MATHEMATICS OF OPERATIONS RESEARCH
Vol. 19, No. 2, May 1994
Printed in U.S.A.

BOUNDS FOR TWO-STAGE STOCHASTIC PROGRAMS WITH FIXED RECOURSE

N. C. P. EDIRISINGHE AND W. T. ZIEMBA

This paper develops upper and lower bounds on two-stage stochastic linear programs using limited moment information. The case considered is when both the right-hand side as well as the objective coefficients of the second stage problem are random. Random variables are allowed to have arbitrary multivariate probability distributions with bounded support. First, upper and lower bounds are obtained using first and cross moments, from which we develop bounds using only first moments. The bounds are shown to solve the respective general moment problems.

1. Introduction. A stochastic program arises when some or all of the problem parameters of a deterministic mathematical program are subject to random variation. These problems often involve optimizing an expectation functional of multivariate random variables which may be stochastically dependent. Multidimensional numerical integration as a computational strategy is prohibitively expensive. This is not only because it is at the outer limits of computer technology, but also because it may become necessary to evaluate the expectation functional as many times as would be necessary in an iterative search of an optimal set of decisions for the problem at hand. Inevitably, therefore, one is forced to look for approximations, such as upper and lower bounds to the stochastic program. A possible solution strategy is to compute bounds for the expectation functional which can eliminate the need for multidimensional numerical integration and then to use these bounding functions in the inner optimization. This paper is a contribution in this direction.

We are concerned with the two-stage stochastic linear programming problem with fixed recourse, see, for instance, Kall (1976) and Wets (1982). The model is

$$Z^* := \min_x \{c'x + E_\omega[\phi(x,\xi,\eta)] : Ax = b,\ x \geqslant 0\} \tag{1}$$

where

$$\phi(x,\xi,\eta) := \min_y \{q(\eta)'y : Wy = h(\xi) - T(\xi)x,\ y \geqslant 0\}, \tag{2}$$

with primes (′) denoting transposition of vectors and E_ω denoting mathematical expectation with respect to the joint random vector $\omega := (\xi, \eta)$. The domain of the random vector ξ is denoted by $\Xi \subset \Re^K$ and that of η by $\Theta \subset \Re^L$. The joint domain of the random $(K+L)$-vector ω is denoted by Ω. The recourse problem (2) has fixed recourse because the $(m_2 \times n_2)$-matrix W is deterministic. A is a fixed $(m_1 \times n_1)$-matrix and c is a fixed vector in $\Re^{n_1}$. T is a $(m_2 \times n_1)$-dimensional stochastic matrix and q is a n_2-dimensional random vector. Define the set

$$X^1 := \{x \in \Re^{n_1} : Ax = b,\ x \geqslant 0\}. \tag{3}$$

Let the first stage decisions x induced by the second stage recourse problem (2) (for almost sure satisfaction of its constraints) be

$$X^2 := \{x \in \Re^{n_1} : \exists y \geqslant 0 \text{ s.t. } Wy = h(\xi) - T(\xi)x,\ \ \forall \xi \in \Xi\}. \tag{4}$$

The set of feasible solutions of the stochastic program (1)–(2) is

$$X := X^1 \cap X^2. \tag{5}$$

Assume that:
(A1) $X \neq \varnothing$,
(A2) $\{\pi \in \Re^{m_2} : \pi' W \leqslant q(\eta),\ \eta \in \Theta\} \neq \varnothing$,
(A3) $q(\eta) \equiv q_0 + Q\eta$,
$h(\xi) \equiv h_0 + H\xi$,
$T(\xi) \equiv T_0 + \sum_{k=1}^K T_k \xi_k$,
where the vectors q_0, h_0 and matrices Q, H, T_k are all fixed and defined with appropriate dimensions,
(A4) $\Omega = \Xi \times \Theta$, and
(A5) Ξ and Θ are convex, compact subsets of $\Re^K$ and $\Re^L$, respectively.

Assumption (A1) ensures that the stochastic program (1)–(2) is feasible with probability one while (A2) ensures that the recourse problem (2) is never unbounded. Furthermore, due to (A3), the recourse function $\phi(x,\xi,\eta)$ is a convex-concave saddle function in (ξ,η) for fixed $x \in X$. Moreover, from the linear programming theory, $\phi(x,\xi,\eta)$ is polyhedral in (ξ,η) for fixed $x \in X$, thus ensuring continuity of ϕ for fixed $x \in X$. The convexity assumption in (A5) is not restrictive for, if not, one could work with the convex hulls of the domains. In particular, consider these convex hulls as being generated by a finite set of extreme points, which we shall follow subsequently in order to make the exposition easy.

Properties of the program (1)–(2) are well known, see Wets (1974). For example, the set X is convex. Moreover, since integration with respect to a probability measure preserves order, convexity of ϕ in x implies that the deterministic equivalent program of (1)–(2) is a convex program. However, direct application of convex programming algorithms is difficult (except, for example, when W has special structure such as simple recourse) due to the multidimensional integration required for evaluating (1), of which the integrand is the optimization in (2). Such procedures become even more prohibitive computationally as one would need to evaluate gradients as well. A standard approach to alleviate this difficulty is to develop approximations.

This paper is concerned with determining tight upper and lower approximations to (1)–(2) using limited moment information on the random vector ω. This is accomplished by developing upper and lower bounds on the expectation of the recourse function $\phi(x,\xi,\eta)$ for a fixed first stage decision $x \in X$. The literature on bounds for such problems is extremely sparse. Frauendorfer (1988b, c, d, 1989) obtained upper and lower bounds for the expectation of the convex-concave recourse function. These bounds have been derived when domains are either multidimensional rectangles or simplices. In the former case, approximating extremal discrete distributions are determined on vertices of the domains, and thus, is exponential in the dimensions of the random vectors, while in the latter case, bounding computations can be afforded

Received March 30, 1990; revised May 25, 1992.
AMS 1980 subject classification. Primary: 90C42.
IAOR 1973 subject classification. Main: Programming, probabilistic.
OR/MS Index 1978 subject classification. Primary: 663 Programming/Stochastic.
Key words. Stochastic programs with recourse, generalized moment problems, bounds on expectation

0364-765X/94/1902/0292/$01.25

within a quadratic amount of function evaluations. In §2, we extend the ideas to determine tight upper and lower bounds using only first and cross moments of the random vector ω, with no independence assumption and requiring only that the random vectors have bounded domains, more precisely, polytopes.

Using these bounds, in §3, we obtain tight upper and lower bounds using only first moments and without any independence assumption. These bounds may be viewed as a generalization of the known bounds in the pure convex case, that is, the special case when randomness is only present in h and T of (2). Bounding techniques for this special case has been developed by, for example, Birge and Wets (1986), Dula (1992), Frauendorfer (1988a), and Gassmann and Ziemba (1986), using limited moment information.

In §4, we apply these bounding strategies to the stochastic linear program (1)–(2) to obtain finite mathematical programming formulations to compute the bounds. In particular, we obtain upper and lower bounds which are convex polyhedral in the first stage decisions x so that the bounding computations do not require any more sophistication than solving linear programs.

Tightness of bounds is discussed in the context of generalized moment problems. Specifically, we investigate if the bounds solve the dual of the moment problem. For this purpose, the required duality relationships are summarized in the remainder of this section. Notation is introduced as it becomes necessary.

1.1. Bounds using moment problems. We define the generalized moment problem without any reference to a stochastic program and subsequently duality relationships are specified.

Consider a set of finite measurable functions $f_i(\omega)$: $\Re^n \to \Re$, for $i = 1, \ldots, m+1$, with ω being a random vector mapping the measurable space $(\Omega, \mathscr{B})$ to $\Re^n$ and $\mathscr{B}$ the Borel-sigma field of the events in Ω. Suppose the (*generalized*) moments $E_\omega[f_i(\omega)]$ for $i = 1, \ldots, m$ are given, where $E_\omega[\cdot]$ denotes mathematical expectation with respect to the true probability measure on $(\Omega, \mathscr{B})$. For simplicity, $E_\omega[f_i(\omega)]$ is written as Ef_i. We are concerned with upper bounds on Ef_{m+1}. Given the generalized moments $Ef_1, \ldots, Ef_m$, we say, a *tight* upper bound is determined for Ef_{m+1} when one solves the following moment problem (GMP), see Kemperman (1968):

$$(6)\qquad \begin{aligned} Ef_{m+1} \leqslant V(\text{GMP}) &:= \sup_{P \in \mathscr{P}} \int_\Omega f_{m+1}(\omega)\, P(d\omega) \\ \text{s.t.} &\quad \int_\Omega f_i(\omega)\, P(d\omega) = Ef_i, \qquad i = 1, \ldots, m. \end{aligned}$$

where $\mathscr{P}$ is the set of all probability measures on Ω. The dual formulation of (6) is the semi-infinite linear program (SIP):

$$(7)\qquad \begin{aligned} V(\text{SIP}) &:= \inf_y \; y_0 + \sum_{i=1}^m (Ef_i) y_i \\ \text{s.t.} &\quad y_0 + \sum_{i=1}^m f_i(\omega) y_i \geqslant f_{m+1}(\omega), \quad \forall \omega \in \Omega. \end{aligned}$$

While the weak duality for (6) and (7) is immediate, i.e., $V(\text{GMP}) \leqslant V(\text{SIP})$, to ensure strong duality, certain mild conditions are required, in general.

