Math. Program., Ser. A
DOI 10.1007/s10107-008-0226-9

FULL LENGTH PAPER

Lifting inequalities: a framework for generating strong cuts for nonlinear programs

Jean-Philippe P. Richard · Mohit Tawarmalani

Received: 23 March 2007 / Accepted: 18 April 2008

Abstract In this paper, we introduce the first generic lifting techniques for deriving strong globally valid cuts for nonlinear programs. The theory is geometric and provides insights into lifting-based cut generation procedures, yielding short proofs of earlier results in mixed-integer programming. Using convex extensions, we obtain conditions that allow for sequence-independent lifting in nonlinear settings, paving a way for efficient cut-generation procedures for nonlinear programs. This sequence-independent lifting framework also subsumes the superadditive lifting theory that has been used to generate many general-purpose, strong cuts for integer programs. We specialize our lifting results to derive facet-defining inequalities for mixed-integer bilinear knapsack sets. Finally, we demonstrate the strength of nonlinear lifting by showing that these inequalities cannot be obtained using a single round of traditional integer programming cut-generation techniques applied on a tight reformulation of the problem.

Keywords Nonlinear mixed-integer programming · Cutting planes · Bilinear knapsacks · Convex extensions · Sequence-independent lifting · Elementary closures

Mathematics Subject Classification (2000) 90C26 · 90C30 · 90C11

J.-P. P. Richard
School of Industrial Engineering, Purdue University, 315 N. Grant Street,
West Lafayette, IN 47907-2023, USA
e-mail: jprichar@purdue.edu

M. Tawarmalani (✉)
Krannert School of Management, Purdue University, 100 S. Grant Street,
West Lafayette, IN 47907-2076, USA
e-mail: mtawarma@purdue.edu

1 Introduction

The use of cutting planes in branch-and-cut algorithms has proven to be very effective in solving linear integer programs quickly [7,8,19]. There are primarily two types of approaches for generating cutting planes for unstructured programs. The first approach is based on generating tight relaxations using disjunctive arguments [4,5,32,34] by injecting the problem in a higher-dimensional space. The second approach is to generate cutting planes without the addition of variables. Lifting is a successful integer programming technique that generates such cuts. It is the process of converting an inequality valid for a subset of the feasible region, here referred to as the *seed inequality*, to be globally valid [27,42]. Typically, the seed inequality is derived under the assumption that some of the variables are fixed at certain values. Then, the inequality is made globally valid by sequentially relaxing these restrictions. The case where the derived inequality is independent of the sequence in which restrictions are relaxed [16,43] is of particular interest since it has been shown to yield many families of effective and strong cuts for integer programs that are also computationally efficient to generate. Certain cuts such as the Gomory mixed-integer cut can be obtained via both disjunctive arguments [26] and via lifting arguments [25].

Although disjunctive programming techniques have found generalizations to nonlinear programming [9,35–38] and inequalities for special-purpose global optimization have been obtained recently via lifting techniques (see [13] for linear programs with complementarity and [20,40] for nonconvex quadratic programs), a general approach to lifting or sequence-independent lifting in nonlinear optimization has not been derived. One possible reason is that it can no longer be assumed that the seed inequality, even if linear, will remain so after lifting. Another possible explanation is that lifting techniques in mixed-integer programming have, for a significant time, been limited to integer variables. Only recently has the process been generalized to fixing and then relaxing continuous linear variables [23,30,31].

In this paper, we study lifting as the process that extends affine minorants of a function restricted to an affine subspace to yield affine minorants of the unrestricted function. This approach generalizes the traditional definition of lifting from integer programming to nonlinear programming through the use of the indicator function for the underlying set. In contrast to defining the lifting function in terms of perturbations to the right-hand-sides of the constraints, our lifting function is defined in the space of the restricted variables. The new definition is necessary since the traditional lifting function does not lend itself to a straightforward extension to nonlinear programs. Furthermore, for integer programs, the two definitions are related to each other via an affine transformation. One advantage of our derivation is that it admits an appealing geometric interpretation, which reveals short proofs and generalizations of earlier lifting results for continuous variables in the mixed-integer knapsack sets studied in [30,31]. A second advantage is that it provides a general definition for sequence-independence that is motivated by the geometry of the lifting function, and yields as a special case the superadditive theory of lifting for integer programs [43]. The main purpose of this paper is, however, to illustrate the applicability of the proposed methods in nonlinear programming. Towards this end, we study mixed-integer bilinear knapsack sets. For these sets, we derive facet-defining inequalities using sequence-dependent and sequence-independent lifting. The geometric interpretation plays a significant role in this derivation. We then investigate the strength of the resulting inequalities. First, we define an aggregation-tightening procedure which subsumes typical integer programming cut generation procedures. Then, we introduce a non-inclusion certificate, whose existence proves that an inequality cannot be obtained using one round of the aggregation-tightening procedure. For the lifted facet-defining inequalities mentioned earlier, we construct such a certificate and thus show that our inequalities are not easy to derive using traditional integer programming techniques on an equivalent linear formulation of the nonlinear set. We also prove that the lifted inequalities are not rank-one split cuts. This serves to illustrate the value of nonlinear lifting even when the underlying sets can be effectively linearized.

