A Tighter Variant of Jensen's Lower Bound for Stochastic Programs and Separable Approximations to Recourse Functions

Huseyin Topaloglu

School of Operations Research and Information Engineering,
Cornell University, Ithaca, New York 14853, USA
topaloglu@orie.cornell.edu

October 1, 2008

Abstract

In this paper, we propose a new method to compute lower bounds on the optimal objective value of a stochastic program and show how this method can be used to construct separable approximations to the recourse functions. We show that our method yields tighter lower bounds than Jensen's lower bound and it requires a reasonable amount of computational effort even for large problems. The fundamental idea behind our method is to relax certain constraints by associating dual multipliers with them. This yields a smaller stochastic program that is easier to solve. We particularly focus on the special case where we relax all but one of the constraints. In this case, the recourse functions of the smaller stochastic program are one dimensional functions. We use these one dimensional recourse functions to construct separable approximations to the original recourse functions. Computational experiments indicate that our lower bounds can significantly improve Jensen's lower bound and our recourse function approximations can provide good solutions.

Stochastic programs form a powerful tool to model a variety of situations in which the decisions are made over time without completely knowing the realizations of the future random quantities; see Birge and Louveaux (1997) and Ruszczynski and Shapiro (2003). Despite their wide applicability, however, stochastic programs often pose significant optimization challenges. If the problem involves continuous random variables, then it is quite difficult to come up with tractable solution methods. If the random variables take on finitely many possible realizations, then the problem can be formulated as an equivalent deterministic linear program, but the size of this linear program grows exponentially with the number of time periods and the number of possible realizations. Very often, one has to be content with a small number of possible realizations for the random variables, use Monte Carlo techniques or resort to approximation strategies.

In this paper, we propose a new method to compute lower bounds on the optimal objective value of a stochastic program and demonstrate how this method can be used to construct separable piecewise linear approximations to the recourse functions. We consider stochastic programs that include random variables only on the right side of their constraints. The fundamental idea behind our method is to relax some of these constraints by associating dual multipliers with them. This yields smaller stochastic programs that are hopefully easier to solve. We particularly focus on the special case where we relax all but one of the constraints. In this case, the recourse functions of the smaller stochastic programs are one dimensional functions, roughly capturing the "curvature" of the original recourse functions along different directions. We use these one dimensional recourse functions to construct separable piecewise linear approximations to the original recourse functions.

A traditional method to compute lower bounds on the optimal objective value of a stochastic program is to formulate a deterministic problem by replacing all of the random variables with their expected values and to use Jensen's inequality. We show that the lower bounds obtained by our method improve the ones that are obtained by using Jensen's inequality. Also, our method is quite fast and requires a reasonable amount of computational effort even for large problems. In addition, it allows the user to specify for which random variables to use the full distribution information and for which random variables to use only the first moment information. Finally, it turns out that the decisions made under the guidance of the separable piecewise linear recourse function approximations that are obtained by our method can provide significantly better performance when compared with the common engineering practice of ignoring the uncertainty in the problem and assuming that the random variables take on their expected values.

There has been a long line of research for computing bounds on the optimal objective values of stochastic programs. Jensen's lower bound mentioned above is a common approach and one way to visualize this lower bound is to assume that the probability distributions in the problem are replaced with degenerate distributions that put mass only on the expected values of the random variables; see Kall and Wallace (1994). On the other hand, Edmundson Madansky upper bound is obtained by replacing the probability distributions in the problem with two point distributions that put mass only on the extreme points of the supports of the random variables; see Madansky (1959). An important shortcoming of Edmundson Madansky upper bound is that the computational effort for it grows exponentially as the

number of random variables in the problem increases. Wallace (1987) uses restrictions of stochastic programs with network structure to develop separable piecewise linear upper bounds on the recourse functions, and Birge and Wallace (1988) extend this approach to general stochastic programs. Following very similar ideas, Birge and Wets (1989) use the sublinearity of the recourse functions to develop separable piecewise linear upper bounds on the recourse functions. The computational effort for the last three methods grows linearly as the number of random variables in the problem increases, but the numerical experiments in Birge and Wallace (1988) and Birge and Wets (1989) demonstrate that the upper bounds that are obtained by these methods can be less tight than Edmundson Madansky upper bound. Morton and Wood (1999) construct upper and lower bounds on the recourse functions by relaxing some of the constraints in the problem and penalizing the violations of the relaxed constraints. Their approach has similarities with ours, but if we work with stochastic programs that include random variables only on the right side of their constraints, then the lower bounds obtained by Morton and Wood (1999) collapse to Jensen's lower bound, whereas our lower bounds have the potential to improve Jensen's lower bound. Cheung and Powell (1996) and Godfrey and Powell (2001) exclusively focus on stochastic programs with network structure. Cheung and Powell (1996) use relaxations to develop separable piecewise linear lower bounds on the recourse functions. Godfrey and Powell (2001) propose a Monte Carlo simulation method to construct approximations to the recourse functions, which are neither lower nor upper bounds.

In this paper, we make the following research contributions. 1) We propose a tractable method to compute lower bounds on the optimal objective value of a stochastic program that includes random variables only on the right side of its constraints. 2) We show that the lower bounds that are obtained by our method are tighter than Jensen's lower bound. 3) The fundamental idea behind our method is to relax certain constraints by associating dual multipliers with them. We show the intuitive property that we can obtain progressively tighter lower bounds by relaxing fewer constraints. This shows how we can trade off computational tractability with accuracy. 4) By focusing on the special case where we relax all but one of the constraints, we demonstrate how we can use our method to construct separable piecewise linear approximations to the recourse functions. 5) Computational experiments indicate that the lower bounds that are obtained by our method can significantly improve Jensen's lower bound. We also compare the quality of the solutions that are obtained by our separable piecewise linear recourse function approximations with the quality of the solutions that are obtained under the assumption that the random variables take on their expected values. This is an important comparison that tests the viability of our approach to obtain good solutions. Majority of the literature tests the tightness of the lower bounds but not the quality of the solutions that are obtained by using the lower bounds as recourse function approximations.

The rest of the paper is organized as follows. Section 1 describes a method to compute lower bounds on the optimal objective value of a stochastic program. Section 2 shows that the lower bounds obtained by this method are at least as tight as Jensen's lower bound and our method can obtain progressively tighter lower bounds by relaxing fewer constraints. Section 3 illustrates how our method can be used to construct separable piecewise linear approximations to the recourse functions. Section 4 presents computational experiments.

1   BOUNDING METHOD

In this section, we describe a method to compute lower bounds on the optimal objective value of a stochastic program. All of the results in this section hold for multi-stage stochastic programs, but we illustrate them on two-stage stochastic programs to minimize the notational clutter. Nevertheless, we emphasize that when presenting our computational experiments in Section 4, we go back to working with multi-stage stochastic programs.