THEOREM 1.1. *If*
(i) f_i, $i = 1, \ldots, m$, *are continuous*,
(ii) Ω *is compact, and*
(iii) *there exists a probability measure which achieves the given moments* Ef_i*s on* Ω,
then (GMP) *is solvable and* $V(GMP) = V(SIP)$.

For a proof of the above results, see Kall (1987) or Kemperman (1968). The issue of strong duality for semi-infinite linear programs and the associated dual problems has been addressed in a general sense by Glashoff and Gustafson (1983) who also develop this theory from first principles.

It is assumed that the information on limited moments of the random vector ω is such that there exists a true probability measure that achieves these. Moreover, for the stochastic program defined before, together with the assumptions listed therein, the sufficient conditions in Theorem 1.1 are satisfied. Consequently, we use strong duality for the underlying moment problem and its semi-infinite dual, without any ambiguity.

2. Bounds for the expectation of the recourse function. Consider the stochastic program (1)–(2) under Assumptions (A1)–(A5). The recourse function $\phi(x, \xi, \eta)$ is a convex-concave saddle function in $(\xi, \eta) := \omega$, for fixed $x \in X$. Let us fix the following notation with regard to domains of the random vectors: Ξ is a polytope with extreme points $u^1, \ldots, u^I$ and Θ is a polytope with extreme points $v^1, \ldots, v^J$. The first moments of ξ and η, respectively, denoted by $\bar{\xi}$ and $\bar{\eta}$, as well as the cross moments $E[\xi_k \eta_l]$ (for $k = 1, \ldots, K$; $l = 1, \ldots, L$), denoted by m_{kl}, are all assumed to exist and be finite. The expectation of the recourse function $\phi(x, \xi, \eta)$, for fixed $x \in X$, is denoted by $\bar{\phi}(x)$.

We are concerned with developing upper and lower bounds to $\bar{\phi}(x)$ using limited moment information of the random vector ω and without making any independence assumption with regard to the random vector ω. We take an approach similar to that used by Birge and Wets (1986, §5) in their work in the pure convex case. The basic idea is to consider each realization of the random vector as a convex combination of the extreme points of the domain. Thus, by the convexity of Ξ, for given $\xi \in \Xi$ there exist scalars $\lambda(\xi)$ that satisfy

$$(8)\qquad \sum_{i=1}^I u^i \lambda_i(\xi) = \xi, \qquad \sum_{i=1}^I \lambda_i(\xi) = 1, \qquad \lambda_i(\xi) \geqslant 0, \quad i = 1, \ldots, I.$$

Similarly, by the convexity of Θ, for given $\eta \in \Theta$ there exist scalars $\mu(\eta)$ satisfying

$$(9)\qquad \sum_{j=1}^J v^j \mu_j(\eta) = \eta, \qquad \sum_{j=1}^J \mu_j(\eta) = 1, \qquad \mu_j(\eta) \geqslant 0, \quad j = 1, \ldots, J.$$

Then by the saddle property of $\phi(x, \ldots)$,

$$(10)\qquad \sum_{j=1}^J \mu_j(\eta) \phi(x, \xi, v^j) \leqslant \phi(\xi, \eta) \leqslant \sum_{i=1}^I \lambda_i(\xi) \phi(x, u^i, \eta)$$
$$\text{for all } (\xi, \eta) \in \Xi \times \Theta = \Omega.$$

Upon integration of (10), with respect to the true probability measure, bounding inequalities for $\bar{\phi}(x)$ are

$$(11)\qquad \sum_{j=1}^{J} E\left[\mu_j(\eta)\phi(x,\xi,v^j)\right] \leqslant \bar{\phi}(x) \leqslant \sum_{i=1}^{I} E\left[\lambda_i(\xi)\phi(x,u^i,\eta)\right].$$

The lower and upper bounds in (11) are difficult to compute since the scalars λ and μ are to be determined according to (8) and (9), respectively, and moreover, product terms involving these scalars and nonlinear functional forms of ϕ need to be integrated with respect to the true probability measure. To alleviate this difficulty, we use a two-step procedure in the next section. First, each of the functions $\phi(x,\xi,v^j)$ and $\phi(x,u^i,\eta)$ are linearized at some arbitrary points in the respective domains, similar in spirit to what was done in Frauendorfer (1988c). Bounds are then determined by considering the resulting inequalities in expectation, an idea originally due to Birge and Wets (1986). Since the bounds from this procedure would depend on the chosen linearizations, in the second step, we optimize over all possible linearizations to tighten the bounds. Finally, we will show that these bounds solve the respective generalized moment problems, hence, the bounds are tight.

2.1. Tight bounds using first and cross moments. In the ensuing development, the first stage decision x $(\in X)$ is fixed. First, considering the lower bounding inequality in (11), linearize each of the convex functions $\phi(x,\xi,v^j)$, being convex in $\xi\in\Xi$, at some known point $\hat{\xi}^j\in\Xi$, i.e., consider the supporting hyperplane $\langle\pi^j,\xi\rangle + \sigma^j$, for fixed coefficients $\pi^j\in\Re^K$ and $\sigma^j\in\Re$, to the epigraph of $\phi(x,\xi,v^j)$ at $(\hat{\xi}^j,\phi(x,\hat{\xi}^j,v^j))$. Dependence of the latter coefficients on x is suppressed for simplicity. The coefficient vector π^j is a subgradient of the recourse function at $\xi=\hat{\xi}^j$ and may be interpreted as an optimal dual solution of the linear program:

$$(12)\qquad \phi(x,\hat{\xi}^j,v^j) = \min_y\left\{q(v^j)'y\colon Wy = h(\hat{\xi}^j) - T(\hat{\xi}^j)x,\ y\geqslant 0\right\}.$$

Similarly, linearize each of the concave functions $\phi(x,u^i,\eta)$ in the upper bounding inequality in (11), being concave in $\eta\in\Theta$, at some known point $\hat{\eta}^i\in\Theta$ to obtain the linear support $\langle\alpha^i,\eta\rangle+\beta^i$. $\alpha^i\in\Re^L$ is a subgradient of $\phi(x,u^i,\eta)$ at $\eta=\hat{\eta}^i$ and is interpreted as a primal solution of the linear program:

$$(13)\qquad \phi(x,u^i,\hat{\eta}^i) = \min_y\left\{q(\hat{\eta}^i)'y\colon Wy = h(u^i) - T(u^i)x,\ y\geqslant 0\right\}.$$

Therefore,

$$(14)\qquad \langle\alpha^i,\eta\rangle + \beta^i \geqslant \phi(x,u^i,\eta),\qquad \forall\eta\in\Theta \text{ and } i=1,\ldots,I,$$

$$(15)\qquad \langle\pi^j,\xi\rangle + \sigma^j \leqslant \phi(x,\xi,v^j),\qquad \forall\xi\in\Xi \text{ and } j=1,\ldots,J.$$

Substituting (14)–(15) in (11), the resulting lower and upper bounds to the expectation functional are

$$(16)\qquad \sum_{j=1}^{J} E\left[\mu_j(\eta)\langle\pi^j,\xi\rangle + \mu_j(\eta)\sigma^j\right] \leqslant \bar{\phi}(x) \leqslant \sum_{i=1}^{I} E\left[\lambda_i(\xi)\langle\alpha^i,\eta\rangle + \lambda_i(\xi)\beta^i\right]$$

where λ's and μ's satisfy (8) and (9), respectively.

Denote the kth element of the K-vector u^i by u_k^i and that of π^j by π_k^j. Also, the lth element of the L-vector v^j is denoted by v_l^j and that of α^i by α_l^i. Since Ξ is compact, for each component random variable ξ_k, $k=1,\ldots,K$, there exist (finite) lower and upper bounds denoted by $\hat{a}_{k0}$ and $\hat{a}_{k1}$, respectively. Similarly, those for components η_l, $l=1,\ldots,L$, are denoted by $\tilde{a}_{l0}$ and $\tilde{a}_{l1}$, respectively.

Proposition 2.1.

$$(17)\qquad \begin{aligned} \bar{\phi}(x) \leqslant \max_{t,p}\ & \sum_{i=1}^{I}\langle\alpha^i,t^i\rangle + \beta^i p_i \\ \text{s.t.}\ & \sum_{i=1}^{I} u_k^i t_l^i = m_{kl},\quad \forall k=1,\ldots,K;\quad \forall l=1,\ldots,L, \\ & \sum_{i=1}^{I} t_l^i = \bar{\eta}_l,\quad \forall l, \\ & \sum_{i=1}^{I} u_k^i p_i = \bar{\xi}_k,\quad \forall k, \\ & \sum_{i=1}^{I} p_i = 1,\quad p_i\geqslant 0, \\ & \tilde{a}_{l0}p_i \leqslant t_l^i \leqslant \tilde{a}_{l1}p_i,\quad \forall l,i. \end{aligned}$$

Proof. From every feasible solution $\lambda(\xi)$ of (8), a feasible solution of (17) can be constructed by setting

$$p_i := E\left[\lambda_i(\xi)\right]\quad\text{and}\quad t_l^i := E\left[\lambda_i(\xi)\eta_l\right]$$

for all $i=1,\ldots,I$ and $l=1,\ldots,L$. Moreover, under this construction, the upper bound in (16) develops to be the objective function in (17). Hence, the right-hand side upper bound in (16) is obtained as a feasible point lower bound to the maximization in (17). □

In an analogous manner, the following lower bound is derived.

