# Stochastic Programming Approach to Optimization Under Uncertainty

**A. Shapiro**
School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA

Theory of Reinforcement Learning Boot Camp
September 2020

Risk neutral $T$-stage stochastic programming problem:

$$\begin{array}{ll} \min\limits_{x_1, x_2(\cdot), \ldots, x_T(\cdot)} & f_1(x_1) + \mathbb{E}\left[\sum_{t=2}^{T} f_t\left(x_t, \xi_t\right)\right] \\ \text{s.t.} & x_1 \in \mathcal{X}_1,\ x_t \in \mathcal{X}_t(x_{t-1}, \xi_t),\ t = 2, \ldots, T. \end{array}$$

The optimization is aimed at minimization of the total cost on average, and performed over the set of policies (also called strategies, decision rules) $\pi = (x_1, x_2(\xi_{[2]}), ..., x_T(\xi_{[T]}))$, satisfying the feasibility constraints.

The values of the decision vector $x_t \in \mathbb{R}^{n_t}$ may depend on the information (history) $\xi_{[t]} := (\xi_1, .., \xi_t)$ available up to time $t$, but not on the future observations. The data process $\xi_t \in \mathbb{R}^{d_t}$, $t = 1, 2, ...$, is modeled as *random* (stochastic), first stage decision $x_1$ is deterministic, made before observing any realization of the random data process ($\xi_1$ is deterministic, used for uniformity of notation).

In the linear case, the costs $f_t(x_t, \xi_t) := c_t^\top x_t$ are linear and

$$\mathcal{X}_t(x_{t-1}, \xi_t) := \{x_t : B_t x_{t-1} + A_t x_t = b_t,\ x_t \geq 0\},\ \ t = 2, ..., T.$$

That is

$$\begin{array}{ll} \min\limits_{x_1, x_2(\cdot), ..., x_T(\cdot)} & c_1^\top x_1 + \mathbb{E}\left[\sum_{t=2}^T c_t^\top x_t\right] \\ \text{s.t.} & A_1 x_1 = b_1,\ x_1 \geq 0, \\ & B_t x_{t-1} + A_t x_t = b_t,\ x_t \geq 0,\ t = 2, \ldots, T, \end{array}$$

where components of $(c_t, b_t, B_t, A_t)$ are functions of $\xi_t$, $t = 2, ..., T$.

What does it mean to solve the stochastic programming problem?

Is it computationally feasible to solve such problems?

What is a value of solving **multistage** problem?

How do we know probability distribution of the data process?

Why do we minimize (optimise) on average?

What does it mean that our solution is **time consistent**?

Why do we optimize the total cost?

In order to solve stochastic programming problems numerically the (continuous) distribution of the data process should be discretized by generating a finite number of realizations of the data process (the scenarios approach). Size of the deterministic equivalent problem is proportional to the number of generated scenarios.

If the number of realizations (scenarios) of the process $\xi_t$ is finite, then the above (linear) problem can be written as one large (linear) programming problem (the deterministic equivalent).

The two stage ($T = 2$) case,

$$\begin{array}{ll} \min\limits_{x, y(\cdot)} & c_1^\top x + \mathbb{E}\left[c_2^\top y\right] \\ \text{s.t.} & A_1 x = b_1,\ x \geq 0, \\ & B_2 x + A_2 y = b_2,\ y \geq 0. \end{array}$$

When the number of scenarios (realizations) $\xi^1, ..., \xi^K$ of the second stage random variable is finite with the respective probabilities $p_k > 0$, the deterministic equivalent:

$$\begin{array}{ll} \min\limits_{x, y_1, \ldots, y_K} & c_1^\top x + \sum_{k=1}^K p_k (c^k)^\top y_k \\ \text{s.t.} & A_1 x = b_1,\ x \geq 0, \\ & B^k x + A^k y_k = b^k,\ y_k \geq 0,\ k = 1, \ldots, K. \end{array}$$

Here $x$ is the first stage decision, $y_1, ..., y_K$ are second stage decisions - one copy for every scenario, and $(c^k, b^k, B^k, A^k)$ is realization of the respective parameters for scenario $\xi^k$.

## Distributionally robust - risk averse approach

$$\min_{x,y(\cdot)} \quad f_1(x) + \mathcal{R}\left[f_2\left(y,\xi\right)\right] \quad s.t. \quad x \in \mathcal{X}_1,\ y \in \mathcal{X}_2(x,\xi).$$

Functional $\mathcal{R}(\cdot)$ is defined on a space of random variables

$$\mathcal{R}(Z) = \sup_{P \in \mathfrak{M}} \mathbb{E}_P[Z],$$

where $\mathfrak{M}$ is a specified set of probability measures (distributions).

Alternative formulation

$$\min_{x \in \mathcal{X}_1} f_1(x) + \mathcal{R}[Q(x,\xi)]$$

where $Q(x,\xi)$ is the optimal value of the second stage problem:

$$\min_{y} f_2(y,\xi) \quad s.t.\ y \in \mathcal{X}_2(x,\xi).$$

Relatively complete recourse: for every $x \in \mathcal{X}_1$ the second stage problem has feasible solution for almost every $\xi$.

Without the relatively complete recourse it could happen that for some $x \in \mathcal{X}_1$ and a realization of the random vector $\xi$ the second stage problem is infeasible and $Q(x, \xi) = +\infty$.

Time consistency question: are the two formulations equivalent? By solving the first formulation we obtain solution $\bar{x}$ and $\bar{y}(\xi)$. Is it true that $\bar{y}(\xi)$ is an optimal solution of the second stage problem for $x = \bar{x}$ and a.e. $\xi$? Should we recalculate the second stage decision after observing realization of the random vector $\xi$?

The functional $\mathcal{R}(\cdot)$ is monotone, i.e., if $Z, Z'$ are random variables such that $Z \geq Z'$, then $\mathcal{R}(Z) \geq \mathcal{R}(Z')$. It is said that $\mathcal{R}$ is *strictly* monotone, if $Z \geq Z'$ and $Z \neq Z'$, then $\mathcal{R}(Z) > \mathcal{R}(Z')$.

For example let $\Omega = \{\omega_1, ..., \omega_m\}$ be finite and $Z, Z' : \Omega \to \mathbb{R}$. Then $\mathfrak{M}$ is a subset of

$$\Delta_m = \left\{ p \in \mathbb{R}^m : \sum_{i=1}^{m} p_i = 1,\ p \geq 0 \right\}.$$

Suppose that $\mathfrak{M}$ is convex and closed. The functional

$$\mathcal{R}(Z) = \sup_{p \in \mathfrak{M}} \sum_{i=1}^{m} p_i Z(\omega_i)$$

is strictly monotone iff for every $p \in \mathfrak{M}$ it follows that $p_i > 0$, $i = 1, ..., m$.

