

## OPTIMIZATION SIMULATION: THE CASE OF MULTI-STAGE STOCHASTIC DECISION MODELS

Suvrajeet Sen

Data Driven Decisions Laboratory  
Integrated Systems Engineering  
The Ohio State University  
Columbus, OH 43210

Zhihong Zhou

SIE Department  
University of Arizona,  
Tucson, AZ 85721

### ABSTRACT

In this paper we present a new approach to solving multi-stage stochastic decision models in the presence of constraints. The models themselves are stochastic linear programs (SLP), but we presume that their deterministic equivalent problems are too large to be solved exactly. We seek an asymptotically optimum solution by simulating the stochastic decomposition (SD) algorithmic process, originally designed for two-stage SLPs. When SD is implemented in a time-staged manner the algorithm begins to take the flavor of a simulation leading to what we refer to as optimization simulation. Among its major advantages, it can work directly with sample paths, and this feature makes the new algorithm much easier to integrate within a simulation. We also overcome certain limitations such as a stage-wise independence assumption required by other sampling-based algorithms for multi-stage stochastic programming. Finally, we also discuss how these methods can be interpreted as close relatives of approximate dynamic programming.

### 1 INTRODUCTION

Solution of constrained stochastic decision models remains the holy grail of optimization. Despite decades of research on optimization models and algorithms, our ability to solve constrained stochastic optimization problems remains a continuing challenge. One might lay the blame on the well known “curse-of-dimensionality” associated with Dynamic Programming (DP, e.g. Powell 2007). However, the difficulties go beyond DP. Other approaches to stochastic optimization problems (e.g. Shapiro 2006) also present a rather bleak outlook for constrained stochastic optimization. One might even trace the difficulties to a result of Dyer and Stougie (2006) who show that the problem of optimizing a multi-dimensional expected value functional is inherently difficult because its decision counterpart is #P-complete, in the worst case. Despite these difficulties, practitioners and specialists in the art of modeling do not have the luxury of setting these problems aside; on a daily basis, decisions are made in the face of uncertainty while accommodating physical and financial constraints. Moreover, these systems are inherently dynamic, with the simplest models resulting in two-stage problems. We will focus on multi-stage problems, and in this setting, the demands on optimization algorithms rise to a whole new level. In this paper, we introduce the notion of optimization by simulating the algorithmic process, and hence we refer to this process as “Optimization Simulation.”

The natural question in these circumstances is “what differentiates *Optimization Simulation* from traditional *Simulation Optimization* ?” Whereas the latter refers to *optimizing a simulation model*, the former refers to *simulating an optimization algorithm*. What seems to be only a subtle distinction at first sight is really a significant conceptual difference: the latter (i.e. simulation optimization) operates on a simulation model, whereas, the former simulates an optimization algorithm, seeking a solution to an optimization model. Examples of simulation optimization problems are available at [www.simopt.org](http://www.simopt.org), whereas, an example of optimization simulation is the stochastic decomposition algorithm (SD, Hige and Sen

1991, 1994). The SD algorithm is an optimization simulation method that extends the scope of the L-shaped method (Van Slyke and Wets 1969, also known as Benders' decomposition, Benders 1962) to cases with continuous random variables, or even to discrete random variables in which the number of outcomes are too many to enumerate efficiently. A similar approach based on importance sampling has been proposed by Dantzig and Glynn (1991) and Infanger (1994). Another version of optimization simulation is the Abridged Benders' Decomposition (Donahue and Birge 2006) algorithm which derives from its parent algorithm, the Nested Benders' Decomposition algorithm, as well as dual decomposition method of Pinto and Pereira (1991). These algorithms are designed for multi-stage SLP, and for a general overview of SLP algorithms, we refer the reader to Birge and Louveaux (1997). In any event, the above sample-based methods for multi-stage problems work with a sampled sub-tree, and as shown by Linowski and Philpott (2005), and more recently by Shapiro (2010), these methods can provide asymptotic optimality (wpl) provided certain conditions are satisfied.

There are two handicaps associated with the above mentioned OS methods for multi-stage SP: a) most of the methods do not work with individual sample paths, and while the two-stage SD algorithm does operate on one sample path at a time, it is restricted to two stage SLP, b) the assumptions of independence severely limits the power that simulation can bring to bear on these models. The expressive capability of simulation and its ability to represent complicated random phenomena are some of the attractive features that would provide a rich class of optimization models in SLP. But the current algorithmic limitations are unsatisfying. The current paper is devoted to an algorithm that overcomes both these handicaps.

## 2 MULTI-STAGE STOCHASTIC DECISION MODELS

In our formulation, time will march ahead from now ( $t = 0$ ) to the end of horizon ( $t = T$ ), where  $T$  is a given positive integer. With this notation, there will be  $T + 1$  stages in the formulation, and thus a two-stage formulation will have stages indexed by  $t = 0$  and  $t = 1$ .

