An Importance Sampling Approach for Stochastic Dual Dynamic Programming with Correlated Uncertainty

Felipe Cordera, Marcelo Olivares, Fernando Ordóñez (alphabetically)

Faculty of Physical and Mathematical Sciences, University of Chile, Chile

1 Introduction

Multistage stochastic linear programs are often intractable in practice, due to the presence of continuous (or also very fine discrete) uncertainty. Sometimes it is possible to solve them approximately via sampling methods, in a SAA-fashion, such that the uncertainty at each stage is discretized in a manageable size, and a scenario tree is formed. However, two issues are faced:

1) It turns necessary to have an idea about the accuracy of the approximation.
2) The approximate problem may still be computationally challenging to compute due to the nested structure and usually large number of scenarios.

Shapiro (2003) discussed statistical inference of sample average approximations for multistage stochastic linear programs. The author concluded that: "in order to construct a consistent statistical lower bound one needs to employ the conditional sampling scheme" (to build the scenario tree of the approximate problem). On the other hand, any implementable and feasible policy yields a valid statistical upper bound.

The tightness of both bounds is related to the accuracy of the approximation and the quality of the solution found through that approximation. Which brings us to the second issue: how well can we solve the approximate problem?

A Benders decomposition type solution method for linear problems is introduced by Birge (1985). It decomposes the approximate problem by stages and scenarios, and the recourse (or cost-to-go) function at each subproblem is iteratively approximated by *Benders cuts* that yield an outer-linearization.

Since the number of scenarios grows exponentially with the number of stages, solving the complete scenario tree usually becomes intractable. The Stochastic Dual Dynamic Programming (SDDP) algorithm, also based on Benders decomposition, incorporates *Monte Carlo* sampling to select only a subset of the total scenarios at each iteration. This algorithm was introduced in the context of the hydro-thermal scheduling problem (Pereira & Pinto, 1991) and is used to solve large-scale multistage stochastic linear programs.

When the uncertainty is stagewise-independent the recourse function (at a given stage) does not depend on the current scenario. Hence, cuts generated for a subproblem are also valid for any other subproblem at the same stage. This is known as the *cut-sharing*

property, which is critical for an efficient implementation of SDDP (Infanger & Morton, 1996; Queiroz & Morton, 2013). This property ensures that every subproblem will have cuts even if its corresponding scenario was not sampled in previous iterations.

Furthermore, under stagewise-independence assumption, independent random samples may be generated at each stage and used in a conditional sampling scheme to derive the approximate problem by means of a recombining scenario tree. Therefore, statistical properties of bounds are well-established in this case.

Although SDDP is a popular method to solve problems with stagewise-independent uncertainty, two approaches have been suggested to incorporate dependence (Philpott & Matos, 2012; Shapiro et al., 2013; Lohndorf & Shapiro, 2019):

- One approach is to reformulate the uncertainty as a linear-autoregressive process with stagewise-independent errors and treat the original uncertain parameters as additional state variables. To preserve linearity, this approach is restricted to the right-hand side parameters, and does not allow the use of non-linear transformations (e.g., logarithmic or Box-cox) on the autoregressive process, typical in the analysis of highly skewed data or non-negative random variables, such as demands, commodity prices or resources. As it is a reformulation where stagewise-independence still holds, this approach inherits the convergence guarantees of bounds.
- The other approach is based on Markov Chain discretization of the uncertainty process using optimal quantization. It admits a representation of much richer stochastic models, however convergence of bounds is not guaranteed since it breaks with the conditional sampling scheme.

Following the notation of Lohndorf & Shapiro (2019), we refer to both approaches as TS-SDDP and MC-SDDP, respectively.

In this paper we present an alternative approach to incorporate dependence, which makes convergence guarantees with respect to the true problem and allows the usage of a broad range of stochastic models. It is based on Importance Sampling to build the approximate problem. We refer to this approach as IS-SDDP.

In Section 2 we introduce the proposed approach and demonstrate its convergence properties to approximate the real problem. In Section 3 we describe the SDDP-type algorithm to solve the resulting approximate problem. In Section 4 we include some numerical experiments on an inventory problem to test (in terms of statistical convergence and quality of solutions) the proposed approach. In section 5 we conclude.

2 Proposed approach

2.1 Problem statement

In this paper we consider a multistage stochastic linear program of the form:

$$(P)\ \min_{A_1x_1=b_1} c_1x_1 + \mathbb{E}_{\xi_2|\xi_1}\left[\min_{B_2x_1+A_2x_2=b_2} c_2x_2 + \mathbb{E}_{\xi_3|\xi_2}\left[\ldots + \mathbb{E}_{\xi_T|\xi_{T-1}}\left[\min_{B_Tx_{T-1}+A_Tx_T=b_T} c_Tx_T\right]\right]\right]$$

Where $\xi_t = (c_t, A_t, B_t, b_t),\ t = 2, \dots, T$ is a random vector, with $\xi_1 = (c_1, A_1, b_1)$ being deterministic. We assume relatively complete recourse condition. For simplicity, we assume that ξ_t only depends on ξ_{t-1} (i.e., lag one dependency) but can be easily extended with a reformulation.

It is common to refer to this problem as the "real" problem in contrast to its "approximate" counterpart (that we will see later). The latter has the same structure as the former, however every expectation is discretized in a manageable size, resulting in a scenario-tree structure.

Problem (P) may be decomposed by stages. Then, for $t = T, \dots, 1$

$$Q_t(x_{t-1}, \xi_t) = \min_{B_t x_{t-1} + A_t x_t = b_t} c_t x_t + \mathcal{Q}_{t+1}(x_t, \xi_t)$$

Where

$$\mathcal{Q}_{t+1}(x_t, \xi_t) \coloneqq \begin{cases} \mathbb{E}_{\xi_{t+1}|\xi_t}[Q_{t+1}(x_t, \xi_{t+1})], & t = T-1, \dots, 1 \\ 0, & t = T \end{cases}$$

We refer to $\mathcal{Q}_{t+1}$ as the recourse function (sometimes called future value or cost-to-go function). Notice that $\mathcal{Q}_{t+1}(\cdot, \xi_t)$ is convex for fixed ξ_t, but the same is not necessarily true for $\mathcal{Q}_{t+1}(\cdot, \cdot)$.

