An Integer Programming Approach for Linear Programs with Probabilistic Constraints *†

James Luedtke
Shabbir Ahmed
George Nemhauser
Georgia Institute of Technology
765 Ferst Drive, Atlanta, GA, USA
luedtke@gatech.edu

May 3, 2007

Abstract

Linear programs with joint probabilistic constraints (PCLP) are difficult to solve because the feasible region is not convex. We consider a special case of PCLP in which only the right-hand side is random and this random vector has a finite distribution. We give a mixed-integer programming formulation for this special case and study the relaxation corresponding to a single row of the probabilistic constraint. We obtain two strengthened formulations. As a byproduct of this analysis, we obtain new results for the previously studied mixing set, subject to an additional knapsack inequality. We present computational results which indicate that by using our strengthened formulations, instances that are considerably larger than have been considered before can be solved to optimality.

Keywords: Stochastic Programming, Integer Programming, Probabilistic Constraints, Chance Constraints, Mixing Set

1 Introduction

Consider a linear program with a probabilistic or chance constraint

$$(PCLP) \quad \min\Big\{cx : x \in X,\ P\{\tilde{T}x \geq \xi\} \geq 1-\epsilon\Big\} \tag{1}$$

*This research has been supported in part by the National Science Foundation under grants DMI-0121495 and DMI-0522485.

†An extended abstract of this paper has been published in the Proceedings of the Twelfth Conference on Integer Programming and Combinatorial Optimization (IPCO 2007) [13].

where $X = \{x \in \mathbb{R}^d_+ : Ax = b\}$ is a polyhedron, $c \in \mathbb{R}^d$, $\tilde{T}$ is an $m \times d$ random matrix, ξ is a random vector taking values in $\mathbb{R}^m$, and ϵ is a confidence parameter chosen by the decision maker, typically near zero, e.g., $\epsilon = 0.01$ or $\epsilon = 0.05$. Note that in (1) we enforce a single probabilistic constraint over *all* rows, rather than requiring that each row independently be satisfied with high probability. Such a constraint is known as a *joint probabilistic constraint*, and is appropriate in a context in which it is important to have all constraints satisfied simultaneously and there may be dependence between random variables in different rows.

Problems with joint probabilistic constraints have been extensively studied; see [17] for background and an extensive list of references. Probabilistic constraints have been used in various applications including supply chain management [11], production planning [15], optimization of chemical processes [9, 10] and surface water quality management [20]. Unfortunately, linear programs with probabilistic constraints are still largely intractable except for a few very special cases. There are two primary reasons for this intractability. First, in general, for a given $x \in X$, the quantity $P\{\tilde{T}x \geq \xi\}$ is hard to compute, as it requires multi-dimensional integration. Second, the feasible region defined by a probabilistic constraint generally is not convex.

In this work, we demonstrate that by using integer programming techniques, instances of PCLP that are considerably larger than have been considered before can be solved to optimality under the following two simplifying assumptions:

(A1) Only the right-hand side vector ξ is random; the matrix $\tilde{T} = T$ is deterministic.

(A2) The random vector ξ has a finite distribution.

Despite its restrictiveness, the special case given by assumption A1 has received considerable attention in the literature, see, e.g., [5, 6, 17]. A notable result for this case is that if the distribution of the right-hand side is log-concave, then the feasible region defined by the joint probabilistic constraint is convex [16]. This allows problems in which the dimension, m, of the random vector is small to be solved to optimality, but higher dimensional problems are still intractable due to the difficulty in checking feasibility of the probabilistic constraint. Specialized methods have been developed in [4, 6, 17] for the case in which assumption A1 holds and the random vector has discrete but not necessarily finite distribution. These methods rely on the enumeration of certain efficient points of the distribution, and hence do not scale well with m since the number of efficient points generally grows exponentially with m.

Assumption A2 may also seem very restrictive. However, if the possible values for ξ are generated by taking Monte Carlo samples from a general distribution, we can think of the resulting problem as an approximation of the problem with general distribution. Under some reasonable assumptions we can show that the optimal solution of the sampled problem converges exponentially fast to the optimal solution of the original problem as the number of scenarios increases. Also, the optimal objective of the sampled problem can be used to develop statistical lower bounds on the optimal objective of the original problem. See [1, 3, 19] for some related results. It seems that the reason such a sampling approach has not been seriously considered for PCLP in the past is that the resulting sampled problem has a non-convex feasible region, and thus is generally intractable. Our contribution is to demonstrate that, at least under assumption A1, it is nonetheless possible to solve the sampled problem.

Under assumption A2 it is possible to write a mixed-integer programming (MIP) formulation for PCLP, as has been done, for example, in [18]. In the general case, such a formulation requires the introduction of "big-M" type constraints, and hence is difficult to solve. However, by restricting attention to the case of assumption A1, we are able to develop strong mixed-integer programming formulations. Our approach in developing these formulations is to consider the relaxation obtained from a single row in the probabilistic constraint. This yields a system similar to the *mixing set* introduced by Günlük and Pochet [8], subject to an additional knapsack inequality. We are able to derive strong valid inequalities for this system by first using the knapsack inequality to "preprocess" the mixing set and then applying the mixing inequalities of [8]; see also [2, 7]. We also derive an extended formulation, equivalent to one given by Miller and Wolsey in [14]. Making further use of the knapsack inequality, we are able to derive more general classes of valid inequalities for both the original and extended formulations. If all scenarios are equally likely, the knapsack inequality reduces to a cardinality restriction. In this case, we are able to characterize the convex hull of feasible solutions to the extended formulation for the single row case. Although these results are motivated by the application to PCLP, they can be used in any problem in which a mixing set appears along with a knapsack constraint.

The remainder of this paper is organized as follows. In Section 2 we verify that PCLP remains NP-hard even under assumptions A1 and A2, and present the standard MIP formulation. In Section 3 we analyze this MIP and present classes of valid inequalities that make the formulation strong. In Section 4 we present an extended formulation, and a new class of valid inequalities and show that in the equi-probable scenarios case, these inequalities define the convex hull of the single row formulation. In Section 5 we present computational results using the strengthened formulations, and we close with concluding remarks in Section 6.

2 The MIP Formulation

We now consider a probabilistically constrained linear programming problem with random right-hand side given by

$$(PCLPR) \quad \begin{array}{ll} \min & cx \\ \text{s.t.} & Ax = b \\ & P\{Tx \geq \xi\} \geq 1 - \epsilon \\ & x \geq 0. \end{array} \tag{2}$$

Here A is an $r \times d$ matrix, $b \in \mathbb{R}^r$, T is an $m \times d$ matrix, ξ is a random vector in $\mathbb{R}^m$, $\epsilon \in (0,1)$ (typically small) and $c \in \mathbb{R}^d$. We assume that ξ has finite support, that is there exist vectors, $\xi_i \in \mathbb{R}^m, i = 1, \ldots, n$ such that $P\{\xi = \xi_i\} = \pi_i$ for each i where $\pi_i > 0$ and $\sum_{i=1}^n \pi_i = 1$. We refer to the possible outcomes as scenarios. We assume without loss of generality that $\xi_i \geq 0$ and $\pi_i \leq \epsilon$ for each i. We also define the set $N = \{1, \ldots, n\}$.

Before proceeding, we note that PCLPR is NP-hard even under assumptions A1 and A2.

Theorem 1. *PCLPR is NP-hard, even in the special case in which $\pi_i = 1/n$ for all $i \in N$, the constraints $Ax = b$ are not present, T is the $m \times m$ identity matrix, and $c = (1, \ldots, 1) \in \mathbb{R}^m$.*

Proof. Let $K = \lceil (1-\epsilon)n \rceil$. Then, under the stated conditions PCLPR can be written as

$$\min_{I \subseteq N} \left\{ \sum_{j=1}^{m} \max_{i \in I} \{\xi_{ij}\} : |I| \geq K \right\}.$$

We show that the associated decision problem:

> (DPCLP) Given non-negative integers ξ_{ij} for $i = 1, \ldots, n$, $j = 1, \ldots, m$, $K \leq n$ and B, is there an $I \subseteq N$ such that $|I| \geq K$ and $\sum_{j=1}^m \max_{i \in I} \{\xi_{ij}\} \leq B$?

is NP-complete by reduction from the NP-complete problem CLIQUE. Consider an instance of CLIQUE given by graph $G = (V, E)$, in which we wish to decide whether there exists a clique of size C. We construct an instance of DPCLP by letting $\{1, \ldots, m\} = V$, $N = E$, $B = C$, $K = C(C-1)/2$ and $\xi_{ij} = 1$ if edge i is incident to node j and $\xi_{ij} = 0$ otherwise. The key observation is that for any $I \subseteq E$, and $j \in V$,

$$\max_{i \in I} \{\xi_{ij}\} = \begin{cases} 1 & \text{if some edge } i \in I \text{ is incident to node } j \\ 0 & \text{otherwise.} \end{cases}$$

Hence, if there exists a clique of size C in G then we have a subgraph of G consisting of C nodes and $C(C-1)/2$ edges. Thus there exists $I \subseteq N$ with $|I| = C(C-1)/2 = K$ and

$$\sum_{j=1}^{m} \max_{i \in I} \{\xi_{ij}\} = C = B$$

and the answer to DPCLP is yes.