Proposition 2.2.

$$(18)\qquad \begin{aligned} \bar{\phi}(x) \geqslant \min_{t,p}\ & \sum_{j=1}^{J}\langle\pi^j,t^j\rangle + \sigma^j p_j \\ \text{s.t.}\ & \sum_{j=1}^{J} v_l^j t_k^j = m_{kl},\quad \forall k=1,\ldots,K;\quad \forall l=1,\ldots,L, \\ & \sum_{j=1}^{J} t_k^j = \bar{\xi}_k,\quad \forall k, \\ & \sum_{j=1}^{J} v_l^j p_j = \bar{\eta}_l,\quad \forall l, \\ & \sum_{j=1}^{J} p_j = 1,\quad p_j\geqslant 0, \\ & \hat{a}_{k0}p_j \leqslant t_k^j \leqslant \hat{a}_{k1}p_j,\quad \forall k,j. \end{aligned}$$

The bounds in (17) and (18) require solution of linear programs. Although these are valid upper and lower bounds to $\bar{\phi}(x)$ using only first and cross moments and without any independence assumption, even tighter bounds may be obtained as follows. Moreover, these tightened bounds also serve to develop the *best* possible upper and lower bounds under the *linearization* procedure.

PROPOSITION 2.3.

$$\bar{\phi}(x) \leq \max_{\rho \in \mathscr{C}} \sum_{i,j} (\langle \alpha^i, v^j \rangle + \beta^i)\rho_{ij} \tag{19}$$

and

$$\bar{\phi}(x) \geq \min_{\rho \in \mathscr{C}} \sum_{i,j} (\langle \pi^j, u^i \rangle + \sigma^j)\rho_{ij} \tag{20}$$

where the feasible set $\mathscr{C}$ *is the polytope defined by*

$$\mathscr{C} := \Big\{ \rho \in \Re^{I \times J}: \ \sum_{i,j} (u^i, v^j)\rho_{ij} = (\bar{\xi}, \bar{\eta}),\ \sum_{i,j} u^i_k v^j_l \rho_{ij} = m_{kl}, \tag{21}$$

$$\forall k = 1, \dots, K;\ l = 1, \dots, L,\ \sum_{i,j} \rho_{ij} = 1,\quad \rho_{ij} \geq 0 \Big\}.$$

PROOF. For any feasible solution $\lambda(\xi)$ of (8) and $\mu(\eta)$ of (9), define

$$\rho_{ij} := E\big[\lambda_i(\xi)\mu_j(\eta)\big].$$

For this $\lambda(\xi)$ and $\mu(\eta)$, defining p_i and t^i_l, $i = 1, \dots, I$ and $l = 1, \dots, L$, as in the proof of Proposition 2.1, observe that $\sum_j v^j \rho_{ij} = t^i$, $\sum_j \rho_{ij} = p_i$, and $\rho_{ij} \geq 0$ hold. Substituting these in the upper bound in (17), the upper bound in (19) follows. Similarly, one obtains the lower bound in (20). □

COROLLARY 2.4. *If the domains* Ξ *and* Θ *are multidimensional rectangles of the form* $\Xi := \times_{k=1}^{K} [\hat{a}_{k0}, \hat{a}_{k1}]$ *and* $\Theta := \times_{l=1}^{L} [\tilde{a}_{l0}, \tilde{a}_{l1}]$, *then the upper and lower bounds in* (17) *and* (18) *coincide with those in* (19) *and* (20), *respectively*.

PROOF. Considering the upper bound in (19), from any of its feasible solutions, a feasible solution for the linear program in (17) may be constructed having the same objective value, i.e., the bound in (19) is never worse than that in (17). Consider a feasible solution (t^i, p_i), $\forall i$, of (17). Since Θ is a multidimensional interval, $t^i/p_i \in \Theta$ for $p_i \neq 0$. Thus, for each i, there exist convex combinations θ_{ij} such that $t^i = (\sum_j v^j \theta_{ij})p_i$. The construction $\rho_{ij} := p_i \theta_{ij}$ results in a feasible solution to (19), implying that the upper bound in (17) is as good as that in (19), hence the result. The lower bounding result follows analogously. □

If the supports are multidimensional intervals, due to Corollary 2.4, we may continue the ensuing discussion with the bounds in (17) and (18). However, we shall keep the discussion general by requiring only that the supports Ξ and Θ be convex polyhedral sets. Consequently, we shall be concerned with the bounds given in (19) and (20). The latter upper and lower bounds, respectively, depend on the *linearizing* points which determine the supporting hyperplanes in (14) and (15). Therefore, it is of interest to determine those $\hat{\eta}^{*i}$, $i = 1, \dots, I$, that lead to the *best* linearization upper bound. That is, one needs to determine that linearization of $\phi(x, u^i, \eta)$ for each i such that the upper bound in (19) becomes the smallest of all possible linearizations. Likewise, we obtain that linearization of $\phi(x, \xi, v^j)$ for each j such that the lower bound in (20) becomes the largest of all possible linearizations. For this development, first note the following result.

LEMMA 2.5. *Let* $\rho \in \mathscr{C}$ *where* $\mathscr{C}$ *is as defined in* (21). *Then*,

(i) $\forall i = 1, \dots, I$: $\quad \sum_j \rho_{ij} = 0$ *implies* $\sum_j v^j \rho_{ij} = 0$,

$\quad \sum_j \rho_{ij} > 0$ *implies* $\dfrac{\sum_j v^j \rho_{ij}}{\sum_j \rho_{ij}} \in \Theta$.

(ii) $\forall j = 1, \dots, J$: $\quad \sum_i \rho_{ij} = 0$ *implies* $\sum_i u^i \rho_{ij} = 0$,

$\quad \sum_i \rho_{ij} > 0$ *implies* $\dfrac{\sum_i u^i \rho_{ij}}{\sum_i \rho_{ij}} \in \Xi$.

PROOF. For $\rho \in \mathscr{C}$, since ρ_{ij} is nonnegative, $\sum_j \rho_{ij} = 0$ (for some i) implies $\rho_{ij} = 0$ for all j, and thus, $\sum_j v^j \rho_{ij} = 0$ for such i. If $\sum_j \rho_{ij} \neq 0$ for some i, then

$$\frac{\sum_j v^j \rho_{ij}}{\sum_j \rho_{ij}} = \sum_j v^j \left(\frac{\rho_{ij}}{\sum_j \rho_{ij}} \right) = \sum_j v^j \nu_j$$

where $\nu_j := \rho_{ij}/\sum_j \rho_{ij}$. Since $\sum_j \nu_j = 1$ and $\nu_j \geq 0$ for all j, $\sum_j v^j \nu_j \in \Theta$ proving the first assertion. The second assertion follows similarly. □

THEOREM 2.5. *The best-linearization upper bound from* (19) *is*

$$\begin{aligned} \phi_U(x) := \max_{\rho \in \mathscr{C}} \quad & \sum_{i,j} \rho_{ij} \phi(x, u^i, \hat{\eta}^i(\rho)) \\ \text{s.t.} \quad & \sum_j v^j \rho_{ij} = \hat{\eta}^i(\rho) \sum_j \rho_{ij}, \qquad i = 1, \dots, I, \end{aligned} \tag{22}$$

where $\mathscr{C}$ *is defined in* (21).

PROOF. The best-linearization upper bound from (19) is obtained from

$$\inf_{\alpha, \beta} \left[\sup_{\rho \in \mathscr{C}} \left\{ \sum_{i,j} (\alpha^i v^j + \beta^i)\rho_{ij} : \alpha^i \eta + \beta^i \geq \phi(x, u^i, \eta),\ \eta \in \Theta,\ \ \forall i \right\} \right]. \tag{23}$$

In (23), transposes have been removed for simplicity. Observe that (23) is equivalent to:

$$\sup_{\rho \in \mathscr{C}} \left[\inf_{\alpha, \beta} \left\{ \sum_{i,j} (\alpha^i v^j + \beta^i)\rho_{ij} : \alpha^i \eta + \beta^i \geq \phi(x, u^i, \eta),\ \eta \in \Theta,\ \ \forall i \right\} \right]. \tag{24}$$

For a given feasible solution $\rho \in \mathscr{C}$, consider the inner infimum in (24), the value of which is denoted by $Z(\rho)$. This problem decomposes to I subproblems of the form

$$(25)\quad Z^i(\rho) := \inf_{\alpha^i, \beta^i} \left\{ \alpha^i \sum_j v^j \rho_{ij} + \beta^i \sum_j \rho_{ij} : \alpha^i \eta + \beta^i \geq \phi(x, u^i, \eta),\ \eta \in \Theta \right\}$$

and $Z(\rho) = \sum_i Z^i(\rho)$. First consider the case when $\sum_j \rho_{ij} \neq 0$. Let (α^i, β^i) be a feasible pair to (25), the existence of such being ensured by the concavity of $\phi(x, u^i, \eta)$ in η. Then multiplying the constraints in (25) by $\sum_j \rho_{ij}$, since

$$\hat{\eta}^i(\rho) := \frac{\sum_j v^j \rho_{ij}}{\sum_j \rho_{ij}} \in \Theta$$