Computational complexity of solving **two-stage** linear stochastic programs (deterministic point of view).

Suppose that the components of the second stage random vector $\xi = (\xi_1, ..., \xi_d) \in \mathbb{R}^d$ are independent and discretized with $m$ discretization points per component $\xi_i$, $i = 1, ..., d$. Then the total number of scenarios is $m^d$.

*The approximate solutions, with a sufficiently high accuracy, of linear two-stage stochastic programs with fixed recourse are $\#P$-hard even if the random problem data is governed by independent uniform distributions* (Dyer and Stougie (2006), Hanasusanto, Kuhn and Wiesemann (2016)).

## Sample complexity of solving stochastic programs

Static (two stage) stochastic problem

$$\min_{x \in \mathcal{X}} \{f(x) = \mathbb{E}[F(x, \xi)]\}.$$

Generate an iid sample $\xi^j$, $j = 1, ..., N$, of random vector $\xi$ and approximate the expectation $\mathbb{E}[F(x, \xi)]$ by the respective sample average.

$$\widehat{f}_N(x) = \frac{1}{N} \sum_{j=1}^{N} F(x, \xi^j).$$

This leads to the following so-called Sample Average Approximation (SAA) of the 'true' problem

$$\min_{x \in \mathcal{X}} \widehat{f}_N(x).$$

Rate of convergence: For fixed $x$ by the Central Limit Theorem, $N^{1/2}\left(\widehat{f}_N(x) - f(x)\right)$ converges in distribution to normal $N(0, \sigma^2(x))$ with $\sigma^2(x) = \mathrm{Var}[F(x, \xi)]$. That is, $\widehat{f}_N(x)$ converges to $f(x)$ at a rate of $O_p(N^{-1/2})$.

Under certain regularity conditions

$$\widehat{\vartheta}_N = \inf_{x \in \mathcal{X}^*} \widehat{f}_N(x) + o_p(N^{-1/2}),$$

where $\mathcal{X}^*$ is the set of optimal solutions of the true problem and $\widehat{\vartheta}_N$ is the optimal value of the SAA problem. In particular if the true problem has unique optimal solution $x^*$, then

$$\widehat{\vartheta}_N = \widehat{f}_N(x^*) + o_p(N^{-1/2}).$$

## Large Deviations type bounds.

Suppose that $\varepsilon > \delta \geq 0$, the set $\mathcal{X} \subset \mathbb{R}^n$ is of finite diameter $D$ and

$$\left|F(x',\xi) - F(x,\xi)\right| \leq L\|x' - x\|, \quad x', x \in \mathcal{X} \text{ and } a.e.\ \xi.$$

Then (under certain regularity conditions) for the sample size

$$N \geq \left(\frac{O(1)LD}{\varepsilon - \delta}\right)^2 \left[n \log\left(\frac{O(1)DL}{\varepsilon - \delta}\right) + \log\left(\frac{1}{\alpha}\right)\right],$$

we are guaranteed that $\Pr\left(\widehat{\mathcal{S}}_N^{\delta} \subset \mathcal{S}^{\varepsilon}\right) \geq 1 - \alpha$. Here $\widehat{\mathcal{S}}_N^{\delta}$ and $\mathcal{S}^{\varepsilon}$ are the sets of $\delta$-optimal and $\varepsilon$-optimal solutions of the SAA and true problems respectively (Shapiro (2003)).

How to extend these type of results to distributionally robust - risk averse settings?

## Sample complexity of $T$-stage stochastic programs

In order for the optimal value and solutions of the SAA problem to converge to their true counterparts all sample sizes $N_2, ..., N_T$ should tend to infinity. Furthermore, available estimates of the sample sizes required for a first stage solution of the SAA problem to be $\varepsilon$-optimal for the true problem, with a given confidence (probability), sums up to a number of scenarios which grows as $O(\varepsilon^{-2(T-1)})$ with decrease of the error level $\varepsilon > 0$. *This indicates that from the point of view of the number of scenarios, complexity of multistage programming problems grows exponentially with increase of the number of stages* (Shapiro, Nemirovski (2005)).

These are upper bounds for sample complexity, which guarantee (under certain regularity conditions) that for obtained estimates of the sample sizes, the SAA problem provides an $\varepsilon$-optimal solution of the true problem with high probability. Are these bounds tight?

Can Central Limit Theorem type results be formulated for the multistage programs?

How to extend this to distributionally robust - risk averse settings?

## Dynamic programming equations approach

Recall that $\xi_{[t]} := (\xi_1, .., \xi_t)$ denotes history of the data process. Going recursively backwards in time. At stage $T$ consider

$$Q_T(x_{T-1}, \xi_T) := \inf_{x_T \in \mathcal{X}_T(x_{T-1}, \xi_T)} f_T(x_T, \xi_T).$$

At stages $t = T-1, ..., 2$, consider

$$Q_t(x_{t-1}, \xi_{[t]}) := \inf_{x_t \in \mathcal{X}_t(x_{t-1}, \xi_t)} f_t(x_t, \xi_t) + \underbrace{\mathbb{E}_{|\xi_{[t]}} \left[ Q_{t+1}(x_t, \xi_{[t+1]}) \right]}_{\mathcal{Q}_{t+1}(x_t, \xi_{[t]})},$$

where $\mathbb{E}_{|\xi_{[t]}}$ is the conditional expectation. At the first stage solve:

$$\operatorname*{Min}_{x_1 \in \mathcal{X}_1} f_1(x_1) + \mathbb{E}[Q_2(x_1, \xi_2)].$$

Suppose that the random data process is stagewise independent, i.e., $\xi_{t+1}$ is independent of $\xi_{[t]}$, $t = 1, ..., T-1$. Then

$$\mathcal{Q}_T(x_{T-1}) := \mathbb{E}_{|\xi_{[T-1]}}[Q_T(x_{T-1}, \xi_T)] = \mathbb{E}[Q_T(x_{T-1}, \xi_T)]$$

does not depend on $\xi_{[T-1]}$. By induction going backward in time it is possible to show that the expected value cost-to-go functions

$$\mathcal{Q}_{t+1}(x_t) = \mathbb{E}[Q_{t+1}(x_t, \xi_{t+1})]$$

do not depend on $\xi_{[t]}$, and an optimal policy $\bar{x}_t = \bar{x}_t(\bar{x}_{t-1}, \xi_t)$ is given by

$$\bar{x}_t \in \underset{x_t \in \mathcal{X}_t(\bar{x}_{t-1}, \xi_t)}{\arg\min} f_t(x_t, \xi_t) + \mathcal{Q}_{t+1}(x_t).$$

## Difficulties in solving dynamic programming equations

(stagewise independent case)

- Computing the expectation $\mathbb{E}[Q_{t+1}(x_t, \xi_{t+1})]$ - discretize (marginal) distribution of $\xi_{t+1}$, e.g., by sampling (the SAA approach).

