A Hierarchy of Bounds for Stochastic Mixed-Integer Programs[1]

Burhaneddin Sandıkçı
The University of Chicago Booth School of Business, Chicago, IL 60637
burhan@chicagobooth.edu

Nan Kong
Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN 47907
nkong@purdue.edu

Andrew J. Schaefer
Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, PA 15261
schaefer@ie.pitt.edu

Abstract

Strong relaxations are critical for solving deterministic mixed-integer programs. As solving stochastic mixed-integer programs (SMIPs) is even harder, it is likely that strong relaxations will also prove essential for SMIPs. We consider general two-stage SMIPs with recourse, where integer variables are allowed in both stages of the problem and randomness is allowed in the objective function, the constraint matrices (i.e., the technology matrix and the recourse matrix), and the right-hand side. We develop a hierarchy of lower and upper bounds for the optimal objective value of an SMIP by generalizing the wait-and-see (WS) solution and the expected result of using the expected value (EEV) solution. These bounds become progressively stronger but, generally, more difficult to compute. Our numerical study indicates that the bounds developed in this paper can be very strong relative to those provided by stochastic linear programming relaxations.

[1]We gratefully acknowledge the helpful comments of Osman Y. Özaltın and Amin Dehghanian from the University of Pittsburgh. Nan Kong was supported by Air Force Office of Scientists Research Grant FA9550-08-1-0268. Andrew J. Schaefer was supported by National Science Foundation grants CMMI-0620780 and CMMI-0826141.

1 Introduction

We consider the following two-stage stochastic mixed-integer program (SMIP):

$$(RP) \qquad \min_{x \in \mathbb{X}} \quad c^T x + \mathbb{E}_{\tilde{\xi}}\left[Q\left(x, \tilde{\xi}\right)\right]$$
$$\text{s.t.} \quad Ax = b,$$

where

$$Q\left(x, \tilde{\xi}\right) = \min_{y_{\tilde{\xi}} \in \mathbb{Y}} \left\{ d_{\tilde{\xi}}^T y_{\tilde{\xi}} \;\middle|\; T_{\tilde{\xi}} x + W_{\tilde{\xi}} y_{\tilde{\xi}} = h_{\tilde{\xi}} \right\}$$

and $c \in \mathbb{R}^{n_1}$ and $b \in \mathbb{R}^{m_1}$ are known column vectors, $A \in \mathbb{R}^{m_1 \times n_1}$ is a known matrix, and $\mathbb{X} = \mathbb{R}_+^{n_1 - k_1} \times \mathbb{Z}_+^{k_1}$. We assume $\tilde{\xi}$ is a random vector with finite support Ξ. Furthermore, without loss of generality, we assume that $|\Xi| = K + 1$, the scenarios in Ξ are indexed by $0, 1, \ldots, K$, and scenario i occurs with probability p_i. For each realization $\xi \in \Xi$ of the random vector $\tilde{\xi}$, $\mathbb{Y} = \mathbb{R}_+^{n_2 - k_2} \times \mathbb{Z}_+^{k_2}$ and all random problem parameters become known, i.e., $d_\xi \in \mathbb{R}^{n_2}$, $h_\xi \in \mathbb{R}^{m_2}$, $T_\xi \in \mathbb{R}^{m_2 \times n_1}$ (the *technology matrix*), and $W_\xi \in \mathbb{R}^{m_2 \times n_2}$ (the *recourse matrix*).

The recourse problem (RP) can be written in its *extensive form* as:

$$(EF) \qquad \min \quad c^T x + \sum_{i=0}^{K} p_i d_i^T y_i$$
$$\text{s.t.} \quad Ax = b,$$
$$T_i x + W_i y_i = h_i, \qquad i = 0, 1, \ldots, K,$$
$$x \in \mathbb{X},$$
$$y_i \in \mathbb{Y}, \qquad i = 0, 1, \ldots, K.$$

Define the optimization problem associated with one particular scenario $\xi \in \Xi$ as:

$$
\begin{aligned}
(P1) \qquad\qquad & \min \quad z(x,\xi) = c^T x + d_\xi^T y_\xi \\
& \text{s.t.} \qquad\quad Ax = b, \\
& \qquad T_\xi x + W_\xi y_\xi = h_\xi, \\
& \qquad\qquad\quad x \in \mathbb{X}, \\
& \qquad\qquad\quad y_\xi \in \mathbb{Y}.
\end{aligned}
$$

We may rewrite the recourse problem (RP) as:

$$RP = \min_x \left\{ \mathbb{E}_{\tilde{\xi}} \left[z\left(x, \tilde{\xi}\right) \right] \right\}. \tag{1}$$

Solving problem ($P1$) for each realization $\xi \in \Xi$ yields a perfect information solution. The expected objective value of this solution, known in the literature as the *wait-and-see* (WS) solution [14], is:

$$WS = \mathbb{E}_{\tilde{\xi}} \left[\min_x z\left(x, \tilde{\xi}\right) \right]. \tag{2}$$

When $\tilde{\xi}$ is replaced by its expected value (i.e., by $\bar{\xi} \equiv \mathbb{E}[\tilde{\xi}]$), the problem ($P1$) is called the *expected value problem*. The expected result of using an optimal first-stage solution from the expected value problem is denoted by EEV and given by:

$$EEV = \mathbb{E}_{\tilde{\xi}} \left[z^* \left(\bar{x}\left(\bar{\xi}\right), \tilde{\xi} \right) \right], \tag{3}$$

where $\bar{x}(\bar{\xi})$ is an optimal first-stage solution from the expected value problem and $z^*\left(\bar{x}\left(\bar{\xi}\right), \tilde{\xi}\right)$ is the optimal objective function value of problem $(P1)$ for a given $\bar{x}(\bar{\xi})$ and a given realization of the random vector $\tilde{\xi}$. If $\bar{x}(\bar{\xi})$ is infeasible for any scenario $\xi \in \Xi$, then it means that the recourse problem (RP) is itself infeasible for the given first-stage solution $\bar{x}(\bar{\xi})$ and EEV for such a case is defined to yield the trivial upper bound (i.e., $+\infty$).

We may now compare WS and EEV to RP. Madansky [14] established the following inequalities for stochastic linear programs (SLPs):

$$WS \leq RP \leq EEV. \quad (4)$$

The WS bound on RP follows from observing that for each realization $\xi \in \Xi$, the optimal objective value of $(P1)$ must be at least as good as (i.e., $\leq$ for a minimization problem) the optimal objective value of (RP). Taking the expectation of both sides with respect to $\tilde{\xi}$ yields the first inequality in (4). The EEV bound on RP follows trivially from equations (1) and (3). It is easy to see that inequalities (4) also apply to SMIPs.

The current literature uses two different measures to quantify the value associated with uncertainty in the parameters: (1) the *expected value of perfect information* ($EVPI$), which measures the maximum amount a decision maker would pay in return for accurate information about the future, and (2) the *value of the stochastic solution* (VSS), which measures the value of obtaining a solution considering the randomness in the problem parameters. Formally, $EVPI = RP - WS$ and $VSS = EEV - RP$.

In practice, WS and EEV seldom provide satisfactory lower and upper bounds for the optimal objective value of a stochastic program. Tighter bounds are often required when there is a need to provide a better sense of the optimal value. These tighter bounds may also lead to significant computational reduction in obtaining an exact solution to (RP) in many cases, particularly in SMIPs.

The evaluation of $\mathcal{Q}(x) = \mathbb{E}_{\xi}[Q(x,\xi)]$ can be quite complicated. Even with discrete support Ξ, the problem becomes an optimization problem defined over a potentially large number of scenarios. As the number of scenarios grows, it becomes too large to solve using general-purpose algorithms. Hence, it generally requires some form of approximation that exploits special structures inherent in (RP). The mathematical programming literature on approximations for solving large-scale SLPs is extensive [4, 11, 16].

The most common approximation procedure for SLPs is to find some relatively small set of realizations that somehow represent a good approximation of the real underlying distribution and to solve a deterministic problem that is constructed with this small set of realizations. Many bounds [6, 7, 8, 12] generalize Jensen's inequality [9] for lower bounding and the Edmundson-Madansky inequality [5, 13] for upper bounding. These bounds rely on the fact that the optimal objective function of the SLP relaxation of $(P1)$, $f(\xi) = \min_x z(x, \xi)$, is convex with respective to ξ. Another approach that is used in obtaining approximations for SLPs [3, 15] aggregates constraints and variables in the extensive-form problems, and thus solves a number of smaller problems instead of a large-scale SLP. These approximations follow the results given in [17, 18]. To generate these bounds, both primal and dual solutions to some aggregate problems must be attainable.

The aforementioned bounds typically do not apply to SMIPs. This is mainly because of the fact that a stochastic mixed-integer program, in general, lacks the required convexity properties and hence the Jensen and Edmundson-Madansky inequalities do not apply and the dual solutions are not attainable. Our bounding approach, on the other hand, works for general two-stage SMIPs as well as SLPs.

Our approach extends that of Birge [2]. In our approach, we solve deterministic mixed-integer programs of sizes larger than that of $(P1)$, the one used to compute WS and EEV. Each of these larger problems, termed a *group subproblem*, is constructed based on some *reference scenario* and a subset of scenarios from the support. A reference scenario could be any one of the scenarios from the support or it could be chosen as a scenario outside of the support (e.g., the mean of the random variable). More details regarding the reference scenario are provided in Section 2. A group subproblem can be viewed as a truncated extensive-form problem with a conditional probability distribution over the new scenario set. Thus, $(P1)$ can be viewed as a special group subproblem where the updated probability distribution forms a degenerate distribution, where a probability of 1 is assigned to the selected scenario and 0 to all other scenarios (including the reference scenario).

To construct tighter lower and upper bounds, we generalize the idea in computing WS and EEV. To compute WS, one solves $(P1)$ for each scenario and then takes expectation over all individual scenarios. Similarly, to obtain tighter lower bounds on RP, we solve group subproblems for each scenario group and then take expectation over all scenario groups. To compute EEV, one solves ($P1$) for some scenario (frequently the expected value) and then evaluate the expected objective of ($P1$) given the obtained first-stage decisions. Similarly, to obtain tighter upper bounds on RP, we solve group subproblems for each scenario group and then evaluate the expected objective of ($P1$) given the obtained first-stage decisions. Finally, we select the best expected objective over all scenario groups.

