On bounded pitch inequalities for the min-knapsack polytope

Yuri Faenza∗  Igor Malinović†  Monaldo Mastrolilli‡  Ola Svensson†

January 29, 2018

Abstract

In the min-knapsack problem one aims at choosing a set of objects with minimum total cost and total profit above a given threshold. In this paper, we study a class of valid inequalities for min-knapsack known as bounded pitch inequalities, which generalize the well-known unweighted cover inequalities. While separating over pitch-1 inequalities is NP-hard, we show that approximate separation over the set of pitch-1 and pitch-2 inequalities can be done in polynomial time. We also investigate integrality gaps of linear relaxations for min-knapsack when these inequalities are added. Among other results, we show that, for any fixed $t$, the $t$-th CG closure of the natural linear relaxation has the unbounded integrality gap.

Keywords: bounded pitch, integrality gap, min-knapsack

1 Introduction

The min-knapsack problem (MinKnap)

$$\min c^T x \quad \text{s.t.} \quad p^T x \geq 1, \ x \in \{0,1\}^n$$

is the variant of the max knapsack problem (MaxKnap) where, given a cost vector $c$ and a profit vector $p$, we want to minimize the total cost given a lower bound on the total profit. MinKnap is known to be NP-Complete, even when $p = c$. Moreover, it is easy to see that the classical FPTAS for MaxKnap [9, 13] can be adapted to work for MinKnap, thus completely settling the complexity of MinKnap.

However, pure knapsack problems rarely appear in applications. Hence, one aims at developing techniques that remain valid when less structured constraints are added on top of the original knapsack one. This can be achieved by providing strong linear relaxations for the problem: then, any additional linear constraint can be added to the formulation, providing a good starting point for any branch-and-bound procedure. The most common way to measure the strength of a linear relaxation is by measuring its integrality gap, i.e. the maximum ratio between the optimal solutions of the linear and the integer programs (or of its inverse if the problem is in minimization form) over all the objective functions.

∗IEOR, Columbia University, USA. Email: yf2414@columbia.edu
†SB/IC, EPFL, Switzerland. Email: firstname.lastname@epfl.ch
‡IDSIA, USI-SUPSI, Switzerland. Email: monaldo@idsia.ch
Surprisingly, if we aim at obtaining linear relaxations with few inequalities and bounded integrality gap, \textsc{MinKnap} and \textsc{MaxKnap} seem to be very different. Indeed, the standard linear relaxation for \textsc{MaxKnap} has integrality gap 2, and this can be boosted to \((1 + \epsilon)\) by an extended formulation with \(n^{O(1/\epsilon)}\) many variables and constraints, for \(\epsilon > 0\) [2]. Conversely, the standard linear relaxation for \textsc{MinKnap} has unbounded integrality gap, and this remains true even after \(\Theta(n)\) rounds of the Lasserre hierarchy [12]. No linear relaxation for \textsc{MinKnap} with polynomially many constraints and constant integrality gap can be obtained in the original space [6]. It is an open problem whether an extended relaxation with these properties exists.

Recent results showed the existence [1] and gave an explicit construction [7] of a linear relaxation for \textsc{MinKnap} of quasi-polynomial size with integrality gap \(2 + \epsilon\). This is obtained by giving an approximate formulation for \textit{Knapsack Cover inequalities} (KC) (see [4] and the references therein). Adding those (exponentially many) inequalities gives an integrality gap of 2, and can be approximately separated [4]. The bound on the integrality gap is tight, even in the simpler case when \(p = c\). One can then look for other classes of well-behaved inequalities that can be added to further reduce the integrality gap. A prominent family is given by the so called \textit{bounded pitch inequalities} [3] defined in Section 2. Here, we remark that the \textit{pitch} is a parameter measuring the complexity of an inequality, and the associated separation problem is NP-Hard already for pitch-1. The pitch-1 inequalities are often known in the literature as \textit{unweighted cover inequalities} (see e.g. [1]).

In this paper, we study structural properties and separability of bounded pitch inequalities for \textsc{MinKnap}, and the strength of linear relaxations for \textsc{MinKnap} when they are added. Let \(\mathcal{F}\) be the set given by pitch-1, pitch-2, and inequalities from the linear relaxation of (1). We first show that, for any arbitrarily small precision, we can solve in polynomial time the \textit{weak separation problem} for the set \(\mathcal{F}\). Even better, our algorithm either certifies that the given point \(x^*\) violates an inequality from \(\mathcal{F}\), or outputs a point that satisfies all inequalities from \(\mathcal{F}\) and whose objective function value is arbitrarily close to that of \(x^*\). We define such an algorithm as a \((1 + \epsilon)\)-oracle in Section 2; see Section 3 for the construction. A major step of our procedure is showing that non-redundant pitch-2 inequalities have a simple structure.

It is then a natural question whether bounded pitch inequalities can help to reduce the integrality gap below 2. We show that, when \(p = c\), if we add to the linear relaxation of pitch-1 and pitch-2 inequalities, the integrality gap is bounded by 3/2; see Section 4.1. However, this is false in general. Indeed, we also prove that KC plus bounded pitch inequalities do not improve upon the integrality gap of 2; see Section 4.3. Moreover, bounded pitch alone can be much weaker than KC: we show that, for each fixed \(k\), the integrality gap may be unbounded even if all pitch-\(k\) inequalities are added. Using the relation between bounded pitch and Chvátal-Gomory (CG) closures established in [3], this implies that, for each fixed \(t\), the integrality gap of the \(t\)-th CG closure can be unbounded; see Section 4.2. For an alternative proof that having all KC inequalities bounds the integrality gap to 2 see Section 4.3.

