Mathematical Programming
https://doi.org/10.1007/s10107-018-1339-4

FULL LENGTH PAPER

Series B

Two-stage linear decision rules for multi-stage stochastic programming

Merve Bodur[1] · **James R. Luedtke**[2]

Received: 15 January 2017 / Accepted: 28 September 2018

Abstract
Multi-stage stochastic linear programs (MSLPs) are notoriously hard to solve in general. Linear decision rules (LDRs) yield an approximation of an MSLP by restricting the decisions at each stage to be an affine function of the observed uncertain parameters. Finding an optimal LDR is a static optimization problem that provides an upper bound on the optimal value of the MSLP, and, under certain assumptions, can be formulated as an explicit linear program. Similarly, as proposed by Kuhn et al. (Math Program 130(1):177–209, 2011) a lower bound for an MSLP can be obtained by restricting decisions in the dual of the MSLP to follow an LDR. We propose a new approximation approach for MSLPs, *two-stage LDRs*. The idea is to require only the state variables in an MSLP to follow an LDR, which is sufficient to obtain an approximation of an MSLP that is a *two-stage stochastic linear program* (2SLP). We similarly propose to apply LDR only to a subset of the variables in the dual of the MSLP, which yields a 2SLP approximation of the dual that provides a lower bound on the optimal value of the MSLP. Although solving the corresponding 2SLP approximations exactly is intractable in general, we investigate how approximate solution approaches that have been developed for solving 2SLP can be applied to solve these approximation problems, and derive statistical upper and lower bounds on the optimal value of the MSLP. In addition to potentially yielding better policies and bounds, this approach requires many fewer assumptions than are required to obtain an explicit reformulation when using the standard static LDR approach. A computational study on two example problems demonstrates that using a two-stage LDR can yield significantly better primal policies and modestly better dual policies than using policies based on a static LDR.

Keywords Multi-stage stochastic programming · Linear decision rules · Two-stage approximation

Mathematics Subject Classification 90C15 · 90C39

✉ Merve Bodur
bodur@mie.utoronto.ca

Extended author information available on the last page of the article

Published online: 15 October 2018

1 Introduction

We present a new approach for approximately solving multi-stage stochastic linear programs (MSLPs). MSLPs model dynamic decision-making processes in which a decision is made, a stochastic outcome is observed, another decision is made, and so on, for T stages. At each stage, the decision vectors are constrained by linear constraints that depend on the history of observed stochastic outcomes. A solution of an MSLP is a policy, which defines the decisions to be made at each stage as a function of the observed outcomes up to that stage. The objective in an MSLP is to choose a policy that minimizes the expected cost over all stages. Although MSLPs can be used to model a wide variety of problems (e.g., [61]), they are notoriously hard to solve in general [17,57].

There are a variety of methods available for MSLPs in the case that the stochastic process is represented by a scenario tree [12,28,32,55]. Such algorithms include nested Benders decomposition [8,10,22], progressive hedging [49], and aggregation and partitioning [2,9], and enable the solution of MSLPs with possibly very large scenario trees. Unfortunately, as discussed in [57], the size of a scenario tree needed to obtain even a modestly accurate approximation grows exponentially in the number of stages. For example, a 10-stage problem in which the uncertainty in each stage is represented by just 50 realizations would yield a scenario tree having nearly 2×10^{15} scenarios, making any approach that requires even a single pass through the scenario tree impossible.

Under some conditions, including stage-wise independence of the random variables, stochastic dual dynamic programming (SDDP) [42] can overcome the difficulty in exploding scenario tree size, by constructing a single value function approximation for each stage. The SDDP algorithm converges almost surely on a finite scenario tree [14,53] (see also [24,26,44] for related results). In some cases, such as additive dependence [33] and Markov dependence [43], the assumption of stage-wise independence can be satisfied via an appropriate reformulation, e.g., see Example 10 in [54] (see also [25] for other types of dependence). However, such reformulations are not applicable for stage-wise dependence of random recourse matrices (i.e., in the coefficients of the constraints) or objective coefficients. Similar approaches that exploit stage-wise independence and value function approximations include multi-stage stochastic decomposition [51] and approximate dynamic programming [47].

An alternative approach to handling the complexity of MSLP is to restrict the functional form of the policy. One such approach is the use of *linear decision rules* (LDRs). The idea of an LDR is to require that all decisions made in each stage be a linear (or affine) function of the observed random outcomes up to that stage. The problem then reduces to a static problem of finding the best LDR, whose expected cost then yields an upper bound on the optimal value of the MSLP. In this paper, we refer this use of an LDR as a *static LDR*. While LDRs have a long history (see, e.g., [21]), they have recently gained renewed interest in the mathematical optimization literature after their application to adjustable robust optimization in [4]. The adaptation of this approach to MSLP was presented in [57], and Kuhn et al. [35] analyzed the application of a static LDR in the dual of the MSLP, which yields a lower bound on the optimal value of the MSLP. Moreover, under certain assumptions, the static approximations

obtained after restricting the primal and dual policies to be an LDR are both tractable linear programs, as shown in [35,57], respectively. The assumptions include stage-wise independence (or a slight generalization), compact and polyhedral support, and that uncertainty is limited to the right-hand side of the constraints. While in some cases static LDR policies provide high quality approximations to MSLP, they have potential to be significantly suboptimal. Better policies (primal and dual) can be obtained by considering more flexible (nonlinear) rules such as (static) piecewise linear decision rules [13] and polynomial decision rules [3].

We propose a new use of an LDR, which we refer to as *two-stage linear decision rules*. The key idea is to partition the decision variables into *state* and *recourse* decision variables, with the property that if the state variables are fixed, then the problem decouples into a separate problem for each stage, involving only recourse decision variables. If one applies an LDR *only* to the state variables, then the problem reduces to a *two-stage* stochastic linear program (2SLP), in contrast to a *static* problem which is obtained when using a static LDR. The advantage of two-stage LDRs is that they free the recourse variables from the LDR requirement, thus allowing for a potentially improved policy. Indeed, there exist feasible 2SLPs that are *infeasible* if one enforces an LDR on the recourse variables [21]. This idea of reducing an MSLP to a 2SLP is similar to that proposed by Ahmed [1], except that in [1] the state variables are completely decided in the first-stage and fixed, whereas we allow them to vary according to an LDR. We also consider applying a two-stage LDR in the dual of an MSLP, exploiting the observation that imposing an LDR restriction *only* on the dual variables associated with the *state equations* is sufficient to obtain a 2SLP that approximates the multi-stage dual problem. We investigate how approximate solutions to the associated primal and dual approximation problems can be used to obtain feasible policies with associated statistical estimates on the optimality gap. Our analysis suggests that this can be done under mild assumptions, for example that the primal problem exhibits relatively complete recourse (i.e., for any current state there exists a feasible next decision and state) and has a bounded feasible region with probability 1. We illustrate the two-stage LDR approach on two example problems: an inventory planning problem similar to that studied in [4,35], and a capacity expansion problem proposed in [15]. We find that, for these problems, using two-stage LDRs yields significantly better primal policies (upper bounds), and modestly improves on the lower bounds, when compared to using static LDRs. For the capacity expansion problem, we also compare the two-stage LDR policies and bounds to those obtained using the SDDP algorithm, when run for a similar amount of computational time. We find that the SDDP algorithm yields similar lower bounds and better policies for this problem, as expected since the SDDP algorithm is known to converge to an optimal solution. Thus, the two-stage LDR approximation is expected to be useful primarily for problems where the SDDP algorithm does not apply.

A significant challenge to using two-stage LDRs is that the resulting 2SLP is in general intractable to solve exactly. Indeed, 2SLP is $\#P$-hard [17,27] due to the difficulty in evaluating the expectation of the recourse function. However, as argued in [57], under mild conditions Monte Carlo sampling-based methods can provide solutions of modest accuracy to a 2SLP (such a statement cannot be made for MSLP). Thus, an important benefit of the two-stage LDR approach is that it enables the appli-

cation of the long history of research into solving 2SLPs to the multi-stage setting. While using a sampling-based method may lead to a suboptimal solution of the 2SLP approximations, our hope is that this suboptimality may be more than offset by the improvement gained by eliminating the LDR requirement on the recourse decisions that is imposed when using a static LDR. In addition, when using a sampling-based approach, the assumptions that are required for obtaining a tractable reformulation when applying a static LDR are no longer needed. In particular, the random variables need not have polyhedral (or even bounded) support, the constraint matrices may be random and dependent across time stages, and the LDR may be based on nonlinear functions of the random variables.

The two-stage LDR approach we propose can also be applied to certain multi-stage stochastic *integer programs*, in which some of the decision variables are required to be integer valued. In particular, for the primal problem, the approach applies directly provided integrality restrictions are imposed only on the recourse variables. When the state variables have integrality restrictions as well, the form of the decision rule applied to the state variables must be modified, but the two-stage approach still applies. We refer to [7] for one possible such decision rule based on piecewise-linear binary functions. We remark that combining our approach with that of [7] would have potential benefit in terms of both tractability and policy quality, as removing the piecewise-linear binary decision rule requirement from the recourse variables both eliminates the need to design such a rule, and gives those decisions more flexibility.

The rest of this paper is organized as follows. Section 2 defines the MSLP, reviews the static LDR approach, and presents the proposed two-stage LDR approach, including discussion of how to solve the approximate problem and obtain statistical upper bounds on the original MSLP. Section 3 conducts a similar analysis for the dual of an MSLP, yielding an approach for finding statistical lower bounds on an MSLP. We present illustrative applications in Sects. 4 and 5, and make concluding remarks in Sect. 6.

2 Primal two-stage linear decision rules

We formulate an MSLP with $T \geq 2$ stages as follows, where throughout the paper, for integers $a \leq b$, $[a, b] := \{a, a+1, \ldots, b\}$ and $[b] := \{1, \ldots, b\}$:

$$\min_{x,s} \ \mathbb{E}\Big[\sum_{t\in[T]} c_t(\xi^t)^\top x_t(\xi^t) + h_t(\xi^t)^\top s_t(\xi^t) \Big] \tag{1a}$$

$$\text{s.t.} \quad A_t(\xi^t)s_t(\xi^t) + B_t(\xi^t)s_{t-1}(\xi^{t-1}) + C_t(\xi^t)x_t(\xi^t) = b_t(\xi^t), \quad t \in [T], \ \mathbb{P}\text{-a.s.}, \tag{1b}$$

$$(x_t(\xi^t), s_t(\xi^t)) \in X_t(\xi^t), \quad t \in [T], \ \mathbb{P}\text{-a.s.} \tag{1c}$$

where for $t \in [T]$

$$X_t(\xi^t) := \{x_t \in \mathbb{R}^{p_t}, \ s_t \in \mathbb{R}^{q_t} : D_t(\xi^t)s_t + E_t(\xi^t)x_t \geq d_t(\xi^t)\}.$$

Here, $\{\xi_t\}_{t=1}^T$ is a stochastic process with probability distribution $\mathbb{P}$ and support Ξ, where $\xi_1 = 1$ for all $\xi \in \Xi$ (i.e., data in stage 1 is deterministic), ξ_r is a random vector taking values in $\mathbb{R}^{\ell_r}$ for $r \in [2, T]$, and $\xi^t := (\xi_1, \ldots, \xi_t)$ for $t \in [T]$. Letting $\ell_1 = 1$, we denote $\ell^t := \sum_{r=1}^t \ell_r$ for $t \in [T]$. The s and x variables are referred to as *state* and *recourse* variables, respectively. Similarly, (1b) and (1c) are referred to as *state equations* and *recourse constraints*, respectively. The objective is to minimize the expected total cost. The functions $b_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{m_t}$, $d_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{n_t}$, $A_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{m_t \times q_t}$, $B_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{m_t \times q_{t-1}}$, $C_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{m_t \times p_t}$, $D_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{n_t \times q_t}$, and $E_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{n_t \times p_t}$ define the random coefficients as a function of ξ^t. Frequently in the literature, these are assumed to be affine functions of ξ^t, but we will not need this assumption in this work. In (1b) for $t = 1$, we adopt the convention that $s_0(\xi^0) = 0$. The constraints are required to be almost surely satisfied with respect to the distribution of the stochastic process, denoted by "$\mathbb{P}$-a.s." We note that any MSLP can be brought into the form of (1) by introducing additional variables and constraints. Throughout the paper, we assume that (1) is feasible and has an optimal solution, and we denote its optimal objective value as z^{MSLP}.

2.1 Static linear decision rules

A tractable approximation of MSLP can be obtained by restricting the decision policies to a certain form, i.e., by restricting the decisions to be a special function of the uncertain parameters. A linear decision rule is a policy in which the decisions at each stage t are restricted to be a linear function of the observed random variables ξ^t up that stage. We refer to the policies in which all the decisions are required to follow an LDR as *static LDR policies*. Specifically, a static LDR policy has the form:

$$s_t(\xi^t) = \beta_t \Phi_t(\xi^t), \tag{2a}$$

$$x_t(\xi^t) = \Theta_t \Phi_t(\xi^t), \tag{2b}$$

where $\Theta_t \in \mathbb{R}^{p_t \times K_t}$ and $\beta_t \in \mathbb{R}^{q_t \times K_t}$ are free parameters of the LDR, and $\Phi_t(\xi^t) = (\Phi_{t1}(\xi^t), \ldots, \Phi_{tK_t}(\xi^t)) : \mathbb{R}^{\ell^t} \to \mathbb{R}^{K_t}$ for all $t \in [T]$ is a vector of given *LDR basis functions*. We refer to the Θ and β variables as the *LDR variables*. We assume $K_1 = 1$ and $\Phi_{t1}(\xi^t) \equiv 1$ for all $t \in [T]$. Often, the basis functions are the uncertain parameters themselves, i.e., $K_t = \ell^t$ and $\Phi_{tk}(\xi^t) = (\xi^t)_k$, where $(\xi^t)_k$ denotes the k^{th} component of ξ^t vector. In this case, we refer to the basis functions as the *standard basis functions*. Note that the convention $\xi_1 \equiv 1$ implies that the decisions made at stage t are actually *affine* functions of the random variables $(\xi_2, \ldots, \xi_t)$. Finally, for notational convenience we adopt the convention that $\Phi_0 \equiv 0$, so that any term involving Φ_0 disappears.

