Policy with guaranteed risk-adjusted performance for multistage stochastic linear problems

Lucas Merabet[1†], Bernardo Freitas Paulo da Costa[2*†], Vincent Leclère[3†]

[1*]METRON, 6 rue Cambacérès, Paris, 75008, France.
[2]Applied Mathematics School, Getulio Vargas Foundation, Praia de Botafogo, 190, Rio de Janeiro, 22250–900, RJ, Brazil.
[3]CERMICS, Ecole des Ponts, Marne-la-Vallée, 77455, France.

*Corresponding author(s). E-mail(s): bernardo.paulo@fgv.br;
Contributing authors: lucas.merabet@gmail.com; vincent.leclere@enpc.fr;
[†]These authors contributed equally to this work.

Abstract

Risk-averse multi-stage problems and their applications are gaining interest in various fields of applications. Under convexity assumptions, the resolution of these problems can be done with trajectory following dynamic programming algorithms like Stochastic Dual Dynamic Programming (SDDP) to access a deterministic lower bound, and dual SDDP for deterministic upper bounds.
In this paper, we leverage the dual SDDP algorithm to compute a policy with guaranteed risk-adjusted performance for multistage stochastic linear problems.

Keywords: Duality, Inner approximation, SDDP, primal-dual methods

Contents

1 Introduction

Multi-stage stochastic optimization problems have a wide range of applications, from energy management to portfolio optimization. When the noise can be assumed stagewise-independent, or at least Markovian, these problems can be tackled by dynamic programming. However, when the state and control dimension increase, standard approaches like stochastic dynamic programming fall short due to the curses of dimensionality. To overcome this limitation, Stochastic Dual Dynamic Programming (SDDP) has been pioneered in the 90's to tackle hydro-thermal generation problems in Brazil (see Pereira and Pinto (1991)). It assumes convex costs and linear dynamics, and iteratively refines outer approximations of the Bellman functions (a.k.a cost-to-go function), by adding linear cuts. This approach yields a converging lower bound of the multi-stage problem.

The aforementioned problems mostly concern risk-neutral minimization. In recent years, risk-averse optimization has gained interest for use cases where extreme scenarios need to be specifically accounted for. Convex risk measures like the average value at risk (AV@R, see Rockafellar and Uryasev (2000))[1] have been used in the context of optimization, including multistage stochastic optimization. The SDDP algorithm can easily be extended to risk-averse settings (Philpott and De Matos, 2012; Shapiro et al., 2013; Philpott et al., 2013; Guigues, 2016), as finding cuts, and thus lower

[1] Also known as Conditional Value at Risk, Tail Value at Risk, Expected Shortfall or superquantile.

bounds for composed risk measures (see Ruszczyński (2010) for definition) is as easy as for risk-neutral problem.

However, computing upper bounds is more challenging. In deterministic optimization, an upper bound is often obtained by evaluating an admissible solution. In the risk-neutral setting, this upper bound can be estimated via Monte Carlo sampling of noise scenarios, but, in the risk-averse case, the composed risk-measures make this procedure not numerically tractable.

To tackle this issue, we introduced a dual version of SDDP in Leclère et al. (2020). The idea is to run SDDP on a dual formulation of the problem, getting lower approximations of the Fenchel transform of the Bellman value functions. Taking again the Fenchel transform provides an upper approximation of the Bellman value function. Thus, the algorithm computes at each iteration a deterministic upper bound of the problem. The idea was extended to the risk-averse setting in da Costa and Leclère (2023) and in a slightly different form in Guigues et al. (2023).

In this paper, we show that this upper bound on the optimal (risk-adjusted) value of the primal problem is also an upper bound on a computable policy. More precisely, using the outer approximation in the dual, we derive an inner approximation in the primal, and show that the cost of the policy induced by these inner approximations is no greater than the current upper bound. We thus extend the inner-approximation result of Leclère et al. (2020) to the risk-averse setting.

There are numerous tools to model risk-aversion. As common in the stochastic programming litterature, we focused on coherent risk measures Artzner et al. (1999), and more specifically on composed risk measures Ruszczyński (2010) to enable dynamic programming tools. In da Costa and Leclère (2023), we showed the duality results for composed risk measures with polyhedral step-measures. Finally, for the sake of simplicity, we consider a linear setting, but the results can be extended to convex problems, with adequate qualifications conditions.

Contributions

The main contributions of this paper are the following: i) we provide a new derivation of the dual SDDP algorithm for risk-averse linear problem based on duality of Bellman operator, ii) we show how to use the dual value functions to derive an inner approximation of the primal value functions, and iii) we show that the cost of the policy induced by these inner approximations is no greater than the current upper bound. Additionaly, we present a new numerical application, and study some numerical ideas to accelerate the convergence of the dual SDDP algorithm.

Outline

The remainder of the paper is organized as follows. In Section 2, we introduce the mathematical setting and assumptions. In Section 3, we recall and develop the duality results of da Costa and Leclère (2023) that are required to run a risk-averse dual SDDP algorithm, which we briefly present. These will all be combined in Section 4, where we show that the dual value functions can be used to derive an inner approximation of the primal value functions, and induce a policy with guaranteed risk-adjusted performance. Finally, Section 5 presents numerical results on two examples. To preserve the flow of the arguments, most proofs are deferred to appendix B, complemented by results on convex analysis in appendix A.

Notations

For a vector $\bar{x} \in \mathbb{R}^n$, we extend the interval notation $[0, \bar{x}]$ to indicate the the box $\{ 0 \leq x_i \leq \bar{x}_i \}$. For an integer n, we use $[n]$ to denote the set $\{1, \ldots, n\}$. This is used in particular in the case $x_{[t]}$, which denotes the vector $(x_1, \ldots, x_t)$ with the t first coordinates. Inequality between vectors is understood component-wise.

Random variables are denoted by boldface letters, and their realizations by the corresponding ligthface font. Equality and inequalities between random variables are understood almost surely (a.s.). $\boldsymbol{x}_t \preceq \boldsymbol{\xi}_{[t]}$ means that $\boldsymbol{x}_t$ is measurable with respect to the σ-algebra generated by $\boldsymbol{\xi}_{[t]}$. In a mathematical program, we sometimes specify between brackets the notation of the dual multiplier associated with a constraint.

2 Mathematical setting

We consider, for all $t \in [T]$ a finitely supported random variable $\boldsymbol{\xi}_t$, taking values in Ξ_t. Further ξ_0 is deterministic. These noises will be assumed to be stagewise independent throughout the paper.

2.1 Risk measures

A *coherent risk measure*, as defined by Artzner et al. (1999), is a function mapping the set of random variables to the extended real line, that is monotonous, convex, translation equivariant, and positively homogeneous. If, in addition, it is lower semi-continuous, then there exists a closed convex set $\mathcal{Q}$ of probability measures on $(\Xi_t, \mathcal{F}_t)$ such that

$$\rho(\boldsymbol{X}) = \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[\boldsymbol{X}]. \tag{1}$$

We say that a coherent risk measure ρ is *polyhedral* if the risk set $\mathcal{Q}$ is a polyhedron.

Example 1. The Average Value at Risk (AV@R), defined in Rockafellar and Uryasev (2000) as

$$\mathrm{AV@R}_\alpha(\boldsymbol{X}) = \inf_{z \in \mathbb{R}} \left\{ z + \frac{1}{\alpha} \mathbb{E}\left[(\boldsymbol{X} - z)^+ \right] \right\}, \tag{2}$$

is a coherent risk measure. Note that if Ξ_t is finite, then $\mathrm{AV@R}_\alpha(\boldsymbol{X})$ is polyhedral.

To mitigate the risk aversion of AV@R, we often consider a convex combination of AV@R and expectation, *i.e.,* $\rho = \beta\mathbb{E} + (1-\beta)\mathrm{AV@R}_\alpha$ with $\alpha, \beta \in (0, 1]$, which is also a coherent risk measure that is polyhedral if Ξ_t is finite.