Compared to Birge [2], we present an alternative way of forming the group subproblems. In addition, we combine the solutions from group subproblems differently to construct the tighter lower bounds. Furthermore, we show that the lower bounding formulas in Birge [2] do not necessarily yield correct lower bounds nor do they observe monotonicity as the scenario group size increases. On the other hand, we show that our bounds are indeed lower bounds and that they ensure monotonicity. Also note that the SLP relaxation of an SMIP problem provides a lower bound and this bound is the initial lower bound for several algorithms for SMIPs. Our computational results, based on instances from the literature, demonstrate that the bounds developed in this paper are often much stronger than SLP-based bounds. Therefore, we believe these new bounds have the potential to improve several algorithms for solving SMIPs.

The remainder of this paper is organized as follows. Section 2 introduces our definition of group subproblems and compares it to the definition given in Birge [2]. Section 3 develops refined lower bounds for the optimal objective value of an SMIP and discusses properties of these lower bounds. It also provides a counterexample to show that the bounds introduced in Birge [2] do not necessarily satisfy the aforementioned properties. Section 4 develops upper bounds for the optimal objective value of an SMIP. The results of computational tests for a variety of stochastic programming problems from the literature are given in Section 5. Finally, the paper concludes in Section 6 by providing a discussion and pointing out possible extensions.

2 Group Subproblem

In this section, we first introduce our definition of a group subproblem and then compare it to the definition given in Birge [2]. For ease of exposition, we always use scenario ξ^0 as the *reference scenario*, which may or may not be in Ξ. If the reference scenario is not in Ξ, then $p_0 = 0$ and $|\Xi| = K$. Let $S \equiv \{1, 2, \ldots, K\}$ be the index set of scenarios excluding the reference scenario, and $\mathcal{P}(S)$ be the power set of S, i.e., the set that contains all subsets of S. For $k = 1, 2, \ldots, K$, let $\mathcal{P}_k(S) \equiv \{\Gamma \in \mathcal{P}(S) : |\Gamma| = k\}$, i.e., $\mathcal{P}_k(S)$ contains only those elements in $\mathcal{P}(S)$ that have cardinality k. Clearly, $\bigcap_{k\in S} \mathcal{P}_k(S) = \emptyset$ but $\bigcup_{k\in S} \mathcal{P}_k(S) = \mathcal{P}(S)$. Finally, define $\rho(\Gamma) \equiv \sum_{j\in\Gamma} p_j$ for any $\Gamma \in \mathcal{P}(S)$, i.e., $\rho(\Gamma)$ is the sum of the probabilities of the scenarios included in the set Γ.

Definition 1 *The group subproblem for any given scenario index set Γ is defined as:*

$$
\begin{aligned}
(GR(\Gamma)) \qquad z^*(\Gamma) \;=\; \min \;\; & c^T x + p_0 d_0^T y_0 + (1-p_0)\cdot\left(\sum_{i\in\Gamma} \frac{p_i}{\rho(\Gamma)} d_i^T y_i\right) \\
s.t. \quad & Ax = b, \\
& T_0 x + W_0 y_0 = h_0, \\
& T_i x + W_i y_i = h_i, \qquad i \in \Gamma, \\
& x \in \mathbb{X}, \\
& y_0, y_i \in \mathbb{Y}, \qquad i \in \Gamma.
\end{aligned}
$$

Given a scenario group Γ, the group subproblem $GR(\Gamma)$ can be viewed as a truncated extensive form over scenarios ξ^0 and ξ^i for all $i \in \Gamma$, where the probability of the reference scenario remains p_0 but the remaining probability $1-p_0$ is distributed to scenarios ξ^i for $i \in \Gamma$ according to their relative share in the group (i.e., according to $p_i/\rho(\Gamma)$, the conditional probability of scenario i in the given scenario group Γ). Next we compare our definition with the one given in Birge [2].

Definition 2 *(Birge [2]) The group subproblem for any given scenario index set* Ω *is defined as:*

$$
\begin{aligned}
(GR^B(\Omega)) \qquad z^B(\Omega) \;=\; \min \;\; & c^T x + p_0 d^T y_0 + \sum_{i\in\Omega}(1-p_0)d^T y_i \\
s.t. \quad & Ax = b, \\
& Tx + Wy_0 = h_0, \\
& Tx + Wy_i = h_i, \qquad i \in \Omega, \\
& x, y_0, y_i \geq 0 \qquad i \in \Omega.
\end{aligned}
$$

It should be noted that Birge [2] restricted himself to $\xi^0 = \bar{\xi} \equiv \mathbb{E}\left[\tilde{\xi}\right]$ but mentioned that specific knowledge about a problem may lead to a different choice for ξ^0. Furthermore, he only discussed stochastic linear programs (i.e., $k_1 = k_2 = 0$ in (RP)) where the random component is limited to second-stage right-hand sides as presented in Definition 2, but stated that randomness in other components can easily be included at the expense of increasing the problem size. In these respects, considering the $GR(\Gamma)$ problem with $k_1 = k_2 = 0$, we observe that the constraint sets in both programs $GR(\Gamma)$ and $GR^B(\Omega)$ are identical when $\Gamma = \Omega$. Also note that both Definitions 1 and 2 require Γ and Ω, respectively, to be a set of *distinct* scenario indices. However, the objective functions of the two programs are slightly different. While Birge [2] puts all the remaining mass $(1 - p_0)$ from the reference scenario to each of the other scenarios in the group, we distribute this remaining mass to each of the other scenarios in the group according to their conditional probabilities.

When $k = 1$, we pair a scenario $\xi^i \in \Xi$ with the *reference scenario* ξ^0. With the selected pair of scenarios ξ^0 and ξ^i, we construct the so-called *pairs subproblem*, introduced in [2], as:

$$(P2) \qquad \begin{aligned} \min \quad z(x,\xi^0,\xi^i) &= c^T x + p_0 d_0^T y_0 + (1-p_0) d_i^T y_i \\ \text{s.t.} \quad Ax &= b, \\ T_0 x + W_0 y_0 &= h_0, \\ T_i x + W_i y_i &= h_i, \\ x &\in \mathbb{X}, \\ y_0, y_i &\in \mathbb{Y}. \end{aligned}$$

It is easy to check that our definition of the pairs subproblem, $GR(\{i\})$, and Birge's definition, $GR^B(\{i\})$, are identical for SLPs (i.e., when $k_1 = k_2 = 0$).

3 Lower bounds from group subproblems

In this section, we first introduce a lower bounding formula for general SMIPs based on the solution of the group subproblems introduced in Definition 1. The construction of our bounding formula is also inspired by Birge [2]. We then show that these lower bounds satisfy the monotonic nondecreasing property as the scenario group size increases. Finally, we present the bounds developed by Birge [2] and show that these bounds are not necessarily correct nor do they satisfy the monotonicity property.

Definition 3 *Given an integer k such that $1 \leq k \leq K$, the expected value of group subproblem objective functions with k scenarios in each group, $EGSO(k)$, is defined as:*

$$EGSO(k) \equiv \frac{1}{\binom{K-1}{k-1}(1-p_0)} \left[\sum_{\Gamma \in \mathcal{P}_k(S)} \rho(\Gamma) \cdot z^*(\Gamma) \right], \tag{5}$$

where $z^*(\Gamma)$ is the optimal objective function value of the $GR(\Gamma)$ problem for $\Gamma \in \mathcal{P}_k(S)$ (see Definition 1).

$EGSO(k)$ requires solving all possible group subproblems with k scenarios and taking the expected value of the optimal objective values. In taking this expectation, $EGSO(k)$ assigns a probability distribution to the group subproblems, where the probability assigned to a group subproblem $GR(\Gamma)$ is simply the sum of the probabilities of the scenarios in Γ (i.e., $\rho(\Gamma)$). Note that, for $1 \leq k \leq K$,

$$\sum_{\Gamma \in \mathcal{P}_k(S)} \rho(\Gamma) = \binom{K-1}{k-1} \cdot \rho(S) = \binom{K-1}{k-1} \cdot \sum_{i \in S} p_i = \binom{K-1}{k-1}(1-p_0),$$

where the first equality follows from Lemma 1 (see Section 3.1) with $n = K$ and $m = k$. Therefore, the quantity $\left[\binom{K-1}{k-1}(1-p_0)\right]^{-1}$ can be viewed as a normalization factor. As a result, the $EGSO$ formula can be viewed as a conditional expectation of the optimal objective values of the group subproblems.

Note that the definition of $EGSO(k)$ depends on the choice of the reference scenario. However, for notational simplicity, we suppress such dependency on the left-hand side of equation (5).

A special case of $EGSO(k)$ is when $k = 1$. We let $EPSO = EGSO(1)$ and call it the *expected value of pairs subproblem objective functions*. It is a weighted sum of the optimal objective values of problem ($P2$) for each possible pair (ξ^0, ξ^i):

$$EPSO = \frac{1}{(1-p_0)} \sum_{i \in S} p_i \cdot \min z(x, \xi^0, \xi^i) = \frac{1}{(1-p_0)} \sum_{i \in S} p_i \cdot z^*(\{i\}),$$

where $z^*(\{i\})$ is the optimal objective function value of the group subproblem $GR(\{i\})$ for $i \in S$. Clearly, when $k_1 = k_2 = 0$, $EPSO$ is identical to the so-called *sum of pairs expected values* ($SPEV \equiv SGEV(1)$) introduced by Birge [2].

3.1 Properties of $EGSO$

In this section, we prove that $EGSO(k)$ provides a lower bound for RP for any $k = 1, \ldots, K$; that all these lower bounds are at least as good as the WS bound; and that

$EGSO(k)$ is monotonically nondecreasing in k.