Substituting the LDRs of the form (2) into MSLP given in (1) yields the following approximation of MSLP, which we call P-LDR:

$$\min_{\Theta, \beta} \mathbb{E}\left[\sum_{t \in [T]} c_t(\xi^t)^\top \Theta_t \Phi_t(\xi^t) + h_t(\xi^t)^\top \beta_t \Phi_t(\xi^t)\right]$$

$$
\begin{aligned}
\text{s.t. } & A_t(\xi^t)\beta_t\Phi_t(\xi^t) + C_t(\xi^t)\Theta_t\Phi_t(\xi^t) \\
& \quad + B_t(\xi^t)\beta_{t-1}\Phi_{t-1}(\xi^{t-1}) = b_t(\xi^t), && t \in [T],\ \mathbb{P}\text{-a.s.}, \\
& D_t(\xi^t)\beta_t\Phi_t(\xi^t) + E_t(\xi^t)\Theta_t\Phi_t(\xi^t) \geq d_t(\xi^t), && t \in [T],\ \mathbb{P}\text{-a.s.}, \\
& \Theta_t \in \mathbb{R}^{p_t \times K_t},\ \beta_t \in \mathbb{R}^{q_t \times K_t}, && t \in [T]. \qquad (3)
\end{aligned}
$$

We let z^{LDR} denote the optimal value of P-LDR, where here and elsewhere, we adopt the convention that if a minimization (maximization) problem is infeasible, the associated optimal value is defined to be $+\infty$ ($-\infty$). Note that all the decision variables are deterministic, i.e., they have to be determined before observing any random outcomes, and hence this problem is a static problem. P-LDR is a semi-infinite program having infinitely many constraints. It is observed in [57] (see also [13,35]) that P-LDR can be reformulated as a linear program (LP) using robust optimization techniques under the following assumptions:

A1. The standard basis functions are used.
A2. For all $t \in [T]$, the constraint matrices, A_t, B_t, C_t, D_t, E_t, are independent of the random vector ξ^T, and $b_t(\xi^t)$ and $d_t(\xi^t)$ are affine functions of ξ^t.
A3. The support, Ξ, is a nonempty compact polyhedron.

The LP reformulation has constraints of the form $(\beta, \Theta, w) \in \mathcal{D}$, where w are auxiliary variables and $\mathcal{D}$ is an explicitly given polyhedron. The size of this LP scales well (typically grows only quadratically) with the number of stages T. Moreover, the LP does not require any discretization of $\mathbb{P}$ (e.g., by Monte Carlo sampling), and instead only uses a polyhedral description of Ξ and the second order moment matrix of the random variables. These results have been generalized in [23] to the case of conic support, where A3 is replaced by an assumption that Ξ is described by a finite set of conic inequalities, in which case P-LDR (3) is reformulated as a conic program.

As P-LDR (3) is a restriction of MSLP, it provides an upper bound to MSLP. However, the benefit of tractability comes at the expense of loss of optimality. That is, the obtained upper bound can be substantially far from the optimal value of MSLP. Indeed, for 2SLPs, the optimal recourse decisions are very rarely linear in the random variables, but there always exists an optimal *piecewise linear* decision rule [21].

2.2 Two-stage linear decision rules

We propose *two-stage LDRs* which yield upper bounds to MSLP that cannot be worse than the ones obtained by P-LDR (3). The key idea is to apply an LDR only on the state variables to obtain a *two-stage* approximation of MSLP, rather than a static approximation. Substituting the LDR of the form (2a) into the MSLP given in (1) yields

$$
\begin{aligned}
\min_{x,\beta} & \sum_{t\in[T]} \mathbb{E}\big[c_t(\xi^t)^\top x_t(\xi^t)\big] + \sum_{t\in[T]} \mathbb{E}\big[h_t(\xi^t)^\top \beta_t\Phi_t(\xi^t)\big] \\
\text{s.t. } & A_t(\xi^t)\beta_t\Phi_t(\xi^t) + C_t(\xi^t)x_t(\xi^t) \\
& \quad + B_t(\xi^t)\beta_{t-1}\Phi_{t-1}(\xi^{t-1}) = b_t(\xi^t), && t \in [T],\ \mathbb{P}\text{-a.s.},
\end{aligned}
$$

$$
\begin{aligned}
& D_t(\xi^t)\beta_t\Phi_t(\xi^t) + E_t(\xi^t)x_t(\xi^t) \geq d_t(\xi^t), && t \in [T],\ \mathbb{P}\text{-a.s.},\\
& x_t(\xi^t) \in \mathbb{R}^{p_t}, && t \in [T],\ \mathbb{P}\text{-a.s.},\\
& \beta_t \in \mathbb{R}^{q_t \times K_t}, && t \in [T]. \qquad (4)
\end{aligned}
$$

We denote this problem as P-LDR-2S and the optimal value of this problem as z^{2S}. This problem is a 2SLP, which can be equivalently written as follows, where we drop dependence of the first-stage variables on $\xi^1 \equiv 1$ and use $\beta_1 = s_1(\xi^1) = \Phi_1(\xi^1)\beta_1$,

$$
\begin{aligned}
z^{2S} = \min_{x_1,\beta}\ & c_1^\top x_1 + \sum_{t\in[T]} \mathbb{E}\left[h_t(\xi^t)^\top \beta_t \Phi_t(\xi^t)\right] + \mathbb{E}\left[\mathcal{Q}(\beta, \xi^T)\right]\\
\text{s.t. } & A_1\beta_1 + C_1 x_1 = b_1,\\
& D_1\beta_1 + E_1 x_1 \geq d_1,\\
& x_1 \in \mathbb{R}^{p_1},\\
& \beta_t \in \mathbb{R}^{q_t \times K_t},\ t \in [T],
\end{aligned}
$$

where $\mathcal{Q}(\beta, \xi^T) := \sum_{t\in[2,T]} Q_t(\beta, \xi^t)$ and for $t \in [2, T]$,

$$
\begin{aligned}
Q_t(\beta, \xi^t) := \min_{x_t}\ & c_t(\xi^t)^\top x_t && (5a)\\
\text{s.t. } & C_t(\xi^t)x_t = b_t(\xi^t) - A_t(\xi^t)\beta_t\Phi_t(\xi^t)\\
& \qquad\qquad - B_t(\xi^t)\beta_{t-1}\Phi_{t-1}(\xi^{t-1}), && (5b)\\
& E_t(\xi^t)x_t \geq d_t(\xi^t) - D_t(\xi^t)\beta_t\Phi_t(\xi^t), && (5c)\\
& x_t \in \mathbb{R}^{p_t}. && (5d)
\end{aligned}
$$

The following proposition, immediate from the definitions of the associated problems, summarizes the relationship between the optimal values of MSLP, P-LDR (3), and P-LDR-2S (4).

Proposition 1 *The following inequalities hold:*

$$z^{MSLP} \leq z^{2S} \leq z^{LDR}.$$

The difference between z^{LDR} and z^{2S} can be arbitrarily large. In particular, an example is given in [21] of a 2SLP having relatively complete recourse for which P-LDR (3) is infeasible ($z^{LDR} = \infty$), while the 2SLP [and hence P-LDR-2S, (4)] is feasible.

Unfortunately, the techniques used to derive a static approximation of P-LDR (3) do not yield an efficiently computable reformulation of P-LDR-2S (4), even under assumptions A1–A3. In the next section, we review approaches for obtaining an approximate solution, say $\hat{\beta}$, of P-LDR-2S (4). Then, in Sect. 2.4 we discuss techniques for obtaining a feasible policy (and hence estimating an upper bound on z^{MSLP}) using such a solution.

2.3 Approximate solution of P-LDR-2S

There is a huge literature on (approximately) solving 2SLP problems. In this section, we present a brief overview of relevant approaches, with a focus on identifying the required assumptions. We refer the reader to [55] for more details.

A common approach for approximately solving a 2SLP is *sample average approximation*, in which $\mathbb{P}$ is approximated by a discrete probability measure $\hat{\mathbb{P}}$ that assigns positive weights only to a finite (relatively small) number of realizations of ξ^T which are called *scenarios*. In this way, the intractable expectation term is replaced with a sum. Scenarios may be constructed by a variety of techniques, such as Monte Carlo, quasi-Monte Carlo, and Latin hypercube sampling (e.g., [30,34,38,41,56]). For the purpose of this paper, we consider only the conceptually simplest case in which scenarios are generated via independent Monte Carlo sampling.

Let ξ_j^T, $j = 1, \ldots, N$, be an independent and identically distributed (i.i.d.) random sample of the random vector ξ^T, and define the sample average approximation (SAA) problem:

$$\hat{z}_N^{2S} := \min_{x_1, \beta} \; c_1^\top x_1 + \sum_{t \in [T]} \mathbb{E}\left[h_t(\xi^t)^\top \beta_t \Phi_t(\xi^t)\right] + \frac{1}{N} \sum_{j \in [N]} \mathcal{Q}\left(\beta, \xi_j^T\right) \tag{6a}$$

$$\text{s.t. } A_1\beta_1 + C_1 x_1 = b_1, \tag{6b}$$

$$D_1\beta_1 + E_1 x_1 \geq d_1, \tag{6c}$$

$$x_1 \in \mathbb{R}^{p_1}, \tag{6d}$$

$$\beta_t \in \mathbb{R}^{q_t \times K_t}, \; t \in [T]. \tag{6e}$$

Once the sample is fixed, the SAA problem can be solved by any approach for solving the above, now deterministic, problem. In particular the L-shape decomposition algorithm [59] or a regularized variant [36,50] can be applied, with the further advantage that the subproblem obtained with fixed $\hat{\beta}$ decomposes by both scenario and stage due to the relationship $\mathcal{Q}(\beta, \xi^T) = \sum_{t \in [2,T]} Q_t(\beta, \xi^t)$. The coefficients on β_t in the objective function (6a) can also be estimated by sampling in case the terms $\mathbb{E}[h_{tj}(\xi^t)\Phi_{tk}(\xi^t)]$ cannot be computed efficiently.

If (i) there exists a $\bar{\beta}$ such that $\mathbb{E}[\mathcal{Q}(\beta, \xi^T)] < \infty$ for all β in a neighborhood of $\bar{\beta}$, and (ii) the set of optimal solutions to P-LDR-2S (4) is nonempty and bounded, then because $\mathcal{Q}(\cdot, \xi^T)$ is a convex function for all $\xi^T \in \Xi$, Theorem 5.4 of [55] applies and implies that

$$\hat{z}_N^{2S} \to z^{2S} \quad \text{with probability 1 as } N \to \infty \tag{7}$$

and also that the set of optimal solutions to (6) converges to the set of optimal solutions of P-LDR-2S (4).

Stronger results on the convergence of $\hat{z}_N^{2S}$ to z^{2S} require additional assumptions. For example, a central limit theorem result (e.g., Theorem 5.7 in [55]) can be obtained under the assumptions that $\mathbb{E}[\mathcal{Q}(\bar{\beta}, \xi^T)^2] < \infty$ for some $\bar{\beta}$, and that there exists a measurable function $f : \Xi \to \mathbb{R}_+$ such that $\mathbb{E}[f(\xi^T)^2]$ is finite and

$$\left|\mathcal{Q}(\beta, \xi^T) - \mathcal{Q}(\beta', \xi^T)\right| \leq f(\xi^T)\|\beta - \beta'\|$$

for all β, β' and almost every $\xi^T \in \Xi$. Bounds on the sample size required for (6) to yield an ϵ-optimal solution to P-LDR-2S (4) with probability at least $1 - \alpha$ are derived in [52,55–57]. These bounds scale linearly with the dimension of the first-stage variables, β and x_1 in this case. Dependence on the confidence α is $\ln(1/\alpha)$ so that high confidence can be achieved, but the dependence on ϵ is $O(1/\epsilon^2)$, which is why sampling is limited to obtaining "medium accuracy" solutions [57]. These stronger results all require, at least, that $\mathcal{Q}(\beta, \xi^T)$ is finite for every first-stage solution β and almost every $\xi^T \in \Xi$.

In order to facilitate the solution of P-LDR-2S (4) via a sampling procedure, we also consider adding additional constraints $\beta \in \mathcal{B} \subseteq \mathbb{R}^{\tau_P}$, where $\tau_P := \sum_{t\in[T]} q_t K_t$, to P-LDR-2S (4) [and to the SAA (6)]. For example, some of the convergence results require the first-stage feasible region to be bounded, in which case we may define $\mathcal{B}$ by limiting the absolute value of each component of β to be less than a large constant. If the constant is not chosen large enough, then this may degrade the quality of the solution obtained, but this could be detected after solving the SAA problem by determining if any of the bound constraints are tight. More significantly, most of the SAA convergence results require the following relatively complete recourse assumption on the set $\mathcal{B}$:

Assumption 1 For all $\beta \in \mathcal{B}$, $\mathcal{Q}(\beta, \xi^T) < +\infty$ $\mathbb{P}$-a.s..

If P-LDR-2S (4) already has relatively complete recourse, then we can take $\mathcal{B} = \mathbb{R}^{\tau_P}$, and hence impose no additional constraints. Otherwise, adding the constraints $\beta \in \mathcal{B}$ has the potential to make the approximation more conservative. Derivation of a set $\mathcal{B}$ that satisfies Assumption 1 is a difficult task in general. However, relatively complete recourse can often be achieved by appropriate modeling, e.g., by introducing "artificial" variables that allow violation of a constraint, where the violation amount is then penalized in the objective function. Derivation of a set $\mathcal{B}$ satisfying Assumption 1 may then be possible using ad hoc techniques. We provide an example of how this can be done in an inventory planning problem in Sect. 4.2. Another possibility, if assumptions A1-A3 hold, is to use the robust optimization techniques used in [35,57] to derive a tractable set $\mathcal{D}$ such that (β, Θ) is feasible to P-LDR (3) if and only if there are values of auxiliary variables w such that $(\beta, \Theta, w) \in \mathcal{D}$. Then, $\mathcal{B} = \text{proj}_\beta(\mathcal{D})$ would satisfy Assumption 1. This construction of $\mathcal{B}$ is more conservative than necessary, because it restricts β to values for which there is also a Θ that makes the static LDR policy defined in (2) feasible $\mathbb{P}$-a.s.. However, the resulting policy could still potentially be better (and for sure would not be worse) than the static LDR policy obtained from P-LDR (3), since enforcing $\beta \in \text{proj}_\beta(\mathcal{D})$ would not require the recourse decisions to follow an LDR policy (it only requires existence of a feasible LDR policy).