2.2 Risk-averse multi-stage problem

We consider the risk-averse multi-stage problem

$$\inf_{\boldsymbol{x}_1, \boldsymbol{y}_1} \quad \rho_1\Big(\boldsymbol{c}_1^\top \boldsymbol{y}_1 + \inf_{\boldsymbol{x}_2, \boldsymbol{y}_2} \rho_2\big(\ldots + \inf_{\boldsymbol{x}_T, \boldsymbol{y}_T} \rho_T\big(\boldsymbol{c}_T^\top \boldsymbol{y}_T + g(\boldsymbol{x}_T) \big) \big) \Big) \tag{3a}$$

$$\text{s.t.} \quad \boldsymbol{A}_t \boldsymbol{x}_t + \boldsymbol{B}_t \boldsymbol{x}_{t-1} + \boldsymbol{T}_t \boldsymbol{y}_t = \boldsymbol{d}_t, \qquad \forall t \in [T], \tag{3b}$$

$$
\begin{aligned}
&0 \le \boldsymbol{x}_t \le \bar{x}_t && \forall t \in [T], && (3c)\\
&0 \le \boldsymbol{y}_t \le \bar{y}_t && \forall t \in [T], && (3d)\\
&\boldsymbol{x}_t, \boldsymbol{y}_t \preceq \boldsymbol{\xi}_{[t]} && \forall t \in [T], && (3e)
\end{aligned}
$$

where, x_0 is a given deterministic vector, for all $t \in [T]$, $(\boldsymbol{A}_t, \boldsymbol{B}_t, \boldsymbol{T}_t, \boldsymbol{d}_t, \boldsymbol{c}_t)$ are random variables measurable with respect to $\boldsymbol{\xi}_t$, and ρ_t is a polyhedral risk measure on Ξ_t as introduced in section 2.1.

In the following, we consider a terminal cost function g of the form $g(x) = K^\top (\tilde{x}_0 - x)^+$, where $K \in \mathbb{R}^n_+$, and $(.)^+$ acts component-wise on vectors.

Constraint (3e) is the non-anticipativity condition, stating that the decision at time t can only depend on the past realizations of the noises.

To solve this multistage problem, we define the *primal Bellman operator* associated with (3) as

$$
\begin{aligned}
\mathcal{B}_t(V)(x_t) = \inf_{\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}} \quad & \rho_t\Big(\boldsymbol{c}_{t+1}^\top \boldsymbol{y}_{t+1} + V(\boldsymbol{x}_{t+1})\Big) \\
\text{s.t.} \quad & \boldsymbol{A}_{t+1}\boldsymbol{x}_{t+1} + \boldsymbol{B}_{t+1}x_t + \boldsymbol{T}_{t+1}\boldsymbol{y}_{t+1} = \boldsymbol{d}_{t+1} \\
& 0 \le \boldsymbol{x}_{t+1} \le \bar{x}_{t+1} \\
& 0 \le \boldsymbol{y}_{t+1} \le \bar{y}_{t+1}
\end{aligned} \tag{4}
$$

Finally, for $t \in [T]$, we define the Bellman function V_t recursively as

$$
\begin{cases}
V_{T+1} = g \\
V_t = \mathcal{B}_t(V_{t+1}) \text{ for } t \in [\![0, T-1]\!].
\end{cases} \tag{5}
$$

In particular, $V_0(x_0)$ is the value of the problem (3).

2.3 Technical assumptions

In this paper we discuss version of the SDDP algorithm for the problem (3). This algorithm requires some assumptions on the problem, namely stagewise independence (for Dynamic Programming), relatively complete recourse and compactness of the state space.

In order to state those technical assumptions, we introduce the notion of *reachable set* of the problem (3), that is the set of states that can be reached, at least for one scenario, from the initial state x_0:

$$
\begin{aligned}
X_0 &= \{x_0\} && && (6a)\\
\mathcal{X}_{t,j}(x_{t-1}) &= \big\{\ 0 \le x_t \le \bar{x}_t \mid \exists y_t,\ 0 \le y_t \le \bar{y}_t, && && (6b)\\
&\qquad A_t^j x_t + B_t^j x_{t-1} + T_t^j y_t = d_t^j \big\} && \forall t \in [T], \forall j \in [J] && \\
X_t &= \bigcup_{j \in [J]} \bigcup_{x_{t-1} \in X_{t-1}} \mathcal{X}_{t,j}(x_{t-1}) && \forall t \in [T], && (6c)
\end{aligned}
$$

where, for $j \in [J]$, $A_t^j, B_t^j, T_t^j, d_t^j$ describe the possible realizations of the noise.

We make the following assumptions on the problem (3):

1. **Stagewise independence:** The noises $\boldsymbol{\xi}_t$ are stagewise independent;
2. **Compactness:** the reachable state set X_t is compact for all $t \in [T]$;
3. **Relatively complete recourse (RCR):** for all $x_{t-1} \in X_{t-1}$, and all $j \in [J]$, $\mathcal{X}_t(x_{t-1})$ is non-empty.

With these assumptions, we get two important results. First, that the SDDP algorithm converges to the optimal value of the problem (3), *i.e.*, $V_0(x_0)$ (see *e.g.*, Forcier and Leclere (2023)). Further, that the value functions are Lipschitz continuous, see proposition 3.1.

3 Duality theory for risk-averse multistage problems

Standard results and definitions of convex analysis are regrouped in appendix A. We start this section with more specific results that will be useful in the sequel, and then proceed to build the dual SDDP algorithm.

3.1 Convex analysis results

Using relatively complete recourse, polyhedrality of the Bellman operators and the Hoffman lemma, it can be shown (see, for instance (Shapiro et al., 2009, Proposition 2.7)) that we can propagate Lipschitz constants through the dynamic programming recursion:

Proposition 3.1. *If V is L_{t+1}-Lipschitz for the L^1 norm on $\mathcal{X}_{t+1}$, and $\mathcal{B}_t$ is a Linear Bellman Operator with relatively complete recourse, then there exists L_t, depending only on $\mathcal{B}_t$ and L_{t+1} such that $\mathcal{B}_t(V)$ is L_t-Lipschitz on $\mathcal{X}_t$.*

In particular, since g is Lipschitz, all the V_t defined in (5) are Lipschitz on X_t.

We will need to extend the value functions V_t to the whole space $\mathbb{R}^n$. The Lipschitz regularization of a function f (also known as the Pasch-Hausdorff envelope of f) is defined, for $L > 0$, as the inf-convolution[2] of f with the function $g(x) = L\|x\|_1$. The following proposition (see (Rockafellar and Wets, 2009, Example 9.11) for a proof) justifies the interpretation as a Lipschitz regularization:

Proposition 3.2. *If f is a proper function on $\mathbb{R}^n$ and $L > 0$, then $f \,\square\, (L\|\cdot\|_1)$ is the largest L-Lipschitz function on $\mathbb{R}^n$ that is at most equal to f (or is identically $-\infty$ if such a function does not exist). In particular, if f is L-Lipschitz on its domain, then it coincides with its regularization on its domain.*

Note that if L is too small then $f \,\square\, (L\|\cdot\|_1)$ can be identically $-\infty$. This does not happen if f is bounded from below.

The *coperspective* of a convex function f, introduced by da Costa and Leclère (2023) and denoted $f^{\boxtimes} := \widetilde{(f^*)}$, is the perspective of its Fenchel transform. For $\gamma > 0$, it is given by

$$f^{\boxtimes} : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R} : (\pi, \gamma) \mapsto \sup_x \pi^\top x - \gamma f(x). \tag{7}$$

[2]The inf-convolution and f and g is defined as $f \,\square\, g : x \mapsto \inf_y f(x) + g(x - y)$.

