Totally Unimodular Stochastic Programs

Nan Kong
Department of Industrial and Management Systems Engineering, University of South Florida
4202 East Fowler Ave, ENB 118, Tampa, FL 33620, USA, kong@eng.usf.edu
Andrew J. Schaefer
Department of Industrial Engineering, University of Pittsburgh
1048 Benedum Hall, Pittsburgh, PA 15261, USA, schaefer@ie.pitt.edu
Shabbir Ahmed
School of Industrial and Systems Engineering, Georgia Institute of Technology
765 Ferst Drive, Atlanta, GA 30332, USA, sahmed@isye.gatech.edu

We consider *totally unimodular stochastic programs*, that is, stochastic programs whose extensive-form constraint matrix is totally unimodular. We generalize the notion of total unimodularity to apply to sets of matrices and provide properties of such sets. Using this notion, we give several sufficient conditions for stochastic programs to be totally unimodular, and provide necessary conditions for specific classes of problems. When solving such problems using the L-shaped method it is not clear whether the integrality restrictions should be imposed on the master problem. Such restrictions will make each master problem more difficult to solve. On the other hand, solving the linear relaxation of the master typically means sending fractional (and unlikely optimal) solutions to the subproblems, perhaps leading to more iterations. Our computational results investigate this trade-off and provide insight into which strategy is preferable under a variety of circumstances.

Key words: Stochastic Integer Programming; Total Unimodularity; L-shaped Method

---

1. Introduction

Consider the extensive form of a two-stage stochastic mixed-integer program where all first-stage decision variables are integers.

$$\min c^T x \; + \; \sum_{k=1}^{K} p^k (d^k)^T y^k \tag{1}$$

subject to

$$T^k x \; + \; W^k y^k \; \geq \; h^k, \; 1 \; \leq \; k \; \leq \; K, \tag{2}$$

$$x \in X = \mathbb{Z}_+^{n_1}, \; y^k \in Y = \mathbb{R}_+^l \times \mathbb{Z}_+^{n_2 - l}, \; 1 \; \leq \; k \; \leq \; K. \tag{3}$$

The vector c is a known vector in $\mathbb{R}^{n_1}$, and for every realization (or *scenario*) k, d^k is a vector in $\mathbb{R}^{n_2}$, h^k is a vector in $\mathbb{R}^m$, T^k, the *technology matrix*, is a matrix in $\mathbb{R}^{m\times n_1}$, and W^k, the *recourse matrix*, is a matrix in $\mathbb{R}^{m\times n_2}$. Scenario k occurs with probability p^k. Assume that there are a finite number of scenarios. Without loss of generality, we assume that any explicit constraints on the first-stage decisions x have been incorporated into all of the T^k matrices, where the corresponding rows are $\mathbf{0}$ in the W^k matrices.

The *deterministic equivalent* of this stochastic program is given by

$$\min c^T x \ + \ \mathcal{Q}(x) \tag{4}$$

subject to

$$x \in X = \mathbb{Z}_+^{n_1}, \tag{5}$$

where the *expected recourse function*, $\mathcal{Q}(x) \ = \ \sum_{k=1}^{K} p^k Q(x,k)$, and

$$Q(x,k) = \min (d^k)^T y \tag{6}$$

subject to

$$W^k y \ \geq \ h^k \ - T^k x, \ 1 \ \leq \ k \ \leq \ K, \tag{7}$$

$$y \in Y = \mathbb{R}_+^l \times \mathbb{Z}_+^{n_2 - l}. \tag{8}$$

In the case of stochastic linear programs, it is well known [4] that $\mathcal{Q}(x)$ is a convex function, and algorithms for solving stochastic linear programs, such as the so-called "L-shaped" method [19] and its variants, exploit this property. Wollmer [20] showed that when $X = \mathbb{B}^{n_1}$ and $Y = \mathbb{R}_+^{n_2}$, a slight modification of the L-shaped method applies. In contrast, when Y contains integrality restrictions, $\mathcal{Q}(x)$ is in general nonconvex and discontinuous [17]. As a result, stochastic programs with integer recourse are difficult to solve. Recent approaches to such problems include Ahmed et al. [1], Kong et al. [9], and Sen and Higle [16]. Carøe [3], Klein Haneveld and van der Vlerk [7], and Schultz [15] gave comprehensive surveys of the state of stochastic integer programming.

Define the matrix B to be the constraint matrix of the extensive form of the stochastic program described in (1) - (3),

$$B \ = \ \begin{pmatrix} T^1 & W^1 & & & \\ T^2 & & W^2 & & \\ \vdots & & & \ddots & \\ T^K & & & & W^K \end{pmatrix}. \tag{9}$$

Let $M = m \times K$ and $N = n_1 + n_2 \times K$, so that B is an $M \times N$ matrix. This paper addresses the question of when can the stochastic mixed-integer program defined in (1) - (3) be solved as a stochastic linear program for any integer right-hand side h^k. It is well known that this is true if and only if B is *totally unimodular* (TU).