Lemma 1 *Given integers m and n such that $1 \leq m \leq n \leq K$ and a scenario index set $\Gamma_n \in \mathcal{P}_n(S)$:*

$$\sum_{\Gamma_m \in \mathcal{P}_m(\Gamma_n)} \rho(\Gamma_m) = \binom{n-1}{m-1} \rho(\Gamma_n). \tag{6}$$

Proof. For $m = 1, 2, \ldots, n$, consider the set $\mathcal{P}_m(\Gamma_n)$. Observe that a scenario index $i \in \Gamma_n$ is contained in exactly $\binom{n-1}{m-1}$ elements in $\mathcal{P}_m(\Gamma_n)$. This observation follows from the fact that given n values to fill in an m-tuple, if we fix one of the elements of the tuple to a particular value i, then we are left with $n-1$ values to choose from for the remaining $m-1$ positions. The intended result follows directly from this observation. □

Lemma 2 *Given integers ℓ, m, and n such that $1 \leq \ell \leq m \leq n \leq K$, a scenario index set $\Gamma_n \in \mathcal{P}_n(S)$, and a function $f(\cdot) : \mathcal{P}_\ell(\Gamma_n) \rightarrow \mathbb{R}$:*

$$\sum_{\Gamma_m \in \mathcal{P}_m(\Gamma_n)} \sum_{\Gamma_\ell \in \mathcal{P}_\ell(\Gamma_m)} \rho(\Gamma_\ell) f(\Gamma_\ell) = \binom{n-\ell}{m-\ell} \sum_{\Gamma_\ell \in \mathcal{P}_\ell(\Gamma_n)} \rho(\Gamma_\ell) f(\Gamma_\ell). \tag{7}$$

Proof. Similar to the proof of Lemma 1. □

Corollary 1 *Given integers m and n such that $1 \leq m \leq n \leq K$, a scenario index set $\Gamma_n \in \mathcal{P}_n(S)$ and a function $f(\cdot) : \mathcal{P}_1(\Gamma_n) \rightarrow \mathbb{R}$:*

$$\sum_{\Gamma_m \in \mathcal{P}_m(\Gamma_n)} \sum_{i \in \Gamma_m} p_i f(\{i\}) = \binom{n-1}{m-1} \sum_{i \in \Gamma_n} p_i f(\{i\}). \tag{8}$$

Proof. Setting $\ell = 1$ in Lemma 2 yields the desired result. □

Lemma 3 *Given an integer ℓ such that $1 \leq \ell \leq K$ and a scenario index set $\Gamma_\ell \in \mathcal{P}_\ell(S)$,*

$$(\ell - 1) \cdot \rho(\Gamma_\ell) \cdot z^*(\Gamma_\ell) \geq \sum_{\Gamma_{\ell-1} \in \mathcal{P}_{\ell-1}(\Gamma_\ell)} \rho(\Gamma_{\ell-1}) \cdot z^*(\Gamma_{\ell-1}). \tag{9}$$

Proof. For $\Gamma_\ell = \{i_1, \ldots, i_\ell\}$, where $1 \leq i_1 < \cdots < i_\ell \leq K$, let $(\tilde{x}, \tilde{y}_0, \tilde{y}_{i_1}, \ldots, \tilde{y}_{i_\ell})$ be an optimal solution to $GR(\Gamma_\ell)$. For any scenario group $\Gamma_{\ell-1} = \{j_1, \ldots, j_{\ell-1}\} \in \mathcal{P}_{\ell-1}(\Gamma_\ell)$, it is clear that $(\tilde{x}, \tilde{y}_0, \tilde{y}_{j_1}, \ldots, \tilde{y}_{j_{\ell-1}})$ is a feasible solution to $GR(\Gamma_{\ell-1})$. Let $(\hat{x}, \hat{y}_0, \hat{y}_{j_1}, \ldots, \hat{y}_{j_{\ell-1}})$ be an optimal solution to $GR(\Gamma_{\ell-1})$. Therefore,

$$\begin{aligned} c^T\tilde{x} + p_0 d_0^T \tilde{y}_0 \quad &+ \quad (1-p_0) \sum_{i \in \Gamma_{\ell-1}} \frac{p_i}{\rho(\Gamma_{\ell-1})} d_i^T \tilde{y}_i \\ &\geq \quad c^T\hat{x} + p_0 d_0^T \hat{y}_0 + (1-p_0) \sum_{i \in \Gamma_{\ell-1}} \frac{p_i}{\rho(\Gamma_{\ell-1})} d_i^T \hat{y}_i \\ &= \quad z^*(\Gamma_{\ell-1}). \end{aligned} \tag{10}$$

Multiplying inequality (10) by $\rho(\Gamma_{\ell-1})$, we obtain

$$\rho(\Gamma_{\ell-1}) \cdot \left(c^T\tilde{x} + p_0 d_0^T \tilde{y}_0\right) + (1-p_0) \sum_{i \in \Gamma_{\ell-1}} p_i d_i^T \tilde{y}_i \geq \rho(\Gamma_{\ell-1}) \cdot z^*(\Gamma_{\ell-1}). \tag{11}$$

Summing inequality (11) for all $\Gamma_{\ell-1} \in \mathcal{P}_{\ell-1}(\Gamma_\ell)$, we obtain

$$\begin{aligned} \sum_{\Gamma_{\ell-1} \in \mathcal{P}_{\ell-1}(\Gamma_\ell)} \rho(\Gamma_{\ell-1}) \cdot \left(c^T\tilde{x} + p_0 d_0^T \tilde{y}_0\right) \quad &+ \quad (1-p_0) \cdot \left(\sum_{\Gamma_{\ell-1} \in \mathcal{P}_{\ell-1}(\Gamma_\ell)} \sum_{i \in \Gamma_{\ell-1}} p_i d_i^T \tilde{y}_i \right) \\ &\geq \quad \sum_{\Gamma_{\ell-1} \in \mathcal{P}_{\ell-1}(\Gamma_\ell)} \rho(\Gamma_{\ell-1}) \cdot z^*(\Gamma_{\ell-1}). \end{aligned} \tag{12}$$

Lemma 1 with $m = \ell - 1$ and $n = \ell$ implies that $\sum_{\Gamma_{\ell-1} \in \mathcal{P}_{\ell-1}(\Gamma_\ell)} \rho(\Gamma_{\ell-1}) = (\ell - 1)\rho(\Gamma_\ell)$. Furthermore, Corollary 1 with $m = \ell - 1$, $n = \ell$, and $f(\{i\}) = d_i^T \tilde{y}_i$ implies that $\sum_{\Gamma_{\ell-1} \in \mathcal{P}_{\ell-1}(\Gamma_\ell)} \sum_{i \in \Gamma_{\ell-1}} p_i d_i^T \tilde{y}_i = (\ell - 1) \cdot \sum_{i \in \Gamma_\ell} p_i d_i^T \tilde{y}_i$. With these results, we can rewrite inequality (12) as:

$$\begin{aligned} (\ell-1)\rho(\Gamma_\ell) \left(c^T\tilde{x} + p_0 d_0^T \tilde{y}_0\right) \quad &+ \quad (1-p_0)(\ell-1) \sum_{i \in \Gamma_\ell} p_i d_i^T \tilde{y}_i \\ &\geq \quad \sum_{\Gamma_{\ell-1} \in \mathcal{P}_{\ell-1}(\Gamma_\ell)} \rho(\Gamma_{\ell-1}) \cdot z^*(\Gamma_{\ell-1}). \end{aligned} \tag{13}$$

Finally, factoring out $(\ell-1)\rho(\Gamma_\ell)$ from the left-hand-side of inequality (13), we obtain

$$(\ell-1)\rho(\Gamma_\ell)\cdot\left[c^T\tilde{x}+p_0d_0^T\tilde{y}_0+(1-p_0)\cdot\sum_{i\in\Gamma_\ell}\frac{p_i}{\rho(\Gamma_\ell)}d_i^T\tilde{y}_i\right]$$
$$\geq\sum_{\Gamma_{\ell-1}\in\mathcal{P}_{\ell-1}(\Gamma_\ell)}\rho(\Gamma_{\ell-1})\cdot z^*(\Gamma_{\ell-1}). \quad (14)$$

Note that the term in the square brackets on the left-hand side of inequality (14) is $z^*(\Gamma_\ell)$, which yields the desired result. □

Given a scenario group $\Gamma_\ell\in\mathcal{P}_\ell(S)$, Lemma 3 provides a lower bound for the optimal objective function value of the group subproblem associated with Γ_ℓ by computing the optimal objective function values of the group subproblems in $\mathcal{P}_{\ell-1}(\Gamma_\ell)$.

Theorem 1, one of the main results of this paper, constructs progressively tighter lower bounds for RP by increasing the number of scenarios in the group subproblems, and shows that the worst of these bounds is at least as good as the WS bound.

Theorem 1 *For any chosen reference scenario,*

$$WS\leq EPSO=EGSO(1)\leq EGSO(2)\leq\cdots\leq EGSO(K-1)\leq EGSO(K)=RP.$$

Proof. The proof will be completed in three steps by showing that:

$$(i)\quad WS\leq EPSO=EGSO(1),$$
$$(ii)\quad EGSO(k)\leq EGSO(k+1),\qquad \text{for } k=1,\ldots,K-1,$$
$$(iii)\quad EGSO(K)=RP.$$

Step (*i*). Birge and Louveaux [4] proved the inequality $WS\leq SPEV$ for stochastic linear programs (i.e., when $k_1=k_2=0$ in (RP)). When $k_1=k_2=0$, $EGSO(1)=EPSO=SPEV=SGEV(1)$ by definition. We also note that the proof given in Birge and Louveaux [4] trivially extends to general SMIPs (i.e., their proof is also valid for the program (RP) even when $k_1>0, k_2>0$). Hence, we conclude $WS\leq EPSO$.