Note that it is jointly convex in (π, γ), lower semicontinuous, and positively homogeneous of degree 1. Moreover, the coperspective of a polyhedral function is polyhedral. Finally, for a proper, lsc, convex function f, we have $f = (f^{\boxtimes}(\cdot, 1))^*$.

The Lipschitz regularization is well-behaved for Fenchel duality, and also for the coperpective with an extra regularity assumption, as shown in the following lemma. Its proof can be found in appendix B.

Lemma 3.3. *If f is a proper convex lsc function and $L \geq 0$, then*

$$(f \,\square\, L \|\cdot\|_1)^* = f^* + (L\|\cdot\|_1)^* = f^* + \mathbb{I}_{LB_\infty}, \tag{8}$$

where $B_\infty = \{\pi \in \mathbb{R}^n \mid \|\pi\|_\infty \leq 1\}$. So, for $x \in \mathbb{R}^n$,

$$f \,\square\, (L\|\cdot\|_1)(x) = \sup_{\|\pi\|_\infty \leq L} \pi^\top x - f^*(\pi). \tag{9}$$

If f is also bounded below, then, for $(\pi, \gamma) \in \mathbb{R}^n \times \mathbb{R}_+$,

$$\left[f \,\square\, (L\|\cdot\|_1)\right]^{\boxtimes}(\pi, \gamma) = f^{\boxtimes}(\pi, \gamma) + \mathbb{I}_{\{\|\pi\|_\infty \leq \gamma L\}}. \tag{10}$$

We will use a variation of the Lipschitz regularization for coperspectives; in our setting, it will not be symmetric, but one-sided. We define it more generally for functions $F : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$, but the reader should think of the case where $F = f^{\boxtimes}$. For $x \in \mathbb{R}^n_+$, we define the *x-rectification of F* by

$$\mathcal{R}_x F : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R} : (\pi, \gamma) \mapsto \min_\psi F(\psi, \gamma) + x^\top (\pi - \psi)^+. \tag{11}$$

Therefore, the rectification is a partial inf-convolution with a multiple of the positive part function, evaluated coordinate-wise. [3] For future use, we will denote the function $\pi \mapsto x^\top(\pi)^+$ by $x^\top \mathrm{ReLU}$, by analogy with a layer of ReLU functions.

Rectification preserves convexity and ensures boundedness of the gradients with respect to π. Thus, it is analogous to the Lipschitz regularization, leaving unchanged functions whose π-gradients are bounded. We state this fact more formally in the following lemma, whose proof can be found in appendix B:

Lemma 3.4. *For $x \in \mathbb{R}^n_+$ and $F : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ proper, jointly convex and with domain $\mathbb{R}^n \times \mathbb{R}_+$, we have:*

1. *$\mathcal{R}_x F$ is jointly convex in π and γ, proper, and has domain $\mathbb{R}^n \times \mathbb{R}_+$;*
2. *$\partial_\pi(\mathcal{R}_x F)(\pi, \gamma) \subset [0, x]$ for all $(\pi, \gamma) \in \mathbb{R}^n \times \mathbb{R}_+$;*
3. *If, for all $(\pi, \gamma) \in \mathbb{R}^n \times \mathbb{R}_+$, $\partial_\pi F(\pi, \gamma) \subset [0, x]$, then $\mathcal{R}_x F = F$ on $\mathbb{R}^n \times \mathbb{R}_+$.*

3.2 Dual Bellman operators

As shown in Leclère et al. (2020); da Costa and Leclère (2023), the recursion $V_t = \mathcal{B}_t(V_{t+1})$ leads to a dual recursion between the coperspectives: $V_t^{\boxtimes} = \mathcal{B}_t^{\boxtimes}(V_{t+1}^{\boxtimes})$,

[3] The fact that $|x| = \max\{(x)^+, (-x)^+\}$ justifies "one-sided".

since $\mathcal{B}_t^{\boxtimes}(V^{\boxtimes}) = \left[\mathcal{B}_t(V)\right]^{\boxtimes}$ for proper, convex, lsc functions V. However, $\mathcal{B}_t^{\boxtimes}$ does not necessarily satisfy relatively complete recourse, and the outgoing state variables are not explicitly bounded which are common assumptions for the convergence of the SDDP algorithm. To remedy that, we consider instead the recursion between the modified value functions W_t, which are defined by

$$W_t := (L_t \left\|\cdot\right\|_1) \,\square\, (\mathbb{I}_{[0,\bar{x}_t]} + V_t). \tag{12}$$

Since the optimization problem for $\mathcal{B}_t$ only evaluates states in $[0, \bar{x}_{t+1}]$, and the value function V_{t+1} is L_{t+1}-Lipschitz on $[0, \bar{x}_{t+1}]$, the regularization W_{t+1} coincides with V_{t+1} there, so $\mathcal{B}_t(W_{t+1}) = \mathcal{B}_t(V_{t+1}) = V_t$. The recursion for W_t is therefore

$$W_t = (L_t \left\|\cdot\right\|_1) \,\square\, (\mathbb{I}_{[0,\bar{x}_t]} + \mathcal{B}_t(W_{t+1})).$$

Taking conjugates on (12), which interchanges sums and infimal convolutions (see (37) and (38)), and using that $\mathcal{B}_t(W_{t+1})$ is defined over $X_t \subset [0, \bar{x}_t]$ (by relatively complete recourse), we get

$$W_t^* = \mathbb{I}_{L_t B_\infty} + (\bar{x}_t^\top \mathrm{ReLU} \,\square\, V_t^*). \tag{13}$$

Therefore

$$W_t^{\boxtimes}(\pi, \gamma) = \mathbb{I}_{\|\pi\|_\infty \le \gamma L_t} + (\mathcal{R}_{\bar{x}_t} V_t^{\boxtimes})(\pi, \gamma). \tag{14}$$

This is immediate for $\gamma > 0$, using the definition of the perspective and the rectification (11). For $\gamma = 0$, notice that W_t^* has bounded and non-empty domain, so its recession function is $\mathbb{I}_0$. The same holds for $\mathbb{I}_{L_t B_\infty}$, and the recession function of $\bar{x}_t^\top \mathrm{ReLU} \,\square\, V_t^*$ is defined at zero (since the function is not identically $+\infty$).

Now, let $D_t := \mathcal{R}_{\bar{x}_t} V_t^{\boxtimes}$, and substitute the dual recursion $V_t^{\boxtimes} = \mathcal{B}_t^{\boxtimes}(V_{t+1}^{\boxtimes}) = \mathcal{B}_t^{\boxtimes}(W_{t+1}^{\boxtimes})$ to get

$$D_t = \mathcal{R}_{\bar{x}_t} \mathcal{B}_t^{\boxtimes}(\mathbb{I}_{\|\pi\|_\infty \le \gamma L_{t+1}} + D_{t+1}) \,, \tag{15}$$

which yields a modified dual recursion, which we write as $D_t = \mathcal{D}_t(D_{t+1})$. The rectification on the outside ensures relatively complete recourse, and the indicator function ensures boundedness of the outgoing state variables, which can be seen in the explicit form of the Bellman operator $\mathcal{D}_t$:

$$\begin{aligned} D_t(\pi_t, \gamma_t) = \inf_{\substack{\gamma_j, \zeta, \xi_j \ge 0 \\ \pi_j, \lambda_j}} \quad & \zeta_t^\top \bar{x}_t + \textstyle\sum_{j \in [J_t]} \lambda_j^\top d_{t+1}^j + \xi_j^\top \bar{y}_{t+1} + D_{t+1}(\pi_j, \gamma_j) \\ \text{s.t.} \quad & \zeta_t + \textstyle\sum_{j \in [J]} B_{t+1}^{j\top} \lambda_j \ge \pi_t \\ & (\gamma_j)_j \in \gamma_t \Gamma_t \\ & A_{t+1}^{j\top} \lambda_j + \pi_j \ge 0 && \forall j \in [J] \\ & \xi_j + T_{t+1}^{j\top} \lambda_j + \gamma_j c_{t+1}^j \ge 0 && \forall j \in [J] \\ & |\pi_j| \le \gamma_j L_{t+1} && \forall j \in [J] \end{aligned} \tag{16}$$