P-LDR-2S (4), with the additional constraints $\beta \in \mathcal{B}$, can also be approximately solved by stochastic approximation [48] or one of its robust extensions, e.g., [40,46], when Assumption 1 holds and $\mathcal{B}$ is bounded.

Finally, we remark that if we cannot derive a set $\mathcal{B}$ satisfying Assumption 1, results about sampling-based approximation of chance-constrained programs derived in [11] can be used to show that an optimal solution of the SAA problem (6) yields a policy that

is feasible for a large fraction of the random outcomes. This has been previously used in [6,7,60] when using sampling to approximately solve static approximations derived from finitely adaptable and piecewise-linear decision rules. Although the two-stage LDR policy itself is not necessarily feasible $\mathbb{P}$-a.s. in this case, in the next section we discuss how an approximate solution $\hat{\beta}$ could still be used to guide a feasible policy.

2.4 Feasible policies and upper bounds on z^{MSLP}

Let $(\hat{x}_1, \hat{\beta})$ be an approximate first-stage solution to P-LDR-2S (4). We discuss how such a solution can be used to obtain a feasible policy for the MSLP (1), which in turn can be used to estimate an upper bound on z^{MSLP}. We consider two possibilities for obtaining such a policy, depending on whether or not a set $\mathcal{B}$ satisfying Assumption 1 is used.

We first consider the case that $\hat{\beta} \in \mathcal{B}$ for a set $\mathcal{B}$ satisfying Assumption 1. In this case, $(\hat{x}_1, \hat{\beta})$ defines a feasible solution to P-LDR-2S (4) and a feasible two-stage LDR policy for MSLP. In particular, at stage t, if the current history is ξ^t, the state variable decisions are given by using $\hat{\beta}$ in the LDR (2a) and the recourse decisions are obtained by solving (5), again substituting $\hat{\beta}$ for β. As this solution defines a feasible policy, the expected cost of this solution provides an upper bound on z^{2S} and z^{MSLP}. The expected cost of the policy defined by $(\hat{x}_1, \hat{\beta})$ can be estimated by generating an independent sample of ξ^T, say $\{\xi_j^T\}_{j=1}^{N'}$, and computing

$$c_1^\top \hat{x}_1 + \sum_{t\in[T]} \mathbb{E}\big[h_t(\xi^t)^\top \hat{\beta}_t \Phi_t(\xi^t)\big] + \frac{1}{N'} \sum_{j\in[N']} \mathcal{Q}(\hat{\beta}, \xi_j^T).$$

Because $\hat{\beta}$ is fixed in this evaluation step, it would generally be computationally feasible to use $N' \gg N$. The values $\mathcal{Q}(\hat{\beta}, \xi_j^T)$ for $j \in [N']$ can also be used to construct a confidence interval on the objective value of $(\hat{x}_1, \hat{\beta})$, and hence a statistical upper bound on z^{MSLP}.

We next consider the case when we do not know $\hat{\beta} \in \mathcal{B}$ for a set $\mathcal{B}$ satisfying Assumption 1, so that we do not know a priori that the two-stage LDR defined by $\hat{\beta}$ defines a feasible policy. To construct a policy in this case, we make the following relatively complete recourse assumption for the original problem MSLP.

Assumption 2 For all $\xi^T \in \Xi$, and each $t \in [2, T]$, if the random vectors $\{(s_r(\xi^r), x_r(\xi^r)\}_{r\in[t-1]}$ satisfy the constraints of MSLP for $r \in [t-1]$, then there exists (s_t, x_t) that satisfies the constraints of MSLP in stage t:

$$A_t(\xi^t)s_t + C_t(\xi^t)x_t = b_t(\xi^t) - B_t(\xi^t)s_{t-1}(\xi^{t-1}),$$
$$(x_t, s_t) \in X_t(\xi^t).$$

In other words, this assumption states that in any stage t, for any value of the previous state variables $s_{t-1}(\xi^{t-1})$ that could be obtained from past realizations of the random outcomes and past feasible decisions, there always exists a feasible set of decisions in the current stage (see e.g., [29]).

Under Assumption 2, we can implement a policy which is guided by $\hat{\beta}$, which we refer to as a *state-target tracking* (*STT*) *policy*. Specifically, at stage $t = 1$, we implement the solution $x_1^{STT} = \hat{x}_1$ and $s_1^{STT} = \hat{\beta}_1$. Then, for each stage $t \in [2, T]$, we first observe ξ_t (thus, we have ξ^t), and then solve the problem (deterministic for this fixed ξ^t):

$$\min_{x_t, s_t} c_t(\xi^t)^\top x_t + h_t(\xi^t)^\top s_t + \rho \|s_t - \hat{\beta}_t \Phi_t(\xi^t)\| \tag{8a}$$

$$\text{s.t. } A_t(\xi^t)s_t + C_t(\xi^t)x_t = b_t(\xi^t) - B_t(\xi^t)s_{t-1}^{STT}(\xi^{t-1}),$$
$$(x_t, s_t) \in X_t(\xi^t), \tag{8b}$$

where $\rho \geq 0$ is a parameter of the policy and $\|\cdot\|$ is any norm, and let the optimal solution be $x_t^{STT}(\xi^t)$, $s_t^{STT}(\xi^t)$. For any $\xi^T \in \Xi$, all problems in this sequence are feasible when Assumption 2 holds, and hence this yields a feasible policy to MSLP. Observe that when $\rho = 0$, the policy reduces to a pure myopic policy that only considers the cost of decisions in each stage, without considering the impact of s_t on future costs. Using larger values of $\rho > 0$ has the effect of encouraging the decisions to be made in a way that keeps the state close to what would have been achieved if we could exactly follow the LDR policy defined by $\hat{\beta}$ on the state variables. The cost of the STT policy under a realization ξ of the stochastic process is

$$\sum_{t \in [T]} \left(c_t(\xi^t)^\top x_t^{STT}(\xi^t) + h_t(\xi^t)^\top s_t^{STT}(\xi^t) \right).$$

The expected cost of the STT policy is an upper bound on the optimal value of MSLP, and a confidence interval on this expected cost can be obtained by simulation with independent replications. We do not know an a priori upper bound on the optimality gap between the expected cost of the STT policy and the optimal value z^{MSLP}. However, the dual two-stage LDR discussed in Sect. 3 may be used to estimate a lower bound on z^{MSLP}, which can be used to provide an a posteriori statistical bound on the optimality gap of the STT policy. The value of the parameter ρ can be selected by estimating the expected cost of the policy under different values of ρ and choosing the most promising value, or by using optimization via simulation techniques [20,31]. For example, in our numerical experiments, we used a fixed relatively small sample ($N' = 100$), and applied a variant of a golden section algorithm to find a value of ρ that approximately minimizes the estimated cost given by this sample. See Sect. 5.2 for more details. Once the value of ρ is chosen, the expected cost of the resulting policy is evaluated using a larger sample. Note that using the STT policy, even the decisions $s_t^{STT}(\xi^t)$ may not necessarily have the form of an LDR. Thus, simulating this policy yields an estimate of an upper bound on z^{MSLP}, but not necessarily on z^{2S}.

3 Dual two-stage linear decision rules

In this section, we apply a two-stage LDR to the dual of MSLP, with the goal of obtaining lower bounds on the optimal value of MSLP. The dual of MSLP, which we

refer to as D-MSLP, is the problem (see [18]):

$$\max_{\lambda,\gamma} \mathbb{E}\Big[\sum_{t\in[T]} b_t(\xi^t)^\top \lambda_t(\xi^t) + d_t(\xi^t)^\top \gamma_t(\xi^t)\Big] \tag{9a}$$

$$\text{s.t. } \mathbb{E}\Big[B_{t+1}(\xi^{t+1})^\top \lambda_{t+1}(\xi^{t+1})\Big|\xi^t\Big]$$

$$+ A_t(\xi^t)^\top \lambda_t(\xi^t) + D_t(\xi^t)^\top \gamma_t(\xi^t) = h_t(\xi^t), \qquad t\in[T],\ \mathbb{P}\text{-a.s.}, \tag{9b}$$

$$C_t(\xi^t)^\top \lambda_t(\xi^t) + E_t(\xi^t)^\top \gamma_t(\xi^t) = c_t(\xi^t), \qquad t\in[T],\ \mathbb{P}\text{-a.s.}, \tag{9c}$$

$$\gamma_t(\xi^t) \geq 0, \qquad t\in[T],\ \mathbb{P}\text{-a.s.}, \tag{9d}$$

$$\lambda_t(\xi^t)\in\mathbb{R}^{m_t},\ \gamma_t(\xi^t)\in\mathbb{R}^{n_t}, \qquad t\in[T],\ \mathbb{P}\text{-a.s.}, \tag{9e}$$

where $B_{T+1}(\xi^{T+1}) = 0$. For $t \in [T]$, the dual decisions $\lambda_t(\cdot)$ (corresponding to constraints (1b) in MSLP) and $\gamma_t(\cdot)$ (corresponding to constraints (1c) in MSLP) are functions of the data ξ^t observed up to stage t. Weak duality holds for MSLP and D-MSLP, i.e., the optimal objective value of D-MSLP provides a lower bound on z^{MSLP}. Moreover, under some conditions, strong duality holds (i.e., optimal value of D-MSLP equals z^{MSLP}) [18], although we only require weak duality.

3.1 Static linear decision rules

In [35] it has been proposed to use a static LDR to obtain a tractable approximation of D-MSLP, and thus an efficiently computable lower bound on z^{MSLP}. Specifically, the idea is to require all the dual decisions to be an LDR, i.e.,

$$\lambda_t(\xi^t) = \Lambda_t \Phi_t(\xi^t), \tag{10a}$$

$$\gamma_t(\xi^t) = \Gamma_t \Phi_t(\xi^t), \tag{10b}$$

where $\Lambda_t \in \mathbb{R}^{m_t\times K_t}$, $\Gamma_t \in \mathbb{R}^{n_t\times K_t}$, for all $t \in [T]$ are the parameters of the decision rule. Imposing (10) yields the following static approximation of D-MSLP, which we call D-LDR:

$$\max_{\Lambda,\Gamma} \mathbb{E}\Big[\sum_{t\in[T]} b_t(\xi^t)^\top \Lambda_t \Phi_t(\xi^t) + d_t(\xi^t)^\top \Gamma_t \Phi_t(\xi^t)\Big]$$

$$\text{s.t. } \mathbb{E}\Big[B_{t+1}(\xi^{t+1})^\top \Lambda_{t+1}\Phi_{t+1}(\xi^{t+1})\Big|\xi^t\Big]$$

$$+ A_t(\xi^t)^\top \Lambda_t \Phi_t(\xi^t) + D_t(\xi^t)^\top \Gamma_t \Phi_t(\xi^t) = h_t(\xi^t), \qquad t\in[T],\ \mathbb{P}\text{-a.s.},$$

$$C_t(\xi^t)^\top \Lambda_t \Phi_t(\xi^t) + E_t(\xi^t)^\top \Gamma_t \Phi_t(\xi^t) = c_t(\xi^t), \qquad t\in[T],\ \mathbb{P}\text{-a.s.},$$

$$\Gamma_t \Phi_t(\xi^t) \geq 0, \qquad t\in[T],\ \mathbb{P}\text{-a.s.},$$

$$\Lambda_t \in \mathbb{R}^{m_t\times K_t},\ \Gamma_t \in \mathbb{R}^{n_t\times K_t}, \qquad t\in[T]. \tag{11}$$

We refer to the optimal value of D-LDR as v^{LDR}. The semi-infinite program D-LDR (11) can be reformulated as an efficiently solvable LP if assumptions A1–A3 stated

in Sect. 2 hold, the problem MSLP is strictly feasible, and the following additional assumption holds [35]:

A4. The conditional expectation $\mathbb{E}(\xi^T|\xi^t)$ is almost surely linear in ξ^t for all $t \in [T]$ (e.g., this occurs when $\{\xi_t\}_{t\in[T]}$ are mutually independent, known as *stage-wise independence*).

If assumption A3 is replaced by an assumption that Ξ is described by conic inequalities, D-LDR (11) can be reformulated as a conic program [23].

3.2 Two-stage linear decision rules

Examining the structure of D-MSLP, given in (9), we observe that if the values $\lambda_t(\xi^t)$ are fixed, then (9) decomposes by stage. We thus propose to apply an LDR only to the λ variables, leaving the decision variables γ as recourse variables. Imposing the LDR of (10a) collapses D-MSLP into the following 2SLP, which we refer to as D-LDR-2S:

$$\begin{aligned} v^{2S} := \max_{\gamma_1,\Lambda}\ & d_1^\top\gamma_1 + \sum_{t\in[T]} \mathbb{E}\big[b_t(\xi^t)^\top\Lambda_t\Phi_t(\xi^t)\big] + \mathbb{E}[\mathcal{G}(\Lambda,\xi^T)] \\ \text{s.t. } & \mathbb{E}\big[B_2(\xi^2)^\top\Lambda_2\Phi_2(\xi^2)\big] + A_1^\top\Lambda_1 + D_1^\top\gamma_1 = h_1, \\ & C_1^\top\Lambda_1 + E_1^\top\gamma_1 = c_1, \\ & \gamma_1 \in \mathbb{R}^{n_1}_+, \\ & \Lambda_t \in \mathbb{R}^{m_t\times K_t}, \quad t\in[T], \end{aligned} \tag{12}$$

where we have dropped the dependence on $\xi_1 \equiv 1$ on the first-stage decision variables. Here, $\mathcal{G}(\Lambda,\xi^T)$ is the second-stage value function,

$$\mathcal{G}(\Lambda,\xi^T) := \sum_{t\in[2,T]} G_t(\Lambda,\xi^t) \tag{13}$$

where for each $t \in [2,T]$

$$G_t(\Lambda,\xi^t) := \max_{\gamma_t}\ d_t(\xi^t)^\top\gamma_t \tag{14a}$$

$$\text{s.t. } D_t(\xi^t)^\top\gamma_t = h_t(\xi^t) - A_t(\xi^t)^\top\Lambda_t\Phi_t(\xi^t) - \mathbb{E}\Big[B_{t+1}(\xi^{t+1})^\top\Lambda_{t+1}\Phi_{t+1}(\xi^{t+1})\,\Big|\,\xi^t\Big], \tag{14b}$$

$$E_t(\xi^t)^\top\gamma_t = c_t(\xi^t) - C_t(\xi^t)^\top\Lambda_t\Phi_t(\xi^t), \tag{14c}$$

$$\gamma_t \in \mathbb{R}^{n_t}_+. \tag{14d}$$

The following proposition, immediate from the definitions of the associated problems, summarizes the relationship between the optimal values of MSLP, D-LDR (11), and D-LDR-2S (12).