The literature on total unimodularity in deterministic optimization is vast. Padberg [13], Schrijver [14], and Nemhauser and Wolsey [12] provided surveys. Several authors have considered total unimodularity within stochastic programming. Birge and Louveaux [2] recognized the utility of the constraint matrix of a stochastic program being TU. They described one sufficient condition for B to be TU and concluded that stochastic programs are unlikely to meet this sufficient condition. van der Vlerk [18] provided a class of convex approximations for complete integer recourse models, and showed that when the recourse matrix is TU these approximations are exact.

The remainder of this paper is organized as follows. In Section 2 we generalize total unimodularity so that it applies to sets of matrices. In Section 3 we give sufficient conditions for the constraint matrix of the extensive form of a two-stage stochastic program to be totally unimodular. For some special classes of stochastic programs, these conditions are also necessary. Section 4 shows the validity of using the L-shaped method with two different approaches to optimize totally unimodular stochastic programs. These two approaches differ with respect to whether the integrality restrictions are imposed on the master problem. We provide computational results in Section 5 that investigate the trade-off between the two approaches, and give conclusions in Section 6.

2. A Generalization of Total Unimodularity

The following theorem is a well known characterization of total unimodularity for a single matrix.

Theorem 1 *(Ghouila-Houri [5]) An $m \times n$ matrix A is TU if and only if for every $J \subseteq N' = \{1, \ldots, n\}$ there exists a partition (J^1, J^2) of J such that*

$$\left| \sum_{j \in J^1} a_{ij} - \sum_{j \in J^2} a_{ij} \right| \leq 1 \text{ for } i = 1, \ldots, m. \tag{10}$$

We will find it convenient to extend the definition of total unimodularity so that it addresses groups of matrices.

Definition 1 *Let $\mathcal{T} = \{A_1, \ldots, A_T\}$ be a set of $(0, \pm 1)$ matrices, each of which is of $m \times n_t$ dimension, $t = 1, \ldots, T$. Also let $v \in \{0, \pm 1\}^m$. The set $\mathcal{T}$ is* TU with respect to v, *denoted by $TU(v)$, if for all column subsets $J_t \subseteq A_t$, $1 \leq t \leq T$, there exist partitions (J_t^1, J_t^2) such that*

$$\sum_{j \in J_t^1} a_{ij}^t - \sum_{j \in J_t^2} a_{ij}^t \in \{0, 1\} \text{ for } v_i = -1, \ 1 \leq t \leq T, \tag{11}$$

$$\sum_{j \in J_t^1} a_{ij}^t - \sum_{j \in J_t^2} a_{ij}^t \in \{0, \pm 1\} \text{ for } v_i = 0, \ 1 \leq t \leq T, \tag{12}$$

and

$$\sum_{j \in J_t^1} a_{ij}^t - \sum_{j \in J_t^2} a_{ij}^t \in \{0, -1\} \text{ for } v_i = 1, \ 1 \leq t \leq T. \tag{13}$$

This generalization of total unimodularity will be used in Section 3 to characterize when a stochastic program is totally unimodular. Note that the matrices in $\mathcal{T}$ do not need to have

We next give several properties of a $TU(v)$ set of matrices, some of which provide ways of constructing other $TU(v)$ sets of matrices from a given set. The proofs of most of these properties are obvious and are omitted.

Proposition 1 *A matrix A is TU if and only if $\{A\}$ is $TU(\mathbf{0})$.*

Proposition 2 *For any $v \in \{0, \pm 1\}^m$, $\mathcal{T}$ is $TU(v) \Rightarrow \mathcal{T}$ is $TU(\mathbf{0})$.*

Proposition 3 *For any $v \in \{0, \pm 1\}^m$, if $\mathcal{T}$ is $TU(v)$, then so is any proper subset of $\mathcal{T}$.*

Corollary 1 *For any $v \in \{0, \pm 1\}^m$, if $\mathcal{T}$ is $TU(v)$, each matrix A in $\mathcal{T}$ is TU.*

Proposition 4 *For any $v \in \{0, \pm 1\}^m$, if $\{A\}$ is $TU(v)$ then $(A|v)$ is TU.*

Proof: Consider any subset J of the columns of $(A|v)$. If $v \notin J$ the result follows from the fact that A is TU. Otherwise, let J^1 and J^2 be as in the definition of $TU(v)$, and let $J^1 = J^1 \cup \{v\}$. Consider any $1 \leq i \leq m$, and suppose $v_i = 1$. Then

$$\sum_{j \in J^1} a_{ij} - \sum_{j \in J^2} a_{ij} \in \{0, -1\}, \tag{14}$$

so

$$v_i + \sum_{j \in J^1} a_{ij} - \sum_{j \in J^2} a_{ij} \in \{0, 1\}. \tag{15}$$

A similar argument holds if $v_i = 0$ or -1. □

Proposition 5 *If $\mathcal{T}$ is $TU(v)$, $\mathcal{T}$ is $TU(-v)$ as well.*

Proof: Interchange the sets J_t^1 and J_t^2 for all $1 \leq t \leq T$. $\square$

Proposition 6 *For any $v \in \{0, \pm 1\}^m$, negating any matrix in a $TU(v)$ set results in a $TU(v)$ set.*

Proof: Interchange the sets J_t^1 and J_t^2 for the negated matrix A_t. $\square$

Proposition 7 *For any $v \in \{0, \pm 1\}^m$, the union of two sets of $TU(v)$ matrices is $TU(v)$.*

Corollary 2 *For any $v \in \{0, \pm 1\}^m$, duplicating matrices in a $TU(v)$ set results in a $TU(v)$ set.*

Corollary 3 *For any $v \in \{0, \pm 1\}^m$, if $\mathcal{T}$ is $TU(v)$, then $\mathcal{T} \cup \{I, -I\}$ is $TU(v)$ as well.*

Two matrices in a set of $TU(v)$ matrices may be combined into a larger TU matrix.