This holds up to $t = 1$. For the value of the first stage problem, we don't have the outer rectification, but on the other hand we have a fixed initial primal state $\tilde{x}_0$. Since we wish to bound $V_0(\tilde{x}_0) = \sup_{\pi_0} \pi_0^\top \tilde{x}_0 - V_0^\boxtimes(\pi_0, 1)$, we replace the first-stage dual problem by

$$-\left\{ \inf_{\pi_0} \mathcal{B}_0^\boxtimes(\mathbb{I}_{\|\pi\|_\infty \leq \gamma L_1} + D_1)(\pi_0, 1) - \pi_0^\top \tilde{x}_0 \right\} , \tag{17}$$

which is feasible since π_0 is not fixed.

The introduction of the functions W_t and D_t, and their relation to the original value functions V_t, will also be useful in section 4, where we show that D_t can be used to construct an inner policy with guaranteed risk-adjusted performance for the original problem.

3.3 Dual SDDP

The dual SDDP algorithm is obtained by applying SDDP to the dual Bellman operators $\mathcal{D}_t$ Leclère et al. (2020); da Costa and Leclère (2023). It is guaranteed to converge, since the operators $\mathcal{D}_t$, defined in (16), satisfy relatively complete recourse and compactness of the reachable set (as the outgoing state variables are bounded).

The recursion starts from $D_{T+1} = \mathcal{R}_{\bar{x}_{T+1}} g^\boxtimes$, which is polyhedral because g is, and can be represented explicitly in the last-stage optimization problem in the epigraphical form:

$$D_{T+1}(\pi, \gamma) = \inf_{\psi, \theta} \quad \bar{x}_{T+1}^\top (\pi - \psi)^+ + \theta \quad \text{s.t.} \quad \theta \geq g^\boxtimes(\psi, \gamma). \tag{18}$$

Next, we must construct the initial lower bounds for the dual value functions D_t. This is done by performing an initial backward pass calculating cuts for D_t at $(\pi = 0, \gamma = 1)$, which are valid trial points as shown in lemma B.3. Another possibility is to require the user to provide a valid upper bound for each V_t, $t = 1, \ldots, T$.

At each iteration, we make a forward pass to construct trial points (π_t, γ_t), and then a backward pass to construct cuts for D_t at these points, improving the approximations $\underline{D}_t^k$. Since the dual problems given by (16) include all noise realizations, the forward pass also provides Lagrange multipliers to build a cut for D_t. This is summarized in algorithm 1.

At each iteration of the algorithm we have $\underline{D}_t^k \leq D_t$, and in particular $\underline{D}_1^k \leq D_1$. Therefore, the optimal value of problem (17) is a *deterministic upper bound* for $V_0(\tilde{x}_0)$.

Implementation details

- The dual value functions D_t are homogeneous of degree 1. Therefore, any cut $C(\pi, \gamma) = f\gamma + x^T\pi + h$ can be assumed to have an offset $h = 0$. This ensures that the outer approximations are also 1-homogeneous.
- Contrary to primal SDDP, where cuts are typically good approximations of the value function locally, a cut approximating D_t at some point (π, γ) means we have a good approximation of D_t anywhere of the half line joining this point to the origin. As a result, during the forward pass, starting from time t, if children j is

Algorithm 1 Dual SDDP

Require: K the number of iterations to perform.

```
Build dual problems for t = T, ..., 0.
Initialize D_{T+1}(π, γ) in the last stage problem.
for t = T, ..., 1 do
    Compute a first cut for D_t by solving (16) at (π = 0, γ = 1).
end for
for k = 1, ..., K do
    Solve the first-stage problem (17).
    Update the current bound for V_0(x̃_0) from the optimal value.
    Select a scenario j and the corresponding outgoing state (π_1, γ_1) for the next stage.
    for t = 1, ..., T − 1 do                                        ▷ Forward pass
        Solve (16) at (π_t, γ_t) sampled from the forward pass.
        Add a cut for D_t from the optimal Lagrange multipliers.
        Select a scenario j and the corresponding outgoing state (π_{t+1}, γ_{t+1}) for the next stage.
    end for
    for t = T, ..., 1 do                                            ▷ Backward pass
        Add another cut for D_t by solving (16) at (π_t, γ_t) for the updated D̲^{k+1}_{t+1}.
    end for
end for
```

chosen leading to (π_j, γ_j), we normalize with respect to the γ component and set $(\pi_{t+1}, \gamma_{t+1}) = (\frac{\pi_j}{\gamma_j}, 1)$ for the next time step, when $\gamma \neq 0$.

- Since the dual problems are not separable by noise realization, forward and backward passes have equivalent computational cost, and both can be used to add cuts to D_t. We propose here to perform backward passes anyway, at least for some iterations. Indeed, when performing forward passes only, the information provided by a cut at time T will take T iterations to be propagated back to $t = 0$. Running a backward pass allows to back-propagate all the cut information at the end of one iteration, which improves convergence when T is large.

4 Policies with guaranteed risk bounds

In this section, we detail how to leverage the value functions computed in the dual SDDP algorithm to provide a feasible (primal) policy. We use the approximations $\underline{D}_t^k$ of the dual value functions to build an *inner approximation* of the primal value functions, which are *compatible* in a way that the risk-adjusted cost of the corresponding *inner strategy* is no greater than the upper bound computed in the dual SDDP algorithm.

4.1 Primal policies

For a sequence of finite polyhedral functions $\left\{\widehat{V}_t\right\}_{0\le t\le T}$, we define the state and control processes $\{\hat{\boldsymbol{x}}_t, \hat{\boldsymbol{y}}_t\}$ as

$$
\begin{aligned}
(\hat{\boldsymbol{x}}_{t+1}, \hat{\boldsymbol{y}}_{t+1}) = \underset{\boldsymbol{x}_{t+1},\boldsymbol{y}_{t+1}}{\operatorname{argmin}} \quad & \rho\Big(\boldsymbol{c}_{t+1}^\top \boldsymbol{y}_{t+1} + \widehat{V}_{t+1}(\boldsymbol{x}_{t+1})\Big) \\
\text{s.t.} \quad & \boldsymbol{A}_{t+1}\boldsymbol{x}_{t+1} + \boldsymbol{B}_{t+1}x_t + \boldsymbol{T}_{t+1}\boldsymbol{y}_{t+1} = \boldsymbol{d}_{t+1} \text{ a.s} \\
& 0 \le \boldsymbol{x}_{t+1} \le \bar{x}_{t+1} \qquad \text{a.s} \\
& 0 \le \boldsymbol{y}_{t+1} \le \bar{y}_{t+1} \qquad \text{a.s}
\end{aligned}
\tag{19}
$$

The risk of this strategy when starting from state $x \in \operatorname{supp}(\hat{\boldsymbol{x}}_t)$ at time t is denoted

$$
\hat{\rho}_t(x) = \rho\Big(\boldsymbol{c}_t^\top \hat{\boldsymbol{y}}_t + \rho\Big(\ldots + \rho\big(\boldsymbol{c}_T^\top \hat{\boldsymbol{y}}_{T-1} + g(\boldsymbol{x}_T)\big)\Big)\Big| \hat{\boldsymbol{x}}_t = x\Big)
\tag{20}
$$

Traditionally, we use $\widehat{V}_t = \underline{V}_t$, where $\underline{V}_t$ is an outer approximation of the exact value function V_t obtained via primal SDDP, to derive a primal strategy. In the risk-neutral setting, it is possible to use Monte-Carlo methods to estimate the average cost of this strategy Pereira and Pinto (1991). This can be used to stop the algorithm when the exact lower bound is close enough to the upper bound of the confidence interval, see (Shapiro, 2011, remark 4). In the risk-averse setting, however, this is more difficult, because of the nested risk measures.