We are interested in computing lower bounds on the optimal objective value of the stochastic program given by

$$\zeta_{SP} = \min \quad c\,x + \mathbb{E}\big\{Q(z,\omega)\big\} \tag{1}$$
$$\text{subject to} \quad A\,x \leq b \tag{2}$$
$$T_i\,x - z_i = 0 \qquad i = 1,\ldots,n \tag{3}$$
$$x \geq 0, \tag{4}$$

where c, x and T_i are vectors of compatible dimensions, $z = \{z_i : i = 1,\ldots,n\}$ is an n dimensional vector, $\omega = \{\omega_i : i = 1,\ldots,n\}$ is an n dimensional random variable, and A and b are respectively a matrix and a vector of compatible dimensions. For notational brevity, we do not distinguish between column and row vectors in the paper and the distinction will always be clear from the context. The recourse function is defined by

$$Q(z,\omega) = \min \quad h\,y \tag{5}$$
$$\text{subject to} \quad W_i\,y \leq z_i + \omega_i \qquad i = 1,\ldots,n \tag{6}$$
$$y \geq 0, \tag{7}$$

where h, y and W_i are vectors of compatible dimensions. It is well-known that $Q(z,\omega)$ is a convex function of ω and Jensen's inequality implies that $Q(z,\mathbb{E}\{\omega\}) \leq \mathbb{E}\{Q(z,\omega)\}$. Therefore, the optimal objective value of the problem

$$\zeta_{JB} = \min \quad c\,x + Q(z,\mathbb{E}\{\omega\})$$
$$\text{subject to} \quad (2)\text{-}(4)$$

provides a lower bound on ζ_{SP}. Using the definition of the recourse function in problem (5)-(7), we can write the last problem above as

$$\zeta_{JB} = \min \quad c\,x + h\,y \tag{8}$$
$$\text{subject to} \quad A\,x \leq b \tag{9}$$
$$W_i\,y - T_i\,x \leq \mathbb{E}\{\omega_i\} \qquad i = 1,\ldots,n \tag{10}$$
$$x,y \geq 0. \tag{11}$$

Therefore, the optimal objective value of problem (8)-(11) is a lower bound on ζ_{SP} and this lower bound is known as Jensen's lower bound in the stochastic programming literature.

The method that we propose in this paper is based on using the optimal values of the dual variables associated with constraints (10) to construct a partial linear approximation to the recourse function and completing this partial linear approximation by solving a smaller stochastic program. Specifically, we let $\{\hat{\mu}_i : i = 1, \ldots, n\}$ be the optimal values of the dual variables associated with constraints (10) in problem (8)-(11). For given $\mathcal{I} \subset \{1, \ldots, n\}$, we let $\bar{\mathcal{I}} = \{1, \ldots, n\} \setminus \mathcal{I}$ for notational brevity and relax some of constraints (6) in problem (5)-(7) to define the $|\mathcal{I}|$ dimensional recourse function

$$Q^{\mathcal{I}}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) = \min \quad \left[h - \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i W_i \right] y \tag{12}$$

$$\text{subject to} \quad W_i y \leq z_i + \omega_i \qquad i \in \mathcal{I} \tag{13}$$

$$y \geq 0, \tag{14}$$

where $z^{\mathcal{I}} = \{z_i : i \in \mathcal{I}\}$ denotes the vector composed of the components of z corresponding to $\mathcal{I}$ and $\omega^{\mathcal{I}} = \{\omega_i : i \in \mathcal{I}\}$ denotes the random variable composed of the components of ω corresponding to $\mathcal{I}$. Comparing problems (5)-(7) and (12)-(14), we note that solving the latter problem can be much easier than solving the former one.

The next proposition shows that $Q^{\mathcal{I}}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i z_i + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i \omega_i$ provides a lower bound on $Q(z, \omega)$. Consequently, $\sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i z_i + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i \omega_i$ can be visualized as a partial linear approximation to the recourse function. We complete this approximation by using $Q^{\mathcal{I}}(z^{\mathcal{I}}, \omega^{\mathcal{I}})$, which is obtained by solving a smaller stochastic program that involves an $|\mathcal{I}|$ dimensional recourse function and an $|\mathcal{I}|$ dimensional random variable.

Proposition 1 *We have* $Q^{\mathcal{I}}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i z_i + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i \omega_i \leq Q(z, \omega)$.

Proof Since $\{\hat{\mu}_i : i = 1, \ldots, n\}$ are the optimal values of the dual variables associated with constraints (10), we have $\hat{\mu}_i \leq 0$ for all $i = 1, \ldots, n$. If we let $\hat{y}$ be an optimal solution to problem (5)-(7), then we have $Q(z, \omega) = h \hat{y} \geq h \hat{y} + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i z_i + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i \omega_i - \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i W_i \hat{y} \geq Q^{\mathcal{I}}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i z_i + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i \omega_i$, where the first inequality follows from the fact that $\hat{\mu}_i \leq 0$ and $W_i \hat{y} \leq z_i + \omega_i$ for all $i = 1, \ldots, n$ and the second inequality follows from the fact that $\hat{y}$ is a feasible but not necessarily an optimal solution to problem (12)-(14). $\square$

We propose solving the problem

$$\zeta_{LB}^{\mathcal{I}} = \min \quad \left[c + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i T_i \right] x + \mathbb{E}\{Q^{\mathcal{I}}(z^{\mathcal{I}}, \omega^{\mathcal{I}})\} + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i \mathbb{E}\{\omega_i\} \tag{15}$$

$$\text{subject to} \quad A x \leq b \tag{16}$$

$$T_i x - z_i = 0 \qquad i \in \mathcal{I} \tag{17}$$

$$x \geq 0, \tag{18}$$

to compute a lower bound on ζ_{SP}. Using Proposition 1, it is easy to see that the optimal objective value of the last problem above indeed provides a lower bound on ζ_{SP}. In particular, if we let $(\hat{x}, \hat{z})$ be

an optimal solution to problem (1)-(4) and $\hat{z}^{\mathcal{I}} = \{\hat{z}_i : i \in \mathcal{I}\}$, then we have

$$
\begin{aligned}
\zeta_{SP} = c\,\hat{x} + \mathbb{E}\{Q(\hat{z}, \omega)\} &\geq c\,\hat{x} + \mathbb{E}\{Q^{\mathcal{I}}(\hat{z}^{\mathcal{I}}, \omega^{\mathcal{I}})\} + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, \hat{z}_i + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, \mathbb{E}\{\omega_i\} \\
&= c\,\hat{x} + \mathbb{E}\{Q^{\mathcal{I}}(\hat{z}^{\mathcal{I}}, \omega^{\mathcal{I}})\} + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, \hat{z}_i + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, \mathbb{E}\{\omega_i\} + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, T_i\,\hat{x} - \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, \hat{z}_i \geq \zeta_{LB}^{\mathcal{I}},
\end{aligned}
\tag{19}
$$

where the first inequality follows from Proposition 1, the second equality follows from the fact that $T_i\,\hat{x} - \hat{z}_i = 0$ for all $i = 1, \ldots, n$ and the second inequality follows from the fact that $(\hat{x}, \hat{z}^{\mathcal{I}})$ is a feasible but not necessarily an optimal solution to problem (15)-(18).

