Department of Applied Mathematics
Faculty of EEMCS

University of Twente
The Netherlands

P.O. Box 217
7500 AE Enschede
The Netherlands

Phone: +31-53-4893400
Fax: +31-53-4893114

Email: memo@math.utwente.nl
www.math.utwente.nl/publications

Memorandum No. 1785

On solving discrete optimization problems with multiple random elements under general regret functions

P.K. Mandal, D. Ghosh[1], and
S. Das[2]

December, 2005

ISSN 0169-2690

[1]P&QM Area, IIM Ahmedabad, Vastrapur, Ahmedabad 380015, INDIA
[2]QM&IS Area, IIM Bangalore, Bannerghatta Road, Bangalore 560076, INDIA

On Solving Discrete Optimization Problems with Multiple Random Elements Under General Regret Functions

Pranab K. Mandal
Dept. of Applied Mathematics, Faculty of EEMCS, University of Twente,
P.O.Box 217, 7500AE Enschede, The NETHERLANDS.

Diptesh Ghosh
P&QM Area, IIM Ahmedabad, Vastrapur, Ahmedabad 380015, INDIA.

Shubhabrata Das
QM&IS Area, IIM Bangalore, Bannerghatta Road, Bangalore 560076, INDIA.

Abstract

In this paper we attempt to find least risk solutions for stochastic discrete optimization problems (SDOP) with multiple random elements, where the feasibility of a solution does not depend on the particular values the random elements in the problem take. While the optimal solution, for a linear regret function, can be obtained by solving an auxiliary (non-stochastic) discrete optimization problem (DOP), the situation is complex under general regret. We characterize a finite number of solutions which will include the optimal solution. We establish through various examples that the results from Ghosh, Mandal and Das (2005) can be extended only partially for SDOPs with additional characteristics. We present a result where in selected cases, a complex SDOP may be decomposed into simpler ones facilitating the job of finding an optimal solution to the complex problem. We also propose numerical local search algorithms for obtaining an optimal solution.

AMS Subject Classification: 90C31
Key words: stochastic discrete optimization problems, min-sum, regret, risk

1 Introduction

In discrete optimization problems (DOPs), some of the problem parameters are often stochastic in nature. In these situations, the traditional notion of optimality (e.g. least cost solutions for minimization problems) does not remain unique. In Ghosh et al. [3], the authors considered min-sum stochastic discrete optimization problems (SDOP) having only one random element, and used a notion of least risk solution (corresponding to a regret function) to define an optimal solution. The notion of risk of a solution introduced in Ghosh et al. [3] remains valid irrespective of the number of random elements. In this work we consider SDOPs with multiple random elements and try to characterize the optimal solutions. Ghosh et al. [3] showed that a SDOP with one random element can be solved by solving a (deterministic) DOP obtained by replacing the random element in SDOP by an element with suitable fixed cost θ. In the auxiliary DOP the cost θ is the mean (median) of the random cost in the SDOP as long as the latter is symmetrically distributed. We show that this result continues to hold for SDOPs with multiple random elements when the regret function is taken to be linear. We also show through various examples that the one dimensional result cannot be extended to any reasonable generality for SDOPs with multiple random elements under general regret functions. Some partial results, however, have been obtained.

The remainder of this paper is organized as follows. Section 2 provides the preliminaries needed for this work. In Section 3 we treat SDOPs under general setup. We characterize a finite number of feasible solutions which will include the optimal solution to the SDOP. Section 4 deals with extension of one dimensional result of Ghosh et al. [3] to SDOPs with multiple random elements and contains some partial characterization for optimal solution for a special class called balanced SDOPs. In Section 5 we consider another special class of SDOPs which can be decomposed into simpler ones under suitable conditions. Obtaining an optimal solution numerically is taken up in Section 6. We consider a few heuristic algorithms and compare their performances through simulation. The article is concluded with a summary in Section 7.

2 Preliminaries

We shall describe in this section all the notations and definitions we would need in this article. Though we shall use the same notations introduced in Ghosh et al. [3], for completeness we mention them here as well. We also restate the main result of that paper.

Definition 1 A *discrete optimization problem (DOP)* is denoted by $\pi = (G, \mathbb{S}, z)$, where G is a finite ground set, with each element $e \in G$ having an associated value c_e (often referred to as the cost of e). The set, $\mathbb{S}$, of feasible solutions is a subset of the power set of G and is usually described by a set of rules that each $S \in \mathbb{S}$ must satisfy. The function $z : \mathbb{S} \to \Re$ is referred to as the objective function (or the cost function), and the optimization problem is one of finding a member of $\arg\min_{S \in \mathbb{S}}\{z(S)\}$.

Definition 2 An element $e \in G$ in π is called *random* (alternatively *fixed*) if the associated cost c_e is random valued (alternatively constant).

Definition 3 A *stochastic discrete optimization problem* (SDOP) is one in which the costs of some of the elements in G are random.

Definition 4 Given any fixed set of values for c_e's, the regret associated with a solution $S \in \mathbb{S}$ is defined by

$$\textit{regret}(S) = r(z(S) - Z^*),$$

where Z^* is the minimum possible value of the objective function for given values of c_e's (and hence is a function of these c_e's) and $r(\cdot)$ is an increasing continuous function on $[0, \infty)$, such that $r(0) = 0$.

Definition 5 For a given regret function $r(\cdot)$ which is increasing and continuous on $[0, \infty)$ with $r(0) = 0$, the *risk* associated with a solution $S \in \mathbb{S}$ is given by

$$R(S) = \mathbb{E}\ \textit{regret}(S) = \mathbb{E}\ r(z(S) - Z^*),$$

where Z^* is the cost of the least cost solution at specific values of the random elements, and hence is random itself. The expectation is taken with respect to the costs of the random elements.

Definition 6 An *optimal* solution (also referred to as a least risk solution) to a SDOP is defined as a feasible solution with minimum risk among all feasible solutions.

In addition we introduce the following definition of *balanced* DOP.

Definition 7 The set of feasible solutions $\mathbb{S}$ in a DOP $\pi = (G, \mathbb{S}, z)$ is said to be *balanced* (equivalently the DOP is called balanced) if

$$S(\subseteq G) \in \mathbb{S}, |S| = m \Rightarrow \tilde{S} \in \mathbb{S} \text{ for any } \tilde{S} \subseteq G \text{ with } |\tilde{S}| = m.$$

We restrict ourselves to SDOPs where all feasible solutions remain feasible, irrespective of the randomness involved. Further, the objective function is taken to be $z(S) = \sum_{e \in S} c_e$ and the probability distributions of the random elements are assumed to be known and unimodal.

We now state the main result of Ghosh et al. [3] for SDOP with one random element.

Theorem 1 *Consider a SDOP with one random element. Suppose the random cost has a (cumulative) distribution function $H(\cdot)$. Then a least risk solution to the SDOP under regret function $r(\cdot)$ can be obtained by solving a non-stochastic DOP obtained by replacing the random cost by a fixed one θ, which is the solution to $\Psi(t) = 0$ where*

$$\Psi(t) = \int_t^{\infty} r(x - t) dH(x) \; - \int_{-\infty}^{t} r(t - x) dH(x). \tag{1}$$

Further, if the random cost is symmetric with mean μ, then $\theta = \mu$.

Proof: See Theorem 1 and Theorem 2 in Ghosh et al. [3].

3 General Framework

In this section we consider a SDOP with r (more than one) random elements. We partition G into $G_R = \{e_1, \ldots, e_r\}$ of random elements, and $G_F = \{e_{r+1}, \ldots, e_{r+f}\}$ of fixed elements. Let $X_1, \ldots, X_r$ be the random variables denoting the values of $c_{e_1}, \ldots, c_{e_r}$ and $H(x_1, \ldots, x_r)$ denote $Pr(X_1 \le x_1, \ldots, X_r \le x_r)$. Note that the objective function value of any solution S can be represented as

$$z(S) = F(S) + \sum_{i: e_i \in S \cap G_R} X_i \tag{2}$$

where $F(S) = \sum_{e \in S \cap G_F} c_e$ is the fixed component of the cost $z(S)$.

If one works with a linear regret function then it turns out out that a SDOP can be solved by solving a (non-stochastic) DOP obtained by replacing the random elements in the former by (non-random) elements with costs equal to the means of their random counterpart.

Theorem 2 *Consider a SDOP π with r random elements having costs $X_1, \ldots, X_r$. Suppose X_i's are random variables having finite means μ_i, $i = 1, \ldots, r$, respectively. Consider the non-stochastic DOP π^* obtained from the SDOP π by replacing the random elements with (non-random) elements having costs $\mu_1, \ldots, \mu_r$, respectively. Then the least cost solution to the DOP π^* will be an optimal solution of the SDOP π in the least risk sense under any linear regret function of the form*

$$r(t) = \alpha + \beta t, \text{where } \beta > 0.$$

Proof: Under the linear regret, minimizing $R(S)$ is equivalent to minimizing $\mathbb{E}z(S)$ which by (2) reduces to minimizing

$$F(S) + \sum_{i:e_i \in S \cap G_R} \mu_i.$$

This is the same as finding the least cost solution of the DOP π^* of the theorem.

When the regret function is nonlinear, solving a SDOP becomes more involved. However, the search for an optimal solution can be reduced considerably as indicated in Theorem 3 below. To that end let $K_1, \ldots, K_{2^r}$ be the 2^r subsets of $K = \{1, \ldots, r\}$. For $i = 1, \ldots, 2^r$, let

$$\mathbb{S}_i = \{S : S \in \mathbb{S}; \quad e_j \in S \quad \forall j \in K_i; \quad e_j \notin S \quad \forall j \in K \setminus K_i, \} \tag{3}$$

constitute a partition of $\mathbb{S}$. In certain problem situations, some of the $\mathbb{S}_i$'s may be empty.

Lemma 1 *If $S^1, S^2 \in \mathbb{S}_i$, for some i, then $z(S^1) - z(S^2)$ is non-random.*

Proof: By construction (3), S^1 and S^2 have the same set of random elements and hence by (2), $z(S^1) - z(S^2) = F(S^1) - F(S^2)$ which is non-random.

In light of Lemma 1 it is easy to see that a least risk solution within $\mathbb{S}_i$ can be obtained by fixing the random costs to any fixed values.

Remark 1 For nonempty $\mathbb{S}_i$, $i = 1, \ldots, 2^r$, denote the least cost (risk) solution within $\mathbb{S}_i$ as S_i, i.e.,

$$z(S_i) = \min_{S \in \mathbb{S}_i} z(S).$$

We now introduce the sets $\{\mathcal{R}_i; 1 \le i \le 2^r\}$ in the r-dimensional Euclidean space ($\Re^r$) as follows. $\mathcal{R}_i = \emptyset$ if $\mathbb{S}_i = \emptyset$, otherwise

$$\mathcal{R}_i = \{(x_1, \ldots x_r) : S_i \text{ is a least cost solution at } (x_1, \ldots x_r)\}. \tag{4}$$

We define a partition of $\Re^r$ through $\{P_i; 1 \le i \le 2^r\}$ where

$$P_1 = \mathcal{R}_1, \quad \text{and } P_i = \mathcal{R}_i \setminus (\bigcup_{j<i} P_j) \quad i = 2, \ldots, 2^r. \tag{5}$$

Notice that for all $i = 1, \ldots, 2^r$, $P_i \subseteq \mathcal{R}_i$.

We are now in a position to prove the main result for SDOPs with multiple random elements.

Theorem 3 *Consider a SDOP with r random elements $e_1, e_2, \ldots, e_r$ having (random) costs $X_1, \ldots, X_r$. Then the least risk solution to the SDOP (under general regret function) will be one of the S_i's, as introduced in Remark 1, and their risks are given by*

$$R(S_i) = \sum_{j=1}^{2^r} \int_{P_j} r(z(S_i) - z(S_j)) dH(\cdot), \tag{6}$$

where P_j's are as defined in (5).