Proposition 2 *The following inequalities hold:*

$$z^{MSLP} \geq v^{2S} \geq v^{LDR}.$$

As D-LDR-2S (12) is a 2SLP, the discussion in Sect. 2.3 of methods to obtain an approximate solution to P-LDR-2S (4) applies also to D-LDR-2S (12). In particular, one possibility is to obtain an i.i.d. sample $\{\xi_j^T\}_{j=1}^N$ of ξ^T and solve the SAA problem:

$$\hat{v}_N^{2S} := \max_{\gamma_1, \Lambda} d_1^\top \gamma_1 + \sum_{t\in[T]} \mathbb{E}\big[b_t(\xi^t)^\top \Lambda_t \Phi_t(\xi^t)\big] + \frac{1}{N}\sum_{j\in[N]} \mathcal{G}(\Lambda, \xi_j^T) \tag{15a}$$

$$\text{s.t. } \mathbb{E}\big[B_2(\xi^2)^\top \Lambda_2 \Phi_t(\xi^2)\big] + A_1^\top \Lambda_1 + D_1^\top \gamma_1 = h_1, \tag{15b}$$

$$C_1^\top \Lambda_1 + E_1^\top \gamma_1 = c_1, \tag{15c}$$

$$\gamma_1 \in \mathbb{R}^{n_1}_+, \tag{15d}$$

$$\Lambda_t \in \mathbb{R}^{m_t \times K_t}, \quad t \in [T]. \tag{15e}$$

As in the primal, note that the SAA problem can be solved by decomposition algorithms as the second-stage problem decomposes by both scenario and by stage due to the relationship (13). The expected value coefficients in the objective and constraints (15b) may be further estimated by sampling in case they cannot be computed efficiently.

3.3 Obtaining lower bounds on z^{MSLP}

Next, we discuss how to use an approximate solution of D-LDR-2S (12) to estimate a lower bound on z^{MSLP}. As in the primal case, in order to assure that we obtain a two-stage LDR policy that is feasible for all possible realizations, we consider the possibility of adding a set of constraints $\Lambda \in \mathcal{L} \subseteq \mathbb{R}^{\tau_D}$ to D-LDR-2S (12) and its SAA counterpart (15) where $\tau_D := \sum_{t\in[T]} K_t m_t$. The following assumption on $\mathcal{L}$ assures that the problem D-LDR-2S (12) has relatively complete recourse when the constraints $\Lambda \in \mathcal{L}$ are enforced.

Assumption 3 For all $\Lambda \in \mathcal{L}$, $\mathcal{G}(\Lambda, \xi^T) > -\infty$ $\mathbb{P}$-a.s..

The following assumption provides a sufficient condition under which the set $\mathcal{L} = \mathbb{R}^{\tau_D}$ satisfies Assumption 3 (i.e., no additional constraints are necessary).

Assumption 4 The set $X_t(\xi^t)$ is bounded for all $t \in [T]$ and $\mathbb{P}$ almost all $\xi^T \in \Xi$.

A special case of this assumption occurs when x and s variables have explicit upper and lower bounds. An important feature of this assumption is that the sets $X_t(\xi^t)$ are not required to be uniformly bounded. For example, bounds on the decision variables of the form $0 \leq x_t(\xi^t) \leq M(\xi^t)$ (and similarly for s variables), are sufficient for satisfying this assumption, even if $M(\xi^t)$ is not bounded over $\xi^T \in \Xi$.

Proposition 3 *If Assumption 4 is satisfied, then $\mathcal{L} = \mathbb{R}^{\tau_D}$ satisfies Assumption 3.*

Proof We show that for any given $\Lambda \in \mathbb{R}^{\tau_D}$ and $\xi^T \in \mathbb{R}^{\ell^T}$, (14) is feasible for any $t \in [T]$. Let $R_b(\xi^t)$ and $R_c(\xi^t)$ denote the right-hand sides of the constraints (14b) and (14c), respectively. Then, the dual of (14)

$$\begin{aligned} \min\ & R_b(\xi^t)^\top s_t(\xi^t) + R_c(\xi^t)^\top x_t(\xi^t) \\ \text{s.t. } & (x_t(\xi^t), s_t(\xi^t)) \in X_t(\xi^t), \\ & x_t(\xi^t) \in \mathbb{R}^{p_t},\ s_t(\xi^t) \in \mathbb{R}^{q_t}, \end{aligned}$$

is bounded due to Assumption 4. It is also feasible as MSLP is assumed to be feasible. This implies that (14) cannot be infeasible. □

Now, suppose we have an approximate solution $(\hat{\gamma}_1, \hat{\Lambda})$ to D-LDR-2S (12), where $\hat{\Lambda} \in \mathcal{L}$ for some set $\mathcal{L}$ that satisfies Assumption 3. In this case, $(\hat{\gamma}_1, \hat{\Lambda})$ defines a feasible solution to D-LDR-2S (12), and hence its objective value provides a lower bound on v^{2S}, and hence is a lower bound on z^{MSLP}. The objective value of $(\hat{\gamma}_1, \hat{\Lambda})$ can be estimated by generating an independent sample of ξ^T, say $\{\xi_j^T\}_{j=1}^{N'}$ (where possibly $N' \gg N$), and computing

$$d_1^\top \hat{\gamma}_1 + \sum_{t\in[T]} \mathbb{E}\big[b_t(\xi^t)^\top \hat{\Lambda}_t \Phi_t(\xi^t)\big] + \frac{1}{N'} \sum_{j\in[N']} \mathcal{G}(\hat{\Lambda}, \xi_j^T).$$

The values $\mathcal{G}(\hat{\Lambda}, \xi_j^T)$ for $j \in [N']$ can also be used to construct a confidence interval on the objective value of $(\hat{\gamma}_1, \hat{\Lambda})$, and hence a statistical lower bound on z^{MSLP}.

We close this section by discussing an approach for estimating the gap between a primal two-stage LDR policy defined by $(\hat{x}_1, \hat{\beta})$ and a dual two-stage LDR policy defined by $(\hat{\gamma}_1, \hat{\Lambda})$. Following [39], the motivation is that if the same sample (common random numbers) is used in estimating the upper and lower bounds, then the variance of the gap estimator can be reduced if the upper and lower bound sample estimates are positively correlated. Specifically, given a sample $\{\xi_j^T\}_{j=1}^{N'}$, the gap observations are then calculated as

$$\begin{aligned} \text{Gap}_j = & \Big[c_1^\top \hat{x}_1 + \sum_{t\in[T]} \mathbb{E}\big[h_t(\xi^t)^\top \hat{\beta}_t \Phi_t(\xi^t)\big] + \mathcal{Q}(\hat{\beta}, \xi_j^T)\Big] \\ & - \Big[d_1^\top \hat{\gamma}_1 + \sum_{t\in[T]} \mathbb{E}\big[b_t(\xi^t)^\top \hat{\Lambda}_t \Phi_t(\xi^t)\big] + \mathcal{G}(\hat{\Lambda}, \xi_j^T)\Big], \end{aligned}$$

for $j \in [N']$. These values can then be used to construct a confidence interval on the gap.

4 Illustrative example: inventory planning

We first present a numerical example on an inventory planning problem to investigate the performance of two-stage LDR policies and bounds, in comparison to static LDR policies and bounds.

4.1 Problem description

We consider a variation of the inventory planning problem used for numerical illustration in [4,35]. The system consists of I factories and a single product type, and the goal is to meet demands over the planning horizon at minimum expected cost. The model is stated as follows:

$$\min \mathbb{E}\left[\sum_{t\in[T]}\sum_{i\in[I]} c_{it}x_{it}(\xi^t)\right] \tag{16a}$$

$$\text{s.t. } s_{t-1}(\xi^t) - s_t(\xi^t) + \sum_{i\in[I]} x_{it}(\xi^t) = \xi_t, \qquad t \in [T],\ \mathbb{P}\text{-a.s.}, \tag{16b}$$

$$\underline{s} \le s_{it}(\xi^t) \le \bar{s} \qquad t \in [T],\ i \in [I],\ \mathbb{P}\text{-a.s.}, \tag{16c}$$

$$0 \le x_{it}(\xi^t) \le \bar{x}_i \qquad t \in [T],\ i \in [I],\ \mathbb{P}\text{-a.s.}. \tag{16d}$$

Here, ξ_t is a scalar random variable representing demand for the product in each $t \in [T]$. The recourse decision variable $x_{it}(\xi^t)$ determines amount of the product to produce in factory i at stage t, while the state variable $s_t(\xi^t)$ represents the inventory level at the end of stage t. Constraints (16b) are the inventory balance equations, (16c) limit the inventory level to be between lower bound $\underline{s}$ and upper bound $\bar{s}$, and (16d) are the limits on production in each stage to be at most $\bar{x}_i$.

The model in [4,35] also has a constraint on the total amount that can be produced from any single factory over all the stages in the planning horizon. Modeling this constraint in our standard model format requires introducing an additional state variable for each factory i, representing the cumulative amount of production from each factory. Imposing an LDR on that state variable would in turn imply that the variables $x_{it}(\xi^t)$ also follow an LDR, and hence for that model the static and two-stage LDR policies are identical. This illustrates an example where there is no benefit to using a two-stage LDR over a static LDR. In the version we consider, the $x_{it}(\xi^t)$ are still flexible when the state variables $s_t(\xi^t)$ follow an LDR, and hence there is potential for a two-stage LDR to yield better solutions.

Following the data in [4,35], we consider an instance with $I = 3$, $\underline{s} = 500$, $\bar{s} = 2000$, and $\bar{x}_i = 567$ for $i \in [I]$. The random demand ξ_t in stage $t \in [T]$ is uniformly distributed in the interval $\Xi_t = [(1-\theta)\xi^*\zeta_t, (1+\theta)\xi^*\zeta_t]$, where $\theta = 0.3$ is the variability parameter, $\xi^* = 1000$ is the nominal demand, and $\zeta_t = 1 + (1/2)\sin(\pi(t-1)/12)$ is the seasonality factor. Finally, the cost coefficients are defined as $c_{it} = \alpha_i\zeta_t$, where $\alpha_1 = 1$, $\alpha_2 = 1.5$, and $\alpha_3 = 2$.

4.2 Implementation details

For both the static and two-stage LDR policies, we use the standard basis functions, ξ^t, in stage t. For the static LDR, we implemented the deterministic reformulations proposed in [35,57] to obtain upper bounds (with a primal LDR policy) and lower bounds (with a dual LDR policy).

For the primal two-stage LDR policy, we first observe that this problem as stated does not satisfy relatively complete recourse, Assumption 2, although we remark that this assumption is satisfied in a slightly modified version of the problem in which variables are introduced to allow some amount of demand to go unserved, with a large penalty. Rather than making this modification, however, we demonstrate how for this problem a set of constraints satisfying Assumption 1 can be derived. Using the standard basis functions, the state variables $s_t(\xi^t)$ take the form

$$s_t(\xi^t) = \beta_t \xi^t$$

where $\beta_t \in \mathbb{R}^{1\times t}$. Thus, the constraints (16c) take the form

$$\underline{s} \leq \beta_t \xi^t \leq \bar{s}, \quad t \in [T], \forall \xi_t \in [1-\theta\xi^*\zeta_t, (1+\theta)\xi^*\zeta_t].$$

These constraints can be reformulated with deterministic linear constraints in an extended variable space using standard robust optimization techniques [5]. To ensure $\beta_t, t \in [T]$ are selected such that constraints (16b) can be satisfied for some $x_t(\xi^t), i \in [I]$ satisfying (16d), it is sufficient to enforce

$$\xi_t - \sum_{i\in[I]} \bar{x}_i \leq \beta_{t-1}\xi^{t-1} - \beta_t \xi^t \leq \xi_t, \quad t \in [T], \forall \xi_t \in [(1-\theta)\xi^*\zeta_t, (1+\theta)\xi^*\zeta_t],$$

where the lower bound is based on the maximum total production and the upper bound is based on the minimum total production in each period. Again, these constraints can be reformulated as deterministic linear constraints using robust optimization techniques.

For both the primal and dual two-stage LDR policy, we use 250 scenarios to construct an SAA, and solve the resulting problem by explicitly solving the determistic equivalent formulation. Given the resulting LDR coefficients, we then use an independent sample of 10^5 scenarios to evaluate the quality of the primal policy and dual bound.