2 Basics

A \textsc{MinKnap} instance is a binary optimization problem of the form (1), where \(p, c \in \mathbb{Q}^n\) and we assume \(0 \leq p_1 \leq p_2 \leq \cdots \leq p_n \leq 1, 0 < c_i \leq 1, \forall i \in [n]\). We will often deal with its \textit{natural linear relaxation}

\[
\begin{align*}
\min c^T x & \quad \text{s.t.} \quad p^T x \geq 1, \quad x \in [0,1]^n.
\end{align*}
\]
The NP-Hardness of MinKnap immediately follows from the fact that MaxKnap is NP-Hard [10], and that a MaxKnap instance

\[
\max v^T x \text{ s.t. } w^T x \leq 1, \ x \in \{0,1\}^n.
\]

(3)
can be reduced into a MinKnap instance \([1]\) as follows: each \(x \in \{0,1\}^n\) is mapped via \(\pi: \mathbb{R}^n \to \mathbb{R}^n\) with \(\pi(x) = 1 - x\); \(p_i = \frac{w_i}{c_i + 1} \) and \(c_i = v_i\) for \(i \in [n]\). Note that the reduction is not approximation-preserving.

We say that an inequality \(w^T x \geq \beta\) with \(w \geq 0\) is dominated by a set of inequalities \(\mathcal{F}\) if \(w^T x \geq \beta'\) can be written as a conic combination of inequalities in \(\mathcal{F}\) for some \(\beta' \geq \beta\) and \(w' \geq w\). \(w^T x \geq \beta\) is undominated if any set of valid inequalities dominating \(w^T x \geq \beta\) contains a positive multiple of it.

Consider a family \(\mathcal{F}\) of inequalities valid for \((1)\). We refer to [8] for the definition of weak separation oracle, which is not used in this paper. We say that \(\mathcal{F}\) admits a \((1 + \epsilon)\)-oracle if, for each fixed \(\epsilon > 0\), there exists an algorithm that takes as input a point \(\bar{x}\) and, in time polynomial in \(n\), either outputs an inequality from \(\mathcal{F}\) that is violated by \(\bar{x}\), or outputs a point \(\bar{y}\), \(\bar{x} \leq \bar{y} \leq (1 + \epsilon)\bar{x}\) that satisfies all inequalities in \(\mathcal{F}\). In particular, if \(\mathcal{F}\) contains the linear relaxation of \((1)\), \(0 \leq \bar{y} \leq \bar{x}\).

Let \(\sum_{i \in T} w_i x_i \geq \beta\) be a valid inequality for \((1)\), with \(w_i > 0\) for all \(i \in T\). Its pitch is the minimum \(k\) such that, for each \(I \subseteq T\) with \(|I| = k\), we have \(\sum_{i \in I} w_i \geq \beta\). Undominated pitch-1 inequalities are of the form \(\sum_{i \in I} x_i \geq 1\). Note that the map from MaxKnap to MinKnap instances defined above gives a bijection between minimal cover inequalities

\[
\sum_{i \in I} x_i \leq |I| - 1
\]

for MaxKnap and undominated pitch-1 inequalities for the corresponding MinKnap instance. Since, given a MaxKnap instance, it is NP-Hard to separate minimal cover inequalities \([11]\), we conclude the following.

**Theorem 1.** It is NP-Hard to decide whether a given point satisfies all valid pitch-1 inequalities for a given MinKnap instance.

Given a set \(S \subseteq [n]\), such that \(\beta := 1 - \sum_{i \in S} p_i > 0\), the Knapsack cover inequality associated to \(S\) is given by

\[
\sum_{i \in [n] \setminus S} \min\{p_i, \beta\} x_i \geq \beta
\]

(4)
and it is valid for \((1)\).

For a set \(S \subseteq [n]\), we denote by \(\chi^S\) its characteristic vector. An \(\epsilon\)-approximate solution for a minimization integer programming problem is a solution \(\bar{x}\) that is feasible, and whose value is at most \((1 + \epsilon)\) times the value of the optimal solution. An algorithm is called a polynomial time approximation scheme (PTAS) if for each \(\epsilon > 0\) and any instance of the problem it returns an \(\epsilon\)-approximate solution in time polynomial in the size of the input. If in addition the running time is polynomial in \(1/\epsilon\), then the algorithm is a fully polynomial time approximation scheme (FPTAS).

Given a rational polyhedron \(P = \{x \in \mathbb{R}^n : Ax \geq b\}\) with \(A \in \mathbb{Z}^{m \times n}\) and \(b \in \mathbb{Z}^m\), the first Chvátal-Gomory (CG) closure \([5]\) of \(P\) is defined as follows:

\[P^{(1)} = \{x \in \mathbb{R}^n : [\lambda^T A]x \geq [\lambda^T b], \ \forall \lambda \in \mathbb{R}^m\}.\]

Equivalently, one can consider all \(\lambda \in [0,1]^m\) such that \(\lambda^T A \in \mathbb{Z}^n\). For \(t \in \mathbb{Z}_{\geq 2}\), the \(t\)-th CG closure of \(P\) is recursively defined as \(P^{(t)} = (P^{(t-1)})^{(1)}\). The CG closure is an important tool for solving integer programs, see again \([5]\).
3 A \( (1 + \epsilon) \)-oracle for pitch-1 and pitch-2 inequalities

In this section, we show the following:

**Theorem 2.** Given a MinKnap instance **(1)**, there exists a \((1 + \epsilon)\)-oracle for the set \( F \) containing: all pitch-1 inequalities, all pitch-2 inequalities and all inequalities from the natural linear relaxation of **(1)**.

We start with a characterization of inequalities of interest for Theorem 2.

**Lemma 3.** Let \( K \) be the set of feasible solutions of a MinKnap instance **(1)**. All pitch-2 inequalities valid for \( K \) are implied by the set composed of:

i) Non-negativity constraints \( x_i \geq 0 \) for \( i \in [n] \);

ii) All valid pitch-1 inequalities;

iii) All inequalities of the form

\[
\sum_{i \in I_1} x_i + 2 \sum_{i \in I_2} x_i \geq 2
\]

where \( I \subseteq [n] \), \( |I| \geq 2 \), \( \beta(I) := 1 - \sum_{i \notin [n]\setminus I} p_i \), \( I_1 := \{i \in I : p_i < \beta(I)\} \neq \emptyset \) and \( I_2 := I \setminus I_1 \).

The inequalities in **(ii)** are pitch-2 and valid.

Proofs of Lemma 5 and Theorem 2 are given in Section 3.1 and Section 3.2, respectively.

