Lifted Inequalities for 0-1 Mixed Integer Programming: Basic Theory and Algorithms*

Jean-Philippe P. Richard[1], Ismael R. de Farias, Jr.[2], and George L. Nemhauser[3]

[1] School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205, USA
jpacc@isye.gatech.edu

[2] Center for Operations Research and Econometrics, 34 Voie du Roman Pays, 1348 Louvain-La-Neuve, Belgium
defarias@core.ucl.ac.be

[3] School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205, USA
george.nemhauser@isye.gatech.edu

Abstract. We study the mixed 0-1 knapsack polytope, which is defined by a single knapsack constraint that contains 0-1 and bounded continuous variables. We develop a lifting theory for the continuous variables. In particular, we present a pseudo-polynomial algorithm for the sequential lifting of the continuous variables. We introduce the concept of superlinear inequalities and show that our lifting scheme can be significantly simplified for them. Finally, we show that superlinearity results can be generalized to nonsuperlinear inequalities when the coefficients of the continuous variables lifted are large.

1 Introduction

Lifted cover inequalities, derived from 0-1 knapsack inequalities, have proven to be a useful family of cuts for solving 0-1 integer programs by branch-and-cut algorithms. The idea was first proposed by Crowder, Johnson and Padberg [5], with the theoretical foundation coming from the studies of the 0-1 knapsack polytope by Balas [3], Balas and Zemel [2], Hammer, Johnson and Peled [15] and Wolsey [23]. A computational study is presented in Gu, Nemhauser and Savelsbergh [13]. In order to extend these ideas to the mixed integer case, in which continuous variables are present as well, it is necessary to develop a lifting theory for the continuous variables. In this paper we study the polytope generated by a mixed 0-1 knapsack inequality with an arbitrary number of bounded continuous variables and we develop a lifting theory for them.

Although we are not aware of any previous study of the mixed 0-1 knapsack polytope, valid inequalities and facets for related polyhedra have been known for quite some time. For example, there are the mixed integer cuts introduced by Gomory [11], the MIR inequalities introduced by Nemhauser and Wolsey

* This research was supported by NSF grants DMI-0100020 and DMI-0121495.

W.J. Cook and A.S. Schulz (Eds.): IPCO 2002, LNCS 2337, pp. 161–175, 2002.

[17], and the mixed disjunctive cuts introduced by Balas, Ceria and Cornuéjols [1]. More closely related to our study is the "0-1 knapsack polyhedron with a single nonnegative continuous variable" introduced by Marchand and Wolsey [16]. However, as we discuss in the next section, our polytope yields more general results with respect to generating cuts.

Given $M = \{1, \dots, m\}$, $N = \{1, \dots, n\}$, the sets of integers $\{a_1, \dots, a_m\}$, $\{b_1, \dots, b_n\}$, and the integer d, let

$$S = \{(x, y) \in \{0,1\}^m \times [0,1]^n \mid \sum_{j\in M} a_j x_j + \sum_{j \in N} b_j y_j \leq d\}.$$

We define the mixed 0-1 knapsack polytope as $PS = conv(S)$. Note that the choice of 1 as an upper bound for the (bounded) continuous variables is without loss of generality, since they can be rescaled. We assume throughout the paper that

Assumption 1. $M \neq \emptyset$ and $N \neq \emptyset$.

Assumption 2. $0 < a_j \leq d \ \forall j \in M$, and $0 < b_j \leq d \ \forall j \in N$.

If Assumption 1 is not satisfied, PS is either trivial ($M = \emptyset$) or has already been studied ($N = \emptyset$). Clearly, there is no loss of generality in Assumption 2.

We now give a few results about PS and its inequality description. Whenever it is clear from the context, we use the term facet to denote the corresponding inequality as well.

Proposition 1. *PS is full-dimensional.* □

Some inequalities of the linear description of PS are easy to explain. We give their characterization in the following proposition.

Proposition 2. *The inequality*

$$x_i \geq 0 \tag{1}$$

is a facet of PS $\forall i \in M$. The inequality

$$x_i \leq 1 \tag{2}$$

is a facet of PS for $i \in M$ iff $\max\{a_j \mid j \in M - \{i\}\} + a_i \leq d$ *and* $a_i < d$. *The inequality*

$$y_i \geq 0 \tag{3}$$

is a facet of PS $\forall i \in N$. The inequality

$$y_i \leq 1 \tag{4}$$

is a facet of PS for $i \in N$ iff $\max\{a_j \mid j \in M\} + b_i \leq d$. □

We call the inequalities of Proposition 2 trivial. Another trivial inequality that may or may not be a facet of PS is the initial knapsack inequality. It does not seem that there exists a simple characterization of when this inequality is a facet except for particular cases, as for example, when the coefficients of the continuous variables are big (see Marchand and Wolsey [16] and Sect. 2). Sufficient conditions can also be derived from Theorem 4 presented later. In the remainder of the paper, we characterize some families of nontrivial facets of PS that can be obtained by sequential lifting. The nontrivial facets of PS satisfy the following proposition.

Proposition 3. *Assume that $\sum_{j\in M} \alpha_j x_j + \sum_{j\in N} \beta_j y_j \leq \delta$ is a facet of PS that is not a multiple of (1) or (3). Then $\alpha_j \geq 0$ for $j \in M$, $\beta_j \geq 0$ for $j \in N$, and $\delta > 0$.* □

In Sect. 2 we discuss why the polytope we study is more general than the polyhedron of Marchand and Wolsey. In Sect. 3 we develop sequential lifting schemes for continuous variables. For the sequential lifting of continuous variables fixed at 0 ("lifting from 0"), we show that the lifting coefficients are 0 almost always (Theorem 2). Finally, we consider the sequential lifting of continuous variables fixed at 1 ("lifting from 1"), which is much more interesting and difficult, and we give a pseudo-polynomial algorithm for it (Theorem 5). In Sect. 4 we present a way to alleviate some of the computational burden associated with our lifting scheme by imposing structure on the inequality to be lifted and we introduce the concept of superlinear lifting (Theorem 7). In Sect. 5 we present another way to reduce the amount of computation in our algorithm by imposing restrictions on the coefficients of the continuous variables to be lifted (Theorem 11). We conclude in Sect. 6 with some discussion on how our lifting theory can be used algorithmically.