(by Lemma 2.5), it follows that $Z^i(\rho) \geq \sum_j \rho_{ij}(x, u^i, \hat{\eta}^i(\rho))$ holds. Furthermore, the continuity and concavity of $\phi(x, u^i, \cdot)$ is sufficient to guarantee that $Z^i(\rho) = \sum_j \rho_{ij} \phi(x, u^i, \hat{\eta}^i(\rho))$ is attained; i.e., choose the coefficients (α^i, β^i) corresponding to the supporting hyperplane at the boundary point $(\hat{\eta}^i(\rho), \phi(x, u^i, \hat{\eta}^i(\rho)))$ of the closed convex set $\{(\eta, \nu): \nu \leq \phi(x, u^i, \eta),\ \eta \in \Theta,\ \nu \in \Re\}$. On the other hand, if $\sum_j \rho_{ij} = 0$, by Lemma 2.5, $\sum_j v^j \rho_{ij} = 0$ and (one may choose any $\hat{\eta}^i \in \Theta$) thus $Z^i(\rho) = 0$ (hence it is not necessary to linearize $\phi(x, u^i, \eta)$). Noting that $\mathscr{C}$ is a closed set and replacing the outer "sup" by "max" completes the proof. □

THEOREM 2.7. *The best-linearization lower bound from* (20) *is*

$$(26)\quad \begin{aligned} \phi_L(x) := \min_{\rho \in \mathscr{C}} \quad & \sum_{i,j} \rho_{ij} \phi(x, \hat{\xi}^j(\rho), v^j) \\ \text{s.t.} \quad & \sum_i u^i \rho_{ij} = \hat{\xi}^j(\rho) \sum_i \rho_{ij}, \qquad j = 1, \ldots, J, \end{aligned}$$

where $\mathscr{C}$ *is defined in* (21).

PROOF. The proof is similar to that of Theorem 2.6. □

It is clear from (22) and (26) that one needs to solve nonlinear programs to obtain the best-linearized upper and lower bounds. However, as follows from Lemma 2.8, these are in fact convex programs.

LEMMA 2.8. *Let* $g: \Re^n \to \Re$ *be convex and define* $f: \Re^n \times (0, \infty) \to \Re$ *by* $f(t, p) = pg(t/p)$. *Then* f *is (jointly) convex.*

PROOF. We show that epi f is a convex cone.

$$\begin{aligned} \text{epi } f &= \left\{ (t, p, \theta): \theta \geq pg\left(\frac{t}{p}\right),\ p > 0 \right\} \\ &= \{(t, p, \theta): \exists \nu > 0 \text{ with } \nu\theta \geq g(\nu t),\ \nu p = 1\} \\ &= \{(t, p, \theta): \exists \nu > 0 \text{ with } \nu(t, p, \theta) \in \mathscr{E}\} \end{aligned}$$

where $\mathscr{E} := \{(t, p, \theta): p = 1, \theta \geq g(t)\}$. Since $g(\cdot)$ is convex, $\mathscr{E}$ is a convex set. Next observe that epi f = rc $\mathscr{E}$, the recession cone of $\mathscr{E}$, and is thus convex. □

Although the bounding expressions $\phi_U(x)$ and $\phi_L(x)$ are convex programming problems, for fixed first stage decisions x, our intent here is not to suggest solution schemes for computing these bounds in the present form. We will use these bounds to derive finite linear programming formulations by using the properties of the recourse function $\phi(x, \xi, \eta)$, being defined by (2), which is the discussion in §4.

2.2. Convexity of the bounds. Observe that the upper and lower bounds are obtained through approximating the true probability distribution by discrete probability distributions defined on the boundary of the joint domain Ω. Moreover, these approximating probability measures have the same first moments and cross moments as the given true measure. However, these approximating measures may depend on the functional form of ϕ, and more importantly, on the first stage decisions x. Nevertheless, as shown next, the upper bounding measure preserves convexity of the bound in the first stage decisions x.

THEOREM 2.9. *The upper bound* $\phi_U(x)$ *in* (22) *is convex in the first stage decisions* $x \in X$.

PROOF. For $x^1, x^2 \in \Re^{n_1}$, such that $x^1 \neq x^2$, and scalar $\nu \in (0, 1)$, defining $x^\nu := \nu x^1 + (1 - \nu) x^2$,

$$\begin{aligned} \phi_U(x^\nu) = \max_{\rho \in \mathscr{C}} \quad & \sum_{i,j} \rho_{ij} \phi(x^\nu, u^i, \hat{\eta}^i(\rho)) \\ \text{s.t.} \quad & \sum_j v^j \rho_{ij} = \hat{\eta}^i(\rho) \sum_j \rho_{ij} \\ \leq \max_{\rho \in \mathscr{C}} \quad & \sum_{i,j} \rho_{ij} \left[\nu \phi(x^1, u^i, \hat{\eta}^i(\rho)) + (1 - \nu) \phi(x^2, u^i, \hat{\eta}^i(\rho)) \right] \\ \text{s.t.} \quad & \sum_j v^j \rho_{ij} = \hat{\eta}^i(\rho) \sum_j \rho_{ij} \\ \leq \max_{\rho^1, \rho^2 \in \mathscr{C}} \quad & \sum_{i,j} \rho^1_{ij} \left[\nu \phi(x^1, u^i, \hat{\eta}^i(\rho^1)) \right] + \sum_{i,j} \rho^2_{ij} \left[(1 - \nu) \phi(x^2, u^i, \hat{\eta}^i(\rho^2)) \right] \\ \text{s.t.} \quad & \sum_j v^j \rho^k_{ij} = \hat{\eta}^i(\rho^k) \sum_j \rho^k_{ij}, \qquad k = 1, 2, \\ = \nu \phi_U(x^1) &+ (1 - \nu) \phi_U(x^2), \end{aligned}$$

i.e., $\phi_U(x)$ is convex in x. □

In §4, when the bounds are formulated explicitly for the stochastic program (1)–(2), it will be shown that $\phi_U(x)$ is in fact polyhedral convex in x. The lower bound $\phi_L(x)$ in (26) may fail to be convex in x. In §4, we use $\phi_L(x)$ to derive a weaker lower bound which is convex in the first stage decisions x, still using first and cross moments.

Nonetheless, in certain special cases, when the domains Ξ and Θ are restricted to be of a certain shape, the bounds $\phi_U(x)$ and $\phi_L(x)$ are obtained with unique (i.e., independent of x or ϕ) discrete measures on the boundary of Ω, i.e., the optimizations in (22) and (26) are trivial, and consequently, convexity of $\phi_L(x)$ is guaranteed. This indeed is the case when Ξ and Θ are multidimensional simplices since every point in a simplex can be represented as a unique convex combination of the extreme points of the simplex, the barycentric coordinates. Consequently, the feasible set $\mathscr{C}$ in (21) implies that the values of $\sum_j \rho_{ij}$, $\sum_i \rho_{ij}$, for any i, and $\sum_j v^j \rho_{ij}$, $\sum_i u^i \rho_{ij}$, for any j, are uniquely determined. That is, the upper and lower bounding discrete probability measures are being determined independently of the function ϕ or the decisions x. This special case is discussed in Frauendorfer (1989).

When ξ and η are *univariate* random variables, hence the domains are one-dimensional simplices, the bounds $\phi_U(x)$ and $\phi_L(x)$ are simplified as follows:

THEOREM 2.10. *If ξ and η are both univariate (with interval domains $\Xi := [\hat{a}_0, \hat{a}_1]$, $\Theta := [\bar{a}_0, \bar{a}_1]$ such that $\hat{a}_0 < \hat{a}_1$ and $\bar{a}_0 < \bar{a}_1$), then denoting the cross moment $C := E[\xi\eta]$, the upper bound $\phi_U(x)$ simplifies to*

$$(27)\quad \phi_U(x) = \left(\frac{\hat{a}_1 - \bar{\xi}}{\hat{a}_1 - \hat{a}_0}\right)\phi\left(x, \hat{a}_0, \frac{\hat{a}_1\bar{\eta} - C}{\hat{a}_1 - \bar{\xi}}\right) + \left(\frac{\bar{\xi} - \hat{a}_0}{\hat{a}_1 - \hat{a}_0}\right)\phi\left(x, \hat{a}_1, \frac{C - \hat{a}_0\bar{\eta}}{\bar{\xi} - \hat{a}_0}\right)$$

and the lower bound $\phi_L(x)$ simplifies to

$$(28)\quad \phi_L(x) = \left(\frac{\bar{a}_1 - \bar{\eta}}{\bar{a}_1 - \bar{a}_0}\right)\phi\left(x, \frac{\bar{a}_1\bar{\xi} - C}{\bar{a}_1 - \bar{\eta}}, \bar{a}_0\right) + \left(\frac{\bar{\eta} - \bar{a}_0}{\bar{a}_1 - \bar{a}_0}\right)\phi\left(x, \frac{C - \bar{a}_0\bar{\xi}}{\bar{\eta} - \bar{a}_0}, \bar{a}_1\right).$$