### 3.1 Restricting the set of valid pitch-2 inequalities

We will build on two auxiliary statements in order to prove Lemma 3.

**Claim 1.** If \( w^T x \geq \beta \) and \( u^T x \geq \beta \) are distinct inequalities valid for and \( u \geq w \), then the latter inequality is dominated by the former.

**Proof.** \( u^T x \geq \beta \) can be obtained summing nonnegative multiples of \( w^T x \geq \beta \) and \( x_i \geq 0 \) for \( i \in [n] \), which are all valid inequalities.

\[ \square \]

**Claim 2.** Let

\[
\sum_{i \in T_1} x_i + 2 \sum_{i \in T_2} x_i \geq 2
\]

be a valid inequality for MinKnap, with \( T_1 \cap T_2 = \emptyset \) and \( T_1, T_2 \subseteq [n] \). Then, **(6)** is dominated by the inequality in **(5)** with \( I = T_1 \cup T_2 \).

**Proof.** One readily verifies that Inequality **(5)** with \( I \) as above is valid. Suppose now that \( i \in T_1 \setminus I_1 \). Then the integer solution that takes all elements in \((n) \setminus I \cup \{i\}\) is feasible for MinKnap, but it does not satisfy **(5)**, a contradiction. Hence \( T_1 \subseteq I_1 \). Since \( T_2 = I \setminus T_1 \supseteq I \setminus I_1 = I_2 \), **(5)** dominates **(6)** componentwise, and the thesis follows by Claim 1.

\[ \square \]
Proof of Lemma 3. The fact that an inequality of the form (5) is pitch-2 and valid is immediate. Because of Claim 2, it is enough to show the thesis with (5) replaced by (6). Consider a pitch-2 inequality valid for $K$:

$$\sum_{i \in T} w_i x_i \geq 1,$$

where $T \subseteq [n]$ is the support of the inequality, $w \in \mathbb{R}^{[T]}_+$. Without loss of generality one can assume that $T = [h]$ for some $h \leq n$ and $w_1 \leq w_2 \leq \cdots \leq w_h$. Since (7) is pitch-2 we have that $w_1 + w_i \geq 1$ for all $i \in [h] \setminus \{1\}$. We can also assume $w_h \leq 1$, since otherwise $\sum_{i \in [h-1]} w_i x_i + x_h \geq 1$ is valid and dominates (7) by Claim 1.

Let $j \in [h]$ be the maximum index such that $w_j < 1$. Note that such $j$ exists, since, if $w_1 \geq 1$, then (7) is a pitch-1 inequality. If $1 - w_1 \leq 1/2$, then, by Claim 1, (7) is dominated by the valid pitch-2 inequality

$$\sum_{i \in [j]} x_i + 2 \sum_{i = j+1}^h x_i \geq 2,$$

which again is of the type (10). Hence $1 - w_1 > 1/2$ and again via Claim 1, (7) is dominated by

$$w_1 x_1 + \sum_{i = 2}^j (1 - w_1) x_i + \sum_{i = j+1}^h x_i \geq 1,$$

since $w_i + w_1 \geq 1$ for all $i \neq 1$, so one has $w_i \geq 1 - w_1 > 1/2$. Thus, we can assume that (7) has the form (9). Note that inequality

$$\sum_{i = 2}^h x_i \geq 1$$

is a valid pitch-1 inequality, since we observed $w_1 < 1$. Therefore, (7) is implied by (8) and (10), taken with the coefficients $w_1$ and $1 - 2w_1$ respectively. Recalling that (8) is a valid pitch-2 inequality of the form (6) concludes the proof. ■

3.2 A $(1 + \epsilon)$-oracle

We will prove Theorem 2 in a sequence of intermediate steps. Our argument extends the weak separation of KC inequalities in [4].

Let $\bar{x}$ be the point we want to separate. Note that it suffices to show how to separate over inequalities i)-ii)-iii) from Lemma 3. Separating over i) is trivial. We first show how to separate over iii).

Claim 3. For $\alpha \in [0,1]$, let $z^\alpha$ be the optimal solution to the following IP $P_\alpha$, and $v(z^\alpha)$ its value:

$$\min_{i \in [n]} \sum_{p_i < \alpha} \bar{x}_i z_i + 2 \sum_{i \in [n]}: p_i \geq \alpha \bar{x}_i z_i \quad s.t. \quad \sum_{i \in [n]} p_i (1 - z_i) \leq 1 - \alpha, \quad z \in \{0,1\}^n.$$

If $v(z^\alpha) < 2$, then $\bar{x}$ violates Inequality (1) with $I := \{i \in [n] : z_i^\alpha = 1\}$, otherwise $\bar{x}$ does not violate any Inequality (5) with $\beta(I) = \alpha$. 

5
\textbf{Proof.} Fix a feasible solution $\bar{\varphi}$ to (11), and let $I := \{i \in [n] : \bar{\varphi}_i = 1\}$. Then:

$$\beta := \beta(I) = 1 - \sum_{i \in [n] \setminus I} p_i = 1 - \sum_{i \in [n]} p_i(1 - \bar{\varphi}_i) \geq \alpha.$$ 

Hence:

$$\sum_{i \in I: p_i < \beta} \bar{\varphi}_i + 2 \sum_{i \in I: p_i \geq \beta} \bar{\varphi}_i = \sum_{i \in [n]: p_i < \beta} \bar{\varphi}_i \bar{\varphi}_i + 2 \sum_{i \in [n]: p_i \geq \beta} \bar{\varphi}_i \bar{\varphi}_i \leq \sum_{i \in [n]: p_i < \alpha} \bar{\varphi}_i \bar{\varphi}_i + 2 \sum_{i \in [n]: p_i \geq \alpha} \bar{\varphi}_i \bar{\varphi}_i = v(\bar{\varphi}),$$

where the central inequality holds at equality if $\alpha = \beta$. Hence, if $v(z^\alpha) < 2$, the inequality with $I := \{i \in [n] : z^\alpha_i = 1\}$ from (11) is violated by $\bar{\varphi}$. Else, all inequalities from (11) with $\beta(I) = \alpha$ are satisfied.