Let $p_{\xi_t}(\cdot\,|\xi_{t-1})$ be the conditional density function of ξ_t given ξ_{t-1}. Thus,

$$\mathbb{E}_{\xi_{t+1}|\xi_t}[Q_{t+1}(x_t, \xi_{t+1})] \coloneqq \int_{\mathcal{D}_{t+1}(\xi_t)} Q_{t+1}(x_t, \xi_{t+1}) p_{\xi_{t+1}}(\xi_{t+1}|\xi_t) d\xi_{t+1}$$

Hereafter, we extend with zero all density functions outside its corresponding support. Besides, whenever we evaluate a conditional density function at a point where it is not defined, we take it as zero. This happens when the density function of the conditioning variable is zero at the evaluated point.

2.2 Importance sampling stochastic dual dynamic programming

The proposed approach works approximating a reformulation of the real problem (P), we call this problem $(\tilde{P})$, which is defined as follows.

Let $\tilde{\xi}_t$ be a stagewise-independent random vector with the same structure as ξ_t (i.e., their components represent the same parameters). Let $p_{\tilde{\xi}_t}(\cdot)$ be its density function. We consider

$$w_t(\cdot\,|\tilde{\xi}_{t-1}) \coloneqq \frac{p_{\xi_t}(\cdot\,|\xi_{t-1})}{p_{\tilde{\xi}_t}(\cdot)}$$

Which is the importance sampling weight (Glynn & Iglehart, 1989; Tokdar & Kass, 2010).

Let $(\tilde{P})$ be the problem

$$\min_{A_1x_1=b_1} c_1x_1 + \mathbb{E}_{\tilde{\xi}_2}\Big[w_2(\tilde{\xi}_2|\tilde{\xi}_1) \min_{B_2x_1+A_2x_2=b_2} c_2x_2 + \mathbb{E}_{\tilde{\xi}_3}\Big[\dots + \mathbb{E}_{\tilde{\xi}_T}\Big[w_T(\tilde{\xi}_T|\tilde{\xi}_{T-1}) \min_{B_Tx_{T-1}+A_Tx_T=b_T} c_Tx_T\Big]\Big]\Big]$$

Which is similar to (P), but is solved for $\tilde{\xi}_t$ instead of ξ_t, and we also introduce a factor $w_t(\cdot\,|\tilde{\xi}_{t-1})$ at each stage. In contrast to (P), the conditioning variable in the expectation at each stage is dropped since $\tilde{\xi}_t$ is assumed stagewise-independent. Problem $(\tilde{P})$ may be decomposed as

$$\tilde{Q}_t(x_{t-1},\tilde{\xi}_t) = \min_{B_tx_{t-1}+A_tx_t=b_t} c_tx_t + \tilde{\mathcal{Q}}_{t+1}(x_t,\tilde{\xi}_t)$$

$$\tilde{\mathcal{Q}}_{t+1}(x_t,\tilde{\xi}_t) \coloneqq \begin{cases} \mathbb{E}_{\tilde{\xi}_{t+1}}\big[w_{t+1}(\tilde{\xi}_{t+1}|\tilde{\xi}_t)\tilde{Q}_{t+1}(x_t,\tilde{\xi}_{t+1})\big], & t = T-1,\dots,1 \\ 0, & t = T \end{cases}$$

The presence of w_{t+1}, makes the recourse function $\tilde{\mathcal{Q}}_{t+1}$ to explicitly depend on $\tilde{\xi}_t$, although $\tilde{\xi}_{t+1}$ is stagewise-independent.

As we will see next, the purpose of w_t is to compensate the effect of solving problem (P) under the random vector $\tilde{\xi}_t$. This is the main idea behind Importance Sampling; approximate the solution of a given stochastic problem using a different distribution. To achieve this correctly, a strict condition must be met.

Proposition 1: Assume that $p_{\tilde{\xi}_t}(\xi_t) > 0$ whenever $p_{\xi_t}(\xi_t|\xi_{t-1})Q_t(x_{t-1},\xi_t) \neq 0$. Then,

$$\mathcal{Q}_{t+1}(x_t,\xi_t) \equiv \tilde{\mathcal{Q}}_{t+1}(x_t,\xi_t), \qquad t = 1,\dots,T$$

Proof: By backward induction. Base case $(t = T)$: $\mathcal{Q}_{T+1}(x_T,\xi_T) \equiv \tilde{\mathcal{Q}}_{T+1}(x_T,\xi_T) \equiv 0$.

Induction step: first notice that if $\mathcal{Q}_{t+2}(x_{t+1},\xi_{t+1}) \equiv \tilde{\mathcal{Q}}_{t+2}(x_{t+1},\xi_{t+1})$ then $Q_{t+1}(x_t,\xi_{t+1}) \equiv \tilde{Q}_{t+1}(x_t,\xi_{t+1})$. Using this

$$\begin{aligned}
\tilde{\mathcal{Q}}_{t+1}(x_t,\xi_t) &= \mathbb{E}_{\tilde{\xi}_{t+1}}\big[w_{t+1}(\tilde{\xi}_{t+1}|\xi_t)\tilde{Q}_{t+1}(x_t,\tilde{\xi}_{t+1})\big] \\
&= \mathbb{E}_{\tilde{\xi}_{t+1}}\big[w_{t+1}(\tilde{\xi}_{t+1}|\xi_t)Q_{t+1}(x_t,\tilde{\xi}_{t+1})\big] \\
&= \int_{\widetilde{\mathcal{D}}_{t+1}} w_{t+1}(\tilde{\xi}_{t+1}|\xi_t)Q_{t+1}(x_t,\tilde{\xi}_{t+1})p_{\tilde{\xi}_{t+1}}(\tilde{\xi}_{t+1})d\tilde{\xi}_{t+1} \\
&= \int_{\widetilde{\mathcal{D}}_{t+1}} \frac{p_{\xi_{t+1}}(\tilde{\xi}_{t+1}|\xi_t)}{p_{\tilde{\xi}_{t+1}}(\tilde{\xi}_{t+1})} Q_{t+1}(x_t,\tilde{\xi}_{t+1})p_{\tilde{\xi}_{t+1}}(\tilde{\xi}_{t+1})d\tilde{\xi}_{t+1} \\
&= \int_{\mathcal{D}_{t+1}(\xi_t)} p_{\xi_{t+1}}(\tilde{\xi}_{t+1}|\xi_t)Q_{t+1}(x_t,\tilde{\xi}_{t+1})d\tilde{\xi}_{t+1} \\
&= \mathbb{E}_{\xi_{t+1}|\xi_t}[Q_{t+1}(x_t,\xi_{t+1})] \\
&= \mathcal{Q}_{t+1}(x_t,\xi_t)
\end{aligned}$$

Notice that the assumption we made allows us to replace $\widetilde{\mathcal{D}}_{t+1}$ by $\mathcal{D}_{t+1}(\xi_t)$ in the last integral. We conclude the demonstration. ■

Specifically, from Proposition 1 we have that $\mathcal{Q}_2(x_1, \xi_1) \equiv \tilde{\mathcal{Q}}_2(x_1, \xi_1)$. Therefore, both (P) and $(\tilde{P})$ have same first stage problems and have the same optimal value (recall that ξ_1 is deterministic). Consequently, any known bound for one problem is valid for the other.