Let  $(\Omega, \mathcal{F}, \wp)$  denote a filtered probability space (i.e  $\mathcal{F}_t \subseteq \mathcal{F}$ , and for  $t = 1, \dots, T$   $\mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}$  for  $t_1 < t_2$ .) As usual,  $\Omega := \Omega_1 \times \dots \times \Omega_T$  ( $\Omega_t \subseteq \mathfrak{R}^{v_t}$  with  $v_t$  a positive integer) and, an outcome consisting of  $t$  elements of the process will be denoted  $\underline{\omega}_t := (\omega_1, \dots, \omega_t)$ , and the corresponding random variable will be denoted  $\underline{\tilde{\omega}}_t$ . The  $\sigma$ -algebras  $\mathcal{F}_t$  represent data available to the decision maker at time  $t$ . We will assume that  $\underline{\tilde{\omega}}_t$  has finite support for all  $t$ ,  $\mathcal{F}_t$  is finite, and so is the filtration  $\{\mathcal{F}_1, \dots, \mathcal{F}_T\}$ . Moreover,  $\mathcal{F}_t$  is generated by a finite partition  $\{\Theta_t^\ell\}$  of  $\Omega$ . The same is true for  $\mathcal{F}_{t+1}$  generated by  $\{\Theta_{t+1}^\ell\}$  with the property that the latter is finer than the former (i.e. for any set  $\bar{\Theta} \in \{\Theta_t^\ell\}$ ,  $\exists$  a collection of sets in  $\{\Theta_{t+1}^\ell\}$ , indexed by  $\mathbb{C}(\bar{\Theta})$ , say, such that  $\bar{\Theta} = \cup_{\ell \in \mathbb{C}(\bar{\Theta})} \Theta_{t+1}^\ell$ ). The relationships between the sets in the partition  $\{\Theta_t^\ell\}$  and their children  $\mathbb{C}(\{\Theta_t^\ell\})$  can be encoded in the form of a tree which is termed as a scenario tree in SP. A node in period  $t$  represents a subset of paths (such as  $\bar{\Theta}$ ) which have the same events in the first  $t$  periods, and record some new event in period  $t + 1$ . Since algorithms will work by using the tree as a road-map, it will be convenient to index nodes on the scenario tree by  $n$ , and denote the (unconditional) probability of reaching node  $n$  by  $p_n$ . We will denote any random child of node  $n$  by the notation  $\tilde{n}+$ , so that the outcomes  $n+ \in \mathbb{C}(n)$ . By the same token, arrival at node  $n$  of a scenario tree implies knowledge of the (unique) parent node which will be denoted by  $n-$ . Finally, letting  $p_{(n+|n)}$  denote the probability of reaching node  $n+$ , given that the process has arrived at node  $n$ , we can simulate the generation of a sample path through the scenario tree. If  $\underline{\tilde{\omega}}_t$  is a continuous stochastic process, then ensuring measurability of decisions requires greater care (Casey and Sen 2005).

As with the two-stage SD algorithm, the multi-stage SD method will work with sample paths, although in this case, we will sample from the scenario tree. The set of nodes of the scenario tree will be denoted  $N$ . In the following statement of the sequential decision model, we associate a time index  $t(n)$  with each node  $n$ . We will use the following convention in the model stated below.

- a) The root node is indexed by node 0.
- b) If node  $n$  belongs to the last stage, the expected value/recourse function of  $n +$  ( the future) is 0.
- c) Finally, we define composite state variables  $\underline{s}_n = (y_n, \underline{\omega}_n)$ . Note that the composite state  $\underline{s}_n$  used here traces the entire history of the data state  $\underline{\omega}_n$ . Thus by using the scenario tree, SP allows us to capture the history of the process. However, the state  $y_n$  is only defined at node  $n$ , and is data and decision dependent, prompting us to refer to them as the “3d” states.

Given a scenario tree, a multi-stage decision model may be stated as follows.

$$\text{Min } \{c_0(x_0) + E[h_{0+}(\underline{s}_{0+})]: x_0 \in X_0, y_1 \in Y_1(x_0, y_0, \omega_1) \text{ a.s.} \} \tag{0}$$

where  $h_n$  are defined recursively for  $n \geq 1$  as

$$h_n(\underline{s}_n) = d_n^T y_n + \text{Min} \left\{ \begin{array}{l} c_n^T x_n + E[h_{n+}(\underline{s}_{n+})] \\ x_n \in X_n(\underline{s}_n) \\ X_n(\underline{s}_n) = \left\{ \begin{array}{l} C_{t(n)} x_n = r_n - \\ D_n y_n, x_n \geq 0 \end{array} \right\} \text{ a.s.} \end{array} \right\} \tag{1}$$

and

$$y_{n+} = q_{n+} + A_{n+} x_n + B_{n+} y_n. \tag{2}$$

Since the dependence of the value function on data can be captured via the node index  $n$ , we drop the dependence on  $(\underline{\omega}_n)$ , and simply write  $(c_n, d_n, A_n, B_n, C_n, D_n, q_n, r_n)$  as data for node  $n$ , and the constraints on the decisions at node  $n$  are written as  $x_n \in X_n(\underline{s}_n)$  a.s. However, in order to appeal to the ideas of two-stage SD (Higle and Sen 1991), and to ease some of the computational burden, we assume  $(C_n) = (C_{t(n)})$ , that is, these instances satisfy a multi-stage version of the fixed-recourse assumption for two-stage problems. All other data elements are allowed to depend on the state of the stochastic process in node  $n$ .

The parallels between the two-stage and the multi-stage stochastic program should be clear. As usual, the above recursive statement is such that the multi-stage model can be conceptually interpreted both as a specialization as well as a generalization of a two stage model. From a computational point of view however, the presence of a nested collection of conditional expectation functions, poses far more serious computational challenges for the multi-stage case.

### 3 OPTIMIZATION SIMULATION WITH MULTI-STAGE STOCHASTIC DECOMPOSITION

We begin by presenting a summary of the algorithm which is discussed in greater depth subsequently.

1. *Simulate a sample path  $(\underline{\omega}_T^k)$  and for all nodes of the scenario tree, update counts as well as frequencies reflecting the number of visits, and the fraction of visits. We will refer to nodes along the sample path by the notation  $n \in \underline{\omega}_T^k$ . (If one initializes all counts to be zero, then one can simply update counts (of visits to nodes) on the sample path observed in the current iteration.)*

1.1. Assuming that approximations  $f_n^k$  are available, obtain incumbent and candidate sequences by solving a sequence of nodal decision simulations (see (3)) for all non-terminal nodes on the sample path (section 3.)