2 A Related Polyhedron

Let

$$T = \{(x, s) \in \{0,1\}^m \times [0, \infty) \mid \sum_{j\in M} a_j x_j \leq d + s\} .$$

Marchand and Wolsey [16] studied the 0-1 knapsack polyhedron with a single nonnegative continuous variable $PT = conv(T)$. The polytope PS and the polyhedron PT are similar. However, as we show in this section, because PS is bounded and has more than one continuous variable, it is possible to derive from it certain cuts for 0-1 mixed integer programming that cannot be derived from PT. We first show that PS with $|N| = 1$ is more general than PT because it is bounded. Then we show that PS with more than one continuous variable is more general than the polytope obtained by aggregating them into a single continuous variable.

When $|N| = 1$, there exists a natural transformation that converts a polytope of the form PS into a polyhedron of the form PT and vice-versa. The simple addition of the constraint $s \le \mu$ transforms PT into PS after adequately scaling and complementing the continuous variable. We call the resulting polytope $PS' = conv\{(x,y) \in \{0,1\}^m \times [0,1] \mid \sum_{j \in M} a_j x_j + \mu y \le d + \mu\}$, where $y = \frac{\mu - s}{\mu}$. When μ is large enough, the facets of PS' are in one-to-one correspondence with the facets of PT (if we remove the facet $y \ge 0$ from the linear description of PS').

Now consider a transformation from PS with a single continuous variable to PT. If we scale and complement the variable y, defining $s = b - by$, we obtain an equivalent polytope where $0 \le s \le b$. Suppose we relax the upper bound on s (i.e. we relax the lower bound on y). Then we obtain a polytope of the form PT that we call $PT' = conv\{(x,s) \in \{0,1\}^m \times [0,\infty) \mid \sum_{j \in M} a_j x_j \le d - b + s\}$. In general, there is no bijection between the facets of PS and those of PT', although valid inequalities for PT' can be turned into valid inequalities of PS by substituting $s = b - by$. However, even when starting from a facet of PT', this procedure is not guaranteed to generate a facet of PS as the following example illustrates.

Example 1. Consider the polytope PS defined by

$$12x_1 + 8x_2 + 8x_3 + 7x_4 + 5y \le 24 \ .$$

Let $s = 5 - 5y$. Introduce s in the previous inequality and relax its upper bound. The inequality becomes

$$12x_1 + 8x_2 + 8x_3 + 7x_4 \le 19 + s,$$

which is in the format studied by Marchand and Wolsey. The inequality

$$8x_1 + 4x_2 + 8x_3 + 7x_4 \le 15 + s$$

is a facet of the polytope PT'. Substituting back $s = 5 - 5y$, we obtain

$$8x_1 + 4x_2 + 8x_3 + 7x_4 + 5y \le 20 \ .$$

This inequality is not a facet of PS since it defines a face whose dimension is only 3. □

Now we show that we need to consider several continuous variables to describe the most general results about facets. Consider the polytopes $PS_1 = conv\{(x,y) \in \{0,1\}^m \times [0,1] \mid \sum_{j \in M} a_j x_j + by \le d\}$ and $PS_n = conv\{(x,y) \in \{0,1\}^m \times [0,1]^n \mid \sum_{j \in M} a_j x_j + \sum_{j \in N} b_j y_j \le d\}$. Note that any polytope of the form PS_n can be turned into a polytope of the fom PS_1 by defining $y = \frac{\sum_{j \in N} b_j y_j}{b}$ and $b = \sum_{j \in N} b_j$. If facets of PS_1 are known, the substitution $y = \sum_{j \in N} b_j y_j$ will turn them into valid inequalities for PS_n. There are two main issues regarding this aggregation procedure. The first is that the inequality generated is not necessarily a facet of PS_n. The second is that there are many facets of PS_n that cannot be obtained in this way. The following example illustrates this last observation.

Example 2. Consider the polytope PS defined by

$$30x_1 + 25x_2 + 23x_3 + 20x_4 + 18x_5 + 17x_6 + 13x_7 + 12x_8 + 16y_1 + 7y_2 \leq 103 \; .$$

The linear description of this polytope was obtained using Porta [4] and contains 3114 inequalities. The following five

$$\begin{array}{rrrrrrrrrrrl}
2x_1 + & 2x_2 + & 2x_3 + & 2x_4 + & x_5 + & x_6 + & x_7 + & x_8 & & & \leq & 8, \quad (5) \\
4x_1 + & 4x_2 + & 4x_3 + & 4x_4 + & 4x_5 + & 4x_6 + & 4x_7 & & +16y_1 & & \leq & 16, \quad (6) \\
6x_1 + & 4x_2 + & 4x_3 + & 4x_4 + & 2x_5 + & 2x_6 + & 2x_7 + & 2x_8 & & +7y_2 & \leq & 23, \quad (7) \\
16x_1 + & 16x_2 + & 12x_3 + & 12x_4 + & 12x_5 + & 8x_6 + & 8x_7 + & 8x_8 & +16y_1 & +7y_2 & \leq & 71, \quad (8) \\
20x_1 + & 20x_2 + & 15x_3 + & 15x_4 + & 15x_5 + & 10x_6 + & 10x_7 + & 10x_8 & +48y_1 & +7y_2 & \leq & 115, \quad (9)
\end{array}$$