To satisfy Proposition 1 assumption, we may take $p_{\tilde{\xi}_t}(\cdot)$ as the marginal distribution of ξ_t. However, other distributions may work as well.

Notice that we could move $w_t(\tilde{\xi}_t|\tilde{\xi}_{t-1})$ inside the *min* operator at each stage and the solution remains the same, since it is just a constant factor for the optimization problem. Doing that we can express $(\tilde{P})$ in the same fashion as (P), whose structure has been widely studied in terms of solution methods and convergence (Shapiro, 2003).

In practice, solving even an approximation of (P) could be quite demanding (if possible); the corresponding Sample Average Approximation tree, which needs to be constructed by conditional sampling (Shapiro, 2003), grows exponentially fast with the number of stages (given fineness of discretization). Under an outer linearization decomposition approach (like L-shaped method or SDDP) it means that every node of the tree must store its own set of cuts.

On the other hand, the SAA tree of $(\tilde{P})$ may be constructed by independent sampling at each stage (since $\tilde{\xi}_t$ is stagewise-independent); and all the nodes of a given stage may have the same descendents. Recall that $\tilde{\mathcal{Q}}_t$ explicitly depends on $\tilde{\xi}_{t-1}$, therefore a different set of cuts must be stored for each $\tilde{\xi}_{t-1}$. Fortunately, this number does not depend on the number of stages but only on the fineness of discretization, which keep things tractable.

2.3 Handling a special case of uncertainty

It may happen in problem (P) that a portion of the random vector does not participate conditioning the next-stage expectation. In that case it is convenient (in algorithmic terms) to handle both portions explicitly. Notice that the above results remain valid.

We denote by η_t and ω_t the portions that do and do not condition the expectation at t+1, respectively, such that

$$\xi_t = (\eta_t, \omega_t)$$

We say that η_t follows a Markovian process, and ω_t is a stagewise-independent process. Nevertheless, we admit ω_t to depend on the Markovian process but not to influence it.

Under this scheme, the expectation operator at each stage turns

$$\mathbb{E}_{\xi_t|\xi_{t-1}}[\cdot] = \mathbb{E}_{\eta_t,\omega_t|\eta_{t-1}}[\cdot]$$

Which, by virtue of the Law of Total Expectation, we accommodate as

$$\mathbb{E}_{\eta_t,\omega_t|\eta_{t-1}}[\cdot] = \mathbb{E}_{\eta_t|\eta_{t-1}}\left[\mathbb{E}_{\omega_t|\eta_t,\eta_{t-1}}[\cdot]\right]$$

And, we let $p_{\eta_t}(\cdot\,|\eta_{t-1})$ and $p_{\omega_t}(\cdot\,|\eta_t, \eta_{t-1})$ be the conditional density functions associated to each of the right-hand side conditional expectations.

On the other hand, we express $\tilde{\xi}_t = (\tilde{\eta}_t, \tilde{\omega}_t)$, with $\tilde{\eta}_t$ and $\tilde{\omega}_t$ stagewise-independent random vectors with density functions $p_{\tilde{\eta}_t}(\cdot)$ and $p_{\tilde{\omega}_t}(\cdot\,|\tilde{\eta}_t)$, respectively, such that

$$p_{\tilde{\xi}_t}\big((\tilde{\eta}_t, \tilde{\omega}_t)\big) = p_{\tilde{\eta}_t}(\tilde{\eta}_t) p_{\tilde{\omega}_t}(\tilde{\omega}_t|\tilde{\eta}_t)$$

And we redefine w_t accordingly, as

$$\begin{aligned} w_t\big(\tilde{\xi}_t|\tilde{\xi}_{t-1}\big) &= w_t(\tilde{\eta}_t, \tilde{\omega}_t|\tilde{\eta}_{t-1}) \\ &= \frac{p_{\eta_t}(\tilde{\eta}_t|\tilde{\eta}_{t-1})}{p_{\tilde{\eta}_t}(\tilde{\eta}_t)} \frac{p_{\omega_t}(\tilde{\omega}_t|\tilde{\eta}_t, \tilde{\eta}_{t-1})}{p_{\tilde{\omega}_t}(\tilde{\omega}_t|\tilde{\eta}_t)} \end{aligned}$$

Where we denote

$$u_t(\tilde{\eta}_t|\tilde{\eta}_{t-1}) \coloneqq \frac{p_{\eta_t}(\tilde{\eta}_t|\tilde{\eta}_{t-1})}{p_{\tilde{\eta}_t}(\tilde{\eta}_t)}$$

$$v_t(\tilde{\omega}_t|\tilde{\eta}_t, \tilde{\eta}_{t-1}) \coloneqq \frac{p_{\omega_t}(\tilde{\omega}_t|\tilde{\eta}_t, \tilde{\eta}_{t-1})}{p_{\tilde{\omega}_t}(\tilde{\omega}_t|\tilde{\eta}_t)}$$

We proceed defining problem $(\tilde{P})$ the same as before, but accommodating the conditional expectations as mentioned

$$\begin{aligned} \min_{A_1 x_1 = b_1} c_1 x_1 + \mathbb{E}_{\tilde{\eta}_2}\Bigg[\mathbb{E}_{\tilde{\omega}_2|\tilde{\eta}_2}\Bigg[u_2(\tilde{\eta}_2|\tilde{\eta}_1) v_2(\tilde{\omega}_2|\tilde{\eta}_2, \tilde{\eta}_1) \min_{B_2 x_1 + A_2 x_2 = b_2} c_2 x_2 \\ + \mathbb{E}_{\tilde{\eta}_3}\Bigg[\mathbb{E}_{\tilde{\omega}_3|\tilde{\eta}_3}\Bigg[\dots \\ + \mathbb{E}_{\tilde{\eta}_T}\Bigg[\mathbb{E}_{\tilde{\omega}_T|\tilde{\eta}_T}\Big[u_T(\tilde{\eta}_T|\tilde{\eta}_{T-1}) v_T(\tilde{\omega}_T|\tilde{\eta}_T, \tilde{\eta}_{T-1}) \min_{B_T x_{T-1} + A_T x_T = b_T} c_T x_T \Big]\Bigg]\Bigg]\Bigg]\Bigg] \end{aligned}$$