1.2. If some nodes  $n \in \omega_T^k$  have not been visited in previous iterations, then perform operations of step 1.1 until no such nodes are available, and then, solve two scenario LPs associated with the remainder of the path. One LP is initialized with the incumbent decision, and the other with the candidate decision. The resulting solutions yield the incumbent and candidate sequences.

2. Solve or approximately solve certain nodal dual approximation (see (4)) for each non-root node  $n \in \omega_T^k$ , for the incumbent sequence  $\{\hat{s}_n\}$ , as well as the candidate sequence  $\{s_n^k\}$ , where once again,  $n \in \omega_T^k$ .

3. Collect information and update approximations for each non-root node  $n \in \omega_T^k$ . To do so, collect information regarding dual vectors of the nodal dual approximation (see (4)), and pass an affine lower bounding approximation to the parent node denoted  $n -$ . Each parent (i.e. non-terminal) node  $n \in \omega_T^k$  updates a lower bounding approximation  $f_n^{k+1}$  (see (13)) using both the incumbent sequence  $\{\hat{x}_n\}$ , as well as the candidate sequence  $\{x_n^k\}$ .

4. Update the incumbent (i.e.  $\hat{x}_0^{k+1} \leftarrow x_0^k$ )  
**if**  $(f_0^{k+1}(x_0^k, \bar{y}_0) - f_0^{k+1}(\hat{x}_0^k, \bar{y}_0)) \leq \gamma [f_0^k(x_0^k, \bar{y}_0) - f_0^k(\hat{x}_0^k, \bar{y}_0)]$   $\hat{x}_0^{k+1} \leftarrow x_0^k$   
otherwise, we continue with  $\hat{x}_0^{k+1} \leftarrow \hat{x}_0^k$ . ( $\gamma \in (0,1)$  is a parameter).  
 Increment the iteration counter  $k$ , and repeat from 1.

It is important to recognize that in any iteration, approximations of the expected recourse functions are developed only for those nodes  $n$  that belong to the sample path generated in iteration  $k$  (i.e.  $\in \omega_T^k$ ). Thus, unlike other approaches for optimization simulation in stochastic programming (e.g. Donahue and Birge 2004, Linowsky and Philpott 2005), the multi-stage SD algorithm creates approximations of the expected recourse function of non-terminal nodes of the sample path  $\omega_T^k$  generated in iteration  $k$ . Further details of each step are provided below.

Borrowing from the case of two-stage SD, we will make the following assumptions: a) the set of first-stage decisions is compact, b) the relatively complete recourse assumption is satisfied at every stage; (i.e. (1) has a finite optimum for any setting of feasible  $y_n$ , and c) zero provides lower bound on all conditional expectations. (This assumption can be easily relaxed as in the two-stage case.) In addition, we reiterate the fixed recourse assumption ( $C_n = C_{t(n)}$ ), as well as the requirement that the stochastic process has finite support.

### 3.1. Simulate a sample path

We assume that the simulation will be consistent with the given stochastic process denoted  $\tilde{\omega}_T$ . In developing the approximations however, we will use the empirical distribution observed for transitions from node  $n$  to a child node  $n+$ . At iteration  $k$ , this estimate ( $p_{(n+|n)}$ ) will be denoted by  $p_{n+}^k$ . Our estimates of probabilities will reflect the number of times the simulation has visited a given node, given that the parent node was visited.

Steps 1.1 and 1.2 are often called forward pass because they are used to generate a sequence of states based on incumbent ( $\hat{x}_n^k$ ) and candidate ( $x_n^k$ ) decisions. Given a decision  $x_n$ , one is able to find the next “3d” states  $\{y_{n+}\}_{n+ \in C_n}$  using the dynamics by traversing the sampled path  $\omega_T^k$  forward in time, by starting with  $n = 0$  and then recursively calculating the “3d” states associated with the candidate sequence using the dynamics:  $y_n^{k+1} = q_n + A_n x_{n-}^{k+1} + B_n y_{n-}^{k+1}$ , followed by the decisions

$$x_n^{k+1} \in \operatorname{argmin} \underbrace{\{f_n^k(x_n, y_n^{k+1}) + \frac{\sigma_k}{2} \|x_n - \hat{x}_n^k\|^2 : x_n \in X_n(y_n^{k+1})\}}_{\text{Nodal Decision Simulation}} \quad (3)$$

where  $n \in \underline{\omega}_T^k$  follows  $n -$ . Similarly, the incumbent trajectory  $\hat{x}_n^k$  may be obtained by using the same process, but starting with  $\hat{x}_0^k$ . We refer to decision problems in (3) as a nodal decision simulation because these forward passes generate decisions during the simulation. The approximate functions  $f_n^k$  are updated from one iteration to the next by using piecewise linear approximations developed during the backward pass (step 2).

Step 1.2 plays a role in those iterations in which the sampling process discovers new nodes of the scenario tree. For such nodes, function approximations have not been created in any previous iteration, and as a result it is not possible to perform (3). What we recommend in this case is that a linear program be solved, starting from the first newly revealed node, with data corresponding to the rest of the sample path. The solution of this LP will provide the dual multiplier estimates to be used in the backward pass (step 2 - see next subsection).

### 3.2. Solve or approximately solve nodal dual approximations

These calculations are carried out backward in time, along the path that was generated in the forward pass. Accordingly, the definitions that follow are best carried out in a recursive manner, starting from a terminal node. One of the main ideas behind the multi-stage SD setup is that it highlights the use of state variables ( $\underline{s}_n := (y_n, \underline{\omega}_n)$ ) in the approximation process.