are some of them. □

Only inequalities (5) and (8) can be derived by the above procedure. But inequalities (6) and (7) of Example 2 suggest a more general aggregation technique to obtain facets of PS_n. Here, we choose $J \subseteq N$ and define $y = \frac{\sum_{j \in J} b_j y_j}{b}$ where $b = \sum_{j \in J} b_j$. PS_n can be converted into PS_1 by dropping all the variables y_j for $j \in N \backslash J$ and replacing $\sum_{j \in J} b_j y_j$ with by. If facets of PS_1 are known, the substitution $y = \frac{\sum_{j \in J} b_j y_j}{b}$ will turn them into valid inequalities of PS_n. The disadvantages of this aggregation procedure are identical to the ones described for the initial one. The inequalities generated are not necessarily facets and there are some facets of PS_n that cannot be explained in this way. Inequality (9) of Example 2 illustrates this last disadvantage. We note, however, that aggregation can be helpful practically since many facets of PS_n can be obtained in this way. Therefore we will refer to inequalities that can be scaled in such a way that coefficients of the continuous variables are either 0 or the given coefficient of the knapsack inequality as having the *ratio property.*

3 Lifting Theory for Continuous Variables

Describing facets of high dimensional polytopes can be a difficult task. One alternative is to reduce the number of variables by fixing some of them at their upper or lower bounds to obtain a polytope for which at least one nontrivial inequality is known. Once such an inequality is available, it is converted progressively into facets of higher dimensional polytopes by the mechanism known as lifting. Lifting was introduced in the context of the group problem by Gomory [12]. Its computational possibilites were emphasized by Padberg [18] and the approach was generalized by Wolsey [24], Zemel [25] and Balas and Zemel [2]. Although lifting techniques have been studied extensively for 0-1 variables, see Gu [14], there has been limited study of lifting for continuous variables (de Farias [10], de Farias, Johnson and Nemhauser [6,8,7] and de Farias and Nemhauser [9] are exceptions). The purpose of this section is to develop algorithms for it. In principle we should investigate the lifting of continuous variables that are fixed at any value within the interval $[0,1]$. However, we have shown (see Richard, de Farias and Nemhauser [20]) that when S is defined by only one 0-1 mixed knapsack inequality, the inequalities obtained by fixing continuous variables at fractional values are only rarely different from the ones that can be generated from the fixing of continuous variables at 0 and 1. Thus we focus only on the lifting of continuous variables fixed to 0 and 1. For the lifting of continuous variables from 0, we show that the lifting coefficients are almost always equal to 0. We also describe when they are not equal to 0 and how to obtain the lifting coefficients in this case. Finally, we study the lifting of continuous variables from 1 and we show that it is much richer and more difficult than the lifting from 0.

3.1 General Lifting Results

Given M_0, M_1, two disjoint subsets of M, and N_0, N_1 two disjoint subsets of N, we define

$$\begin{aligned} S(M_0, M_1, N_0, N_1) = \; & S \cap \{ (x,y) \in \{0,1\}^m \times [0,1]^n \mid x_j = 0 \, \forall j \in M_0, \\ & x_j = 1 \, \forall j \in M_1, \, y_j = 0 \, \forall j \in N_0, \, y_j = 1 \, \forall j \in N_1 \}, \end{aligned}$$

and $PS(M_0, M_1, N_0, N_1) = conv(S(M_0, M_1, N_0, N_1))$. Whenever it is clear from the context, we abbreviate $S(M_0, M_1, N_0, N_1)$ as S^*, $PS(M_0, M_1, N_0, N_1)$ as PS^*, $M-(M_0 \cup M_1)$ as M^*, $N-(N_0 \cup N_1)$ as N^*, and $d - \sum_{j \in M_1} a_j - \sum_{j \in N_1} b_j = d^*$. We also let $m^* = |M^*|$ and $n^* = |N^*|$. Note that PS^* is defined by the constraint

$$\sum_{j \in M^*} a_j x_j + \sum_{j \in N^*} b_j y_j \leq d^*. \tag{10}$$

We represent nontrivial valid inequalities of PS^* by

$$\sum_{j \in M^*} \alpha_j x_j + \sum_{j \in N^*} \beta_j y_j \leq \delta^*. \tag{11}$$

We assume throughout the paper that:

Assumption 3. $i \in M_0$ whenever $a_i > d^*$.

Note that, as with Assumption 2, there is no loss of generality in Assumption 3 and it implies that PS^* is full-dimensional. As a result, it is simpler to lift (11).

In the remainder of this paper, we assume that $r_1, \ldots, r_s$ are distinct elements of N_0 and N_1 and that we wish to sequentially lift the corresponding variables $y_{r_1}, \ldots, y_{r_s}$ in (11). We denote the associated lifting coefficients by $\beta_{r_1}, \ldots, \beta_{r_s}$. Lemma 1 establishes the lifting procedure, see [10,24] for a proof. In the lemma, i indicates the value we lift from, i.e. 0 or 1. Given a polytope Q we let $V(Q)$ be the set of its extreme points.

Lemma 1. *Let (11) be a valid (resp. facet-defining) inequality of PS^*. Let $i \in \{0,1\}$ and let $r_1, \dots, r_s \in N_i$. Define $\tilde{N}_i^q = N_i - \{r_1, \dots, r_q\}$ and $\tilde{N}_{1-i}^q = N_{1-i}$ for $q = 1, \dots, s$. Then,*

$$\sum_{j \in M^*} \alpha_j x_j + \sum_{j \in N^*} \beta_j y_j + \sum_{t=1}^{s} \beta_{r_t} y_{r_t} \le \delta^* + \sum_{t=1}^{s} i \beta_{r_t} \tag{12}$$

is a valid (resp. facet-defining) inequality for $PS(M_0, M_1, \tilde{N}_0^s, \tilde{N}_1^s)$, where β_{r_q} is the optimal value of $L_i(q)$:

$$\beta_{r_q} = (-1)^i \min\, (-1)^i \frac{\sum_{j \in M^*} \alpha_j x_j + \sum_{j \in N^*} \beta_j y_j + \sum_{t=1}^{q-1} \beta_{r_t}(y_{r_t} - i) - \delta^*}{i - y_{r_q}}$$

$$s.t.\ (x,y) \in V(PS(M_0, M_1, \tilde{N}_0^q, \tilde{N}_1^q)) \text{ and } y_{r_q} \in \left[\frac{1-i}{b_{r_q}}, 1 - \frac{i}{b_{r_q}}\right]$$

for $q = 1, \dots, s$. □

Regardless of whether the lifting coefficients of the continuous variables satisfy the ratio property, they satisfy the following relations.