Which may be decomposed as

$$\tilde{Q}_t(x_{t-1}, \tilde{\eta}_t, \tilde{\omega}_t) = \min_{B_t x_{t-1} + A_t x_t = b_t} c_t x_t + \tilde{\mathcal{Q}}_{t+1}(x_t, \tilde{\eta}_t)$$

Where

$$\tilde{\mathcal{Q}}_{t+1}(x_t, \tilde{\eta}_t) \coloneqq \begin{cases} \mathbb{E}_{\tilde{\eta}_{t+1}}\big[u_{t+1}(\tilde{\eta}_{t+1}|\tilde{\eta}_t)\, \tilde{\mathcal{G}}_{t+1}(x_t, \tilde{\eta}_{t+1}, \tilde{\eta}_t)\big], & t = T-1, \dots, 1 \\ 0, & t = T \end{cases}$$

$$\tilde{\mathcal{G}}_{t+1}(x_t, \tilde{\eta}_{t+1}, \tilde{\eta}_t) \coloneqq \mathbb{E}_{\tilde{\omega}_{t+1}|\tilde{\eta}_{t+1}}\big[v_{t+1}(\tilde{\omega}_{t+1}|\tilde{\eta}_{t+1}, \tilde{\eta}_t)\, \tilde{Q}_{t+1}(x_t, \tilde{\eta}_{t+1}, \tilde{\omega}_{t+1})\big], \qquad t = T-1, \dots, 1$$

This explicit handling of both random portions will be useful when evaluating the solution of the approximate counterpart of $(\tilde{P})$.

3 Solution algorithm

In this section we derive the approximate problem $(\tilde{P})$ and we show an algorithm to solve it. This algorithm has been used in previous works, such as Philpott & Matos (2012) and Lohndorf & Shapiro (2019), to solve the approximate problem under the MC-SDDP approach. In this sense, the novelty of our work is not to propose a new solution algorithm but rather a new approximation scheme with convergence guarantees.

3.1 Approximate problem

To approximate problem $(\tilde{P})$ we need to discretize every expectation. As the uncertainty is stagewise-independent we derive a recombining scenario tree. To do so, at each stage, we generate samples $\mathrm{H}_t = \{\tilde{\eta}_t^i\}_{i=1}^M$ taken from $p_{\tilde{\eta}_t}(\cdot)$, that we call Markov states, and for each of them we generate samples $\Omega_t^i = \{\tilde{\omega}_t^{ij}\}_{j=1}^N$ from $p_{\tilde{\omega}_t}(\cdot\,|\tilde{\eta}_t^i)$. Next, we compute $u_t^{ik} \coloneqq u_t(\tilde{\eta}_t^i|\tilde{\eta}_{t-1}^k)$ and $v_t^{ijk} \coloneqq v_t(\tilde{\omega}_t^j|\tilde{\eta}_t^i, \tilde{\eta}_{t-1}^k)$. This way, the approximate problem $(\tilde{P})$ is decomposed into subproblems as follows

$$\tilde{Q}_t(x_{t-1}, \tilde{\eta}_t^k, \tilde{\omega}_t^{kl}) = \min_{B_t x_{t-1} + A_t x_t = b_t} c_t x_t + \tilde{\mathcal{Q}}_{t+1}(x_t, \tilde{\eta}_t^k)$$

For $k = 1, \dots, M$, $l = 1, \dots, N$ and $t = 2, \dots, T$. Where

$$\tilde{\mathcal{Q}}_{t+1}(x_t, \tilde{\eta}_t^k) \coloneqq \begin{cases} \dfrac{1}{M} \displaystyle\sum_i u_{t+1}^{ik}\, \tilde{\mathcal{G}}_{t+1}(x_t, \tilde{\eta}_{t+1}^i, \tilde{\eta}_t^k), & t = T-1, \dots, 1 \\ 0, & t = T \end{cases}$$

$$\tilde{\mathcal{G}}_{t+1}(x_t, \tilde{\eta}_{t+1}^i, \tilde{\eta}_t^k) \coloneqq \frac{1}{N} \sum_j v_{t+1}^{ijk}\, \tilde{Q}_{t+1}(x_t, \tilde{\eta}_{t+1}^i, \tilde{\omega}_{t+1}^{ij}), \qquad t = T-1, \dots, 1$$

Notice that u_t/M and v_t/N are not probabilities, instead they are weighting factors that do not necessarily sum up to one (although it can be shown that they tend to, as the sample size increases). However, the solution algorithm proceeds the same way as if they were probabilities. Therefore, from this point on, the approximate problem is solved as in the MC-SDDP approach.

3.2 Solution algorithm

We implement a SDDP-type solution algorithm, and as such, $\tilde{\mathcal{Q}}_{t+1}$ is progressively (outer) approximated by means of Benders cuts. As we mentioned before, since the uncertainty process is not stagewise-independent we need to store different sets of cuts for each Markov state $\tilde{\eta}_t^k$. On the other hand, $\tilde{\mathcal{Q}}_{t+1}$ does not depend on $\tilde{\omega}_t$, hence the cuts are valid for any realization of it.

It is interesting to note that a multi-cut approach (Birge, 1988) might be implemented, which explicitly approximates $\tilde{\mathcal{G}}_{t+1}$. However, we will only deal with the single-cut approach in this work.

The following pseudo-code describes the solution algorithm. To simplify the cut generation procedure, at each stage we introduce local copies (y_t) of the state variables

(x_{t-1}), and we replace each occurrence of the later with the former. We also introduce linking constraints ($y_t = x_{t-1}$), which become the state equations. We denote by π_t the duals of this constraints, and z_t the optimal value of the subproblem.