4.3 Results

Table 1 provides the results comparing the bounds obtained for this instance, for varying values of $T = 2, \ldots, 10$. The columns under Static LDR provide the lower bound (LB), upper bound (UB), and optimality gap [Gap (%)], respectively, where optimality gap for an instance is calculated as (UB − LB)/UB. For the two-stage LDR policy, 95% confidence intervals for the lower bound (LB CI) and upper bound (UB CI) are provided, along with an estimate of the optimality gap, which is computed by using the lower end of the lower bound confidence interval and the upper end of the upper bound confidence interval. We find that the two-stage LDR policy can yield modestly better lower bound estimates than the static LDR lower bounds, and somewhat more significantly better primal policies. In terms of solution time, the static LDR lower and upper bounds were computed very quickly, less than 0.02 s in all cases. For the two-stage LDR policies, solving the two SAA problems took at most 3.98 s, and evaluating

Table 1 Comparison of static and two-stage LDR policies for inventory problem

	Static LDR			2S LDR		
T	LB	UB	Gap (%)	LB CI	UB CI	Gap (%)
2	1972.4	2026.0	2.65	1974.4 ± 2.7	1993.9 ± 1.9	1.21
3	3825.0	3940.2	2.92	3831.6 ± 4.0	3856.1 ± 3.2	0.82
4	6089.8	6345.0	4.02	6102.4 ± 5.5	6146.9 ± 4.7	0.89
5	8664.4	9021.3	3.96	8669.1 ± 6.6	8737.6 ± 5.9	0.93
6	11482.4	11975.0	4.11	11515.2 ± 10.2	11594.8 ± 7.3	0.84
7	14431.1	15076.3	4.28	14482.3 ± 12.3	14618.8 ± 8.6	1.08
8	17431.6	18200.3	4.22	17527.4 ± 13.7	17660.4 ± 9.9	0.89
9	20251.8	21147.9	4.24	20326.2 ± 14.9	20535.3 ± 10.9	1.14
10	22764.8	23738.3	4.10	22809.5 ± 15.0	23067.0 ± 11.5	1.23

the bounds took at most 5.17 s. Thus, as expected, in this case where the assumptions required for obtaining a deterministic formulation of static LDR apply, the solution time for the static LDR policy are significantly faster than for the two-stage LDR. On the other hand, the solution times for the two-stage LDR policy were still modest, and yielded better policies.

5 Illustrative example: capacity expansion

We next consider a capacity expansion problem. On this problem, we again compare the two-stage LDR policies and bounds to those obtained from static LDR policies, and also compare to the policy and lower bound obtained from using the SDDP algorithm.

5.1 Problem description

We consider a variant of the stochastic capacity expansion problem given in [15]. We wish to determine an investment schedule over T stages for the installation of new capacities of I different power generation technologies, together with some operational decisions to meet demand for power over time. The demand is modeled by a load duration curve, which is approximated by partitioning each stage into J segments (of possibly different length). The demand corresponding to segment $j \in [J]$ in $t \in [T]$ is denoted by d_{tj}. The amount of new capacity of technology $i \in [I]$ added in stage $t \in [T]$ is represented by u_{ti}^{+}, and is assumed to be available for use immediately, i.e., at the beginning of stage t. The unit cost of u_{ti}^{+} is denoted by $c_{ti}^{u^+}$. The state variable s_{ti} represents the current installed capacity of technology $i \in [I]$ in the beginning of stage $t \in [T]$, which incurs holding cost of c_{ti}^{s} per unit. We assume that it is possible to discard (i.e., remove) some capacity of $i \in [I]$ in $t \in [T]$, denoted by u_{ti}^{-}, at a (possibly zero) unit cost of $c_{ti}^{u^-}$. The operating level of $i \in [I]$ at $t \in [T]$ for meeting the demand in segment $j \in [J]$ is represented by the decision variable x_{tij}, while

the amount of unsatisfied demand is represented as z_{tj}, whose unit costs are c^x_{tij} and c^z_{tj}, respectively. Then, the stochastic capacity expansion problem is formulated as an MSLP as follows:

$$\min \mathbb{E} \sum_{t\in[T]} \Bigg[\sum_{i\in[I]} \Big(c^{u^+}_{ti} u^+_{ti}(\xi^t) + c^{u^-}_{ti} u^-_{ti}(\xi^t) + c^s_{ti} s_{ti}(\xi^t) + \sum_{j\in[J]} c^x_{tij} x_{tij}(\xi^t) \Big) + \sum_{j\in[J]} c^z_{tj} z_{tj}(\xi^t) \Bigg] \tag{17a}$$

$$\text{s.t. } s_{ti}(\xi^t) - s_{t-1,i}(\xi^{t-1}) - u^+_{ti}(\xi^t) + u^-_{ti}(\xi^t) = 0, \quad t \in [T],\ \mathbb{P}\text{-a.s.},\ i \in [I], \tag{17b}$$

$$s_{ti}(\xi^t) - x_{tij}(\xi^t) \geq 0, \quad t \in [T],\ \mathbb{P}\text{-a.s.},\ i \in [I],\ j \in [J], \tag{17c}$$

$$\sum_{i\in[I]} x_{tij}(\xi^t) + z_{tj}(\xi^t) \geq d_{tj}(\xi^t), \quad t \in [T],\ \mathbb{P}\text{-a.s.},\ j \in [J], \tag{17d}$$

$$0 \leq z_{tj}(\xi^t) \leq d_{tj}(\xi^t), \quad t \in [T],\ \mathbb{P}\text{-a.s.},\ j \in [J], \tag{17e}$$

$$0 \leq u^+_{ti}(\xi^t) \leq M^{u^+}_{ti}, \quad t \in [T],\ \mathbb{P}\text{-a.s.},\ i \in [I], \tag{17f}$$

$$0 \leq u^-_{ti}(\xi^t) \leq M^{u^-}_{ti}, \quad t \in [T],\ \mathbb{P}\text{-a.s.},\ i \in [I], \tag{17g}$$

$$0 \leq s_{ti}(\xi^t) \leq M^s_{ti}, \quad t \in [T],\ \mathbb{P}\text{-a.s.},\ i \in [I], \tag{17h}$$

$$x_{tij}(\xi^t) \geq 0, \quad t \in [T],\ \mathbb{P}\text{-a.s.},\ i \in [I],\ j \in [J]. \tag{17i}$$

The objective function (17a) minimizes the expected total cost. The constraints (17b) are the only state equations, which keep track of the available capacity of each technology. Constraints (17c) limit the operating levels to the available capacity level, while (17d) ensure that either demand is met, or unmet demand is recorded in the z_{tj} variable values. Constraints (17e)–(17h) represent the bounds on the shortfall, installation, removal, and inventory level variables, respectively. Note that (17h) constitute upper bounds also on the x variables due to (17c), and thus this formulation satisfies Assumption 4. In addition, we assume that $M^{u^-}_{ti} \geq M^s_{t-1,i}$ which ensures that this formulation satisfies relatively complete recourse, Assumption 2, since at stage t, given any feasible value of $s_{t-1,i}(\xi^{t-1})$, it is feasible to set $u^-_{ti}(\xi^t) = s_{t-1,i}(\xi^{t-1})$ for $i \in [I]$, $z_{tj}(\xi^t) = d_{tj}(\xi^t)$ for $j \in [J]$ and all remaining variables to zero.

To the extent possible, we use data from [15], which focuses on a German system, although we extend their 3-stage example to $T = 5, 10, 15, 20$. In [15], there are $I = 3$ technologies (coal-fired power plant, combined cycle gas turbine and open cycle gas turbine). Each stage is divided into $L = 8$ periods, and $W = 5$ wind regimes are considered for each period. We model this as $J = LW = 40$ segments at each stage, corresponding to each period/wind regime pair. For $t \geq 2$, the demand corresponding to the segment $j = (\ell, w) \in [L] \times [W]$ is modeled as

$$d_{tj}(\xi^t) = \max \left\{ d_{0,\ell} \prod_{r=2}^{t} \xi^g_r - \eta_w K^w_t \prod_{r=2}^{t} \xi^w_r, 0 \right\}$$

where $d_{0,\ell}$ is the base demand value of period ℓ, ξ_t^g is a random variable reflecting the demand growth of stage t, η_w is the parameter denoting the wind efficiency, K_t^w is the wind power generation target, and ξ_t^w is a random variable representing the growth in the wind power generation in stage t. The values of $d_{0,\ell}$ and η_w are from Tables 2 and 3 of [15], and are reproduced in "Appendix A". We use $K_2^w = 36.64$ and $K_t^w = 45.75$ for all $t \geq 3$. We assume ξ_t^g has lognormal distribution with $\mu = 0.2$ and $\sigma = 0.1 + 0.01t$, and ξ_t^w has lognormal distribution with $\mu = 0.15$ and $\sigma = 0.25 + 0.025t$. For the first stage, we use the deterministic demand values of $d_{1,j=(\ell,w)} = d_{0,\ell}\mathbb{E}[\xi_1^g] - \eta_w\mathbb{E}[\xi_1^w] = 1.229d_{0,\ell} - 1.207\eta_w$. The units of all demands (and all primal decision variables) are gigawatts.

We assume there are no holding costs and no costs for removing capacity, i.e., we use $c_{ti}^{u^-} = c_{ti}^s = 0$ for all $i \in [I]$, $t \in [T]$. We use discounting to determine the other costs, setting $c_{ti}^{u^+} = 5\iota_i/1.1^t$, $c_{ti,j=(\ell,w)}^x = 0.001c_i\tau_\ell\tau_w/1.1^t$ and $c_{t,j=(\ell,w)}^z = \tau_\ell\tau_w/1.1^t$ where the values of the annualized costs ι_i, operation costs c_i, τ_ℓ and τ_w values are from [15] (see "Appendix A"). All costs are in million of Euros. Finally, we assume the maximum installation per stage is a constant $M_{ti}^{u^+} = C$, and derive redundant upper bounds on s and u^-, i.e., $M_{ti}^s = \sum_{r=1}^t M_{ri}^{u^+}$ and $M_{ti}^{u^-} = M_{t-1,i}^s$. In our experiments we consider two different sets of instances defined using $C = 50$ and $C = 100$.

5.2 Implementation details

We compare the primal and dual bounds obtained using two-stage and static LDR. For the LDR basis functions, for each $t \in [T]$, we let $K_t = 3$, and

$$\Phi_{t1}(\xi^t) = 1, \quad \Phi_{t2}(\xi^t) = \prod_{r=2}^{t} \xi_r^g, \quad \Phi_{t3}(\xi^t) = \prod_{r=2}^{t} \xi_r^w.$$

Because assumptions A3 and A4 do not hold for this problem (the random variables do not have bounded support and $\mathbb{E}(\xi^T|\xi^t)$ is not linear in ξ^t), the reformulation approach from [13,35,57] used for the static LDR cannot be applied to solve P-LDR (3) and D-LDR (11). We therefore use a sampling strategy to approximately solve these problems. Specifically, the sample approximations of P-LDR (3) and D-LDR (11) are identical to P-LDR (3) and D-LDR (11), respectively, except that the infinite set of constraints $\mathbb{P}$-a.s. are replaced by the finite set corresponding to the sample. Approximate solutions to P-LDR-2S (4) and D-LDR-2S (12) are obtained by solving the SAA problems (6) and (15), respectively. We solve all sample approximations using the same sample of size $N = 150T$.

Models P-LDR-2S (4) and D-LDR-2S (12) corresponding to model (17) and its dual are given in "Appendix B". Although the MSLP given in (17) has relatively complete recourse (Assumption 2), its two-stage primal approximation P-LDR-2S (4) does not have relatively complete recourse. Thus, for obtaining a primal policy using P-LDR-2S (4), we implement the STT policy proposed in Sect. 2.4. The parameter ρ is determined by conducting a golden section search using a fixed evaluation sample of 100 scenarios. Specifically, starting with a lower bound of 0 and an upper bound of

1000, a golden section search is performed in which a simulation of the STT policy with these 100 scenarios is used to guide the search. In case the current upper estimate of ρ in the search process yields the minimum estimated cost, the search is restarted with the new lower estimate set to the current upper estimate of ρ, and the new upper estimate set to four times the current upper estimate. The search is terminated when either the upper estimate and lower estimates of ρ differ by less than 1.0, or the difference in the estimated objective values between the upper and lower estimates is less than 10^{-6} times the sum of the two objective estimates. The resulting value of ρ is then used in the simulation with $N' = 5000T$ replications to estimate the quality of the resulting policy. The time to select ρ in this process was vastly dominated by the time to simulate the policy, but is included in all numerical results that follow.

Since Assumption 4 is satisfied, any solution to the SAA problem (15) provides a feasible solution to D-LDR-2S (12), and hence evaluating this solution using N' independent replications yields a statistical lower bound on z^{MSLP}. In our experiments we use $N' = 5000T$ scenarios for estimating the value of this policy. Unfortunately, the sample approximations of P-LDR (3) and D-LDR (11) do not yield policies (primal or dual) that are feasible under all scenarios. Thus, when evaluating these policies with the independent replications, we report two measures: the average objective value over scenarios that are feasible, and the fraction of scenarios that are infeasible. By averaging only over feasible scenarios, these estimates are optimistically biased, i.e., they underestimate the bound on the primal problem, and overestimate the bound for the dual problem. As a result, these estimates do not necessarily provide valid (statistical) upper and lower bounds on z^{MSLP}, but we use them to provide a "best case" estimate when comparing to the estimates obtained from the two-stage LDR policies.

All of our numerical results are carried out using IBM ILOG CPLEX 12.6 as the LP solver. We perform all experiments using a single thread on a Mac OS X 10.12 with 4 GHz Intel Core i7 CPUs and 16 GB RAM.

The SAA problems (6) and (15) are solved with a sample size of $N = 150T$. The primal SAA problem (6) is solved with Benders decomposition, using a single aggregate cut per time-stage. The Benders decomposition is run until no violated cuts are found. The dual SAA problem (15) is solved with the bundle-level method [19,37]. The level method for solving the dual SAA problem is terminated when the relative gap between the lower and upper bounds is less than 10^{-5}. The details of the Benders decomposition and the level method are provided in "Appendices B.1 and B.2", respectively.