Theorem 2 *For any $v \in \{\pm 1\}^m$, if $\{A_1, \ldots, A_T\}$ is $TU(v)$, then the matrix $(A_s|A_t)$ is TU for all $1 \leq s, t \leq T$.*

Proof: Consider any subset $J = (J_s, J_t)$ of columns of $(A_s|A_t)$. Since $\{A_s, A_t\}$ is $TU(v)$, there exist a partition (J_s^1, J_s^2) of J_s and a partition (J_t^1, J_t^2) of J_t that satisfy the definition of $TU(v)$. Consider $J^1 = \{J_s^1, J_t^2\}$ and $J^2 = \{J_s^2, J_t^1\}$. This partition satisfies the condition of Theorem 1. $\square$

3. Characterizing Totally Unimodular Stochastic Programs

Consider the two-stage stochastic programming polyhedron defined by

$$P = \left\{ x \in \mathbb{R}_+^{n_1}, y \in \mathbb{R}_+^{n_2 \times K} \middle| T^k x + W^k y^k \geq h^k, \ 1 \leq k \leq K \right\}. \tag{16}$$

We denote $h = ((h^1)^T, (h^2)^T, \ldots, (h^K)^T)^T$.

Theorem 3 *The extensive-form constraint matrix B of a two-stage stochastic program is TU if and only if the corresponding polyhedron P is integral for all right-hand sides $h \in \mathbb{Z}^M$ for which it is nonempty.*

Theorem 3, as it applies to deterministic optimization problems, was first proved by Hoffman and Kruskal [6].

Corollary 4 *The extensive-form constraint matrix B of a two-stage stochastic program is TU if and only if P has only integer extreme points for any integer right-hand side h for which it is nonempty.*

It should be noted that if B is TU, each matrix $T^1, \dots, W^K$ must be as well, and B must be a $(0, \pm 1)$ matrix. When Theorem 1 is applied to the extensive-form constraint matrix B, we obtain the following.

Theorem 4 *The extensive-form constraint matrix B of a two-stage stochastic program is TU if and only if for every $J = \{J_0, J_1, \dots, J_K\} \subseteq \{1, \dots, N\}$ there exists a partition $(J^1, J^2) = \{(J_0^1, J_1^1, \dots, J_K^1), (J_0^2, J_1^2, \dots, J_K^2)\}$ such that*

$$\left| \sum_{j \in J_0^1} t_{ij}^k + \sum_{j \in J_k^1} w_{ij}^k - \sum_{j \in J_0^2} t_{ij}^k - \sum_{j \in J_k^2} w_{ij}^k \right| \leq 1 \text{ for } i = 1, \dots, m, \; k = 1, \dots, K. \tag{17}$$

Corollary 5 *Consider a two-stage stochastic integer program with fixed technology matrix T and recourse matrix W. If $T = I$ and W is TU, B is TU.*

Proposition 8 *If B is TU, $((T^1)^T | \dots | (T^K)^T)^T$ is a TU matrix.*

Proof: Apply Theorem 4 to the subset of columns corresponding to the first-stage decision variables. □

When the stochastic program has *simple recourse*, i.e., the only recourse action is to incur a linear penalty for shortages or surpluses, the recourse matrix $W^k = (I, -I)$ for all k. In such cases, the condition that $((T^1)^T | \dots | (T^K)^T)^T$ is TU, is also sufficient for the stochastic program to be totally unimodular.

Proposition 9 *If a two-stage stochastic program has simple recourse, B is TU if and only if the matrix $((T^1)^T | \dots | (T^K)^T)^T$ is TU.*

Proof: "⇒": Follows from Proposition 8.

"⇐": Consider any subset $J = \{J_0, \ldots, J_K\}$ of columns of the extensive form. Since $((T^1)^T|\ldots|(T^K)^T)^T$ is TU, there exists a partition (J_0^1, J_0^2) of J_0 such that

$$\left|\sum_{j\in J_0^1} t_{ij}^k - \sum_{j\in J_0^2} t_{ij}^k\right| \leq 1,\ \forall i,\ 1 \leq k \leq K. \tag{18}$$

Initialize $J^1 = J_0^1$ and $J^2 = J_0^2$. Since each column corresponding to a second-stage variable contains exactly one non-zero entry, which is either $+1$ or -1, it is clear that the sets J^1 and J^2 can be completed so that Theorem 4 holds. □

Corollary 6 *If a two-stage stochastic program with fixed technology matrix T has simple recourse, B is TU if and only if T is TU.*

When the stochastic program does not have simple recourse, stronger conditions are needed for the technology and recourse matrices.