Conversely, if the answer to DPCLP is yes, there exists $I \subseteq E$ of size at least $K = C(C-1)/2$ such that the number of nodes incident to I is at most $B = C$. This can only happen if I defines a clique of size C. □

We now formulate PCLPR as a mixed-integer program [18]. To do so, we introduce for each $i \in N$, a binary variable z_i, where $z_i = 0$ guarantees that

$Tx \geq \xi_i$. Observe that because $\epsilon < 1$ we must have $Tx \geq \xi_i$ for at least one $i \in N$, and because $\xi_i \geq 0$ for all i, this implies $Tx \geq 0$ in every feasible solution of PCLPR. Then, letting $v = Tx$, we obtain the MIP formulation of PCLPR

$$\begin{aligned}
(PMIP) \quad \min \quad & cx \\
\text{s.t.} \quad & Ax = b, \ Tx - v = 0 && (3) \\
& v + \xi_i z_i \geq \xi_i \qquad i = 1, \ldots, n && (4) \\
& \sum_{i=1}^{n} \pi_i z_i \leq \epsilon && (5) \\
& x \geq 0, \quad z \in \{0,1\}^n
\end{aligned}$$

where (5) is equivalent to the probabilistic constraint

$$\sum_{i=1}^{n} \pi_i (1 - z_i) \geq 1 - \epsilon.$$

3 Strengthening the Formulation

We begin by considering how the formulation PMIP can be strengthened when the probabilities π_i are general. In Section 3.2 we present results specialized to the case when all π_i are equal.

3.1 General Probabilities

Our approach is to strengthen PMIP by ignoring (3) and finding strong formulations for the set

$$F := \left\{ (v, z) \in \mathbb{R}_+^m \times \{0,1\}^n : (4), (5) \right\}. \tag{6}$$

Note that

$$F = \bigcap_{j=1}^{m} \{(v, z) : (v_j, z) \in G_j\},$$

where for $j = 1, \ldots, m$

$$G_j = \{(v_j, z) \in \mathbb{R}_+ \times \{0,1\}^n : (5), \quad v_j + \xi_{ij} z_i \geq \xi_{ij} \quad i = 1, \ldots, n\}.$$

Thus, a natural first step in developing a strong formulation for F is to develop a strong formulation for each G_j. In particular, note that if an inequality is facet-defining for $\text{conv}(G_j)$, then it is also facet-defining for $\text{conv}(F)$. This follows because if an inequality valid for G_j is supported by $n+1$ affinely independent points in $\mathbb{R}^{n+1}$, then because this inequality will not have coefficients on v_i for any $i \neq j$, the set of supporting points can trivially be extended to a set of $n+m$ affinely independent supporting points in $\mathbb{R}^{n+m}$ by appropriately setting the v_i values for each $i \neq j$.

The above discussion leads us to consider the generic set

$$G = \{(y, z) \in \mathbb{R}_+ \times \{0,1\}^n : (5), \quad y + h_i z_i \geq h_i \quad i = 1, \dots, n\} \tag{7}$$

obtained by dropping the index j and setting $y = v_j$ and $h_i = \xi_{ij}$ for each i. We assume without loss of generality that $h_1 \geq h_2 \geq \cdots \geq h_n$. The relaxation of G obtained by dropping (5) is a *mixing set* given by

$$P = \{(y, z) \in \mathbb{R}_+ \times \{0,1\}^n : y + h_i z_i \geq h_i \quad i = 1, \dots, n\}.$$

This set has been extensively studied, in varying degrees of generality, by Atamtürk et. al [2], Günlük and Pochet [8], Guan et. al [7] and Miller and Wolsey [14]. The *star inequalities* of [2] given by

$$y + \sum_{j=1}^{l} (h_{t_j} - h_{t_{j+1}}) z_{t_j} \geq h_{t_1} \quad \forall T = \{t_1, \dots, t_l\} \subseteq N, \tag{8}$$

where $t_1 < t_2 < \cdots < t_l$ and $h_{t_{l+1}} := 0$ are valid for P. Furthermore, these inequalities can be separated in polynomial time, are facet-defining for P when $t_1 = 1$, and are sufficient to define the convex hull of P [2, 7, 8].

We can tighten these inequalities for G by using the knapsack constraint (5). In particular, let $p := \max\{k : \sum_{i=1}^{k} \pi_i \leq \epsilon\}$. Then, from the knapsack constraint, we cannot have $z_i = 1$ for all $i = 1, \dots, p+1$ and thus we have $y \geq h_{p+1}$. This also implies that the mixed-integer constraints in G are redundant for $i = p+1, \dots, n$. Thus, we can replace the inequalities $y + h_i z_i \geq h_i$ for $i = 1, \dots, n$ in the definition of G by the inequalities

$$y + (h_i - h_{p+1}) z_i \geq h_i \quad i = 1, \dots, p. \tag{9}$$

That is, we have

$$G = \{(y, z) \in \mathbb{R}_+ \times \{0,1\}^n : (5), (9)\}. \tag{10}$$

In addition to yielding a tighter relaxation, the description (10) of G is also more compact. In typical applications, ϵ is near 0, suggesting $p << n$. When applied for each j in the set F, if p is the same for all rows, this would yield a formulation with $mp << mn$ rows.

By applying the star inequalities to (10) we obtain

Theorem 2. *The inequalities*

$$y + \sum_{j=1}^{l} (h_{t_j} - h_{t_{j+1}}) z_{t_j} \geq h_{t_1} \quad \forall T = \{t_1, \dots, t_l\} \subseteq \{1, \dots, p\} \tag{11}$$

with $t_1 < \ldots < t_l$ and $h_{t_{l+1}} := h_{p+1}$, are valid for G. Moreover, (11) is facet-defining for conv(G) *if and only if $h_{t_1} = h_1$.*

Proof. The result follows directly from Proposition 3.4 and Theorem 3.5 of [2] after appropriate reformulation. See also [7, 8]. However, since our formulation differs somewhat, we give a self-contained proof. To prove (11) is valid, let $(y,z) \in G$ and let $j^* = \min\{j \in \{1,\dots,l\} : z_{t_j} = 0\}$. Then $y \geq h_{t_{j^*}}$. Thus,

$$y + \sum_{j=1}^{l} (h_{t_j} - h_{t_{j+1}}) z_{t_j} \geq h_{t_{j^*}} + \sum_{j=1}^{j^*-1} (h_{t_j} - h_{t_{j+1}}) = h_{t_1}.$$

If $h_{t_1} < h_1$, then a stronger inequality can be obtained by including index 1 in the set T, proving that this is a necessary condition for (11) to be facet-defining. Consider the following set of points: $(h_1, e_i), i \in N \setminus T$, $(h_i, \sum_{j=1}^{i-1} e_j), i \in T$ and $(h_{p+1}, \sum_{j=1}^{p} e_j)$, where e_j is the j^{th} unit vector in $\mathbb{R}^n$. It is straightforward to verify that these $n+1$ feasible points satisfy (11) at equality and are affinely independent, completing the proof. □

We refer to the inequalities (11) as the *strengthened star inequalities*. Because the strengthened star inequalities are just the star inequalities applied to a strengthened mixing set, separation can be accomplished using an algorithm for separation of star inequalities [2, 7, 8].

3.2 Equal Probabilities

We now consider the case in which $\pi_i = 1/n$ for all $i \in N$. Thus $p = \max\{k : \sum_{i=1}^{k} 1/n \leq \epsilon\} = \lfloor n\epsilon \rfloor$ and the knapsack constraint (5) becomes

$$\sum_{i=1}^{n} z_i \leq n\epsilon$$

which, by integrality on z_i, can be strengthened to the simple cardinality restriction

$$\sum_{i=1}^{n} z_i \leq p. \tag{12}$$

Thus, the feasible region (10) becomes

$$G' = \{(y,z) \in \mathbb{R}_+ \times \{0,1\}^n : (9), (12)\}.$$

Although the strengthened star inequalities are not sufficient to characterize the convex hull of G', we now show that it is possible to separate over $\text{conv}(G')$ in polynomial time. To obtain this result we first show that for any $(\gamma, \alpha) \in \mathbb{R}^{n+1}$, the problem

$$\min\{\gamma y + \alpha z : (y,z) \in G'\} \tag{13}$$

is easy to solve. For $k = 1, \dots, p$ let

$$\mathcal{S}_k = \{S \subseteq \{k, \dots, n\} : |S| \leq p - k + 1\}$$

and

$$S_k^* \in \arg\min_{S \in \mathcal{S}_k} \left\{ \sum_{i \in S} \alpha_i \right\}.$$

Also, let $S_{p+1}^* = \emptyset$ and $k^* \in \arg\min\{\gamma h_k + \sum_{i \in S_k^*} \alpha_i : k = 1, \ldots, p+1\}$.