Theorem 1. *Let $r \in N_0$ and $s \in N_1$ be distinct.*

1. *If y_{r_1} and y_{r_2} are lifted from 0 in (11) then $\frac{\beta_{r_1}}{b_{r_1}} \ge \frac{\beta_{r_2}}{b_{r_2}}$.*
2. *If y_{r_1} and y_{r_2} are lifted from 1 in (11) then $\frac{\beta_{r_1}}{b_{r_1}} \le \frac{\beta_{r_2}}{b_{r_2}}$.*
3. *If y_r is lifted from 0 and y_s is lifted from 1 in (11) then $\frac{\beta_r}{b_r} \le \frac{\beta_s}{b_s}$ (note that the lifting order is irrelevant in this case).* □

3.2 Lifting Continuous Variables from 0 (L_0)

We show that the lifting coefficients of the continuous variables fixed at 0 are almost always zero. We also develop a pseudo-polynomial algorithm when this is not the case.

Theorem 2. *Assume that (11) is a facet of PS^* that is not a multiple of (10). Then, when lifting continuous variables from 0 in (11), we have that $\beta_{r_1} = \cdots = \beta_{r_s} = 0$.* □

We now consider the lifting of continuous variables from 0 in an inequality that is a multiple (10). First we consider the case with $N^* \neq \emptyset$.

Theorem 3. *Let (10) be a facet of PS^* with $N^* \neq \emptyset$. When lifting continuous variables from 0 in (10) we have that $\beta_{r_q} = b_{r_q}$ for $q = 1, \dots, s$.* □

Now we consider the case where $N^* = \emptyset$. Given a polytope PS^* with $N^* = \emptyset$, let

$$\sigma = d^* - \max \sum_{j \in M^*} a_j x_j$$

$$s.t. \sum_{j \in M^*} a_j x_j \le d^* - 1, \tag{13}$$

$$x_j \in \{0,1\}\, \forall j \in M^* .$$

Theorem 4. *Let (10) be a facet of PS^* with $N^* = \emptyset$. When lifting continuous variables from* 0 *in (10), we have that*

1. *If $b_{r_1} \geq \sigma$ then $\beta_{r_q} = b_{r_q}$ for $q = 1, \ldots, s$*
2. *If $b_{r_1} < \sigma$ then $\beta_{r_1} = \sigma$ and $\beta_{r_q} = 0$ for $q = 2, \ldots, s$.* □

Theorem 4 leads to a simple pseudo-polynomial algorithm to perform the lifting from 0 of continuous variables when (11) is a multiple of (10): we compute σ by dynamic programming, and then we use Theorem 4 to deduce the lifting coefficients.

3.3 Lifting Continuous Variables from 1 (L_1)

In Lemma 1, we described a formal way to lift continuous variables from 1. In order to turn it into a practical scheme, we define the function

$$\begin{aligned} \Lambda(w) = \min & \sum_{j\in M^*} a_j x_j + \sum_{j \in N^*} b_j y_j - d^* \\ s.t. & \sum_{j\in M^*} \alpha_j x_j + \sum_{j \in N^*} \beta_j y_j = \delta^* + w \\ & x_j \in \{0,1\}\, \forall j \in M, y_j \in [0,1]\, \forall j \in N. \end{aligned}$$

The domain of the function Λ is

$$\mathbb{W} = \{w \in \mathbb{R} \,|\, \exists\, (x, y) \in \{0,1\}^{m^*} \times [0,1]^{n^*} \text{s.t.} \sum_{j\in M^*} \alpha_j x_j + \sum_{j\in N^*} \beta_j y_j = \delta^* + w\}.$$

We say that (11) is satisfied at equality at least once, abbreviated SEO, if there is at least one $(x^*, y^*) \in S^*$ such that $\sum_{j\in M^*} \alpha_j x^*_j + \sum_{j\in N^*} \beta_j y^*_j = \delta^*$. For $A \subseteq \mathbb{R}$, we define $A_+ = \{x \in A \,|\, x > 0\}$.

We focus on the case $N^* = \emptyset$, i.e. we will consider the lifting of continuous variables from 1 with respect to

$$\sum_{j\in M^*} \alpha_j x_j \leq \delta^* \tag{14}$$

(which is purely 0-1 and SEO). The reason for restricting ourselves to this case is that Λ is a discrete function and it is simpler to obtain the lifting coefficients. The case $N^* \neq \emptyset$ can be treated similarly, see [20].

Let $\mathbb{W}^q = \{w \in \mathbb{W} \,|\, \Lambda(w) \leq \sum_{j=1}^{q} b_{r_j}\}$, $\mathbb{S}^q = \mathbb{W}^q \backslash \mathbb{W}^{q-1}$, and $\mathbb{T}^q = \{w \in \mathbb{S}^q \,|\, \frac{\beta_{r_{q-1}}}{b_{r_{q-1}}} < \frac{w - \sum_{i=1}^{q-1} \beta_{r_i}}{\Lambda(w) - \sum_{i=1}^{q-1} b_{r_i}}\}$ for $q \in \{1, \ldots, s\}$.