PROOF. Let us consider the upper bound when ξ and η are univariate. Setting $u^1 = \hat{a}_0$, $u^2 = \hat{a}_1$, $v^1 = \bar{a}_0$, and $v^2 = \bar{a}_1$, from the corresponding feasible region $\mathscr{C}$ for ρ_{11}, ρ_{12}, ρ_{21}, and ρ_{22}, one can show that $(\rho_{i1} + \rho_{i2})$ and $(\bar{a}_0\rho_{i1} + \bar{a}_1\rho_{i2})$ are uniquely determined for each $i = 1, 2$ as follows:

$$\rho_{11} + \rho_{12} = \frac{\hat{a}_1 - \bar{\xi}}{\hat{a}_1 - \hat{a}_0} \quad\text{and}\quad \rho_{21} + \rho_{22} = \frac{\bar{\xi} - \hat{a}_0}{\hat{a}_1 - \hat{a}_0},$$

$$\bar{a}_0\rho_{11} + \bar{a}_1\rho_{12} = \frac{\hat{a}_1\bar{\eta} - C}{\hat{a}_1 - \bar{\xi}} \quad\text{and}\quad \bar{a}_0\rho_{21} + \bar{a}_1\rho_{22} = \frac{C - \hat{a}_0\bar{\eta}}{\bar{\xi} - \hat{a}_0}.$$

Moreover, since

$$\frac{\hat{a}_1 - \bar{\xi}}{\hat{a}_1 - \hat{a}_0} > 0 \quad\text{and}\quad \frac{\bar{\xi} - \hat{a}_0}{\hat{a}_1 - \hat{a}_0} > 0,$$

$$\hat{\eta}^1 = \frac{\hat{a}_1\bar{\eta} - C}{\hat{a}_1 - \bar{\xi}} \in \Theta \quad\text{and}\quad \hat{\eta}^2 = \frac{C - \hat{a}_0\bar{\eta}}{\bar{\xi} - \hat{a}_0} \in \Theta.$$

Substituting these in the upper bound $\phi_U(x) = \sum_{i=1}^2\sum_{j=1}^2 \rho_{ij}\phi(x, u^i, \hat{\eta}^i)$ gives (27). Similarly, one can obtain the lower bound in (28). □

The bounds in (27) and (28) for a saddle function in 3-dimensions are depicted in Figure 1.

REMARK. Frauendorfer (1988c) derived bounds for the expectation of a convex-concave function using the following moments (for the case when Ξ and Θ are multidimensional rectangles):

$$(29)\quad \begin{cases} m_{\Lambda,1} := \int_\Theta \left(\prod_{l\in\Lambda} \eta_l\right) P(d\eta) & \text{for all } \Lambda \in B_2, \\ m^l_{\Lambda,2} := \int_\Omega \eta_l \left(\prod_{k\in\Lambda} \xi_k\right) P(d\xi, d\eta) & \text{for all } l = 1, \dots, L \text{ and } \Lambda \in B_1, \\ m_{\Lambda,3} := \int_\Xi \left(\prod_{k\in\Lambda} \xi_k\right) P(d\xi) & \text{for all } \Lambda \in B_1, \\ m^k_{\Lambda,4} := \int_\Omega \xi_k \left(\prod_{l\in\Lambda} \eta_l\right) P(d\xi, d\eta) & \text{for all } k = 1, \dots, K \text{ and } \Lambda \subset B_2, \end{cases}$$

where B_1 is the set of all subsets of $\{1, \dots, K\}$ and B_2 is the set of all subsets of $\{1, \dots, L\}$. The moment information in (29) allows one to determine upper and lower bounding extremal (discrete) distributions uniquely. For the special case of univariate ξ and η, the upper and lower bounds due to Frauendorfer (1988c) can be shown to coincide with those given in (27) and (28), respectively.

2.3. Relation to generalized moment problems. An immediate concern would be whether or not the bounds $\phi_U(x)$ and $\phi_L(x)$ derived in the previous section are tight, in the sense of the corresponding generalized moment problems. We will show that this indeed is the case. First, observe that it is possible to attain these bounds as the exact expectation in certain cases, i.e.,

THEOREM 2.11. *If the convex-concave recourse function $\phi(x, \xi, \eta)$ is bilinear in ξ and η, for fixed x, then $\phi_L(x) = \bar{\phi}(x) = \phi_U(x)$ holds where $\phi_U(x)$ and $\phi_L(x)$ are defined in (22) and (26), respectively.*

PROOF. Bilinearity allows one to simplify the upper bounding formulation in (22) so that the constraints defining the set $\mathscr{C}$ imply that $\phi_U(x) = \bar{\phi}(x)$. The lower bounding result also follows trivially. □

To show that the bounds $\phi_U(x)$ and $\phi_L(x)$ are the *best* using first and cross moments, consider upper and lower bounds on $\bar{\phi}(x)$ by formulating the corresponding moment problems as follows, where the set of all probability measures on Ω is denoted by $\mathscr{P}$:

$$(30)\quad \begin{aligned} \bar{\phi}(x) \leqslant V_U(x) := \sup_{P\in\mathscr{P}} \quad & \int_\Omega \phi(x, \xi, \eta) P(d\xi, d\eta) \\ \text{s.t.} \quad & \int_\Omega \xi P(d\xi, d\eta) = \bar{\xi}, \\ & \int_\Omega \eta P(d\xi, d\eta) = \bar{\eta}, \\ & \int_\Omega \xi_k\eta_l P(d\xi, d\eta) = m_{kl}, \\ & k = 1, \dots, K;\ l = 1, \dots, L, \end{aligned}$$

and

$$(31)\quad \begin{aligned} \bar{\phi}(x) \geqslant V_L(x) := \inf_{P\in\mathscr{P}} \quad & \int_\Omega \phi(x, \xi, \eta) P(d\xi, d\eta) \\ \text{s.t.} \quad & \int_\Omega \xi P(d\xi, d\eta) = \bar{\xi}, \\ & \int_\Omega \eta P(d\xi, d\eta) = \bar{\eta}, \\ & \int_\Omega \xi_k\eta_l P(d\xi, d\eta) = m_{kl}, \\ & k = 1, \dots, K;\ l = 1, \dots, L. \end{aligned}$$

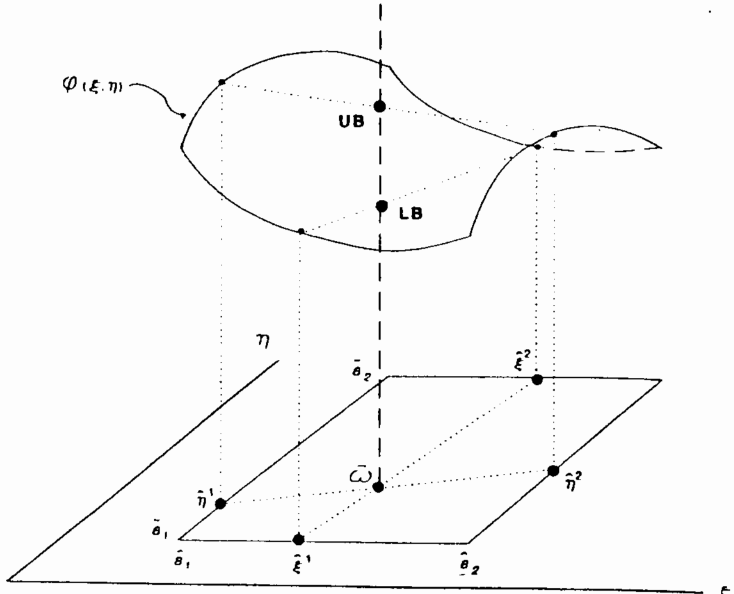

FIGURE 1. Upper and lower bounds using first and cross moments in the univariate case.

THEOREM 2.12. *Assuming that* (30) *and* (31) *are feasible*, $V_U(x) = \phi_U(x)$ *and* $V_L(x) = \phi_L(x)$ *hold, where* $\phi_U(x)$, $\phi_L(x)$, $V_U(x)$, *and* $V_L(x)$ *are defined in* (22), (26), (30), *and* (31), *respectively*.

PROOF. We shall prove the result for the upper bound. The proof for the lower bound is analogous.

Consider the symmetric dual problem of the upper bounding moment problem in (30), which is a semi-infinite linear program. By feasibility of (30), compactness of Ω, and continuity of $\phi(x, \cdot, \cdot)$, it follows by Theorem 1.1 that strong duality holds and thus,

$$
\begin{aligned}
V_U(x) = \inf_{\theta}\ & \theta_0 + \sum_{k=1}^{K} \theta_k^1 \bar{\xi}_k + \sum_{l=1}^{L} \theta_l^2 \bar{\eta}_l + \sum_{k,l} \theta_{kl} m_{kl} \\
\text{s.t.}\ & \theta_0 + \sum_{k=1}^{K} \theta_k^1 \xi_k + \sum_{l=1}^{L} \theta_l^2 \eta_l + \sum_{k,l} \theta_{kl} \xi_k \eta_l \geq \phi(x, \xi, \eta), \\
& \qquad \forall (\xi, \eta) \in \Omega. \\
= \inf_{\theta}\ & \theta_0 + \sum_{k=1}^{K} \theta_k^1 \bar{\xi}_k + \sum_{l=1}^{L} \theta_l^2 \bar{\eta}_l + \sum_{k,l} \theta_{kl} m_{kl} \\
\text{s.t.}\ & \theta_0 + \sum_{k=1}^{K} \theta_k^1 u_k^i + \sum_{l=1}^{L} \theta_l^2 \eta_l + \sum_{k,l} \theta_{kl} u_k^i \eta_l \geq \phi(x, u^i, \eta) \\
& \qquad \forall i = 1, \ldots, I;\ \eta \in \Theta.
\end{aligned}
\tag{32}
$$

The equality in (32) is due to convexity of $\phi(x, \cdot, \eta)$ and the Assumption (A4) that $\Omega = \Xi \times \Theta$. Consider any feasible solution $\rho \in \mathscr{C}$. Due to Lemma 2.5, $\Sigma_l \rho_{il} > 0$ implies

$$\frac{\Sigma_l v^l \rho_{il}}{\Sigma_l \rho_{il}} =: \hat{\eta}^i(\rho) \in \Theta.$$

If $\Sigma_l \rho_{il} = 0$, pick any arbitrary $\hat{\eta}^i(\rho)$ from the domain Θ. It is easy to verify that the pairs $(\Sigma_l \rho_{il}, \hat{\eta}^i(\rho))$, $i = 1, \ldots, I$, are dual feasible to (32)[1], hence $V_U(x) \geq \Sigma_{i,l} \rho_{il} \phi(x, u^i, \hat{\eta}^i(\rho))$ for any such feasible solution $\rho \in \mathscr{C}$. That is, $V_U(x) \geq \phi_U(x)$.