- For a terminal node  $n$ , we define  $E[h_{n+}] = 0$  and let  $h_n^k$  provide a lower bound on  $h_n$ .
- In order to define approximations for non-terminal nodes, we set forth a recursive definition by assuming that for all  $n+ \in \mathbb{C}_n$ , we have approximations  $h_{n+}^k$ , as well as empirical probabilities  $\{p_{n+}^k\}_{n+ \in \mathbb{C}_n}$ . Then, we define  $h_n^k$  as an empirical *lower bounding approximation* of the nodal recourse function at node  $n$  as follows:

$$h_n^k(\underline{s}_n) := d_n^T y_n + \operatorname{Min} \{ c_n^T x_n + \sum_{n+ \in \mathbb{C}_n} p_{n+}^k h_{n+}^k(\underline{s}_{n+}) : C_{t(n)} x_n = r_n - D_n y_n, x_n \geq 0 \}.$$

Of course, if  $n$  is a terminal node, above formula still applies, with  $h_{n+}^k = h_{n+} = 0$ .

For finite  $k$ , our simulation-based successive approximation scheme has two sources of error: the probability distribution (using empirical distributions), and the future value functions  $h_{n+}^k$ . The latter will be approximated using subgradients of approximations. Now, if we could ensure that the collection of approximations satisfy  $\{h_{n+}^k \rightarrow h_{n+}\}_{n+ \in \mathbb{C}_n}$  in some sense (e.g. epi-convergence), then one could (using the law of large numbers) ensure that  $h_n^k \rightarrow h_n$  with probability one. However, it is well known that the approximations generated by SD obeys epi-nesting relative to  $h_n$ , but not epi-convergence to  $h_n$  (Higle and Sen 1992, Rockafellar and Wets 1997). Thus the approximations  $h_n^k$  developed below (using subgradients) should be designed to ensure epi-nesting, which will be sufficient for convergence of state trajectories with probability one.

Suppose that for all  $n+ \in \mathbb{C}_n$ , we have already calculated a subgradient (with respect to  $y_{n+}$ )  $\delta_{n+}^k \in [\partial h_{n+}^k(\underline{s}_{n+}) - d_{n+}]$  and let  $\rho_{n+}^k$  denote the constant term associated with the corresponding supporting hyperplane. Assuming relatively complete recourse, and substituting for the “3d” state variables, that is,  $y_{n+} = q_{n+} + A_{n+} x_n + B_{n+} y_n$ , and replacing the primal value function by its dual representation, we obtain an affine lower bounding approximation as follows:

$$\begin{aligned}
 h_n^k(\underline{s}_n) &:= d_n^\top y_n + \sum_{n+ \in \mathbb{C}_n} p_{n+}^k \{ \rho_{n+}^k + (d_{n+} + \delta_{n+}^k)^\top [q_{n+} + B_{n+} y_n] \} + \\
 \text{Max } \pi_n^\top [r_n - D_n y_n] : C_{t(n)}^\top \pi_n &\leq c_n + \underbrace{\sum_{n+ \in \mathbb{C}_n} p_{n+}^k A_{n+}^\top (d_{n+} + \delta_{n+}^k)}_{\text{Nodal Dual Approximation for node } n} =: c_n^k
 \end{aligned} \tag{4}$$

The LPs in the second line of (4) will be referred to as nodal dual approximations (NDA). In order to instantiate the NDA for node  $n$ , denoted  $NDA(n)$ , note that the right-hand side requires quantities  $p_{n+}^k, A_{n+}, d_{n+}, \delta_{n+}^k$  for  $n+ \in \mathbb{C}_n$ . In order to do the recursion will assume that  $p_{n+}^k, \rho_{n+}^k$ , and  $\delta_{n+}^k$  are estimated and available for all  $n+ \in \mathbb{C}_n$ . Indeed, while we will maintain these quantities for all non-terminal nodes that have been visited at least once, they will only be updated periodically, depending on their “proximity” to the sampled path. The notion of “proximity” as well as the updating procedure will be described subsequently. For now, let us assume that we have all estimates  $\{ \delta_{n+}^k \}_{n+ \in \mathbb{C}_n}$  that are necessary to calculate  $c_n^k$ . Let  $\pi_n^k$  denote solutions obtained by setting  $y_n = y_n^k$  (a candidate) for the LPs in (4). Hence using  $\pi_n^k$ , we obtain

$$h_n^k(\underline{s}_n) \geq (\pi_n^k)^\top r_n + (d_n - D_n^\top \pi_n^k)^\top y_n + \sum_{n+ \in \mathbb{C}_n} p_{n+}^k \{ \rho_{n+}^k + (d_{n+} + \delta_{n+}^k)^\top [q_{n+} + B_{n+} y_n] \} \tag{5}$$

Letting

$$\delta_n^k = -D_n^\top \pi_n^k + \sum_{n+ \in \mathbb{C}_n} p_{n+}^k B_{n+}^\top (d_{n+} + \delta_{n+}^k) \tag{6}$$

$$\text{and } \rho_n^k = (\pi_n^k)^\top r_n + \sum_{n+ \in \mathbb{C}_n} p_{n+}^k \{ \rho_{n+}^k + (d_{n+} + \delta_{n+}^k)^\top q_{n+} \}, \tag{7}$$

we recursively obtain a subgradient  $\delta_n^k \in [\partial h_n^k(\underline{s}_n) - d_n]$  (with respect to  $y_n$ ) and the quantity  $\rho_n^k$  represents the “constant” term of the hyperplane. These quantities will be used for approximations at the parent node  $n -$ . Finally define an indicator vector  $w_{t(n)}^k$  for which the  $i^{th}$  element is 1 if the  $i^{th}$  row of  $NDA(n)$  is tight (i.e.  $C_{i,t(n)}^\top \pi_n^k = c_{i,n}^k$ ), and 0 otherwise. Then, we update

$$W_{t(n)}^k = W_{t(n)}^{k-1} \cup w_{t(n)}^k. \tag{8}$$

For the case in which we wish to update approximations for the incumbent sequence  $y_n = \hat{y}_n^k$ , we use it (the incumbent  $\hat{y}_n^k$ ) in NDA in (4) together with values  $\hat{\rho}_{n+}^k$  and vectors  $\hat{\delta}_{n+}^k$  provided by nodes  $n +$  and future incumbents  $\hat{y}_{n+}^k$ . Upon solving the NDA, the resulting dual multiplier for the incumbent will be denoted  $\hat{\pi}_n^k$ , and accordingly, one defines analogous  $\hat{\delta}_n^k$  and  $\hat{\rho}_n^k$  using the expressions in (6) and (7). As in case of the candidate updates, we will update the indicator vectors as in (8).