Theorem 5. *When lifting continuous variables from* 1 *in (14), we have that*

1. *If $\mathbb{T}^q = \emptyset$ then $\frac{\beta_{r_q}}{b_{r_q}} = \frac{\beta_{r_{q-1}}}{b_{r_{q-1}}}$*
2. *If $\mathbb{T}^q \neq \emptyset$ then $\frac{\beta_{r_q}}{b_{r_q}} = \max\{\frac{w - \sum_{j=1}^{q-1} \beta_{r_j}}{\Lambda(w) - \sum_{j=1}^{q-1} b_{r_j}} \,|\, w \in \mathbb{T}^q\}$*

for $q \in \{1, \ldots, s\}$, *where we define* $\frac{\beta_{r_0}}{b_{r_0}} = 0$. □

Table 1 gives an algorithm that computes the lifting coefficients of continuous variables fixed at 1. The algorithm requires $\mathbb{W}_+$, the function Λ and the initial coefficients b_i of the continuous variables to be lifted. It outputs θ_i and the desired lifting coefficients β_i are computed as $\beta_i = \theta_i b_i$ for $i = 1, \ldots, m$.

Table 1. Algorithm for the lifting of continuous variables from 1

```
Lift(W,Λ,b)

θ_0 = 0, β^old = 0, b^old = 0

For i = 1 : m

        b^new = b^old + b_i
        θ_i = θ_{i-1}

        For Each w ∈ W s.t. b^old < Λ(w) ≤ b^new
                θ_i = max{θ_i, (w − β^old)/(Λ(w) − b^old)}
        End For

        b^old = b^new
        β^old = β^old + θ_i b_i

End For
```

Let w_1 and w_2 be distinct points of $\mathbb{W}_+$. We say that w_1 is dominated by w_2 if $w_2 > w_1$ and $\Lambda(w_1) \geq \Lambda(w_2)$. It can be shown that it suffices in the algorithm of Table 1 to consider the nondominated points in the domain of the function Λ. This result is important because it allows us to compute the sets $\mathbb{T}^q$ of Theorem 5 without performing any sorting of the ordinates of Λ. From this observation, we obtain the complexity of the algorithm.

Theorem 6. *The algorithm presented in Table 1 runs in time* $O(m + |\mathbb{W}_+|)$. □

This algorithm is pseudo-polynomial when the inequalities we lift have no particular structure. However, for structured inequalities like covers, it is polynomial since $|\mathbb{W}_+|$ is polynomially bounded in m.

Example 2 (continued). Consider the polytope PS presented in Example 2. Let $N_1 = \{1, 2\}$. The inequality

$$4x_1 + 4x_2 + 3x_3 + 3x_4 + 3x_5 + 2x_6 + 2x_7 + 2x_8 \leq 12 \tag{15}$$

is valid and defines a facet of $PS(\emptyset, \emptyset, \emptyset, N_1)$. In fact,

$$4x_2 + 3x_3 + 3x_4 + 3x_5 + 2x_6 + 2x_7 + 2x_8 \leq 8$$

is a Chvátal-Gomory inequality with weight $\frac{1}{6}$ in $PS(\emptyset, \{1\}, \emptyset, N_1)$ and can be extended into a valid inequality of $PS(\emptyset, \emptyset, \emptyset, N_1)$ by lifting x_1, leading to (15). It can be shown that (15) is a facet of the associated polytope. So assume that we want to lift from 1 the continuous variables y_2 and y_1 in this order (i.e. $r_1 = 2$ and $r_2 = 1$) using Theorem 5. We first need to compute the function Λ associated with (15). Its values are represented in Table 2.

Table 2. Function Λ associated with (15)

w	1	2	3	4	5	6	7	8	9	10	11
$\Lambda(w)$	5	8	18	25	31	38	48	55	61	∞	78

We compute the first lifting coefficient using Theorem 5. We have that $\mathbb{T}^1 = \{1\}$. Therefore $\frac{\beta_{r_1}}{b_{r_1}} = \max\{\frac{w}{\Lambda(w)} \,|\, w \in \mathbb{T}^1\} = \frac{1}{5}$. Next, we compute $\mathbb{S}^2 = \{2, 3\}$ and $\mathbb{T}^2 = \{2\}$. It follows that

$$\frac{\beta_{r_2}}{b_{r_2}} = \max\{\frac{w - \frac{7}{5}}{\Lambda(w) - 7} \,|\, w \in \mathbb{T}^2\} = \frac{3}{5}.$$

This shows that the inequality

$$4x_1 + 4x_2 + 3x_3 + 3x_4 + 3x_5 + 2x_6 + 2x_7 + 2x_8 + \frac{48}{5} y_1 + \frac{7}{5} y_2 \leq 12 + \frac{48}{5} + \frac{7}{5}$$

is a facet of PS. In fact, this inequality is (9) from Example 2. □

4 Superlinear Lifting

Our algorithm is not polynomial because $|\mathbb{W}_+|$ can be large. There are three directions we can consider to diminish the impact of this pseudo-polynomiality. The first is to consider families of inequalities for which $|\mathbb{W}_+|$ is polynomially bounded (like covers). The second is to determine conditions that would make the function Λ yield an easier lifting process (and a simpler algorithm than the one presented before). The third is to impose conditions on the coefficients of the variables to be lifted so that the lifting algorithm can be simplified. Each of these three approaches is fruitful in its own way. We will describe the second one in this section, leaving the third one for Sect. 5, see [21] for details.