2   COMPARISON OF THE BOUNDS

Since $\mathcal{I}$ is an arbitrary subset of $\{1, \ldots, n\}$, we have some freedom when computing the lower bound provided by problem (15)-(18). If we choose $\mathcal{I} = \{1, \ldots, n\}$, then problem (15)-(18) is equivalent to the original stochastic program and we have $\zeta_{LB}^{\{1, \ldots, n\}} = \zeta_{SP}$. If, on the other hand, we choose $\mathcal{I} = \emptyset$, then it is possible to show that $\zeta_{LB}^{\emptyset} = \zeta_{JB}$. The next proposition shows that $\zeta_{LB}^{\mathcal{I}}$ is always between ζ_{JB} and ζ_{SP} for any choice of $\mathcal{I}$. Therefore, the lower bound provided by problem (15)-(18) is always at least as tight as Jensen's lower bound.

Proposition 2 *We have $\zeta_{JB} \leq \zeta_{LB}^{\mathcal{I}} \leq \zeta_{SP}$.*

Proof The second inequality is already shown in (19) and we only show the first inequality here. If we let $\hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}})$ be an optimal solution to problem (12)-(14) as a function of $z^{\mathcal{I}}$ and $\omega^{\mathcal{I}}$, then we have

$$
\begin{aligned}
Q^{\mathcal{I}}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) &= \big[\, h - \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, W_i \,\big]\, \hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) \\
&\geq \big[\, h - \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, W_i \,\big]\, \hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) + \sum_{i \in \mathcal{I}} \hat{\mu}_i\, z_i + \sum_{i \in \mathcal{I}} \hat{\mu}_i\, \omega_i - \sum_{i \in \mathcal{I}} \hat{\mu}_i\, W_i\, \hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) \\
&= h\,\hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) + \sum_{i \in \mathcal{I}} \hat{\mu}_i\, z_i + \sum_{i \in \mathcal{I}} \hat{\mu}_i\, \omega_i - \sum_{i=1}^{n} \hat{\mu}_i\, W_i\, \hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}),
\end{aligned}
\tag{20}
$$

where the inequality follows from the fact that $\hat{\mu}_i \leq 0$ for all $i = 1, \ldots, n$ and $W_i\, \hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) \leq z_i + \omega_i$ for all $i \in \mathcal{I}$. Similarly, if we let $(\hat{x}, \hat{z}^{\mathcal{I}})$ be an optimal solution to problem (15)-(18), then we have

$$
\begin{aligned}
\zeta_{LB}^{\mathcal{I}} &= \big[\, c + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, T_i \,\big]\, \hat{x} + \mathbb{E}\{Q^{\mathcal{I}}(\hat{z}^{\mathcal{I}}, \omega^{\mathcal{I}})\} + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, \mathbb{E}\{\omega_i\} \\
&= \big[\, c + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, T_i \,\big]\, \hat{x} + \mathbb{E}\{Q^{\mathcal{I}}(\hat{z}^{\mathcal{I}}, \omega^{\mathcal{I}})\} + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\, \mathbb{E}\{\omega_i\} + \sum_{i \in \mathcal{I}} \hat{\mu}_i\, T_i\,\hat{x} - \sum_{i \in \mathcal{I}} \hat{\mu}_i\, \hat{z}_i \\
&\geq c\,\hat{x} + \sum_{i=1}^{n} \hat{\mu}_i\, T_i\,\hat{x} + h\, \mathbb{E}\{\hat{y}(\hat{z}^{\mathcal{I}}, \omega^{\mathcal{I}})\} + \sum_{i=1}^{n} \hat{\mu}_i\, \mathbb{E}\{\omega_i\} - \sum_{i=1}^{n} \hat{\mu}_i\, W_i\, \mathbb{E}\{\hat{y}(\hat{z}^{\mathcal{I}}, \omega^{\mathcal{I}})\},
\end{aligned}
\tag{21}
$$

where the second equality follows from the fact that $T_i\,\hat{x} - \hat{z}_i = 0$ for all $i \in \mathcal{I}$ and the inequality follows from (20). On the other hand, by linear programming duality, since $\{\hat{\mu}_i : i = 1, \ldots, n\}$ are the optimal

5

values of the dual variables associated with constraints (10) in problem (8)-(11), the optimal objective value of the problem

$$
\min \quad c\,x + h\,y + \sum_{i=1}^{n} \hat{\mu}_i\,\mathbb{E}\{\omega_i\} - \sum_{i=1}^{n} \hat{\mu}_i\,W_i\,y + \sum_{i=1}^{n} \hat{\mu}_i\,T_i\,x
$$
$$
\text{subject to} \quad A\,x \le b
$$
$$
x, y \ge 0
$$

is equal to ζ_{JB}. The result follows from the inequality in (21) and the fact that $\big(\hat{x}, \mathbb{E}\{\hat{y}(\hat{z}^{\mathcal{I}}, \omega^{\mathcal{I}})\}\big)$ is a feasible solution to the last problem above. $\qquad\square$

If $\mathcal{I}$ includes many elements, then solving problem (15)-(18) becomes computationally more difficult, since this requires dealing with an $|\mathcal{I}|$ dimensional recourse function and an $|\mathcal{I}|$ dimensional random variable. The next proposition illustrates the tradeoff between the computational effort and the tightness of the lower bounds. In particular, it shows that we can tighten the lower bound $\zeta_{LB}^{\mathcal{I}}$ by adding more elements into $\mathcal{I}$.

Proposition 3 *If the sets $\mathcal{K}$ and $\mathcal{I}$ satisfy $\mathcal{K} \subset \mathcal{I}$, then we have $\zeta_{LB}^{\mathcal{K}} \le \zeta_{LB}^{\mathcal{I}}$.*