To claim that $V_U(x) \leq \phi_U(x)$ holds, note that, by Theorem 1.1, (30) is solvable. Let P^* be a maximizing measure in (30). Now, applying the upper bounding technique in the previous section, with the underlying probability measure being replaced by P^*, yields

$$\int_{\Omega} \phi(x, \xi, \eta) P^*(d\xi, d\eta) \leq \phi_U(x), \tag{33}$$

since P^* has the same first and cross moments as the *true* probability measure. Therefore, $V_U(x) \leq \phi_U(x)$ holds and the proof is completed. □

3. First moment bounds—dependent case. If the random vectors ξ and η are stochastically independent of each other, then Jensen's inequality and the Gassmann-Ziemba (1986) inequality can be utilized simultaneously to obtain upper and lower bounding functions for $\bar{\phi}(x)$, using only first moments, as follows.

$$
\begin{aligned}
\bar{\phi}(x) &= E_\eta\big[E_\xi[\phi(x, \xi, \eta)]\big] \quad \text{(due to independence)} \\
&\geq E_\eta\big[\phi(x, \bar{\xi}, \eta)\big] \quad \text{(due to Jensen's inequality)} \\
&\geq \min_{\mu \geq 0} \left\{ \sum_{j=1}^{J} \phi(x, \bar{\xi}, v^j) \mu_j :\ \sum_{j=1}^{J} v^j \mu_j = \bar{\eta},\ \sum_{j=1}^{J} \mu_j = 1 \right\}.
\end{aligned}
\tag{34}
$$

The last inequality follows from the Gassmann-Ziemba inequality being applied on the concave function $\phi(x, \bar{\xi}, \cdot)$. Similarly, one obtains

$$\bar{\phi}(x) \leq \max_{\lambda \geq 0} \left\{ \sum_{i=1}^{I} \phi(x, u^i, \bar{\eta}) \lambda_i :\ \sum_{i=1}^{I} u^i \lambda_i = \bar{\xi},\ \sum_{i=1}^{I} \lambda_i = 1 \right\}. \tag{35}$$

However, under dependence between ξ and η, the above derivation is no longer valid and thus one has to proceed differently. In this section, we show that the approach in §2 can be used to obtain tight lower and upper bounds for $\bar{\phi}(x)$, in the dependent case, under the first moment information. These bounds could also be viewed as generalizations of Jensen's and the Gassmann-Ziemba bounds, from the pure convex case to the convex-concave case. Most results follow trivially from those in the

[1] We thankfully acknowledge the referee who pointed out the dual feasibility here, thus shortening the proof.

preceding sections. Define the bounded convex polyhedron $\overline{\mathscr{C}}$ by

$$(36)\qquad \overline{\mathscr{C}} := \Big\{\rho \in \Re^{I \cdot J}: \sum_{i,j}(u^i, v^j)\rho_{ij} = (\bar{\xi}, \bar{\eta}),\ \sum_{i,j}\rho_{ij} = 1,\ \rho_{ij} \geq 0\Big\}.$$

Then,

THEOREM 3.1. *An upper bound to $\bar{\phi}(x)$ using only first moments is*

$$(37)\qquad \begin{aligned} \hat{\phi}_U(x) &:= \max_{\rho \in \overline{\mathscr{C}}} \sum_{i,j}\rho_{ij}\phi(x, u^i, \hat{\eta}^i(\rho)) \\ \text{s.t.}\quad & \sum_j v^j\rho_{ij} = \hat{\eta}^i(\rho)\sum_j \rho_{ij}, \qquad i = 1, \ldots, I. \end{aligned}$$

A lower bound to $\bar{\phi}(x)$ using only first moments is

$$(38)\qquad \begin{aligned} \hat{\phi}_L(x) &:= \min_{\rho \in \overline{\mathscr{C}}} \sum_{i,j}\rho_{ij}\phi(x, \hat{\xi}^j(\rho), v^j) \\ \text{s.t.}\quad & \sum_i u^i\rho_{ij} = \hat{\xi}^j(\rho)\sum_i \rho_{ij}, \qquad j = 1, \ldots, J, \end{aligned}$$

where $\overline{\mathscr{C}}$ is defined in (36).

PROOF. $\phi_U(x) \leq \hat{\phi}_U(x)$ and $\phi_L(x) \geq \hat{\phi}_L(x)$ trivially hold since $\mathscr{C} \subseteq \overline{\mathscr{C}}$, where $\phi_U(x)$ and $\phi_L(x)$ are defined in (22) and (26), respectively, hence the result. □

To see that the bounds $\hat{\phi}_U(x)$ and $\hat{\phi}_L(x)$ are tight, formulate the upper and lower bounding problems as generalized moment problems. One obtains

THEOREM 3.2. *Defining $\mathscr{P}$ as the set of all probability measures on the domain Ω,*

$$(39)\qquad \hat{\phi}_U(x) = \sup_{P \in \mathscr{P}}\Big\{\int_\Omega \phi(x, \xi, \eta)P(d\xi, d\eta): \int_\Omega \xi P(d\xi, d\eta) = \bar{\xi},\ \int_U \eta P(d\xi, \eta) = \bar{\eta}\Big\}$$

and

$$(40)\qquad \hat{\phi}_L(x) = \inf_{P \in \mathscr{P}}\Big\{\int_\Omega \phi(x, \xi, \eta)P(d\xi, d\eta): \int_\Omega \xi P(d\xi, d\eta) = \bar{\xi},\ \int_\Omega \eta P(d\xi, \eta) = \bar{\eta}\Big\}$$

hold where $\hat{\phi}_U(x)$ and $\hat{\phi}_L(x)$ are defined in (37) and (38), respectively.

PROOF. Noting that (39) and (40) are always feasible, the proof is similar to that in Theorem 2.12. □

These first moment bounds inherit the following properties:

THEOREM 3.3. (1) *$\hat{\phi}_U(x)$ is convex in x.*
(2) *If $\phi(x, \xi, \eta)$ is jointly linear in ξ and η, then $\hat{\phi}_U(x) = \bar{\phi}(x) = \hat{\phi}_L(x)$.*

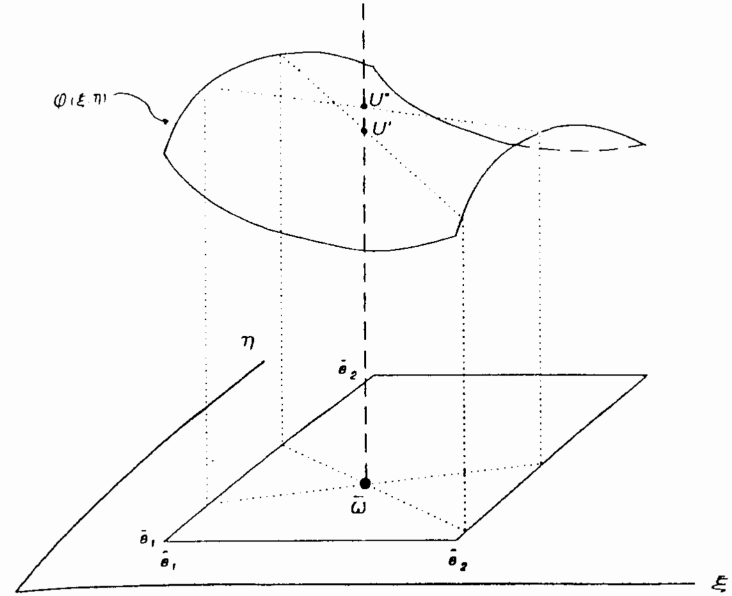

FIGURE 2. First moment bounds in the univariate case. To determine the upper bound, pick the plane through first moments, which gives the maximum height of U', U'', etc.

(3) *If the function ϕ is convex (in random vector ξ), for fixed x, then*
(i) *$\hat{\phi}_L(x)$ simplifies to Jensen's lower bound, and*
(ii) *$\hat{\phi}_U(x)$ simplifies to the Gassmann-Ziemba upper bound.*

PROOF. Straightforward. □

As pointed out in the previous section, when first and cross moments are used, the resulting bounds are trivially determined for the case when domains are multidimensional simplices. However, with the first moment bounds $\hat{\phi}_U(x)$ and $\hat{\phi}_L(x)$, such is not possible. As illustrated in Figure 2, when ξ and η are univariate, the upper bound is determined by choosing a "vertical" cross section of the saddle surface through the point of first moments ($\bar{\omega}$) such that the line joining the points of intersection of the saddle surface has a maximum height at $\bar{\omega}$.