Definition 1. *The function $\Lambda(w)$ associated with (14) in PS^* and the inequality (14) are said to be superlinear if for $w \geq w^*$, $w^*\Lambda(w) \geq w\Lambda(w^*)$ with $w^* = \max argmin\{\Lambda(w) \,|\, w \in \mathbb{W}_+\}$. We call w^* the superlinearity point.*

Example 2 (continued). Inequality (15) is not superlinear. In fact, we can see from Table 2 that $w^* = 1$ and $\Lambda(1) = 5$. In order for the function Λ to be

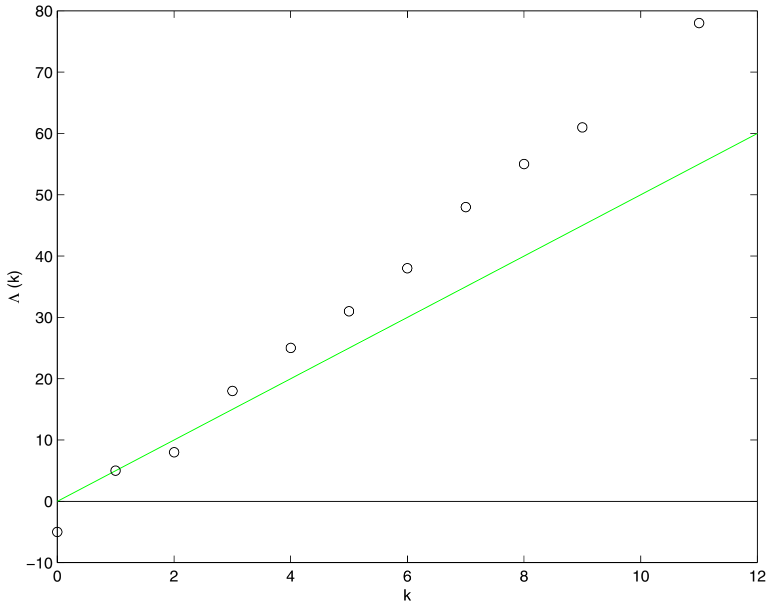

Fig. 1. Function Λ associated with (15)

superlinear, we should have $\Lambda(w) \geq w\Lambda(1)$ which is not true when $w = 2$. In Fig. 1, we see that Λ is not superlinear because the point $(2, \Lambda(2))$ lies beneath the line containing the points $(0, 0)$ and $(w^*, \Lambda(w^*))$. □

Superlinear inequalities play a pivotal role in our study since when continuous variables are lifted from 1, the inequality generated always satisfies the ratio property. Moreover, once their superlinearity point is known, the lifting from 1 of any continuous variable becomes trivial. Let $p = \max\{i \in \{0, \ldots, s\} \mid \sum_{j=1}^{i} b_{r_j} < \Lambda(w^*)\}$.

Theorem 7. *If (14) is superlinear, then in the lifting of continuous variables from* 1 *we obtain* $\frac{\beta_{r_i}}{b_{r_i}} = 0$ *for* $i = 1, \ldots, p$ *and* $\frac{\beta_{r_i}}{b_{r_i}} = \frac{w^*}{\Lambda(w^*) - \sum_{j=1}^{p} b_{r_j}}$ *for* $i = p+1, \ldots, s$. □

Note that Theorem 7 can be used to generate facets of PS. First, we fix all the continuous variables to 0 or 1, then we generate a facet of the resulting pure 0-1 knapsack polytope that we finally lift with respect to the continuous variables. This scheme is restricted by Assumption 3 so that, after we perform the lifting of continuous variables, we are left with a facet of $PS(M_0, \emptyset, \emptyset, \emptyset)$ where $M_0 = \{i \in M \mid a_i \geq d^*\}$. Surprisingly, there exists a closed form expression for lifting the members of M_0 that is presented in the next theorem. For $a \in \mathbb{R}$, let $(a)^+ = \max\{a, 0\}$.

Theorem 8. *Let (N_0, N_1) be a partition of N, $N_1 = \{1, \ldots, n_1\}$ and $M_0 = \{i \in M \,|\, a_i \geq d - \sum_{j \in N_1} a_j\}$. Assume that (14) is superlinear and is not a multiple of (2) and that $p < n_1$. Then*

$$\sum_{j \in M^*} \alpha_j x_j + \sum_{j=p+1}^{n_1} \bar{\theta} b_{r_j} y_{r_j} + \sum_{j \in M_0} (\delta^* + \bar{\theta}(a - d^* - b)^+) x_j \leq \delta^* + \sum_{j=p+1}^{n_1} \bar{\theta} b_{r_j} \quad (16)$$

is a facet of $PS(\emptyset, \emptyset, N_0, \emptyset)$ where $b = \sum_{j=1}^{p} b_{r_j}$ and $\bar{\theta} = \frac{w^*}{\Lambda(w^*) - b}$. □

The practical relevance of Theorem 8 depends on our ability to find families of pure 0-1 facets that are superlinear. We will show next that covers are superlinear. Let $P = conv\{x \in \{0,1\}^m \,|\, \sum_{j \in M} a_j x_j \leq d\}$. Let (C_1, U, C_2) be a partition of M and C_1 be a minimal cover in $P(U, C)$ where $x_i = 0$ for $i \in U$ and $x_i = 1$ for $i \in C_2$. We say that the inequality

$$\sum_{j \in C_1} x_j + \sum_{j \in U} \alpha_j x_j + \sum_{j \in C_2} \alpha_j x_j \leq |C_1| - 1 + \sum_{j \in C_2} \alpha_j \quad (17)$$

is a partitioned cover inequality based on (C_1, U, C_2) if it is obtained from the inequality $\sum_{j \in C_1} x_j \leq |C_1| - 1$ by sequentially lifting the members of U from 0 and then the members of C_2 from 1.

Theorem 9. *Let (17) be a facet of P that is a partitioned cover based on (C_1, U, C_2). Then $\Lambda(w) \geq w\Lambda(1)$ for $w \in \mathbb{W}_+$.* □

We already know, from the algorithm presented in Table 1, how to lift 0-1 covers in polynomial time. Together with Theorem 8, Theorem 9 provides a faster technique to construct facets of PS based on 0-1 covers. We illustrate this technique in the next example.