Lemma 3. *If $\gamma < 0$, then (13) is unbounded. Otherwise, an optimal solution to (13) is given by $y = h_{k^*}$ and $z_i = 1$ for $i \in S_{k^*}^* \cup \{1, \ldots, k^* - 1\}$ and $z_i = 0$ otherwise.*

Proof. Problem (13) is unbounded when $\gamma < 0$ because $(1, \mathbf{0})$ is a feasible direction for G'. Now suppose $\gamma \geq 0$. We consider all feasible values of y, $y \geq h_{p+1}$. First, if $y \geq h_1$, then the z_i can be set to any values satisfying (12), and hence it would yield the minimum objective to set $z_i = 1$ if and only if $i \in S_1^*$ and to set $y = h_1$ since $\gamma \geq 0$. For any $k \in \{2, \ldots, p+1\}$, if $h_{k-1} > y \geq h_k$ then we must set $z_i = 1$ for $i = 1, \ldots, k-1$. The minimum objective in this case is then obtained by setting $y = h_k$ and $z_i = 1$ for $i = 1, \ldots, k-1$ and $i \in S_k^*$. The optimal solution to (13) is then obtained by considering y in each of these ranges. □

Using Lemma 3, we can optimize over G' by first sorting the values of α_i in increasing order, then finding the sets S_k^* by considering at most $p - k + 1$ of the smallest values in this list for each $k = 1, \ldots, p+1$. Subsequently finding the index k^* yields an optimal solution defined by Lemma 3. This yields an obvioius algorithm with complexity $O(n \log n + p^2) = O(n^2)$. It follows that we can separate over $\mathrm{conv}(G')$ in polynomial time. We begin by characterizing the set of valid inequalities for G'.

Theorem 4. *Any valid inequality for G' with nonzero coefficient on y can be written in the form*

$$y \geq \beta + \sum_{i=1}^{n} \alpha_i z_i. \tag{14}$$

Furthermore, (14) is valid for G' if and only if there exists (σ, ρ) such that

$$\beta + \sum_{i=1}^{k-1} \alpha_i + (p-k+1)\sigma_k + \sum_{i=k}^{n} \rho_{ik} \leq h_k \quad k = 1, \ldots, p+1 \tag{15}$$

$$\alpha_i - \sigma_k - \rho_{ik} \leq 0 \qquad i = k, \ldots, n, \; k = 1, \ldots, p+1 \tag{16}$$

$$\sigma \geq 0, \rho \geq 0. \tag{17}$$

Proof. First consider a generic inequality of the form $\gamma y \geq \beta + \sum_{i=1}^{n} \alpha_i z_i$. Since $(1, \mathbf{0})$ is a feasible direction for G', we know this inequality is valid for G' only if $\gamma \geq 0$. Thus, if a valid inequality for G' has nonzero coefficient γ on y, then $\gamma > 0$, and so we can scale the inequality such that $\gamma = 1$, thus obtaining the form (14). Now, since any extreme point of $\mathrm{conv}(G')$ is an optimal solution to (13) for some $(\gamma', \alpha') \in \mathbb{R}_{n+1}$, we know by Lemma 3 that the extreme points of $\mathrm{conv}(G')$ are contained in the set of feasible points given by $y = h_k$, $z_i = 1$

for $i = 1, \ldots, k-1$ and $i \in S$, and $z_i = 0$ otherwise, for all $S \in \mathcal{S}_k$ and $k = 1, \ldots, p+1$. This fact, combined with the fact that $(1, \mathbf{0})$ is the *only* feasible direction for G', implies inequality (14) is valid for G' if and only if

$$\beta + \sum_{i=1}^{k-1} \alpha_i + \max_{S \in \mathcal{S}_k} \sum_{i \in S} \alpha_i \leq h_k \quad k = 1, \ldots, p+1. \tag{18}$$

Note that

$$\begin{aligned} \max_{S \in \mathcal{S}_k} \sum_{i \in S} \alpha_i \quad = \quad & \max_{\omega} \quad \sum_{i=k}^{n} \omega_{ik} \alpha_i \\ & \text{s.t.} \quad \sum_{i=k}^{n} \omega_{ik} \leq p-k+1 \qquad\qquad (19) \\ & \qquad 0 \leq \omega_{ik} \leq 1 \qquad\qquad i = k, \ldots, n \\ = \quad & \min_{\sigma, \rho} \quad (p-k+1)\sigma_k + \sum_{i=k}^{n} \rho_{ik} \\ & \text{s.t.} \quad \sigma_k + \rho_{ik} \geq \alpha_i \qquad\qquad i = k, \ldots, n \qquad (20) \\ & \qquad \sigma_k \geq 0, \rho_{ik} \geq 0 \qquad\qquad i = k, \ldots, n \end{aligned}$$

by linear programming duality since (19) is feasible and bounded and its optimal solution is integral. It follows that condition (18) is satisfied and hence (14) is valid for G' if and only if there exists (σ, ρ) such that the system (15) - (17) is satisfied. □

Using Theorem 4 we can separate over $\text{conv}(G')$ by solving a polynomial size linear program.

Corollary 5. *Suppose* (y^*, z^*) *satisfy* $z^* \in Z := \{z \in [0,1]^n : \sum_{i=1}^n z_i \leq p\}$. *Then,* $(y^*, z^*) \in \text{conv}(G')$ *if and only if*

$$y^* \geq LP^* = \max_{\alpha, \beta, \sigma, \rho} \left\{ \beta + \sum_{i=1}^{n} \alpha_i z_i^* : (15) - (17) \right\} \tag{21}$$

where LP^* *exists and is finite. Furthermore, if* $y^* < LP^*$ *and* (α^*, β^*) *is optimal to* (21)*, then* $y \geq \beta^* + \sum_{i=1}^n \alpha^* z_i$ *is a valid inequality for* G' *which is violated by* (y^*, z^*).

Proof. By Theorem (4), if $y^* \geq LP^*$, then (y^*, z^*) satisfies all valid inequalities for G' which have nonzero coefficient on y. But, $z^* \in Z$ and integrality of the set Z imply (y^*, z^*) also satisfies all valid inequalities which have a zero coefficient on y, showing that $(y^*, z^*) \in \text{conv}(G')$. Conversely, if $y^* < LP^*$, then the optimal solution to (21) defines a valid inequality of the form (14) which is violated by (y^*, z^*).

We next argue that (21) has an optimal solution. First note that it is feasible since we can set $\beta = h_{p+1}$ and all other variables to zero. Next, because $z^* \in Z$, and Z is an integral polytope, we know there exists sets S_j, $j \in J$ for some finite index set J, and $\lambda \in \mathbb{R}_+^{|J|}$ such that $\sum_{j\in J} \lambda_j = 1$ and $z^* = \sum_{j\in J} \lambda_j z^j$ where $z_i^j = 1$ if $i \in S_j$ and 0 otherwise. Hence,

$$\beta + \sum_{i=1}^{n} \alpha_i z_i^* = \beta + \sum_{j\in J} \lambda_j \sum_{i\in S_j} \alpha_i = \sum_{j\in J} \lambda_j (\beta + \sum_{i\in S_j} \alpha_i) \leq \sum_{j\in J} \lambda_j h_1 = h_1$$

where the inequality follows from (18) for $k = 1$ which is satisfied whenever $(\alpha, \beta, \sigma, \rho)$ is feasible to (15) - (17). Thus, the objective is bounded, and so (21) has an optimal solution. □

Although (21) yields a theoretically efficient way to separate over $\text{conv}(G')$, it still may be too expensive to solve a linear program to generate cuts. We would therefore prefer to have an explicit characterization of a class or classes of valid inequalities for G' with an associated combinatorial algorithm for separation. The following theorem gives an example of one such class, which generalizes the strengthened star inequalities.