We now return to the issue of how one approximates the quantities used in defining NDA recursively. Let  $t(n)$  denote time period associated with node  $n$ . Among those nodes  $m$  that satisfy  $t(m) = t(n)$ , and  $p_m^k > 0$ , we calculate the quantities  $\delta_m^k$  in two subgroups.

- If  $m \neq n -$ , then we let  $\rho_m^k = \rho_m^{k-1}$ ,  $\delta_m^k = \delta_m^{k-1}$ . In other words, the estimates are not revised unless node  $m$  is a child of some node along the sample path generated in iteration  $k$ . (This is how we measure “proximity” of a node to a sample path.) In case iteration  $k$  is one

for which the subgradient approximations are to be updated for the incumbent sequence, we also put  $\hat{\rho}_m^k = \hat{\rho}_m^{k-1}$ ,  $\hat{\delta}_m^k = \hat{\delta}_m^{k-1}$ .

- Otherwise (i.e.  $m- = n-$ ), we update  $\delta_m^k$  by using an “argmax” procedure as follows. Let  $\pi_m^k \in \operatorname{argmax} \{ \pi_m^\top (r_m - D_m y_m^k) : C_{i,t(n)}^\top \pi_m = c_{i,m}^k, \forall w_i = 1, C_{i,t(n)}^\top \pi_m \leq c_{i,m}^k, \forall w_i = 0, w \in W_{t(m)}^k \}$ . Note that this procedure involves the solution of a system of linear equations (using equations indexed by  $w_i = 1$ ), followed by dual feasibility verification (for  $w_i = 0$ ), thus reducing the complexity of LP solves. If one is willing to store the factorizations associated with the corresponding bases of  $C_{t(n)}$ , then, one can reduce the computational complexity of this step even further, but at the cost of additional storage. Now the quantities  $\delta_m^k, \rho_m^k$  are estimated analogously to (6) and (7), as follows:

$$\delta_m^k = -D_m^\top \pi_m^k + \sum_{m+ \in \mathbb{C}_m} p_{m+}^k B_{m+}^\top (d_{m+} + \delta_{m+}^k) \tag{9}$$

$$\text{and } \rho_m^k = (\pi_m^k)^\top r_m + \sum_{m+ \in \mathbb{C}_m} p_{m+}^k \{ \rho_{m+}^k + (d_{m+} + \delta_{m+}^k)^\top q_{m+} \}. \tag{10}$$

We also remind the reader that one also needs to record analogous quantities  $(\hat{\delta}_n^k, \hat{\rho}_n^k)$ , for those  $m$  satisfying  $t(m) = t(n)$ , and  $p_m^k > 0$ , so that incumbent cuts can be updated when necessary. Because each iteration needs one subgradient for a candidate and one for an incumbent solution for all nodes that have been visited by the method, it may be worthwhile clustering these vectors whenever the difference are negligible (or acceptable for approximations).

### 3.3 Collect information and update approximations

Using the pairs (6),(7) (and their analogs for incumbent trajectories when necessary), we can now summarize the approximations developed at the candidate state trajectory for the multi-stage SD algorithm:

$$h_n^k(\underline{s}_n) \geq \rho_n^k + (d_n + \delta_n^k)^\top y_n =: g_{n-,n}^k(\underline{s}_n(x_{n-}, y_{n-})). \tag{11}$$

Upon substituting the dynamics (2) in place of  $y_n$  in (11), and setting  $n \leftarrow n-$ , we obtain an affine approximation of the following form

$$G_n^k(x_n, y_n) := \sum_{n+ \in \mathbb{C}_n} p_{n+}^k g_{n,n+}^k(\underline{s}_{n+}(x_n, y_n)) = \alpha_n^k + \beta_n^k x_n + \gamma_n^k y_n, \tag{12}$$

with

$$\alpha_n^k = \sum_{n+ \in \mathbb{C}_n} p_{n+}^k \left[ \rho_{n+}^k + (d_{n+} + \delta_{n+}^k)^\top q_{n+} \right], \quad \beta_n^k = \sum_{n+ \in \mathbb{C}_n} p_{n+}^k (d_{n+} + \delta_{n+}^k)^\top A_{n+},$$

$$\gamma_n^k = \sum_{n+ \in \mathbb{C}_n} p_{n+}^k (d_{n+} + \delta_{n+}^k)^\top B_{n+}$$

Let  $\kappa_n^\ell$  denote the number of visits to node  $n$ , when the  $\ell^{th}$  approximation is created. Then, in iterations  $k$ , the nodal lower bounding approximation for  $n \in \underline{\omega}_T^k$  is given by

$$f_n^k(x_n, y_n) := c_n^\top x_n + d_n^\top y_n + \operatorname{Max}_{1 \leq \ell \leq \kappa_n^k} \frac{\kappa_n^\ell}{\kappa_n^k} G_n^\ell(x_n, y_n). \tag{13}$$

Based on the assumption that the expected value/recourse function has a lower bound of zero, the above “Max” function above has the same form as in two-stage SD.