Markov chain SDDP

0 Initialize:
completed iterations: $it = 0$
sets of cuts: $\mathcal{C}_t^i = \emptyset,\ \forall t = 1, \dots, T-1,\ \forall i = 1, \dots, M$

1 Forward pass:
solve first stage subproblem → $z_1, \bar{x}_1$
$k = 1$
for $t = 2, .., T$
  draw a Markov state $\tilde{\eta}_t^i$ from H_t, with probabilities proportional to u_t^{ik}, given k
  draw a sample $\tilde{\omega}_t^{ij}$ from Ω_t^i, with probabilities proportional to v_t^{ijk}, given (k, i)
  update subproblem at stage t with $\left(\bar{x}_{t-1}, \tilde{\eta}_t^i, \tilde{\omega}_t^{ij}\right)$ and solve → $z_t, \bar{x}_t$
  (set of cuts to use depend on index i)
  $k = i$
end
$lower\ bound = z_1;\ total\ cost = \sum_t z_t$

2 Backward pass:
for $t = T, \dots, 2$
  for $\tilde{\eta}_t^i \in \mathrm{H}_t$
    for $\tilde{\omega}_t^{ij} \in \Omega_t^i$
      update subproblem at stage t with $\left(\bar{x}_{t-1}, \tilde{\eta}_t^i, \tilde{\omega}_t^{ij}\right)$ and solve → z_t^{ij}, π_t^{ij}
      (set of cuts to use depend on index i)
    end
  end
  for $\tilde{\eta}_{t-1}^k \in \mathrm{H}_{t-1}$
    $\bar{z}_t^k = \frac{1}{M}\sum_i u_t^{ik}\left(\frac{1}{N}\sum_j v_t^{ijk}\, z_t^{ij}\right)$
    $\bar{\pi}_t^k = \frac{1}{M}\sum_i u_t^{ik}\left(\frac{1}{N}\sum_j v_t^{ijk}\, \pi_t^{ij}\right)$
    append cut $\tilde{Q}_t\left(x_{t-1}, \tilde{\eta}_{t-1}^k\right) \geq \bar{z}_t^k + \bar{\pi}_t^k (x_{t-1} - \bar{x}_{t-1})$ to set $\mathcal{C}_{t-1}^k$
  end
end
$it = it + 1$

3 Check stopping criterion:
if stopping criterion is not met:
  go to 1
end

The proof of convergence for this algorithm relies on the finiteness of samples across stages. This is similar for a broad class of sampling algorithms, refer to Philpott & Guan (2008) for details. Despite the above, convergence may take too long in practice. In that sense, we did not specify the stopping criterion, but many optional setups may be

considered. However, is out of the scope of this work to analyze them. Just to mention a few: time limit, maximum iterations, lower bound stagnation, and convergence of bounds (see below).

For simplicity we considered a single sample-path in each forward pass. However, this can be easily extended to many. In that case, we can derive a confidence interval for the total cost (upper bound) and evaluate convergence as a hypothesis test (Homem-de-Mello et al., 2011; Matos et al., 2017). Besides, alternative path-sampling strategies may be employed, thus enhancing the performance of the algorithm (e.g., convergence, speed-up, and quality of solutions). Such modifications may be easily implemented without altering the main structure of the algorithm. Both topics above remain active fields of research (Philpott & Guan, 2008; Homem-de-Mello et al., 2011; Shapiro, 2011; de Matos et al., 2015; de Matos et al., 2017).

3.3 Bounds on the true problem

In the previous subsection, we briefly refer to "convergence of bounds". However, this does not necessarily imply convergence to the optimal value of the true problem. For instance, coarse discretization in the approximate problem may lead to fast SDDP-convergence, nonetheless the obtained solution may be far from optimal for the true problem.

To derive a statistical lower bound for the true problem (P), we may solve many different replications of the approximate of $(\tilde{P})$, compute a confidence interval using the resulting collection of lower bounds and take its lower end.

On the other hand, for each replication, the approximations obtained for the recourse functions $\{\mathcal{C}_t^i\}_{t,i}$ can be used to derive an implementable policy for the true problem (P) as follows: for any state $(x_{t-1}, \eta_t, \omega_t)$ find the closest Markov state $\tilde{\eta}_t^i \in \mathrm{H}_t$ (we consider the Mahalanobis distance; a unitless and scale-invariant measure, that considers the correlations in the feature space) and use the corresponding set of cuts $\mathcal{C}_t^i$ to approximate $\mathcal{Q}_{t+1}(x_t, \eta_t)$; that is identical to $\tilde{\mathcal{Q}}_{t+1}$, as we saw in Proposition 1. This way, we can evaluate the policy across many different paths drawn from the true process, compute a confidence interval for the total cost and take its upper end. Finally, the statistical upper bound is defined as the smallest upper end across replications; and the corresponding policy is established as the best-found solution.

4 Numerical experiments

In this section we demonstrate numerically the effect of different sampling approaches in the quality of solutions and evaluate the convergence of bounds for the true problem. We consider a multi-stage inventory problem with stochastic demand for our experiments.

First, we start by looking at the two-stage version of the problem, which corresponds to the well-known Newsvendor problem, whose solution is available analytically. We compare the solutions (and their performance) obtained from different approximations of the demand distribution. These approximate models are motivated by the mathematical assumptions necessary to efficiently extend to the multi-stage setting.

Next, we apply the proposed approach (IS-SDDP) to the multi-stage version of the problem. We evaluate the convergence of bounds for the true problem, and we compare the obtained solution against the TS approach.

4.1 Inventory problem

We consider the model described by Shapiro et al. (2014). Under a planning horizon of length T, at the beginning of each stage the inventory is observed, and an order is placed to raise its level to a desired point, incurring in an ordering cost. The demand is then revealed, and the planner incurs backordering or holding costs whenever the inventory level does not exactly meet the demand. The inventory is observed again, and the cycle repeats all over the planning horizon. The following costs are assumed:

Cost	Value (per unit)
Ordering	1.0
Backordering	9.0
Holding	3.0

4.2 Demand distribution

First, we present the (true) model governing the dynamics of the demand. Next, we show some approximate models that mimic the true dynamics. These are motivated by the stagewise-independency assumption necessary for SDDP-type algorithms.

4.2.1 Mixture model

The product demand distributes according to a gaussian mixture model of two components, where the mixture-weight distributes uniformly and is revealed one stage before the demand does. Each gaussian is truncated at zero to guarantee non-negative demand.

$$d_t \sim (1 - w_{t-1})\, TN(\mu_1, \sigma_1) + w_{t-1}\, TN(\mu_2, \sigma_2)$$

$$w_{t-1} \sim U(0,1)$$

Notice that the first is a sum of densities, and not of random variables. Therefore, the demand has a bi-modal distribution.

The stochastic process is as follows: at each stage are revealed both the demand d_t and the mixture-weight w_t, which in turn determines the next-stage demand d_{t+1} distribution. Figure 1 shows two (truncated) gaussians densities and the corresponding demand distribution functions for some values of the mixture-weight ranging between 0 and 1. We also refer to this model as the "real" model.

This model represents a demand with a (continuous) changing regime, which may arise in situations where consumers preferences depend on some external factor (e.g., weather or fashion trends). Going back to the origins of SDDP, similar phenomena may be observed in hydro-thermal scheduling if water-inflows follow a Hidden Markov model depending on some climate index; see Philpott & Matos (2012) for a related example.