Theorem 5 *If a two-stage stochastic program has fixed technology matrix T, B is TU if there exists a $v \in \{\pm 1\}^m$ such that $\{T, W^1, \ldots, W^K\}$ is $TU(v)$.*

Proof: Consider any subset $J = \{J_0, J_1, \ldots, J_K\}$. Since $\{T, W^1, \ldots, W^K\}$ is $TU(v)$, there exists a partition (J_k^1, J_k^2) of each J_k and a partition of the rows such that

$$\sum_{j\in J_k^1} w_{ij}^k - \sum_{j\in J_k^2} w_{ij}^k \in \{0, 1\} \text{ for } v_i = -1,\ 1 \leq\ k \leq K, \tag{19}$$

$$\sum_{j\in J_k^1} w_{ij}^k - \sum_{j\in J_k^2} w_{ij}^k \in \{0, -1\} \text{ for } v_i = 1,\ 1 \leq\ k \leq K, \tag{20}$$

$$\sum_{j\in J_0^1} t_{ij} - \sum_{j\in J_0^2} t_{ij} \in \{0, 1\} \text{ for } v_i = -1, \tag{21}$$

and

$$\sum_{j\in J_0^1} t_{ij} - \sum_{j\in J_0^2} t_{ij} \in \{0, -1\} \text{ for } v_i = 1. \tag{22}$$

The partition of J given by $J^1 = \{J_0^2, J_1^1, \ldots, J_K^1\}$ and $J^2 = \{J_0^1, J_1^2, \ldots, J_K^2\}$ satisfies the requirements of Theorem 4. □

Of particular interest are those matrices A for which $\{A\}$ is $TU(v)$ for all $v \in \{0, \pm 1\}^m$. The identity matrix is such a matrix.

Theorem 6 *Suppose the matrix* $((T^1)^T|\dots|(T^K)^T)^T$ *is TU, and each matrix* W^k, $1 \le k \le K$*, is such that* $\{W^k\}$ *is* $TU(v)$ *for all* $v \in \{0, \pm 1\}^m$*. Then* B *is TU.*

Proof: Since $((T^1)^T|\dots|(T^K)^T)^T$ is TU, for every subset $J \subseteq \{1, \dots, n_1\}$, there must exist a partition (J^1, J^2) of J and a vector $v^k \in \{0, \pm 1\}^m$ such that

$$\sum_{j \in J^1} t_{ij}^k \; - \; \sum_{j \in J^2} t_{ij}^k = v_i^k, \tag{23}$$

for all $1 \le i \le m$ and every $1 \le k \le K$. Since each set $\{W^k\}$ is $TU(v^k)$, for each subset of columns of W^k there exists a partition such that Theorem 4 is satisfied. □

4. Optimizing Totally Unimodular Stochastic Programs

It is obvious that the special structure of totally unimodular stochastic programs makes them easier to solve than general stochastic mixed-integer programs. We present two approaches to solving these problems with the L-shaped method [19].

Consider the linear relaxation of the recourse problem in the stochastic program described in (1) - (3). Define $Q^{LP}(x, k)$ to be the optimal objective value to the LP relaxation of the recourse problem under scenario k given first-stage solution x. Let $\mathcal{Q}^{LP}(x) = \sum_{k=1}^{K} p^k Q^{LP}(x, k)$.

It is obvious that for any $x \in X$ and all scenarios k, $Q^{LP}(x, k) \le Q(x, k)$ and thus $\mathcal{Q}^{LP}(x) \le \mathcal{Q}(x)$.

Example 1 *Suppose* $K = 1$ *and* $T^1 = W^1 = I^m$*, and* $Y = \mathbb{Z}_+^m$*. Clearly, the matrix* B *is TU. Note that for any* x*,* $y_i^* = \lceil h_i - x_i \rceil^+$ *for* $1 \le i \le m$*. Therefore,* $Q(x, k) = (d^k)^T(\lceil h - x \rceil^+)$*, and so* $\mathcal{Q}(x)$ *is nonconvex and discontinuous. Note that* $\mathcal{Q}^{LP}(x) = (d^k)^T(h - x)^+ < \mathcal{Q}(x)$ *for* $h - x \notin \mathbb{Z}_+^m$*.*

Example 1 shows that even if B is TU, the expected recourse function $\mathcal{Q}(x)$ is in general nonconvex and discontinuous. Furthermore, there may exist $x \in \mathbb{R}_+^{n_1}$ such that $\mathcal{Q}^{LP}(x) < \mathcal{Q}(x)$.