Proof: To prove that one of the S_i's is a least risk solution it suffices to show that

$$R(S) \geq \min_i \{R(S_i)\} \quad \text{for any } S \in \mathbb{S}. \tag{7}$$

To see (7), note that $\exists\ j$ such that $S \in \mathbb{S}_j$. It follows from Remark 1 and Lemma 1 that $z(S) - z(S_j)$ is nonnegative and non-random. Then the fact that $r(\cdot)$ is increasing leads to

$$\begin{aligned} R(S) &= \mathbb{E}\, r(z(S) - Z^*) \;=\; \mathbb{E}\, r(z(S) - z(S_j) + z(S_j) - Z^*) \\ &\geq \mathbb{E}\, r(z(S_j) - Z^*) \;=\; R(S_j) \;\geq\; \min_i \{R(S_i)\}. \end{aligned}$$

The expression (6) follows from the definition of risk and the fact that S_j is the least cost solution when the random cost is in P_j.

Remark 2 An alternative characterization of the partitions P_i's can be obtained as follows. From (2),

$$\begin{aligned} z(S_i) - z(S_j) &= F(S_i) + \sum_{m \in K_i} x_m - \Big[F(S_j) + \sum_{m \in K_j} x_m\Big] \\ &= \Big[\sum_{m \in K_i \setminus K_j} x_m - \sum_{m \in K_j \setminus K_i} x_m\Big] + F(S_i) - F(S_j). \end{aligned} \tag{8}$$

If S_i is a least cost solution at $(x_1, \ldots x_r)$, then for this set of costs, $z(S_i) \leq z(S_j), \quad \forall j = 1, \ldots, 2^r$. Hence, an alternative characterization of P_i is

$$P_i = \Big\{(x_1, \ldots, x_r) : \sum_{m \in K_i \setminus K_j} x_m - \sum_{m \in K_j \setminus K_i} x_m \leq F(S_j) - F(S_i), \quad j = 1, \ldots, 2^r\Big\} \tag{9}$$

4 Partial Characterization for Optimal Solution in Balanced DOP

We have seen in Theorem 2 that solving an SDOP under linear regret function reduces to solving a DOP which is the same as the SDOP except that the random costs were replaced by their respective means.

For nonlinear regret functions, on the other hand, replacing the random costs by their respective means does not always lead to the desired solution, even for SDOPs with one random element, as shown in Ghosh et al. [3]. One can, however, solve a SDOP with one single random element by solving an auxiliary (non-stochastic) DOP as stated in Theorem 1. Furthermore, if the probability distribution of the single random element is symmetric then this auxiliary DOP is obtained from the SDOP by replacing the random cost by its mean. It is then natural to explore if the result can be extended to SDOP with multiple random elements whose costs are symmetrically distributed. In an attempt to provide reasonably complete answer to such possible extensions, we now explore a series of examples and examine the outcomes from corresponding simulation exercises that lead to partial results and conjecture. In these examples, the r random elements are assumed to have costs $X_1, \ldots, X_r$ (with X_i having mean μ_i) as before, while the f fixed elements have costs $c_1 \leq \ldots \leq c_f$. We also use the notation $S^{j_1,\ldots,j_t}_{i_1,\ldots,i_l}$ to denote the least cost (risk) solution among the feasible solutions containing $X_{i_1},\ldots,X_{i_l}$, but not containing any of the $X_{j_1},\ldots,X_{j_t}$. Accordingly, we use notations $\mathcal{R}^{j_1,\ldots,j_t}_{i_1,\ldots,i_l}$ and $P^{j_1,\ldots,j_t}_{i_1,\ldots,i_l}$ for the sets introduced in (4) and (5).

In addition, in many of the examples, we consider a balanced SDOP (see Definition 7), where the minimum cardinality of a feasible solution is k. Note that, in that case, any optimal solution (in any sense) would have cardinality exactly k.

Example 1 Consider a balanced SDOP with $r = 2, f = 1 = k$. Suppose that X_1 has a symmetric triangular distribution on (0,1) with mode 0.5, while $X_2 = 1 - X_1$ and the fixed cost $c_1 \geq 0.5$. Thus the two random costs are strongly dependent but identically and symmetrically distributed random variables with mean 0.5. The three candidates for an optimal (least risk) solution S_1^2, S_2^1 and $S^{1,2}$ have costs X_1, X_2 and c_1 respectively. It can be checked that, with the regret function $r(t) = t^2$, the risks of the three solutions are as follows:

$$R(S_1^2) = R(S_2^1) = \frac{1}{12}; \quad \text{and } R(S^{1,2}) = \left(c_1 - \frac{1}{3}\right)^2 + \frac{1}{72};$$

and hence $S^{1,2}$ is the least risk solution as long as $c_1 \in (\frac{1}{2}, \frac{1}{3} + \sqrt{\frac{5}{72}})$, although it is not the least cost solution of the analogous non-stochastic DOP.

Example 2 Consider a balanced stochastic DOP with $r = f = k = 2$. The two random elements are assumed to have costs that are independently and identically distributed as symmetric triangular distribution on (0,1) with mode 0.5. From Theorem 3 there are four candidates for optimal solutions, viz. $S_{1,2}, S_1^2, S_2^1$ and $S^{1,2}$, and Table 6 reports the risks of these four candidate solutions, for some randomly chosen values of c_1, c_2 with the regret function being $r(t) = t^2$.

These examples show that for SDOP's with multiple random elements, an extension of the result for SDOP's with one random element is, in general, not true, i.e., it is not enough to replace the random costs by the respective average costs to arrive at the optimal solution irrespective of whether the symmetrically distributed random elements are

- dependent or independent of each other;
- identically distributed or otherwise.

The results from the simulation exercises, however, indicate possibility of partial results in the multidimensional case.

CONJECTURE *Consider a balanced stochastic DOP with r random elements, having costs $X_1,\ldots,X_r$ that are independent and symmetric random variables with finite means μ_i, $i = 1,\ldots,r$ respectively, and f fixed elements with costs $c_1 \leq c_2 \leq \ldots \leq c_f$. Then the least risk solution will consist only of the*

- *fixed elements (when feasible)*   *if* $\min_{1\leq i\leq r} \mu_i > c_f$,
- *random elements (when feasible)*   *if* $\max_{1\leq i\leq r} \mu_i < c_1$.

Remark 3 Without the 'balanced'-ness condition the conjecture need not be true. To see this, consider a SDOP with 3 elements e, e_1, e_2, the first of them being random while the other two are fixed. Let their costs be X (random), c_1 and c_2, respectively, such that $c_1 + c_2 > \mu(= E(X)) > c_2 > c_1$. Suppose $\mathcal{S} = \{\{e\}, \{e_1, e_2\}\}$. For symmetric X the least risk solution (see Theorem 1) is then $\{e\}$ which violates the conclusion of Conjecture 4. This SDOP is not balanced.

Remark 4 Only the following cases are relevant for the conjecture:

$$(a)\ r > f = k \quad (b)\ r = k > f \ \text{ and } \ (c)\ r = f = k.$$

To see this, first observe that we may assume, without loss of any generality, $f \leq k$. Further, if $r < k$, then the optimal solution would necessarily contain at least $(k - r)$ fixed elements with costs $c_1, \cdots, c_{k-r}$, and consequently the problem can be redefined in terms of a balanced DOP with $r = k > f$. Finally, if $r > k$ and $f < k$, then neither $S_{1,2,\ldots,r}$ nor $S^{1,2\ldots,r}$ is feasible.

We prove a version of the conjecture in the case of balanced SDOPs with two random elements, i.e, when $r = 2$ in the following subsection. The scope and validity of the conjecture for larger problems with the random elements having possibly different distributions has been examined through fairly extensive simulation study. For illustration, consider Table 7 which shows selected results from a SDOP with $r = f = k = 3$ with the three random costs having three different types of symmetric distributions, but are independent of each other. Table 8 reports selected results from a similar balanced SDOP with $r = f = k = 3$, where the random costs are independent and normally distributed with same variance, but means changing from simulation to simulation (as does the fixed costs). For additional results from simulation exercise, see Das et al. [2].

Special Case: $r = 2$

Consider a SDOP with two random elements with (random) costs X_1 and X_2. Suppose X_1 and X_2 are independent, and symmetrically distributed with means μ_1 and μ_2, respectively. Without loss of generality, we assume $\mu_1 \leq \mu_2$. Further, assume that the regret function $r(\cdot)$ satisfies the growth condition:

$$r(b + \beta) - r(b) \geq r(a + \alpha) - r(a) \quad \text{for } b \geq a \geq 0 \text{ and } \beta \geq \alpha \geq 0. \tag{10}$$

Remark 5 The growth condition (10) is satisfied by commonly used regret functions such as $r(t) = t^n$; $r(t) = (1 + t)^n - 1$ and $r(t) = \exp(\lambda t) - 1, \quad$ where $\lambda > 0$.

Remark 6 Though the following lemmas are stated for independent and symmetrically distributed random costs X_1 and X_2, the statements are true under more relaxed conditions, namely, under the assumption that

$$(U_1, U_2) \stackrel{D}{=} (-U_1, U_2) \stackrel{D}{=} (U_1, -U_2) \stackrel{D}{=} (-U_1, -U_2). \tag{11}$$

where $U_i = X_i - \mu_i$, $i = 1, 2$ and $\stackrel{D}{=}$ implies that the probability distributions are equal. In fact, we shall give the proofs under this relaxed condition.

Lemma 2 *Consider a balanced SDOP with $r = 2$, $f = 1$, and $k = 1$. If $(\mu_1, \mu_2) \in \mathcal{R}^{1,2}$, then $S^{1,2}$ is a least risk solution.*

Proof: Let the fixed element have cost c. Note that, in this case $\mathcal{R}_{1,2} = \emptyset$, $\mathcal{R}_1^2 = \{x_1 \leq \min(c, x_2)\}$, $\mathcal{R}_2^1 = \{x_2 \leq \min(c, x_1)\}$, and $\mathcal{R}^{1,2} = \{c \leq \min(x_1, x_2)\}$. Consequently,

$$(\mu_1, \mu_2) \in \mathcal{R}^{1,2} \qquad \Rightarrow \qquad \mu_2 \geq \mu_1 \geq c. \tag{12}$$

Further, the partitions of $\Re^2$, given in (5), can be taken as $P_{1,2}(x_1, x_2) = \emptyset$, $P_1^2(x_1, x_2) = \{x_1 \leq \min(c, x_2)\}$, $P_2^1(x_1, x_2) = \{x_2 \leq c, x_2 < x_1\}$, and $P^{1,2}(x_1, x_2) = \{c < \min(x_1, x_2)\}$.

Letting $u_i = x_i - \mu_i$, $i = 1, 2$, we can rewrite the partitions with respect to the (u_1, u_2)-space as $Q_{1,2}(u_1, u_2) = \emptyset$, $Q_1^2(u_1, u_2) = \{u_1 \leq \min(c - \mu_1, u_2 + \mu_2 - \mu_1)\}$, $Q_2^1(u_1, u_2) = \{u_2 \leq c - \mu_2, u_2 < u_1 + \mu_1 - \mu_2)\}$, and $Q^{1,2}(u_1, u_2) = \{u_1 > c - \mu_1, u_2 > c - \mu_2\}$.