Step (*ii*). Consider any integer k such that $1 \le k \le K-1$. For any $\Gamma_{k+1} \in \mathcal{P}_{k+1}(S)$, Lemma 3 implies

$$k \cdot \rho(\Gamma_{k+1}) \cdot z^*(\Gamma_{k+1}) \ge \sum_{\Gamma_k \in \mathcal{P}_k(\Gamma_{k+1})} \rho(\Gamma_k) \cdot z^*(\Gamma_k). \tag{15}$$

Summing inequalities (15) over all $\Gamma_{k+1} \in \mathcal{P}_{k+1}(S)$ yields

$$\sum_{\Gamma_{k+1} \in \mathcal{P}_{k+1}(S)} [k\rho(\Gamma_{k+1})z^*(\Gamma_{k+1})] \ge \sum_{\Gamma_{k+1} \in \mathcal{P}_{k+1}(S)} \left[\sum_{\Gamma_k \in \mathcal{P}_k(\Gamma_{k+1})} \rho(\Gamma_k)z^*(\Gamma_k) \right]. \tag{16}$$

By definition of $EGSO(k+1)$, for $k = 1, 2, \dots, K-1$, we can rewrite the quantity on the left-hand side of inequality (16) as:

$$\sum_{\Gamma_{k+1} \in \mathcal{P}_{k+1}(S)} [k\rho(\Gamma_{k+1})z^*(\Gamma_{k+1})] = k \cdot \binom{K-1}{k} \cdot (1-p_0) \cdot EGSO(k+1). \tag{17}$$

Furthermore, applying Lemma 2 with $\ell = k$, $m = k+1$, $n = K$, and $f(\cdot) = z^*(\cdot)$, the right-hand side of inequality (16) can be rewritten as:

$$\begin{aligned} \sum_{\Gamma_{k+1} \in \mathcal{P}_{k+1}(S)} \left[\sum_{\Gamma_k \in \mathcal{P}_k(\Gamma_{k+1})} \rho(\Gamma_k)z^*(\Gamma_k) \right] &= (K-k) \sum_{\Gamma_k \in \mathcal{P}_k(S)} \rho(\Gamma_k)z^*(\Gamma_k) \\ &= (K-k) \cdot \binom{K-1}{k-1} \cdot (1-p_0) \cdot EGSO(k), \end{aligned} \tag{18}$$

where the last equality follows from the definition of $EGSO(k)$.

Finally, substituting the right-hand sides of equalities (17) and (18) into inequality (16), we obtain:

$$k \cdot \binom{K-1}{k} \cdot (1-p_0) \cdot EGSO(k+1) \ge (K-k) \cdot \binom{K-1}{k-1} \cdot (1-p_0) \cdot EGSO(k), \tag{19}$$

which yields the desired result after canceling out the identical terms on both sides.

Step (*iii*). By definition $EGSO(K) = z^*(S)$, where $z^*(S)$ is, in fact, the optimal objective function value to the extensive form of the recourse problem (see problem (*EF*)). Therefore, $EGSO(K) = RP$. $\square$

3.2 Lower Bounds in Birge [2]

Birge [2] developed a sequence of bounds for SLPs based on the group subproblems given in Definition 2. We first state this bounding formula and then compare it with our $EGSO$ formula.

Definition 4 *(Birge [2]) The sum of group expected values with k scenarios, $SGEV(k)$, for a stochastic linear program is defined as:*

$$SGEV(k) \equiv \frac{1}{(1-p_0)^k} \left[\sum_{i_1=1}^{K} \sum_{i_2 \geq i_1}^{K} \cdots \sum_{i_k \geq i_{k-1}}^{K} \left(\prod_{j=1}^{k} p_{i_j} \right) z^B(\Omega) \right], \tag{20}$$

where Ω is the set of distinct indices among $i_1, i_2, \ldots, i_k$, and $z^B(\Omega)$ is the optimal objective function value of the program $GR^B(\Omega)$.

In the $SGEV(k)$ formula, Birge allows the repetition of the scenario indices since $i_j \geq i_{j-1}$ for $j = 2, \ldots, k$. However, when forming the group subproblem $GR^B(\Omega)$ to obtain the value of $z^B(\Omega)$, which is required to compute $SGEV(k)$, he eliminates the repeated indices among $i_1, i_2, \ldots, i_k$.

Furthermore, for any integer $k > 1$, the term $\sum_{i_1=1}^{K} \sum_{i_2 \geq i_1}^{K} \cdots \sum_{i_k \geq i_{k-1}}^{K} \left(\prod_{j=1}^{k} p_{i_j} \right)$ does not, in general, equal $(1-p_0)^k$. And this point is where our $EGSO(k)$ formula distinguishes itself from the $SGEV(k)$ formula. In computing $EGSO(k)$, we assign a probability $\rho(\Gamma) = \sum_{i \in \Gamma} p_i$ to the group subproblem $\Gamma \in \mathcal{P}_k(S)$ and normalize this probability by dividing it to the relative share of the particular group subproblem of interest (Γ) in the entire set of group subproblems ($\mathcal{P}_k(S)$), i.e., normalize with $\sum_{\Gamma \in \mathcal{P}_k(S)} \rho(\Gamma) = \binom{K-1}{k-1}(1-p_0)$.

3.3 A counterexample for $SGEV$

Birge [2] claimed that the $SGEV$ formula presented in Definition 4 above provides successively tighter lower bounds for the recourse problem and that the worst $SGEV$ bound is as good as the WS bound (see Lemma 4.4 in Birge [2]), i.e., the following inequalities hold

$$WS \leq SPEV = SGEV(1) \leq SGEV(2) \leq \cdots \leq SGEV(K-1) \leq RP. \qquad (21)$$

Our experience with the $SGEV$ formula revealed that although the inequality $WS \leq SGEV(1)$ is always correct, the inequalities $SGEV(k) \leq SGEV(k+1)$ for $k = 1, 2, \ldots, K-2$ do not necessarily hold in general. We provide a counterexample that shows that $SGEV(k)$ ($k > 1$) bound can be even worse than the WS bound.

Remark 1 *Lemma 4.4 of Birge [2] is not necessarily correct.*

Consider the following two-stage stochastic linear programming problem as adopted from Birge and Louveaux [4, page 149]:

$$\begin{aligned} \min \quad & 3x_1 + 2x_2 + \mathbb{E}_{\tilde{\xi}}\left[Q\left(x, \tilde{\xi}\right)\right] \\ \text{s.t.} \quad & x_1, x_2 \geq 0, \end{aligned}$$

where

$$\begin{aligned} Q\left(x, \tilde{\xi}\right) = \min \ & -15y_1 - 12y_2 \\ \text{s.t.} \quad & 3y_1 + 2y_2 \leq x_1, \\ & 2y_1 + 5y_2 \leq x_2, \\ & 0.8\xi_1 \leq y_1 \leq \xi_1, \\ & 0.8\xi_2 \leq y_2 \leq \xi_2, \\ & y_1, y_2 \geq 0, \end{aligned}$$

and $\tilde{\xi} = \xi^i \equiv \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$ with probability $p_i = \frac{1}{4}$ for $i = 1, \ldots, 4$, and

$$\xi^1 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \ \xi^2 = \begin{bmatrix} 4 \\ 8 \end{bmatrix}, \ \xi^3 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}, \ \text{ and } \ \xi^4 = \begin{bmatrix} 6 \\ 8 \end{bmatrix}.$$

To obtain the WS bound, we solve problem ($P1$) for each ξ^i. Based on the results summarized in Table 1, we compute WS as

$$WS = \frac{1}{4}\,(4.8 + 17.6 + 0.8 + 13.6) = 9.2.$$

We solve problem (EF) to find $RP = 30.94$.

In order to compute $SGEV(\cdot)$, we first pick a reference scenario, say ξ^4, and set $\xi^0 = \xi^4$ and $p_0 = p_4 = \frac{1}{4}$. Then, for $k = 1, 2, 3$, we form the appropriate group subproblems $GR^B(\Omega)$ and compute $SGEV(k)$. Tables 2, 3, and 4 summarize the results of solving the appropriate group subproblems for computing $SGEV(1)$, $SGEV(2)$, and $SGEV(3)$, respectively.

Based on the results summarized in Table 2, we compute $SGEV(1)$ as

$$SGEV(1) = \frac{1}{\left(1 - \frac{1}{4}\right)} \left[\frac{1}{4}\,(46.6 + 22.12 + 24.1)\right] = 30.94.$$

Based on the results summarized in Table 3, we compute $SGEV(2)$ as

$$SGEV(2) = \frac{1}{\left(1 - \frac{1}{4}\right)^2} \left[\frac{1}{16}\,(46.6 - 58.88 - 56.9 + 22.12 - 81.38 + 24.1)\right] = -11.593.$$

Table 1: Computation of WS

Scenario index (i)	1	2	3	4
p_i	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$\min z(x, \xi^i)$	4.8	17.6	0.8	13.6

Based on the results summarized in Table 4, we compute $SGEV(3)$ as

$$SGEV(3) = \frac{1}{\left(1-\frac{1}{4}\right)^3}\left[\frac{1}{64}\left(46.6 - 58.88 - 56.9 - \cdots - 81.38 + 24.1\right)\right] = -17.181.$$

These results clearly show that $SGEV(2) \ngeq SGEV(1)$, and more surprisingly $SGEV(2) \ngeq WS$. Furthermore, $SGEV(3) \ngeq SGEV(1)$, $SGEV(3) \ngeq SGEV(2)$ and $SGEV(3) \ngeq WS$.

Table 2: Computation of $SGEV(1)$

Scenario index (i_1)	Ω	p_{i_1}	$z^B(\Omega)$
1	1	$\frac{1}{4}$	46.6
2	2	$\frac{1}{4}$	22.12
3	3	$\frac{1}{4}$	24.1

Table 3: Computation of $SGEV(2)$

Scenario indices (i_1, i_2)	Ω	$p_{i_1} \cdot p_{i_2}$	$z^B(\Omega)$
1, 1	1	$\frac{1}{16}$	46.6
1, 2	1, 2	$\frac{1}{16}$	−58.88
1, 3	1, 3	$\frac{1}{16}$	−56.9
2, 2	2	$\frac{1}{16}$	22.12
2, 3	2, 3	$\frac{1}{16}$	−81.38
3, 3	3	$\frac{1}{16}$	24.1

Table 4: Computation of $SGEV(3)$

Scenario indices (i_1, i_2, i_3)	Ω	$p_{i_1} \cdot p_{i_2} \cdot p_{i_3}$	$z^B(\Omega)$
1, 1, 1	1	$\frac{1}{64}$	46.6
1, 1, 2	1, 2	$\frac{1}{64}$	−58.88
1, 1, 3	1, 3	$\frac{1}{64}$	−56.9
1, 2, 2	1, 2	$\frac{1}{64}$	−58.88
1, 2, 3	1, 2, 3	$\frac{1}{64}$	−162.38
1, 3, 3	1, 3	$\frac{1}{64}$	−56.9
2, 2, 2	2	$\frac{1}{64}$	22.12
2, 2, 3	2, 3	$\frac{1}{64}$	−81.38
2, 3, 3	2, 3	$\frac{1}{64}$	−81.38
3, 3, 3	3	$\frac{1}{64}$	24.1

For the same setting, we compute $EGSO$ bounds as follows. Since ξ^4 is used as a reference scenario, we set $\xi^0 = \xi^4$, $p_0 = \frac{1}{4}$, $K = 3$, and hence $S = \{1, 2, 3\}$. Then, for $k = 1, 2, 3$, we form the appropriate group subproblems $GR(\Gamma_k)$ and compute $EGSO(k)$. Tables 5, 6, and 7 summarize the results of solving the appropriate group subproblems for computing $EGSO(1)$, $EGSO(2)$, and $EGSO(3)$, respectively.