We propose here to use an inner approximation derived from the approximations of the dual functions, for which the policy is guaranteed to have a risk no greater than the upper bound computed in the dual SDDP algorithm. This has been shown in Leclère et al. (2020) for risk-neutral problems.

4.2 Inner approximation from $\underline{D}_t$

We define here a primal strategy based on an inner approximation of the primal value functions. Following the discussion on section 3.2, we can use an inner approximation to W_t instead of one for V_t. Since $W_t^{\boxtimes}(\pi, \gamma) = \mathbb{I}_{\|\pi\|_\infty \le \gamma L_t} + D_t(\pi, \gamma)$, evaluating at $\gamma = 1$ and taking the dual (as in section 3.3) suggests the following

Definition 4.1. Let, for $t \in [T]$,

$$
\overline{V}_t^k := W_t^k = (L_t \|\cdot\|_1) \,\square\, \Big[\underline{D}_t^k(.,1)\Big]^*.
\tag{21}
$$

It is an inner approximation of V_t on X_t since $\underline{D}_t^k \le D_t$ and Fenchel conjugacy reverses the ordering.

The dual SDDP algorithm yields lower approximations of the form

$$
\begin{aligned}
\underline{D}_t^k(\pi, \gamma) = \min_z \quad & z \\
\text{s.t} \quad & z \ge f_t^l \gamma + {x_t^l}^\top \pi \qquad \text{for } l = 1, \ldots, k,
\end{aligned}
\tag{22}
$$

where $\{(f_t^l, x_t^l)\}_{1\le l\le k}$ are the cut coefficients calculated for D_t^k.

The following proposition is analogous to (Leclère et al., 2020, Proposition 26) for the risk-averse setting.

Proposition 4.2. *For $0 \le t \le T$,*

1. *$\overline{V}_t^k \ge V_t$ on X_t;*
2. *For $x \in \mathbb{R}^n$,*

$$\overline{V}_t^k(x) = \min_{y\in\mathbb{R}^n, \sigma_l \ge 0} \quad L_t \|x-y\|_1 - \sum_{l=1}^k \sigma_l f_t^l \quad \text{s.t.} \quad \sum_{l=1}^k \sigma_l = 1, \quad \sum_{l=1}^k \sigma_l x_t^l = y \tag{23}$$

Using formula (23) as a representation for $\widehat{V}$ in equation (19) gives a practical way of implementing a primal strategy from $\overline{V}_t^k$. Since $\overline{V}_t^k \ge V_t$, this justifies the name *inner approximation* for $\overline{V}_t^k$.

The lower approximations built by the primal SDDP algorithm satisfy by construction $\underline{V}_t^k \le \mathcal{B}_t(\underline{V}_{t+1}^k)$, since the cuts for that define $\underline{V}_t$ are lower bounds for the expected future value function. The same holds for the dual SDDP algorithm: $\underline{D}_t^k \le \mathcal{D}_t(\underline{D}_{t+1}^k)$. Since $\overline{V}_t^k$ is obtained by duality, we have a result, proved in appendix B.4, with the reversed ordering:

Proposition 4.3. *For all $t \in [\![0, T-1]\!]$ and all $k \ge 0$ we have*

$$\mathcal{B}_t\left(\overline{V}_{t+1}^k\right) \le \overline{V}_t^k \text{ on } X_t. \tag{24}$$

From this, we deduce that the dual upper bound is also an upper bound of this strategy.

Theorem 4.4 (Bounding theorem)**.** *Let $\{\hat{\boldsymbol{x}}_t^k, \hat{\boldsymbol{y}}_t^k\}$ be the state and control processes induced by the inner strategy defined by the sequence $\left\{\overline{V}_t^k\right\}_{0\le t\le T}$, and let $\hat{\rho}_t^k(x)$ be the risk of this strategy starting at time t from state x, as given by equation (20). For all $t \in [\![0, T]\!]$ and $x \in [0, \bar{x}_t]$, we have*

$$\hat{\rho}_t^k(x) \le \overline{V}_t(x). \tag{25}$$

In particular, the risk of the inner strategy is lower than the current upper bound U_k given by the dual SDDP algorithm.

Proof. We proceed by backward induction on time t. The property holds for $t = T$. Assume that $\hat{\rho}_{t+1}^k(\boldsymbol{x}_{t+1}^k) \le \overline{V}_t^k(\boldsymbol{x}_{t+1}^k)$. Let $x \in [0, \bar{x}_t]$. For $t > 0$,

$$\begin{aligned}\hat{\rho}_t^k(x) &= \rho\Big(\boldsymbol{c}_{t+1}^\top \boldsymbol{y}_{t+1}^k + \hat{\rho}_{t+1}^k(\boldsymbol{x}_{t+1}^k)\Big) \\ &\le \rho\Big(\boldsymbol{c}_{t+1}^\top \boldsymbol{y}_{t+1}^k + \overline{V}_{t+1}^k(\boldsymbol{x}_{t+1}^k)\Big) \qquad \text{(by induction and monotonicity of } \rho)\end{aligned}$$

$$= \mathcal{B}_t(\overline{V}^k_{t+1})(x) \qquad \text{(by definition of } \boldsymbol{y}^k_{t+1}\text{)}$$
$$\leq \overline{V}^k_t(x) \qquad \text{(by Proposition 4.3)}$$

For $t = 0$, $\left[\underline{D}^k_0(.,1)\right]^* \square \left(L_0\|.\|_1\right)(\tilde{x}_0) = \mathcal{F}^{L_0}_{\tilde{x}_0}(\underline{D}^k_0)$ per Lemma 3.3, which is equal to U_k. □

5 Numerical experiments

We detail here the numerical experiments we performed to illustrate the performance of the dual SDDP algorithm.

5.1 Interconnected micro-grids

The first model we consider is a chain of 4 interconnected micro-grids denoted MG_i for $i = 1, ..., 4$, optimized over a 24h horizon with $T = 96$ 15-minute time steps. The main grid will be denoted MG_0. Each microgrid MG_i has a battery installed, with state of charge $x_i \in [0, 70]$. During time interval $[t, t+1]$, it can charge an amount Φ^{i+}_{t+1} or discharge an amount Φ^{i-}_{t+1}, with efficiencies ρ^+ and ρ^-, and incurs a residual (and random) electrical demand $d^i_{t+1} \in \mathbb{R}$. The global state of the system is

$$x_t = (x^1_t, x^2_t, x^3_t, x^4_t)^\top \tag{26}$$

and the vector of battery controls is

$$\Phi_{t+1} = \left(\Phi^{1+}_{t+1}, \Phi^{1-}_{t+1}, \Phi^{2+}_{t+1}, \Phi^{2-}_{t+1}, \Phi^{3+}_{t+1}, \Phi^{3-}_{t+1}, \Phi^{4+}_{t+1}, \Phi^{4-}_{t+1}\right)^\top . \tag{27}$$

MG_i can send a non-negative amount of energy e^{ij} to MG_j in the following way:

- MG_1 can buy energy from MG_0 at a unit (deterministic, but time-varying) price p_t. It can also sell to MG_0 at a unit price $0.4p_t$;
- MG_2 can exchange energy with MG_1 and MG_3 with unit exchange cost c_e modeling average losses by Joule effect;
- MG_4 can exchange energy with MG_3 with unit exchange cost c_e.