Theorem 7 *Suppose* B *is TU and* $h \in \mathbb{Z}^M$*. Then for any* $x \in \mathbb{Z}_+^{n_1}$*,* $\mathcal{Q}(x) = \mathcal{Q}^{LP}(x)$*.*

Proof: Consider any scenario k and any $x \in \mathbb{Z}_+^{n_1}$. The matrices W^k are TU and the right-hand side $h^k - T^k x$ is integral, so there exists an optimal solution to the LP relaxation of the recourse problem that is integral. $\square$

Given $x \in X = \mathbb{Z}_+^{n_1}$, let $y^{LP}(x,k)$ be an extreme point optimal solution to $Q^{LP}(x,k)$ for $1 \leq k \leq K$, and let $y^{LP}(x) = ((y^{LP}(x,1))^T, \ldots, (y^{LP}(x,K))^T)^T \in \mathbb{R}_+^{n_2 \times K}$. Let $(x^{IP}, y^{LP}(x^{IP})) \in \mathbb{Z}_+^{n_1} \times \mathbb{R}_+^{n_2 \times K}$ be an optimal solution to the stochastic linear programming relaxation where integrality is relaxed in the second stage.

Theorem 8 *Consider any stochastic mixed-integer program with $h \in \mathbb{Z}^M$, described in (1) - (3), for which P in (16) is nonempty. If B is TU, then $(x^{IP}, y^{LP}(x^{IP}))$ is an optimal solution to the stochastic mixed-integer program.*

Proof: Since B is TU, T^k and W^k are TU for all k, and thus for $1 \leq k \leq K$, an extreme point optimal solution $y^{LP}(x,k)$ to $Q^{LP}(x,k)$ is integral for any feasible first-stage decision $x \in \mathbb{Z}_+^{n_1}$. Therefore, replacing Y by $\mathbb{R}_+^{n_2}$ still yields an optimal solution to the stochastic mixed-integer program. $\square$

Let $(x^{LP}, y^{LP}) \in \mathbb{R}_+^{n_1} \times \mathbb{R}_+^{n_2 \times K}$ be an extreme point optimal solution to the stochastic linear programming relaxation where integrality is relaxed in both stages.

Theorem 9 *Consider any stochastic mixed-integer program with $h \in \mathbb{Z}^M$, described in (1) - (3), for which P in (16) is nonempty. If B is TU, then (x^{LP}, y^{LP}) is an optimal solution to the stochastic mixed-integer program.*

Proof: By Corollary 4, P has only integer extreme points. $\square$

Two-stage stochastic programs with resource are frequently solved with the L-shaped method that is a variant of the Benders' decomposition in stochastic programming. It is based on adding linear cutting planes to build outer approximations of the recourse function, commonly denoted as θ, and solving an iterative master problem that is the first-stage problem plus this approximation. Theorems 8 and 9 indicate two approaches to solving a totally unimodular stochastic program with the application of the L-shaped method. One can either solve the master problem as an LP or an IP. It is clear that applying both

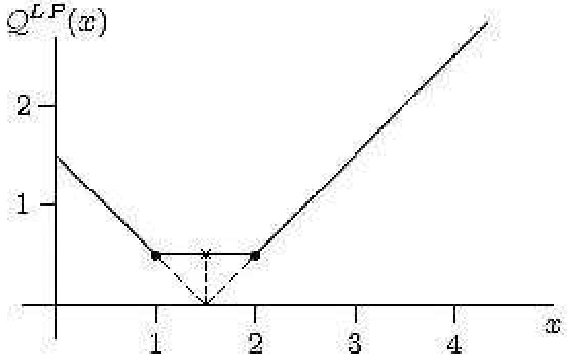

Figure 1: Recourse function for Example 2

approaches, the L-shaped method terminates to the same integer optimal solution if the solution is unique. However, if there are multiple optima, solving the master as a linear program may result in a non-integral solution, as the following example shows.

Example 2 *Suppose $K = 2$ and $T^1 = T^2 = 1$ and $W^1 = W^2 = [1 \ -1]$. Clearly, the matrix B is TU by Proposition 9. Let $c = 0$ and $d^1 = d^2 = [1\ 1]^T$. Hence,*

$$Q^{LP}(x,k) = \begin{cases} h^k - x & \textit{if } x < h^k, \\ x - h^k & \textit{if } x \geq h^k. \end{cases}$$

Suppose $h^1 = 1$ and $h^2 = 2$, each with probability $\frac{1}{2}$. Assume also $0 \leq x \leq 10$.

Because the first-stage objective $c^T x = 0$, $\mathcal{Q}(x)$ is also the function $z(x)$ to be minimized (see Figure 1). Suppose we start the L-shaped method at $x^1 = 0$ and solve the master problem as a linear program. We show the sequence of iterations for the L-shaped method as follows.

1. Iteration 1: x^1 *is not optimal, generate the cut* $\theta \geq 1.5 - x$.

2. Iteration 2: $x^2 = 10$, $\theta^2 = -8.5$ *is not optimal, generate the cut* $\theta \geq x - 1.5$.

3. Iteration 3: $x^3 = 1.5$, $\theta^3 = 0$ *is not optimal, generate the cut* $\theta \geq 0.5$.

4. Iteration 4: $x^4 = 1.5$, $\theta^4 = 0.5$, *which is optimal.*

It should be noted that with adding the last L-shaped optimality cut $\theta \geq 0.5$ to the master problem, the optimal solution x is not unique. As a result, solving the master problem as a linear program may not guarantee that the L-shaped method terminates with a feasible optimal solution to the totally unimodular stochastic program. However, as Theorem 10 shows, if an extreme point solution is obtained when solving the master problem at the last iteration of the L-shaped method, the first-stage decision must be integral.