Table 5: Computation of $EGSO(1)$

Γ_1	$\rho(\Gamma_1)$	$z^*(\Gamma_1)$
1	$\frac{1}{4}$	46.6
2	$\frac{1}{4}$	22.12
3	$\frac{1}{4}$	24.1

Table 6: Computation of $EGSO(2)$

Γ_2	$\rho(\Gamma_2)$	$z^*(\Gamma_2)$
1, 2	$\frac{2}{4}$	34.36
1, 3	$\frac{2}{4}$	35.35
2, 3	$\frac{2}{4}$	23.11

Table 7: Computation of $EGSO(3)$

Γ_3	$\rho(\Gamma_3)$	$z^*(\Gamma_3)$
1, 2, 3	$\frac{3}{4}$	30.94

Based on the results summarized in Table 5, we compute $EGSO(1)$ as

$$EGSO(1) = \frac{1}{\binom{3-1}{1-1}\left(1 - \frac{1}{4}\right)} \left[\frac{1}{4}\left(46.6 + 22.12 + 24.1\right)\right] = 30.94.$$

Based on the results summarized in Table 6, we compute $EGSO(2)$ as

$$EGSO(2) = \frac{1}{\binom{3-1}{2-1}\left(1 - \frac{1}{4}\right)} \left[\frac{2}{4}\left(34.36 + 35.35 + 23.11\right)\right] = 30.94.$$

Based on the results summarized in Table 7, we compute $EGSO(3)$ as

$$EGSO(3) = \frac{1}{\binom{3-1}{3-1}\left(1 - \frac{1}{4}\right)} \left[\frac{3}{4}\left(30.94\right)\right] = 30.94.$$

Alternatively, if we pick $\bar{\xi} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$ as our reference scenario, we set $p_0 = 0$, because $\bar{\xi} \notin \Xi$. With this choice of the reference scenario, we find $SGEV(1) = 14$, $SGEV(2) = -38.4$, and $SGEV(3) = -35.431$. Clearly, again, $SGEV(2) \ngeq SGEV(1)$ and $SGEV(2) \ngeq WS$. Furthermore, $SGEV(3) \ngeq SGEV(1)$ and $SGEV(3) \ngeq WS$. But for this setting, our proposed $EGSO$ bounds yield $EGSO(1) = 14$, $EGSO(2) = 23.78$, and $EGSO(3) = 28.56$.

This example clearly shows that inequalities (21) may not always hold. As a result, Theorem 4.6 in [2], which requires inequalities (21), may not always hold.

4 Upper bounds from group subproblems

Birge [2] used optimal first-stage solutions from his group subproblems given in Definition 2 to obtain upper bounds for RP. For this purpose, let Ω be a set of distinct scenario indices, $x^B(\Omega)$ be an optimal first-stage solution of the group subproblem $GR^B(\Omega)$, and $\mathcal{P}'_k(S) = \mathcal{P}_k(S) \bigcup \mathcal{P}_{k-1}(S) \bigcup \cdots \bigcup \mathcal{P}_1(S) \bigcup \{0\}$. Then, the *expectation of group expected values solution with k scenarios*, denoted $EGEV(k)$, is computed as

$$EGEV(k) = \min_{\Omega \in \mathcal{P}'_k(S)} \mathbb{E}_{\tilde{\xi}} \left[z^* \left(x^B(\Omega), \tilde{\xi} \right) \right].$$

The special case when $k = 1$ is called the *expectation of pairs expected value solutions* ($EPEV$).

Birge [2] showed that when the reference scenario choice in $GR^B(\Omega)$ problems is $\bar{\xi}$, the $EGEV$ formula provides tighter bounds than the EEV bound and that the $EGEV(k)$ bounds become successively tighter as k increases, i.e.,

$$RP \leq EGEV(K-1) \leq EGEV(K-2) \leq \cdots \leq EGEV(1) \leq EEV.$$

Birge and Louveaux [4] generalized the EEV solution concept to incorporate the solutions that can be obtained by different choices for the reference scenario. That is, if we replace $\bar{\xi}$ in problem ($P1$) by any other scenario choice ξ^0 and use the resulting optimal first-stage solution of this new problem, denoted x^0, to evaluate

$$\mathbb{E}_{\tilde{\xi}}\left[z^*\left(x^0,\tilde{\xi}\right)\right], \tag{22}$$

we find a new solution concept called the *expected value of the reference scenario* ($EVRS$) [4]. Similar to computing EEV, if x^0 yields an infeasible problem ($P1$) when computing $z^*(x^0,\xi)$ for any $\xi \in \Xi$, then we set the expectation in (22) to $+\infty$. With this convention, it can be shown that $RP \leq EVRS$ [4].

We adapt these two ideas to construct an upper bounding formula for general SMIPs.

Definition 5 *Given an integer k such that $1 \leq k \leq K$, the expected value of using the first-stage solutions of group subproblems with k scenarios in the group, $EFGS(k)$, is defined as:*

$$EFGS(k) \equiv \min_{\Gamma \in \mathcal{P}_k(S) \bigcup \{0\}} \mathbb{E}_{\tilde{\xi}}\left[z^*\left(x^G\left(\Gamma\right),\tilde{\xi}\right)\right], \tag{23}$$

where $x^G(\Gamma)$ is an optimal first-stage solution to the $GR(\Gamma)$ problem for the given scenario index set $\Gamma \in \mathcal{P}_k(S) \bigcup \{0\}$.

Following the convention used in computing EEV and $EVRS$, if the first-stage solution of a group subproblem associated with any $\Gamma \in \mathcal{P}_k(S) \bigcup \{0\}$ yields an infeasible problem ($P1$) for some $\xi \in \Xi$, then we define the expected value in equation (23) to be $+\infty$, because such a first-stage solution would also be infeasible for the recourse problem. Also note that, as we did in $EGSO(k)$, we suppress the dependency of $EFGS(k)$ to the choice of the reference scenario.

Our formula differs from Birge's $EGEV$ formula mainly in two aspects. First, our formula is based on the group subproblems introduced in Definition 1, whereas the $EGEV$ formula is based on the group subproblems given in Definition 2. The second difference is the number of group subproblems to be solved: for a given group size k, the $EGEV(k)$ formula requires optimizing $|\mathcal{P}'_k(S)|$ group subproblems, whereas our approach requires optimizing only $|\mathcal{P}_k(S)| + 1$ subproblems. Clearly, the difference $|\mathcal{P}'_k(S)| - |\mathcal{P}_k(S)|$ can grow very fast as k increases. Also note that computing the expectations in both upper bounding formulas requires optimizing an additional $|\Xi|$ recourse problems for each group subproblem needed for the bound. Each of these additional problems are identical for both bounding formulas except for may be the first-stage solutions they use. Moreover, although each of these additional problems may be extremely fast to optimize, the overall time required for computing $EGEV(k)$, for any given group size k, will be significantly higher than that required for computing $EFGS(k)$ due to the difference $|\mathcal{P}'_k(S)| - |\mathcal{P}_k(S)|$. Finally, our $EFGS$ formula applies to general SMIPs, but the $EGEV$ formula is presented in [2] for specific SLPs although we believe it could also be generalized to general SMIPs.

Proposition 1 shows that, for any given group size k, the $EFGS(k)$ formula yields a bound that is as strong as the $EVRS$ bounds.

Proposition 1 *For any chosen reference scenario, $EFGS(k) \leq EVRS$ for any integer k such that $1 \leq k \leq K$.*

Proof. Consider any group size k. Let $x^G(\{0\})$ be an optimal first-stage solution to the group subproblem associated with scenario index set $\Gamma = \{0\}$. Then, equation (23) implies that $EFGS(k) \leq \mathbb{E}_{\tilde{\xi}}\left[z^*\left(x^G(\{0\}), \tilde{\xi}\right)\right] = EVRS$. □

Proposition 2 shows that, for any given group size k, the $EFGS(k)$ formula provides an upper bound for RP.

Proposition 2 *For any chosen reference scenario, $RP \leq EFGS(k)$ for any integer k such that $1 \leq k \leq K$.*

Proof. Consider any reference scenario and any integer k such that $1 \leq k \leq K$. The intended result is obvious when the minimum in equation (23) is attained for $\Gamma = \{0\}$.