From (6) we then have

$$\begin{aligned} R(S_1^2) &= E[r(z(S_1^2) - z^*)] = E[r(X_1 - X_2)\mathbb{I}_{P_2^1}(X_1, X_2)] + E[r(X_1 - c)\mathbb{I}_{P^{1,2}}(X_1, X_2)] \\ &= E[r(U_1 + \mu_1 - U_2 - \mu_2)\mathbb{I}_{Q_2^1(U_1,U_2)}] + E[r(U_1 + \mu_1 - c)\mathbb{I}_{Q^{1,2}(U_1,U_2)}] \end{aligned}$$

which, using (11), can be rewritten as

$$= E[r(U_1 + U_2 + \mu_1 - \mu_2)\mathbb{I}_{Q_2^1(U_1,-U_2)}] + E[r(U_1 + \mu_1 - c)\mathbb{I}_{Q^{1,2}(U_1,U_2)}].$$

Similarly

$$\begin{aligned} R(S_2^1) &= E[r(X_2 - X_1)\mathbb{I}_{P_1^2}(X_1, X_2)] + E[r(X_2 - c)\mathbb{I}_{P^{1,2}}(X_1, X_2)] \\ &= E[r(U_1 + U_2 + \mu_2 - \mu_1)\mathbb{I}_{Q_1^2(-U_1,U_2)}] + E[r(U_2 + \mu_2 - c)\mathbb{I}_{Q^{1,2}(U_1,U_2)}], \end{aligned}$$

and

$$\begin{aligned} R(S^{1,2}) &= E[r(c - X_1)\mathbb{I}_{P_1^2}(X_1, X_2)] + E[r(c - X_2)\mathbb{I}_{P_2^1}(X_1, X_2)] \\ &= E[r(U_1 + c - \mu_1)\mathbb{I}_{Q_1^2(-U_1,U_2)}] + E[r(U_2 + c - \mu_2)\mathbb{I}_{Q_2^1(U_1,-U_2)}]. \end{aligned}$$

We need to show that $R(S_1^2)$ and $R(S_2^1)$ both exceed $R(S^{1,2})$. In the remainder of the proof we show that $R(S_1^2) \geq R(S^{1,2})$. That $R(S_2^1) \geq R(S^{1,2})$ can be derived following very similar arguments.

In order to prove that $R(S_1^2) \geq R(S^{1,2})$ we proceed as follows. Using the above expressions for risks we write the difference $R(S_1^2) - R(S^{1,2})$ as sums of integrations (expectations). We identify the parts of integration sets $Q_{(\cdot)}^{(\cdot)}(\cdot)$ with positive and negative contributions to the difference. For each negative contribution we identify (unique) positive contribution so that the total contribution becomes nonnegative and thus making $R(S_1^2) - R(S^{1,2}) \geq 0$. To that end we decompose the sets of integrations $Q_{(\cdot)}^{(\cdot)}(\cdot)$ in terms of the following seven sets:

$$\begin{aligned} A &= \{c - \mu_1 < U_1 < \mu_1 - c, U_2 > c - \mu_2\}, \\ B &= \{U_1 \geq \mu_1 - c, U_2 > c - \mu_2\}, \\ C &= \{U_2 \geq \mu_2 - c, U_1 \geq U_2 - \mu_2 + \mu_1\} \\ C^{(2)} &= \{U_2 \leq c - \mu_2, U_1 \geq -U_2 - \mu_2 + \mu_1\} \\ D &= \{U_1 > \mu_1 - c, U_2 > U_1 - \mu_1 + \mu_2\}, \\ D^{(1)} &= \{U_1 < c - \mu_1, U_2 > -U_1 - \mu_1 + \mu_2\}, \text{ and} \\ E &= \{c - \mu_1 \leq U_1 \leq \mu_1 - c, U_2 \geq \mu_2 - c, U_1 + U_2 > -\mu_1 + \mu_2\}. \end{aligned}$$

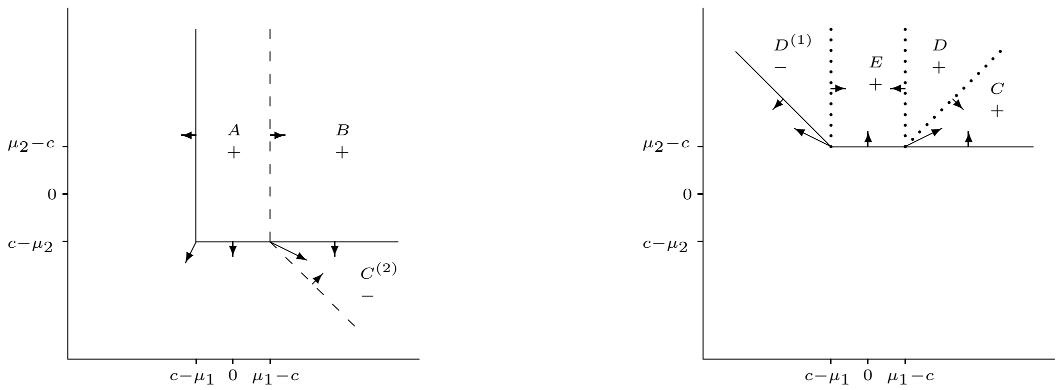

Figure 1: Decompositions of sets of integrations for $R(S_1^2) - R(S^{1,2})$. First plot is for sets involving $r(U_1+\cdots)$ and the second for the (common) set involving $r(U_2+\cdots)$ and $r(U_1+U_2+\cdots)$. Arrows indicate the (non)inclusion of boundary lines/points. Symbols '+'/'−' indicate the contribution from the integral over the corresponding set to the difference $R(S_1^2)-R(S^{1,2})$.

Then it is easy to see that (see Figure 1)

$$\begin{aligned} Q^{1,2}(U_1, U_2) &= A \cup B, \\ Q_1^2(-U_1, U_2) &= B \cup C^{(2)}, \text{ and} \\ Q_2^1(U_1, -U_2) &= D^{(1)} \cup E \cup D \cup C; \end{aligned}$$

and

$$\begin{aligned} &R(S_1^2) - R(S^{1,2}) \\ &\quad= \quad E[r(U_1 + \mu_1 - c)\mathbb{I}_A] \\ &\qquad + E[\{r(U_1 + \mu_1 - c) - r(U_1 + c - \mu_1)\}\mathbb{I}_B] \\ &\qquad - E[r(U_1 + c - \mu_1)\mathbb{I}_{C^{(2)}}] \\ &\qquad + E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\{\mathbb{I}_{D^{(1)}} + \mathbb{I}_E + \mathbb{I}_D + \mathbb{I}_C\}] \\ &\quad= \quad E[r(U_1 + \mu_1 - c)\mathbb{I}_A] \\ &\qquad + E[\{r(U_1 + \mu_1 - c) - r(U_1 + c - \mu_1)\}\mathbb{I}_B] \\ &\qquad + E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\mathbb{I}_C] \\ &\qquad - E[r(U_1 + c - \mu_1)\mathbb{I}_{C^{(2)}}] \\ &\qquad + E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\{\mathbb{I}_D + \mathbb{I}_{D^{(1)}}] \\ &\qquad + E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\mathbb{I}_E]. \end{aligned}$$

Now note that the region $C^{(2)}$ reflected along the u_1-axis (i.e., replacing U_2 by $-U_2$) is same as C and similarly, $D^{(1)}$ reflected along the u_2-axis (i.e., replacing U_1 by $-U_1$) is same as D. Using this observation, and (11) we can rewrite $R(S_1^2) - R(S^{1,2})$ as

$$\begin{aligned} &R(S_1^2) - R(S^{1,2}) \\ &\quad= \quad E[r(U_1 + \mu_1 - c)\mathbb{I}_A] \\ &\qquad + E[\{r(U_1 + \mu_1 - c) - r(U_1 + c - \mu_1)\}\mathbb{I}_B] \\ &\qquad + E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2) - r(U_1 + c - \mu_1)\}\mathbb{I}_C] \end{aligned}$$

$$+ E[\{r(U_1+U_2+\mu_1-\mu_2)+r(-U_1+U_2+\mu_1-\mu_2)-2r(U_2+c-\mu_2)\}\mathbb{I}_D]$$
$$+ E[\{r(U_1+U_2+\mu_1-\mu_2)-r(U_2+c-\mu_2)\}\mathbb{I}_E]. \tag{13}$$

In region A, $c-\mu_1 < U_1$, so that $U_1+\mu_1-c>0$. Therefore

$$E[r(U_1+\mu_1-c)\mathbb{I}_A] \geq 0. \tag{14}$$

Also, since $\mu_1 \geq c$ (see (12)), we have $U_1+\mu_1-c \geq U_1-\mu_1+c$, and subsequently $r(U_1+\mu_1-c) \geq r(U_1-\mu_1+c)$. Therefore

$$E[\{r(U_1+\mu_1-c)-r(U_1+c-\mu_1)\}\mathbb{I}_B] \geq 0. \tag{15}$$

Next, note that the expression $r(U_1+U_2+\mu_1-\mu_2)-r(U_2+c-\mu_2)-r(U_1+c-\mu_1)$ can be rewritten as

$$[r(b+\beta)-r(b)]-[r(a+\alpha)-r(a)]$$

where $a=0$, $\alpha=U_1+c-\mu_1$, $b=U_2+c-\mu_2$, and $\beta=U_1-c+\mu_1$. However, over C, $U_2 \geq \mu_2-c$, i.e., $b=U_2+c-\mu_2 \geq a=0$. In addition, since $U_1 \geq U_2-\mu_2+\mu_1$, $\alpha=U_1+c-\mu_1 \geq 0$. Also, $\beta=U_1-c+\mu_1=U_1+c-\mu_1+2(\mu_1-c) \geq U_1+c-\mu_1=\alpha$, since by (12), $\mu_1-c \geq 0$. Thus, using (10),

$$E[\{r(U_1+U_2+\mu_1-\mu_2)-r(U_2+c-\mu_2)-r(U_1+c-\mu_1)\}\mathbb{I}_C] \geq 0. \tag{16}$$

Similarly, $r(U_1+U_2+\mu_1-\mu_2)+r(-U_1+U_2+\mu_1-\mu_2)-2r(U_2+c-\mu_2)$ can be rewritten as

$$[r(b+\beta)-r(b)]-[r(a+\alpha)-r(a)]$$

where $a=-U_1+U_2+\mu_1-\mu_2$, $b=U_2+c-\mu_2$, $\alpha=U_1+c-\mu_1$, and $\beta=U_1-c+\mu_1$. Using arguments similar to the one above, it can be shown that over region D, $b \geq a \geq 0$ and $\beta \geq \alpha \geq 0$. Therefore, using (10),

$$E[\{r(U_1+U_2+\mu_1-\mu_2)+r(-U_1+U_2+\mu_1-\mu_2)-2r(U_2+c-\mu_2)\}\mathbb{I}_D] \geq 0. \tag{17}$$

Finally, since $r(\cdot)$ is an increasing function, and since in region E, $U_1 \geq c-\mu_1$, therefore $r(U_1+U_2+\mu_1-\mu_2)=r((U_2+c-\mu_2)+(U_1+\mu_1-c)) \geq r(U_2+c-\mu_2)$, which leads to the result

$$E[\{r(U_1+U_2+\mu_1-\mu_2)-r(U_2+c-\mu_2)\}\mathbb{I}_E] \geq 0. \tag{18}$$

Using results (14) through (18) in the risk expression (13), we see that $R(S_1^2) \geq R(S^{1,2})$.

It is worthwhile to observe that when proving $R(S_2^1) \geq R(S^{1,2})$ one may need to decompose the sets of integrations $Q_{(\cdot)}^{(\cdot)}(\cdot)$ differently from the one shown above. Decomposition is to be done according to the procedure mentioned at the beginning of proof of $R(S_1^2) \geq R(S^{1,2})$.

Lemma 3 *Consider a balanced SDOP with $r=2$, $f=1$, and $k=2$. If $(\mu_1,\mu_2) \in R_{1,2}$, then $S_{1,2}$ is a least risk solution.*

Proof: The proof can be obtained using the steps outlined in the proof of Lemma 2. Note that here, $\mathcal{R}^{1,2}=\emptyset$.