Let P^M denote the polyhedral set of the feasible solutions (x, θ) of the master problem when the L-shaped method terminates.

Theorem 10 *Suppose the L-shaped method terminates with an optimal solution* $(\overline{x}, \overline{\theta})$ *that is an extreme point of* P^M*. Then if* B *is TU,* $\overline{x} \in \mathbb{Z}_+^{n_1}$*.*

Proof: We will prove that the solution $(\overline{x}, \overline{y})$, with $\overline{y} = y^{LP}(\overline{x})$, is an extreme point of P. The result then follows from the total unimodularity of B.

Clearly $(\overline{x}, \overline{y}) \in P$. Suppose that $(\overline{x}, \overline{y})$ is not an extreme point of P. Then there exists (x^1, y^1) and (x^2, y^2) in P with $(x^1, y^1) \neq (x^2, y^2)$ such that $\overline{x} = \frac{1}{2}x^1 + \frac{1}{2}x^2$ and $\overline{y} = \frac{1}{2}y^1 + \frac{1}{2}y^2$.

First, consider the case $x^1 = x^2 = \overline{x}$. Then y^1 and y^2 (with $y^1 \neq y^2$) are feasible solutions to the linear program defining $\mathcal{Q}^{LP}(\overline{x})$ whereas $\overline{y}$ is an extreme point of the linear program defining $\mathcal{Q}^{LP}(\overline{x})$. However, $\overline{y} = \frac{1}{2}y^1 + \frac{1}{2}y^2$, and we have a contradiction.

Now, consider the case $x^1 \neq x^2$. Define $\theta^1 = \sum_{k=1}^K p^k (d^k)^T (y^k)^1$ and $\theta^2 = \sum_{k=1}^K p^k (d^k)^T (y^k)^2$. Note that $(y^i)^T = ((y^{1i})^T, \ldots, (y^{Ki})^T)$ for $i = 1, 2$. Then $\theta^1 \geq \mathcal{Q}^{LP}(x^1)$ and $\theta^2 \geq \mathcal{Q}^{LP}(x^2)$, and from the validity of the optimality cuts it follows that $(x^1, \theta^1) \in P^M$ and $(x^2, \theta^2) \in P^M$. Moreover, $\frac{1}{2}\theta^1 + \frac{1}{2}\theta^2 = \sum_{k=1}^K p^k (d^k)^T (\frac{1}{2}(y^k)^1 + \frac{1}{2}(y^k)^2) = \sum_{k=1}^K p^k (d^k)^T \overline{y}^k = \overline{\theta}$, where the last equality follows from the fact that at termination of the L-shaped method $\overline{\theta} = \mathcal{Q}^{LP}(\overline{x}) = \sum_{k=1}^K p^k (d^k)^T \overline{y}^k$. We thus have (x^1, θ^1) and (x^2, θ^2) in P^M, with $x^1 \neq x^2$ such that $\frac{1}{2}(x^1, \theta^1) + \frac{1}{2}(x^2, \theta^2) = (\overline{x}, \overline{\theta})$. This contradicts the hypothesis that $(\overline{x}, \overline{\theta})$ is an extreme point of P^M. □

Theorem 10 shows that one only needs to obtain an extreme point solution of the master problem as the L-shaped method terminates. This is normally ensured by using the simplex method for the master problem.

Applying the L-shaped method to solve totally unimodular stochastic programs may not preserve the integrality property.

Proposition 10 *A master with at least one feasibility or optimality cut is in general not TU.*

Proof: In general, feasibility and optimality cuts contain coefficients not in $\{0, \pm 1\}$. □

Theorems 8 and 9 indicate the trade-off of solving totally unimodular stochastic programs using the L-shaped method. On the one hand, Theorem 8 shows that a minimizer can be found in $\mathbb{Z}_+^{n_1}$. On the other hand, Theorem 9 shows that by solving the master as an LP, we are minimizing $\mathcal{Q}^{LP}(x)$, a convex function, over $\mathbb{R}_+^{n_1}$. Our computational experiments explore when this knowledge can be exploited.

5. Computational Results Comparing Various L-shaped Approaches

As in the case of general stochastic linear programs with continuous second stage, the L-shaped method [19] can be used to exploit the dual block-angular structure existing in totally unimodular stochastic programs. While certain specialized versions of the L-shaped method may apply to certain classes of the problems, e.g., those with simple integer recourse [10], the purpose of our experiments is to explore the trade-off between the two approaches within the most common L-shaped method framework as described in Theorems 8 and 9.

When applying the L-shaped method to totally unimodular two-stage stochastic programs, Theorem 8 and 9 indicate the trade-off in terms of solving the master problem. When integrality restrictions in X are relaxed, the master problem is easier to solve for a given set of cuts. However, in this case fractional solutions x may be generated. Since there exists an optimal solution with x being integral, solving the master over X may be advantageous, since every solution to the master is potentially an optimal solution.

