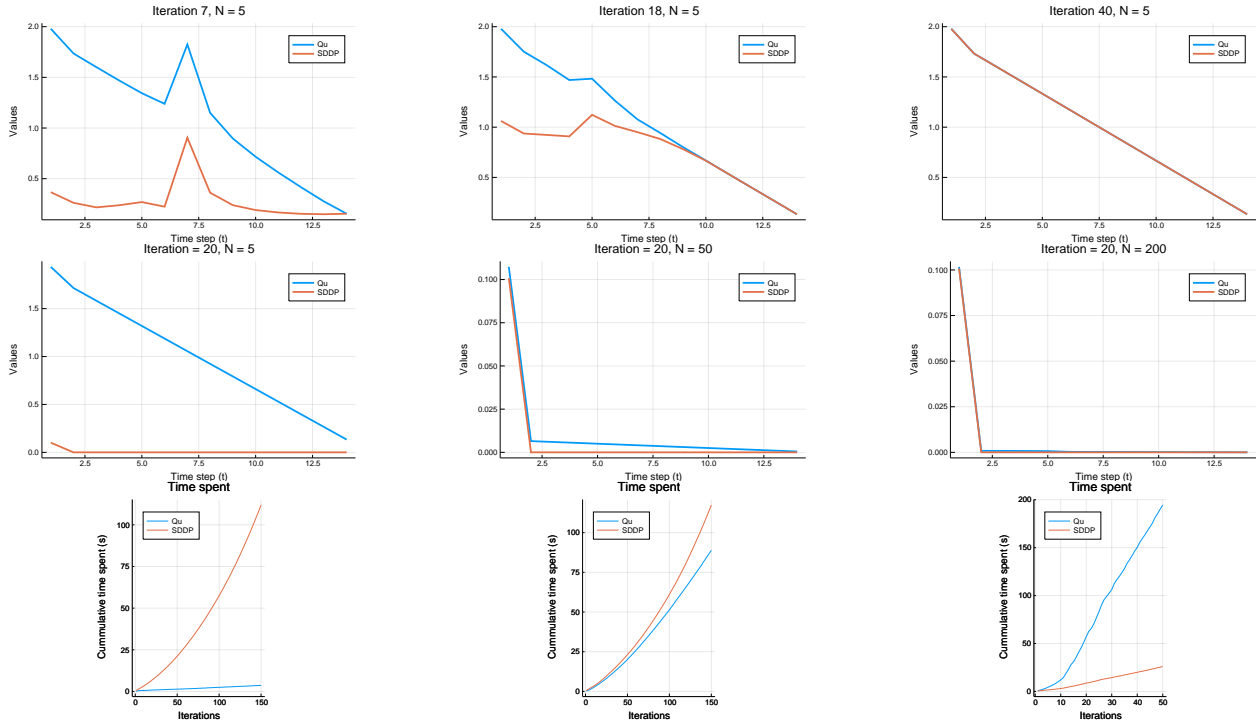


Fig. 2: Top: first example for $\beta = 1$, $\gamma = 5$ with $N = 5$ after 7 iterations (left), 18 iterations (middle) and 40 iterations (right). Middle: second example for $\beta = -3$, $\gamma = 5$ with varying $N = 5$ (left), $N = 50$ (middle) and $N = 200$ (right) after 20 iterations. Bottom: time spent in the first example (left) and the second example when $N = 50$ (middle) and $N = 200$ (right).

to be done in order to compute $\mathcal{B}_t(\phi)(x)$ for some $x \in \mathbb{X}$ and $\phi \in \mathbb{F}_t^k$. Thus, as shown on the bottom left of Figure 2, the computation time of an iteration of Qu's algorithm is small compared to SDDP which does not exploit this property. In the second example, the constrained control is in an interval containing 0. The switching control often changes and this means more computations. A compromise between computational time and precision can be achieved in order to make the computational time of the min-plus algorithm similar to the one of SDDP algorithm.