Lemma 4 *Consider a balanced SDOP with $r = 2$, $f = 2$, and $k = 2$.*

(i) If $(\mu_1, \mu_2) \in \mathcal{R}_{1,2}$, then $S_{1,2}$ is a least risk solution.

(ii) If $(\mu_1, \mu_2) \in \mathcal{R}^{1,2}$, then $S^{1,2}$ is a least risk solution.

Proof: We will prove statement (i) only, since the proof of statement (ii) is similar.

The proof of statement (i) follows in the general lines of the proof of Lemma 2. Let the fixed elements have costs c_1 and c_2, and assume without loss of generality that $c_1 \leq c_2$.

Note that, in this case $\mathcal{R}^{1,2} = \{x_1 \geq c_2,\ x_2 \geq c_2\}$, $\mathcal{R}^1_2 = \{x_1 \geq c_1,\ x_2 \leq \min(c_2,\ x_1)\}$, $\mathcal{R}^2_1 = \{x_1 \leq c_2,\ x_2 \geq \max(c_1,\ x_1)\}$, and $\mathcal{R}_{1,2} = \{x_1 \leq c_1,\ x_2 \leq c_1\}$. Hence,

$$(\mu_1, \mu_2) \in \mathcal{R}^{1,2} \qquad \Rightarrow \qquad c_1 \leq c_2 \leq \mu_1 \leq \mu_2. \tag{19}$$

One corresponding partition of $\Re^2$ can be taken to be $P^{1,2} = \{x_1 \geq c_2,\ x_2 \geq c_2\}$, $P^1_2 = \{x_1 \geq c_1,\ x_2 < \min(c_2,\ x_1)\}$, $P^2_1 = \{x_1 < c_2,\ x_2 \geq \max(c_1,\ x_1)\}$, and $P_{1,2} = \{x_1 < c_1,\ x_2 < c_1\}$.

As in Lemma 2, we rewrite the partitions with respect to the (u_1, u_2)-space, where $u_i = x_i - \mu_i,\ i = 1, 2$ as $Q^{1,2}(u_1, u_2) = \{u1 \geq c_2 - \mu_1, u2 \geq c_2 - \mu_2\}$, $Q^1_2(u_1, u_2) = \{u_1 \geq c_1 - \mu_1, u_1 > u_2 - \mu_1 + \mu_2, u_2 < c_2 - \mu_2\}$, $Q^2_1(u_1, u_2) = \{u_2 \geq c_1 - \mu_2, u_2 \geq u_1 + \mu_1 - \mu_2, u_1 < c_2 - \mu_1)\}$, and $Q_{1,2}(u_1, u_2) = \{u_1 < c_1 - \mu_1, u_2 < c_1 - \mu_2\}$.

We need to show that $R(S^{1,2}) \leq \min\left(R(S^1_2), R(S^2_1), R(S_{1,2})\right)$. In this proof we only show that $R(S^{1,2}) \leq R(S^1_2)$. That $R(S^{1,2}) \leq R(S^2_1)$ and $R(S^{1,2}) \leq R(S_{1,2})$ can be shown using similar arguments.

From (6) we have

$$\begin{aligned} R(S^{1,2}) &= E[r(z(S^{1,2}) - z^*)] \\ &= E[r(c_2 - X_2)\mathbb{I}_{P^1_2}] + E[r(c_2 - X_1)\mathbb{I}_{P^2_1}] + E[r(c_1 + c_2 - X_1 - X_2)\mathbb{I}_{P_{1,2}}] \\ &= E[r(c_2 - \mu_2 - U_2)\mathbb{I}_{Q^1_2(U_1,U_2)}] + E[r(c_2 - \mu_1 - U_1)\mathbb{I}_{Q^2_1(U_1,U_2)}] \\ &\qquad + E[r(c_1 + c_2 - \mu_1 - \mu_2 - U_1 - U_2)\mathbb{I}_{Q_{1,2}(U_1,U_2)}] \end{aligned}$$

which, using (11), can be rewritten as

$$\begin{aligned} R(S^{1,2}) \quad &= \quad E[r(U_2 + c_2 - \mu_2)\mathbb{I}_{Q^1_2(U_1,-U_2)}] + E[r(U_1 + c_2 - \mu_1)\mathbb{I}_{Q^2_1(-U_1,U_2)}] \\ &\qquad + E[r(U_1 + U_2 + c_1 + c_2 - \mu_1 - \mu_2)\mathbb{I}_{Q_{1,2}(-U_1,-U_2)}] \end{aligned}$$

Similarly,

$$\begin{aligned} R(S^1_2) \quad &= \quad E[r(X_2 - c_2)\mathbb{I}_{P^{1,2}}] + E[r(X_2 - X_1)\mathbb{I}_{P^2_1}] + E[r(c_1 - X_1)\mathbb{I}_{P_{1,2}}] \\ &= \quad E[r(U_2 + \mu_2 - c_2)\mathbb{I}_{Q^{1,2}(U_1,U_2)}] + E[r(U_2 + U_1 + \mu_2 - \mu_1)\mathbb{I}_{Q^2_1(-U_1,U_2)}] \\ &\qquad + E[r(c_1 + U_1 - \mu_1)\mathbb{I}_{Q_{1,2}(-U_1,-U_2)}] \end{aligned}$$

Next, we decompose the sets of integrations $Q^{(\cdot)}_{(\cdot)}(\cdot)$ in terms of the following sets:

$$\begin{aligned} A \quad &= \quad \{U_1 > \mu_1 - c_1, U_2 > \mu_2 - c_1\}, \\ B \quad &= \quad \{\mu_1 - c_2 < U_1 \leq \mu_1 - c_1, U_2 > U_1 - \mu_1 + \mu_2\}, \\ B^{(1)} \quad &= \quad \{c_1 - \mu_1 \leq U_1 < c_2 - \mu_1, U_2 > -U_1 - \mu_1 + \mu_2\}, \end{aligned}$$

$$
\begin{aligned}
C &= \{\mu_2 - c_2 < U_2 \leq \mu_2 - c_1, U_1 \geq U_2 + \mu_1 - \mu_2\}, \\
C^{(1)} &= \{c_1 - \mu_2 \leq U_2 < c_2 - \mu_2, U_1 \geq -U_2 + \mu_1 - \mu_2\}, \\
D &= \{U_1 > \mu_1 - c_2, c_2 - \mu_2 \leq U_2 \leq \mu_2 - c_2\}, \\
E &= \{U_1 \geq c_2 - \mu_1, U_2 > \mu_2 - c_2\}, \text{ and} \\
F &= \{U_1 \geq c_2 - \mu_1, c_2 - \mu_2 \leq U_2 \leq \mu_2 - c_2\},
\end{aligned}
$$

and rewrite the risks as

$$
\begin{aligned}
R(S^{1,2}) = \; & E[r(U_1 + c_2 - \mu_1)\{\mathbb{I}_A + \mathbb{I}_B + \mathbb{I}_C + \mathbb{I}_{C^{(1)}} + \mathbb{I}_D\}] \\
& + E[r(U_2 + c_2 - \mu_2)\{\mathbb{I}_{B^{(1)}} + \mathbb{I}_E\}] \\
& + E[r(U_1 + U_2 + c_1 + c_2 - \mu_1 - \mu_2)\mathbb{I}_A]
\end{aligned}
$$

and

$$
\begin{aligned}
R(S_2^1) = \; & E[r(U_1 + c_1 - \mu_1)\mathbb{I}_A] + E[r(U_2 + \mu_2 - c_2)\{\mathbb{I}_E + \mathbb{I}_F\}] \\
& + E[r(U_1 + U_2 + \mu_2 - \mu_1)\{\mathbb{I}_A + \mathbb{I}_B + \mathbb{I}_C + \mathbb{I}_{C^{(1)}} + \mathbb{I}_D\}].
\end{aligned}
$$

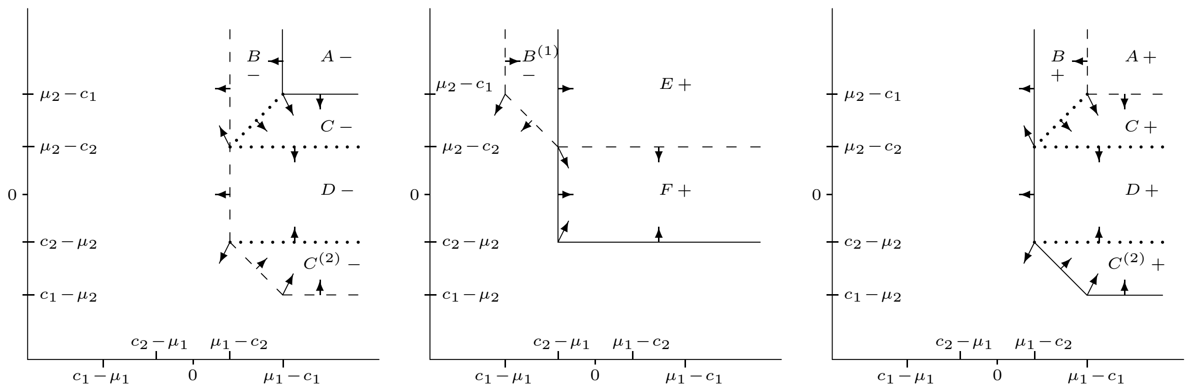

Figure 2: Decompositions of sets of integrations for $R(S_2^1)$ (solid line) and $R(S^{1,2})$ (dashed line). First plot is for sets involving $r(U_1 + \cdots)$, second for sets involving $r(U_2 + \cdots)$ and third for sets involving $r(U_1 + U_2 + \cdots)$. Arrows indicate the (non)inclusion of boundary lines/points. Symbols '+'/'−' indicate the contribution from the integral over the corresponding set to the difference $R(S_2^1) - R(S^{1,2})$.

Then

$$
\begin{aligned}
& R(S_2^1) - R(S^{1,2}) \\
= \; & E[(r(U_1 + c_1 - \mu_1) + r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) \\
& \qquad - r(U_1 + U_2 + c_1 + c_2 - \mu_1 - \mu_2))\mathbb{I}_A] \\
& + E[(r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1))\mathbb{I}_B] \\
& - E[r(U_2 + c_2 - \mu_2)\mathbb{I}_{B^{(1)}}] \\
& + E[(r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1))\mathbb{I}_C] \\
& + E[(r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1))\mathbb{I}_{C^{(1)}}] \\
& + E[(r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1))\mathbb{I}_D]
\end{aligned}
$$

$$+ E[(r(U_2 + \mu_2 - c_2) - r(U_2 + c_2 - \mu_2))\mathbb{I}_E]$$
$$+ E[r(U_2 + \mu_2 - c_2)\mathbb{I}_F].$$

Note that $B^{(1)}$ is a reflection of B along the u_2 axis, (i.e., replacing U_1 by $-U_1$) while $C^{(1)}$ is a reflection of C along the u_1 axis (i.e., replacing U_2 by $-U_2$).Then using (11) the expression above can be rewritten as

$$\begin{aligned}
& R(S_2^1) - R(S^{1,2}) \\
= \quad & E[(r(U_1 + c_1 - \mu_1) + r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) \\
& \qquad - r(U_1 + U_2 + c_1 + c_2 - \mu_1 - \mu_2))\mathbb{I}_A] \\
& + E[(r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) - r(U_2 + c_2 - \mu_2))\mathbb{I}_B] \\
& + E[(r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) \\
& \qquad + r(U_1 - U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1))\mathbb{I}_C] \\
& + E[(r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1))\mathbb{I}_D] \\
& + E[(r(U_2 + \mu_2 - c_2) - r(U_2 + c_2 - \mu_2))\mathbb{I}_E] \\
& + E[r(U_2 + \mu_2 - c_2)\mathbb{I}_F]. \qquad (20)
\end{aligned}$$