We solved four classes of totally unimodular two-stage stochastic programs. The first two classes of problems comprise instances in which T is a fixed TU matrix and W^k is $[I, -I]$ for each scenario k. The next two classes comprise instances in which T is identity matrix and W is a fixed TU matrix. Table 1 summarizes these four classes. We used the capacitated network generator NETGEN [8, 11] to generate TU matrices.

Tables 2 and 3 show several parameters used in the TU matrix generation. In each class, we generated 100 totally unimodular stochastic program instances whose generating parameters are also reported in Tables 2 and 3. In Class 2, we specified that the second-stage objective function coefficients are the only stochastic component for simplicity in the generation. For each instance, we varied the number of scenarios, i.e., $K = 1000, 2000, \ldots, 10000$, and tested two solution approaches in the L-shaped method. One approach was to solve the master problem itself as an IP at each iteration. The other one was to solve the LP-relaxation

Table 1: Characteristics of test problems

	1st-stage	2nd-stage
Class 1	MCNFP*	Simple Recourse
Class 2	Assignment	Simple Recourse
Class 3	Identity Matrix	Assignment
Class 4	Identity Matrix	Transportation

*MCNFP: Multicommodity Network Flow Problem.

Table 2: Generation of instances in Classes 1 and 2

	T matrix							SIP	
	# of	# of	Total	Arc cost		# of	# of	2nd-stage	2nd-stage
Class	nodes	arcs	supply	min	max	sources	sinks	obj	rhs
1	40	65	30	1	10	2	2	$\sim U(5, 10)$	$\sim U(0.9h_0, 1.1h_0)^*$
2	70	100	35	1	20	35	35	$\sim U(1, 10)$	1

$^*h_0$: average supply/demand generated by NETGEN.

of the master problem at each iteration.

All the instances were solved on a Pentium IV PC with a 2.40 GHz CPU and their computational results are reported in Tables 4 – 6. In Table 4, we present the average ratio of the number of cuts required by the approach of solving each master problem as an IP to the number of cuts required by the approach of solving each master problem as an LP. In Classes 1 and 2, only optimality cuts are needed since the stochastic integer programs have relatively complete recourse. Regardless of the number of scenarios, solving the master problem as an IP requires fewer cuts on average than solving its LP-relaxation. In Table 5, we present the number of integer incumbent solutions obtained in the L-shaped method with the approach of solving each master problem as an LP, for 10 instances randomly selected from Classes 3 and 4. These results show that the L-shaped method with the LP master approach converges along a much different path than the IP master approach. For some instances with given numbers of scenarios, e.g. Instance 1 from Class 3 with 1000 scenarios, it generates only one integer solution, namely the optimal solution. Since the generated fractional solutions are unlikely optimal, these results indicate that the LP master approach tends to require more iterations than the IP master approach.

Table 6 reports the percentage of instances in which the approach of solving each master

Table 3: Generation of instances in Classes 3 and 4

	W matrix							SIP	
	# of	# of	Total	Arc cost		# of	# of	1st-stage	2nd-stage
Class	nodes	arcs	supply	min	max	sources	sinks	obj	obj
3	50	90	25	1	20	25	25	$\sim U(1, 20)$	$\sim U(0.5d_0, 1.5d_0)^*$
4	80	300	500	1	10	40	40	$\sim U(1, 5)$	$\sim U(0.5d_0, 1.5d_0)^*$

$^*d_0$: average arc cost genearated by NETGEN.

Table 4: Average ratio for the number of cuts (LP master vs. IP master)

		Number of Scenarios									
Class	Cuts	1000	2000	3000	4000	5000	6000	7000	8000	9000	10000
1	opt.	1.311	1.304	1.323	1.335	1.312	1.321	1.332	1.300	1.304	1.271
2	opt.	1.701	1.695	1.675	1.647	1.681	1.718	1.689	1.653	1.709	1.715
3	fea.	1.163	1.151	1.135	1.158	1.133	1.125	1.104	1.149	1.117	1.157
	opt.	1.068	1.149	1.099	1.082	1.067	1.079	1.111	1.122	1.082	1.112
	total	1.159	1.164	1.133	1.143	1.125	1.120	1.119	1.153	1.120	1.151
4	fea.	1.058	1.083	1.057	1.049	1.063	1.049	1.052	1.049	1.064	1.072
	opt.	1.541	1.543	1.517	1.495	1.387	1.529	1.506	1.515	1.508	1.536
	total	1.333	1.347	1.326	1.299	1.227	1.326	1.314	1.323	1.321	1.340

Table 5: Integer incumbent solution in the L-shaped method with the LP master approach (# of integer incumbent solutions / # of iterations)