Theorem 6. *Let $m \in \{1, \ldots, p\}$, $T = \{t_1, \ldots, t_l\} \subseteq \{1, \ldots, m\}$ and $Q = \{q_1, \ldots, q_{p-m}\} \subseteq \{p+1, \ldots, n\}$. For $m < p$, define $\Delta_1^m = h_{m+1} - h_{m+2}$ and*

$$\Delta_i^m = \max \left\{ \Delta_{i-1}^m, h_{m+1} - h_{m+i+1} - \sum_{j=1}^{i-1} \Delta_j^m \right\} \quad i = 2, \ldots, p-m.$$

Then, with $h_{t_{l+1}} := h_{m+1}$,

$$y + \sum_{j=1}^{l} (h_{t_j} - h_{t_{j+1}}) z_{t_j} + \sum_{j=1}^{p-m} \Delta_j^m (1 - z_{q_j}) \geq h_{t_1} \tag{22}$$

is valid for G'.

Proof. First note that if $m = p$, we recover the strengthened star inequalities. Now, let $m < p$ and T, Q satisfy the conditions of the theorem and let $(y, z) \in G'$ and $S = \{i \in N : z_i = 1\}$. Suppose first there exists $t_j \in T \setminus S$ and let $j^* = \min\{j \in \{1, \ldots, l\} : t_j \notin S\}$. Then, $z_{t_{j^*}} = 0$ and so $y \geq h_{t_{j^*}}$. Hence,

$$\begin{aligned} y + \sum_{j=1}^{l} (h_{t_j} - h_{t_{j+1}}) z_{t_j} &\geq h_{t_{j^*}} + \sum_{j=1}^{j^*-1} (h_{t_j} - h_{t_{j+1}}) \\ &= h_{t_1} \geq h_{t_1} - \sum_{j=1}^{p-m} \Delta_j^m (1 - z_{q_j}) \end{aligned}$$

since $\Delta_j^m \geq 0$ for all j.

Next, suppose $T \subseteq S$. Now let $k = \sum_{i\in Q} (1 - z_i)$ so that, because $|Q| = p-m$, $0 \leq k \leq p-m$ and $\sum_{i\in Q} z_i = p-m-k$. Because $Q \subseteq \{p+1, \ldots, n\}$, we know

$\sum_{i=1}^{p} z_i + \sum_{i \in Q} z_j \leq p$ and hence $\sum_{i=1}^{p} z_i \leq k+m$. It follows that $y \geq h_{k+m+1}$. Next, note that by definition, $\Delta_1^m \leq \Delta_2^m \leq \cdots \Delta_{p-m}^m$. Thus

$$\begin{aligned}
\sum_{j=1}^{p-m} \Delta_j^m (1 - z_{q_j}) \geq \sum_{j=1}^{k} \Delta_j^m = \Delta_k^m + \sum_{j=1}^{k-1} \Delta_j^m \\
\geq (h_{m+1} - h_{m+k+1} - \sum_{j=1}^{k-1} \Delta_j^m) + \sum_{j=1}^{k-1} \Delta_j^m \\
= h_{m+1} - h_{m+k+1}. \qquad (23)
\end{aligned}$$

Using (23), $y \geq h_{k+m+1}$ and the fact that $T \subseteq S$ we have

$$\begin{aligned}
y + \sum_{j=1}^{l} (h_{t_j} - h_{t_{j+1}}) z_{t_j} \geq h_{k+m+1} + \sum_{j=1}^{l} (h_{t_j} - h_{t_{j+1}}) \\
= h_{k+m+1} + h_{t_1} - h_{m+1} \geq h_{t_1} - \sum_{j=1}^{p-m} \Delta_j^m (1 - z_{q_j})
\end{aligned}$$

completing the proof. □

In [12] it is shown that the inequalities given by (22) are facet-defining for $\mathrm{conv}(G')$ when $t_1 = 1$.

Example 1. Let $n = 10$ and $\epsilon = 0.4$ so that $p = 4$ and suppose $h_{1-5} = \{20, 18, 14, 11, 6\}$. The formulation of G' for this example is

$$\begin{aligned}
y + 14z_1 \geq 20 \\
y + 12z_2 \geq 18 \\
y + 8z_3 \geq 14 \\
y + 5z_4 \geq 11 \\
\sum_{i=1}^{10} z_i \leq 4, \quad z_i \in \{0,1\} \quad i = 1, \ldots, 10.
\end{aligned}$$

Let $m = 2, T = \{1, 2\}$ and $Q = \{5, 6\}$. Then, $\Delta_1^2 = 3$ and $\Delta_2^2 = \max\{3, 8-3\} = 5$ so that (22) yields

$$y + 2z_1 + 4z_2 + 3(1 - z_5) + 5(1 - z_6) \geq 20.$$ □

Separation of inequalities (22) can be accomplished by a simple modification to the routine for separating the strengthened star inequalities. We have also identified other classes of valid inequalities [12], but have not yet been able to find a general class that characterizes the convex hull of G'.

4 A Strong Extended Formulation

4.1 General Probabilities

Let

$$FS = \{(y,z) \in \mathbb{R}_+ \times [0,1]^n : (5),(11)\}.$$

FS represents the polyhedral relaxation of G, augmented with the strengthened star inequalities. Note that the inequalities (9) are included in FS by taking $T = \{i\}$, so that enforcing integrality in FS would yield a valid formulation for the set G. Our aim is to develop a reasonably compact extended formulation which is equivalent to FS. To do so, we introduce variables $w_1, \ldots, w_p$ and let

$$EG = \Big\{(y,z,w) \in \mathbb{R}_+ \times \{0,1\}^{n+p} : (24) - (27)\Big\}$$

where

$$w_i - w_{i+1} \geq 0 \quad i = 1,\ldots,p \tag{24}$$

$$z_i - w_i \geq 0 \quad i = 1,\ldots,p \tag{25}$$

$$y + \sum_{i=1}^{p}(h_i - h_{i+1})w_i \geq h_1 \tag{26}$$

$$\sum_{i=1}^{n} \pi_i z_i \leq \epsilon \tag{27}$$

and $w_{p+1} := 0$. The variables w_i can be interpreted as deciding whether or not scenario i is satisfied for the single row under consideration. The motivation for introducing these variables is that because they are specific to the single row under consideration, the ordering on the h_i values implies that the inequalities (24) can be safely added. Note that this is not the case for the original variables z_i for $i \in N$ since they are common to all rows in the formulation. The inequalities (25) ensure that if a scenario is infeasible for the single row under consideration, then it is infeasible overall. Because of the inequalities (24), the inequalities (9) used in the description (10) of G can be replaced by the single inequality (26). We now show that EG is a valid formulation for G.

Theorem 7. $\mathrm{Proj}_{(y,z)}(EG) = G$.

Proof. First, suppose $(y,z,w) \in EG$. Let $l \in \{1,\ldots,p+1\}$ be such that $w_i = 1,\ i = 1,\ldots,l-1$ and $w_i = 0,\ i = l,\ldots,p$. Then, $y \geq h_1 - (h_1 - h_l) = h_l$. For $i = 1,\ldots,l-1$ we have also $z_i = 1$ and hence,

$$y + (h_i - h_{p+1})z_i \geq h_l + (h_i - h_{p+1}) \geq h_i$$

and for $i = l,\ldots,n$ we have $y + (h_i - h_{p+1})z_i \geq h_l \geq h_i$ which establishes that $(y,z) \in G$. Now, let $(y,z) \in G$ and let $l = \min\{i : z_i = 0\}$. Then, $y + (h_l - h_{p+1})z_l = y \geq h_l$. Let $w_i = 1,\ i = 1,\ldots,l-1$ and $w_i = 0,\ i = l,\ldots,p$. Then, $z_i \geq w_i$ for $i = 1,\ldots,p$, w_i are non-increasing, and $y \geq h_l = h_1 - \sum_{i=1}^{p}(h_i - h_{i+1})w_i$ which establishes $(y,z,w) \in EG$. □

An interesting result is that the linear relaxation of this extended formulation is as strong as having all strengthened star inequalities in the original formulation. A similar type of result has been proved in [14]. Let EF be the polyhedron obtained by relaxing integrality in EG.

Theorem 8. $\mathrm{Proj}_{(y,z)}(EF) = FS$.

Proof. First suppose $(y, z) \in FS$. We show there exists $w \in \mathbb{R}^p_+$ such that $(y, z, w) \in EF$. For $i = 1, \dots, p$ let $w_i = \min\{z_j : j = 1, \dots, i\}$. By definition, $1 \geq w_1 \geq w_2 \geq \cdots w_p \geq 0$ and $z_i \geq w_i$ for $i = 1, \dots, p$. Next, let $T := \{i = 1, \dots, p : w_i = z_i\} = \{t_1, \dots, t_l\}$, say. By construction, we have $w_i = w_{t_j}$ for $i = t_j, \dots, t_{j+1} - 1$, $j = 1, \dots, l$ $(t_{p+1} := p + 1)$. Thus,

$$\sum_{i=1}^{p}(h_i - h_{i+1})w_i = \sum_{j=1}^{l}(h_{t_j} - h_{t_{j+1}})w_{t_j} = \sum_{j=1}^{l}(h_{t_j} - h_{t_{j+1}})z_{t_j}$$

implying that $y + \sum_{i=1}^{p}(h_i - h_{i+1})w_i \geq h_1$ as desired.