Note that $P_\alpha$ is a MINKNAP instance, hence we can use the appropriate FPTAS to find, for each $\epsilon > 0$, an $\epsilon$-approximate solution for it.

Since all data are rationals, we can assume there exists $q \in \mathbb{N}$ such that, for each $i \in [n]$, $p_i = r_i/q$ for some $r_i \in \mathbb{N}$.

\textbf{Claim 4.} Let $r \in \{r_i + 1 : i \in [n]\}$ and, for $\alpha = r/q$, let $\bar{\varphi}^\alpha$ be the solution output by the FPTAS for problem $P_\alpha$ and $\bar{\varphi}^\alpha$ its objective function value. If $\bar{\varphi}^\alpha < 2$ for some $\alpha$, then $\bar{x}$ violates Inequality (11) with $I = \{i \in [n] : \bar{\varphi}^\alpha_i = 1\}$. Else, $(1 + \epsilon)\bar{x}$ satisfies all inequalities in Lemma 11.

\textbf{Proof.} Let $r = r_i + 1$ for some $i \in [n]$. If $\bar{\varphi}^\alpha < 2$ then by Claim 8 $\bar{x}$ violates the corresponding inequality (11). Otherwise, $\bar{\varphi}^\alpha \geq 2/(1 + \epsilon)$, and $(1 + \epsilon)\bar{x}$ is feasible for any pitch-2 Inequality (11) induced by $I$ with $\beta(I) = \alpha$.

Now let $I^* \subseteq [n]$ with $\beta(I^*) = \frac{r^*}{q} < 1$. There exists $i^* \in [n]$ such that $r_{i^*} < r^* \leq r_{i^* + 1} \leq q$ (with $r_{n+1} = q$). Let $\alpha := \frac{r_{i^* + 1}}{q} \leq \alpha^* := \beta(I^*)$. The set of feasible solutions of $P_\alpha$ contains that of $P_{\alpha^*}$, and $\{i \in [n] : p_i < \alpha\} = \{i \in [n] : p_i < \alpha^*\}$. Hence, $\bar{\varphi}^{\alpha^*} < 2$ implies $\bar{\varphi}^\alpha < 2$. Thus, for separating all inequalities in Lemma 11, it suffices to check (11) for all $\alpha = \frac{r}{q}$ as in the statement of the claim.

The following claim follows in a similar fashion to the previous one by observing that, for $\beta(I^*) = \frac{1}{q}$, (11) separates over undominated pitch-1 inequalities.

\textbf{Claim 5.} Let $\alpha = 1/q$, and $\bar{\varphi}^\alpha$ be the solution output by the FPTAS for problem $P_\alpha$, and $\bar{\varphi}^\alpha$ its objective function value. If $\bar{\varphi}^\alpha < 2$, then $\bar{x}$ violates the pitch-1 inequality with support $I = \{i \in [n] : \bar{\varphi}^\alpha_i = 1\}$. Else, $(1 + \epsilon)\bar{x}$ satisfies all valid pitch-1 inequalities.

Next claim shows how to round a point in the unit cube that almost satisfies all pitch-1 and pitch-2 inequalities, to one that satisfies them and is still contained in the unit cube.

\textbf{Claim 6.} Let $\bar{x} \in [0, 1]^n$ be such that $(1 + \epsilon)\bar{x}$ satisfies all inequalities from Lemma 8 and define $\bar{y} \in \mathbb{R}^n$ as follows: $\bar{y}_i = \min\{1, \frac{1 + \epsilon}{\bar{\varphi}_i} \}$ for $i \in [n]$. Then $\bar{y} \in [0, 1]^n$ and $\bar{y}$ satisfies all inequalities from Lemma 8.
Clearly, we are only interested in valid inequalities. Hence, we can assume that the left-hand side of the inequality computed in \( y \) is lower bounded by \( \sum_{i \in J} \alpha_i \), where \( \alpha_i \) is the coefficient of \( x_i \). First assume there exists \( j \in J \cap I_2 \). Then \( \sum_{i \in J} \alpha_i \geq \alpha_j = 2 \). Similarly, if \( j, j' \in J \), then \( \sum_{i \in J} \alpha_i \geq \alpha_j + \alpha_{j'} \geq 2 \). In both cases, \( y \) satisfies the pitch-2 inequality. Hence, we can assume \( J = \{ j \} \subseteq I_1 \). Then:

\[
\sum_{i \in I} \alpha_i \bar{x}_i \geq \frac{2}{1+\epsilon},
\]

from which we deduce

\[
\sum_{i \in I \setminus \{ j \}} \alpha_i \bar{x}_i \geq \frac{2}{1+\epsilon} - \bar{x}_j \geq \frac{2}{1+\epsilon} - 1 = \frac{1-\epsilon}{1+\epsilon}
\]

and

\[
\sum_{i \in I} \alpha_i \bar{y}_i = \sum_{i \in I \setminus \{ j \}} \alpha_i \bar{y}_i + 1 = \frac{1+\epsilon}{1-\epsilon} \sum_{i \in I \setminus \{ j \}} \alpha_i \bar{x}_i + 1 \geq 2,
\]

as required. A similar (simpler) argument shows that \( \bar{y} \) also satisfies all pitch-1 inequalities \( \sum_{i \in I} x_i \geq 1 \).

Proof of Theorem 2. We can now sum up our \((1+\epsilon)\)-oracle, see Algorithm 1. Correctness and polynomiality follow from the discussion above.

Algorithm 1

1: Let \( \epsilon' = \frac{1}{2\epsilon(1+\epsilon)} \).
2: For \( r \in \{ r_i + 1 : i \in [n] \} \) and for \( \alpha = r/q \), run the FPTAS for \( P_\alpha \) with approximation factor \( \epsilon' \). If any of the output solution \( \bar{z}^\alpha \) has value \( \bar{v}^\alpha < 2 \), output inequality (5) with \( I = \{ i \in [n] : \bar{z}^\alpha = 1 \} \) and stop.
3: For \( \alpha = 1/q \), run the FPTAS for \( P_\alpha \) with approximation factor \( \epsilon' \). If the output solution \( \bar{z}^\alpha \) has value \( \bar{v}^\alpha < 2 \), output inequality \( \sum_{i \in I} x_i \geq 1 \) and stop.
4: Output point \( \bar{y} \) constructed as in Claim 6 with \( \epsilon' \) and stop. Note that \( \bar{x} \leq \bar{y} \leq \frac{1+\epsilon'}{1-\epsilon'} \bar{x} = (1+\epsilon)\bar{x} \).