We note

$$e_{t+1} = \left(e^{01}_{t+1}, e^{10}_{t+1}, e^{12}_{t+1}, e^{21}_{t+1}, e^{23}_{t+1}, e^{32}_{t+1}, e^{34}_{t+1}, e^{43}_{t+1}\right)^\top \tag{28}$$

the vector of possible energy exchanges and

$$T_e = \begin{pmatrix} 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \tag{29}$$

the corresponding exchange matrix.

For convenience, we denote bdiag(a,b) the block diagonal matrix

$$\text{bdiag}(a,b) = \begin{pmatrix} a & b & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & b & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a & b & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a & b \end{pmatrix}, \tag{30}$$

and define $T_\Phi = \text{bdiag}(-1,1)$, $T_\rho = \text{bdiag}(\rho^+, -\rho^-)$. We write

$$c_{t+1} = (0,0,0,0,0,0,0,0,p_t,-0.4p_t,c_e,c_e,c_e,c_e,c_e,c_e)^\top \tag{31}$$

the cost vector on time interval $[t,t+1]$ and

$$\boldsymbol{d}_{t+1} = \left(0,0,0,0,\boldsymbol{d}^1_{t+1},\boldsymbol{d}^2_{t+1},\boldsymbol{d}^3_{t+1},\boldsymbol{d}^4_{t+1}\right)^\top \tag{32}$$

the random demand vector with $|\text{supp}(\boldsymbol{d}_{t+1})| = 50$ realizations. Then the system dynamics can be put in the form (3) with

$$A = I_{8,4}, \quad B = -I_{8,4}, \quad T = \begin{pmatrix} T_\rho & 0_{4,8} \\ T_\Phi & T_e \end{pmatrix} \quad \text{and } y_{t+1} = \begin{pmatrix} \Phi_{t+1} \\ e_{t+1} \end{pmatrix} \tag{33}$$

Note that the matrices are deterministic and static.

5.2 Brazilian Hydro-thermal model

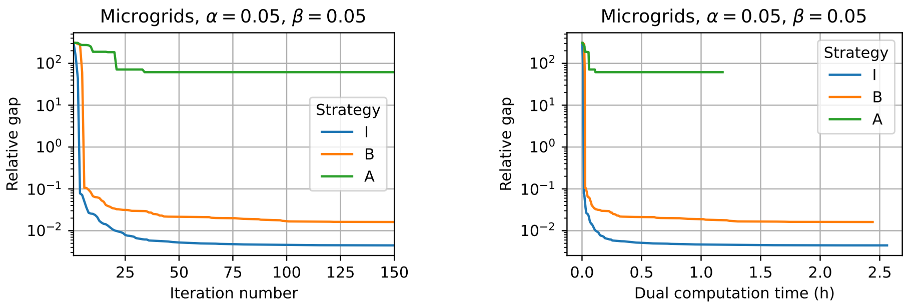
The second experiment we consider is a hydro-thermal problem over a one year horizon with $T = 12$ monthly stages and 82 noise realizations per time step. It is described in detail in da Costa and Leclère (2023).

5.3 Numerical results

We used three different strategies for the forward sampling:

- **Importance sampling (I)**: the child j for the next state is chosen at random from the distribution induced by the weights γ_j computed during the solution of the forward pass;
- **Base probability (B)**: the child j for the next state is drawn from the base probability $\mathbb{P}$;
- **Alternating (A)**: at each iteration, we perform a forward and backward pass of the primal SDDP algorithm, and use the subgradient π from the cut at stage t as the state (with $\gamma = 1$) used to calculate cuts for the dual SDDP;

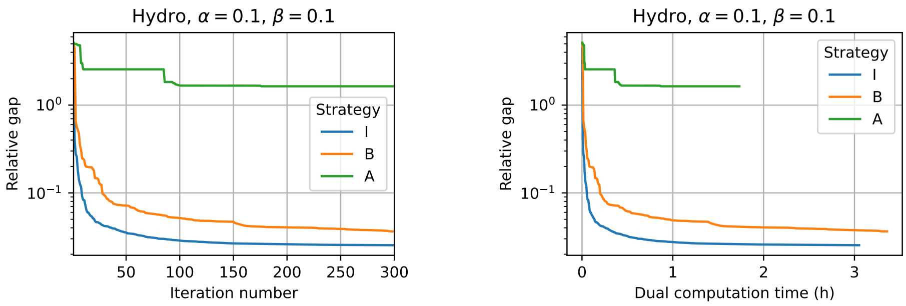
We present the evolution of the gap between the upper bounds of the dual SDDP and the lower bounds of the primal SDDP along the iterations and along the time in Figures 1a and 1b for the multigrid problem, and in Figures 2a and 2b for the hydro-thermal problem. These experiments were run on an Intel(R) Xeon(R) Platinum 8370C CPU @ 2.80GHz, using the CPLEX solver for the LPs.



(a) Gap along iterations

(b) Gap along time

Fig. 1: Relative gap between dual SDDP upper bounds and primal SDDP lower bounds on the microgrid problem, for different forward sampling strategies: (I): Importance sampling; (B): Base probability; (A): Alternating.



(a) Gap along iterations

(b) Gap along time

Fig. 2: Relative gap between dual SDDP upper bounds and primal SDDP lower bounds on the hydro-thermal problem, for different forward sampling strategies: (I): Importance sampling; (B): Base probability; (A): Alternating.

For both problems, the importance sampling procedure is the one with smallest gap, with similar time compared to the base probability sampling. This is especially noticeable in very risk-averse parameter settings, and also in the multigrid problem. Also, we observe that the alternating strategy is not competitive with the other two.

References

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

Rockafellar, R.T., Uryasev, S.: Optimization of conditional value-at-risk. Journal of Risk **2**, 21–42 (2000)

Philpott, A.B., De Matos, V.L.: Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. European Journal of operational research **218**(2), 470–483 (2012)

Shapiro, A., Tekaya, W., Costa, J.P., Soares, M.P.: Risk neutral and risk averse stochastic dual dynamic programming method. European Journal of Operational Research **224**(2), 375–391 (2013)

Philpott, A., Matos, V., Finardi, E.: On solving multistage stochastic programs with coherent risk measures. Operations Research **61**(4), 957–970 (2013)

Guigues, V.: Convergence analysis of sampling-based decomposition methods for risk-averse multistage stochastic convex programs. SIAM Journal on Optimization **26**(4), 2468–2494 (2016)

Ruszczyński, A.: Risk-averse dynamic programming for markov decision processes. Mathematical programming **125**, 235–261 (2010)

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

Costa, B.F.P., Leclère, V.: Dual SDDP for risk-averse multistage stochastic programs. Operations Research Letters **51**(3), 332–337 (2023)

Guigues, V., Shapiro, A., Cheng, Y.: Duality and sensitivity analysis of multistage linear stochastic programs. European Journal of Operational Research **308**(2), 752–767 (2023)

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Mathematical finance **9**(3), 203–228 (1999)

Forcier, M., Leclere, V.: Convergence of trajectory following dynamic programming algorithms for multistage stochastic problems without finite support assumptions. Journal of Convex Analysis, (2023)

Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on Stochastic Programming: Modeling and Theory. MPS-SIAM Series on Optimization. SIAM, Philadelphia (2009)

Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Grundlehren der mathematischen Wissenschaften, vol. 317. Springer, Berlin (2009)

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

Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)

Combettes, P.L.: Perspective functions: Properties, constructions, and examples. Set-Valued and Variational Analysis **26**(2), 247–264 (2018)

A Convex analysis

We recall here some standard definitions and results from convex analysis, which can be found in Rockafellar (1970); Rockafellar and Wets (2009); Combettes (2018).