Proof If we let $\hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}})$ be an optimal solution to problem (12)-(14) as a function of $z^{\mathcal{I}}$ and $\omega^{\mathcal{I}}$, then we have

$$
\begin{aligned}
Q^{\mathcal{I}}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) &= \Big[\, h - \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\,W_i \,\Big]\, \hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) \\
&\ge \Big[\, h - \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\,W_i \,\Big]\, \hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) + \sum_{i \in \mathcal{I} \backslash \mathcal{K}} \hat{\mu}_i\,z_i + \sum_{i \in \mathcal{I} \backslash \mathcal{K}} \hat{\mu}_i\,\omega_i - \sum_{i \in \mathcal{I} \backslash \mathcal{K}} \hat{\mu}_i\,W_i\,\hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) \\
&= h\,\hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) + \sum_{i \in \mathcal{I} \backslash \mathcal{K}} \hat{\mu}_i\,z_i + \sum_{i \in \mathcal{I} \backslash \mathcal{K}} \hat{\mu}_i\,\omega_i - \sum_{i \in \bar{\mathcal{K}}} \hat{\mu}_i\,W_i\,\hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) \\
&\ge Q^{\mathcal{K}}(z^{\mathcal{K}}, \omega^{\mathcal{K}}) + \sum_{i \in \mathcal{I} \backslash \mathcal{K}} \hat{\mu}_i\,z_i + \sum_{i \in \mathcal{I} \backslash \mathcal{K}} \hat{\mu}_i\,\omega_i, \quad\quad (22)
\end{aligned}
$$

where the first inequality follows from the fact that $\hat{\mu}_i \le 0$ for all $i = 1, \ldots, n$ and $W_i\,\hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}}) \le z_i + \omega_i$ for all $i \in \mathcal{I}$, the second equality follows from the fact that $\mathcal{K} \subset \mathcal{I}$ and the second inequality follows from the fact that $\hat{y}(z^{\mathcal{I}}, \omega^{\mathcal{I}})$ is a feasible but not necessarily an optimal solution when we solve problem (12)-(14) after replacing $\mathcal{I}$ with $\mathcal{K}$. On the other hand, if we let $(\hat{x}, \hat{z}^{\mathcal{I}})$ be an optimal solution to problem (15)-(18), then we have

$$
\begin{aligned}
\zeta_{LB}^{\mathcal{I}} &= \Big[\, c + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\,T_i \,\Big]\, \hat{x} + \mathbb{E}\{Q^{\mathcal{I}}(\hat{z}^{\mathcal{I}}, \omega^{\mathcal{I}})\} + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\,\mathbb{E}\{\omega_i\} \\
&= \Big[\, c + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\,T_i \,\Big]\, \hat{x} + \mathbb{E}\{Q^{\mathcal{I}}(\hat{z}^{\mathcal{I}}, \omega^{\mathcal{I}})\} + \sum_{i \in \bar{\mathcal{I}}} \hat{\mu}_i\,\mathbb{E}\{\omega_i\} + \sum_{i \in \mathcal{I} \backslash \mathcal{K}} \hat{\mu}_i\,T_i\,\hat{x} - \sum_{i \in \mathcal{I} \backslash \mathcal{K}} \hat{\mu}_i\,\hat{z}_i \\
&\ge \Big[\, c + \sum_{i \in \bar{\mathcal{K}}} \hat{\mu}_i\,T_i \,\Big]\, \hat{x} + \mathbb{E}\{Q^{\mathcal{K}}(\hat{z}^{\mathcal{K}}, \omega^{\mathcal{K}})\} + \sum_{i \in \bar{\mathcal{K}}} \hat{\mu}_i\,\mathbb{E}\{\omega_i\} \ge \zeta_{LB}^{\mathcal{K}},
\end{aligned}
$$

where the second equality follows from the fact that $T_i\,\hat{x} - \hat{z}_i = 0$ for all $i \in \mathcal{I}$, the first inequality follows from (22) and the fact that $\mathcal{K} \subset \mathcal{I}$ and the last inequality follows from the fact that $(\hat{x}, \hat{z}^{\mathcal{K}})$ is a feasible but not necessarily an optimal solution when we solve problem (15)-(18) after replacing $\mathcal{I}$ with $\mathcal{K}$. $\qquad\square$

3   A SEPARABLE RECOURSE FUNCTION APPROXIMATION

In this section, we focus on the lower bound $\zeta_{LB}^{\mathcal{I}}$ under the assumption that $\mathcal{I}$ is a singleton and show how we can use problem (15)-(18) to construct a separable approximation to the recourse function.

If $\mathcal{I}$ is a singleton, then we only need to deal with a one dimensional recourse function and a one dimensional random variable in problem (15)-(18), and this makes problem (15)-(18) relatively easy to solve. Indeed, in many practical settings, it may be possible to solve problem (15)-(18) only when $\mathcal{I}$ is a singleton. Furthermore, if we have $\mathcal{I} = \{i\}$ for some $i \in \{1, \ldots, n\}$, then $\mathbb{E}\{Q^{\{i\}}(z^{\{i\}}, \omega^{\{i\}})\}$ is a one dimensional piecewise linear function and it is straightforward to store one dimensional piecewise linear functions in a computer environment. Finally, if we compute $\zeta_{LB}^{\{i\}}$ for all $i \in \{1, \ldots, n\}$, then we have $\zeta_{LB}^{\{i\}} \leq \zeta_{SP}$ for all $i \in \{1, \ldots, n\}$ and we can use $\max_{i \in \{1, \ldots, n\}} \zeta_{LB}^{\{i\}}$ as a lower bound on ζ_{SP}. This is the tightest lower bound that we can obtain by using singletons.

The method that we use to construct a separable approximation to the recourse function is based on the observation that we can use Proposition 1 to visualize the lower bound $Q^{\{i\}}(z^{\{i\}}, \omega^{\{i\}}) + \sum_{j \in \{\bar{i}\}} \hat{\mu}_j z_j + \sum_{j \in \{\bar{i}\}} \hat{\mu}_j \omega_j$ as an approximation to $Q(z, \omega)$. Roughly speaking, the piecewise linear component $Q^{\{i\}}(z^{\{i\}}, \omega^{\{i\}})$ in this approximation captures the "curvature" of $Q(z, \omega)$ in the direction of the i-th component of z and ω. On the other hand, the linear component $\hat{\mu}_j z_j + \hat{\mu}_j \omega_j$ captures the slope of $Q(z, \omega)$ in the direction of the j-th component of z and ω for $j \in \{\bar{i}\}$. Since the latter component is linear in z and ω, it does not capture "curvature" information.

To be able to capture "curvature" information in all directions, we propose computing the lower bound $Q^{\{i\}}(z^{\{i\}}, \omega^{\{i\}}) + \sum_{j \in \{\bar{i}\}} \hat{\mu}_j z_j + \sum_{j \in \{\bar{i}\}} \hat{\mu}_j \omega_j$ for all $i \in \{1, \ldots, n\}$. In this case, we can average over all $i \in \{1, \ldots, n\}$ and use

$$\frac{1}{n} \sum_{i=1}^{n} Q^{\{i\}}(z^{\{i\}}, \omega^{\{i\}}) + \frac{1}{n} \sum_{i=1}^{n} \sum_{j \in \{\bar{i}\}} \hat{\mu}_j z_j + \frac{1}{n} \sum_{i=1}^{n} \sum_{j \in \{\bar{i}\}} \hat{\mu}_j \omega_j \tag{23}$$