- At every stage $t$, in order to compute the cost-to-go function $Q_t(x_{t-1}, \xi_t)$ there is a need to solve optimization problem

$$\min_{x_t \in \mathcal{X}_t(x_{t-1}, \xi_t)} f_t(x_t, \xi_t) + \mathcal{Q}_{t+1}(x_t).$$

## Curse of dimensionality of dynamic programming

One of the main difficulties in solving the dynamic programming equations (of the SAA problem) is how to represent the cost-to-go functions $\mathcal{Q}_{t+1}(x_t)$ in a computationally feasible way.

For dimension of $x_t$ say greater than 3 and large number of stages it is practically impossible to solve the dynamic programming equations with high accuracy. Several alternatives were suggested in the literature.

## Approximate dynamic programming

Basic idea is to approximate the cost-to-go functions $\mathcal{Q}_t(\cdot)$ by a class of computationally manageable functions. Consider

$$\begin{array}{cl} \min\limits_{x_1, x_2(\cdot), \ldots, x_T(\cdot)} & f_1(x_1) + \mathbb{E}\left[\sum_{t=2}^{T} f_t\left(x_t, \xi_t\right)\right] \\ \text{s.t.} & A_1 x_1 = b_1,\ x_1 \geq 0, \\ & B_t x_{t-1} + A_t x_t = b_t,\ x_t \geq 0,\ t = 2, \ldots, T. \end{array}$$

If the functions $f_t\left(x_t, \xi_t\right)$ are convex in $x_t$, then the cost-to-go functions are convex in $x_t$. Then it is natural to approximate $\mathcal{Q}_t(\cdot)$ by piecewise linear functions given by maximum of cutting hyperplanes.

Stochastic Dual Dynamic Programming (SDDP) method (Kelley (1960), Birge (1985), Pereira and Pinto (1991))

Consider the stagewise independent case. For trial decisions $\bar{x}_t$, $t = 1, ..., T-1$, at the backward step of the SDDP algorithm, piecewise linear approximations $\mathfrak{Q}_t(\cdot)$ of the cost-to-go functions $\mathcal{Q}_t(\cdot)$ are constructed by solving problems

$$\operatorname*{Min}_{x_t \in \mathbb{R}^{n_t}} (c_t^j)^{\mathsf{T}} x_t + \mathfrak{Q}_{t+1}(x_t) \text{ s.t. } B_t^j \bar{x}_{t-1} + A_t^j x_t = b_t^j,\ x_t \geq 0,$$

$j = 1, ..., N_t$, and their duals, going backward in time $t = T, ..., 1$.

By construction

$$\mathcal{Q}_t(\cdot) \geq \mathfrak{Q}_t(\cdot),\ t = 2, ..., T.$$

Therefore the optimal value of

$$\underset{x_1 \in \mathbb{R}^{n_1}}{\text{Min}}\ c_1^{\mathsf{T}} x_1 + \mathfrak{Q}_2(x_1)\ \text{ s.t. } A_1 x_1 = b_1,\ x_1 \geq 0$$

gives a lower bound for the optimal value $\widehat{v}_N$ of the SAA problem.

We also have that

$$v^0 \geq \mathbb{E}[\widehat{v}_N].$$

Therefore *on average* $\widehat{v}_N$ is also a lower bound for the optimal value of the true problem.

The approximate cost-to-go functions $\mathfrak{Q}_2, ..., \mathfrak{Q}_T$ and a feasible first stage solution $\bar{x}_1$ define a feasible policy. That is for a realization (sample path) $\xi_1, ..., \xi_T$ of the data process, $\bar{x}_t = \bar{x}_t(\xi_{[t]})$ are computed recursively in $t = 2, ..., T$ as a solution of

$$\underset{x_t \geq 0}{\text{Min}}\, c_t^{\mathsf{T}} x_t + \mathfrak{Q}_{t+1}(x_t) \text{ s.t. } B_t \bar{x}_{t-1} + A_t x_t = b_t.$$

In the *forward step* of the SDDP algorithm $M$ sample paths (scenarios) are generated and the corresponding $\bar{x}_t$, $t = 2, ..., T$, are used as trial points in the next iteration of the backward step. At the same time this allows to construct a statistical upper bound for the optimal value of the corresponding multistage problem - SAA or the "true" problem depending from what data process the random scenarios were generated.

## Convergence of the SDDP algorithm

It is possible to show that, under mild regularity conditions, the SDDP algorithm converges as the number of iterations go to infinity. However, the convergence can be very slow.

What is the convergence rate of SDDP algorithm as the number of stages increases?

In what situations the method works? Nobody solved the "curse of dimensionality" problem.

In the two stage case ($T = 2$) this becomes just Kelley's cutting plane method. How more efficient regularized type algorithms can be adapted to the multistage setting?

When the data process is not stagewise independent, how to use Markovian structure of the data process in an efficient way?

## Distributionally robust - risk averse approach to multistage programming

Consider a set $\mathfrak{M}$ of probability measures on sample space $(\Omega, \mathcal{F})$ and functional

$$\mathcal{R}(Z) := \sup_{P \in \mathfrak{M}} \mathbb{E}_P[Z]$$

defined on a linear space $\mathcal{Z}$ of $\mathcal{F}$-measurable functions (random variables) $Z : \Omega \to \mathbb{R}$.

Popular approaches to define the ambiguity set $\mathfrak{M}$: (i) by distributions in some sense close to a specified reference distribution $\mathbb{P}$, or (ii) by moment constraints.

In case (i), the set $\mathfrak{M}$ is assumed to consist of probability measures absolutely continuous with respect to $\mathbb{P}$, and

$$\mathfrak{A} = \{\zeta : dP/d\mathbb{P},\ P \in \mathfrak{M}\},$$

is the corresponding set of densities.

In that case it is assumed that $\mathcal{Z} = L_p(\Omega, \mathcal{F}, \mathbb{P})$, the set $\mathfrak{A}$ of densities is a subset of its dual space $\mathfrak{A} \subset \mathcal{Z}^*$, $\mathcal{Z}^* = L_q(\Omega, \mathcal{F}, \mathbb{P})$, and

$$\langle \zeta, Z \rangle = \int_\Omega \zeta(\omega) Z(\omega) d\mathbb{P}(\omega)$$

is well defined for $Z \in \mathcal{Z}$ and $\zeta \in \mathcal{Z}^*$.

In case of moment constraints there is no reference probability measure and the above duality framework cannot be applied.