### 3.4 Update incumbent solutions

As with the two-stage regularized SD algorithm (Higle and Sen 1994), the choice of the incumbent is based on objective value estimates at the root node; that is, we estimate whether there is a reduction in the value  $f_0^k(x^k; \bar{y}_0)$ ; that is,  $\hat{x}_0^{k+1} \leftarrow x_0^k$  if

$$f_0^{k+1}(x_0^k, \bar{y}_0) - f_0^{k+1}(\hat{x}_0^k, \bar{y}_0) \leq \gamma [f_0^k(x_0^k, \bar{y}_0) - f_0^k(\hat{x}_0^k, \bar{y}_0)]$$

Note if any sequence of stage 1 incumbent decisions converge ( $\hat{x}_0^k \rightarrow \hat{x}_0$ ), then with probability one, the state trajectories also converge. Then the stability of the nodal decision simulations (see (3)) implies that the nodal decisions must also converge with probability one. The fact that these approximations satisfy epi-nesting follows from the observation that the sequence of sets  $W_t^k$  (for all  $t$ ) grows in cardinality as iterations proceed. Because there are finitely many constraints in any stage, these sets cannot grow indefinitely. Hence these “argmax” steps must eventually provide accurate estimates of solutions of the nodal dual problems. Moreover, the finite support assumption (of the scenario tree) ensures that every node of the scenario tree must be visited infinitely often. Thus, with probability one the approximations become accurate at the accumulation points of the sequence of nodal decisions. Hence, the method satisfies sufficient conditions for asymptotic convergence as discussed in Higle and Sen (1992) and Rockafellar and Wets (1997). Other regularizing tools may also be applied (see Lemarechal, Nemirovski and Nesterov 1995).

## 4 PRELIMINARY COMPUTATIONS

We present our preliminary computations within the context of a production and inventory instance available from H. Gassmann’s website (<http://myweb.dal.ca/gassmann/>). First, we run a small instance, which has 5 stages and every node in the scenario tree has three children nodes except the nodes at last stage. Therefore, there are totally 81 scenarios in this instance. Figure 1 reports the computational results of the first stage solution as iterations proceed. One can observe that after 32 iterations, the first stage solution converges to (8,0). This solution is verified to be optimal by solving the corresponding deterministic equivalent linear program.



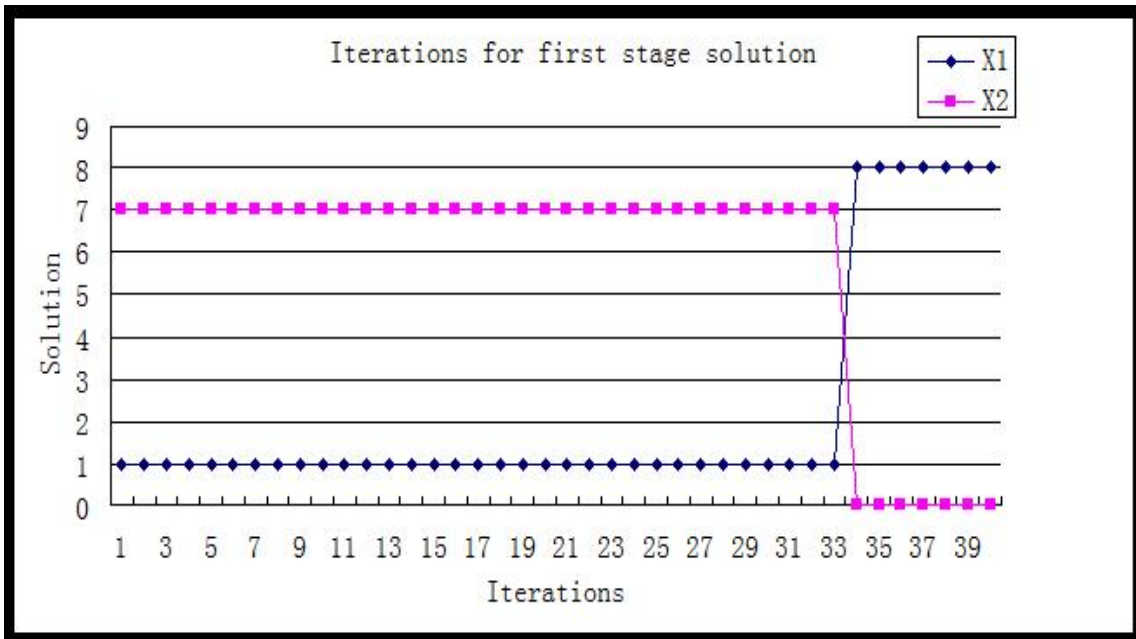


Figure 1: Sequence of first-stage solutions for a small production-inventory instance (81 scenarios)