as a separable piecewise linear approximation to $Q(z, \omega)$. Noting that $Q^{\{i\}}(z^{\{i\}}, \omega^{\{i\}}) + \sum_{j \in \{\bar{i}\}} \hat{\mu}_j z_j + \sum_{j \in \{\bar{i}\}} \hat{\mu}_j \omega_j$ is a lower bound on $Q(z, \omega)$ for all $i = 1, \ldots, n$, it is clear that the approximation in (23) is also a lower bound on $Q(z, \omega)$. In this case, we can replace $Q(z, \omega)$ in problem (1)-(4) with the approximation in (23) and solve the problem

$$\min \quad c\,x + \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\{Q^{\{i\}}(z^{\{i\}}, \omega^{\{i\}})\} + \frac{1}{n} \sum_{i=1}^{n} \sum_{j \in \{\bar{i}\}} \hat{\mu}_j z_j + \frac{1}{n} \sum_{i=1}^{n} \sum_{j \in \{\bar{i}\}} \hat{\mu}_j \mathbb{E}\{\omega_j\} \tag{24}$$

$$\text{subject to} \quad (2)\text{-}(4) \tag{25}$$

to obtain an approximate solution to problem (1)-(4). We do not have a performance guarantee for the solution that is obtained by solving problem (24)-(25), but since the approximation in (23) is a lower bound on the recourse function, the optimal objective value of problem (24)-(25) is also a lower bound on the optimal objective value of problem (1)-(4). We also note that the solution obtained by solving problem (24)-(25) is feasible to the original problem, since we impose constraints (2)-(4) in problem (24)-(25). Finally, we observe that the term $\frac{1}{n} \sum_{i=1}^{n} \sum_{j \in \{\bar{i}\}} \hat{\mu}_j \mathbb{E}\{\omega_j\}$ in the objective function is a constant and it can be dropped without changing the optimal solution to problem (24)-(25).

4 Computational Experiments

In this section, we numerically compare the lower bounds obtained by our method with Jensen's lower bounds and test the quality of the solutions obtained by solving problem (24)-(25). We work with four problem classes. The first problem class is a multi-stage stochastic program that takes place over a time-staged network with random node supplies. The second problem class is a multi-stage stochastic program that models the airline network revenue management problem. The third problem class is a two-stage stochastic program that arises in capacity expansion planning. The fourth problem class is a two-stage stochastic program that models an assemble to order production line.

For the four problem classes, our computational results indicate that the lower bounds obtained by our method can noticeably improve Jensen's lower bounds and the solutions obtained by solving problem (24)-(25) can perform quite well. This is especially the case when the underlying random variables have high variability. Nevertheless, we emphasize that it is possible to construct small pathological problem instances where the gaps among ζ_{SP}, $\zeta_{LP}^{\mathcal{I}}$ and ζ_{JB} are arbitrarily large. Therefore, the reader should be cautious about deriving broad generalizations from our computational results and realize that it is important to test our method for the particular problem class that he or she is interested in.

4.1 Time-Staged Network with Random Node Supplies

In this problem, we are interested in generating profits by relocating boxcars along the arcs of the directed graph $(\mathcal{N}, \mathcal{A})$. The problem takes place over the planning horizon $\{1, \ldots, \tau\}$. A random number of boxcars become available at each node at the beginning of each time period. We use ω_{it} to denote the random boxcar supply at node i at time period t. We assume that the supply random variables at different time periods are independent and the boxcars available at a particular node remain at this node until they are shipped elsewhere. We use r_{ijt} and u_{ijt} to respectively denote the profit and upper bound associated with arc (i, j) at time period t. The objective is to maximize the total expected profit over the planning horizon, but for consistency with the earlier development, we formulate the problem to minimize the total expected cost. Letting z_{it} be the number of boxcars available at node i at time period t due to the earlier decisions and x_{ijt} be the number of cars that move from node i to node j at time period t, the problem can be formulated as a multi-stage stochastic program as

$$Q_t(z_t, \omega_t) = \min \quad \sum_{(i,j) \in \mathcal{A}} -r_{ijt} \, x_{ijt} + \mathbb{E}\big\{ Q_{t+1}(z_{t+1}, \omega_{t+1}) \big\} \tag{26}$$

$$\text{subject to} \quad \sum_{j:(i,j) \in \mathcal{A}} x_{ijt} = z_{it} + \omega_{it} \qquad\qquad i \in \mathcal{N} \tag{27}$$

$$\sum_{i:(i,j) \in \mathcal{A}} x_{ijt} - z_{j,t+1} = 0 \qquad\qquad j \in \mathcal{N} \tag{28}$$

$$0 \leq x_{ijt} \leq u_{ijt} \qquad\qquad (i, j) \in \mathcal{A} \tag{29}$$

$$z_{i,t+1} \geq 0 \qquad\qquad i \in \mathcal{N}, \tag{30}$$

with $Q_{\tau+1}(\cdot, \cdot) = 0$. We assume that $z_{i1} = 0$ for all $i \in \mathcal{N}$ and let $\zeta_{SP} = \mathbb{E}\big\{ Q_1(z_1, \omega_1) \big\}$, which is the optimal total expected cost over the planning horizon. We can compute Jensen's lower bound

by replacing $Q_t(z_t, \omega_t)$ and $\mathbb{E}\{Q_{t+1}(z_{t+1}, \omega_{t+1})\}$ in the objective function and ω_{it} in constraints (27) in the problem above respectively with $Q_t(z_t, \mathbb{E}\{\omega_t\})$, $Q_{t+1}(z_{t+1}, \mathbb{E}\{\omega_{t+1}\})$ and $\mathbb{E}\{\omega_{it}\}$. This simply corresponds to replacing all random variables with their expected values and solving a deterministic problem. We let $\{\hat{\mu}_{it} : i \in \mathcal{N}, \ t = 1, \ldots, \tau\}$ be the optimal values of the dual variables associated with constraints (27) when we compute Jensen's lower bound.

Problem (26)-(30) is a multi-stage stochastic program, but as mentioned at the beginning of Section 1, it is possible to follow the same ideas in the paper to compute a lower bound on ζ_{SP} by relaxing some of constraints (27). We denote this lower bound by $\zeta_{LB}^{\mathcal{I}}$, where $\mathcal{I}$ is the subset of constraints (27) that we choose to leave in problem (26)-(30). For computational tractability, when computing $\zeta_{LB}^{\mathcal{I}}$ in this and the next three sections, we only concentrate on the case where $\mathcal{I}$ is a singleton. To have a feel for the quality of the solutions that are obtained by the separable approximation strategy described in Section 3, we replace $Q_{t+1}(z_{t+1}, \omega_{t+1})$ in the objective function of problem (26)-(30) with $\frac{1}{|\mathcal{N}|} \sum_{i \in \mathcal{N}} Q_{t+1}^{\{i\}}(z_{t+1}^{\{i\}}, \omega_{t+1}^{\{i\}}) + \frac{1}{|\mathcal{N}|} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N} \setminus \{i\}} \hat{\mu}_{j,t+1} z_{j,t+1}$ and solve this problem to make the decisions at time period t.