Now suppose $(y, z, w) \in EF$. Let $T = \{t_1, \dots, t_l\} \subseteq \{1, \dots, p\}$. Then,

$$\begin{aligned} y + \sum_{j=1}^{l}(h_{t_j} - h_{t_{j+1}})z_{t_j} &\geq y + \sum_{j=1}^{l}(h_{t_j} - h_{t_{j+1}})w_{t_j} \\ &\geq y + \sum_{j=1}^{l}\sum_{i=t_j}^{t_{j+1}-1}(h_i - h_{i+1})w_i \\ &= y + \sum_{i=t_1}^{p}(h_i - h_{i+1})w_i. \end{aligned}$$

But also, $y + \sum_{i=1}^{p}(h_i - h_{i+1})w_i \geq h_1$ and so

$$y + \sum_{i=t_1}^{p}(h_i - h_{i+1})w_i \geq h_1 - \sum_{i=1}^{t_1-1}(h_i - h_{i+1})w_i \geq h_1 - (h_1 - h_{t_1}) = h_{t_1}.$$

Thus, $(y, z) \in FS$. □

Because of the knapsack constraint (27), formulation EF does not characterize the convex hull of feasible solutions of G. We therefore investigate what other valid inequalities exist for this formulation. We introduce the notation

$$f_k := \sum_{i=1}^{k}\pi_i, \quad k = 0, \dots, p.$$

Theorem 9. *Let $k \in \{1, \dots, p\}$ and let $S \subseteq \{k, \dots, n\}$ be such that $\sum_{i \in S}\pi_i \leq \epsilon - f_{k-1}$. Then,*

$$\sum_{i \in S}\pi_i z_i + \sum_{i \in \{k,\dots,p\} \setminus S}\pi_i w_i \leq \epsilon - f_{k-1} \tag{28}$$

is valid for EG.

Proof. Let $l = \max\{i : w_i = 1\}$ so that $z_i = w_i = 1$ for $i = 1, \ldots, l$ and hence $\sum_{i=l+1}^{n} \pi_i z_i \leq \epsilon - f_l$. Suppose first $l < k$. Then, $\sum_{i \in \{k,\ldots,p\} \setminus S} \pi_i w_i = 0$ and the result follows since, by definition of the set S, $\sum_{i \in S} \pi_i \leq \epsilon - f_{k-1}$. Next, suppose $l \geq k$. Then,

$$\sum_{i \in S} \pi_i z_i \leq \sum_{i \in S \cap \{k,\ldots,l\}} \pi_i z_i + \sum_{i=l+1}^{n} \pi_i z_i \leq \sum_{i \in S \cap \{k,\ldots,l\}} \pi_i + \epsilon - f_l$$

and also $\sum_{i \in \{k,\ldots,p\} \setminus S} \pi_i w_i = \sum_{i \in \{k,\ldots,l\} \setminus S} \pi_i$. Thus,

$$\sum_{i \in S} \pi_i z_i + \sum_{i \in \{k,\ldots,p\} \setminus S} \pi_i w_i \leq \sum_{i \in S \cap \{k,\ldots,l\}} \pi_i + \epsilon - f_l + \sum_{i \in \{k,\ldots,l\} \setminus S} \pi_i = \epsilon - f_{k-1}.$$

□

4.2 Equal Probabilities

Now, consider the case in which $\pi_i = 1/n$ for $i = 1, \ldots, n$. Then the extended formulation becomes

$$EG' = \left\{ (y, z, w) \in \mathbb{R}_+ \times \{0,1\}^{n+p} : (12) \text{ and } (24) - (26) \right\}.$$

The inequalities (28) become

$$\sum_{i \in S} z_i + \sum_{i \in \{k,\ldots,p\} \setminus S} w_i \leq p - k + 1 \qquad \forall S \in \mathcal{S}_k,\ k = 1, \ldots, p. \tag{29}$$

Example 2. (Example 1 continued.) The extended formulation EG' is given by

$$\begin{aligned} w_1 \geq w_2 \geq w_3 \geq w_4& \\ z_i \geq w_i& \quad i = 1, \ldots, 4 \\ y + 2w_1 + 4w_2 + 3w_3 + 5w_4 \geq 20& \\ \sum_{i=1}^{10} z_i \leq 4,\ z \in \{0,1\}^{10},\ w \in \{0,1\}^4.& \end{aligned}$$

Let $k = 2$ and $S = \{4, 5, 6\}$. Then (29) becomes

$$z_4 + z_5 + z_6 + w_2 + w_3 \leq 3.$$

□

Next we show that in the equal probabilities case, the inequalities (29) together with the inequalities defining EG' are sufficient to define the convex hull of the extended formulation EG'. Let

$$EH' = \left\{ (y, z, w) \in \mathbb{R}_+ \times [0,1]^{n+p} : (12), (24) - (26) \text{ and } (29) \right\}$$

be the linear relaxation of the extended formulation, augmented with this set of valid inequalities.

Theorem 10. $EH' = \text{conv}(EG')$.

Proof. That $EH' \supseteq \text{conv}(EG')$ is immediate by validity of the extended formulation and the inequalities (29).

To prove $EH' \subseteq \text{conv}(EG')$ we first show that it is sufficient to prove that the polytope

$$H = \{(z, w) \in [0, 1]^{n+p} : (12), (24), (25) \text{ and } (29)\}$$

is integral. Indeed, suppose H is integral, and let $(y, z, w) \in EH'$. Then, $(z, w) \in H$, and hence there a exists a finite set of integral points $(z^j, w^j), j \in J$, each in H, and a weight vector $\lambda \in \mathbb{R}_+^{|J|}$ with $\sum_{j \in J} \lambda_j = 1$ such that $(z, w) = \sum_{j \in J} \lambda_j (z^j, w^j)$. For each $j \in J$ define $y^j = h_1 - \sum_{i=1}^p (h_i - h_{i+1}) w_i^j$ so that $(y^j, z^j, w^j) \in EG'$ and also

$$\sum_{j \in J} \lambda_j y^j = h_1 - \sum_{i=1}^{p} (h_i - h_{i+1}) w_i \le y.$$

Thus, there exists $\mu \ge 0$ such that $(y, z, w) = \sum_{j \in J} \lambda_j (y^j, z^j, w^j) + \mu(1, \mathbf{0})$ where each $(y^j, z^j, w^j) \in EG'$ and $(1, \mathbf{0})$ is a feasible direction for EG', which establishes that $(y, z, w) \in \text{conv}(EG')$.

We now move to proving the integrality of H, or equivalently, that $H = \text{conv}(H^I)$ where $H^I = H \cap \{0, 1\}^{n+p}$. Thus, if $(z, w) \in H$, we aim to prove $(z, w) \in \text{conv}(H^I)$. We do this in two steps. First we establish a sufficient condition for $(z, w) \in \text{conv}(H^I)$, and then show that if $(z, w) \in H$ it satisfies this condition.

A sufficient condition for $(z, w) \in \text{conv}(H^I)$.

First observe that the feasible points of H^I are given by $w_j = 1$ for $j = 1, \ldots, k-1$ and $w_j = 0$ for $j = k, \ldots, p$ and

$$z_j = \begin{cases} 1 & j = 1, \ldots, k-1 \text{ and } j \in S \\ 0 & j \in \{k, \ldots, n\} \setminus S \end{cases}$$

for all $S \in \mathcal{S}_k$ and $k = 1, \ldots, p+1$. Thus, an inequality

$$\sum_{j=1}^{n} \alpha_j z_j + \sum_{j=1}^{p} \gamma_j w_j - \beta \le 0 \tag{30}$$

is valid for $\text{conv}(H^I)$ if and only if

$$\sum_{j=1}^{k-1} (\alpha_j + \gamma_j) + \max_{S \in \mathcal{S}_k} \sum_{j \in S} \alpha_j - \beta \le 0 \qquad k = 1, \ldots, p+1. \tag{31}$$