APPENDIX I STRUCTURAL ASSUMPTIONS

- 1) **Common regularity:** for every $t \in \llbracket 0, T \rrbracket$, there exists a common (local) modulus of continuity of all $\phi \in \mathbb{F}_t$.
- 2) **Final condition:** the value function V_T at time T is a pointwise optimum for some given subset F_T of \mathbb{F}_T .
- 3) **Stability by the Bellman operators:** for every $t \in \llbracket 0, T - 1 \rrbracket$, if $\phi \in \mathbb{F}_{t+1}$, then $\mathcal{B}_t(\phi)$ belongs to \mathbb{F}_t .
- 4) **Stability by pointwise optimum:** if $F_t \subset \mathbb{F}_t$ then $\mathcal{V}_{F_t} \in \mathbb{F}_t$.
- 5) **Stability by pointwise convergence:** if a sequence of functions $(\phi^k)_{k \in \mathbb{N}} \subset \mathbb{F}_t$ converges pointwise to ϕ on the domain of V_t , then $\phi \in \mathbb{F}_t$.
- 6) **Order preserving operators:** if $\phi, \varphi \in \mathbb{F}_{t+1}$ are such that $\phi \leq \varphi$, then $\mathcal{B}_t(\phi) \leq \mathcal{B}_t(\varphi)$.
- 7) **Existence of the value functions:** the solution $(V_t)_t$ to the Bellman system of equations (3) are proper functions.

- 8) **Existence of optimal sets:** for every compact $K_t \subset \text{dom}(V_t)$, every function $\phi \in \mathbb{F}_{t+1}$ and constant $\lambda \in \mathbb{R}$, there exists a compact set $K_{t+1} \subset \text{dom}(V_{t+1})$ such that $\mathcal{B}_t(\phi + \lambda + \delta_{K_{t+1}}) \leq \mathcal{B}_t(\phi + \lambda) + \delta_{K_t}$.
- 9) **Additively subhomogeneous operators:** given a compact set K_t , there exists $M_t > 0$ such that for every constant λ and every $\phi \in \mathbb{F}_{t+1}$, we have $\mathcal{B}_t(\phi + \lambda) + \delta_{K_t} \leq \mathcal{B}_t(\phi) + \lambda M_t + \delta_{K_t}$.

REFERENCES

- [1] R. Bellman, "The theory of dynamic programming," *Bulletin of the American Mathematical Society*, vol. 60, no. 6, pp. 503–515, Nov. 1954.
- [2] P. Girardeau, V. Leclere, and A. B. Philpott, "On the Convergence of Decomposition Methods for Multistage Stochastic Convex Programs," *Mathematics of Operations Research*, vol. 40, no. 1, pp. 130–145, Feb. 2015.
- [3] M. V. F. Pereira and L. M. V. G. Pinto, "Multi-stage stochastic optimization applied to energy planning," *Mathematical Programming*, vol. 52, no. 1-3, pp. 359–375, May 1991.
- [4] A. Shapiro, "Analysis of stochastic dual dynamic programming method," *European Journal of Operational Research*, vol. 209, no. 1, pp. 63–72, Feb. 2011.
- [5] Z. Qu, "A max-plus based randomized algorithm for solving a class of HJB PDEs," in *53rd IEEE Conference on Decision and Control*, Dec. 2014, pp. 1575–1580.
- [6] W. M. McEneaney, "A Curse-of-Dimensionality-Free Numerical Method for Solution of Certain HJB PDEs," *SIAM Journal on Control and Optimization*, vol. 46, no. 4, pp. 1239–1276, Jan. 2007.
- [7] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, ser. Oxford Science Publications. Oxford University Press, 1995.
- [8] M. Akian, J.-P. Chancelier, and B. Tran, "A stochastic algorithm for deterministic multistage optimization problems," *arXiv:1810.12870 [math]*, Oct. 2018.
- [9] D. Antunes and W. P. Maurice Heemels, "Linear Quadratic Regulation of Switched Systems Using Informed Policies," *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 2675–2688, Jun. 2017.