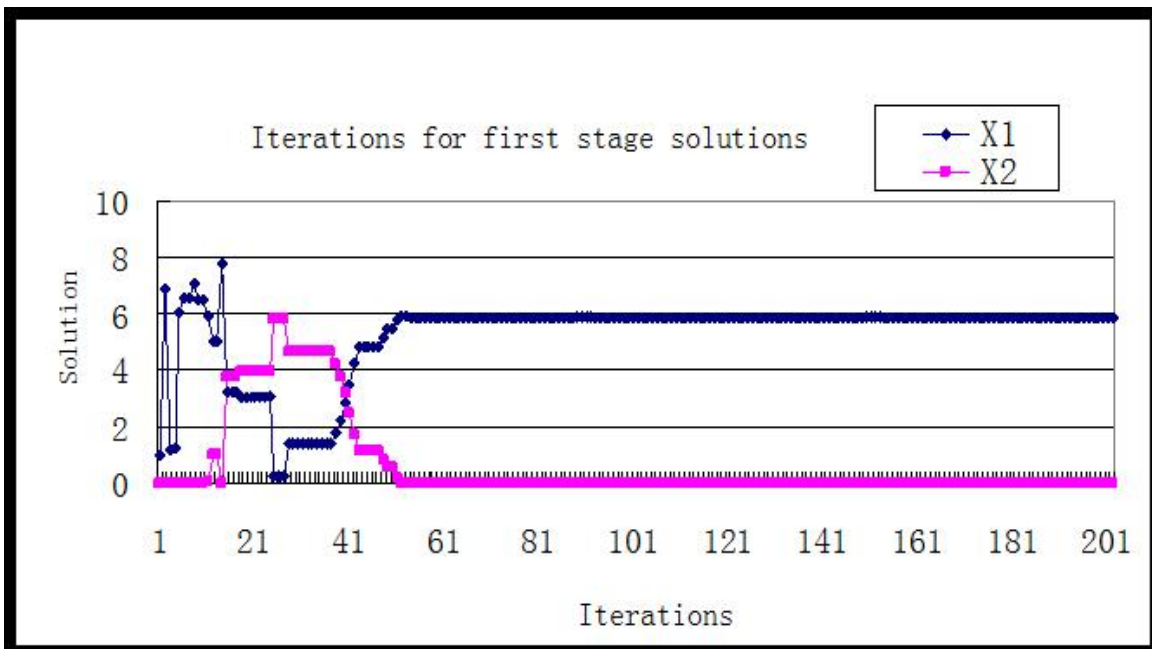


Figure 2: Sequence of first-stage solutions for a larger production-inventory instance ( $32^5$  scenarios)

We also replicated the solution of this (small) instance by using 20 different seeds to run this algorithm. A summary of the outcomes of these runs are provided in Table 1.

Table 1: Results from Replications of the Multi-stage Stochastic Decomposition Algorithm

First-stage Solution	Sample Average	Std. Deviation	Confidence Interval Half Width ( $\alpha = 0.05$ )
$x_1$	7.978	0.141	0.062
$x_2$	0.068	0.032	0.019

In order to study whether the method would scale to problems with more scenarios, we run a similar production-inventory instance, but for this one we allow 6 stages and every node has 32 children nodes. This results in an instance with  $32^5$  scenarios. Furthermore, we fix the number of iterations as 500. Figure 2 reports the computational results of this instance. It is not difficult to observe that the solutions stabilize after approximately 50 iterations. Since it is not possible to solve the deterministic equivalent problem in this instance, we use the scenario tree generated by Multi-stage SD method and solve the corresponding deterministic linear program. Once again, we find that solutions by solving deterministic problem are identical to that obtained by running the Multi-stage SD algorithm.

## 5 COMPARISONS AND CONCLUSIONS

The multi-stage SD algorithm can be specialized or generalized in several ways: a) The process is well suited for special structures (e.g. networks and generalized networks for the matrices  $C_t$ ); b) the method is applicable to cases such as model-predictive control in which objectives are typically defined using quadratic terms; c) single-stage dependence (as in Markov Chains) and more general dependencies and lags are also allowed within a finite horizon setting. These are clearly important areas for which the proposed methodology is well suited. We compare our methodology with certain variants of dynamic programming, as well as sampled versions of Nested Benders' decomposition in greater detail below.

### 5.1. Connections with Variants of Dynamic Programming

The statement of a multi-stage stochastic programming (SP) problem (see 0-2) is patterned after the DP principle of optimality. However, the algorithmic process used to solve such problems go beyond the DP methodology, through its use of numerical approximation techniques. Just as one states differential equations as models, one might liken the statement in 0 as a model, whose solution via analytical DP is untenable. Instead, as in the case of differential equations, we resort to numerical approximations, and this is where the power of convexity and variational methods (King and Wets 1991, Rockafellar and Wets 1997) are brought to bear on this problem.

One might also think of our approach as a non-smooth and stochastic analog of smooth deterministic methods such as Differential Dynamic Programming (DDP) proposed by Jacobson and Mayne (1970). In this context however, the reader might be reminded of a popular commercial campaign slogan "This is not your father's,..." DP or DDP!. The approach we have presented accommodates randomness, constraints, and non-smooth objectives. It is also the only sampling-based method that works with directly with sample paths, thus making it a convenient tool for integration within a simulation.

In a sense, one can also think of SD as an approximate DP (ADP) method in which stochastic estimates of subgradients are used to create and update the value function approximations recursively. To see this, note that the subgradient of the future nodes appear in the right hand side of the NDA via  $c_n^k$ . In the spirit of ADP, one could interpret SD as using forecasts of future prices as the basis for value function approximations. For illustrative purposes, consider the two-stage case. Suppose that  $\{\pi_w\}_{w \in W_1^k}$  denote the collection of future prices (or dual variables) available for use in forming the first-stage approximation. Then, the approximate value function for state variable  $s_1$  can be expressed as  $h_1^k(s_1) =$

$\max \left\{ \sum_{w \in W_1^k} \lambda_w \phi_w(s_1) : \sum_{w \in W_1^k} \lambda_w = 1, \lambda_w \geq 0 \right\}$ , where  $\phi_w(s_1) = \pi_w^\top [r_1 - D_1 y_1]$ , where the quantities  $(r_1, D_1, y_1)$  are specified by the state  $s_1$ , which in turn is specified by the outcome  $\omega_1$  pointing to  $(r_1, D_1)$ , and the state  $y_1$ . Clearly, the “max” above is attained for that dual price for which  $\phi_w(s_1)$  assumes the highest value, which is the same as that given in 0. Note that in SD, the collection of functions indexed by  $w$  grows with iterations, and ultimately, the dual LP must stop generating new functions because the total number of dual bases is finite. Similar analogs exist in the multi-stage case. However, unlike ADP, multi-stage SD requires both forward as well as backward passes. It is these passes that guarantee epi-nesting, and hence convergence (with probability one).