For a scenario index set $\Gamma \in \mathcal{P}_k(S)$, let $x^G(\Gamma)$ be an optimal first-stage solution to the group subproblem associated with Γ. Substituting $x^G(\Gamma)$ into problem $(P1)$, we calculate the expectation in equation (23) as

$$\begin{aligned}\mathbb{E}_{\tilde{\xi}}\left[z^*\left(x^G(\Gamma),\tilde{\xi}\right)\right] &= \sum_{j=0}^{K} p_j \cdot \left[c^T x^G(\Gamma) + d_j^T y_j^*(\Gamma)\right] \\ &= c^T x^G(\Gamma) + \sum_{j=0}^{K} p_j d_j^T y_j^*(\Gamma), \end{aligned} \tag{24}$$

where $y_j^*(\Gamma)$ denotes an optimal second-stage completion in problem $(P1)$ for the given first-stage solution $x^G(\Gamma)$ and scenario j. Observe that we can construct a solution $\zeta = \left(x^G(\Gamma), y_0^*(\Gamma), y_1^*(\Gamma), \cdots, y_K^*(\Gamma)\right)$ that is feasible for the problem (RP) and that the right-hand side of equation (24) is identical to the objective function of the problem (RP) evaluated at the constructed solution ζ. Therefore, we conclude that $RP \le \mathbb{E}_{\tilde{\xi}}\left[z^*\left(x^G(\Gamma),\tilde{\xi}\right)\right]$ for any $\Gamma \in \mathcal{P}_k(S)$, from which the intended result follows. $\square$

Proposition 3 shows that when all of the available scenario information is used, the $EFGS$ formula yields the true optimal objective value for the (RP) problem. This result, however, is only of theoretical interest as one would never use this approach to find the tightest bound in practice, because such a computation would require even more effort than that required for solving the problem (RP) itself.

Proposition 3 *For any chosen reference scenario,* $EFGS(K) = RP$.

Proof. Consider any reference scenario. $RP \le EFGS(K)$ follows from Proposition 2 with $k = K$.

To prove the reverse inequality, consider the group subproblem associated with the scenario index set $\Gamma \in \mathcal{P}_K(S) = \{S\}$, denoted $GR(S)$, and let $\psi = (\bar{x}, \bar{y}_0, \bar{y}_1, \cdots, \bar{y}_K)$ be an optimal solution to the (RP) problem. Clearly, the solution ψ is also an optimal solution to $GR(S)$. Furthermore, the pair $(\bar{x}, \bar{y}_j)$ is a feasible solution to $(P1)$ when $\xi = \xi^j$ for $j = 0, 1, \ldots, K$. Therefore, $c^T\bar{x} + d_j^T\bar{y}_j \ge z^*(\bar{x}, \xi^j)$ for $\xi^j \in \Xi$. Multiplying both sides of this inequality by p_j and summing the resulting inequalities yields

$$c^T\bar{x} + \sum_{j=0}^{K} p_j d_j^T \bar{y}_j \geq \sum_{j=0}^{K} p_j z^*(\bar{x}, \xi^j).$$

Clearly, the left hand-side of the last inequality is RP, whereas the right hand-side is $\mathbb{E}_{\tilde{\xi}}\left[z^*\left(x^G(\Gamma), \tilde{\xi}\right)\right] = EFGS(K)$. $\square$

We are also interested in whether the $EFGS$ formula introduced in Definition 5 also admits a monotonicity property with respect to the group size (similar to the result in Theorem 1). That is, do the following inequalities hold?

$$RP = EFGS(K) \leq EFGS(K-1) \leq \cdots \leq EFGS(2) \leq EFGS(1). \tag{25}$$

We were not able to prove analytically that the $EFGS$ bound given in Definition 5 admits such a property, but, as we will discuss in Section 5, all the numerical experiments we have tested do not violate inequalities (25). On the other hand, we can prove that such a monotonicity property exists for a slightly modified version of the original $EFGS$ definition, which we provide in Definition 6. The only difference between this modified formula and the $EGEV$ formula is in the group subproblems used.

Definition 6 *Given an integer k such that $1 \leq k \leq K$, the modified $EFGS(k)$, denoted $EFGS'(k)$, is given by:*

$$EFGS'(k) \equiv \min_{\Gamma \in \mathcal{P}'_k(S)} \mathbb{E}_{\tilde{\xi}}\left[z^*\left(x^G(\Gamma), \tilde{\xi}\right)\right], \tag{26}$$

where $x^G(\Gamma)$ is an optimal first-stage solution to the $GR(\Gamma)$ problem for the given scenario index set $\Gamma \in \mathcal{P}'_k(S)$.

Corollary 2 *For any given group size k,*

$$EFGS'(k) = \min_{j \leq k}\{EFGS(j), EVRS\}.$$

Proof. Consider any group size k. The definition of the set $\mathcal{P}'_k(S)$ implies that $EFGS'(k) \leq$

$\mathbb{E}_{\tilde{\xi}}\left[z^*\left(x^G(\{0\}),\tilde{\xi}\right)\right] = EVRS$, where $x^G(\{0\})$ is an optimal first-stage solution to the group subproblem associated with scenario index set $\Gamma = \{0\}$. Furthermore, for integer $j \le k$, $EFGS'(k) \le EFGS(j)$ since $\mathcal{P}'_k(S) \supset \mathcal{P}_j(S)$. □

Theorem 2 shows that the $EFGS'$ formula provides upper bounds for RP that are at least as strong as the $EVRS$ bound and that the $EFGS'$ bounds become successively tighter as k increases. Note also that if $\bar{\xi}$ is chosen as the reference scenario, the inequalities in Theorem 2 still hold with $EVRS$ re-labeled as EEV.

Theorem 2 *For any chosen reference scenario,*

$$RP = EFGS'(K) \le EFGS'(K-1) \le \cdots \le EFGS'(2) \le EFGS'(1) \le EVRS.$$

Proof. Follows directly from Proposition 2 and Corollary 2. □

5 Computational experiments

This section summarizes our computational experiments on a number of instances obtained from the existing literature. Our goal is to gain an understanding of the computational performance of the $EGSO$ and $EFGS$ bounds proposed in Sections 3 and 4, respectively.

We implement the computation of the $EGSO$ and $EFGS$ bounds in the C programming language and use the callable library of CPLEX 11.0 to solve all the necessary group subproblems. We test all instances on a 64-bit machine with 16GB of RAM and a 2.33GHz processor.

We first test our implementation on the simplified example given in Section 3.3. For convenience, we denote this problem set as `bl` and consider the original problem and its 3 variations. The original problem is denoted as `bl_a`. The first variation, denoted `bl_b`, modifies `bl_a` by allowing uncertainty in the second-stage objective function coefficients (d^T), the technology matrix (T), and the recourse matrix (W). The possible values for these

Table 8: Possible values of the additional random components in problem `bl_b`

Scenario	d^T	T	W^1
1	$\begin{bmatrix} -15 & -12 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 3 \\ 2 \end{bmatrix}$
2	$\begin{bmatrix} -10 & -9 \end{bmatrix}$	$\begin{bmatrix} -0.5 & 0 \\ 0 & -0.5 \end{bmatrix}$	$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$
3	$\begin{bmatrix} -20 & -15 \end{bmatrix}$	$\begin{bmatrix} -1.5 & 0 \\ 0 & -1.5 \end{bmatrix}$	$\begin{bmatrix} 4 \\ 3 \end{bmatrix}$
4	$\begin{bmatrix} -17 & -13 \end{bmatrix}$	$\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$	$\begin{bmatrix} 5 \\ 4 \end{bmatrix}$

Note. W^1 denotes the coefficients of y_1 in the first two constraints.

additional random components are summarized in Table 8. The second variation, denoted `bl_c`, modifies `bl_a` such that $p_1 = 1/10$, $p_2 = 2/10$, $p_3 = 3/10$, and $p_4 = 4/10$. Finally, the third variation, denoted `bl_d`, modifies `bl_a` by adding integrality restrictions to the second stage decision variables y_1 and y_2.

Table 9 summarizes the results for the `bl` problem set. The row labeled K indicates the total number of scenarios in the problem, ξ^0 indicates the reference scenario used in calculations, and the remainder of the table summarizes the computed quantities.

The results in Table 9 confirm our findings from Sections 3 and 4: (i) both $EGSO$ and $EFGS$ formulas provide true lower and upper bounds, respectively, for RP, and (ii) both bounds monotonically improve with increased group size. We indeed observe that the worst $EGSO$ bound (i.e., $EGSO(1)$) is significantly better than the widely used WS bound for all four instances. Similar observation holds comparing $EFGS(1)$ to the widely used EEV (or, $EVRS$) bound. Finally, it is interesting to observe that the $EFGS$ bounds significantly outperform the $EGSO$ bounds. In 3 of the 4 instances, $EFGS$ yields the true optimal objective value even when $k = 1$. For `bl_c`, $EFGS$ converges to the true optimum earlier than $EGSO$. This observation regarding the comparative performances of $EGSO$ and $EFGS$ formulas is further supported with other test instances that we discuss later in this section.

Table 9: Results for the bl problem set

Instance	bl_a	bl_b	bl_c	bl_d
K	4	4	4	4
ξ^0	1	1	1	1
EEV	$+\infty$	$+\infty$	$+\infty$	$+\infty$
$EVRS$	$+\infty$	$+\infty$	$+\infty$	$+\infty$
$EFGS(1)$	30.940	110.100	32.200	39.750
$EFGS(2)$	30.940	110.100	24.952	39.750
$EFGS(3)$	30.940	110.100	24.952	39.750
RP	**30.940**	**110.100**	**24.952**	**39.750**
$EGSO(3)$	30.940	110.100	24.952	39.750
$EGSO(2)$	28.237	67.398	23.347	35.417
$EGSO(1)$	18.280	22.522	14.156	23.167
WS	9.200	8.217	9.680	12.000

Table 10: Effect of the reference scenario choice

Instance	ξ^0	$EGSO(1)$	$EGSO(2)$	$EGSO(3)$	$EFGS(3)$	$EFGS(2)$	$EFGS(1)$
bl_a	1	18.280	28.237	30.940	30.940	30.940	30.940
	2	23.313	27.577	30.940	30.940	30.940	30.940
	3	15.607	27.487	30.940	30.940	30.940	30.940
	4	30.940	30.940	30.940	30.940	30.940	30.940
bl_b	1	22.522	67.398	110.100	110.100	110.100	110.100
	2	110.100	110.100	110.100	110.100	110.100	110.100
	3	22.400	67.399	110.100	110.100	110.100	110.100
	4	21.967	66.518	110.100	110.100	110.100	110.100
bl_c	1	14.156	23.347	24.952	24.952	24.952	32.200
	2	20.186	23.244	24.952	24.952	24.952	24.952
	3	17.707	23.236	24.952	24.952	24.952	24.952
	4	24.952	24.952	24.952	24.952	24.952	24.952
bl_d	1	23.167	35.417	39.750	39.750	39.750	39.750
	2	31.083	35.417	39.750	39.750	39.750	39.750
	3	20.667	35.417	39.750	39.750	39.750	39.750
	4	38.750	39.750	39.750	39.750	39.750	39.750