Now, note that the expression $r(U_1 + c_1 - \mu_1) + r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) - r(U_1 + U_2 + c_1 + c_2 - \mu_1 - \mu_2)$ can be rewritten as

$$[r(b + \beta) - r(b)] - [r(a + \alpha) - r(a)],$$

where $a = U_1 + c_1 - \mu_1$, $\alpha = c_2 - c_1$, $b = U_1 + U_2 + c_1 + c_2 - \mu_1 - \mu_2$, and $\beta = 2\mu_2 - c_1 - c_2$. It is easy to see that $a > 0$ over region A where $U_1 > \mu_1 - c_1$. Also, from (19) and the fact that $U_2 > \mu_2 - c_1$ over A, we have $b - a = U_2 + c_2 - \mu_2 \geq U_2 + c_1 - \mu_2 \geq 0$. That $\alpha \geq 0$ and $\beta - \alpha = 2(\mu_2 - c_2) \geq 0$ follow from (19). Hence, by condition (10), $r(U_1 + c_1 - \mu_1) + r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) - r(U_1 + U_2 + c_1 + c_2 - \mu_1 - \mu_2) \geq 0$ over A, so that

$$E[(r(U_1 + c_1 - \mu_1) - r(U_1 + c_2 - \mu_1) + r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + U_2 + c_1 + c_2 - \mu_1 - \mu_2))\mathbb{I}_A] \geq 0. \qquad (21)$$

Next, consider the expression $r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) - r(U_2 + c_2 - \mu_2)$. This can be rewritten as

$$[r(b + \beta) - r(b)] - [r(a + \alpha) - r(a)],$$

where $a = 0$, $\alpha = U_2 + c_2 - \mu_2$, $b = U_1 + c_2 - \mu_1$, and $\beta = U_2 - c_2 + \mu_2$. It is easy to check that over B, $b \geq a = 0$ and $\beta \geq \alpha \geq 0$. Then by (10) we have

$$E[(r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) - r(U_2 + c_2 - \mu_2))\mathbb{I}_B] \geq 0. \qquad (22)$$

In a similar fashion we can rewrite the expression $r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) + r(U_1 - U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1)$ as

$$[r(b + \beta) - r(b)] - [r(a + \alpha) - r(a)],$$

where $a = U_1 - U_2 - \mu_1 + \mu_2$, $\alpha = U_2 + c_2 - \mu_2$, $b = U_1 + c_2 - \mu_1$, and $\beta = U_2 - c_2 + \mu_2$ and show that over C, $b \geq a \geq 0$ and $\beta \geq \alpha \geq 0$, so that

$$E[(r(U_1 + U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1) + r(U_1 - U_2 + \mu_2 - \mu_1) - r(U_1 + c_2 - \mu_1))\mathbb{I}_C] \geq 0. \qquad (23)$$

Note that $(U_1 + U_2 - \mu_1 + \mu_2) - (U_1 + c_2 - \mu_1) = U_2 - c_2 + \mu_2 \geq 0$ over D. Since $r(\cdot)$ is a non-decreasing function, therefore

$$E[(r(U_1 + U_2 - \mu_1 + \mu_2) - r(U_1 + c_2 - \mu_1))\mathbb{I}_D] \geq 0. \quad (24)$$

Further, by assumption (19) we have $U_2 - c_2 + \mu_2 \geq U_2 + c_2 - \mu_2$, which implies that

$$E[(r(U_2 - c_2 + \mu_2) - r(U_2 + c_2 - \mu_2))\mathbb{I}_E] \geq 0. \quad (25)$$

Finally, since $U_2 \geq c_2 - \mu_2$ over F,

$$E[r(U_2 - c_2 + \mu_2)\mathbb{I}_F] \geq 0. \quad (26)$$

Expressions (21) through (26) show that the risk difference as expressed in (20) is non-negative.

In light of Remark 4, Lemmas 2, 3, and 4 lead to the following version of the conjecture with $r = 2$.

Theorem 4 *Consider a 2-dimensional balanced SDOP. Suppose the random costs X_1, and X_2 are independent and symmetrically distributed with mean μ_1 and μ_2, respectively. Further assume that the regret function satisfies the growth condition (10). Then the following holds:*

(i) if $(\mu_1, \mu_2) \in \mathcal{R}^{1,2}$ then $S^{1,2}$ is a least risk solution.

(ii) if $(\mu_1, \mu_2) \in \mathcal{R}_{1,2}$ then $S_{1,2}$ is a least risk solution.

Remark 7 The decompositions of the sets of integrations in Lemmas 2 through 4 do not follow any obvious rules; hence the theorem cannot be extended elegantly to SDOPs with more random elements.

Remark 8 The reverse statements of Theorem 4 are not true, that is, $(\mu_1, \mu_2) \in P_1^2$ (respectively P_2^1) does not necessarily imply that S_1^2 (respectively S_2^1) is a least risk solution. Also, the theorem does not hold true if the two random elements are dependent on each other. These can be validated through computer simulations (see Das et al. [2]).

5 Decomposition into smaller DOPs

In this section we look at decomposing a complex SDOP having multiple random elements (which are independent and symmetrically distributed random variable) into several simpler ones, each having one random element. To illustrate, consider a problem in which one needs to traverse all nodes in a graph in a certain (pre-specified) order. Between each pair of nodes in the graph, there exists a pair of edges. One of the edges in the pair is fixed, and the other is random. An optimal solution has to be decided before the journey begins. The following theorem says that if the random costs are independent of each other and symmetrically distributed then the solution incorporating the better among each pair of edges is optimal.

Theorem 5 *Consider a SDOP with n random elements having costs X_i's and n fixed elements with costs c_i's, for $i = 1, \ldots, n$. Suppose that a solution is feasible if and only if it contains one (and only one) element from each of the pairs having costs $(X_1, c_1), \ldots, (X_n, c_n)$. If the X_i's are independent random variables having symmetric distributions with mean μ_i's, then a feasible solution containing only those X_i's satisfying $\mu_i < c_i$ is a least risk solution.*

Remark 9 The theorem holds in a more general setup when $U_i := X_i - \mu_i,\ i = 1, 2, \cdots n$ satisfy for any $i_j = 0, 1,\ j = 1, 2, \ldots, n$,

$$((-1)^{i_1}U_1, (-1)^{i_2}U_2, \ldots, (-1)^{i_n}U_n) \stackrel{D}{=} (U_1, U_2, \ldots, U_n) \tag{27}$$

which is true when X_i's are independent and symmetrically distributed with mean μ_i.

Remark 10 The assumption about symmetrical distribution is critical for the validity of Theorem 5. First of all without symmetry the surrogate parameters are no longer (necessarily) the means μ_i's. Perhaps more critically, as the following example shows, without the symmetry assumption, one may not get an optimal solution by combining the optimal solutions of the (independent) one-dimensional stochastic DOP's (i.e., having single random element). While this may appear to be counter-intuitive, this is a natural manifestation of nonlinear regret function. To illustrate this take $n = 2$ and consider

$$\pi = (G = \{X_1, c_1, X_2, c_2\}, \mathcal{S} = \{\{X_1, X_2\}, \{X_1, c_2\}, \{c_1, X_2\}, \{c_1, c_2\}\}, z)$$

where $X_1 \sim$ Beta$(1, 1.5)$; $c_1 = 0.401$; $X_2 \sim$ Beta$(1, 2)$; $c_2 = 0.35$. Now consider the two sub-DOP's each with single random element, viz.,

$$\pi_i = (G_i = \{X_i, c_i\}, \mathcal{S}_i = \{\{X_i\}, \{c_i\}\}, z),\ i = 1, 2;\ r(t) = t^2.$$

Then

$$R_1(c_1) = 0.0306,\ R_1(X_1) = 0.0380,\ R_1(c_2) = 0.0261,\ R_1(X_2) = 0.0298,$$

i.e., c_1 is better than X_1 and c_2 is better than X_2 but for SDOP π $\{c_1, X_2\}$ is optimal, since

$$R_2(c_1, c_2) = 0.0809,\ R_2(c_1, X_2) = 0.0808,\ R_2(X_1, c_2) = 0.0881,\ R_2(X_1, X_2) = 0.0881.$$

Proof of Theorem 5: We prove the theorem under the general setup of Remark 9, i.e., under (27). Note that the set of feasible solutions $\mathcal{S} = \{S_I : I \subseteq \{1, 2, \ldots, n\} =: \mathbb{N}\}$, where $X_i \in S_I$, if $i \in I$, otherwise $c_i \in S_I$. Then $z_I := z(S_I) = X^I + c^{\bar{I}}$, where $\bar{I} = \mathbb{N} \setminus I$, the complement of I, and we use a^I to denote $\sum_{i \in I} a_i$, if $I \neq \phi$ and 0, otherwise.

Then defining $d_i = c_i - \mu_i$ we have

$$\begin{aligned}
\mathcal{R}_I &:= \{z_I = \min_{J \subset \mathbb{N}} z_J\} = \cap_J\{z_I \le z_J\} = \cap_J\{X^I + c^{\bar{I}} \le X^J + c^{\bar{J}}\} \\
&= \cap_J\{X^I - X^J \le c^{\bar{J}} - c^{\bar{I}} = c^I - c^J\} = \cap_J\{U^I - U^J \le d^I - d^J\} \\
&= \cap_J\{U^{I\setminus J} - U^{J\setminus I} \le d^{I\setminus J} - d^{J\setminus I}\} = \{U_i \le d_i, i \in I \text{ and } U_i \ge d_i, i \notin I\}.
\end{aligned}$$

One set of corresponding partition can be then taken to be

$$P_I = \{U_i < d_i, i \in I \text{ and } U_i \ge d_i, i \notin I\}, \quad I \subset \mathbb{N}.$$

Then from (27), we have

$$\begin{aligned}
R_I &:= R(S_I) = \sum_{J \neq I} E[r(z_I - z_J)\mathbb{I}_{P_J}] = \sum_{J \neq I} E[r(U^{I\setminus J} - U^{J\setminus I} - (d^{I\setminus J} - d^{J\setminus I}))\mathbb{I}_{P_J}] \\
&= \sum_{J \neq I} E[r(U^{I\setminus J} + U^{J\setminus I} - (d^{I\setminus J} - d^{J\setminus I}))\mathbb{I}_{Q_J}] = \sum_{J \neq I} E[r(U^{I\Delta J} - (d^{I\setminus J} - d^{J\setminus I}))\mathbb{I}_{Q_J}] \\
&= \sum_{\substack{K \subset \mathbb{N} \\ K \neq \phi}} E\left[r\left(U^K - \left(d^{KI} - d^{K\bar{I}}\right)\right)\mathbb{I}_{Q_{I\Delta K}}\right],
\end{aligned}$$

where

$$Q_I = \{U_i > -d_i, i \in I \text{ and } U_i \geq d_i, i \notin I\}, \quad I \subset \mathbb{N}.$$