3.3 Separating inequalities of pitch \( k \geq 3 \) with fixed support

Here, we give an example showing that inequalities of pitch-3 and higher do not have the nice structure of pitch-2. Let

\[
P = \{ x \in [0, 1]^7 : 5x_1 + 6x_2 + 11x_3 + 16x_4 + 17x_5 + 18x_6 + 21x_7 \geq 41 \}.
\]

Inequality \( x_1 + x_3 + x_4 + 2x_5 + x_6 + 2x_7 \geq 3 \) is a facet of the first CG closure \( P^{(1)} \) (although not of the integer hull of \( P \)) and thus a valid pitch-3. Observe that the coefficient of \( x_5 \) is higher than \( x_6 \) in this pitch-3, while it is the opposite in \( P^{(2)} \). Such situations we call inversions and they do not occur in (relevant) pitch-2 inequalities. Inequality \( x_1 + x_2 + 2x_3 + 3x_4 + 4x_5 + 5x_6 + 4x_7 \geq 8 \) is an inverted facet of both the integer hull and the first CG closure.

For later use (in Section 4.4), we observe here that when \( I \subseteq [n] \) is fixed, we can efficiently and exactly solve the separation problem over inequalities with support \( I \) just by solving an LP. Clearly, we are only interested in valid inequalities \( \alpha^T x \geq 1 \) with \( \alpha \geq 0 \) and points \( 0 \leq x^* \leq 1 \).
Let \( \beta = 1 - p([n] \setminus I) \). We can assume \( \beta > 0 \), otherwise there is no valid inequality as above with support \( I \). Call \( J \subseteq I \) massive if \( \sum_{i \in J} p_i \geq \beta \). Consider the following LP:

\[
\begin{align*}
\min & \quad \sum_{i \in I} \alpha_i x_i^* \\
\text{s.t.} & \quad \sum_{i \in J} \alpha_i \geq 1 \quad \text{for all massive } J \subseteq I \\
& \quad \alpha_i \geq 0.
\end{align*}
\]

(13)

Note that, for each feasible solution \( \bar{\alpha} \) to the previous LP, we have that \( \bar{\alpha}^T x \geq 1 \) is a valid inequality for the original MinKnap instance, and conversely that all inequalities with support \( I \) can be obtained in this way. Hence, let \( \alpha^* \) be the optimal solution to the previous LP. If \( (\alpha^*)^T x^* < 1 \), we obtain an inequality whose support is contained in \( I \), that is violated by \( x^* \). The support of the inequality can be extended to \( I \) by setting \( \alpha_i = \epsilon \) for all \( i \in I \) with \( \alpha_i = 0 \).

On the other hand, if \( (\alpha^*)^T x^* \geq 1 \), \( x^* \) satisfies all inequalities with support \( I \).

4 Integrality gap for MinKnap with bounded pitch inequalities

4.1 When \( p = c \)

Theorem 4. Consider an instance of MinKnap with \( p = c \). Denote by \( K \) the linear relaxation of (1) to which all pitch-1 and pitch-2 inequalities have been added. The integrality gap of \( K \) is at most \( 3/2 \).

Proof. Let \( p = c \), and let \( \bar{x} \) be the optimal integer solution to (1). We can assume \( p^T \bar{x} > 1.5 \), else we are done.

Claim 7. The support of \( \bar{x} \) has size 2.

Proof. Let \( k \) be the size of the support of \( \bar{x} \). If \( k = 1 \), then \( \bar{x} \) is also the optimal fractional solution. Now assume \( k \geq 3 \). Remove from \( \bar{x} \) the cheapest item as to obtain \( \bar{x}' \). We have

\[
p^T x' \geq \left( 1 - \frac{1}{k} \right) p^T \bar{x} > \frac{2}{3} \cdot 1.5 = 1,
\]

contradicting the fact that \( \bar{x} \) is the optimal integral solution.

Hence, we can assume that the support of \( \bar{x} \) is given by \( \{i, j\} \), with \( 0 < p_i \leq p_j \leq 1 \). Since \( p_i + p_j > 1.5 \), we deduce \( p_j > .75 \). Since \( p_j \leq 1 \), we deduce \( p_i > .5 \).

Claim 8. Let \( \ell < j \) and \( \ell \neq i \). Then \( p_\ell < .25 \).

Proof. Recall that for \( S \subseteq [n] \) we denote its characteristic vector with \( \chi^S \). If \( 0.25 \leq p_\ell < p_i \), then \( \chi^{\{\ell,j\}} \) is a feasible integral solution of cost strictly less than \( \bar{x} \). Else if \( 0.5 < p_i \leq p_\ell < p_j \), then \( \chi^{\{\ell,i\}} \) is a feasible integral solution of cost strictly less than \( \bar{x} \). In both cases we obtain a contradiction.

Because of the previous claim, we can assume w.l.o.g. \( j = i + 1 \).

Claim 9. \( p_n + \sum_{\ell=1}^{i-1} p_\ell < 1. \)
Proof. Suppose \( p_n + \sum_{\ell=1}^{i-1} p_\ell \geq 1 \). Since \( p_\ell < .25 \) for all \( \ell = 1, \ldots, i - 1 \), there exists \( k \leq i - 1 \) such that \( x_n + \sum_{\ell=1}^{k} p_\ell \in [1, 1.25] \). Hence \( x^{\{1,\ldots,k,n\}} \) is a feasible integer solution of cost at most 1.25, a contradiction.

Because of the previous claim, the pitch-2 inequality \( \sum_{k=i}^{n} x_k \geq 2 \) is valid. The fractional solution of minimum cost that satisfies this inequality is the one that sets \( x_i = x_j = 1 \) (since \( j = i + 1 \)) and all other variables to 0. This is exactly \( \overline{x} \).