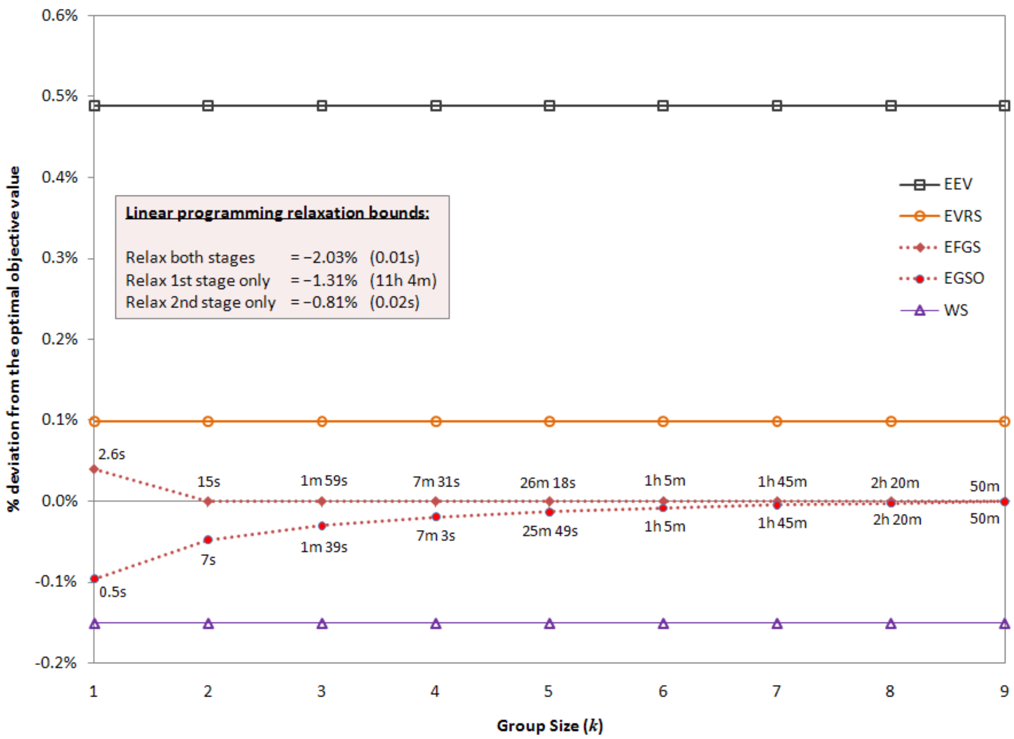

Figure 1: Results for `sizes10` test instance

We arbitrarily choose scenario 1 as the reference scenario. For any given group size k, choosing alternative scenarios as the reference scenario can potentially change $EGSO(k)$ and $EFGS(k)$. The choice of the best reference scenario (i.e., the one that gives the tightest bounds) is desirable, however it is not obvious which scenario would be best *a priori* (see the results summarized in Table 10 for the performance of choosing alternative reference scenarios for the `bl` problem set). On the other hand, knowledge of special structures in a particular instance may lead to better choices for the reference scenario.

Next, we test the `sizes` problem set of Jorjani et al. [10]. This problem set consists of two-stage multi-period stochastic mixed-integer program instances arising in product planning (or, lot-sizing) applications. These instances have continuous and integer variables in both stages.

Figure 1 shows the results for a 10-scenario instance of the `sizes` problem, labeled `sizes10`. The extensive form of this test instance has 341 constraints and 825 variables of which 110 are binary. The horizontal axis indicates the number of scenarios included in the group subproblems and the vertical axis indicates the percent deviation from the true optimal objective value for the stochastic program. Observe that all of the $EFGS$ bounds do a much better job than the $EVRS$ bound in terms of estimating the optimal objective value. Specifically, the worst $EFGS$ bound (i.e., $EFGS(1)$), which can be computed in 2.6 seconds, overestimates the optimal value by only 0.04% and $EFGS(2)$ yields the true optimal value in 15 seconds. Computing $EFGS(k)$ for larger group sizes (i.e., $k = 3, \ldots, 9$) also yields the true optimal value but at significantly longer computational times. Although the monotonicity of the upper bounds with respect to increased group size is proven in Theorem 2 for the $EFGS'$ formula, we observe that $EFGS$ also exhibits monotonicity with respect to k.

On the other hand, the worst lower bound obtained via the $EGSO$ formula (i.e., $EGSO(1)$) is better than the widely used WS value and is computed in essentially the same amount of time. We also observe that the lower bounds by the $EGSO(k)$ formula improve with larger group size (k) as proved in Theorem 1, but, of course, at the cost of increased computational time. Clearly, despite both upper and lower bounds improve monotonically as we increase the group size, the computational time to compute these bounds may not always increase with increased group size. In Figure 1, the computational time increases until $k = 8$ and decreases for $k = 9$. This is because of the fact that the bounds for $k = 9$ require optimizing exactly one group subproblem, whereas bounds for $k = 8$ require optimizing nine marginally easier group subproblems. As a result, the total time required for $k = 8$ significantly exceeds that for $k = 9$.

We also compare the $EGSO$ bounds to those obtained from relaxing the integrality restrictions in the problem, which is frequently used as starting points in solving stochastic integer programming applications. All of the relaxed problems are solved using CPLEX 11.0 at its default settings. The lower bounds that are obtained through relaxing integrality restrictions at different stages of the problem perform relatively poorly, where the worst $EGSO$ bound (i.e., $EGSO(1)$) is more than 8 times better than the best relaxation bound

Table 11: Results for the `sizes` problem set

Instance	K	ξ^0	k	WS	$EGSO(k)$	$EFGS(k)$	$EVRS$	EEV
`sizes3`	3	1	1	224067	224227	224276	224571	224890
`sizes3`	3	1	2	224067	224275	224275	224571	224890
`sizes5`	5	1	1	224084	224242	224413	224446	225113
`sizes5`	5	1	2	224084	224305	224344	224446	225113
`sizes5`	5	1	3	224084	224329	224335	224446	225113
`sizes5`	5	1	4	224084	224335	224335	224446	225113

and can be obtained in roughly the same amount of time.

Another observation we make from Figure 1 regarding the behavior of $EGSO$ and $EFGS$ bounds is that the marginal benefit of increasing the group size (k) is decreasing for both bounds. That is, we observe that

$$EGSO(k) - EGSO(k-1) \geq EGSO(k+1) - EGSO(k) \qquad \text{for } k = 2, 3, \ldots, K-1.$$

This observation is also valid for the `bl` instances shown in Table 9 as well as other instances of the `sizes` problem set as shown in Table 11. However, we were not able to establish this observation analytically.

Table 11 displays the results associated with other `sizes` instances. All the quantities in this table can effectively be computed in just a few seconds. All of the observations we have noted until now are also valid with these results, that is, (i) $EGSO$ and $EFGS$, respectively, provide true lower and upper bounds for RP, (ii) both bounds are monotonically improving in increased group size, (iii) both bounds, respectively, outperform the widely used WS and EEV bounds, (iv) $EGSO$ significantly outperforms stochastic linear programming relaxation bounds (the performance of the relaxation bounds for the instances listed in Table 11 are almost identical to those for `sizes10` and, therefore, exact values are omitted from the table to save space), (v) $EFGS$ tend to yield tighter bounds than $EGSO$, and (vi) the marginal benefit of increasing the group size is decreasing for both bounds.

The final set of problems that we test is the `dcap` problem set of Ahmed and Garcia [1]. This problem set consists of stochastic mixed-integer programs arising in dynamic capacity acquisition and allocation applications. The problems have mixed-integer first-stage variables and pure binary second-stage variables.

Table 12 lists the bounds obtained via linear relaxations (LR) for the 12 instances of the `dcap` problem set along with WS, EEV, and $EVRS$. LR^0 indicates relaxing the integrality restrictions in both stages of the problem, whereas LR^2 indicates relaxing integrality only for the second-stage decision variables. The computational time required to obtain these results are negligibly small (typically < 5 seconds). All but two of the `dcap` instances, namely, `dcap233_200` and `dcap243_200`, are not solvable to optimality within the allowed runtime of 12 hours when only the first-stage integer variables are relaxed. These `dcap` instances show how poor the easy-to-solve relaxation bounds (e.g., LR^0 and LR^2) could be: LR^0, on average, underestimates the true optimal value by a striking 58% and similarly for LR^2. The lower bounds obtained by the WS formula are significantly better than the relaxation bounds: WS, on average, underestimates the optimal value by only 3.2%. The upper bounds $EVRS$ and EEV yield unsatisfactory results: $EVRS$ (EEV), on average, overestimates the optimal value by a striking 33% (93%).