5.3 Comparison between static and two-stage LDR

Tables 2 and 3 present 95% confidence intervals (CIs) on the expected costs of primal policies and dual lower bounds, respectively, obtained using the two-stage and static LDR policies. These results are reported only for the instances having $T = 5, 10$. The costs are normalized such that for each instance, the estimated lower bound obtained by the two-stage LDR policy has value 100.0. In these tables, the CIs are presented with their mean and half-width ($\pm$). In Table 2, the upper end of the CI the two-stage

Table 2 Confidence intervals for expected costs of the primal policies

		2S LDR		Static LDR			
C	T	Mean	±	Mean	±	Inf. (%)	%UΔ
50	5	100.8	0.3	138.0	0.2	3.0	36.7
	10	114.6	0.6	232.9	0.4	3.8	102.7
100	5	101.3	0.3	138.4	0.2	2.9	36.5
	10	109.2	0.4	195.8	0.3	4.0	78.8

Table 3 Confidence intervals for expected costs of the dual policies

		2S LDR		Static LDR			
C	T	Mean	±	Mean	±	Inf. (%)	%LΔ
50	5	100.0	0.3	98.7	0.3	2.4	− 1.3
	10	100.0	0.5	97.1	0.4	3.5	− 2.9
100	5	100.0	0.3	100.0	0.3	2.1	0.0
	10	100.0	0.4	98.1	0.4	3.2	− 1.8

LDR policy is an upper bound on the expected cost of using that policy, and hence is a statistical upper bound on z^{MSLP}. We also report under column 'Inf. (%)' the percentage of the scenarios (out of 5000T evaluated scenarios) for which the static LDR policy is infeasible. To give an idea of the relative improvement in the expected policy cost obtained by using the two-stage LDR, the column '%UΔ', presents the percentage increase in the upper bound on the cost obtained with the static policy over the upper bound on the cost obtained with the two-stage LDR policy. We observe that the expected cost of the static LDR policy is between 36 and 102% higher than that of the static LDR policy, with the most significant differences occurring with larger time stages. We also observe that the static LDR policy is frequently infeasible. Finally, although not presented in the table, we find that the estimated expected cost of the STT policies was consistently similar (within 2.3%) to the objective value of the SAA problem (6), indicating that the STT policy is effectively "tracking" the obtained two-stage LDR policy.

Considering the CIs of the lower bounds obtained from using two-stage and static LDR policies presented in Table 3, we again find that the static LDR policy is often infeasible. Column '%LΔ' presents the percentage difference between the lower end of the CI obtained from the static and two-stage LDR policies, and indicates that the (95% confidence) lower bounds obtained by the static LDR range from being similar to 2.9% lower than those obtained by the two-stage LDR policy.

5.4 Comparison with SDDP

We next compare the two-stage LDR approximation with the results obtained using SDDP. In order to apply SDDP, we need a formulation having a finite number of scenarios per stage and stage-wise independent random variables. We obtain a model with stage-wise independence by introducing new state variables v_t^g and v_t^w to repre-

sent $\prod_{r=2}^{t} \xi_r^g$ and $\prod_{r=2}^{t} \xi_r^w$, respectively, which is implemented by adding the state equations

$$v_t^g = \xi_t^g v_{t-1}^g, \quad v_t^w = \xi_t^w v_{t-1}^w, \qquad t \in [2, T] \tag{18}$$

and $v_1^g = v_1^w = 1$. With these state variables, the demand in stage $t \geq 2$ is then represented as $\max\{d_{0,\ell} v_t^g - \eta_w K_t^w v_t^w, 0\}$. In particular, the right-hand side of constraints (17d) are replaced with the expression $d_{0,\ell} v_t^g - \eta_w K_t^w v_t^w$, and the redundant upper bounds $z_{tj}(\xi^t) \leq d_{tj}(\xi^t)$ in (17e) are removed. Thus the only random variables appearing in stage t constraints are ξ_t^g and ξ_t^w, which are stage-wise independent. We use SAA to construct scenario trees with a finite number of outcomes per stage. In an SAA problem, we approximate the joint distribution of ξ_t^g and ξ_t^w with 200 scenarios, obtained by independent Monte Carlo sampling. Note that the SAA approximation has 200^{T-1} total sample paths. The number of scenarios per stage was determined based on initial experiments solving multiple replications of the SAA problem, and was found to provide a good trade-off between difficulty in solving each individual SAA problem by SDDP and the variability of the SAA estimates. The optimal value of an SAA problem is random because it is defined by a random sample. The expected value of this optimal value is a lower bound on the true optimal value [39]. Thus, by solving multiple SAA problems with independent samples, a confidence interval on the expected value of the SAA problem, and hence a lower bound on the true optimal value, can be obtained. We thus generate 25 independently generated SAA problems, and for each one we obtain a lower bound by solving it with SDDP for a limited time. These replication values are then used to construct a confidence interval on the lower bound on z^{MSLP}.

We use the SDDP implementation `sddp.jl` [16] to solve each SAA problem. This algorithm is implemented in Julia. In benchmarks reported in [16], it was found that the computation times for `sddp.jl` were about 30% higher than those for the C++ code DOASA [45], on a test instance for which DOASA was designed for. The code `sddp.jl` does not directly support having random constraint coefficients, as in (18). However, the algorithm does support solving a problem with an underlying state evolving according to a Markov chain, and with parameters in the constraints dependent on the state of the Markov chain. Thus, we model the stochastic process as a Markov chain having 200 states corresponding to the 200 scenarios of joint realizations of (ξ_t^g, ξ_t^w) in each stage $t \in [2, T]$. The transition probability from each state in stage t to each state in stage $t+1$ is $1/200$. To limit the risk that the cutting plane models used in the SDDP algorithm grow too large, we set the parameter "cut_selection_frequency" to 50, which means that after every 50 iterations of the SDDP algorithm, cuts that are not currently binding are removed. Finally, to be consistent with the implementation of the two-stage LDR approximation, we run `sddp.jl` serially, although we note that both `sddp.jl` and the two-stage LDR approximation have significant potential for speedup via parallelization.

The time limit for each SDDP replication is set as follows. We let t_{LDR} be the total time required to solve the SAA problems (6) and (15), and evaluate the value of the obtained dual policy with an independent sample of size $N' = 5000T$. We run the SDDP algorithm on each of the 25 SAA replications with two time limits: TL:=1.5$*t_{\text{LDR}}$/25 and 10$*$TL. The first time limit is used to approximately match the

Table 4 Comparison of lower bounds obtained form two-stage LDR and SDDP algorithm

			D-LDR-2S		SDDP TL			SDDP 10X TL		
C	T	t_{LDR}	Mean	±	Mean	±	%LΔ	Mean	±	%LΔ
50	5	188	100.0	0.3	100.6	0.7	0.2	100.8	0.7	0.4
	10	866	100.0	0.5	101.5	1.3	0.6	103.2	1.3	2.4
	15	1728	100.0	1.3	94.3	2.6	− 7.1	103.9	2.9	2.3
	20	2897	100.0	1.5	80.8	3.4	− 21.4	94.2	3.8	− 8.2
100	5	171	100.0	0.3	101.5	0.7	1.1	101.7	0.7	1.3
	10	1094	100.0	0.4	100.8	1.1	0.1	101.6	1.2	0.9
	15	2235	100.0	0.9	102.4	2.3	1.0	108.6	2.5	7.1
	20	3827	100.0	1.8	88.3	3.4	− 13.5	101.4	3.8	− 0.6

total time (over all replications) allotted to the SDDP algorithm with the time used by the two-stage LDR approach (where the factor 1.5 is used to compensate for the fact that `sddp.jl` is implemented in Julia whereas the two-stage LDR approach is implemented in C++). The second time limit is used to demonstrate the potential of SDDP to obtain improved lower bounds and policies when given more time. Estimating the expected cost of the SDDP and STT policies requires a separate simulation of these policies, which has very similar computational effort for the two policies, and thus this time is excluded from t_{LDR}.

The lower bound results are reported in Table 4, in which again the objective values are scaled such that the estimated lower bound obtained by the two-stage LDR algorithm is 100.0. In the table, t_{LDR} is rounded to the nearest second. In aggregate, 40% of this time is spent solving (6), 48% is spent solving (15), and 12% is spent evaluating the dual bound with the independent sample. The table also presents the mean and half-width (±) of the lower bound obtained using two-stage LDR and the SDDP algorithm given time limits TL and 10∗TL. The columns %LΔ present the percentage difference between the lower end of the 95% CI on the lower bound obtained by the SDDP algorithm and that obtained by the two-stage LDR algorithm. Here a negative number indicates the lower bound was smaller (worse), and a positive number indicates an improvement over two-stage LDR. We find that when given a time limit similar to the time used by the two-stage LDR approximation, the SDDP algorithm obtains slightly better lower bounds on instances with fewer time stages, but somewhat worse lower bounds on the instances with more time stages. On the other hand, when given more time, the SDDP algorithm is able to achieve noticeably better lower bounds on instances with the fewer time stages, and closes much of the gap on the instances with more stages.

We next compare estimates of the expected cost of policies obtained with the two-stage LDR and SDDP methods. For the two-stage LDR policy, the policy and estimate of associated upper bound are determined as described in Sect. 5.2. For the SDDP algorithm, a policy can be obtained by first solving a (single) SAA approximation problem, and then using the resulting value-function approximation to drive a policy that is then evaluated via forward simulation replications using independently gen-

Table 5 Comparison of approximate upper bounds obtained form two-stage LDR and SDDP algorithm

			P-LDR-2S		SDDP TL			SDDP 10X TL		
C	T	t_{EVAL}	Mean	±	Mean	±	%UΔ	Mean	±	%UΔ
50	5	53	100.8	0.3	100.7	0.7	0.3	100.8	0.7	0.4
	10	239	114.6	0.6	104.3	1.4	− 8.2	104.1	1.3	− 8.4
	15	491	121.0	1.3	109.5	3.0	− 8.0	108.4	3.0	− 8.8
	20	932	102.9	1.4	100.6	4.1	0.4	100.4	4.0	0.1
100	5	53	101.3	0.3	101.7	0.7	0.8	101.7	0.7	0.7
	10	238	109.2	0.4	102.0	1.2	− 5.9	102.0	1.2	− 5.9
	15	495	133.2	1.2	113.4	2.6	− 13.6	112.4	2.6	− 14.4
	20	900	116.6	1.6	109.6	4.0	− 3.9	109.0	3.9	− 4.5

erated values of the random variables (i.e., independent from those used in the SAA approximation). Unfortunately, the ability to run a forward simulation using samples different from those used to solve the SDDP problem is not supported in `sddp.jl`. To obtain an estimate of the value of the policy that can be obtained using SDDP, for each of the 25 SAA replications solved by SDDP, we simulated the resulting policy using the sample distribution used in the SAA problem to estimate the expected cost of that policy. We then constructed a 95% confidence interval of the resulting upper bounds, and these are the values reported in Table 5. The column 't_{EVAL}' in this table presents the time, in seconds, to estimate the expected cost of the STT policy. The remaining columns present the confidence intervals of the estimated upper bounds in format similar to Table 4. As we see from the columns %UΔ, the estimated expected cost of the SDDP policies is in many cases significantly lower than the estimated expected cost of the two-stage LDR policy, suggesting that SDDP obtains significantly better primal policies for this problem.

In summary, for this problem, we find that SDDP provides similar, or slightly worse, lower bounds, and significantly better primal policies, in a comparable amount of time as the two-stage LDR approximation, and the lower bounds can be improved by running SDDP for more time. Thus, for this problem, SDDP is clearly favored over the two-stage approximation. Thus, LDR approximations (both static and two-stage) may be most useful for problems in which the assumptions required to apply SDDP do not hold. For example, in a hydropower planning case study presented in [58], the time series of water inflows, X_t, was modeled as $X_t = e^{Y_t}$, where Y_t follows a first order $AR(1)$ autoregressive time series, making the model nonlinear in X_t, and hence not solvable directly by SDDP.

6 Concluding remarks

We propose two-stage LDRs, a new approximate solution method for MSLPs. This approach has two advantages over staic LDRs. Due to the flexibility in the recourse decisions, our method potentially yields better (at least not worse) bounds and policies

than standard static LDR policies. In addition, as our approach is based on sampling and 2SLP, it works with very mild assumptions and can take advantage of existing literature on methods for approximately solving 2SLP problems. We illustrate the new approach on two example problems, an inventory planning problem and a capacity planning problem, which indicate that two-stage LDR policies have potential to yield significantly better policies than static LDR policies.

In future research it will be interesting to test the use of two-stage LDR policies on more problems, and to investigate if there are problem classes where two-stage LDR policies are provably optimal or near-optimal.

In the primal problem, a two-stage LDR can be directly applied to multi-stage stochastic *mixed integer programs*, provided the integrality restrictions are imposed only on the recourse variables. Since availability of algorithms for multi-stage stochastic mixed integer programs is very limited, it will be interesting to explore this extension further, in particular possibly using ideas from [7] to obtain a decision rule in the case the state variables also have integrality constraints.

Acknowledgements The research of James Luedtke has been supported in part by the National Science Foundation under Grant CMMI-1634597, and by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program under Contract Number DE-AC02-06CH11347. The research of Merve Bodur has been supported in part by the Natural Sciences and Engineering Research Council of Canada under the Discovery Grant RGPIN-2018-04984. We would like to thank the reviewers for their careful and constructive comments that have significantly improved the paper. Also, we would like to thank Angelos Georghiou for providing data and clarifications on the inventory planning example, and Oscar Dowson for his assistance using the sddp.jl code.

A Data for the capacity expansion problem

See Tables 6, 7 and 8.