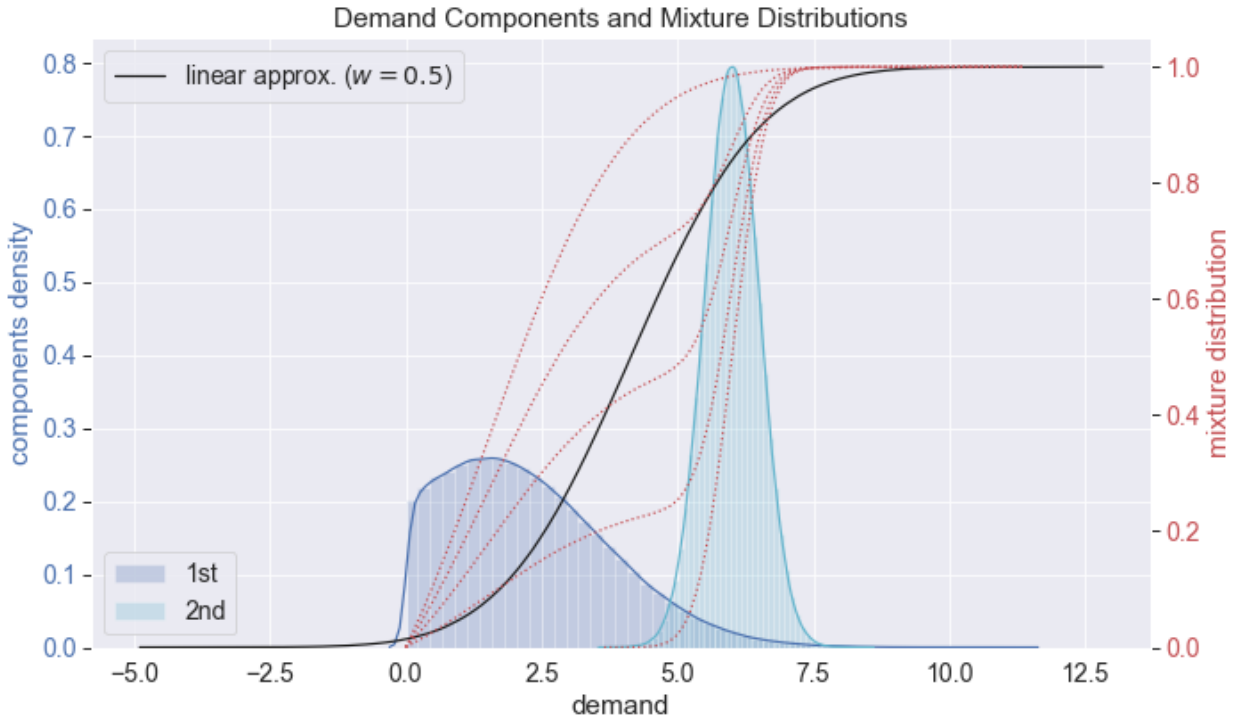

Figure 1: Mixture-demand components and linear approximation.

4.2.2 Linear model

The model described above does not meet the stagewise-independent assumption nor can be linearly reformulated exactly, because of the non-linear stagewise-dependency of d_t on w_{t-1}. However, this relationship can be linearly approximated as follows:

$$d_t = a\, w_{t-1} + b + \xi_t$$

$$\xi_t \sim N(0, \sigma)$$

Where a, b and σ are calibrated statistically using samples (d_t, w_{t-1}) drawn from the real model.

4.2.3 Marginal model

A different approach might be to drop the stagewise-dependency and rely on a purely stagewise-independent model. To do that we can marginalize-out the mixture weight in the demand distribution as:

$$f(d_t) = \int f(d_t|w_{t-1}) f(w_{t-1}) dw_{t-1}$$

Where f denotes the corresponding density function. In this way, the resulting model is:

$$d_t \sim \frac{1}{2}\big(\, TN(\mu_1, \sigma_1) + TN(\mu_2, \sigma_2)\big)$$

Which is equivalent to the real model with a fixed value of 0.5 for the mixture-weight.

4.2.4 Importance-sampling model

In the context of IS-SDDP we use samples drawn from the marginal model but re-weighted according to the density of the real model (which incorporates the effect of the observed mixture-weight).

We consider normalized importance-sampling (i.e., ordinary weights are normalized to sum one). This introduces some bias (although it converges asymptotically to zero) compared to ordinary importance-sampling, but the variance is generally reduced. In this case, the resulting weights may be interpreted as probabilities.

4.3 Two-stage setting

The optimal solution for the two-stage setting is obtained as (Shapiro et al., 2014):

$$\bar{x} = H^{-1}(\kappa), \qquad \kappa = \frac{b - c}{b + h}$$

Where b, c and h denote the backordering, ordering, and holding costs; and $H(\cdot)$ denotes the demand cdf. We refer to κ as the optimal quantile and $\bar{x}$ the optimal solution. This solution holds for both the continuous and discrete cases.

In Figure 2 we compare the solutions obtained by the above models under different values for the mixture-weight. To do so, we rely on the empirical (discrete) cdf obtained by extensive sampling in each case. Then, the performance of solutions is evaluated under the true model.