A.1 Fenchel conjugate

Let $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$ be a function. Recall that its domain is $\mathrm{dom}(\varphi) = \{x \in \mathbb{R}^n \mid \varphi(x) < +\infty\}$, and its epigraph is $\mathrm{epi}(\varphi) = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid \varphi(x) \leq \alpha\}$. φ is *convex* if its epigraph is a convex set, *lower semicontinuous (lsc)* if its epigraph is closed, and *proper* if its domain is non-empty (*i.e.,* it takes at least one finite value) and never takes the value $-\infty$.

For a proper function $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$, its *Fenchel transform* is the lsc convex function

$$\varphi^* \; : \; \mathbb{R}^n \to \bar{\mathbb{R}} \; : \; \pi \; \mapsto \; \sup_{x \in \mathbb{R}^n} \pi^\top x - \varphi(x). \tag{34}$$

Importantly, the Fenchel transform is *order-reversing*: for proper functions φ and ψ,

$$\varphi \leq \psi \quad \implies \quad \psi^* \leq \varphi^*. \tag{35}$$

Finally, if φ is proper, convex and lower semicontinuous, then $\varphi^{**} = \varphi$.

A.2 Infimal convolution

Let f and g be two proper functions on $\mathbb{R}^n$. Their *infimal convolution* is

$$f \,\square\, g : x \mapsto \inf_y f(y) + g(x - y). \tag{36}$$

The infimal convolution is commutative and associative. However, it does not necessarily preserve properness.

The infimal convolution can be used to regularize a function f. The most common is the Moreau-Yosida regularization, which is the infimal convolution with $\frac{1}{\mu}\|\cdot\|_2^2$ for some $\mu > 0$. Another example is the L-Lipschitz regularization of f for the L^1 norm, also know as Pasch-Hausdorff envelope of f, which is the infimal convolution with $g = L\left\|\cdot\right\|_1$. More precisely, let f be a proper convex function and $L > 0$, then $f \square L\left\|\cdot\right\|_1$ is either $\equiv -\infty$ or the largest L-Lipschitz function that minorizes f (Rockafellar and Wets, 2009, Example 9.11).

Finally, the Fenchel transform of the infimal convolution of proper convex functions $\varphi_1, \ldots, \varphi_m$ is the sum of the Fenchel transforms (Rockafellar, 1970, Theorem 16.4):

$$(\varphi_1 \,\square \cdots \square\, \varphi_m)^* = \varphi_1^* + \cdots + \varphi_m^*. \tag{37}$$

To obtain the reverse equality, we need an additional assumption over the domain of φ_i. More precisely, if $(\varphi_i)_{i \in [n]}$ are proper, convex functions, such that, for $i \leq k$, φ_k

is polyhedral, we have that (see (Rockafellar, 1970, Theorem 20.1))

$$\bigcap_{i=1}^{k} \operatorname{dom}(\varphi_i) \cap \bigcap_{i=k+1}^{n} \operatorname{ri}\operatorname{dom}(\varphi_i) \neq \emptyset \quad \Longrightarrow \quad (\varphi_1 + \cdots + \varphi_k)^* = \varphi_1^* \,\square \cdots \square\, \varphi_k^*. \tag{38}$$

A.3 Perspective function

For a convex function φ, its *perspective* is the lsc convex function $\widetilde{\varphi}$ defined by

$$\widetilde{\varphi} : \mathbb{R}^n \times \mathbb{R} \to]-\infty, +\infty] : (\gamma, x) \mapsto \begin{cases} \gamma\varphi(x/\gamma), & \text{if } \gamma > 0 \\ (\operatorname{rec}\varphi)(x), & \text{if } \gamma = 0 \\ +\infty, & \text{otherwise} \end{cases} \tag{39}$$

where $\operatorname{rec}\varphi$ is the recession function of φ, defined by, for any $z \in \operatorname{dom}(\varphi)$,

$$\operatorname{rec}\varphi(x) = \lim_{\gamma \to +\infty} \frac{\varphi(z + \gamma x)}{\gamma}. \tag{40}$$

B Proofs of technical results

We recall the statements before the proofs, for ease of reference.

B.1 Proof of Lemma 3.3

Lemma B.1. *If f is a proper convex lsc function and $L \geq 0$, then*

$$(f \,\square\, L\,\|\cdot\|_1)^* = f^* + (L\,\|\cdot\|_1)^* = f^* + \mathbb{I}_{LB_\infty}, \tag{41}$$

where $B_\infty = \{\pi \in \mathbb{R}^n \mid \|\pi\|_\infty \leq 1\}$. So, for $x \in \mathbb{R}^n$,

$$f \,\square\, (L\,\|\cdot\|_1)(x) = \sup_{\|\pi\|_\infty \leq L} \pi^\top x - f^*(\pi). \tag{42}$$

If f is also bounded below, then, for $(\pi, \gamma) \in \mathbb{R}^n \times \mathbb{R}_+$,

$$\left[f \,\square\, (L\,\|\cdot\|_1)\right]^{\boxtimes}(\pi, \gamma) = f^{\boxtimes}(\pi, \gamma) + \mathbb{I}_{\{\|\pi\|_\infty \leq \gamma L\}}. \tag{43}$$

Proof. The first equation follows by eq. (37). As $f \,\square\, L\,\|\cdot\|_1$ is proper convex and lsc, taking the conjugate of eq. (41) yields the second result.

Taking the perspective of both sides of eq. (41) yields eq. (43) for $\gamma > 0$. If $\gamma = 0$, both sides yield $\mathbb{I}_0(\pi)$. Indeed, for the right hand side, $\operatorname{rec} f^* \not\equiv -\infty$ since f is proper, convex and lsc, and as $\operatorname{dom}(f^*) \neq \emptyset$, we have $\operatorname{rec} f^*(0) = \lim_{\gamma \to +\infty} \frac{f^*(\pi_0)}{\gamma} = 0$. For the left hand side of eq. (43), since f is bounded below, we have that $f \,\square\, L\,\|\cdot\|_1$ is also bounded below, so its conjugate is not identically $+\infty$, and therefore $\operatorname{rec}\,[f \,\square\, L\,\|\cdot\|_1]^* \not\equiv +\infty$.

□

B.2 Proof of Lemma 3.4

Lemma B.2. *For $x \in \mathbb{R}^n_+$ and $F : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ proper, jointly convex and with domain $\mathbb{R}^n \times \mathbb{R}_+$, we have:*

1. *$\mathcal{R}_x F$ is jointly convex in π and γ, proper, and has domain $\mathbb{R}^n \times \mathbb{R}_+$;*
2. *$\partial_\pi(\mathcal{R}_x F)(\pi, \gamma) \subset [0, x]$ for all $(\pi, \gamma) \in \mathbb{R}^n \times \mathbb{R}_+$;*
3. *If, for all $(\pi, \gamma) \in \mathbb{R}^n \times \mathbb{R}_+$, $\partial_\pi F(\pi, \gamma) \subset [0, x]$, then $\mathcal{R}_x F = F$ on $\mathbb{R}^n \times \mathbb{R}_+$.*

Proof. 1. For $(\pi, \gamma) \in \mathbb{R}^n \times \mathbb{R}_+$, we have

$$\mathcal{R}_x F(\pi, \gamma) = \inf_{\substack{\zeta + \psi \geq \pi \\ \zeta \geq 0}} x^\top \zeta + F(\psi, \gamma) = \inf_\psi F(\psi, \gamma) + \inf_{\substack{\zeta \geq \pi - \psi \\ \zeta \geq 0}} x^\top \zeta, \tag{44}$$

and the second infimum is equal to $x^\top(\pi - \psi)^+$ since $x \geq 0$.