Table 1 shows the computational results. Our test problems take place over 10 time periods and we label them by $(|\mathcal{N}|, v)$, where $|\mathcal{N}|$ is the number of nodes and v is the average coefficient of variation for the supply random variables. The second and third columns in Table 1 show ζ_{JB} and $\max_{i \in \mathcal{N}} \zeta_{LB}^{\{i\}}$, which are respectively Jensen's lower bound and the tightest lower bound that can be obtained by our method by using singletons. The fourth column shows the percent gap between the second and third columns. The fifth column shows the CPU seconds required to compute the lower bound $\max_{i \in \mathcal{N}} \zeta_{LB}^{\{i\}}$ on a Pentium IV desktop PC with 2.6 GHz CPU and 1 GB RAM. The sixth column shows the total expected cost incurred by the strategy that replaces $\mathbb{E}\{Q_{t+1}(z_{t+1}, \omega_{t+1})\}$ in the objective function of problem (26)-(30) with $Q_{t+1}(z_{t+1}, \mathbb{E}\{\omega_{t+1}\})$ and solves this problem to make the decisions at time period t. This strategy corresponds to the common engineering practice of ignoring the uncertainty in the problem and assuming that the random variables take on their expected values. We note that the total expected cost in the sixth column is different from Jensen's lower bound. To see this, we emphasize that ignoring the uncertainty in the problem and assuming that the random variables take on their expected values yields a suboptimal policy and the total expected cost incurred by doing so is naturally greater than or equal to the optimal total expected cost. On the other hand, Jensen's lower bound is less than or equal to the optimal total expected cost. The seventh column shows the total expected cost incurred by the strategy that replaces $Q_{t+1}(z_{t+1}, \omega_{t+1})$ in the objective function of problem (26)-(30) with $\frac{1}{|\mathcal{N}|} \sum_{i \in \mathcal{N}} Q_{t+1}^{\{i\}}(z_{t+1}^{\{i\}}, \omega_{t+1}^{\{i\}}) + \frac{1}{|\mathcal{N}|} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N} \setminus \{i\}} \hat{\mu}_{j,t+1} z_{j,t+1}$. This is the separable approximation strategy described in Section 3 and in the paragraph above. Since it is impossible to compute the total expected costs in the sixth and seventh columns explicitly, we estimate them through simulation. The last column shows the percent gap between the sixth and seventh columns. We emphasize that even the smallest test problem in Table 1 is too large to be solved as a multi-stage stochastic program through an equivalent deterministic linear programming formulation.

The profit parameters, upper bounds and expected values of the supply random variables in the first and last five consecutive test problems in Table 1 are the same. Therefore, Jensen's lower bounds

Problem			ζ_{JB}		Determ.	Sep.	Determ.
$(\lvert\mathcal{N}\rvert, v)$	ζ_{JB}	ζ_{LB}	vs. ζ_{LB}	CPU	app.	app.	vs. sep.
$(2, 0.7)$	-213,124	-208,120	2.3	0.5	-201,450	-203,982	1.3
$(2, 1.0)$	-213,124	-204,523	4.0	0.1	-192,126	-197,592	2.8
$(2, 1.4)$	-213,124	-202,999	4.8	0.4	-186,288	-192,761	3.5
$(2, 3.0)$	-213,124	-198,011	7.1	0.6	-143,704	-154,676	7.6
$(2, 4.4)$	-213,124	-191,905	10.0	0.2	-116,268	-129,923	11.7
$(4, 0.7)$	-964,696	-962,257	0.3	1.5	-955,452	-958,532	0.3
$(4, 2.0)$	-964,696	-956,409	0.9	2.7	-909,207	-919,672	1.2
$(4, 3.0)$	-964,696	-953,372	1.2	3.1	-819,629	-829,495	1.2
$(4, 4.4)$	-964,696	-951,526	1.4	1.0	-727,496	-739,876	1.7
$(4, 5.4)$	-964,696	-949,531	1.6	2.2	-666,863	-679,476	1.9

Table 1: Computational results for time-staged network with random node supplies.

for these test problems are also the same. As the average coefficient of variation for the supply random variables increases, the gap between Jensen's lower bound and the lower bound obtained by our method gets larger. The lower bound obtained by our method can be up to 10.0% tighter than Jensen's lower bound. Similarly, if the variability in the supply random variables is small, then the performances of the deterministic and separable approximation strategies are close to each other. We use common random numbers when simulating the performances of the two strategies and the performance gap between the two strategies is statistically significant at 95% level for all of the test problems. As expected, as the average coefficient of variation for the supply random variables increases, the performance of the deterministic approximation strategy deteriorates and the performance gap between the deterministic and separable approximation strategies gets larger.

4.2 Airline Network Revenue Management

This problem is the traditional airline network revenue management problem; see Talluri and van Ryzin (2004). We have a set of flight legs that can be used to satisfy the itinerary requests that arrive randomly over time. Accepted itinerary requests generate revenues and consume capacities on the relevant flight legs, whereas rejected itinerary requests simply leave the system. The problem takes place over the planning horizon $\{1, \ldots, \tau\}$. The set of flight legs is $\mathcal{L}$ and the set of itineraries is $\mathcal{J}$. If we accept a request for itinerary j, then we generate a revenue of r_j and consume a_{ij} units of capacity on flight leg i. We use ω_{jt} to denote the random amount of demand for itinerary j at time period t, and assume that the demand random variables at different time periods are independent. We are interested in maximizing the total expected revenue over the planning horizon. Letting z_{it} be the remaining capacity on flight leg i at time period t and x_{jt} be the amount of demand for itinerary j that we satisfy at time period t, the problem can be formulated as a multi-stage stochastic program as

$$Q_t(z_t, \omega_t) = \min \quad \sum_{j \in \mathcal{J}} -r_j\, x_{jt} + \mathbb{E}\big\{Q_{t+1}(z_{t+1}, \omega_{t+1})\big\}$$

$$\text{subject to} \quad \sum_{j \in \mathcal{J}} a_{ij}\, x_{jt} + z_{i,t+1} = z_{it} \qquad i \in \mathcal{L}$$