Example 2 (continued). Let $M_0 = \{3, 4\}$, $M_1 = \{1, 2\}$, $N_0 = \{1\}$ and $N_1 = \{2\}$. The cover inequality $x_5 + x_6 + x_7 + x_8 \leq 3$ is a facet of PS^*. It can be turned into the partitioned cover facet of $PS(\emptyset, \emptyset, N_0, N_1)$

$$3x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 + x_6 + x_7 + x_8 \leq 8$$

by lifting the variables x_3, x_4, x_1 and x_2 in this order. By Theorem 9, this inequality is superlinear and we have also that $\Lambda(1) = 2$. Therefore,

$$3x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 + x_6 + x_7 + x_8 + \frac{7}{2} y_2 \leq 8 + \frac{7}{2}$$

is a facet of PS, which is inequality (7). □

We have shown that the lifting of continuous variables from 1 in a superlinear inequality always leads to an inequality that satisfies the ratio property. Moreover, as stated in the next theorem, these are the only inequalities for which this can be guaranteed.

Theorem 10. *Assume (14) is not superlinear. Let $\bar{\Lambda} = \min\{\Lambda(w) \,|\, w^* \Lambda(w) < w\Lambda(w^*)$ and $w \in \mathbb{W}_+\}$ and $\bar{w} = \max\, argmin\{\Lambda(w) \,|\, w^* \Lambda(w) < w\Lambda(w^*)$ and $w \in \mathbb{W}_+\}$. If $b_{r_1} = \Lambda(w^*)$ and $b_{r_2} = \bar{\Lambda} - \Lambda(w^*)$ then $0 < \frac{\beta_{r_1}}{b_{r_1}} < \frac{\beta_{r_2}}{b_{r_2}}$.* □

5 Pseudo-superlinear Lifting

Although lifted nonsuperlinear inequalities do not necessarly have the ratio property, we would like to understand when they do. Refer to Table 2 and the lifting of (15). The only reason we obtain two different lifting ratios is because b_{r_1} is too small which forces the point $(2, \Lambda(2))$ out of $\mathbb{S}^1$. This suggests that the ratio property can possibly be salvaged if the first continuous variable lifted has a big coefficient.

Definition 2. *The pseudo-superlinearity ratio of the function $\Lambda(w)$ associated with (14) in PS^* and the inequality (14) is* $\max\{\frac{w}{\Lambda(w)} \mid w \in \mathbb{W}_+\}$. *Their pseudo-superlinearity point $\tilde{w}$ is* $\min\{w \in \mathbb{W}_+ \mid \frac{w}{\Lambda(w)} = \tilde{\theta}\}$.

The pseudo-superlinearity point can be seen to be a substitute for the superlinearity point when the inequality is not superlinear. This statement is supported by the following characterization of superlinear inequalities.

Proposition 4. *$\Lambda(w)$ is superlinear if and only if $w^* = \tilde{w}$.* □

Therefore, we can generalize Theorem 8 using $\tilde{w}$ as a substitue for w^*.

Theorem 11. *Let (N_0, N_1) be a partition of N and $M_0 = \{i \in M \mid a_i \geq d - \sum_{j \in N_1} a_j\}$. Assume that (14) is not a multiple of (2), $\sum_{j=1}^{p+1} b_{r_j} \geq \Lambda(\tilde{w})$ and $p < n_1$. Then (16) is a facet of $PS(\emptyset, \emptyset, N_0, \emptyset)$ where $b = \sum_{j=1}^{p} b_{r_j}$, $\bar{\theta} =$* $\max\{\frac{w}{\Lambda(w)-b} \mid w \in \mathbb{W}_+ \cap [w^*, \tilde{w}]\}$ *and* $\bar{w} = \min\{w \in \mathbb{W}_+ \cap [w^*, \tilde{w}] \mid \frac{w}{\Lambda(w)-b}\}$. □

Note that Theorem 11 is a result about asymptotic lifting. In fact, if we assume that $b_{r_1} \geq \Lambda(\tilde{w})$, we can conclude that $\frac{\beta_{r_i}}{b_{r_i}} = \tilde{\theta}$ for $i = 1, \ldots, s$. On the other hand, if the inequality we start from is superlinear, we have $w^* = \bar{w}$ and Theorem 11 reduces to Theorem 8 since the condition $\sum_{j=1}^{p+1} b_{r_j} \geq \Lambda(\tilde{w})$ becomes void and $\bar{\theta} = \frac{w^*}{\Lambda(w^*)-b}$. Next, we present an example of the asymptotic version of Theorem 11.

Example 2 (continued). Consider inequality (15), which we already observed is not superlinear. We can easily establish that its pseudo-superlinearity ratio is $\tilde{\theta} = \frac{1}{4}$ and that $\Lambda(\tilde{w}) = 8$. Now suppose we want to sequentially lift from 1 the continuous variables y_1 and y_2 in this order (i.e. $r_1 = 1$ and $r_2 = 2$). Since $b_{r_1} \geq \Lambda(\tilde{w})$, we can apply Theorem 11 to obtain that inequality (8) from Example 2,

$$4x_1 + 4x_2 + 3x_3 + 3x_4 + 3x_5 + 2x_6 + 2x_7 + 2x_8 + \frac{1}{4}(16y_1 + 7y_2) \leq 12 + \frac{23}{4}$$

is a facet of PS. □

Note that sometimes it is easy to find the pseudo-superlinearity ratio. As an example, the weight inequalities introduced by Weismantel [22] have $\tilde{\theta} = 1$. The lifting of continuous variables is therefore easy for them provided that the coefficients of the variables to be lifted are large. Large coefficients arise, for example, in the transformation of PT into PS presented in Sect. 2. This simple observation gives rise to the following refinement of a theorem by Marchand and Wolsey [16]. Let M_0, M_1 be nonintersecting subsets of M and define $PT(M_0, M_1) = \{x \in \{0,1\}^m \,|\, \sum_{j \in M} a_j x_j \leq d, x_j = 0 \,\forall j \in M_0, x_j = 1 \,\forall j \in M_1\}$.