Now suppose that $\mu_i < c_i, i \in M$ and $\mu_i \geq c_i, i \notin M$, i.e., $d_i > 0, i \in M$ and $d_i \leq 0, i \notin M$. For the theorem we need to show that $R(S_M) \leq R(S_I)$ for all $I \subset \mathbb{N}$. Note that for any $I \subset \mathbb{N}$,

$$\begin{aligned}
R(S_I) - R(S_M) &= \sum_{\substack{K \subset \mathbb{N} \\ K \neq \phi}} E\left[r\left(U^K - d^{KI} + d^{K\bar{I}}\right)\mathbb{I}_{Q_{I\Delta K}} - r\left(U^K - d^{KM} + d^{K\bar{M}}\right)\mathbb{I}_{Q_{M\Delta K}}\right] \\
&= \sum_{\substack{K \subset \mathbb{N} \\ K \neq \phi}} E\left[\left\{r\left(U^K - d^{KI} + d^{K\bar{I}}\right) - r\left(U^K - d^{KM} + d^{K\bar{M}}\right)\right\}\mathbb{I}_{Q_{I\Delta K} \cap Q_{M\Delta K}}\right] \\
&\quad + \sum_{\substack{K \subset \mathbb{N} \\ K \neq \phi}} E\left[r\left(U^K - d^{KI} + d^{K\bar{I}}\right)\mathbb{I}_{Q_{I\Delta K} \setminus Q_{M\Delta K}}\right] \\
&\quad - \sum_{\substack{K \subset \mathbb{N} \\ K \neq \phi}} E\left[r\left(U^K - d^{KM} + d^{K\bar{M}}\right)\mathbb{I}_{Q_{M\Delta K} \setminus Q_{I\Delta K}}\right]
\end{aligned}$$

Note that

$$\begin{aligned}
(U^K - d^{KI} + d^{K\bar{I}}) - (U^K - d^{KM} + d^{K\bar{M}}) &= d^{K\bar{I}} - d^{KI} + d^{KM} - d^{K\bar{M}} \\
&= d^{K\bar{I}M} + d^{K\bar{I}\bar{M}} - d^{KIM} - d^{KI\bar{M}} + d^{KIM} + d^{K\bar{I}M} - d^{KI\bar{M}} - d^{K\bar{I}\bar{M}} \\
&= 2(d^{K\bar{I}M} - d^{KI\bar{M}}) \geq 0, \text{ since } d_i > 0, i \in M, \text{ and } d_i \leq 0, i \notin M.
\end{aligned}$$

Hence the first sum in $R(S_I) - R(S_M)$ is nonnegative. Now let us look at the sets $Q_{I\Delta K} \setminus Q_{M\Delta K}$ and $Q_{M\Delta K} \setminus Q_{I\Delta K}$. Below we describe the structure of these sets.

for	$Q_{I\Delta K}$	$Q_{M\Delta K}$	$Q_{I\Delta K} \setminus Q_{M\Delta K}$	$Q_{M\Delta K} \setminus Q_{I\Delta K}$
$i \in KMI$	$d_i \leq u_i$	$d_i \leq u_i$	$d_i \leq u_i$	$d_i \leq u_i$
$i \in KM\bar{I}$	$-d_i < u_i$	$d_i \leq u_i$	"$-d_i < u_i < d_i$"	$d_i \leq u_i$
$i \in K\bar{M}I$	$d_i \leq u_i$	$-d_i < u_i$	"$d_i \leq u_i \leq -d_i$"	$-d_i < u_i$
$i \in K\bar{M}\bar{I}$	$-d_i < u_i$	$-d_i < u_i$	$-d_i < u_i$	$-d_i < u_i$
$i \in \bar{K}MI$	$-d_i < u_i$	$-d_i < u_i$	$-d_i < u_i$	$-d_i < u_i$
$i \in \bar{K}M\bar{I}$	$d_i \leq u_i$	$-d_i < u_i$	$d_i \leq u_i$	"$-d_i < u_i < d_i$"
$i \in \bar{K}\bar{M}I$	$-d_i < u_i$	$d_i \leq u_i$	$-d_i < u_i$	"$d_i \leq u_i \leq -d_i$"
$i \in \bar{K}\bar{M}\bar{I}$	$d_i \leq u_i$	$d_i \leq u_i$	$d_i \leq u_i$	$d_i \leq u_i$

In the table if a region is within quotation marks, this implies that some (but not necessarily all) of these type of boundary conditions would appear in the corresponding set. The exact construction is as follows. For nonempty K^*, $K \subset \mathbb{N}$,

$$Q_{I\Delta K^*} \setminus Q_{M\Delta K^*} = \bigcup_{J^* \subset K^*M\bar{I}} \bigcup_{L^* \subset K^*\bar{M}I} \left\{\begin{array}{cl} d_i \leq u_i & i \in K^*MI \cup \bar{K}^*\bar{M}\bar{I} \\ -d_i < u_i & i \in K^*\bar{M}\bar{I} \cup \bar{K}^*MI \\ -d_i < u_i < d_i & i \in J^* \\ d_i \leq u_i & i \in \bar{J}^*K^*M\bar{I} \cup \bar{K}^*M\bar{I} \\ d_i \leq u_i \leq -d_i & i \in L^* \\ -d_i < u_i & i \in \bar{L}^*K^*\bar{M}I \cup \bar{K}^*\bar{M}I \end{array}\right\}$$

$$= \bigcup_{J^*\subset K^*M\bar{I}} \bigcup_{L^*\subset K^*\bar{M}I} Q(I\setminus M, K^*, J^*, L^*), \text{ say,}$$

$$Q_{M\Delta K}\setminus Q_{I\Delta K} = \bigcup_{J\subset \bar{K}M\bar{I}} \bigcup_{L\subset \bar{K}\bar{M}I} \left\{ \begin{array}{cl} d_i \le u_i & i\in KMI\cup\bar{K}\bar{M}\bar{I} \\ -d_i < u_i & i\in K\bar{M}\bar{I}\cup\bar{K}MI \\ -d_i < u_i < d_i & i\in J \\ d_i\le u_i & i\in \bar{J}\bar{K}M\bar{I}\cup KM\bar{I} \\ d_i\le u_i\le -d_i & i\in L \\ -d_i<u_i & i\in \bar{L}\bar{K}\bar{M}I\cup K\bar{M}I \end{array}\right\}$$

$$= \bigcup_{J\subset \bar{K}M\bar{I}} \bigcup_{L\subset \bar{K}\bar{M}I} Q(M\setminus I, K, J, L).$$

Then

$$\begin{aligned} & R(S_I)-R(S_M) \\ \ge & \sum_{\substack{K\subset\mathbb{N}\\K\ne\phi}} E\left[r\left(U^K-d^{KI}+d^{K\bar{I}}\right)\mathbb{I}_{Q_{I\Delta K}\setminus Q_{M\Delta K}}\right] - \sum_{\substack{K\subset\mathbb{N}\\K\ne\phi}} E\left[r\left(U^K-d^{KM}+d^{K\bar{M}}\right)\mathbb{I}_{Q_{M\Delta K}\setminus Q_{I\Delta K}}\right] \\ = & \sum_{\substack{K^*\subset\mathbb{N}\\K^*\ne\phi}} \sum_{J^*\subset K^*M\bar{I}} \sum_{L^*\subset K^*\bar{M}I} E\left[r\left(U^{K^*}-d^{K^*I}+d^{K^*\bar{I}}\right)\mathbb{I}_{Q(I\setminus M,K^*,J^*,L^*)}\right] \\ & -\sum_{\substack{K\subset\mathbb{N}\\K\ne\phi}} \sum_{J\subset\bar{K}M\bar{I}} \sum_{L\subset\bar{K}\bar{M}I} E\left[r\left(U^K-d^{KM}+d^{K\bar{M}}\right)\mathbb{I}_{Q(M\setminus I,K,J,L)}\right] \end{aligned}$$

We shall show that for each negative term in the 2nd sum above there is a positive term in the 1st sum so that the net contribution from these two terms becomes nonnegative, and thus proving that $R(S_I)\ge R(S_M)$. So fix a negative term, i.e., fix $\phi\ne K\subset\mathbb{N}$, $J\subset\bar{K}M\bar{I}$, $L\subset\bar{K}\bar{M}I$. Consider $K^*=K\cup J\cup L$, $J^*=J$, $L^*=L$.

First note that $J\subset M\bar{I}$ and $J\subset K^* \Rightarrow J^*=J\subset K^*M\bar{I}$, and similarly $L^*=L\subset K^*\bar{M}I$.

Next $J\cup L\subset (I\Delta M)\Rightarrow$

$$\bar{K}^*\bar{M}\bar{I}=\bar{K}\cap\overline{(J\cup L)}\cap\bar{M}\bar{I}=\bar{K}\bar{M}\bar{I}, \qquad \bar{K}^*MI=\bar{K}\cap\overline{(J\cup L)}\cap MI=\bar{K}MI,$$

$$K^*MI=KMI\cup[(J\cup L)\cap MI]=KMI \qquad K^*\bar{M}\bar{I}=K\bar{M}\bar{I}\cup[(J\cup L)\cap\bar{M}\bar{I}]=K\bar{M}\bar{I}.$$

Further, note that $J\subset\bar{K}M\bar{I}\Rightarrow\bar{J}\supset K\cup\bar{M}\cup I\supset KM\bar{I}$. Then

$$L\subset\bar{M}I \quad\Rightarrow\quad \bar{J}^*K^*M\bar{I}=\bar{J}\cap(J\cup K\cup L)\cap M\bar{I}=\bar{J}\cap K\cap M\bar{I}=KM\bar{I}.$$

In a similar fashion one can see that $\bar{L}^*K^*\bar{M}I=K\bar{M}I$. Finally note that

$$L\subset\bar{M}I \quad\Rightarrow\quad \bar{L}\supset M\bar{I} \quad\Rightarrow\quad \bar{K}^*M\bar{I}=\bar{K}\cap\bar{J}\cap\bar{L}\cap M\bar{I}=\bar{J}\bar{K}M\bar{I},$$

and similarly $\bar{K}^*\bar{M}I=\bar{L}\bar{K}\bar{M}I$.

Hence $Q(M\setminus I,K,J,L) = Q(I\setminus M,K^*,J^*,L^*)$, and

$$\begin{aligned} & E\left[r\left(U^{K^*}-d^{K^*I}+d^{K^*\bar{I}}\right)\mathbb{I}_{Q(I\setminus M,K^*,J^*,L^*)}\right]-E\left[r\left(U^K-d^{KM}+d^{K\bar{M}}\right)\mathbb{I}_{Q(M\setminus I,K,J,L)}\right] \\ = & E\left[\left\{r\left(\underbrace{U^{K^*}-d^{K^*I}+d^{K^*\bar{I}}}_{b}\right)-r\left(\underbrace{U^K-d^{KM}+d^{K\bar{M}}}_{a}\right)\right\}\mathbb{I}_{Q(M\setminus I,K,J,L)}\right] \\ \ge & 0, \end{aligned}$$

because from $K^*MI = KMI,\ K^*\bar{M}\bar{I} = K\bar{M}\bar{I}$, and the facts that $d_i > 0,\ i \in M;\ d_i \leq 0,\ i \notin M$, we have

$$\begin{aligned}
b - a &= U^{K^*} - d^{K^*I} + d^{K^*\bar{I}} - U^K + d^{KM} - d^{K\bar{M}} \\
&= \quad U^K + U^J + U^L \quad - (d^{K^*MI} + d^{K^*\bar{M}I}) \quad + (d^{K^*M\bar{I}} + d^{K^*\bar{M}\bar{I}}) \\
&\quad - U^K \qquad\qquad\qquad + (d^{KMI} + d^{KM\bar{I}}) \qquad - (d^{K\bar{M}I} + d^{K\bar{M}\bar{I}}) \\
&= U^J + U^L - d^{K^*\bar{M}I} + d^{K^*M\bar{I}} + d^{KM\bar{I}} - d^{K\bar{M}I} \\
&= U^J + U^L - d^L - d^{\bar{L}K^*\bar{M}I} + d^J + d^{\bar{J}K^*M\bar{I}} + d^{KM\bar{I}} - d^{K\bar{M}I} \\
&= (U^J + d^J) + (U^L - d^L) + (d^{\bar{J}K^*M\bar{I}} + d^{KM\bar{I}}) - (d^{\bar{L}K^*\bar{M}I} + d^{K\bar{M}I}) \\
&= \sum_{i \in J}(u_i + d_i) + \sum_{i \in L}(u_i - d_i) + \sum_{i \in (\bar{J}K^*M\bar{I} \cup KM\bar{I}) \subset M} d_i - \sum_{i \in (\bar{L}K^*\bar{M}I \cup K\bar{M}I) \subset \bar{M}} d_i \\
&\geq 0,
\end{aligned}$$

if $(U_1, \ldots, U_n) \in Q(M \setminus I, K, J, L)$.