Representing the term $\max\left\{\sum_{j\in S}\alpha_j : S\in\mathcal{S}_k\right\}$ as a linear program and taking the dual, as in (19) and (20) in the proof of Theorem 4, we obtain that (31) is satisfied and hence (30) is valid if and only if the the system of inequalities

$$\sum_{j=1}^{k-1}(\alpha_j+\gamma_j)+\sum_{j=k}^{n}\rho_{jk}+(p-k+1)\sigma_k-\beta\le 0 \tag{32}$$

$$\alpha_j-\sigma_k-\rho_{jk}\le 0 \quad j=k,\dots,n \tag{33}$$

$$\sigma_k\ge 0,\ \rho_{jk}\ge 0 \quad j=k,\dots,n \tag{34}$$

has a feasible solution for $k=1,\dots,p+1$. Thus, $(w,z)\in\operatorname{conv}(H^I)$ if and only if

$$\max_{\alpha,\beta,\gamma,\sigma,\rho}\left\{\sum_{j=1}^{n}\alpha_j z_j+\sum_{j=1}^{p}\gamma_j w_j-\beta : (32)-(34),\ k=1,\dots,p+1\right\}\le 0.$$

For $k=1,\dots,p+1$, associate with (32) the dual variable δ_k, and with (33) the dual variables η_{jk} for $j=k,\dots,n$. Then, applying Farkas' lemma to (32) - (34) and the condition

$$\sum_{j=1}^{n}\alpha_j z_j+\sum_{j=1}^{p}\gamma_j w_j-\beta>0$$

we obtain that $(w,z)\in\operatorname{conv}(H^I)$ if and only if the system

$$\sum_{k=j+1}^{p+1}\delta_k=w_j \quad j=1,\dots,p \tag{35}$$

$$\sum_{k=j+1}^{p+1}\delta_k+\sum_{k=1}^{j}\eta_{jk}=z_j \quad j=1,\dots,p \tag{36}$$

$$\sum_{k=1}^{p+1}\eta_{jk}=z_j \quad j=p+1,\dots,n \tag{37}$$

$$(p-k+1)\delta_k-\sum_{j=k}^{n}\eta_{jk}\ge 0 \quad k=1,\dots,p+1 \tag{38}$$

$$\delta_k-\eta_{jk}\ge 0 \quad j=k,\dots,n,\ k=1,\dots,p+1 \tag{39}$$

$$\sum_{k=1}^{p+1}\delta_k=1 \tag{40}$$

$$\delta_k\ge 0,\eta_{jk}\ge 0 \quad j=k,\dots,n,\ k=1,\dots,p+1 \tag{41}$$

has a feasible solution, where constraints (35) are associated with variables γ, (36) and (37) are associated with α, (38) are associated with σ, (39) are associated with ρ, and (40) is associated with β. Noting that (35) and (40) imply $\delta_k = w_{k-1} - w_k$ for $k = 1, \dots, p+1$, with $w_0 := 1$ and $w_{p+1} := 0$, we see that $(w, z) \in \text{conv}(H^I)$ if and only if $w_{k-1} - w_k \geq 0$ for $k = 1, \dots, p+1$ and the system

$$\sum_{k=1}^{\min\{j,p+1\}} \eta_{jk} = \theta_j \qquad j = 1, \dots, n \tag{42}$$

$$\sum_{j=k}^{n} \eta_{jk} \leq (p-k+1)(w_{k-1} - w_k) \qquad k = 1, \dots, p+1 \tag{43}$$

$$0 \leq \eta_{jk} \leq w_{k-1} - w_k \qquad j = k, \dots, n,\ k = 1, \dots, p+1 \tag{44}$$

has a feasible solution, where $\theta_j = z_j - w_j$ for $j = 1, \dots, p$ and $\theta_j = z_j$ for $j = p+1, \dots, n$.

Verification that $(z, w) \in H$ implies $(z, w) \in \text{conv}(H^I)$.

Let $(z, w) \in H$ and consider a network G with node set given by $V = \{u, v, r_k \text{ for } k = 1, \dots, p+1,\ m_j \text{ for } j \in N\}$. This network has arcs from u to r_k with capacity $(p-k+1)(w_{k-1}-w_k)$ for all $k = 1, \dots, p+1$, arcs from r_k to m_j with capacity $w_{k-1} - w_k$ for all $j = k, \dots, n$ and $k = 1, \dots, p+1$, and arcs from m_j to v with capacity θ_j for all $j \in N$. An example of this network with $n = 4$ and $p = 2$ is given in Figure 1. The labels on the arcs in this figure represent the capacities. For the arcs from nodes r_k to nodes m_j, the capacity depends only on the node r_k, so only the first outgoing arc from each r_k is labeled. It is easy to check that if this network has a flow from u to v of value $\sum_{j \in N} \theta_j$, then the system (42) - (44) has a feasible solution. We will show that $(z, w) \in H$ implies the minimum $u - v$ cut in the network is at least $\sum_{j \in N} \theta_j$, and by the max-flow min-cut theorem, this guarantees a flow of this value exists, proving that $(z, w) \in \text{conv}(H^I)$.

Now, consider a minimum $u - v$ cut in the network G, defined by a node set $U \subset V$ with $u \in U$ and $v \notin U$. Let $S = \{m_j : j \in N \setminus U\}$. Note that if $r_k \notin U$ we obtain an arc in the cut, from u to r_k, with capacity $(p-k+1)(w_{k-1} - w_k)$, whereas if $r_k \in U$, we obtain a set of arcs in the cut, from r_k to m_j for $j \in S$ such that $j \geq k$, with total capacity

$$\sum_{j \in S \cap \{k, \dots, n\}} (w_{k-1} - w_k) = |S \cap \{k, \dots, n\}\,|(w_{k-1} - w_k).$$

Thus, because $w_{k-1} \geq w_k$ we can assume that in this minimum $u - v$ cut $r_k \in U$ if and only if $|S \cap \{k, \dots, n\}\,| < p - k + 1$. Hence, if we let $l = \min\{k = 1, \dots, p+1 : |S \cap \{k, \dots, n\}\,| \geq p–k+1\}$ then we can assume $r_k \in U$ for $1 \leq k < l$ and $r_k \notin U$ for $l \leq k \leq p+1$.

We now show that $S \subseteq \{l, \dots, n\}$. Indeed, suppose $j < l$. If $j \in S$ then the cut includes arcs from r_k to m_j with capacity $(w_{k-1} - w_k)$ for all $1 \leq k \leq j$ yielding a total capacity of $1 - w_j$. If $j \notin S$, then the cut includes an arc from

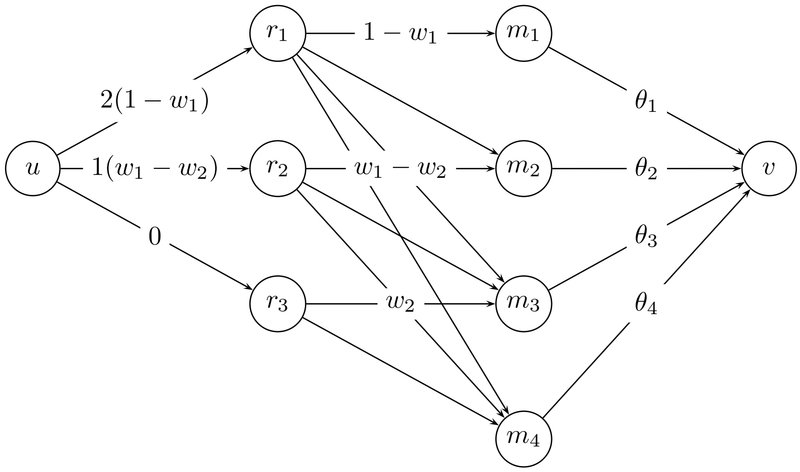

Figure 1: Example of network G with $p = 2$ and $n = 4$.

m_j to v with capacity $\theta_j = z_j - w_j$. Because $z_j \leq 1$, this implies we can assume that in this minimum $u - v$ cut if $j < l$, then $j \notin S$.

Now suppose that $l = 1$, which occurs if $|S| \geq p$. Then the value of the minimum cut is given by

$$\sum_{k=1}^{p+1}(p-k+1)(w_{k-1}-w_k) + \sum_{j \notin S}\theta_j = p - \sum_{k=1}^{p} w_k + \sum_{j \notin S}\theta_j$$
$$\geq (p - \sum_{j \in S} z_j) + \sum_{j \in N}\theta_j \geq \sum_{j \in N}\theta_j$$

since $\sum_{j \in N} z_j \leq p$. Thus, in this case, the value of the minimum cut is at least $\sum_{j \in N}\theta_j$. So now assume $l > 1$. In this case, we claim that $|S| = p - l + 1$. Indeed, if not, then $|S| > p - l + 1$, and so $|S \cap \{l-1, \dots, n\}| \geq p - (l-1) - 1$, contradicting the minimality in the definition of l since $l - 1$ also satisfies the condition in the definition.