The functional $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ satisfies the following conditions
(i) Subadditivity:

$$\mathcal{R}(Z_1 + Z_2) \leq \mathcal{R}(Z_1) + \mathcal{R}(Z_2),\ Z_1, Z_2 \in \mathcal{Z}.$$

(ii) Monotonicity: If $Z_1, Z_2 \in \mathcal{Z}$ and $Z_1 \succeq Z_2$, then $\mathcal{R}(Z_1) \geq \mathcal{R}(Z_2)$.
(iii) Translation Equivariance: If $a \in \mathbb{R}$ and $Z \in \mathcal{Z}$, then $\mathcal{R}(Z + a) = \mathcal{R}(Z) + a$.
(iv) Positive Homogeneity:

$$\mathcal{R}(\alpha Z) = \alpha \mathcal{R}(Z),\ \ Z \in \mathcal{Z},\ \alpha > 0.$$

Conversely if $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ satisfies these conditions, then by Fenchel - Moreau Theorem it can be represented in the dual form (Ruszczyński and Shapiro, 2006). Functionals $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ satisfying these conditions (axioms) were called *coherent risk measures* in Artzner, Delbaen, Eber and Heath (1999).

Expectation operator has the following property (recall that $\xi_1$ is deterministic, so $\mathbb{E}_{|\xi_1} = \mathbb{E}$)

$$\mathbb{E}[Z] = \mathbb{E}_{|\xi_1}\left[\mathbb{E}_{|\xi_2}\left[\cdots \mathbb{E}_{|\xi_{[T-1]}}[Z]\right]\right].$$

Consider the nested functional

$$\begin{aligned}\mathfrak{R}(Z) &:= \sup_{P\in\mathfrak{M}} \mathbb{E}_P\left[\sup_{P\in\mathfrak{M}} \mathbb{E}_{P|\xi_{[2]}}\left[\cdots \sup_{P\in\mathfrak{M}} \mathbb{E}_{P|\xi_{[T-1]}}[Z]\right]\right] \\ &= \mathcal{R}\left[\mathcal{R}_{|\xi_{[2]}}\left[\cdots \mathcal{R}_{|\xi_{[T-1]}}[Z]\right]\right],\ Z\in\mathcal{Z}.\end{aligned}$$

Note that

$$\begin{aligned}\mathcal{R}(Z) &= \sup_{P\in\mathfrak{M}} \mathbb{E}_P\left[\mathbb{E}_{P|\xi_{[2]}}\left[\cdots \mathbb{E}_{P|\xi_{[T-1]}}[Z]\right]\right] \\ &\le \sup_{P\in\mathfrak{M}} \mathbb{E}_P\left[\sup_{P\in\mathfrak{M}} \mathbb{E}_{P|\xi_{[2]}}\left[\cdots \sup_{P\in\mathfrak{M}} \mathbb{E}_{P|\xi_{[T-1]}}[Z]\right]\right] = \mathfrak{R}(Z).\end{aligned}$$

Rigourous definition of the conditional functional (?)

$$\mathcal{R}_{|\xi_{[t]}}(Z) = \sup_{P\in\mathfrak{M}} \mathbb{E}_{P|\xi_{[t]}}[Z]$$

Nested functional $\mathfrak{R}(\cdot)$ satisfies the axioms (i) - (iv) and hence can be represented in the dual form

$$\mathfrak{R}(Z) = \sup_{P \in \widehat{\mathfrak{M}}} \mathbb{E}_P[Z]$$

for some set $\widehat{\mathfrak{M}}$ of probability measures.

In general $\mathfrak{M} \neq \widehat{\mathfrak{M}}$ even in the rectangular case when

$$\mathfrak{M} = \{P = P_2 \times \cdots \times P_T : P_t \in \mathfrak{M}_t,\ t = 1, ..., T\}$$

where $\mathfrak{M}_t$ is a set of marginal distributions of $\xi_t$.

Constructive description of the set $\widehat{\mathfrak{M}}$?

## Minimization of nested risk functionals

$$\min_{\pi \in \Pi} f_1(x_1) + \mathfrak{R}\Big[\sum_{t=2}^{T} f_t(x_t^\pi, \xi_t)\Big],$$

where $\Pi$ is the set of feasible policies and ($\mathcal{R}_{|\xi_1} = \mathcal{R}$ since $\xi_1$ is deterministic) $\mathfrak{R}(\cdot) := \mathcal{R}_{|\xi_1}\Big[\cdots \mathcal{R}_{|\xi_{[T-1]}}(\cdot)\Big]$.
Note that unlike the expectation, the risk functional $\mathcal{R}$ does not have the decomposition property, i.e. $\mathfrak{R}(\cdot) \neq \mathcal{R}(\cdot)$.

Since $\mathfrak{R}(\cdot) = \sup_{P \in \widehat{\mathfrak{M}}} \mathbb{E}_P[\cdot]$, we can write this risk averse problem as

$$\min_{\pi \in \Pi} \sup_{P \in \widehat{\mathfrak{M}}} \mathbb{E}_P\Big[f_1(x_1) + \sum_{t=2}^{T} f_t(x_t^\pi, \xi_t)\Big].$$

The dual of this problem is

$$\max_{P \in \widehat{\mathfrak{M}}} \inf_{\pi \in \Pi} \mathbb{E}_P\Big[f_1(x_1) + \sum_{t=2}^{T} f_t(x_t^\pi, \xi_t)\Big].$$

For the nested formulation it is possible to write dynamic programming equations with the respective cost-to-go (value) functions $Q_t(x_{t-1}, \xi_{[t]})$ given by the optimal value of the problem

$$\begin{array}{ll} \min\limits_{x_t \in \mathcal{X}_t} & c_t(x_t, \xi_t) + \underbrace{\sup\limits_{P \in \mathfrak{M}} \mathbb{E}_{P|\xi_{[t]}}}_{\mathcal{R}_{|\xi_{[t]}}}[Q_{t+1}(x_t, \xi_{[t+1]})] \\ \text{s.t.} & B_t x_{t-1} + A_t x_t = b_t. \end{array}$$

At the first stage the following problem is supposed to be solved

$$\begin{array}{ll} \min\limits_{x_1 \in \mathcal{X}_1} & c_1(x_1) + \sup\limits_{P \in \mathfrak{M}} \mathbb{E}_P[V_2(x_1, \xi_2)] \\ \text{s.t.} & A_1 x_1 = b_1. \end{array}$$

In the rectangular (stagewise independence) case these equations simplify to

$$\begin{aligned} \min_{x_t \in \mathcal{X}_t} \quad & c_t(x_t, \xi_t) + \underbrace{\sup_{P_{t+1} \in \mathfrak{M}_{t+1}} \mathbb{E}_{P_{t+1}}}_{\mathcal{R}}[Q_{t+1}(x_t, \xi_{t+1})] \\ \text{s.t.} \quad & B_t x_{t-1} + A_t x_t = b_t. \end{aligned}$$

It is possible to extend the SDDP method to nested risk functionals. Assuming the stagewise independence, replace $\mathbb{E}[Q_{t+1}(x_t, \xi_{t+1})]$ in the respective dynamic programming equations with $\mathcal{R}[Q_{t+1}(x_t, \xi_{t+1})]$ (Shapiro, Tekaya, da Costa, Pereira Soares (2013)).