$$0 \le x_{jt} \le \omega_{jt} \qquad j \in \mathcal{J}$$

$$z_{i,t+1} \ge 0 \qquad i \in \mathcal{L},$$

Problem $(|\mathcal{L}|, |\mathcal{J}|, \alpha, \kappa)$	ζ_{JB}	ζ_{LB}	ζ_{JB} vs. ζ_{LB}	CPU	Determ. app.	Sep. app.	Determ. vs. sep.
$(4, 12, 1.0, 4)$	-11,114	-10,951	1.5	0.3	-10,537	-10,654	1.1
$(4, 12, 1.0, 8)$	-17,558	-17,367	1.1	0.3	-16,598	-16,955	2.1
$(4, 12, 1.2, 4)$	-10,163	-10,029	1.3	0.2	-9,475	-9,729	2.7
$(4, 12, 1.2, 8)$	-16,607	-16,440	1.0	0.2	-15,194	-15,976	5.1
$(4, 12, 1.6, 4)$	-8,827	-8,698	1.5	0.2	-8,033	-8,381	4.3
$(4, 12, 1.6, 8)$	-15,271	-15,103	1.1	0.2	-13,524	-14,587	7.9
$(8, 40, 1.0, 4)$	-21,531	-21,289	1.1	1.2	-19,780	-19,970	1.0
$(8, 40, 1.0, 8)$	-34,571	-34,281	0.8	1.2	-31,632	-32,265	2.0
$(8, 40, 1.2, 4)$	-19,882	-19,670	1.1	1.0	-17,873	-18,357	2.7
$(8, 40, 1.2, 8)$	-32,922	-32,656	0.8	1.0	-29,022	-30,536	5.2
$(8, 40, 1.6, 4)$	-17,530	-17,325	1.2	0.8	-15,345	-15,945	3.9
$(8, 40, 1.6, 8)$	-30,570	-30,306	0.9	0.8	-26,086	-27,968	7.2

Table 2: Computational results for airline network revenue management.

with $Q_{\tau+1}(\cdot, \cdot) = 0$. The initial capacities on the flight legs are given by $\{z_{i1} : i \in \mathcal{L}\}$ and they are known quantities.

Table 2 shows the computational results in the same format as Table 1. All of our test problems take place over 200 time periods. We label our test problems by $(|\mathcal{L}|, |\mathcal{J}|, \alpha, \kappa)$, where $|\mathcal{L}|$ is the number of flight legs, $|\mathcal{J}|$ is the number of itineraries, α is the ratio of the total expected demand to the total initial capacity and κ is the ratio of the revenues associated with the most and least expensive itineraries.

For a majority of the test problems, the lower bound obtained by our method is more than 1% tighter than Jensen's lower bound. This is a significant gap for the network revenue management setting. More importantly, especially for the test problems with tight leg capacities and large revenue differences between the most and least expensive itineraries, the performance gap between the deterministic and separable approximation strategies can be as high as 7.9%. We note that the regret associated with making an "incorrect" decision increases when the leg capacities are tight and there is a large revenue difference between the most and least expensive itineraries. In particular, if the leg capacities are tight and we accept an itinerary request that we should not have accepted, then this "incorrect" decision limits our ability to accept a more desirable itinerary request in the future. Similarly, if the revenue difference between the most and least expensive itineraries is large, then if we accept a request for a cheap itinerary that we should not have accepted, then this "incorrect" decision limits our ability to accept a future request that may bring substantially higher revenue. Thus, it is fundamental to observe the tradeoff between when to accept the requests for cheap itineraries so that the flight legs do not take off empty and when to reserve the seats for the future requests for expensive itineraries so that we make the most revenue from the seats. These tradeoffs become more obvious when the leg capacities are tight and there is a large revenue difference between the most and least expensive itineraries. It is encouraging that separable approximation strategy performs significantly better in these cases.

4.3 CAPACITY EXPANSION PLANNING

We take this problem from Higle and Sen (1994). We are interested in purchasing production capacity on several machines to satisfy the demand for parts. The set of machines is $\mathcal{M}$ and the set of parts

| Problem | | | ζ_{JB} | | Determ. | Sep. | Determ. | |
v	ζ_{JB}	ζ_{LB}	vs. ζ_{LB}	CPU	app.	app.	vs. sep.	ζ_{SP}
0.68	90,200	150,974	67.4	0.0	366,898	373,281	-1.7	355,159
0.82	90,200	181,193	100.9	0.0	442,065	442,065	0.0	422,948
0.94	90,200	221,264	145.3	0.0	504,679	491,600	2.6	482,855
1.00	90,200	237,300	163.1	0.0	537,794	514,258	4.4	514,252

Table 3: Computational results for capacity expansion planning.

is $\mathcal{P}$. The current capacity on machine j is h_j hours and it costs c_j to purchase one hour of capacity. Machine j cannot be used for more than u_j hours and it requires t_j hours of maintenance for every hour of usage. The total maintenance time for all of the machines cannot exceed T hours. Part i can be produced on machine j with a production rate of a_{ij} units per hour and at a cost of q_{ij} per hour. The random demand for part i is ω_i. All demand has to be satisfied and it is possible to purchase part i from a subcontractor at a premium price of p_i per unit. Letting x_j and z_j respectively be the hours of production capacity purchased and used on machine j, y_{ij} be the hours of production capacity used on machine j to produce part i and s_i be the units of part i purchased from the subcontractor, the problem can be formulated as a two-stage stochastic program as

$$
\begin{aligned}
\min \quad & \sum_{j\in\mathcal{M}} c_j\, x_j + \mathbb{E}\big\{Q(z,\omega)\big\} \\
\text{subject to} \quad & -x_j + z_j \le h_j && j \in \mathcal{M} \\
& \sum_{j\in\mathcal{M}} t_j\, z_j \le T \\
& 0 \le z_j \le u_j && j \in \mathcal{M} \\
& x_j \ge 0 && j \in \mathcal{M},
\end{aligned}
$$

where the recourse function is given by

$$
\begin{aligned}
Q(z,\omega) = \min \quad & \sum_{i\in\mathcal{P}}\sum_{j\in\mathcal{M}} q_{ij}\, y_{ij} + \sum_{i\in\mathcal{P}} p_i\, s_i \\
\text{subject to} \quad & \sum_{j\in\mathcal{M}} a_{ij}\, y_{ij} + s_i \ge \omega_i && i \in \mathcal{P} \\
& \sum_{i\in\mathcal{P}} y_{ij} \le z_j && j \in \mathcal{M} \\
& y_{ij}, s_i \ge 0 && i \in \mathcal{P},\ j \in \mathcal{M}.
\end{aligned}
$$

Table 3 shows the computational results in the same format as Table 1. We label our test problems by v, where v is the average coefficient of variation for the demand random variables. The first test problem in Table 3 corresponds to the original problem parameters used by Higle and Sen (1994). Since the test problems are not too large, the last column in Table 3 shows the optimal total expected cost, which we obtain by solving the stochastic program above through an equivalent deterministic linear programming formulation.