		Number of Scenarios									
Class	Instance	1000	2000	3000	4000	5000	6000	7000	8000	9000	10000
	1	1/61	4/64	5/57	2/47	2/63	1/55	1/68	1/68	1/59	2/61
	2	5/52	3/65	1/62	2/70	1/56	6/60	1/47	2/54	7/67	2/60
3	3	1/56	1/53	1/47	1/62	1/58	5/55	2/51	2/51	1/52	2/57
	4	1/43	1/55	1/47	1/31	2/36	1/31	6/49	1/33	4/47	1/36
	5	1/41	1/57	2/42	1/56	1/55	3/51	1/51	2/56	2/54	4/59
	1	3/185	5/165	4/178	7/175	2/152	4/159	4/177	5/163	5/164	6/163
	2	3/154	4/155	3/162	1/160	2/140	2/163	7/169	2/160	3/153	12/165
4	3	2/164	5/147	8/156	7/155	6/162	6/147	3/153	5/160	3/152	6/154
	4	1/144	6/143	3/141	5/136	5/128	2/138	1/131	2/136	5/138	1/128
	5	7/119	1/121	6/117	5/121	3/120	7/127	7/114	9/132	7/115	11/132

problem as an IP takes less time than the approach of solving each master problem as an LP. In all four classes, solving the master as an IP is more likely to be superior as the number of scenarios increases. This observation can be interpreted as follows. When imposing the integrality restrictions on the master problem, the L-shaped method tends to require fewer iterations, which compensates the extra computational burden caused by solving the IP master.

Table 6: Percentage of instances where solving IP masters is faster than solving LP masters

	Number of Scenarios									
Class	1000	2000	3000	4000	5000	6000	7000	8000	9000	10000
1	29	51	59	69	69	73	78	73	75	78
2	0	10	17	20	28	33	37	39	47	43
3	39	56	58	60	56	58	58	64	62	68
4	11	33	45	58	61	72	71	75	79	80

6. Conclusions

This paper considers two-stage stochastic programs with totally unimodular extensive-form constraint matrices. Such stochastic programs enjoy the integrality property in the sense that every extreme point solution is integral for any integer right-hand sides.

We provide sufficient conditions for a stochastic program to be totally unimodular. Our computational experiments explore under which conditions such stochastic programs should be solved with an integer master or a linear master. Our computational results indicate that when there are relatively few scenarios, solving an integer master is not worth the extra effort. However, as the number of scenarios grows solving fewer master problems becomes more important, and the IP master approach is more effective.

References

[1] S. Ahmed, M. Tawarmalani, and N. V. Sahinidis. A finite branch-and-bound algorithm for two-stage stochastic integer programs. *Mathematical Programming*, 100(2):355–377, 2004.

[2] J. R. Birge and F. V. Louveaux. *Introduction to Stochastic Programming.* Springer, New York, New York, 1997.

[3] C. C. Carøe. *Decomposition in stochastic integer programming.* PhD thesis, University of Copenhagen, 1998.

[4] G. B. Dantzig. Linear programming under uncertainty. *Management Science*, 1:197–206, 1955.

[5] A. Ghouila-Houri. Caracterisation des matrices totalement unimodulaires. *C. R. Academy of Sciences of Paris*, 254:1192–1194, 1962.

[6] A. J. Hoffman and J. B. Kruskal. Integral boundary points of convex polyhedra. In H. W. Kuhn and A. W. Tucker, editors, *Linear Inequalities and Related Systems*, pages 223–246. Princeton University Press, 1956.

[7] W. K. Klein Haneveld and M. H. van der Vlerk. Stochastic integer programming: General models and algorithms. *Annals of Operations Research*, 85:39–57, 1999.

[8] D. Klingman, A. Napier, and J. Stutz. NETGEN: A program for generating large scale capacitated assignment, transportation, and minimum cost network flow problems. *Management Science*, 20:814–821, 1974.

[9] N. Kong, A. J. Schaefer, and B. K. Hunsaker. Two-stage integer programs with stochastic right-hand sides: A superadditive dual approach. *To appear in Mathematical Programming*, 2005. Available from Stochastic Programming E-Print Series (http://www.speps.info).

[10] F. V. Louveaux and M.H. van der Vlerk. Stochastic programming with simple integer recourse. *Mathematical Programming*, 61(3):301–325, 1993.

[11] MP-Testdata. Capacitated network generator Netgen. Available from http://elib.zib.de/pub/packages/mp-testdata/generator/netgen/.

[12] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization*. Wiley, 1988.

[13] M. W. Padberg. Characterizations of totally unimodular, balanced and perfect matrices. In B. Roy, editor, *Combinatorial Programming: Methods and Applications*, pages 275–284. Reidel, 1975.

[14] A. Schrijver. *Linear and Integer Programming*. Wiley, 1986.

[15] R. Schultz. Stochastic programming with integer variables. *Mathematical Programming*, 97(1-2):285–309, 2003.

[16] S. Sen and J. L. Higle. The C^3 theorem and a D^2 algorithm for large scale stochastic integer programming: Set convexification. *Stochastic Programming E-Print Series*, 2000. Available from http://www.speps.info.

[17] L. Stougie. *Design and analysis of methods for stochastic integer programming*. PhD thesis, University of Amsterdam, 1987.

[18] M. H. van der Vlerk. Convex approximations for complete integer recourse models. *Mathematical Programming*, 99(2):297–310, 2004.

[19] R. Van Slyke and R. J-B. Wets. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, 17:638–663, 1969.

[20] R. Wollmer. Two stage linear programming under uncertainty with 0-1 integer first stage variables. *Mathematical Programming*, 19(3):279–288, 1980.