## Periodical infinite horizon multistage programs

Shapiro and Ding (2019)

Infinite horizon problem with discount factor $\gamma \in (0,1)$:

$$\min_{\pi \in \Pi} \; f_1(x_1) + \mathfrak{R}\left[\sum_{t=2}^{\infty} \gamma^{t-1} f_t\left(x_t, \xi_t\right)\right],$$

where $\mathfrak{R}$ is the nested functional and $\Pi$ is the set of policies satisfying the feasibility constraints

$$x_t \in \mathcal{X}_t, \; B_t x_{t-1} + A_t x_t = b_t.$$

Suppose that the data process $\xi_t$ is stagewise independent, and the problem has periodic structure with period $\mathsf{m} \in \mathbb{N}$:

• The random vectors $\xi_t$ and $\xi_{t+\mathsf{m}}$ have the same distribution, with support $\Xi \subset \mathbb{R}^d$, for $t \geq 2$ (recall that $\xi_1$ is deterministic).

• The functions $b_t(\cdot)$, $B_t(\cdot)$, $A_t(\cdot)$ and $f_t(\cdot,\cdot)$ have period $\mathsf{m}$, i.e., are the same for $t = \tau$ and $t = \tau + \mathsf{m}$, $t = 2, ...,$ and the sets $\mathcal{X}_t$ are nonempty and $\mathcal{X}_t = \mathcal{X}_{t+\mathsf{m}}$ for all $t$.

Under these assumptions the value functions $\mathcal{Q}_t(\cdot)$ and $\mathcal{Q}_{t+\mathsf{m}}(\cdot)$ of the dynamic equations are the same for all $t \geq 2$. This leads to the following periodical variant of Bellman equations for the value functions $\mathcal{Q}_2(\cdot), ..., \mathcal{Q}_{\mathsf{m}+1}(\cdot)$:

$$\mathcal{Q}_\tau(x_{\tau-1}) = \mathcal{R}[Q_\tau(x_{\tau-1}, \xi_\tau)],$$

$$Q_\tau(x_{\tau-1}, \xi_\tau) = \inf_{x_\tau \in \mathcal{X}_\tau} \left\{ \begin{array}{c} f_\tau(x_\tau, \xi_\tau) + \gamma \mathcal{Q}_{\tau+1}(x_\tau) : \\ B_\tau x_{\tau-1} + A_\tau x_\tau = b_\tau \end{array} \right\},$$

for $\tau = 2, ..., \mathsf{m}+1$, and $\mathcal{Q}_{\mathsf{m}+2}$ replaced by $\mathcal{Q}_2$ for $\tau = \mathsf{m}+1$. Consequently for $t \geq \mathsf{m}+2$ the corresponding value functions are defined recursively as $Q_t(\cdot, \xi_t) = Q_{t-\mathsf{m}}(\cdot, \xi_t)$, and hence $\mathcal{Q}_t(\cdot) = \mathcal{Q}_{t-\mathsf{m}}(\cdot)$.

In order to show that Bellman equations have a solution the standard approach is to show that the corresponding operator is a contraction mapping and hence has a unique fixed point. That is, suppose for the sake of simplicity that the period length $\mathsf{m} = 1$ (in that case we remove the subscript $t$ from the data). Bellman equation takes the form

$$\mathcal{Q}(x) = \mathcal{R}[Q(x, \xi)],$$

$$Q(x, \xi) = \inf_{x' \in \mathcal{X}} \left\{ f(x', \xi) + \gamma \mathcal{Q}(x') : B(\xi)x + A(\xi)x' = b(\xi) \right\}.$$

Let $\mathbb{B}(\mathcal{X})$ be the space of bounded functions $g : \mathcal{X} \to \mathbb{R}$ with the sup-norm $\|g\| = \sup_{x\in\mathcal{X}} |g(x)|$. Consider mapping $\mathfrak{T} : \mathbb{B}(\mathcal{X}) \to \mathbb{B}(\mathcal{X})$ defined as

$$\mathfrak{T}(g)(x) = \mathcal{R}[\Psi_g(x, \xi)],$$

$$\Psi_g(x, \xi) = \inf_{x'\in\mathcal{X}} \{f(x', \xi) + \gamma g(x') : B(\xi)x + A(\xi)x' = b(\xi)\}.$$

Then $\mathcal{Q}(\cdot)$ is a solution of Bellman equation if $\mathcal{Q}$ is a fixed point of $\mathfrak{T}$. The mapping $\mathfrak{T}$ is a contraction mapping, i.e.,

$$\|\mathfrak{T}(g) - \mathfrak{T}(g')\| \leq \gamma\|g - g'\|, \ \forall g, g' \in \mathbb{B}(\mathcal{X}).$$

This can be extended to period $\mathsf{m} \geq 1$, and also to risk verse problems with expectation operator $\mathbb{E}$ replaced by a coherent law invariant risk measure $\varrho$ with the respective conditional analogues.

## Sample complexity of Bellman equations

Consider the case of the period length $\mathsf{m} = 1$ and expectation functional $\mathcal{R} = \mathbb{E}$. Let $P$ be true distribution of the random vector $\xi$ and $\widehat{P}_N = N^{-1} \sum_{j=1}^{N} \delta_{\xi^j}$ be its empirical counterpart based on sample $\xi^1, ..., \xi^N$. Let $\mathcal{Q}$ and $\widehat{\mathcal{Q}}_N$ be solutions of Bellman equations associated with $P$ and $\widehat{P}_N$, respectively. Then

$$\|\mathcal{Q} - \widehat{\mathcal{Q}}_N\| \leq (1-\gamma)^{-1} \left\| \mathbb{E}_P[\Psi(x,\xi)] - \mathbb{E}_{\widehat{P}_N}[\Psi(x,\xi)] \right\|_\infty$$

where $\|g\|_\infty = \sup_{x \in \mathcal{X}} |g(x)|$ is the sup-norm,

$$\Psi(x,\xi) = \inf_{x' \in \mathcal{X}} \left\{ f(x',\xi) + \gamma \mathcal{Q}(x') : B(\xi)x + A(\xi)x' = b(\xi) \right\}$$

and $\mathbb{E}_{\widehat{P}_N}[\Psi(x,\xi)] = N^{-1} \sum_{j=1}^{N} \Psi(x,\xi^j)$.