4.2 CG closures of bounded rank of the natural MinKnap relaxation

For \( t \in \mathbb{N} \), let \( K^t \) be the linear relaxation of the (1) given by: the original knapsack inequality; non-negativity constraints; all pitch-\( k \) inequalities, for \( k \leq t \).

Lemma 5. For \( t \geq 2 \), the integrality gap of \( K^t \) is at least \( \max\{\frac{1}{2}, \frac{t-2}{t-1}\} \) times the integrality gap of \( K^{t-1} \).

Proof. Fix \( t \geq 2 \), and let \( C \) be the cost of the optimal integral solution to (1). Let \( C/v' \) be the integrality gap of \( K^t \). Since \( v' \) is the optimal value of \( K^t \), by the strong duality theorem (and Caratheodory’s theorem), there exist nonnegative multipliers \( \alpha, \alpha_1, \ldots, \alpha_n, \gamma_1, \ldots, \gamma_{n+1} \) such that the inequality \( c'x \geq v' \) can be obtained as a conic combination of the original knapsack inequality (with multiplier \( \alpha \)), non-negativity constraints (with multipliers \( \alpha_1, \ldots, \alpha_n \)), and at most \( n+1 \) inequalities of pitch at most \( t \) (with multipliers \( \gamma_1, \ldots, \gamma_{n+1} \)). By scaling, we can assume that the rhs of the latter inequalities is 1. Hence \( v' = \alpha + \sum_{t=1}^{t} \gamma_t \).

Claim 10. Let \( d^T x \geq 1 \) be a valid pitch-\( t \) inequality for (1), and assume w.l.o.g. that \( d_1 \leq d_2 \leq \cdots \leq d_n \). Then inequality \( \sum_{i=2}^{n} d_i x_i \geq \max\{\frac{1}{2}, \frac{t-2}{t-1}\} \) is a valid inequality of pitch at most \( t - 1 \) for (1).

Proof. The inequality \( \sum_{i=2}^{n} d_i x_i \geq 1 - d_1 \) is a valid inequality for (1), and by construction it is of pitch at most \( t - 1 \). If \( t \geq 3 \), we obtain \( \sum_{i=2}^{t-1} d_i < 1 \) and consequently \( d_1 < 1/(t - 1) \), from which we deduce

\[
1 - d_1 > 1 - \frac{1}{t - 1} = \frac{t - 2}{t - 1} > \frac{1}{2}.
\]

If conversely \( t = 2 \), by Lemma 3 we can assume w.l.o.g. that \( d_1 = 1/2 \), and we can conclude \( 1 - d_1 = \frac{1}{2} > \frac{t-2}{t-1} \).

Now consider the conic combination with multipliers \( \alpha, \alpha_1, \ldots, \alpha_n, \gamma_1, \ldots, \gamma_{n+1} \) given above, where each inequality of pitch-\( t \) is replaced with the inequality of pitch at most \( t - 1 \) obtained using Claim 10. We obtain an inequality \( (c')^T x \geq v'' \), where one immediately checks that \( c' \leq c \) and

\[
v'' \geq \alpha + \sum_{i=2}^{n+1} \gamma_i \max\left\{\frac{1}{2}, \frac{t-2}{t-1}\right\} \geq \max\left\{\frac{1}{2}, \frac{t-2}{t-1}\right\} \left(\alpha + \sum_{i=1}^{n+1} \gamma_i\right)
\]

Hence the integrality gap of \( K^t \) is

\[
\frac{C}{v'} \geq \frac{C}{v''} \max\left\{\frac{1}{2}, \frac{t-2}{t-1}\right\}
\]

and the thesis follows since the integrality gap of \( K^{t-1} \) is at most \( C/v'' \).
Lemma 6. For a fixed $\epsilon > 0$ and square integers $n \geq 4$, consider the MinKnap instance $K$ defined as follows:

$$\min_{st} \quad \epsilon y + \sqrt{n}z + \sum_{i=1}^{n} x_i$$

$$(n - \sqrt{n})y + \frac{\epsilon}{2}z + \sum_{i=1}^{n} x_i \geq n$$

For every fixed $t \in \mathbb{N}$, the integrality gap of $K^t$ is $\Omega(\sqrt{n})$.

Proof. Because of Lemma 5, it is enough to show that the integrality gap of $K^t$ is $\Omega(\sqrt{n})$. Clearly, the value of the integral optimal solution of the instance is $\sqrt{n} + \epsilon$. We claim that the fractional solution

$$(\bar{y}, \bar{x}, \bar{z}) = \left( \frac{1}{n - \sqrt{n} + 1}, \ldots, \frac{1}{n - \sqrt{n} + 1}, \frac{2}{\sqrt{n}} \right)$$

is a feasible point of $K^t$. Since $\epsilon \bar{y} + \sqrt{n} \bar{z} + \sum_{i=1}^{n} \bar{x}_i = \epsilon + 2 + \frac{\epsilon}{n - \sqrt{n} + 1}$, the thesis follows.

Observe that $(n - \sqrt{n}) \bar{y} + \frac{\epsilon}{2} \bar{z} + \sum_{i=1}^{n} \bar{x}_i = (n - \sqrt{n}) + \frac{\epsilon}{2} + \frac{n}{n - \sqrt{n} + 1} > n$, hence $(\bar{y}, \bar{x}, \bar{z})$ satisfies the original knapsack inequality.

Now consider a valid pitch-1 inequality whose support contains $y$. Since $\bar{y} = 1$, $(\bar{y}, \bar{x}, \bar{z})$ satisfies this inequality. Hence, the only pitch-1 inequalities of interest do not have $y$ in the support. Note that such inequalities must have $z$ in the support, and some of the $x_i$. Hence, all those inequalities are dominated by the valid pitch-1 inequalities

$$z + \sum_{i \in I} x_i \geq 1 \quad \forall I \subseteq [n], |I| = n - \sqrt{n} + 1,$$

which are clearly satisfied by $(\bar{y}, \bar{x}, \bar{z})$.