Table 6 Fixed annual cost and operation cost (Table 1 in [15])

	$i=1$	$i=2$	$i=3$
ι_i (k€/MW)	245.8	113.9	57.8
c_i (€/MWh)	41.9	58.9	90.8

Table 7 Initial demand (Table 2 in [15])

	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$	$\ell=6$	$\ell=7$	$\ell=8$
$d_{0,\ell}$ (GW)	77.1	71.4	65.7	60.1	54.4	48.8	43.1	37.4
τ_ℓ (h)	68	677	1585	1781	1367	1688	1289	305

Table 8 Wind regimes (Table 3 in [15])

	$w=1$	$w=2$	$w=3$	$w=4$	$w=5$
η_w (%)	92.9	81.1	54.9	21.2	0.0
τ_w (%)	19.8	21.78	18.2	26.7	13.5

B Additional models for the capacity expansion example

B.1 Primal model and Benders decomposition

P-LDR-2S of the capacity expansion model is obtained by substituting

$$s_{ti}(\xi^t) = \sum_{k\in[K_t]} \Phi_{tk}(\xi^t)\beta_{tki}$$

in (17). Dropping ξ^t dependences for variables to simplify the notation, we obtain

$$\begin{aligned}
\min\ & \sum_{i\in[I]} \left(c^{u^+}_{1i} u^+_{1i} + c^{u^-}_{1i} u^-_{1i} + \sum_{j\in[J]} c^x_{1ij} x_{1ij}(\xi^t)\right) + \sum_{j\in[J]} c^z_{1j} z_{1j} \\
& \quad + \sum_{t\in[T]} \sum_{i\in[I]} c^s_{ti} \sum_{k\in[K_t]} \mathbb{E}\big[\Phi_{tk}(\xi^t)\big]\beta_{tki} + \sum_{t\in[2,T]} \mathbb{E}[Q_t(\beta,\xi^t)] && \text{(19a)}\\
\text{s.t. } & \beta_{11i} - u^+_{1i} + u^-_{1i} = 0, && i\in[I], && \text{(19b)}\\
& \beta_{11i} - x_{1ij} \geq 0, && i\in[I],\ j\in[J], && \text{(19c)}\\
& \sum_{i\in[I]} x_{1ij} + z_{1j} \geq d_{1j}, && j\in[J], && \text{(19d)}\\
& 0 \leq z_{1j} \leq d_{1j}, && j\in[J], && \text{(19e)}\\
& 0 \leq u^+_{1i} \leq M^{u^+}_{1i},\ 0 \leq u^-_{1i} \leq M^{u^-}_{1i},\ 0 \leq \beta_{11i} \leq M^s_{1i}, && i\in[I], && \text{(19f)}\\
& x_{1ij} \geq 0, && i\in[I],\ j\in[J], && \text{(19g)}
\end{aligned}$$

where, for $t \in [2,T]$, $Q_t(\beta,\xi^t)$ is defined as the optimal objective value of the following problem:

$$\begin{aligned}
\min\ & \sum_{i\in[I]} \left(c^{u^+}_{ti} u^+_i + c^{u^-}_{ti} u^-_i + \sum_{j\in[J]} c^x_{tij} x_{ij}\right) + \sum_{j\in[J]} c^z_{tj} z_j && \text{(20a)}\\
\text{s.t. } & u^+_i - u^-_i = \sum_{k\in[K_t]} \Phi_{tk}(\xi^t)\beta_{tki} - \sum_{k\in[K_{t-1}]} \Phi_{tk}(\xi^{t-1})\beta_{t-1,k,i}, && i\in[I], && \text{(20b)}\\
& x_{ij} \leq \sum_{k\in[K_t]} \Phi_{tk}(\xi^t)\beta_{tki}, && i\in[I],\ j\in[J], && \text{(20c)}\\
& \sum_{i\in[I]} x_{ij} + z_j \geq d_{tj}(\xi^t), && j\in[J], && \text{(20d)}\\
& 0 \leq z_j \leq d_{tj}(\xi^t), && j\in[J], && \text{(20e)}\\
& 0 \leq u^+_i \leq M^{u^+}_{ti}, && i\in[I], && \text{(20f)}\\
& 0 \leq u^-_i \leq M^{u^-}_{ti}, && i\in[I], && \text{(20g)}\\
& x_{ij} \geq 0, && i\in[I],\ j\in[J], && \text{(20h)}\\
& 0 \leq M^s_{ti} - \sum_{k\in[K_t]} \Phi_{tk}(\xi^t)\beta_{tki}, && i\in[I], && \text{(20i)}\\
& 0 \leq \sum_{k\in[K_t]} \Phi_{tk}(\xi^t)\beta_{tki}, && i\in[I], && \text{(20j)}
\end{aligned}$$

We note that P-LDR-2S of the capacity expansion model does not have relatively complete recourse since the recourse constraints (20i) and (20j) might be violated under some scenarios.

Let $\xi_n^T, n \in [N]$, be an independent and identically distributed (i.i.d.) random sample of the random vector ξ^T. We solve the obtained primal SAA problem with Benders decomposition, using a single aggregate cut per time-stage. That is, we have a master problem of the following form:

$$\min \sum_{i\in[I]} \left(c_{1i}^{u^+} u_{1i}^{+} + c_{1i}^{u^-} u_{1i}^{-} + \sum_{j\in[J]} c_{1ij}^{x} x_{1ij}(\xi^t)\right) + \sum_{j\in[J]} c_{1j}^{z} z_{1j} + \sum_{t\in[T]}\sum_{i\in[I]} c_{ti}^{s} \sum_{k\in[K_t]} \mathbb{E}\big[\Phi_{tk}(\xi^t)\big]\beta_{tki} + \sum_{t\in[2,T]} \eta_t \tag{21a}$$

$$\text{s.t. (19b)} - \text{(19g)}, \tag{21b}$$

$$(\eta_t, \beta_{t11}, \cdots, \beta_{tK_tI}) \in \mathcal{O}^t, \qquad t \in [2,T], \tag{21c}$$

$$(\beta_{t11}, \cdots, \beta_{tK_tI}) \in \mathcal{F}^t, \qquad t \in [2,T], \tag{21d}$$

$$0 \le \sum_{k\in[K_t]} \Phi_{tk}(\xi_n^t)\beta_{tki} \le M_{ti}^s, \qquad t \in [2,T], i \in [I],\ n \in [N], \tag{21e}$$

$$\eta_t \ge 0, \qquad t \in [2,T]. \tag{21f}$$

The variable η_t represents the expected second-stage cost at period $t \in [2,T]$, i.e., $\frac{1}{N}\sum_{n\in[N]}[Q_t(\beta, \xi_n^t)]$. Note that since all the original decision variables are defined to be nonnegative, and all the cost parameters are assumed to be nonnegative, $Q_t(\beta, \xi^t) \ge 0$ for any given β, thus (21f) are valid. Constraints (21c) and (21d) correspond to the set of Benders optimality and feasibility cuts, respectively. As β variables belong to the master problem, we add constraints (20i) and (20j) for each scenario in the sample to the master problem as (21e) which can be seen as an additional set of feasibility cuts.

The subproblem decomposes not only by scenario but also by stage. For $t \in [2,T]$ and $n \in [N]$, we have the corresponding subproblem (20a)–(20h), denoted by SP(t,n).

At every iteration of the Benders decomposition algorithm, we solve the master problem, get a candidate β solution which is fixed in the subproblems, and solve all the subproblems. For $t \in [2,T]$, if there is at least one index $n \in [N]$ for which SP(t,n) is infeasible, then we generate a Benders feasibility cut and add it to the master problem. Otherwise, we generate a Benders optimality cut, but add it to the master problem only if it is violated at the current master problem solution. We repeat this procedure until all the subproblems are feasible and no violated optimality cuts are found.

B.2 Dual model and level method

Let $\lambda, \gamma, \theta^+, \theta^-, \Gamma^{u^+}, \Gamma^{u^-}, \Gamma^s$ be the dual variables associated with the constraints (17b)–(17h) in (17), respectively. Then, the dual of (17) is:

$$\max \mathbb{E}\sum_{t\in[T]}\Big[\sum_{j\in[J]} d_{tj}(\xi^t)\big(\theta^+_{tj}(\xi^t)-\theta^-_{tj}(\xi^t)\big) - \sum_{i\in[I]}\big(M^{u^+}_{ti}\Gamma^{u^+}_{ti}(\xi^t)+M^{u^-}_{ti}\Gamma^{u^-}_{ti}(\xi^t)+M^{s}_{ti}\Gamma^{s}_{ti}(\xi^t)\big)\Big] \tag{22a}$$

$$\text{s.t. } \lambda_{ti}(\xi^t)-\mathbb{E}[\lambda_{t+1,i}(\xi^{t+1})\mid\xi^t] + \sum_{j\in[J]}\gamma_{tij}(\xi^t)-\Gamma^s_{ti}(\xi^t)\le c^s_{ti}, \quad t\in[T],\ \mathbb{P}\text{-a.s.},\ i\in[I], \tag{22b}$$

$$-\Gamma^{u^+}_{ti}(\xi^t)-\lambda_{ti}(\xi^t)\le c^{u^+}_{ti}, \quad t\in[T],\ \mathbb{P}\text{-a.s.},\ i\in[I], \tag{22c}$$

$$-\Gamma^{u^-}_{ti}(\xi^t)+\lambda_{ti}(\xi^t)\le c^{u^-}_{ti}, \quad t\in[T],\ \mathbb{P}\text{-a.s.},\ i\in[I], \tag{22d}$$

$$\theta^+_{tj}(\xi^t)-\theta^-_{tj}(\xi^t)\le c^z_{tj}, \quad t\in[T],\ \mathbb{P}\text{-a.s.},\ j\in[J], \tag{22e}$$

$$\theta^+_{tj}(\xi^t)-\gamma_{tij}(\xi^t)\le c^x_{tij}, \quad t\in[T],\ \mathbb{P}\text{-a.s.},\ i\in[I],\ j\in[J], \tag{22f}$$

$$\gamma_{tij}(\xi^t)\ge 0, \quad t\in[T],\ \mathbb{P}\text{-a.s.},\ i\in[I],\ j\in[J], \tag{22g}$$

$$\theta^+_{tj}(\xi^t),\theta^-_{tj}(\xi^t)\ge 0, \quad t\in[T],\ \mathbb{P}\text{-a.s.},\ j\in[J], \tag{22h}$$

$$\Gamma^{u^+}_{ti}(\xi^t),\Gamma^{u^-}_{ti}(\xi^t),\Gamma^s_{ti}(\xi^t)\ge 0 \quad t\in[T],\ \mathbb{P}\text{-a.s.},\ i\in[I]. \tag{22i}$$

Observing that θ^- variables are redundant, we remove them to simplify the model. D-LDR-2S of the capacity expansion model is obtained by substituting

$$\lambda_{ti}(\xi^t)=\sum_{k\in[K_t]}\Phi_{tk}(\xi^t)\Lambda_{tki}$$

in (22). Dropping ξ^t dependences for variables to simplify the notation, we obtain

$$\max \sum_{j\in[J]} d_{1j}\theta^+_{1j}-\sum_{i\in[I]}\big(M^{u^+}_{1i}\Gamma^{u^+}_{1i}+M^{u^-}_{1i}\Gamma^{u^-}_{1i}+M^{s}_{1i}\Gamma^{s}_{1i}\big)+\sum_{t\in[2,T]}\mathbb{E}[G_t(\Lambda,\xi^t)] \tag{23a}$$

$$\text{s.t. } \Lambda_{11i}-\sum_{k\in[K_2]}\mathbb{E}\big[\Phi_{2k}(\xi^2)\big]\Lambda_{2ki}+\sum_{j\in[J]}\gamma_{1ij}-\Gamma^s_{1i}\le c^s_{1i}, \quad i\in[I], \tag{23b}$$

$$-\Gamma^{u^+}_{1i}-\Lambda_{11i}\le c^{u^+}_{1i}, \quad i\in[I], \tag{23c}$$

$$-\Gamma^{u^-}_{1i}+\Lambda_{11i}\le c^{u^-}_{1i}, \quad i\in[I], \tag{23d}$$

$$\theta^+_{1j}\le c^z_{1j}, \quad j\in[J], \tag{23e}$$

$$\theta^+_{1j}-\gamma_{1ij}\le c^x_{1ij}, \quad i\in[I],\ j\in[J] \tag{23f}$$

$$\gamma_{1ij}\ge 0, \quad i\in[I],\ j\in[J] \tag{23g}$$

$$\theta^+_{1j}\ge 0, \quad j\in[J], \tag{23h}$$

$$\Gamma^{u^+}_{1i},\Gamma^{u^-}_{1i},\Gamma^s_{1i}\ge 0 \quad i\in[I], \tag{23i}$$

where for $t\in[2,T]$, $G_t(\Lambda,\xi^t)$ is defined as the optimal objective value of the following problem:

$$\max \sum_{j\in[J]} d_{tj}\theta_j^+ - \sum_{i\in[I]} \left(M_{ti}^{u^+}\Gamma_i^{u^+} + M_{ti}^{u^-}\Gamma_i^{u^-} + M_{ti}^{s}\Gamma_i^{s}\right) \tag{24a}$$

$$\text{s.t.} \sum_{j\in[J]} \gamma_{ij} - \Gamma_i^s \le c_{ti}^s - \sum_{k\in[K_t]} \Phi_{tk}(\xi^t)\Lambda_{tki} + \sum_{k\in[K_{t+1}]} \mathbb{E}\left[\Phi_{t+1,k}(\xi^{t+1})\middle|\xi^t\right]\Lambda_{t+1,k,i}, \quad i \in [I], \tag{24b}$$

$$-\Gamma_i^{u^+} \le c_{ti}^{u^+} + \sum_{k\in[K_t]} \Phi_{tk}(\xi^t)\Lambda_{tki}, \quad i \in [I], \tag{24c}$$

$$-\Gamma_i^{u^-} \le c_{ti}^{u^-} - \sum_{k\in[K_t]} \Phi_{tk}(\xi^t)\Lambda_{tki}, \quad i \in [I], \tag{24d}$$

$$\theta_j^+ \le c_{tj}^z, \quad j \in [J], \tag{24e}$$

$$\theta_j^+ - \gamma_{ij} \le c_{tij}^x, \quad i \in [I],\ j \in [J], \tag{24f}$$

$$\gamma_{ij} \ge 0, \quad i \in [I],\ j \in [J], \tag{24g}$$

$$\theta_j^+ \ge 0, \quad j \in [J], \tag{24h}$$

$$\Gamma_i^{u^+}, \Gamma_i^{u^-}, \Gamma_i^s \ge 0 \quad i \in [I]. \tag{24i}$$

Let ξ_n^T, $n \in [N]$, be an independent and identically distributed (i.i.d.) random sample of the random vector ξ^T. We solve the obtained dual SAA problem with the bundle-level method, because the Benders decomposition method converged slowly for this problem. We use cuts aggregated over scenarios, thus introduce ζ_t variable to represent the expected second-stage cost value, i.e., $\frac{1}{N}\sum_{n\in[N]} G_t(\Lambda, \xi_n^t)$, for $t \in [2, T]$.