When ξ and η are known to be stochastically independent, the bounds in (35) and (34) are generally tighter than those in (37) and (38), respectively.

4. Application to stochastic programming. In this section, the bounds derived using first and cross moments are applied to the stochastic program (1)–(2) to obtain finite mathematical programming formulations. The intent here is to develop upper and lower bounding formulations which are amenable to efficient solution techniques, such as linear programming.

Consider the upper and lower bounds $\phi_U(x)$ and $\phi_L(x)$ in (22) and (26), respectively, on the expectation $\bar{\phi}(x)$, using first and cross moments of the random vector. With Z^* being defined by (1), the resulting upper and lower bounding approximations for Z^* are

$$(41)\qquad Z^* \leq Z_U^* := \min_{x \in X}\{c'x + \phi_U(x)\}$$

and

$$Z^* \geq Z_L^* := \min_{x \in X}\{c'x + \phi_L(x)\}. \tag{42}$$

PROPOSITION 4.1. *For the stochastic program in* (1)–(2), *under the Assumptions* (A1)–(A5), $\phi_U(x)$ *is convex polyhedral in* $x \in X$, *where* $\phi_U(x)$ *is defined by* (22).

PROOF. That $\phi_U(x)$ is convex in x was shown in Theorem 2.9. Here we show that it is convex polyhedral, when applied to the recourse function definition in (2). For any $i = 1, \dots, I$,

$$\phi(x, u^i, \hat{\eta}^i(\rho)) := \min_{y^i}\{q(\hat{\eta}^i(\rho))'y^i : Wy^i = h(u^i) - T(u^i)x,\ y^i \geq 0\}. \tag{43}$$

Denote the feasible set on y^i, $i = 1, \dots, I$ (simply referred to as y), for any given x by $\mathscr{Y}(x)$. Thus,

$$\phi_U(x) = \max_{\rho \in \mathscr{C}} \min_{y \in \mathscr{Y}(x)} G(y, \rho) := \sum_{i,j} \rho_{ij} q(\hat{\eta}^i(\rho))'y^i. \tag{44}$$

By Assumption (A3) in §1, $G(y, \rho) = \sum_{i,j} q(v^j)'y^i\rho_{ij}$, i.e., $G(y, \rho)$ is bilinear in y and ρ. Moreover, since $\mathscr{C}$ and $\mathscr{Y}(x)$ are polyhedral sets, interchanging maximization and minimization in (44),

$$\phi_U(x) = \min_{y \in \mathscr{Y}(x)} \max_{\rho \in \mathscr{C}} G(y, \rho). \tag{45}$$

Using linear programming strong duality for the inner maximization in (45),

$$\max_{\rho \in \mathscr{C}} G(y, \rho) = \min_{w} \quad w^0 + w^1\bar{\xi} + w^2\bar{\eta} + \sum_{k,l} m_{kl}w^3_{kl}$$

$$\text{s.t.} \quad w^0 + w^1u^i + w^2v^j + \sum_{k,l} u^i_k v^j_l w^3_{kl} \geq q(v^j)'y^i, \quad \forall i, j.$$

Therefore,

$$\phi_U(x) = \min_{y,w} \quad w^0 + w^1\bar{\xi} + w^2\bar{\eta} + \sum_{k,l} m_{kl}w^3_{kl} \tag{46}$$

$$\text{s.t.} \quad Wy^i = h(u^i) - T(u^i)x, \quad \forall i,$$

$$w^0 + w^1u^i + w^2v^j + \sum_{k,l} u^i_k v^j_l w^3_{kl} - q(v^j)'y^i \geq 0, \quad \forall i, j,$$

$$y^i \geq 0, \quad \forall i,$$

and thus is convex polyhedral in x. □

Since $\phi_U(x)$ is convex polyhedral, as follows from the above proof, the upper bounding approximation Z_U^* in (41) can be solved as a linear program. On the contrary, the lower bounding approximation $\phi_L(x)$ may fail to be convex. To see this rewrite $\phi_L(x)$ using the recourse function definition in (2) as

$$\phi_L(x) = \min_{\rho \in \mathscr{C}, y \geq 0}\left\{\sum_{i,j} \rho_{ij} q(v^j)'y^j : Wy^j + T(\hat{\xi}^j(\rho))x - h(\hat{\xi}^j(\rho)) = 0, \quad \forall j\right\}. \tag{47}$$

Due to our Assumptions (A2) and (A3):

$$\phi_L(x) = \min_{\rho \in \mathscr{C}, y \geq 0} \quad \sum_{j=1}^{J} q(v^j)'y^j$$

$$\text{s.t.} \quad Wy^j + (T_0x - h_0)\sum_i \rho_{ij} \tag{48}$$

$$+ \sum_{k=1}^{K} (T_kx - h_k)\left(\sum_i u^i_k\rho_{ij}\right) = 0, \quad \forall j.$$

Although for any given $x \in X$, $\phi_L(x)$ could be evaluated by solving a linear program, observe that it cannot be concluded from (48) that $\phi_L(x)$ is convex in x. Therefore, in general, the lower bounding approximation Z_L^* in (42) involves nonconvex optimization. To overcome this computational difficulty, we derive below an alternative convex polyhedral lower bound $\bar{\phi}_L(x)$, using first and cross moments, which is generally weaker than $\phi_L(x)$.

Construct the $(K \times L)$-dimensional matrix M of cross moments, with the entry at the kth row and the lth column being m_{kl}.

PROPOSITION 4.2.

$$\bar{\phi}(x) \geq \phi_L(x) \geq \bar{\phi}_L(x) \tag{49}$$

where

$$\bar{\phi}_L(x) := \min_{y} \quad \sum_j q(v^j)'y^j \tag{50}$$

$$\text{s.t.} \sum_j Wy^j = h(\bar{\xi}) - T(\bar{\xi})x,$$

$$\sum_j v^j_l Wy^j = h(M_{(l)}) - T(M_{(l)})x + (h_0 - T_0x)(\bar{\eta}_l - 1),$$

$$l = 1, \dots, L, \phi y^j \geq 0, \quad j = 1, \dots, J,$$

and $M_{(l)}$ *is the* l*th column of the cross moment matrix* M. *Moreover*, $\bar{\phi}_L(x)$ *is convex polyhedral in* x.

PROOF. Summing the constraints (involving variables y^j) of (48) for all $j = 1, \dots, J$,

$$\sum_j Wy^j = h(\bar{\xi}) - T(\bar{\xi})x \tag{51}$$

follows since $\rho \in \mathscr{C}$. Multiplying the constraints (involving variables y^l) of (48) by c_j^l (for some $l = 1, \ldots, L$) and summing the resulting equalities together for all $j = 1, \ldots, J$,

$$(52)\qquad \sum_l c_j^l W y^l = (h_0 - T_0 x)\bar{\eta}_l + \sum_{k=1}^{K} (h_k - T_k x) m_{kl}$$

$$= h(M_l) - T(M_l)x + (h_0 - T_0 x)(\bar{\eta}_l - 1).$$

Since the constraints are aggregated, (51)–(52) provide a relaxation to the feasible set of (48), hence a lower bound which is convex polyhedral in x. □

The constraint aggregation procedure above should be viewed as a rather natural way to obtain a convex polyhedral lower bound from the generally nonconvex bound in (48), as evident from the next proposition.

PROPOSITION 4.3. *If Θ, the domain of the random vector η, is an L-dimensional simplex, then $\phi_l(x) = \bar{\phi}_l(x)$ for all $x \in X$.*

PROOF. Since $\phi_l(x) \geqslant \bar{\phi}_l(x)$ holds trivially from (49), we only need to show that $\phi_l(x) \leqslant \bar{\phi}_l(x)$ holds when Θ is a simplex.

For every feasible $(L + 1)$-tuple $(\Sigma_l y_n^l, \Sigma_l c_j^l y_n^l, l = 1, \ldots, L)$ of (50), where $n = 1, \ldots, n_2$, observe that Θ being a simplex implies that y_n^l, $\forall j$, $\forall n$ are completely and uniquely determined. It is easy to verify that this unique solution y^l, for the chosen feasible $(L + 1)$-tuple, is feasible in (48) since $\Sigma_l \rho_{ll}$ and $\Sigma_l u^l \rho_{ll}$ are uniquely determined for Θ being a simplex. □

Therefore, not only is the lower bound $\bar{\phi}_L(x)$ in (50) tight in the sense of Proposition 4.3, but it also allows one to solve the resulting lower bounding approximation

$$(53)\qquad Z_{LL}^* := \min_{x \in X} \{c'x + \bar{\phi}_L(x)\}$$

as a linear program. We illustrate these bounds on a simple numerical example below.