Theorem 7. For a fixed $q \in \mathbb{N}$, let $CG^q(K)$ be the $q$-th CG closure of the MinKnap instance $K$ as defined in Lemma 4. The integrality gap of $CG^q(K)$ is $\Omega(\sqrt{n})$.

Proof. We will use the following fact, proved (for a generic covering problem) in [3]. Let $t,q \in \mathbb{N}$ and suppose $(\bar{y}, \bar{z}, \bar{x}) \in K^t$. Define point $(y', x', z')$, where each component is the minimum between 1 and $(\frac{t+1}{t})^q$ times the corresponding component of $(\bar{y}, \bar{z}, \bar{x})$. Then $(y', x', z') \in CG^q(K)$. Now fix $t,q$. We have therefore that

$$\epsilon y' + \sqrt{n}z' + \sum_{i=1}^{n} x_i' \leq \left( \frac{t+1}{t} \right)^q \left( \epsilon \bar{y} + \sqrt{n} \bar{z} + \sum_{i=1}^{n} \bar{x}_i \right)$$

and the claim follows in a similar fashion to the proof of Lemma 6.

4.3 When all knapsack cover inequalities are added

In this section we consider the min-knapsack formulation with knapsack cover inequalities (we use KC to denote this LP formulation). In [4] it is shown that the integrality gap of KC is 2. In the following we provide a simpler proof.

Let $\bar{x}$ be a feasible fractional solution for KC of cost $C(\bar{x})$. Starting from $\bar{x}$, we show a simple and fast rounding procedure to obtain a feasible integral solution of cost at most $2C(\bar{x})$. 

10
The rounding procedure: Let \( S = \{ i \in [n] : \bar{x}_i \geq 1/2 \} \). Set \( x_i = 1 \) for any \( i \in S \). Consider the residual variables \( \bar{S} := [n] \setminus S \). The problem is to assign integral values to the residual variables. We call this problem the residual problem (RP). By abusing notation, from now on, let \( \bar{x} \) denote the fractional solution for KC restricted to residual variables. Consider the following residual relaxation (RR):

\[
\min \sum_{i \in \bar{S}} C_i x_i \\
\text{s.t.} \sum_{i \in \bar{S}} p'_i x_i \geq b' \\
0 \leq x_i \leq 1/2
\]

(14) \hspace{3cm} (15) \hspace{3cm} (16)

where \( p'_i = \min\{p_i, b - p(S)\} \) and \( b' = b - p(S) \). Note that \( \bar{x} \) satisfies (RR). So if \( x^* \) is the optimal solution of (RR) then it follows that

\[
\sum_{i \in S} C_i x^*_i \leq \sum_{i \in \bar{S}} C_i \bar{x}_i
\]

Therefore, if it exists an integral solution \( x^{int} \) to (RP) of cost \( C(x^{int}) \leq 2C(x^*) \) then \( C(x^{int}) \leq 2C(\bar{x}) \) and we are done. We can rewrite the residual relaxation (RR) in the following equivalent way:

\[
\min \sum_{i \in \bar{S}} C_i y_i \\
\text{s.t.} \sum_{i \in \bar{S}} p'_i y_i \geq b' \\
0 \leq y_i \leq 1
\]

(17) \hspace{3cm} (18) \hspace{3cm} (19)

Clearly the optimal fractional solution to (RR) can be obtained by ordering the variables according to their densities. W.l.o.g., assume that \( \frac{C_1}{p'_1} \leq \frac{C_2}{p'_2} \leq \ldots \leq \frac{C_{|\bar{S}|}}{p'_t} \) and let \( t + 1 \in [|\bar{S}|] \) be the smallest integer such that \( \sum_{i=1}^{t+1} p'_i \geq b' \) and therefore (recall \( p'_i \leq b' \)):

\[
\sum_{i=1}^{t} p'_i \geq 2b' - p'_{t+1} \geq b'
\]

(20)

Note that the optimal fractional solution to (RR) picks the first \( t \) variables integrally and the last \( t + 1 \) potentially fractional (but it could be integral). It follows that:

\[
\sum_{i=1}^{t} \frac{C_i}{2} \leq \sum_{i \in \bar{S}} C_i x^*_i
\]

(21)

It follows that the integral solution \( x^{int} \) obtained by setting \( x_1 = 1 \) for \( i \in [t] \) and zero otherwise is feasible by (20) and of cost at most twice the optimal fractional of (RR) by (21).
4.4 When all bounded pitch and knapsack cover inequalities are added

Consider the following MinKnap instance with $\epsilon_n = \frac{1}{\sqrt{n}}$:

$$\min \sum_{i \in [n]} x_i + \frac{1}{\sqrt{n}} \sum_{j \in [n]} z_j$$

$$\text{s.t. } \sum_{i \in [n]} x_i + \frac{1}{n} \sum_{j \in [n]} z_j \geq 1 + \epsilon_n$$

(22)

Furthermore, for $k$ be a valid KC or pitch-componentwise the point generated at step $k - 1$, and the latter by induction hypothesis satisfies all inequalities of pitch at most $k - 1$. Let

$$\sum_{i \in I} w_i x_i + \sum_{j \in J} w_j z_j \geq \beta$$

(23)

be a valid KC or pitch-k inequality with support $I \cup J$, which gives that $w_i, w_j \in \mathbb{R}_{>0}$, $\forall i \in I$, $\forall j \in J$ and $\beta > 0$. Observe the following.

Claim 11. $|I| \geq n - 1$. In addition, $|I| = n - 1$ or $|J| \leq n(1 - \epsilon_n)$ implies $w_i \geq \beta$, $\forall i \in I$.

Proof. Since all coefficients in (23) are strictly positive and $\beta > 0$, $|I| \leq n - 2$ gives that the feasible solution $(\chi^{[n]\setminus I}, \bar{0})$ for (22) is cut off by (23), a contradiction.