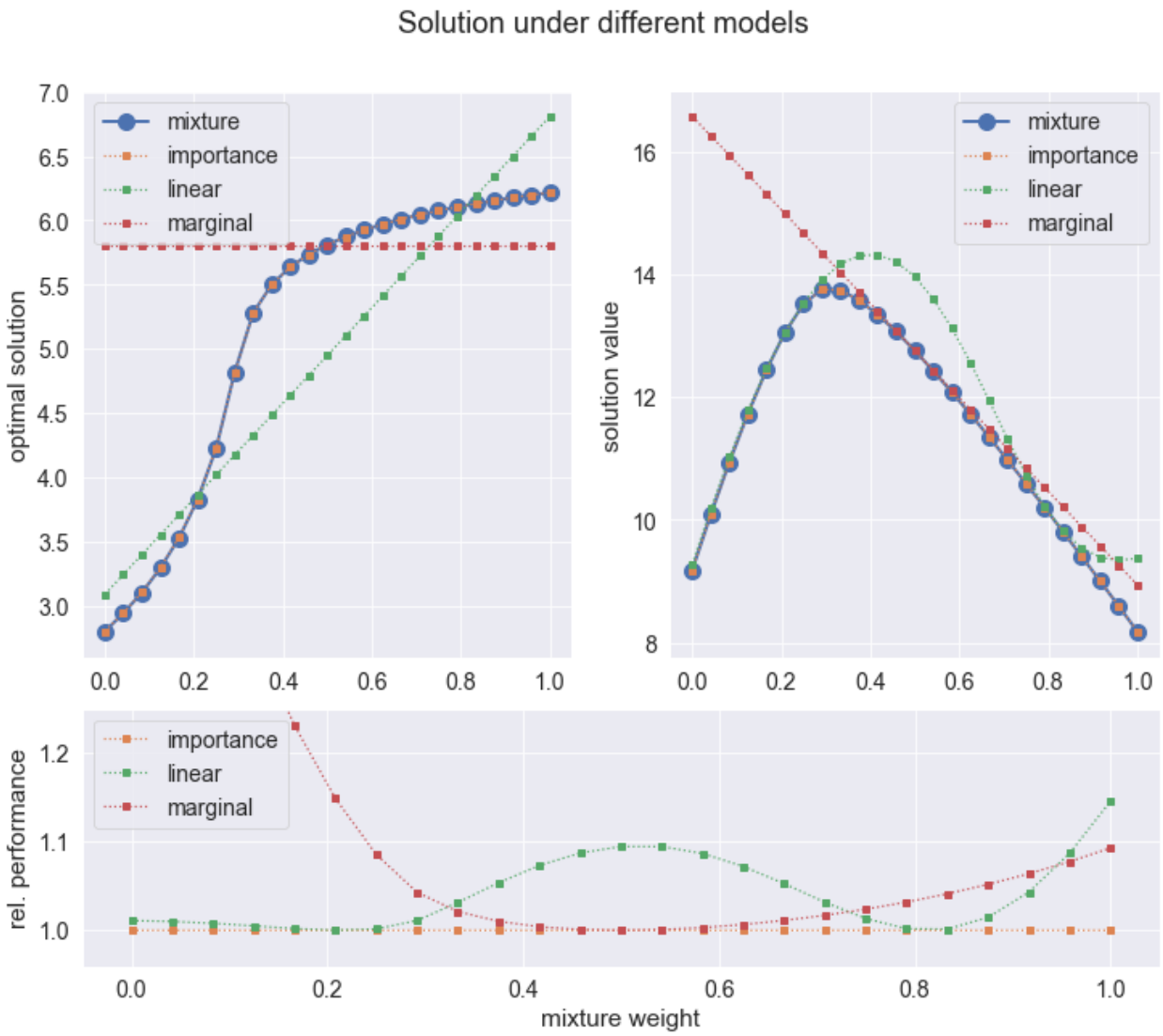

Figure 2: Two-stage setting solution.

We notice the importance-sampling model sticks to the correct solution (recall the mixture model is the true model) in every case, whereas the linear and marginal models present a suboptimal behavior which varies depending on the mixture-weight. As expected, the same occurs for the solution's values, where the relative performance can be worse than 10 or 20% of the optimal values in some cases. This illustrates the (negative) effect of approximate models to represent the demand distribution.

4.4 Multi-stage setting

Now we move to the multi-stage case of the Inventory problem and perform similar analysis, however we only focus on the linear and importance sampling demand models since they allow us to incorporate stagewise-dependency efficiently in this case. The former model leads to the TS-SDDP, while the latter leads to IS-SDDP. We compare both approaches in terms of quality of solutions and evaluate the convergence to the optimal value.

Contrary to the two-stage setting, the exact solution in this case is not known, hence we solve in a SAA-fashion through replications-analysis (see Section 3.3). Theoretically, only IS-SDDP correctly approximates the true problem, therefore we use its corresponding statistical lower bound to evaluate convergence to optimality.

We solve for a 4-stage horizon and consider 5 replications for each approach. Notice under IS-SDDP, the uncertainty has the structure described in Section 2.3; we consider $M = 10$ and $N = 100$. Under TS-SDDP, the uncertainty is purely stagewise-independent; we consider $N = 1{,}000$. We run each SDDP instance for a maximum of 200 iterations, evaluating (in-sample) every 50 iterations under 5,000 scenarios. The purpose of evaluating in between the optimization process is two-fold: (i) keep track of the value of the best solution found so far since the process of deriving cuts does not produce solutions of decreasing value necessarily, and (ii) evaluate SDDP convergence gap, which we set to 0.1%. When optimization ends, the best solution is evaluated under 10,000 (out-of-sample) scenarios.