EXAMPLE 4.4. Consider the following two-stage stochastic programming (complete recourse) problem with random right-hand side, technology matrix, and cost coefficients:

$$Z^* = \min_{x \geqslant 0} 2x_1 + 2x_2 + E_{\xi,\eta}\Big[\min_{y \geqslant 0} \quad 2\eta_1 y_1 + 3\eta_2 y_2 + y_3\Big]$$

$$\text{s.t. } 2x_1 - x_2 \leqslant 1, \quad \text{s.t.} \quad 2y_1 + y_2 - 3y_3 = 2 + 3\xi_1 - 3(\xi_1 + \xi_2)x_1 + 4\xi_1 x_2,$$

$$x_1 + x_2 \leqslant 1, \qquad -y_1 + 3y_2 + y_3 = 4 + 2\xi_2 - x_1 - 3\xi_2 x_2$$

where $\Xi = [0, 1] \times [0, 1] = \Theta$ and the joint distribution of $(\xi_1, \xi_2, \eta_1, \eta_2)$ is given by

$$f_k(\xi_k, \eta_k) = \begin{cases} \dfrac{2\eta_k}{\xi_k} & \text{if } \eta_k \leqslant \xi_k, \\[2mm] \dfrac{2(1 - \eta_k)}{1 - \xi_k} & \text{if } \eta_k \geqslant \xi_k, \end{cases}$$

for $k = 1, 2$ with (ξ_1, η_1) being independent of (ξ_2, η_2). One obtains $\bar{\xi}_k = \bar{\eta}_k = 0.5$ for $k = 1, 2$ and the cross moments $m_{11} = m_{22} = \frac{5}{18}$ and $m_{12} = m_{21} = 0.25$.

Using the lower bounding approximation in (53), solving only a linear program, yields $Z_{LL}^* = 3.6369$ and the approximate solution $x_L = (0.5, 0)$. Using the upper bounding approximation in (41) and solving only a linear program, one obtains $Z_U^* = 3.7977$ and the approximate solution $x_U = (0, 0)$. Thus, the relative error of the bounds is

$$\frac{3.7977 - 3.6369}{3.6369} < 5\%.$$

Moreover, from the dual solution of the upper bounding linear program, one obtains that the true distribution is approximated by a discrete probability measure on the boundary of the joint domain as $\omega^1 = (0, 0, 0.4580, 0.5420)$, $p(\omega^1) = 0.3306$; $\omega^2 = (1, 0, 0.5821, 0.2538)$, $p(\omega^2) = 0.1694$; $\omega^3 = (1, 1, 0.5420, 0.6261)$, $p(\omega^3) = 0.3306$; and $\omega^4 = (0, 1, 0.4180, 0.4180)$, $p(\omega^4) = 0.1694$; where $p(\omega)$ denotes the probability assigned to the point $\omega = (\xi_1, \xi_2, \eta_1, \eta_2)$.

Since the stochastic program above has complete recourse, using the lower bounding solution x_l in the upper bounding function $\phi_U(x)$ in (46), one obtains a less tight upper bound to Z^* as 4.1226.

Since (ξ_1, η_1) is independent of (ξ_2, η_2), the univariate bounding inequalities (27) and (28) may be used successively to determine tighter bounds than above. Under this pairwise independence, one obtains the bounds $3.6369 \leqslant Z^* \leqslant 3.7973$. Using the lower bounding solution, one computes a slightly weaker upper bound (under pairwise independence) as 3.8874.

5. **Concluding remarks.** This paper presents a general approach for bounding the expectation of a saddle function using limited moment information when random vectors have compact domains. The procedure formalizes concepts from Birge and Wets (1986) and Frauendorfer (1988d) in an attempt to formulate bounds as finite mathematical programs. Solutions of these mathematical programs provide discrete probability measures on the boundary of the joint domain of the random vectors to determine upper and lower approximations for the given stochastic program. Furthermore, these measures are shown to solve the respective moment problems, thus in the sense of Kemperman (1968), the derived bounds are tight. In fact, it follows that with using first and cross moment conditions, the *linearization* strategy is *optimal*, in the sense of the specific moment problem. This is also the case with the bounds due to Frauendorfer (1988b, c, d, 1989). However, if other joint moment conditions such as $E[\xi_i \xi_k]$ are used, then linearization may not be an optimal strategy and one needs to use higher order, such as quadratic, function approximations which is due to the nature of the moment problem of concern. In the latter case, even with a linearization strategy, our approach can be shown to yield valid bounds, although they may not generally solve the moment problem of interest.

It is easy to prove that the bounds $\phi_U(x)$ and $\phi_L(x)$ behave monotonously under partitioning the domain, an approach generally used to solve stochastic programs to ϵ-optimality, see Birge and Wallace (1986). In fact, these bounds converge to the optimal value of the stochastic program as long as the diameters of the partitioning cells (the cells being polytopes themselves) are sufficiently small, and furthermore, one can even claim convergence of approximate solutions to a true optimal solution, relying on epi-convergence arguments, see Birge and Wets (1986). Nevertheless, such assertions do not hold, in general, regarding the *relaxed* lower bound $\bar{\phi}_L(x)$, the exception being when Θ is a simplex. Despite this shortcoming, the convex polyhedral lower bound $\bar{\phi}_L(x)$ is numerically attractive because even with larger dimensions, the number of constraints grows only linearly in the dimension L of η since the linear program defining $\bar{\phi}_L(x)$ has only $L + 1$ blocks of constraints. The number of variables, however, depends on the chosen shape for domain Θ.

Towards solving stochastic programs to ϵ-optimality, what remains to be addressed is the choice of such partitioning schemes so that the computational burden does not become excessive; for instance, how to refine partitions based on information gained from previous partitions. Frauendorfer (1988c) addressed the issues when domains are rectangles. More recently, simplicial domains have been investigated by Frauendorfer (1991).

An extension of the results in this paper to the case when domains are possibly unbounded sets appears in Edirisinghe and Ziemba (1994).

Acknowledgements. The authors wish to thank the anonymous referees for their useful comments on earlier versions of the paper, which have improved the presentation of the paper significantly. This research was partially supported by the Natural Sciences and Engineering Research Council of Canada and the National Science Foundation. An earlier version of this paper was presented at the ORSA/TIMS meeting in Vancouver, May 1989 and at the Fifth International Conference on Stochastic Programming, University of Michigan, Ann Arbor, Michigan, August 1989.

References

Birge, J. R. and Wallace, S. W. (1986). Refining Bounds for Stochastic Linear Programs with Linearly Transformed Independent Random Variables. *Oper. Res. Lett.* **5** 73–77.

——— and Wets, R. J-B. (1986). Designing Approximation Schemes for Stochastic Optimization Problems, in Particular for Stochastic Programs with Recourse. *Math. Programming Stud.* **27** 54–102.

Dula, J. H. (1991). An Upper Bound on the Expectation of Simplicial Functions of Multivariate Random Variables. *Math. Programming.* **55** 69–80.

Edirisinghe, N. C. P. and Ziemba, W. T. (1994). Bounding the Expectation of a Saddle Function with Application to Stochastic Programming. *Math. Oper. Res.* **19** 314–340.

Frauendorfer, K. (1988a). SLP Recourse Problems with Arbitrary Multivariate Distributions—the Dependent Case. *Math. Oper. Res.* **13** 377–394.

——— (1988b). SLP Problems: Objective and Right-Hand Side being Stochastically Dependent—Part I. Proceedings of the 12th IMACS World Congress in Paris. (to appear).

——— (1988c). SLP Problems: Objective and Right-Hand Side being Stochastically Dependent—Part II. Manuscript. IOR, University of Zurich.

——— (1988d). Solving SLP Recourse Problems: The Case of Stochastic Technology Matrix, RHS and Objective. *Proceedings of the 13th IFIP Conference on System Modelling and Optimization.* Springer-Verlag, (to appear).

——— (1989). A Simplicial Approximation Scheme for Convex Two-Stage Stochastic Programming Problems. Presented at the *Fifth International Conference on Stochastic Programming.* University of Michigan, Ann Arbor, Michigan.

——— (1991). A Barycentric Approximation Scheme for Convex Stochastic Optimization Problems (Algorithmic Design and Computational Results). Presentation at *ORSA Applied Probability Conference*, Monterey, CA, January 9–11.

Gassmann, H. and Ziemba, W. T. (1986). A Tight Upper Bound for the Expectation of a Convex Function of a Multivariate Random Variable. *Math. Programming Stud.* **27** 39–53.

Glashoff, K. and Gustafson, S.-A. (1983). *Linear Optimization and Approximation.* Springer-Verlag, NY.

Kall, P. (1976). *Stochastic Linear Programming.* Springer-Verlag, Berlin.

——— (1987). Stochastic Programming with Recourse: Upper Bounds and Moment Problems—A Review. In *Advances in Mathematical Optimization and Related Topics* (J. Guddat, P. Kall, K. Lommatzsch, M. Vlach, and K. Zimmermann, Eds.). Akademie-Verlag, Belrin.

Kemperman, J. (1968). The General Moment Problem. A Geometric Approach. *Ann. Math. Statist.* **39** 93–112.

Wets, R. J.-B. (1974). Stochastic Programs with Fixed Recourse: The Equivalent Deterministic Problem. *SIAM Rev.* **16** 309–339.

——— (1982). Stochastic Programming: Solution Techniques and Approximation Schemes. In *Mathematical Programming: State-of-the-Art* (A. Bachem, M. Groetschel, B. Korte, Eds.) Springer-Verlag, Berlin, pp. 566–603.

N. C. P. Edirisinghe: Management Science Program, College of Business Administration, University of Tennessee, Knoxville, Tennessee 37996; e-mail: pa144918@utkvm1.bitnet

W. T. Ziemba: Faculty of Commerce and Business Administration, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Y8