Furthermore, if $|I| = n - 1$ and $w_{i^*} < \beta$ for some $i^* \in I$, then $(\chi^{[n]\setminus I \cup \{i^*\}}, \bar{0})$ is cut off, again a contradiction. Finally, if $|I| = n$, $|J| \leq n(1 - \epsilon_n)$ and $w_{i^*} < \beta$ for some $i^* \in I$, then $(\chi^{\{i^*\}}, \chi^{[n]\setminus J})$ does not satisfy (23), but it is feasible in (22).

We first show the statement for (23) being a KC. By the definition of KC: $\beta = 1 + \epsilon_n - \frac{|I| - [n] \setminus I}{n}$, $w_i = \min\{1, \beta\}$, $\forall i \in I$ and $w_j = \min\{\frac{1}{n}, \beta\}$, $\forall j \in J$. If $I = [n]$, then

$$\sum_{i \in I} w_i x_i = \min\{1, \beta\} \cdot (1 + \epsilon_n) \geq \beta$$

since $\beta \leq 1 + \epsilon_n$. Otherwise, $|[n] \setminus I| = 1$ so $w_i = \beta$, $\forall i \in I$ and

$$\sum_{i \in I} w_i x_i = \beta(n - 1) \approx \beta$$

for sufficiently large $n$.

Conversely, let (23) be a valid pitch-k inequality. By Claim 11 if $|I| = n - 1$ or $|J| \leq n(1 - \epsilon_n)$ then $w_i \geq \beta$, $\forall i \in I$ so the proof is analogous to the one for KC. Otherwise, $|I| = n$ and $|J| > n(1 - \epsilon_n)$. Consider the LP (13) in Section 3 specialized for our case – that is, we want to detect if $(\bar{x}, \bar{z})$ can be separated via an inequality with support $I \cup J$. Since $(\chi^{\{i\}}, \bar{0})$ is feasible in (22) for $i, \bar{i} \in I$, then

$$\alpha_i + \alpha_{\bar{i}} \geq 1.$$  \hspace{1cm} (24)

Furthermore, for $n$ large enough one has $|J| > n(1 - \epsilon_n) \geq k$ so

$$\sum_{j \in K} \alpha_j \geq 1.$$  \hspace{1cm} (25)
for any $k$-subset $K$ of $J$. We claim that the minimum in (13) is attained at $\bar{\alpha}_i = 1/2$, $\forall i \in I$ and $\bar{\alpha}_j = 1/k$, $\forall j \in J$. Indeed, the objective function of (13) computed in $\bar{\alpha}$ is given by

$$|I| \cdot \frac{1}{2} \cdot \frac{1 + \epsilon}{n} + |J| \cdot \frac{1}{k} \cdot \frac{k}{n} = \frac{1 + \epsilon_n}{2} + \frac{|J|}{n}.$$ 

On the other hand, by summing (24) for all possible pairs with multipliers $\frac{1 + \epsilon_n}{2(n-1)}$ and (25) for all subsets of $J$ of size $k$ with multipliers $\left(\frac{|J|}{k} - 1\right)^{-1} \frac{k}{n}$, simple linear algebra calculations lead to

$$\sum_{i \in I} \bar{x}_i \alpha_i + \sum_{j \in J} \bar{z}_j \alpha_j \geq \frac{1 + \epsilon_n}{2} + \frac{|J|}{n},$$

showing the optimality of $\bar{\alpha}$. Recalling $|J| > n(1 - \epsilon_n)$, we conclude that

$$\sum_{i \in I} \bar{x}_i \bar{\alpha}_i + \sum_{j \in J} \bar{z}_j \bar{\alpha}_j = \frac{1 + \epsilon_n}{2} + \frac{|J|}{n} > \frac{1 + \epsilon_n}{2} + 1 - \epsilon_n > 1,$$

hence $(\bar{x}, \bar{z})$ satisfies all inequalities with support $I \cup J$. \hfill \blacksquare

**Acknowledgments.** Some of the work was done when the second and the third author visited the IEOR department of Columbia University, partially funded by a gift of the Swiss National Science Foundation.

**References**

[1] Abbas Bazzi, Samuel Fiorini, Sangxia Huang, and Ola Svensson. Small extended formulation for knapsack cover inequalities from monotone circuits. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA ’17, pages 2326–2341, Philadelphia, PA, USA, 2017. Society for Industrial and Applied Mathematics.

[2] Daniel Bienstock. Approximate formulations for 0-1 knapsack sets. *Oper. Res. Lett.*, 36(3):317–320, May 2008.

[3] Daniel Bienstock and Mark Zuckerberg. Approximate fixed-rank closures of covering problems. *Mathematical Programming*, 105(1):9–27, Jan 2006.

[4] Robert D. Carr, Lisa K. Fleischer, Vitus J. Leung, and Cynthia A. Phillips. Strengthening integrality gaps for capacitated network design and covering problems. In *Proc. of SODA ’00*, pages 106–115, 2000.

[5] Michele Conforti, Gerard Cornuejols, and Giacomo Zambelli. *Integer Programming*. Springer Publishing Company, Incorporated, 2014.

[6] Szymon Dudycz and Carsten Moldenhauer. Approximated extended formulations for the knapsack cover problem. *Technical report, EPFL*, 2015.

[7] Samuel Fiorini, Tony Huynh, and Stefan Weltge. Strengthening convex relaxations of 0/1-sets using boolean formulas. *arXiv preprint arXiv:1711.01358*, 2017.

[8] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag, New York, 1988.
[9] Oscar H. Ibarra and Chul E. Kim. Fast approximation algorithms for the knapsack and sum of subset problems. *J. ACM*, 22(4):463–468, October 1975.

[10] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer US, 1972.

[11] Diego Klabjan, George L. Nemhauser, and Craig A. Tovey. The complexity of cover inequality separation. *Oper. Res. Lett.*, 23(1-2):35–40, August 1998.

[12] Adam Kurpisz, Samuli Leppänen, and Monaldo Mastrolilli. On the hardest problem formulations for the 0/1 lasserre hierarchy. *Math. Oper. Res.*, 42(1):135–143, 2017.

[13] Eugene L. Lawler. Fast approximation algorithms for knapsack problems. *Math. Oper. Res.*, 4(4):339–356, November 1979.