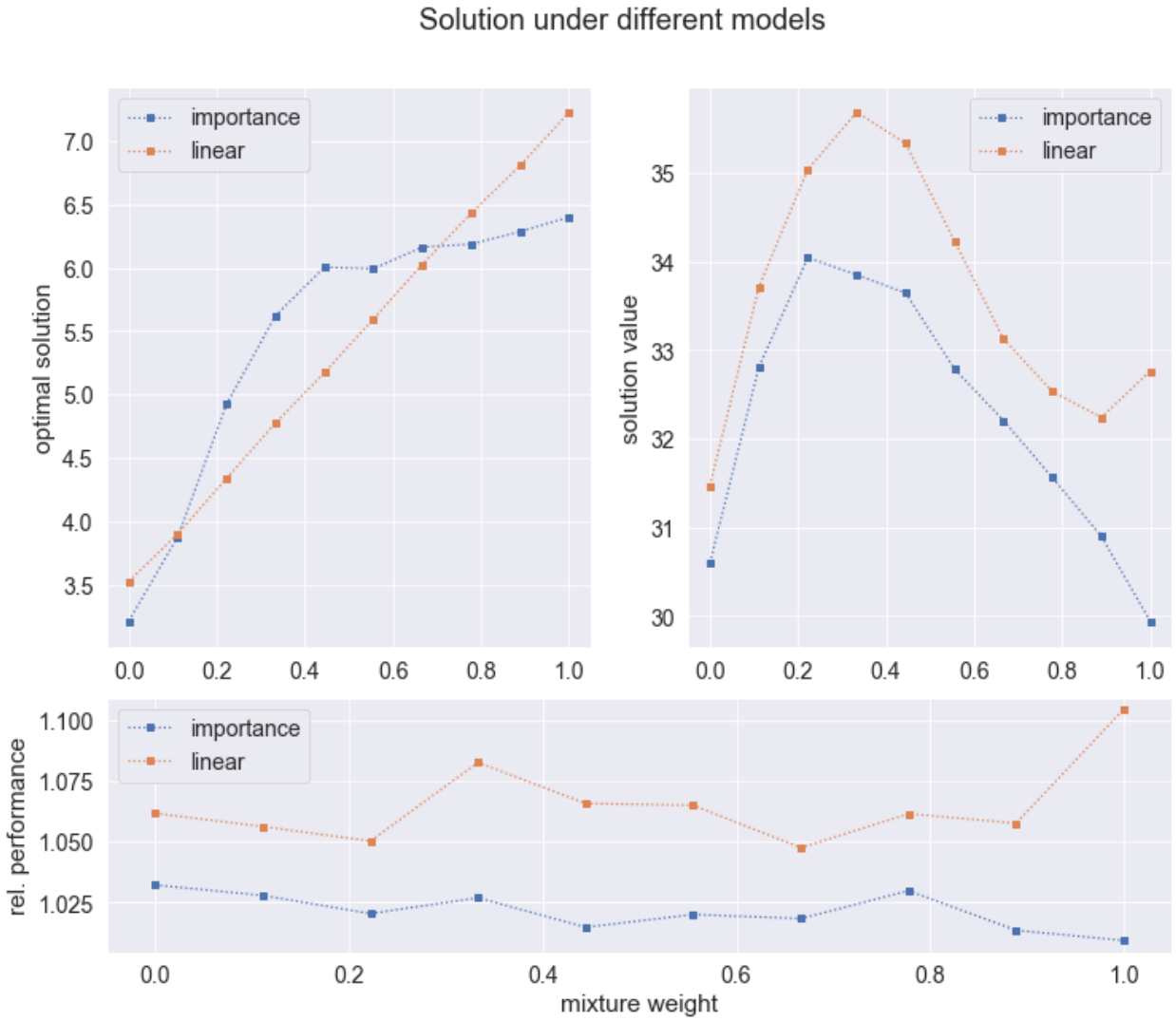

Figure 3: Multi-stage setting solution. ("importance" refers to IS-SDDP, "linear" refers to TS-SDDP)

Figure 3 shows the results under both SDDP approaches. As in the two-stage setting, the problem is solved for different initial values for the mixture-weight; each run considers the above configuration. We compare the first stage optimal solution reported by both approaches and the value of the corresponding policy when evaluated under the true demand process on the planification horizon. It is observed that IS-SDDP consistently outperforms TS-SDDP in terms of solution value. The relative performance (to the statistical lower bound) ranges between 5%-10% for TS-SDDP, while it is less than 3% for IS-SDDP.

5 Discussion

We presented IS-SDDP, an alternative SDDP approach based on Importance Sampling to incorporate stagewise-dependence for a broad range of stochastic models compared to TS-SDDP. It makes convergence guarantees with respect to the true problem. Moreover, it provides a statistical lower bound on the true problem optimal value, which is extremely useful to determine quality of solutions since multi-stage programs are known to be solvable only approximately in practice.

We tested the underlying sampling scheme on a two-stage problem with theoretically known solution. We observed that it satisfactorily approaches the optimum, whereas other sampling strategies used in TS-SDDP does not.

Under the multi-stage setting we observed similar behavior, which is expected since IS-SDDP correctly approximates the true problem. However, care must be taken in practice, since coarse levels of discretization may lead to poor results and misleading comparisons despite theoretical guarantees, since these apply for increasing levels of discretization.

Future work should evaluate the quality of solutions and convergence as a function of discretization level on a broad range of problems.

6 Acknowledgement

We want to thank to CONICYT Chile for the financial support in this work. (CONICYT-PFCHA/MagísterNacional/2017-22170944)

7 References

Pereira, M. V., & Pinto, L. M. (1991). Multi-stage stochastic optimization applied to energy planning. Mathematical programming, 52(1-3), 359-375.

Pereira, M. V. F., & Pinto, L. M. V. G. (1985). Stochastic optimization of a multireservoir hydroelectric system: a decomposition approach. Water resources research, 21(6), 779-792.

Infanger, G., & Morton, D. P. (1996). Cut sharing for multistage stochastic linear programs with interstage dependency. Mathematical Programming, 75(2), 241-256.

De Queiroz, A. R., & Morton, D. P. (2013). Sharing cuts under aggregated forecasts when decomposing multi-stage stochastic programs. Operations Research Letters, 41(3), 311-316.

Homem-de-Mello, T., De Matos, V. L., & Finardi, E. C. (2011). Sampling strategies and stopping criteria for stochastic dual dynamic programming: a case study in long-term hydrothermal scheduling. Energy Systems, 2(1), 1-31.

De Matos, V. L., & Finardi, E. C. (2012). A computational study of a stochastic optimization model for long term hydrothermal scheduling. International Journal of Electrical Power & Energy Systems, 43(1), 1443-1452.

De Matos, V. L., Larroyd, P. V., & Finardi, E. C. (2014). Assessment of the Long-Term Hydrothermal Scheduling operation polices with alternative inflow modeling. In Power Systems Computation Conference (PSCC), 2014 (pp. 1-7). IEEE.

De Matos, V. L., Philpott, A. B., & Finardi, E. C. (2015). Improving the performance of stochastic dual dynamic programming. Journal of Computational and Applied Mathematics, 290, 196-208.

De Matos, V. L., Morton, D. P., & Finardi, E. C. (2017). Assessing policy quality in a multistage stochastic program for long-term hydrothermal scheduling. Annals of Operations Research, 253(2), 713-731.

Birge, J. R. (1985). Decomposition and partitioning methods for multistage stochastic linear programs. Operations research, 33(5), 989-1007.

Birge, J. R., & Louveaux, F. V. (1988). A multicut algorithm for two-stage stochastic linear programs. European Journal of Operational Research, 34(3), 384-392.

Löhndorf, N., & Shapiro, A. (2019). Modeling time-dependent randomness in stochastic dual dynamic programming. European Journal of Operational Research, 273(2), 650-661.

Philpott, A. B., & Guan, Z. (2008). On the convergence of stochastic dual dynamic programming and related methods. Operations Research Letters, 36(4), 450-455.

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

Shapiro, A. (2003). Inference of statistical bounds for multistage stochastic programming problems. Mathematical Methods of Operations Research, 58(1), 57-68.

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

Shapiro, A., Tekaya, W., da Costa, J. P., & Soares, M. P. (2013). Risk neutral and risk averse stochastic dual dynamic programming method. European journal of operational research, 224(2), 375-391.

Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2014). Lectures on stochastic programming: modeling and theory. Society for Industrial and Applied Mathematics.

Glynn, P. W., & Iglehart, D. L. (1989). Importance sampling for stochastic simulations. Management science, 35(11), 1367-1392.

Tokdar, S. T., & Kass, R. E. (2010). Importance sampling: a review. Wiley Interdisciplinary Reviews: Computational Statistics, 2(1), 54-60.