What happens with sample complexity as the discount factor $\gamma$ approaches one?

## Duals of periodical linear programs

Dual approach to construction of upper bounds was initiated in Leclére, Carpentier, Chancelier, Lenoir, Pacaud (2019), we follow here the approach of Guigues, Shapiro, Cheng (2019)

Consider linear (risk neutral) multistage stochastic program

$$\begin{array}{ll} \min\limits_{x_t \geq 0} & \mathbb{E}\left[\sum_{t=1}^T c_t^\top x_t\right] \\ \text{s.t.} & A_1 x_1 = b_1, \\ & B_t x_{t-1} + A_t x_t = b_t,\ t = 2, ..., T. \end{array}$$

Dualization of the the feasibility constraints

The Lagrangian

$$L(x,\pi) = \mathbb{E}\left[\sum_{t=1}^{T} c_t^\top x_t + \pi_t^\top (b_t - B_t x_{t-1} - A_t x_t)\right]$$

in variables $x = (x_1(\xi_{[1]}), \ldots, x_T(\xi_{[T]}))$ and $\pi = (\pi_1(\xi_{[1]}), \ldots, \pi_T(\xi_{[T]}))$ with the convention that $x_0 = 0$. Dualization of the feasibility constraints leads to the following dual

$$\begin{array}{ll} \max\limits_{\pi} & \mathbb{E}\Big[\sum\limits_{t=1}^{T} b_t^\top \pi_t\Big] \\ \text{s.t.} & A_T^\top \pi_T \le c_T, \\ & A_{t-1}^\top \pi_{t-1} + \mathbb{E}_{|\xi_{[t-1]}}\left[B_t^\top \pi_t\right] \le c_{t-1},\ t = 2, ..., T. \end{array}$$

The optimization is over policies $\pi_t = \pi_t(\xi_{[t]})$, $t = 1, ..., T$.

## Dynamic programming equations for the dual problem

Assume the stagewise independence condition and finite number of scenarios $N_t$ per stage and respective probabilities $p_{tj}$. At the last stage $t = T$ we have the following problem

$$\begin{array}{ll} \max\limits_{\pi_{T1},\ldots,\pi_{TN_T}} & \mathbb{E}[b_T^\top \pi_T] = \sum\limits_{j=1}^{N_T} p_{Tj} b_{Tj}^\top \pi_{Tj} \\ \text{s.t.} & A_{Tj}^\top \pi_{Tj} \leq c_{Tj},\ j = 1, ..., N_T, \\ & A_{T-1}^\top \pi_{T-1} + \sum\limits_{j=1}^{N_T} p_{Tj} B_{Tj}^\top \pi_{Tj} \leq c_{T-1}. \end{array}$$

The optimal value $V_T(\pi_{T-1}, \xi_{T-1})$ and an optimal solution $(\bar{\pi}_{T1}, \ldots, \bar{\pi}_{TN_T})$ of that problem are functions of vectors $\pi_{T-1}$ and $c_{T-1}$ and matrix $A_{T-1}$.

And so on going backward in time we can write the respective dynamic programming equations for $t = T-1, ..., 2$, as

$$\begin{aligned} \max_{\pi_{t1},\ldots,\pi_{tN_t}} \quad & \sum_{j=1}^{N_t} p_{tj} \left[ b_{tj}^\top \pi_{tj} + V_{t+1}(\pi_{tj}, \xi_{tj}) \right] \\ \text{s.t.} \quad & A_{t-1}^\top \pi_{t-1} + \sum_{j=1}^{N_t} p_{tj} B_{tj}^\top \pi_{tj} \leq c_{t-1}, \end{aligned}$$

with $V_t(\pi_{t-1}, \xi_{t-1})$ being the optimal value of the above problem. Finally at the first stage the following problem should be solved

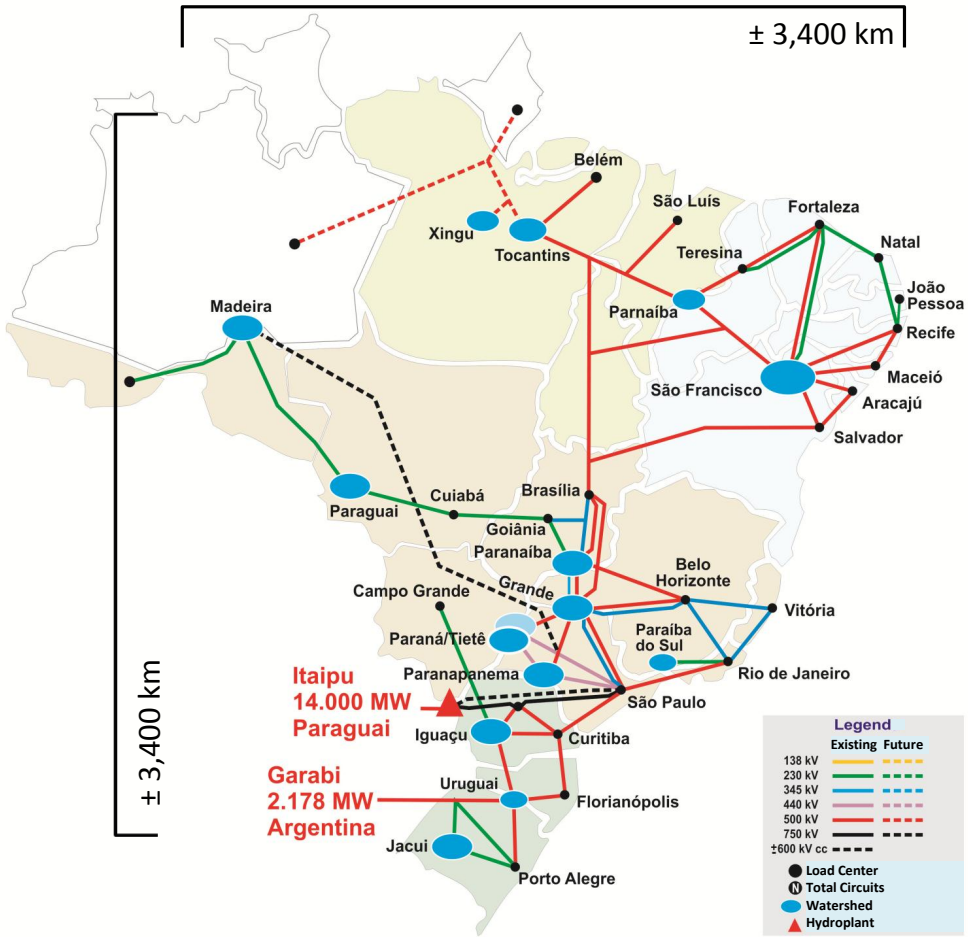
$$\max_{\pi_1} b_1^\top \pi_1 + V_2(\pi_1, \xi_1).$$

## The Brazilian hydro power operation planning problem

The Brazilian power system generation is hydro dominated (about 75% of the installed capacity) and characterized by large reservoirs presenting multi-year regulation capability, arranged in complex cascades over several river basins. The hydro plants use store water in the reservoirs to produce energy in the future, replacing fuel costs from the thermal units. Since the water inflows depend on rainfalls, the amount of future inflows is uncertain and cannot be predicted with a high accuracy.