The cost parameters, maintenance and production times and expected values of the demand random variables in all of our test problems are the same. Therefore, Jensen's lower bounds for all of our

| Problem | | | ζ_{JB} | | Determ. | Sep. | Determ. | |
v	ζ_{JB}	ζ_{LB}	vs. ζ_{LB}	CPU	app.	app.	vs. sep.	ζ_{SP}
0.21	-8,300	-7,955	4.2	0.0	-7,062	-7,366	4.3	-7437
0.33	-8,300	-7,650	7.8	0.0	-6,008	-6,723	11.9	-6797
0.49	-8,300	-7,025	15.4	0.0	-4,298	-5,179	20.5	-5521
0.71	-8,300	-6,563	20.9	0.0	-1,644	-3,752	128.2	-3979

Table 4: Computational results for assemble to order production line.

test problems are also the same. The capacity expansion problem is well-known in the literature for its loose Jensen's lower bounds and the lower bounds obtained by our method substantially improve Jensen's lower bounds. When the average coefficient of variation for the demand random variables is small, the performance of the deterministic approximation strategy is better than that of the separable approximation strategy. As expected, as the variability in the demand random variables increases, the separable approximation strategy tends to perform better. We note that the performance of the separable approximation strategy is very close to optimal for the last test problem.

4.4 ASSEMBLE TO ORDER PRODUCTION LINE

In this problem, we are interested in serving the demand for different products by assembling the products from different components. The set of components is $\mathcal{C}$ and the set of products is $\mathcal{P}$. The cost of purchasing one unit of component i is c_i and the revenue from one unit of product j is r_j. Product j is assembled from a_{ij} units of component i. The random demand for product j is ω_j. Letting z_i be the units of component i purchased and y_j be the units of product j assembled, the problem can be formulated as a two-stage stochastic program as

$$\min \quad \sum_{i \in \mathcal{C}} c_i \, z_i + \mathbb{E}\big\{Q(z, \omega)\big\}$$
$$\text{subject to} \quad z_i \geq 0 \qquad i \in \mathcal{C},$$

where the recourse function is given by

$$Q(z, \omega) = \min \quad \sum_{j \in \mathcal{P}} -r_j \, y_j$$
$$\text{subject to} \quad \sum_{j \in \mathcal{P}} a_{ij} \, y_j \leq z_i \qquad i \in \mathcal{C}$$
$$0 \leq y_j \leq \omega_j \qquad j \in \mathcal{P}.$$

Table 4 shows the computational results in the same format as Table 1. All of our test problems involve three components and seven products. We label our test problems by v, where v is the average coefficient of variation for the demand random variables. Similar to Table 3, since the test problems are not too large, the last column in Table 4 shows the optimal total expected cost and we obtain the optimal total expected cost by solving the stochastic program above through an equivalent deterministic linear programming formulation.

13

The cost and revenue parameters and the expected values of the demand random variables in all of our test problems are the same. Therefore, Jensen's lower bounds for all of our test problems are also the same. The results indicate that Jensen's lower bound can be quite loose when the average coefficient of variation is high. As the average coefficient of variation increases, the solutions that are obtained by the separable approximation strategy can perform substantially better than the solutions that are obtained under the assumption that the demand random variables take on their expected values. For example, the performance gap between the two solutions can be as high as 128.2% when the average coefficient of variation is 0.71. For many of the test problems, the performance of the solutions that are obtained by the separable approximation strategy is quite close to the optimal total expected cost, but the performance of the solutions that are obtained under the assumption that the demand random variables take on their expected values can be quite poor.

5   CONCLUSIONS

In this paper, we proposed a tractable method to compute lower bounds on the optimal objective value of a stochastic program. We showed that the lower bounds obtained by our method improve Jensen's lower bounds. We also showed how a special case of our method can be used to construct separable piecewise linear approximations to the recourse functions. A unifying observation from all of our computational experiments is that when the underlying random variables have high coefficient of variation, the performance of the solutions that are obtained by our separable approximation strategy is significantly better than the performance of the solutions that are obtained under the assumption that the random variables take on their expected values. This can be observed by noting the gaps in the eight column in Tables 1, 3 and 4. Therefore, our separable approximation strategy can provide significant advantages when the randomness in the problem plays a substantial role. Similarly, for the airline network revenue management test problems, our separable approximation strategy performs better than the deterministic approximation strategy when the leg capacities are tight and the revenue difference between the most and least expensive itineraries is large. As discussed at the end of Section 4.2, these test problems capture the case where the regret associated with making an "incorrect" decision is high and one has to address the stochastic and dynamic elements of the problem more carefully. As a result, our separable approximation strategy appears to provide noticeable improvements when dealing with problems where uncertain and dynamic elements play a crucial role.

There are two possible directions for future research. First, the quality of the solutions obtained by the separable approximation strategy should be investigated within the context of different problem classes. The results for the time-staged network with random node supplies and airline network revenue management settings are quite encouraging. Our separable approximation strategy has the potential to make an impact in finding good solutions for practical stochastic programs. Second, our computational experiments focus on the case where $\mathcal{I}$ is a singleton, but as Proposition 3 shows, it is possible to tighten the lower bounds by adding more elements into $\mathcal{I}$. If $\mathcal{I}$ is to include k elements, then there are $\binom{n}{k}$ possible choices. In this case, a complete enumeration of $\binom{n}{k}$ possible choices to obtain the tightest lower bound can be prohibitive and a method to predict which subset of $\{1, \ldots, n\}$ yields the tightest lower bound can be useful.

REFERENCES

Birge, J. R. and Louveaux, F. (1997), *Introduction to Stochastic Programming*, Springer-Verlag, New York.

Birge, J. R. and Wallace, S. W. (1988), 'A separable piecewise linear upper bound for stochastic linear programs', *SIAM Journal of Control and Optimization* **26**(3), 725–739.

Birge, J. and Wets, R. (1989), 'Sublinear upper bounds for stochastic programs with recourse', *Mathematical Programming* **43**, 131–149.

Cheung, R. K. and Powell, W. B. (1996), 'An algorithm for multistage dynamic networks with random arc capacities, with an application to dynamic fleet management', *Operations Research* **44**(6), 951–963.

Godfrey, G. A. and Powell, W. B. (2001), 'An adaptive, distribution-free approximation for the newsvendor problem with censored demands, with applications to inventory and distribution problems', *Management Science* **47**(8), 1101–1112.

Higle, J. L. and Sen, S. (1994), 'Finite master programs in regularized stochastic decomposition', *Mathematical Programming* **67**, 143–168.

Kall, P. and Wallace, S. W. (1994), *Stochastic Programming*, John Wiley and Sons, New York.

Madansky, A. (1959), 'Bounds on the expectation of a convex function of a multivariate random variable', *Annals Math. Stat.* **30**, 743–746.

Morton, D. and Wood, K. (1999), 'Restricted-recourse bounds for stochastic linear programming', *Operations research* **47**, 943–956.

Ruszczynski, A. and Shapiro, A., eds (2003), *Handbooks in Operations Research and Management Science, Volume on Stochastic Programming*, North Holland, Amsterdam.

Talluri, K. T. and van Ryzin, G. J. (2004), *The Theory and Practice of Revenue Management*, Kluver Academic Publishers.

Wallace, S. W. (1987), 'A piecewise linear upper bound on the network recourse function', *Mathematical Programming* **38**, 133–146.