2. By definition, $\mathcal{R}_x F(\pi, \gamma) = F(\cdot, \gamma) \,\square\, (x^\top \mathrm{ReLU})$. Since the conjugate of $x^\top\mathrm{ReLU}$ is $\mathbb{I}_{[0,x]}$, and using that F is proper and convex, equation (37) shows that the Fenchel conjugate of $\mathcal{R}_x F(\cdot, \gamma)$ is finite only on (a subset of) $[0, x]$.
3. Assume that all the π-subgradients of F are in $[0, x]$. Let $(\pi, \gamma) \in \mathbb{R}^n \times \mathbb{R}_+$. By convexity of F,

$$F(\psi, \gamma) \geq F(\pi, \gamma) + g^\top(\psi - \pi) \tag{45}$$

where $g \in \partial_\pi F(\pi, \gamma)$. Thus

$$F(\psi, \gamma) + x^\top(\pi - \psi)^+ \geq F(\pi, \gamma) + g^\top(\pi - \psi)^- + (x - g)^\top(\pi - \psi)^+ \tag{46}$$

where $(.)^- : y \mapsto \max(0, -y)$. By assumption on g, both products in (46) are non-negative. We deduce that

$$F(\psi, \gamma) + x^\top(\pi - \psi)^+ \geq F(\pi, \gamma) \tag{47}$$

for all ψ, with equality when $\psi = \pi$.

□

B.3 Proof of feasibility of trial points

Lemma B.3. *The point $(\pi = 0, \gamma = 1)$ is feasible for the dual recursion (15).*

Proof. Substituting $\gamma = 1$ yields

$$\mathcal{D}_t(\underline{D}^0_{t+1})(\pi, 1) = \left[\bar{x}_t^\top \mathrm{ReLU} \,\square\, \left(\mathcal{B}_t^{\boxtimes}(\mathbb{I}_{\|\pi\| \leq \gamma L_{t+1}} + \underline{D}^0_{t+1})\right)(\cdot, 1)\right](\pi) =: \phi_t(\pi),$$

which defines a proper convex polyhedral function ϕ_t. It suffices to show that $\phi_t(0)$ is finite. By Fenchel duality:

$$\begin{aligned}\phi_t(0) = \phi_t^{**}(0) &= \sup_x 0^\top x - \phi_t^*(x) \\ &= -\inf_x\, (\bar{x}_t^\top \mathrm{ReLU})^*(x) + \left[\mathcal{B}_t^{\boxtimes}(\mathbb{I}_{\|\pi\| \leq \gamma L_{t+1}} + \underline{D}^0_{t+1})(\cdot, 1)\right]^*(x)\end{aligned}$$

$$= -\inf_x \ \mathbb{I}_{[0,\bar{x}_t]} + \mathcal{B}_t(\big[(\mathbb{I}_{\|\pi\|\leq \gamma L_{t+1}} + \underline{D}^0_{t+1})(\cdot,1)\big]^*)(x)$$
$$= -\inf_x \ \mathbb{I}_{[0,\bar{x}_t]} + \mathcal{B}_t(L_{t+1}\left\|\cdot\right\|_1 \square (\underline{D}^0_{t+1}(\cdot,1))^*)(x).$$

The equalities are, respectively: the definition of the convex conjugate; the conjugate of the infimal convolution; the evaluation of the conjugates, using $\mathcal{B}_t^{\boxtimes}(V^{\boxtimes}) = [\mathcal{B}_t(V)]^{\boxtimes}$; the conjugate of the sum. For convenience, we define

$$W_t^k = L_t \left\|\cdot\right\|_1 \square (\underline{D}_t^k(\cdot,1))^*. \tag{48}$$

In the first iteration, $\underline{D}^0_{T+1} = \mathcal{R}_{\bar{x}_{T+1}} g^{\boxtimes} \not\equiv +\infty$, and for all other $1 \leq t \leq T$ we have that $\underline{D}^0_{t+1}$ is a linear function. So, in both cases the argument for $\mathcal{B}_t$ will be L_{t+1}-Lipschitz, and by the RCR hypothesis $\mathcal{B}_t(W^0_{t+1})$ will be L_t-Lipschitz over $X_t \subset [0, \bar{x}_t]$. Therefore, the minimum is finite, which shows that we can construct an L_t-Lipschitz cut for D_t. □

B.4 Proof of proposition 4.3

Proposition B.4. *For all $t \in [\![0, T-1]\!]$ and all $k \geq 0$ we have*

$$\mathcal{B}_t\left(\overline{V}^k_{t+1}\right) \leq \overline{V}^k_t \ \text{on } \mathcal{X}_t. \tag{49}$$

Proof. By definition, the coperspective of $\mathcal{B}_t(\overline{V}^k_{t+1})$ is

$$\begin{aligned}
\left[\mathcal{B}_t\left(L_{t+1}\left\|\cdot\right\|_1 \square \left[\underline{D}^k_{t+1}(.,1)\right]^*\right)\right]^{\boxtimes} &= \mathcal{B}_t^{\boxtimes}\left(\mathbb{I}_{\|\pi\|_\infty\leq\gamma L_{t+1}} + \left[\left[\underline{D}^k_{t+1}(.,1)\right]^*\right]^{\boxtimes}\right) \\
&= \mathcal{B}_t^{\boxtimes}\left(\mathbb{I}_{\|\pi\|_\infty\leq\gamma L_{t+1}} + \widetilde{\left[\underline{D}^k_{t+1}(.,1)\right]^{**}}\right) \\
&= \mathcal{B}_t^{\boxtimes}\left(\mathbb{I}_{\|\pi\|_\infty\leq\gamma L_{t+1}} + \underline{D}^k_{t+1}\right)
\end{aligned} \tag{50}$$

by proper polyhedrality and 1-homogeneity of $\underline{D}^k_{t+1}$.

Since $\mathcal{D}_t$ performs a regularization, we have that

$$\begin{aligned}
\mathcal{D}_t\left(\underline{D}^k_{t+1}\right) &= \mathcal{R}_{\bar{x}_t}\mathcal{B}_t^{\boxtimes}\left(\mathbb{I}_{\|\pi\|_\infty\leq\gamma L_{t+1}} + \underline{D}^k_{t+1}\right) \\
&\leq \mathcal{B}_t^{\boxtimes}\left(\mathbb{I}_{\|\pi\|_\infty\leq\gamma L_{t+1}} + \underline{D}^k_{t+1}\right) = \left(\mathcal{B}_t(\overline{V}^k_{t+1})\right)^{\boxtimes}.
\end{aligned}$$

By the construction of cuts for D_t, we get inductively

$$\underline{D}^k_t \leq \mathcal{D}_t\left(\underline{D}^k_{t+1}\right) \leq \left[\mathcal{B}_t\left(\overline{V}^k_{t+1}\right)\right]^{\boxtimes}. \tag{51}$$

Evaluating it at $(\pi, 1)$ for $\pi \in \mathbb{R}^n$, we have

$$\left[\mathcal{B}_t\left(\overline{V}_{t+1}^k\right)\right]^*(\pi) \geq \underline{D}_t^k(\pi, 1). \tag{52}$$

By applying the Fenchel transform to (52), and exploiting the proper polyhedrality of $\mathcal{B}_t\left(\overline{V}_{t+1}^k\right)$, we obtain

$$\mathcal{B}_t\left(\overline{V}_{t+1}^k\right) \leq \left[\underline{D}_t^k(., 1)\right]^*. \tag{53}$$

By Proposition 3.2, $\overline{V}_{t+1}^k$ is L_{t+1}-Lipschitz on $\mathbb{R}^n$. Hence, by Proposition 3.1, $\mathcal{B}_t\left(\overline{V}_{t+1}^k\right)$ is L_t-Lipschitz on $\mathcal{X}_t$. Regularizing both sides and using Lemma 3.4, we get that $\mathcal{B}_t\left(\overline{V}_{t+1}^k\right) \leq \left[\underline{D}_t^k(., 1)\right]^* \square \left(L_t\|.\|_1\right)$ on $\mathcal{X}_t$. $\square$