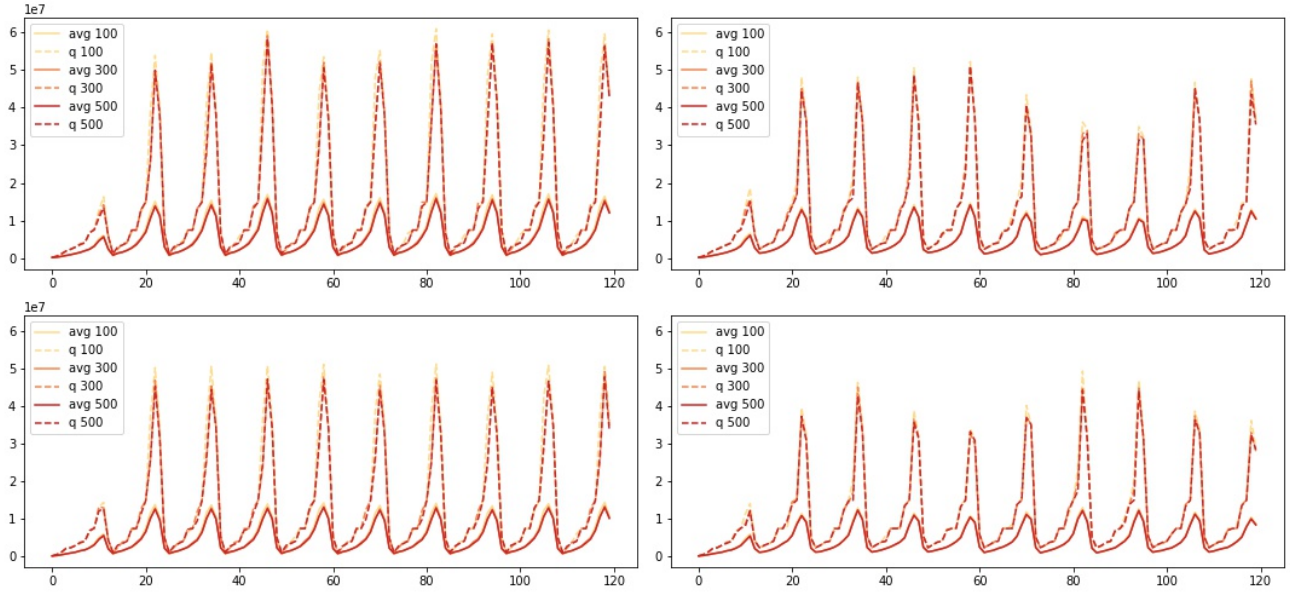
The purpose of hydrothermal system operation planning is to define an operation strategy which, for each stage of the planning period, given the system state at the beginning of the stage, produces generation targets for each plant.

The Brazilian hydro power operation planning problem is a multistage, large scale (more than 200 power plants, of which 141 are hydro plants), stochastic optimization problem. On a high level, planning is for 5 years on monthly basis together with 5 additional years to smooth out the end of horizon effect. This results in 120-stage stochastic programming problem. Four energy equivalent reservoirs are considered, one in each one of the four interconnected main regions, SE, S, N and NE. The resulting policy obtained with the aggregate representation can be further refined, so as to provide decisions for each of the hydro and thermal power plants.

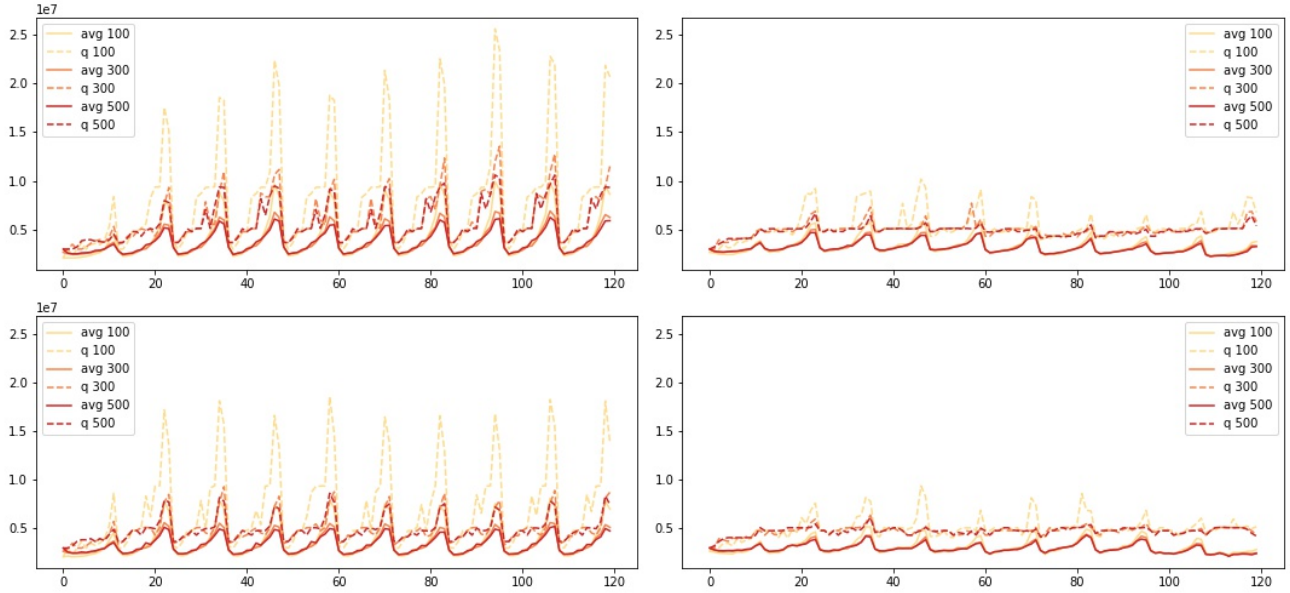

± 3,400 km
± 3,400 km
Belém
São Luís
Fortaleza
Xingu
Tocantins
Teresina
Natal
João Pessoa
Parnaíba
Madeira
Recife
São Francisco
Maceió
Aracajú
Salvador
Brasília
Cuiabá
Paraguai
Goiânia
Paranaíba
Belo Horizonte
Grande
Campo Grande
Vitória
Paraná/Tietê
Paraíba do Sul
Itaipu
14.000 MW
Paraguai
Paranapanema
Rio de Janeiro
São Paulo
Iguaçu
Curitiba
Garabi
2.178 MW
Argentina
Uruguai
Florianópolis
Jacui
Porto Alegre
Legend
Existing Future
138 kV
230 kV
345 kV
440 kV
500 kV
750 kV
±600 kV cc
Load Center
Total Circuits
Watershed
Hydroplant