The capacity of this minimum $u - v$ cut is

$$C = \sum_{k=1}^{l-1}(p-k+1)(w_{k-1}-w_k) + \sum_{k=l}^{p+1}|S|(w_{k-1}-w_k) + \sum_{j \in N \setminus S}\theta_j.$$

Since,

$$\sum_{k=1}^{l-1}(p-k+1)(w_{k-1}-w_k) = \sum_{k=l}^{p}\sum_{j=k}^{p}(w_{k-1}-w_k) = \sum_{j=l}^{p}(w_{l-1}-w_j)$$

it follows that

$$C = (p-l+1)w_{l-1} - \sum_{k=l}^{p} w_k + (1-w_{l-1})|S| + \sum_{j\in N\setminus S} \theta_j$$
$$= (p-l+1) - \sum_{k=l}^{p} w_k + \sum_{j\in N\setminus S} \theta_j \geq \sum_{j\in N} \theta_j$$

by (29) for $k = l$ since $S \subseteq \{l, \ldots, n\}$ and $|S| = p - l + 1$. □

We close this section by noting that inequalities (29) can be separated in polynomial time. Indeed, suppose we wish to separate the point (z^*, w^*). Then separation can be accomplished by calculating

$$V_k^* = \max_{S\in\mathcal{S}_k} \left\{ \sum_{i\in S} z_i^* + \sum_{i\in\{k,\ldots,p\}\setminus S} w_i^* \right\} = \max_{S\in\mathcal{S}_k} \left\{ \sum_{i\in S} \theta_i^* \right\} + \sum_{i=k}^{p} w_i^*$$

for $k = 1, \ldots, p$ where $\theta_i^* = z_i^* - w_i^*$ for $i = 1, \ldots, p$ and $\theta_i^* = z_i^*$ for $i = p+1, \ldots, n$. If $V_k^* > p{-}k{+}1$ for any k, then a violated inequality is found. Hence, a trivial separation algorithm is to first sort the values θ_i^* in non-increasing order, then for each k, find the maximizing set $S \in \mathcal{S}_k$ by searching this list. This yields an algorithm with complexity $O(n\log n + p^2) = O(n^2)$. However, the complexity can be improved to $O(n\log n)$ as follows. Start by storing the p largest values of θ_i^* over $i \in \{p+1, \ldots, n\}$ in a heap, and define $V_{p+1}^* = 0$. Then, for $k = p, \ldots, 1$ do the following. First insert θ_k^* into this heap. Next remove the largest value, say $\theta_{\max}^*$, from the heap and finally calculate V_k^* by

$$V_k^* = V_{k+1}^* + \max\{\theta_{\max}^*, 0\} + w_k^*.$$

The initial heap construction is accomplished with complexity $O(n\log n)$, and the algorithm then proceeds through p steps, each requiring insertion into a heap and removal of the maximum value from a heap, which can each be done with $O(\log p)$ complexity, yielding overall complexity of $O(n\log n)$. For general probabilities π_i, (heuristic) separation of inequalities (28) can be accomplished by (heuristically) solving p knapsack problems.

5 Computational Experience

We performed computational tests on a probabilistic version of the classical transportation problem. We have a set of suppliers I and a set of customers D with $|D| = m$. The suppliers have limited capacity M_i for $i \in I$. There is a transportation cost c_{ij} for shipping a unit of product from supplier $i \in I$ to customer $j \in D$. The customer demands are random and are represented by a random vector $\tilde{d} \in \mathbb{R}_+^m$. We assume we must choose the shipment quantities before the customer demands are known. We enforce the probabilistic constraint

$$P\{\sum_{i\in I} x_{ij} \geq \tilde{d}_j, j = 1, \ldots, m\} \geq 1 - \epsilon \tag{45}$$

where $x_{ij} \geq 0$ is the amount shipped from supplier $i \in I$ to customer $j \in D$. The objective is to minimize distribution costs subject to (45), and the supply capacity constraints

$$\sum_{j \in D} x_{ij} \leq M_i, \quad \forall i \in I.$$

We randomly generated instances with the number of suppliers fixed at 40 and varying numbers of customers and scenarios. The supply capacities and cost coefficients were generated using normal and uniform distributions respectively. For the random demands, we experimented with independent normal, dependent normal and independent Poisson distributions. We found qualitatively similar results in all cases, but the independent normal case yielded the most challenging instances, so for our experiments we focus on this case. For each instance, we first randomly generated the mean and variance of each customer demand. We then generated the number n of scenarios required, independently across scenarios and across customer locations, as Monte Carlo samples with these fixed parameters. In most instances we assumed all scenarios occur with probability $1/n$, but we also did some tests in which the scenarios have general probabilities, which were randomly generated. CPLEX 9.0 was used as the MIP solver and all experiments were done on a computer with two 2.4 Ghz processors (although no parallelism is used) and 2.0 Gb of memory. We set a time limit of one hour. For each problem size we generated 5 random instances and, unless otherwise specified, the computational results reported are averages over the 5 instances.

5.1 Comparison of Formulations

In Table 1 we compare the results obtained by solving our instances using

1. formulation PMIP given by (3) - (5),
2. formulation PMIP with strengthened star inequalities (11), and
3. the extended formulation of Sect. 4, but without (28) or (29).

When the strengthened star inequalities are not used, we still used the improved formulation of G corresponding to (10). Recall that the strengthened star inequalities subsume the rows defining the formulation PMIP; therefore, when using these inequalities we initially added only a small subset of the mp inequalities in the formulation. Subsequently separating the strengthened star inequalities as needed guarantees the formulation remains valid. For formulation PMIP without strengthened star inequalities, we report the average optimality gap that remained after the hour time limit was reached, where we define the optimality gap as the difference between the final upper and lower bounds, divided by the upper bound. For the other two formulations, which we refer to as the strong formulations, we report the geometric average of the time to solve the instances to optimality. We used $\epsilon = 0.05$ and $\epsilon = 0.1$, reflecting the natural assumption that we want to meet demand with high probability.

Table 1: Average solution times for different formulations.

				PMIP	PMIP+Star		Extended
Probabilities	ϵ	m	n	Gap	Cuts	Time(s)	Time(s)
Equal	0.05	100	1000	0.18%	734.8	7.7	1.1
		100	2000	1.29%	1414.2	31.8	4.6
		200	2000	1.02%	1848.4	61.4	12.1
		200	3000	2.56%	2644.0	108.6	12.4
	0.10	100	1000	2.19%	1553.2	34.6	12.7
		100	2000	4.87%	2970.2	211.3	41.1
		200	2000	4.48%	3854.0	268.5	662.2
		200	3000	5.84%	5540.8	812.7	490.4
General	0.05	100	1000	0.20%	931.8	9.0	3.9
		100	2000	1.04%	1806.6	55.2	13.2
	0.10	100	1000	1.76%	1866.0	28.7	52.5
		100	2000	4.02%	3686.2	348.5	99.2

The first observation from Table 1 is that formulation PMIP without the strengthened star inequalities failed to solve these instances within an hour, often leaving large optimality gaps, whereas the instances are solved efficiently using the strong formulations. The number of nodes required to solve the instances for the strong formulations is very small. The instances with equi-probable scenarios were usually solved at the root, and even when branching was required, the root relaxation usually gave an exact lower bound. Branching in this case was only required to find an integer solution which achieved this bound. The instances with general probabilities required slightly more branching, but generally not more than 100 nodes. Observe that the number of strengthened star inequalities that were added is small relative to the number of rows in the formulation PMIP itself. For example, with equi-probable scenarios, $\epsilon = 0.1$, $m = 200$ and $n = 3,000$, the number of rows in PMIP would be $mp = 60,000$, but on average, only $5,541$ strengthened star inequalities were added. Next we observe that in most cases the computation time using the extended formulation is significantly less than the formulation with strengthened star inequalities. Finally, we observe that the instances with general probabilities take somewhat longer to solve than those with equi-probable scenarios but can still be solved efficiently.

5.2 The Effect of Increasing ϵ

The results of Table 1 indicate that the strong formulations can solve large instances to optimality when ϵ is small, which is the typical case. However, it is still an interesting question to investigate how well this approach works for larger ϵ. Note first that we should expect solution times to grow with ϵ if only because the formulation sizes grow with ϵ. However, we observe from the chart in Figure 2 that the situation is much worse than this. This chart shows the root LP solve times and optimality gaps achieved after an hour of computation time for an example instance with equi-probable scenarios, $m = 50$ rows and $n = 1,000$ scenarios at increasing levels of ϵ, using the formulation PMIP with strengthened star inequalities. Root LP solve time here refers to the time until no further strengthened star inequalities could be separated. We see that the time to solve the root linear programs does indeed grow with ϵ as expected, but the optimality gaps achieved after an hour of computation time deteriorate drastically with growing ϵ. This is explained by the increased time to solve the linear programming relaxations *combined with* an apparent weakening of the relaxation bound as ϵ increases.