We observe that the subproblem (24) can be further decomposed into two: one problem including only the u^+ and u^- variables, and the other problem including the remaining set of variables.

$$(\text{DSP}^{upart}): \max -\sum_{i\in[I]} \left(M_{ti}^{u^+}\Gamma_i^{u^+} + M_{ti}^{u^-}\Gamma_i^{u^-}\right)$$
$$\text{s.t. (24c), (24d), (24i)}$$
$$(\text{DSP}^{rest}): \max \sum_{j\in[J]} d_{tj}\theta_j^+ - \sum_{i\in[I]} M_{ti}^s\Gamma_i^s$$
$$\text{s.t. (24b), (24e)}-\text{(24h)}$$

We exploit this decomposition to disaggregate the optimality cuts in the master problem. Thus, we introduce additional variables ζ_t^{upart} and ζ_t^{rest} for $t \in [2, T]$ and obtain the following master problem:

$$(\text{MP}): \max \sum_{j\in[J]} d_{1j}\theta_{1j}^+ - \sum_{i\in[I]} \left(M_{1i}^{u^+}\Gamma_{1i}^{u^+} + M_{1i}^{u^-}\Gamma_{1i}^{u^-} + M_{1i}^{s}\Gamma_{1i}^{s}\right) + \sum_{t\in[2,T]} \zeta_t \tag{25a}$$

$$\text{s.t. (23b)} - \text{(23i)}, \tag{25b}$$

$$\zeta_t = \zeta_t^{upart} + \zeta_t^{rest}, \qquad t \in [2, T], \tag{25c}$$

$$(\zeta_t^{upart}, \Lambda_{t11}, \ldots, \Lambda_{tK_tI}) \in \mathcal{U}^t, \qquad t \in [2, T], \tag{25d}$$

$$(\zeta_t^{rest}, \Lambda_{t11}, \ldots, \Lambda_{tK_tI}) \in \mathcal{R}^t, \qquad t \in [2, T], \tag{25e}$$

$$\zeta_t^{upart} \leq 0, \qquad t \in [2, T], \tag{25f}$$

$$\zeta_t^{rest} \leq \frac{1}{N} \sum_{n\in[N]} \sum_{j\in[J]} d_{tj}(\xi_n^t) c_{tj}^z, \qquad t \in [2, T], \tag{25g}$$

where $\mathcal{U}^t$ and $\mathcal{R}^t$ represent the optimality cuts derived from problems (DSP^{upart}) and (DSP^{rest}), respectively. Moreover, we introduce the upper bounds on the new auxiliary variables, which are derived from the subproblems (DSP^{upart}) and (DSP^{rest}).

The level method also uses a quadratic program for regularization which projects the previous iterate on the level set of the current approximation of the objective function. We use the following problem for this projection:

$$\begin{aligned} \text{(QP)}: \ & \max ||\Lambda - \hat{\Lambda}||_2^2 \\ & \text{s.t. (25b)}-\text{(25g)}, \\ & \sum_{j\in[J]} d_{1j}\theta_{1j}^+ - \sum_{i\in[I]} \left(M_{1i}^{u^+}\Gamma_{1i}^{u^+} + M_{1i}^{u^-}\Gamma_{1i}^{u^-} + M_{1i}^{s}\Gamma_{1i}^{s}\right) \geq L, \end{aligned}$$

where $\hat{\Lambda}$ and L denote the current Λ solution (i.e., the previous iterate) and the level target, respectively. The optimal solution values of Λ variables determine the next iterate.

The details of the level method are provided in Algorithm 1 where LB and UB denote lower bound and upper bound, respectively.

Algorithm 1 : Level Algorithm

1. Initialize $\hat{\Lambda} = 0$, LB $= -\infty$, UB $= \infty$

2. Solve all the subproblems, i.e., (DSP^{upart}) and (DSP^{rest}) for all $t \in [2, T]$ and $n \in [N]$.
   Generate Benders optimality cuts, add them to both (MP) and (QP).
   Compute the objective value of the current iterate, and set it as LB.

3. do
   Solve (MP). Update UB if (MP) optimal objective value is lower than UB.
   Set $L = 0.3 \times$ UB $+ 0.7 \times$ LB.
   Solve (QP) with updated level constraint, to obtain iterate $\hat{\Lambda}$.
   Solve all the subproblems at current iterate.
   Generate Benders optimality cuts, and add violated cuts to both (MP) and (QP).
   Compute the objective value of the current iterate; if it is larger than LB, update LB.
   until |UB - LB| / UB $< 10^{-5}$

References

1. Ahmed, S.: Multistage stochastic optimization (2016). https://www.ima.umn.edu/materials/2015-2016/ND8.1-12.16/25386/mssp.pdf. New Directions Short Course on Mathematical Optimization
2. Bakir, I., Boland, N., Dandurand, B., Erera, A.: Scenario set partition dual bounds for multistage stochastic programming: a hierarchy of bounds and a partition sampling approach (2016). http://www.optimization-online.org/DB_FILE/2016/01/5311.pdf
3. Bampou, D., Kuhn, D.: Scenario-free stochastic programming with polynomial decision rules. In: 2011 50th IEEE Conference on Decision and Control and European Control Conference, pp. 7806–7812. IEEE (2011)
4. Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A.: Adjustable robust solutions of uncertain linear programs. Math. Program. **99**(2), 351–376 (2004)
5. Ben-Tal, A., Nemirovski, A.: Robust optimization—methodology and applications. Math. Program. **92**, 453–480 (2002)
6. Bertsimas, D., Caramanis, C.: Adaptability via sampling. In: 2007 46th IEEE Conference on Decision and Control, pp. 4717–4722. IEEE (2007)
7. Bertsimas, D., Georghiou, A.: Design of near optimal decision rules in multistage adaptive mixed-integer optimization. Oper. Res. **63**(3), 610–627 (2015)
8. Birge, J.: Decomposition and partitioning methods for multistage stochastic linear programs. Oper. Res. **33**(5), 989–1007 (1985)
9. Birge, J.R.: Aggregation bounds in stochastic linear programming. Math. Program. **31**(1), 25–41 (1985)
10. Birge, J.R., Donohue, C.J., Holmes, D.F., Svintsitski, O.G.: A parallel implementation of the nested decomposition algorithm for multistage stochastic linear programs. Math. Program. **75**(2), 327–352 (1996)
11. Calafiore, G., Campi, M.: The scenario approach to robust control design. IEEE Trans. Automat. Contr. **51**, 742–753 (2006)
12. Casey, M.S., Sen, S.: The scenario generation algorithm for multistage stochastic linear programming. Math. Oper. Res. **30**(3), 615–631 (2005)
13. Chen, X., Sim, M., Sun, P., Zhang, J.: A linear decision-based approximation approach to stochastic programming. Oper. Res. **56**(2), 344–357 (2008)
14. Chen, Z.L., Powell, W.: Convergent cutting-plane and partial-sampling algorithm for multistage stochastic linear programs with recourse. J. Optim. Theory Appl. **102**(3), 497–524 (1999)
15. de Maere d'Aertrycke, G., Shapiro, A., Smeers, Y.: Risk exposure and Lagrange multipliers of nonanticipativity constraints in multistage stochastic problems. Math. Methods Oper. Res. **77**(3), 393–405 (2013)
16. Dowson, O., Kapelevich, L.: SDDP.jl: a Julia package for stochastic dual dynamic programming. Optimization Online (2017). http://www.optimization-online.org/DB_HTML/2017/12/6388.html
17. Dyer, M., Stougie, L.: Computational complexity of stochastic programming problems. Math. Program. **106**(3), 423–432 (2006)
18. Eisner, M., Olsen, P.: Duality for stochastic programming interpreted as L.P. in L_p-space. SIAM J. Appl. Math **28**(4), 779–792 (1975)
19. Fábián, C.I., Szőke, Z.: Solving two-stage stochastic programming problems with level decomposition. Comput. Manag. Sci. **4**, 313–353 (2007)
20. Fu, M.C.: Feature article: optimization for simulation: theory vs. practice. INFORMS J. Comput. **14**(3), 192–215 (2002)
21. Garstka, S., Wets, R.: On decision rules in stochastic programming. Math. Program. **7**(1), 117–143 (1974)
22. Gassmann, H.I.: Mslip: a computer code for the multistage stochastic linear programming problem. Math. Program. **47**(1), 407–423 (1990)
23. Georghiou, A., Wiesemann, W., Kuhn, D.: Generalized decision rule approximations for stochastic programming via liftings. Math. Program. **152**(1–2), 301–338 (2015)
24. Girardeau, P., Leclere, V., Philpott, A.: On the convergence of decomposition methods for multistage stochastic convex programs. Math. Oper. Res. **40**, 130–145 (2015)
25. Guigues, V.: SDDP for some interstage dependent risk-averse problems and application to hydro-thermal planning. Comput. Optim. Appl. **57**, 167–203 (2014)
26. Guigues, V.: Convergence analysis of sampling-based decomposition methods for risk-averse multistage stochastic convex programs. SIAM J. Optim. **26**, 2468–2494 (2016)

27. Hanasusanto, G.A., Kuhn, D., Wiesemann, W.: A comment on "computational complexity of stochastic programming problems". Math. Program. **159**(1), 557–569 (2016)
28. Heitsch, H., Römisch, W.: Scenario tree modeling for multistage stochastic programs. Math. Program. **118**(2), 371–406 (2009)
29. Higle, J., Sen, S.: Multistage stochastic convex programs: duality and its implications. Ann. Oper. Res. **142**(1), 129–146 (2006)
30. Homem de Mello, T.: On rates of convergence for stochastic optimization problems under non-independent and identically distributed sampling. SIAM J. Optim. **19**, 524–551 (2008)
31. Hong, L.J., Nelson, B.L.: A brief introduction to optimization via simulation. In: Winter Simulation Conference, WSC '09, Winter Simulation Conference, pp. 75–85 (2009)
32. Høyland, K., Wallace, S.: Generating scenario trees for multistage decision problems. Manag. Sci. **47**(2), 295–307 (2001)
33. Infanger, G., Morton, D.: Cut sharing for multistage stochastic linear programs with interstage dependency. Math. Program. **75**(2), 241–256 (1996)
34. Koivu, M.: Variance reduction in sample approximations of stochastic programs. Math. Program. **103**, 463–485 (2005)
35. Kuhn, D., Wiesemann, W., Georghiou, A.: Primal and dual linear decision rules in stochastic and robust optimization. Math. Program. **130**(1), 177–209 (2011)
36. Lemaréchal, C., Nemirovskii, A., Nesterov, Y.: New variants of bundle methods. Math. Program. **69**, 111–147 (1995)
37. Lemaréchal, C., Nemirovskii, A., Nesterov, Y.: New variants of bundle methods. Math. Program. **69**, 111–147 (1995)
38. Linderoth, J., Shapiro, A., Wright, S.: The empirical behavior of sampling methods for stochastic programming. Ann. Oper. Res. **142**, 215–241 (2006)
39. Mak, W.K., Morton, D., Wood, R.: Monte Carlo bounding techniques for determining solution quality in stochastic programs. Oper. Res. Lett. **24**, 47–56 (1999)
40. Nemirovski, A., Juditsky, A., Lan, G., Shapiro, A.: Robust stochastic approximation approach to stochastic programming. SIAM J. Optim. **19**(4), 1574–1609 (2009)
41. Pennanen, T.: Epi-convergent discretizations of multistage stochastic programs via integration quadratures. Math. Program. **116**, 461–479 (2009)
42. Pereira, M., Pinto, L.: Multi-stage stochastic optimization applied to energy planning. Math. Program. **52**(1–3), 359–375 (1991)
43. Philpott, A., De Matos, V.: Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. Eur. J. Oper. Res. **218**(2), 470–483 (2012)
44. Philpott, A., Guan, Z.: On the convergence of stochastic dual dynamic programming and related methods. Oper. Res. Lett. **36**(4), 450–455 (2008)
45. Philpott, A., Pritchard, G.: Emi-doasa. Technical report, Electric Power Optimization Centre (2013). http://www.emi.ea.govt.nz/Content/Tools/Doasa/DOASA%20paper%20by%20SOL.pdf
46. Polyak, B., Juditsky, A.: Acceleration of stochastic approximation by averaging. SIAM J. Control Optim. **30**, 838–855 (1992)
47. Powell, W.: Approximate Dynamic Programming: Solving the Curses of Dimensionality, vol. 703. Wiley, New York (2007)
48. Robbins, H., Monro, S.: A stochastic approximation method. Ann. Math. Stat. **22**, 400–407 (1951)
49. Rockafellar, R., Wets, R.B.: Scenarios and policy aggregation in optimization under uncertainty. Math. Oper. Res. **16**(1), 119–147 (1991)
50. Ruszczyński, A.: A regularized decomposition method for minimizing a sum of polyhedral functions. Math. Program. **35**(3), 309–333 (1986)
51. Sen, S., Zhou, Z.: Multistage stochastic decomposition: a bridge between stochastic programming and approximate dynamic programming. SIAM J. Optim. **24**(1), 127–153 (2014)
52. Shapiro, A.: Stochastic programming approach to optimization under uncertainty. Math. Program. **112**, 183–220 (2008)
53. Shapiro, A.: Analysis of stochastic dual dynamic programming method. Eur. J. Oper. Res. **209**(1), 63–72 (2011)
54. Shapiro, A.: Topics in stochastic programming. Universite Catholique de Louvain, CORE Lecture Series (2011)
55. Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on stochastic programming: modeling and theory, vol. 16. SIAM (2014)

56. Shapiro, A., Homem-de Mello, T.: On the rate of convergence of optimal solutions of monte carlo approximations of stochastic programs. SIAM J. Optim. **11**(1), 70–86 (2000)
57. Shapiro, A., Nemirovski, A.: On complexity of stochastic programming problems. In: Jeyakumar, V., Rubinov, A. (eds.) Continuous Optimization, pp. 111–146. Springer, Heidelberg (2005)
58. Shapiro, A., Tekaya, W., da Paulo Costa, J., Pereira Soares, M.: Risk neutral and risk averse stochastic dual dynamic programming method. Eur. J. Oper. Res. **224**(2), 375–391 (2013)
59. Van Slyke, R.M., Wets, R.: L-shaped linear programs with applications to optimal control and stochastic programming. SIAM J. Appl. Math. **17**(4), 638–663 (1969)
60. Vayanos, P., Kuhn, D., Rustem, B.: A constraint sampling approach for multi-stage robust optimization. Automatica **48**(3), 459–471 (2012)
61. Wallace, S.W., Ziemba, W.T.: Applications of Stochastic Programming, vol. 5. SIAM, Philadelphia (2005)

Affiliations

Merve Bodur[1] · **James R. Luedtke**[2]

James R. Luedtke
jim.luedtke@wisc.edu

1 Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada

2 Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, USA