Comparison of the classical and periodical SDDP (with 8 state variables, period $m = 12$), Shapiro and Ding (2019)

Stored energy (in average value and 0.9 quantile) by periodical SDDP (on the left) and classical SDDP (on the right) for the SAA discretization problem (100 samples per stage) and the true problem (on the bottom) for the risk neutral case with discount factor $\gamma = 0.8$



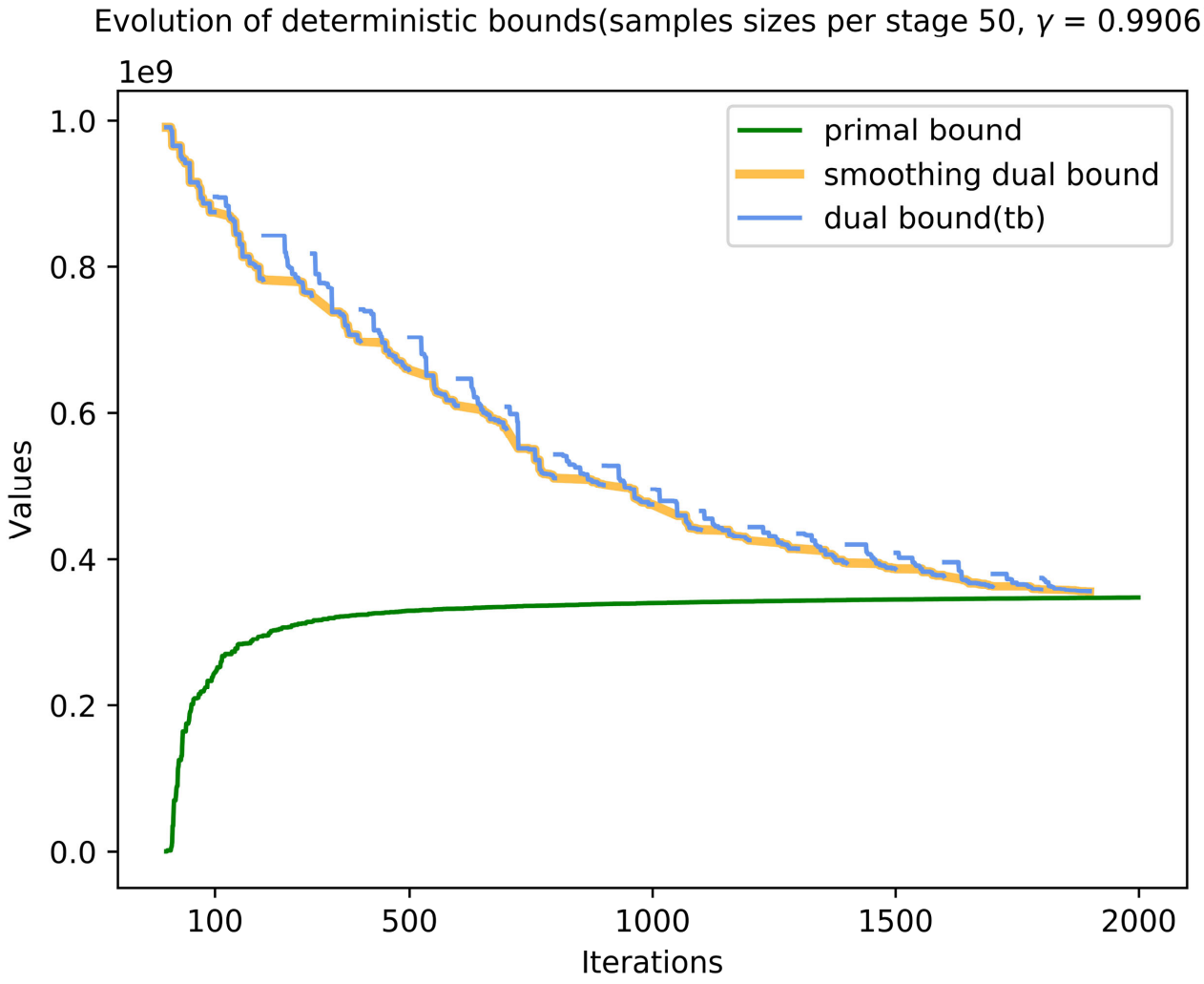
Individual stage costs (in average value and 0.9 quantile) by periodical SDDP (on the left) and classical SDDP (on the right) for the SAA discretization problem (on the above) and the true problem (on the bottom) for the risk neutral case with discount factor $\gamma = 0.9906$ (this $\gamma$ corresponds to the annual discount rate of 12%, that is $1/\gamma^{12} = 1.12$)



## Dual bounds for periodical problem

Shapiro and Cheng (2020)

Hydro-thermal problem with 4 state variables, 50 samples per stage, discount factor $\gamma = 0.9906$ and period $\mathtt{m} = 12$. Evolution of deterministic bounds of primal and dual periodical programs.


Evolution of deterministic bounds(samples sizes per stage 50, $\gamma$ = 0.9906)
1e9
1.0
0.8
0.6
0.4
0.2
0.0
Values
100
500
1000
1500
2000
Iterations
primal bound
smoothing dual bound
dual bound(tb)

## Some references


Shapiro, A. , Dentcheva, D. & Ruszczynski, A. (2014). Lectures on stochastic programming: modeling and theory (2nd). Philadelphia: SIAM .

Shapiro, A. and Nemirovski, A., On complexity of stochastic programming problems, Continuous Optimization: Current Trends and Applications, pp. 111-144, V. Jeyakumar and A.M. Rubinov (Eds.), Springer, 2005.

Shapiro, A. and Ding, L., Periodical Multistage Stochastic Programs, *SIAM Journal on Optimization*, vol. 30, pp. 2083 - 2102, 2020.

Ding, L., Ahmed, S. and Shapiro, A., A Python package for multi-stage stochastic programming,
`http://www.optimization-online.org/DB_HTML/2019/05/7199.html`