The results associated with $EGSO$ and $EFGS$ for the `dcap` problem set are given in Table 13. This table presents the results for all 12 `dcap` instances listed in Table 12. For instance, the $EGSO(1)$ lower bound for the instance labeled `dcap243_500` is 2.61% off of its true optimal objective function value and it took 11.3 seconds to compute this lower bound. Several observations are worth noting. First, the worst $EGSO$ bounds (i.e., $EGSO(1)$) noticably outperform the relaxation bounds shown in Table 12. $EGSO(1)$ bounds, averaged across all 12 instances, underestimate the optimal value by only 3.13% (compare to the 58% relaxation bound) and these can effectively be obtained in less than 15 seconds. Similarly strong upper bounds are obtained by $EFGS(1)$, where the average overestimation across all 12 instances is only 3.33% (compare to 93% for the EEV bound). However, also note that these improved upper bounds come at a significant cost in runtimes. The average

Table 12: Summary of the results for the dcap problem set

Instance	K	LR^0	LR^2	WS	$EGSO$	$EFGS$	$EVRS$	EEV
dcap233_200	200	−52.17%	−51.90%	−2.81%	−1.38%	0.15%	12.58%	39.98%
dcap243_200	200	−37.69%	−37.65%	−2.42%	−1.49%	2.79%	32.89%	14.39%
dcap332_200	200	−76.23%	−76.18%	−7.30%	−3.76%	0.40%	37.04%	119.82%
dcap342_200	200	−57.96%	−57.86%	−2.31%	−0.83%	0.53%	35.39%	45.47%
dcap233_300	300	−55.09%	−54.99%	−2.54%	−1.63%	0.50%	45.65%	182.16%
dcap243_300	300	−38.15%	−38.12%	−1.86%	−1.29%	0.08%	1.61%	146.71%
dcap332_300	300	−76.90%	−76.84%	−6.95%	−4.68%	6.25%	28.93%	168.58%
dcap342_300	300	−60.45%	−60.44%	−2.31%	−1.30%	2.97%	48.57%	213.00%
dcap233_500	500	−54.69%	−54.65%	−2.51%	−1.60%	2.84%	44.16%	59.61%
dcap243_500	500	−39.73%	−39.72%	−2.64%	−2.05%	2.44%	17.45%	16.18%
dcap332_500	500	−82.87%	−82.84%	−2.99%	−1.97%	6.61%	24.65%	80.98%
dcap342_500	500	−60.37%	−60.24%	−2.28%	−1.34%	2.18%	61.46%	27.84%
Average		−57.69%	−57.62%	−3.24%	−1.94%	2.31%	32.53%	92.89%
Standard Deviation		15.11%	15.11%	1.84%	1.13%	2.23%	16.75%	70.27%

Note. The reported $EGSO$ and $EFGS$ bounds are the best bounds available within the allowed runtime of 12 hours. Details are provided in Table 13.

runtime for $EFGS(1)$ is about 18 minutes, whereas the average runtime for $EGSO(1)$ is only about 6 seconds. These bounds display similar performance when group size (k) is 2 or 3. (Indeed, when $k = 2$ or 3, we were not able to solve all the small optimization problems required for computing $EFGS$ within the allotted 12 hours of runtime.) Increasing the group size to 2 improves the averaged lower bound (across all 12 instances) by $EGSO(2)$ to 2.14% but requires an average runtime of about 24 minutes. When $k = 3$, the 12 hours time limit allowed computing $EGSO(3)$ for instances with only 200 scenarios, for which the average bound (across the 4 instances) underestimated the optimal value by 1.87% (compare to 2.47% for $EGSO(2)$ and 3.57% for $EGSO(1)$). Note that the individual bounds in Table 13 for all 12 instances get monotonically better with increased k (recall Theorems 1 and 2). Furthermore, these instances also support our previous conjecture: the marginal benefit of increasing the group size is decreasing in k. Note that we allowed two instances (namely, dcap233_300 and dcap243_300) to run over the 12-hour time limit to compute $EGSO(3)$ and these results also reassure this conjecture. Another observation to note is the effect of increasing the number of scenarios (K) in the problem. As noted earlier and as evidenced from the results in Table 13, increasing K can cause significant increase in runtime requirements (especially for the $EFGS$ formula.)

6 Discussion and future work

We develop a sequence of lower and upper bounds on the optimal objective function value of two-stage SMIPs. We prove that these bounds ensure monotonicity as the number of selected realizations increases. Unlike previous work, which applied to SLPs only, our approach applies to general SMIPs. Furthermore, the inherent structure of computing these bounds naturally leads itself to taking advantage of parallel computing. That is, the individual group subproblems can be optimized independently on separate processors and only the optimal solutions can be passed to a central processor to merge the results.

We tested this approximation procedure on a number of SMIP instances from the literature. Our computational results report that these bounds are superior to the LP relaxation bounds for many problems. Meanwhile, it appears to be not expensive to compute them in many cases. They show that this approximation procedure may be a beneficial addition to the algorithmic development of SMIPs.

A key to excel the applicability of this approximation procedure is the ability to compute these bounds with larger group sizes. We conjecture from our study that both the lower and upper bounds have decreasing marginal benefit in increased group size. We plan to explore more analytical properties of these bounds. For example, under what conditions, a sequence of $EGSO$ ($EFGS$) bounds are guaranteed to have decreasing marginal benefit in the group size? With such result, we can better predict the performance of the bounds a priori. Another example question is that under what conditions, can a sequence of $EFGS$ bounds be computed with just optimizing the group subproblems associated with scenario index set $\Gamma \in \mathcal{P}_k(S)$ (as opposed to $\Gamma \in \mathcal{P}'_k(S)$) and yet guarantee the monotonicity property as in equation (25). With such result, we can substantially save computational time.

The other direction is to statistically estimate $EGSO$ and $EFGS$ bounds for problems with large scenario sets. Instead of exhaustively considering $\mathcal{P}_k(S)$, which includes all scenario subsets of S having cardinality k, we can construct estimates with a sample of scenario subsets in $\mathcal{P}_k(S)$ and develop statistical measures on the deviation of these estimates to the true value of $EGSO$ or $EFGS$.

Finally, extensions of this bounding approach to other stochastic programs such as multistage SMIPs and stochastic nonlinear programs (possibly with integer variables) would be an interesting direction for future research.

Table 13: Detailed results for the `dcap` problem set

K	ξ^0	Bound	Instance				Average
			`dcap233_`K	`dcap243_`K	`dcap332_`K	`dcap342_`K	
200	1	$EGSO(1)$	−2.64% (4.3s)	−2.39% (3.8s)	−7.03% (3.5s)	−2.23% (3.7s)	−3.57% (3.8s)
		$EFGS(1)$	2.07% (5m 49s)	2.79% (7m 21s)	6.10% (5m 11s)	5.10% (5m 56s)	4.02% (6m 4s)
300	1	$EGSO(1)$	−2.45% (5.7s)	−1.74% (5.6s)	−6.72% (5.0s)	−2.26% (5.1s)	−3.29% (5.4s)
		$EFGS(1)$	0.50% (12m 58s)	0.08% (15m 33s)	6.25% (11m 14s)	2.97% (13m 5s)	2.45% (13m 13s)
500	1	$EGSO(1)$	−2.42% (9.8s)	−2.61% (11.3s)	−2.80% (9.0s)	−2.25% (9.8s)	−2.52% (10.0s)
		$EFGS(1)$	2.84% (35m 50s)	2.44% (44m 6s)	6.61% (28m 2s)	2.18% (36m 39s)	3.52% (36m 10s)
200	1	$EGSO(2)$	−1.81% (7m 42s)	−1.81% (7m 51s)	−4.97% (6m 31s)	−1.28% (7m 25s)	−2.47% (7m 22s)
		$EFGS(2)$	0.15% (9h 59m)	NA ($>$ 12h)	0.40% (9h 16m)	0.53% (10h 7m)	NA ($>$ 12h)
300	1	$EGSO(2)$	−1.63% (16m 50s)	−1.29% (18m 36s)	−4.68% (14m 37s)	−1.30% (16m 14s)	−2.22% (16m 34s)
		$EFGS(2)$	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)
500	1	$EGSO(2)$	−1.60% (46m 52s)	−2.05% (55m 56s)	−1.97% (43m 55s)	−1.34% (48m 29s)	−1.74% (48m 48s)
		$EFGS(2)$	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)
200	1	$EGSO(3)$	−1.38% (9h 34m)	−1.49% (10h 33m)	−3.76% (8h 27m)	−0.83% (9h 57m)	−1.87% (9h 37m)
		$EFGS(3)$	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)
300	1	$EGSO(3)$	−1.20% (39h 38m)	−1.07% (49h 38m)	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)
		$EFGS(3)$	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)	NA ($>$ 12h)

`NA`: Results are `not available` within the allowed runtime of 12 hours.

References

[1] Ahmed, S., Garcia, R.: Dynamic capacity acquisition and assignment under uncertainty. Annals of Operations Research **124**, 267–283 (2003)

[2] Birge, J.R.: The value of the stochastic solution in stochastic linear programs with fixed recourse. Mathematical Programming **24**, 314–325 (1982)

[3] Birge, J.R.: Aggregation bounds in stochastic linear programming. Mathematical Programming **31**, 25–41 (1985)

[4] Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer-Verlag, New York (1997)

[5] Edmundson, H.P.: Bounds on the expectation of a convex function of a random variable. Tech. rep., RAND Corporation, Santa Monica, CA (1956)

[6] Frauendorfer, K.: Solving SLP recourse problems with binary multivariate distributions – The dependent case. Mathematics of Operations Research **13**(3), 377–394 (1988)

[7] Huang, C.C., Vertinsky, I., Ziemba, W.T.: Sharp bounds on the value of perfect information. Operations Research **25**(1), 128–139 (1977)

[8] Huang, C.C., Ziemba, W.T., Ben-Tal, A.: Bounds on the expectation of a convex function of a random variable: With applications to stochastic programming. Operations Research **25**(2), 315–325 (1977)

[9] Jensen, J.L.: Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Mathematica **30**(1), 175–193 (1906)

[10] Jorjani, S., Scott, C.H., Woodruff, D.L.: Optimal selection of subsets of sizes for production. International Journal of Production Research **37**(16), 3697–3710 (1999)

[11] Kall, P., Mayer, J.: Stochastic Linear Programming: Models, Theory, and Computation. Springer, Berlin, Germany (2005)

[12] Kall, P., Stoyan, D.: Solving stochastic programming problems with recourse including error bounds. Mathematische Operationsforschung und Statistik, Ser. Optimization **13**, 431–447 (1982)

[13] Madansky, A.: Bounds on the expectation of a convex function of a multivariate random variable. Annals of Mathematical Statistics **30**(3), 743–746 (1959)

[14] Madansky, A.: Inequalities for stochastic linear programming problems. Management Science **6**(2), 197–204 (1960)

[15] Rosa, C.H., Takriti, S.: Improving aggregation bounds for two-stage stochastic programs. Operations Research Letters **24**(3), 127–137 (1999)

[16] Ruszczyński, A. and Shapiro, A. (eds.): Stochastic Programming. Handbooks in Operations Research and Management Science, Vol. 10. Elsevier, Amsterdam, The Netherlands (2003)

[17] Zipkin, P.H.: Bounds for row-aggregation in linear programming. Operations Research **28**(4), 903–916 (1980)

[18] Zipkin, P.H.: Bounds on the effect of aggregating variables in linear programming. Operations Research **28**(2), 403–418 (1980)