Theorem 12. *Let $M_0 = \{i \in M \,|\, a_i \geq d\}$. Assume (14) is a facet of $PT(M_0, \emptyset)$ which is not a multiple of (2). Then the inequality*

$$\sum_{j \in M^*} \frac{\alpha_j}{\tilde{\theta}} x_j + \sum_{j \in M_0} \left(\frac{\delta^*}{\tilde{\theta}} + (a_j - d^*)^+\right) x_j \leq \frac{\delta^*}{\tilde{\theta}} + s$$

is a facet of PT. □

Note that $\tilde{\theta} = 1$ for the weight inequalities and $\tilde{\theta} = \frac{1}{\Lambda(1)}$ for partitioned covers.

6 Concluding Remarks and Further Research

We investigated the lifting of continuous variables for the mixed 0-1 knapsack polytope. Although our algorithm is in general only pseudo-polynomial, we showed that it is polynomial for a particular class of well known and extensively used facets of the 0-1 knapsack polytope and we showed how to decrease its computational burden by introducing the concept of superlinearity. We also have shown how the lifting is made easier when the continuous variables have large coefficients. Currently, we are evaluating the practical impact of these results in the context of cutting plane algorithms. Our mixed integer cuts are not only strong but robust in the sense that a violated mixed integer cut of the type given in this paper can always be found when a basic solution to the linear programming relaxation violates integrality. This observation leads to a simplex-based algorithm for which finite convergence can be proved [19]. We are currently testing the significance of using these inequalities in a branch-and-cut algorithm designed to solve general 0-1 mixed integer programs.

References

1. E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Mathematical Programming*, 58:295–324, 1993.
2. E. Balas and E. Zemel. Facets of the knapsack polytope from minimal covers. *SIAM Journal on Applied Mathematics*, 34:119–148, 1978.
3. E. Balas. Facets of the knapsack polytope. *Mathematical Programming*, 8:146–164, 1975.
4. T. Christof and A. Löbel. PORTA : A POlyhedron Representation Transformation Algorithm. `http://www.zib.de/Optimization/Software/Porta/`, 1997.
5. H.P. Crowder, E.L. Johnson, and M.W. Padberg. Solving large-scale zero-one linear programming problems. *Operations Research*, 31:803–834, 1983.

6. I.R. de Farias, Johnson E.L., and Nemhauser G.L. A polyhedral study of the cardinality constrained knapsack problem. Technical Report 01-05, Georgia Institute of Technology, 2001.
7. I.R. de Farias, E.L. Johnson, and G.L. Nemhauser. A generalized assignment problem with special ordered sets : A polyhedral approach. *Mathematical Programming*, 89:187–203, 2000.
8. I.R. de Farias, E.L. Johnson, and G.L. Nemhauser. Facets of the complementarity knapsack polytope. *To appear in Mathematics of Operations Research.*
9. I.R. de Farias and G.L. Nemhauser. A family of inequalities for the generalized assignment polytope. *Operations Research Letters*, 29:49–51, 2001.
10. I.R. de Farias. *A Polyhedral Approach to Combinatorial Problems.* PhD thesis, School of Industrial and Systems Engineering, Georgia Institute of Technology, 1995.
11. R.E. Gomory. An algorithm for the mixed integer problem. Technical Report RM-2597, RAND Corporation, 1960.
12. R.E. Gomory. Some polyhedra related to combinatorial problems. *Linear Algebra and Its Applications*, 2:451–558, 1969.
13. Z. Gu, G.L. Nemhauser, and M.W.P. Savelsbergh. Lifted cover inequalities for 0-1 integer programs: Computation. *INFORMS Journal on Computing*, 10:427–437, 1998.
14. Z. Gu. *Lifted Cover Inequalities for 0-1 and Mixed 0-1 Integer Programs.* PhD thesis, School of Industrial and Systems Engineering, Georgia Institute of Technology, 1995.
15. P.L. Hammer, E.L. Johnson, and U.N. Peled. Facets of regular 0-1 polytopes. *Mathematical Programming*, 8:179–206, 1975.
16. H. Marchand and L.A. Wolsey. The 0-1 knapsack problem with a single continuous variable. *Mathematical Programming*, 85:15–33, 1999.
17. G.L. Nemhauser and L.A. Wolsey. A recursive procedure for generating all cuts for 0-1 mixed integer programs. *Mathematical Programming*, 46:379–390, 1990.
18. M.W. Padberg. On the facial structure of set packing polyhedra. *Mathematical Programming*, 5:199–215, 1973.
19. J.-P. P. Richard, I.R. de Farias, and G.L. Nemhauser. A simplex based algorithm for 0-1 mixed integer programming. Technical Report 01-09, Georgia Institute of Technology, 2001.
20. J.-P. P. Richard, I.R. de Farias, and G.L. Nemhauser. Lifted inequalities for 0-1 mixed integer programming : Basic theory and algorithms. Technical report, Georgia Institute of Technology, (in preparation).
21. J.-P. P. Richard, I.R. de Farias, and G.L. Nemhauser. Lifted inequalities for 0-1 mixed integer programming : Superlinear lifting. Technical report, Georgia Institute of Technology, (in preparation).
22. R. Weismantel. On the 0/1 knapsack polytope. *Mathematical Programming*, 77:49–68, 1997.
23. L.A. Wolsey. Faces for a linear inequality in 0-1 variables. *Mathematical Programming*, 8:165–178, 1975.
24. L.A. Wolsey. Facets and strong valid inequalities for integer programs. *Operations Research*, 24:367–372, 1976.
25. E. Zemel. Lifting the facets of zero-one polytopes. *Mathematical Programming*, 15:268–277, 1978.