Finally, it is clear that this mapping from negative term to positive term is injective. Hence we do not use the same positive term for two different negative terms, i.e., each negative term is compensated by a unique positive term. Hence the result is proved.

6 Computational Experience

In this section we report our experiences with algorithms that we designed to obtain minimum risk solutions to combinatorial optimization problems. We chose to implement our algorithms to obtain minimum risk solutions to the 0-1 knapsack problem (01KP; see Martello and Toth [5]). The algorithms were coded in C and compiled and run on a 2.8GHz personal computer with 512MB RAM running Linux.

6.1 Description of the Problem Sets

For the 01KP, we chose problems with r random elements, and 10 fixed (non-random) elements. The marginal distributions for each of the r random element were discrete, each having a pre-specified number P support points. Therefore, each of our problems therefore have 2^r candidate optimal solutions, and P^r support points in the joint distribution of the random elements.

In our experiments, we considered two situations: one with r *independent* random elements and the other with r *dependent* random elements. To facilitate comparison, for each problem instance in the dependent case, the joint distributions of the random elements were generated keeping the marginal distribution for each random element identical with the marginal distribution of the element with the same index in the corresponding independent case.

We experimented with the (r, P) pairs (6, 6), (6, 8), (6, 10), (8, 4), and (8, 6). Problems smaller than these were too trivial to be interesting computationally, while problems larger than these took exorbitant amounts of solution time. For each of the (r, P) pairs that we chose, we generated ten instances. Each instance consists of the profit and cost values of all the fixed elements, the non-random cost values of the random elements, and the joint distribution of the profits of the random elements. The collection of these ten instances is called a set. The performance of an algorithm on any instance is measured by the suboptimality (defined by

(28) later) of the solution it generates — the higher the value of suboptimality, the worse the performance. The performance of an algorithm on any of the sets is measured by the average of the performances of the algorithm on all the ten problem instances in the set. Table 1 presents the size of the search problem for our chosen values of r and P.

Table 1: Size of the search problem

r	P	2^r	P^r
6	6	64	46656
6	8	64	262144
6	10	64	1000000
8	4	256	65536
8	6	256	1679616

6.2 Description of the Algorithms

For knapsack problems, we do not have an efficient representation of the regions in the solution space in which a particular solution is the minimum cost one. We are therefore unable to make use of (6) in our implementations; instead we adopt one of the two methods described below to compute the risk of any solution.

Generate All Support Points (GASP) In this method, all support points of the joint distribution are generated. The objective function values of each of the candidate optimum solutions are obtained for the support point, and the maximum of the solution values is retained as the best solution value achievable at that support point. In order to compute the risk associated with any solution, the objective function value of the solution is computed at each point in the support of the joint distribution, and the suboptimality of the solution at that support point is computed, making use of the retained best solution value at that support point. The expected value of the suboptimality of the solution is then computed as the risk of the solution.

Monte-Carlo (MC) In this process (see, e.g., Casella and Robert [1]), a simple random sample of a pre-specified number (s) of points are generated in the support of the joint distribution. As in GASP, the objective function values of each of the candidate optimum solutions are obtained for the support points in the sample, and the maximum of the objective values is retained as the best solution value achievable at that support point. In order to compute the risk associated with any solution, the objective function value of the solution is computed at each of the sampled support points, and the suboptimality of the solution at that support point is computed. The suboptimalities of the solution at each sampled support point are added up and appropriately scaled to provide a measure of the risk of the solution. Needless to say, this is an approximate value of the risk of the solution.

We also implement the following two ways of searching for a least risk solution among the candidate optimal solutions.

Complete Enumeration (CE) In this method, we simply evaluate the risk associated with each of the 2^r candidate optimal solutions, and choose one with the minimum risk value. This method is extremely time consuming, and is appropriate only for very small problems. However, it is also an assured method of finding a least risk solution when combined with GASP, and can be used to benchmark the performance of other algorithms.

Tabu Search (TS) Tabu Search is a well-known method (see, e.g., Glover and Laguna [4]) of obtaining high-quality solutions to large combinatorial optimization problem. It is an extension of the local improvement algorithm. The pseudocode below describes the procedure.

Procedure Tabu Search

Step 0 (Initialize) : Choose a solution as the *current* solution. Generate an empty list $TABU$. Let $BestSolution \leftarrow current$.

Step 1 (Terminate) : If a user-specified termination condition is reached, output $BestSolution$ and exit.

Step 2 (Search) : Search the neighborhood of *current*. Let $BestNontabu$ be the best solution in the neighborhood that can be reached from the current solution without the aid of any move in the $TABU$ list. Also let $BestTabu$ be the best solution in the neighborhood that can be reached from the current solution using a move in the $TABU$ list.

Step 3a (Aspirate) : If $BestTabu$ is better than both $BestNontabu$ and $BestSolution$ then let $BestSolution \leftarrow BestTabu$, $current \leftarrow BestTabu$, and empty the $TABU$ list. Go to Step 1.

Step 3b (Move) : Choose a value of *tenure* and add the move from *current* to $BestNontabu$ to the $TABU$ list for a period of *tenure* iterations. Let $current \leftarrow BestNontabu$. If $BestNontabu$ is better than $BestSolution$ then let $BestSolution \leftarrow BestNontabu$.

Step 4 (Update) : Update the $TABU$ list by removing the moves that have already been in the $TABU$ list for their prescribed tenure. Go to Step 1. ◇

In our implementations, two solutions are said to be neighbors if the sets of random elements in the two solutions differ by at most two elements. The solution chosen as the initial *current* solution is an optimal solution to the 01KP instance obtained by setting the profit value of each of the random elements to the expected value of its marginal distribution. The *tenure* value is chosen as a random integer between $\frac{N}{2}$ and $\frac{2N}{3}$. The termination criterion was based on the execution time alloted for the search, and was set between 15 CPU seconds and 250 CPU seconds depending on the r and P values.

Given that we have two methods for computing the risk of a solution and two methods for search a least risk solution, we define four algorithms, GASP-CE, which uses GASP to compute the risk of each individual solution, and CE to obtain a least risk solution; GASP-TS, which uses GASP to compute the risk of each individual solution, and TS to obtain a least risk solution; MC-CE, which uses MC to compute the risk of each individual solution, and CE to obtain a least risk solution; and MC-TS, which uses MC to compute the risk of each individual solution, and TS to obtain a least risk solution. Of these we recommend GASP-CE for instances with low r and P values, GASP-TS for instances with low P values and moderate

r values, MC-CE for instances with moderate r and P values, and MC-TS for instances with high r and P values.

6.3 Results from Computations

We first report our experience with the execution times required by the four algorithms on our problem sets. The time required by the algorithms can be broken up into two components, the time required by the algorithms to generate the support points (i.e., the GASP and the MC components) and the time required to search for the least risk solution among the candidate solutions (i.e., the CE and TS components). Table 2 reports the times taken by the components of the GASP-CE algorithm, while Table 3 reports the times required by the CE component of the MC-CE algorithm with different cardinalities of the support (denoted by s). (Since the s values are very small compared to the P^r values for these problem instances, the time required by the MC component is very small, and is not reported.) The time required by any of the components on an instance in the dependent case was never found to be significantly different from the time required by that component on the corresponding problem instance in the independent case; therefore we report an average of these times in our tables.

Table 2: Execution times required by GASP-CE (in CPU seconds)

r	P	GASP	CE
6	6	1.625	2.475
6	8	9.094	13.871
6	10	34.759	53.005
8	4	11.558	18.641
8	6	280.255	433.200

Table 3: Execution times required by the CE component of the MC-CE (in CPU seconds)

r	P	CE		
		s=3000	s=4000	s=5000
6	6	0.037	0.056	0.078
6	8	0.037	0.055	0.079
6	10	0.037	0.057	0.077
8	4	0.154	0.236	0.330
8	6	0.155	0.240	0.329

In our experiments with GASP-TS and MC-TS, we found that GASP-TS and MC-TS used the same time as GASP-CE and MC-CE respectively to generate all support points and compute the maximum profit solution at each support point. The TS procedure took less that 0.001 seconds in all cases for GASP-TS and the maximum time limit set for MC-TS to generate the solution that the algorithms finally output.

Note that GASP-CE is an exact algorithm in the sense that it outputs the least risk solution to the problem. The three other algorithms output solutions which are not necessarily minimum risk (but hopefully low risk). The suboptimality of a solution output by any algorithm (say $\mathcal{A}$) to an instance is computed as

$$\text{suboptimality} = \frac{R(x^A) - R(x^*)}{R(x^*)} \tag{28}$$

where x^* is the least risk solution for the instance, and x^A is the solution output by the algorithm $\mathcal{A}$.

The performance of GASP-TS was very encouraging for the problems we tested it on. Each of the solutions that it output was found to be optimal. Hence we can conclude that, at least for these problem sizes, GASP-TS clearly outperforms GASP-CE in terms of execution times, without sacrificing solution quality.

In case of MC-CE and MC-TS algorithms we chose three values of s for our experiments, viz. 3000, 4000, and 5000. Since we take a random sample for the MC procedure, we performed 25 runs for each problem instance and chose the average of the suboptimality values over all 25 runs as the suboptimality of the MC-CE algorithm for each instance.

Table 4: Quality of solutions output by MC-CE and MC-TS when random elements are independent

r	P	s	MC-CE		MC-TS	
			mean	s.d.	mean	s.d.
6	6	3000	1.7354	0.0060	0.4826	0.0000
6	8	3000	1.9721	0.0048	0.5169	0.0000
6	10	3000	1.9564	0.0010	0.4205	0.0000
8	4	3000	0.0230	0.0100	0.0770	0.0000
8	6	3000	0.0194	0.0020	0.1775	0.0000
6	6	4000	1.7347	0.0056	0.4588	0.0000
6	8	4000	1.9697	0.0022	0.4964	0.0000
6	10	4000	1.9562	0.0000	0.4205	0.0000
8	4	4000	0.0230	0.0107	0.0659	0.0001
8	6	4000	0.0198	0.0034	0.2039	0.0000
6	6	5000	1.7337	0.0032	0.2672	0.0000
6	8	5000	1.9716	0.0038	0.5169	0.0000
6	10	5000	1.9564	0.0010	0.4205	0.0000
8	4	5000	0.0236	0.0083	0.0770	0.0000
8	6	5000	0.0191	0.0025	0.2039	0.0000

Tables 4 and 5 contain the results of our experiments with MC-CE and MC-TS on the problem sets. The mean suboptimality value for solutions output by MC-CE for problems with 6 random elements is seen to be high when compared to those for solutions output by MC-TS. On inspection of the results for individual instances, this high mean suboptimality is seen to be caused by one problem instance in the set. If we remove this problem instance from the sets, then the mean suboptimality values of the solutions output by MC-CE are seen to be of the same order as those for the solutions output by MC-TS. Changing the set of points in the sample of support points chosen by MC-CE for this problem however, did not remove this anomaly.