**5.2 Connections with Sampled Nested Benders’ Decomposition**

Finally, we present the connections between sampled versions of Nested Benders’ decomposition, and the algorithmic framework of this paper (i.e Multi-stage SD).

Table 2: Comparison of Optimization Simulation Algorithms

<b>Sampled Nested Benders Decomposition</b>	<b>Multi-stage Stochastic Decomposition</b>
Works with sampled sub-trees	Works with individual sample paths
Allows general constraint matrices for all data	Requires fixed recourse ( $C_t$ ) in each period
Requires stage-wise independence of random variables	Overcomes the independence requirement
Forward and backward passes solve LPs with similar structures, both of which include cuts that may destroy specialized matrices (e.g. network flows)	Forward pass includes cuts representing piecewise linear value functions, but the backward pass has no cuts. The backward pass preserves special structure.
Every iteration updates approximations over the sampled sub-tree, and as the size of the sub-tree grows, the updates also grow	Every iteration updates approximations along the sample path, and some approximations are updated depending upon the proximity of nodes to the sampled path
Not designed to interface with a simulator	Designed specifically to interface with a simulator

**ACKNOWLEDGMENTS**

We are grateful to the National Science Foundation for its continued support of this line of research through CMMI grant 0900070. The first author is also grateful to Warren Powell on interesting discussions regarding DP, ADP and SP.

**REFERENCES**

Benders, J.F. 1962. “Partitioning procedures for solving mixed integer variables programming problems,” *Numerische Methemattick*, 4: 238-252.  
 Birge, J.R. and F. Louveaux 1997. *Introduction to Stochastic Programming*, Springer-Verlag.  
 Casey, M. and S. Sen 2005. “The Scenario Generation Algorithm for Multi-stage Stochastic Linear Programming,” *Mathematics of Operations Research*, 30:615-631.  
 Dantzig, G.B. and Glynn, P.W. 1990. “Parallel Processors for Planning Under Uncertainty,” *Annals of Operations Research* 22:1-21.  
 Donohue, C.J. and J. R. Birge. 2006. “The abridged nested decomposition method for multistage stochastic linear programs with relatively complete recourse,” *Algorithmic Operations Research*, 1:20-30.  
 Dyer, M, and L. Stougie. 2006. “Computational Complexity of Stochastic Programming Problems,” *Mathematical Programming*, 106:423-432.

- Higle, J.L. and S. Sen. 1991. "Stochastic Decomposition: An algorithm for two stage linear programs with recourse," *Mathematics of Operations Research*, vol. 16:650-669.
- Higle, J.L. and S. Sen. 1992. "On the convergence of algorithms with implications for stochastic and nondifferentiable optimization," *Mathematics of Operations Research*, 17:112-131.
- Higle, J.L. and S. Sen. 1994. "Finite master programs in stochastic decomposition," *Mathematical Programming*, 67:143-168.
- Jacobson, D.H. and D. Q. Mayne. 1970. *Differential Dynamic Programming*, Elsevier, Netherlands.
- Infanger, G. 1994. *Planning Under Uncertainty – Solving Large-Scale Stochastic Linear Programs*, The Scientific Press Series, Boyd and Fraser.
- King, A.J. and R. Wets. 1991. "Epi-consistency of convex stochastic programs," *Stochastics and Statistics Reports*, 34: 83-92.
- Linowsky, K. and A. B. Philpott. 2005. "On the convergence of sampling-based decomposition algorithms for multistage stochastic programs," *Journal of Optimization Theory and Applications*, 125: 349-366.
- Lemarechal, C., A. Nemirovskii, A., and Y. Nesterov. 1995. "New variants of bundle methods," *Mathematical Programming*, 69:111-147, 1995.
- Pereira, M.V. and L. M. Pinto. 1991. "Multi-stage stochastic optimization applied to energy planning," *Mathematical Programming*, 52:359-375.
- Powell, W. B. 2007. *Approximate Dynamic Programming: Solving the curses of dimensionality*, Wiley Series in Probability and Statistics.
- Rockafellar, R.T. and R.J-B. Wets. 1997. *Variational Analysis*, Springer-Verlag, Berlin, Germany.
- Shapiro, A. 2010. "Analysis of Stochastic Dual Dynamic Programming Method," *Optimization Online*.
- Shapiro, A. 2006. "On complexity of multistage stochastic programs," *Operations Research Letters*, 34: 1-8.
- Van Slyke, R. and R. Wets. 1969. "L-shaped linear programs with applications to optimal control and stochastic linear programs," *SIAM Journal on Applied Mathematics*, 17: 638-663.

## AUTHOR BIOGRAPHIES

**SUVRAJEET SEN** is Professor of Industrial and Systems Engineering, and Director of the Data-Driven Decisions Lab at the Ohio State University. Until recently, he also served as the Director of the College's Center for Energy, Sustainability, and the Environment. Prior to joining OSU, he served on the faculty at the University of Arizona, and he also served as a program director at NSF where he was responsible for the Operations Research, and the Service Enterprise Engineering programs. Professor Sen is a Fellow of INFORMS. He has served on the editorial board of several journals, including Operations Research as Area Editor for Optimization, and as Associate Editor in INFORMS Journal on Computing, Operations Research, and Journal of Telecommunications Systems. Professor Sen is the past-Chair of the INFORMS Telecommunications Section and founded the INFORMS Optimization Section.

**ZHIHONG ZHOU** is a Ph.D. student at the University of Arizona. His research interests are in stochastic programming models and methods.