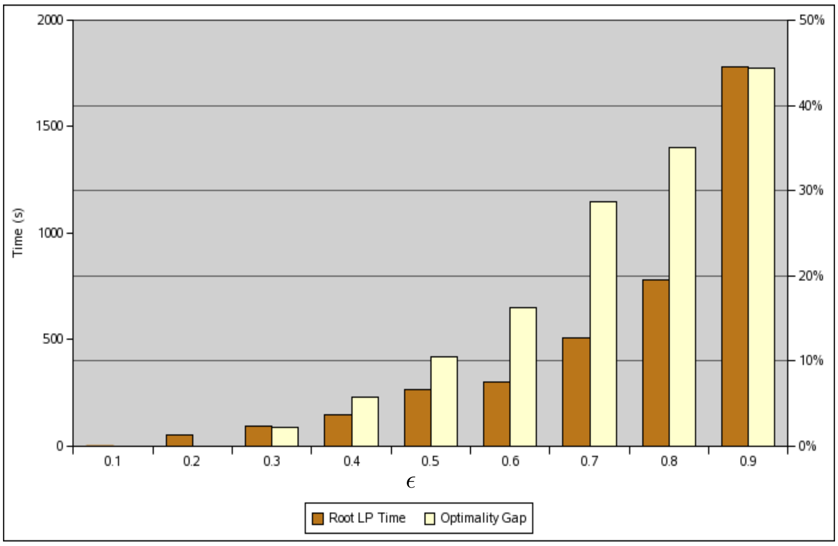

Figure 2: The effect of increasing ϵ.

5.3 Testing Inequalities (29)

With small ϵ the root relaxation given by the extended formulation is extremely tight, so that adding the inequalities (29) is unlikely to have a positive impact on computation time. However, for larger ϵ, we have seen that formulation PMIP, augmented with the strengthened star inequalities, and hence also the extended formulation, may have a substantial optimality gap. We therefore investigated whether using inequalities (29) in the extended formulation can improve solution time in this case. In Table 2 we present results comparing solution times and node counts with and without inequalities (29) for instances with larger ϵ. We performed these tests on smaller instances since these instances are already hard for these values of ϵ. We observe that adding inequalities (29) at the root can decrease the root optimality gap significantly. For the instances that could be solved in one hour, this leads to a significant reduction in the number of nodes

Table 2: Results with and without inequalities (29).

			Root Gap		Nodes		Time(s) or Gap	
m	ϵ	n	Ext	+(29)	Ext	+(29)	Ext	+(29)
25	0.3	250	1.18%	0.67%	276.9	69.0	121.2	93.9
	0.3	500	1.51%	0.58%	455.0	165.8	750.6	641.3
	0.35	250	2.19%	1.50%	1259.4	409.0	563.2	408.4
	0.35	500	2.55%	1.61%	2297.6	968.8	0.22%	0.06%
50	0.3	500	2.32%	2.00%	991.8	238.6	1.37%	1.41%
	0.3	1000	2.32%	1.75%	28.3	8.5	1.98%	1.66%
	0.35	500	4.10%	3.31%	650.4	92.9	3.03%	2.66%
	0.35	1000	4.01%	3.23%	22.7	6.2	3.58%	3.17%

explored, and a moderate reduction in solution time. For the instances which were not solved in one hour, the remaining optimality gap was usually, but not always, lower when the inequalities (29) were used. These results indicate that when ϵ is somewhat larger, inequalities (29) may be helpful on smaller instances. However, they also reinforce the difficulty of the instances with larger ϵ, since even with these inequalities, only the smallest of these smaller instances could be solved to optimality within an hour.

6 Concluding Remarks

We have presented strong integer programming formulations for linear programs with probabilistic constraints in which the right-hand side is random with finite distribution. In the process we made use of existing results on mixing sets, and have introduced new results for the case in which the mixing set additionally has a knapsack restriction. Computational results indicate that these formulations are extremely effective on instances in which reasonably high reliability is enforced, which is the typical case. However, instances in which the desired reliability level is lower remain difficult to solve, partly due to increased size of the formulations, but more significantly due to the weakening of the formulation bounds. Moreover, these instances remain difficult even when using the inequalities which characterize the single row relaxation convex hull. This suggests that relaxations which consider multiple rows simultaneously need to be studied to yield valid inequalities which significantly improve the relaxation bounds for these instances.

Our future work in this area will focus on addressing the two assumptions we made at the beginning of this paper. The finite distribution assumption can be addressed by using the results about the statistical relationship between a problem with probabilistic constraints and its Monte Carlo sample approximation to establish methods for generating bounds on the optimal value of the original problem. Computational studies will need to be performed to establish the practicality of this approach. We expect that relaxing the assumption that only the right-hand side is random will be more challenging. A natural first step in this direction will be to extend results from the *generalized* mixing set [14, 21] to the case in which an additional knapsack constraint is present.

References

[1] Ahmed, S., Shapiro, A.: The sample average approximation method for stochastic programs with integer recourse (2002). Preprint available at www.optimization-online.org

[2] Atamtürk, A., Nemhauser, G.L., Savelsbergh, M.W.P.: The mixed vertex packing problem. Math. Program. **89**, 35–53 (2000)

[3] Atlason, J., Epelman, M.A., Henderson, S.G.: Call center staffing with simulation and cutting plane methods. Ann. Oper. Res. **127**, 333–358 (2004)

[4] Beraldi, P., Ruszczyński, A.: A branch and bound method for stochastic integer programs under probabilistic constraints. Optim. Methods Softw. **17**, 359–382 (2002)

[5] Cheon, M.S., Ahmed, S., Al-Khayyal, F.: A branch-reduce-cut algorithm for the global optimization of probabilistically constrained linear programs. Math. Program. **108**, 617–634 (2006)

[6] Dentcheva, D., Prékopa, A., Ruszczyński, A.: Concavity and efficient points of discrete distributions in probabilistic programming. Math. Program. **89**, 55–77 (2000)

[7] Guan, Y., Ahmed, S., Nemhauser, G.L.: Sequential pairing of mixed integer inequalities. Discrete Optim. **4**, 21–39 (2007)

[8] Günlük, O., Pochet, Y.: Mixing mixed-integer inequalities. Math. Program. **90**, 429–457 (2001)

[9] Henrion, R., Li, P., Möller, A., Steinbach, M.C., Wendt, M., Wozny, G.: Stochastic optimization for operating chemical processes under uncertainty. In: M. Grötschel, S. Krunke, J. Rambau (eds.) Online Optimization of Large Scale Systems, pp. 457–478 (2001)

[10] Henrion, R., Möller, A.: Optimization of a continuous distillation process under random inflow rate. Computers and Mathematics with Applications **45**, 247–262 (2003)

[11] Lejeune, M.A., Ruszczyński, A.: An efficient trajectory method for probabilistic inventory-production-distribution problems. Oper. Res. (Forthcoming, 2007)

[12] Luedtke, J.: Integer programming approaches to two non-convex optimization problems. Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA (2007)

[13] Luedtke, J., Ahmed, S., Nemhauser, G.: An integer programming approach for linear programs with probabilistic constraints. In: M. Fischetti, D. Williamson (eds.) IPCO 2007, Lect. Notes Comput. Sci., pp. 410–423. Springer-Verlag, Berlin (2007)

[14] Miller, A.J., Wolsey, L.A.: Tight formulations for some simple mixed integer programs and convex objective integer programs. Math. Program. **98**, 73–88 (2003)

[15] Murr, M.R., Prékopa, A.: Solution of a product substitution problem using stochastic programming. In: S.P. Uryasev (ed.) Probabilistic Constrained Optimization: Methodology and Applications, pp. 252–271. Kluwer Academic Publishers (2000)

[16] Prékopa, A.: On probabilistic constrained programmming. In: H.W. Kuhn (ed.) Proceedings of the Princeton Symposium on Mathematical Programming, pp. 113–138. Princeton University Press, Princeton, N.J. (1970)

[17] Prékopa, A.: Probabilistic programming. In: A. Ruszczyński, A. Shapiro (eds.) Stochastic Programming, *Handbooks in Operations Research and Management Science*, vol. 10, pp. 267–351. Elsevier (2003)

[18] Ruszczyński, A.: Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. Math. Program. **93**, 195–215 (2002)

[19] Shapiro, A., Homem-de-Mello, T.: On the rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. SIAM J. Optim. **11**, 70–86 (2000)

[20] Takyi, A.K., Lence, B.J.: Surface water quality management using a multiple-realization chance constraint method. Water Resources Research **35**, 1657–1670 (1999)

[21] Van Vyve, M.: The continuous mixing polyhedron. Math. Oper. Res. **30**, 441–452 (2005)