Table 5: Quality of solutions output by MC-CE and MC-TS when random elements are dependent

r	P	s	MC-CE		MC-TS	
			mean	s.d.	mean	s.d.
6	6	3000	1.7395	0.0077	0.4835	0.0000
6	8	3000	1.9724	0.0052	0.5169	0.0000
6	10	3000	1.9563	0.0010	0.4205	0.0000
8	4	3000	0.0221	0.0076	0.0766	0.0000
8	6	3000	0.0188	0.0044	0.1773	0.0000
6	6	4000	1.7380	0.0058	0.4595	0.0000
6	8	4000	1.9700	0.0027	0.4965	0.0000
6	10	4000	1.9561	0.0000	0.4205	0.0000
8	4	4000	0.0206	0.0106	0.0657	0.0001
8	6	4000	0.0197	0.0027	0.2005	0.0000
6	6	5000	1.7376	0.0056	0.2661	0.0000
6	8	5000	1.9717	0.0038	0.5169	0.0000
6	10	5000	1.9563	0.0010	0.4205	0.0000
8	4	5000	0.0226	0.0079	0.0766	0.0000
8	6	5000	0.0182	0.0024	0.2005	0.0000

From the standard deviation values seen in the tables, it seems that MC-CE is sensitive to the samples of support points chosen by the Monte-Carlo method, while MC-TS is not. However, on examination of the results for individual problem instances, MC-CE is seen to be sensitive to the choice of sample points in 1 or 2 of the instances in the sets with 6 random elements, and in 5 or 6 of the instances in the sets with 8 random elements.

We do not see any consistent improvement in the quality of solutions when the sample size is increased. This is surprising, although we think that such an improvement will be observed when the sample size increases significantly. The quality of solutions output by the algorithms when the random elements are independent is not significantly different from when the random elements are dependent. This is expected, since the algorithms do not make use of the property of independence (or otherwise) of the marginal distributions.

7 Summary and Concluding Remarks

In this paper we consider stochastic discrete optimization problems (SDOP) with multiple random elements. Though the problem of solving a SDOP is quite involving we have found a way to reduce the search for optimal solutions by characterizing the candidate optimal solutions. A complete characterization of the optimal solution is still to be found. We have used numerical algorithms based on local search heuristics to find the optimal solution of stochastic binary knapsack problems with six to eight random elements, whose joint distributions are discrete. Our results suggest that when the number of random elements are in this range, tabu search is a more efficient choice than complete enumeration for solving problems.

References

[1] G. Casella and C. P. Robert, *Monte Carlo Statistical Methods*, Springer, 2005.

[2] S. Das, D. Ghosh, and P. K. Mandal., *On solving some stochastic discrete optimization problems under general regret function*, Working Paper 238, IIM Bangalore, India, 2005.

[3] D. Ghosh, P. K. Mandal, and S. Das, *On Solving Discrete Optimization Problems with One Random Element Under General Regret Functions*, Unpublished manuscript, 2005.

[4] F. Glover and M. Laguna, *Tabu Search*, Kluwer Academic Publishers, 1998.

[5] S. Martello and P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*, Wiley, Chichester, West Sussex, England, 1990.

A Tables

Table 6: Table for Example 2: Least Risk solutions (with $r(t) = (1+t)^2 - 1$) in balanced DOP with $r = 2, f = 2, k = 2$. DOP symmetric when both random elements have Triangular distribution on (0,1).

		Risk of candidate solutions				Optimal solution	
c_1	c_2	S_{12}	S_1^2	S_2^1	S^{12}	stochastic	non-stochastic
0.5020	0.5216	0.0516	0.0541	0.0541	0.0635	1	1
0.5160	0.5304	0.0461	0.0548	0.0548	0.0695	1	1
0.4260	0.4350	0.1010	0.0587	0.0587	0.0292	4	4
0.4630	0.4831	0.0723	0.0545	0.0545	0.0450	4	4
0.4771	0.5888	0.0521	0.0491	0.0491	0.0870	2 or 3	2 or 3
0.4872	0.5235	0.0562	0.0528	0.0528	0.0616	2 or 3	2 or 3
0.4937	0.5844	0.0471	0.0504	0.0504	0.0881	1	2 or 3
0.4069	0.5816	0.0806	0.0461	0.0461	0.0743	2 or 3	2 or 3
0.4052	0.5064	0.0931	0.0501	0.0501	0.0462	4	2 or 3
0.4920	0.5945	0.0467	0.0499	0.0499	0.0926	1	2 or 3
0.4381	0.5541	0.0704	0.0483	0.0483	0.0663	2 or 3	2 or 3
0.4739	0.5049	0.0639	0.0532	0.0532	0.0532	4	2 or 3

Among the candidate solutions, '1' stands for $S_{1,2}$, '2' for S_1^2, '3' for S_2^1, '4' for $S^{1,2}$.

Table 7: Table Reporting Simulation Results: Least Risk solutions in balanced DOP with $r = 3, f = 3, k = 3$, when the random elements are independent variables having (symmetric) marginal distributions V-shaped (mean 10.0, supported on 10 ± 3) Triangular shaped (mean 10.1, supported on 10.1 ± 3.0) and Normal distribution with mean 10.11 and variance 1.

fixed costs			Risk of candidate solutions (with $r(t) = (1+t)^2 - 1$)								Optimal solution	
c_1	c_2	c_3	S_{123}	S^3_{12}	S^2_{13}	S^1_{23}	S^{23}_1	S^{13}_2	S^{12}_3	S^{123}	stoch.	non-stoch.
10.13	10.27	10.62	7.98	8.23	8.33	9.26	9.38	10.34	10.44	13.91	1	1
9.55	9.94	10.42	12.84	9.23	9.28	9.96	8.25	8.90	8.95	10.66	5	5
9.65	10.40	10.99	10.16	7.66	7.76	8.66	9.49	10.45	10.55	16.76	2	2
9.64	9.85	10.39	12.72	9.62	9.65	10.31	8.09	8.69	8.73	10.25	5	5
10.08	10.52	10.93	7.39	7.50	7.63	8.65	10.22	11.33	11.46	17.55	1	2
9.13	9.53	9.99	18.94	11.42	11.40	11.86	7.69	8.03	8.00	7.00	8	8
9.69	10.65	10.83	9.23	7.15	7.27	8.27	10.52	11.62	11.76	17.14	2	2
10.10	10.36	10.45	7.92	8.05	8.15	9.09	9.73	10.72	10.83	13.17	1	2
9.54	9.74	9.88	14.69	10.59	10.59	11.09	8.21	8.63	8.63	7.11	8	8
10.40	10.72	10.76	5.51	7.34	7.50	8.70	11.43	12.75	12.93	18.05	1	1
9.52	9.78	10.23	14.03	10.01	10.03	10.63	8.00	8.53	8.55	9.04	5	5
9.35	10.79	10.90	10.96	6.84	6.97	7.99	11.06	12.21	12.36	18.43	2	2
9.22	9.52	9.88	18.45	11.61	11.58	12.00	7.80	8.11	8.07	6.48	8	8
9.20	10.35	10.55	13.44	7.84	7.92	8.76	9.29	10.18	10.27	13.16	2	2
9.08	9.11	9.48	23.85	14.81	14.71	14.87	7.93	7.88	7.76	4.18	8	8
9.40	9.43	9.87	17.62	12.09	12.06	12.46	7.73	7.99	7.95	6.31	8	8
9.53	10.63	10.64	10.28	7.20	7.32	8.29	10.42	11.49	11.62	15.45	2	2
9.61	10.25	11.07	10.89	8.03	8.12	8.96	8.97	9.84	9.93	16.57	2	2
9.29	9.75	9.96	16.36	10.43	10.44	10.95	8.09	8.53	8.53	7.42	8	8
9.60	9.62	10.24	14.40	10.76	10.77	11.32	7.72	8.18	8.18	8.63	5	5
9.32	10.22	10.84	12.91	8.12	8.19	9.00	8.83	9.67	9.74	14.59	2	2
9.33	9.95	10.99	14.03	8.97	9.02	9.73	8.11	8.78	8.84	14.44	5	5
10.05	10.66	10.97	7.24	7.21	7.35	8.41	10.76	11.93	12.08	18.66	2	2
9.73	10.01	10.46	11.31	8.99	9.05	9.78	8.46	9.18	9.23	11.29	5	5
9.52	9.85	10.05	13.94	9.92	9.94	10.52	8.25	8.78	8.80	8.25	8	5

Among the candidate solutions, ‘1’ stands for S_{123}, ‘2’ stands for S^3_{12}, ‘3’ stands for S^2_{13}, ‘4’ stands for S^1_{23}, ‘5’ stands for S^{23}_1, ‘6’ stands for S^{13}_2, ‘7’ stands for S^{12}_3, ‘8’ stands for S^{123}.

Table 8: Table Reporting Simulation Results: Least Risk solutions in balanced DOP with $r = 3, f = 3, k = 3$, when the random elements are independent and normally distributed with unit variance, but means and c_i's change during simulation.

means of random costs			fixed costs			Risk of candidate solutions with $r(t) = (1+t)^2 - 1)$								Optimal solution	
μ_1	μ_2	μ_3	c_1	c_2	c_3	S_{123}	S_{12}^3	S_{13}^2	S_{23}^1	S_1^{23}	S_2^{13}	S_3^{12}	S^{123}	stoch.	non-stoch.
9.71	10.07	9.06	10.07	10.44	10.86	4.34	10.76	4.55	6.61	13.39	16.46	8.77	23.27	1	1
9.18	9.26	9.61	9.78	10.58	10.79	3.79	5.02	7.05	7.53	13.31	13.94	16.96	28.04	1	1
10.25	9.98	10.37	10.58	10.70	10.99	4.88	6.20	8.58	6.88	10.87	8.99	11.72	16.44	1	1
9.00	9.57	9.62	9.55	10.91	10.96	4.73	5.00	5.28	8.64	13.10	17.93	18.40	31.73	1	2
9.23	9.05	9.72	9.66	9.88	10.14	5.14	4.98	8.99	7.84	10.02	8.82	13.68	16.89	2	2
9.16	10.46	9.41	9.45	10.07	10.37	8.74	9.57	3.38	11.36	6.81	16.49	8.39	15.16	3	3
9.89	10.46	9.62	9.03	9.90	10.00	15.01	11.26	5.76	9.34	7.38	11.28	5.75	7.90	7	3
10.07	9.69	9.99	10.07	10.61	10.76	5.48	6.28	8.12	5.83	12.24	9.48	11.64	17.56	1	4
10.63	9.06	9.39	9.33	10.16	10.98	9.65	10.01	12.37	2.77	18.65	6.66	8.72	20.83	4	4
9.43	9.48	10.81	9.21	9.47	10.64	15.44	4.79	13.73	14.16	4.99	5.28	14.33	11.88	2	5
10.21	10.83	10.89	10.21	10.65	10.69	10.94	6.78	7.19	11.36	5.64	9.51	9.97	8.44	5	5
10.99	9.74	10.36	9.09	9.32	9.77	26.06	14.45	20.05	9.81	11.47	3.85	7.36	3.82	8	6
9.82	10.37	9.03	9.47	9.76	9.85	10.40	14.12	5.01	8.31	9.48	13.59	4.65	9.55	7	7
10.24	9.72	9.71	9.08	9.33	9.76	16.52	11.80	11.75	8.09	9.23	5.97	5.93	5.91	8	7
10.28	10.15	10.41	9.03	9.30	9.59	24.77	12.61	14.68	13.67	6.93	6.20	7.76	3.11	8	8
10.72	10.90	10.56	9.07	9.57	9.81	31.67	17.03	14.18	15.64	7.22	8.33	6.27	2.35	8	8
10.33	9.89	10.94	9.24	9.62	9.88	22.12	8.60	16.59	12.88	7.00	4.48	10.96	4.25	8	8

Among the candidate solutions, '1' stands for S_{123}, '2' stands for S_{12}^3, '3' stands for S_{13}^2, '4' stands for S_{23}^1, '5' stands for S_1^{23}, '6' stands for S_2^{13}, '7' stands for S_3^{12}, '8' stands for S^{123}.