In Sect. 2 we describe how the lifting theory for integer programs can be extended to nonlinear programming. In Sect. 3, we apply the lifting theory developed in Sect. 2 to mixed-integer nonlinear knapsack sets and, in particular, mixed-integer bilinear knapsack sets. Our nonlinear knapsack results generalize superadditivity conditions known for 0–1 knapsack and fixed-charge single-node flow models. For the mixed-integer bilinear set, we use the proposed lifting theory to generate, in closed-form, a large family of facet-defining lifted cover inequalities. In Sect. 4, we introduce a cut-generation framework that subsumes many commonly used techniques for finding valid inequalities for integer programs. Then, using this framework, we show that our lifted cover inequalities are strictly inside the elementary closures of many integer programming relaxation techniques. Finally, in Sect. 5, we give concluding remarks and directions for future research.

2 A convex analysis perspective on lifting

Consider a function $\gamma(x)$. We denote the convex envelope of $\gamma(x)$ by conv $(\gamma(x))$, the lower envelope of the closure of the epigraph of $\gamma(x)$ by cl $(\gamma(x))$, the largest positively homogeneous convex underestimator of $\gamma(x)$ by $\hat{\gamma}(x)$, the effective domain of γ by $\text{dom}(\gamma) = \{x \mid \gamma(x) \neq \infty\}$, the subdifferential of $\gamma(x)$ at $\bar{x}$ by $\partial\gamma(\bar{x})$, and the projection of a set $S(x, y)$ in the space of y by $\text{proj}_y(S(x, y))$. The conjugate of a function $\gamma(x)$ is defined as $\gamma^*(s) = \sup_x \{\langle s, x\rangle - \gamma(x)\}$; see Rockafellar [32]. For a function $f(x, y)\colon \mathbb{R}^{n+m} \mapsto \mathbb{R}$, the restricted function obtained when x is fixed at $\bar{x}$ will be denoted by $f(\bar{x}, \cdot)$. Consistent with the above notation, $f(\bar{x}, \cdot)^*$ is the conjugate of the restricted function and (cl conv $f(\bar{x}, \cdot))(y)$ is the convex envelope of the restricted function evaluated at y. The indicator function of a set S is defined to be 0 over S and ∞ otherwise.

The function of interest in this discussion is $f(x, y)\colon \mathbb{R}^{p+n} \to \mathbb{R}$. We assume throughout that f has an affine minorant and $f \not\equiv \infty$. The process of lifting will take an inequality valid for $f(x, y)$ when $y = 0$ and lift it over all y. Observe that fixing y to 0 is without loss of generality. Assume the inequality we lift is $f(x, 0) \geq \langle\bar{\alpha}, x\rangle - \delta$, where we choose δ as small as possible. Then, $\delta = \sup_x \{\langle\bar{\alpha}, x\rangle - f(x, 0)\} = f(\cdot, 0)^*(\bar{\alpha})$. Therefore, the inequality we lift is $f(x, 0) \geq \langle\bar{\alpha}, x\rangle - f(\cdot, 0)^*(\bar{\alpha})$.

Theorem 1 *Assume that* $f(x, y): \mathbb{R}^{p+n} \rightarrow \mathbb{R}$ *has an affine minorant and that* $\mathrm{dom}(f) \neq \emptyset$. *Define* $g(\alpha, y) = -\sup_x \{\langle \alpha, x\rangle - f(x, y)\}$. *Then,*

$$\mathrm{cl\,conv}\, f(x, y) = \sup_{\alpha} \{\langle \alpha, x\rangle + \mathrm{cl\,conv}\,(g(\alpha, \cdot))\,(y)\}. \tag{1}$$

Proof By definition,

$$\begin{aligned} f^*(\alpha, \nu) &= \sup_{x,y} \{\langle (\alpha, \nu), (x, y)\rangle - f(x, y)\} \\ &= \sup_{y} \left\{ \langle \nu, y\rangle + \sup_{x} \{\langle \alpha, x\rangle - f(x, y)\} \right\} \\ &= \sup_{y} \{\langle \nu, y\rangle - g(\alpha, y)\} \\ &= g(\alpha, \cdot)^*(\nu). \end{aligned} \tag{2}$$

Further,

$$\begin{aligned} f^{**}(x, y) &= \sup_{\alpha,\nu} \big\{\langle (\alpha, \nu), (x, y)\rangle - f^*(\alpha, \nu)\big\} \\ &= \sup_{\alpha} \left\{ \langle \alpha, x\rangle + \sup_{\upsilon} \big\{\langle \nu, y\rangle - f^*(\alpha, \nu)\big\} \right\} \\ &= \sup_{\alpha} \big\{\langle \alpha, x\rangle + g(\alpha, \cdot)^{**}(y)\big\}. \end{aligned} \tag{3}$$

Note that $f^{**}(x, y) = \mathrm{cl\,conv}(f(x, y))$ (whenever f has an affine minorant and $f \neq \infty$) (see Theorem 1.3.5 in [18]). Clearly, $f(x, y) \geq \langle \alpha, x\rangle + g(\alpha, y)$ for all α. Since

$$\begin{aligned} f^{**}(x, y) = \mathrm{cl\,conv}\, f(x, y) &\geq \sup_{\alpha} \{\langle \alpha, x\rangle + \mathrm{cl\,conv}\,(g(\alpha, \cdot))\,(y)\} \\ &\geq \sup_{\alpha} \big\{\langle \alpha, x\rangle + g(\alpha, \cdot)^{**}(y)\big\}, \end{aligned}$$

the equality holds throughout. □

In other words, to form the convex envelope of $f(x, y)$, we choose linear functionals in the x-space and form the conjugate considering the y variables to be fixed. Then, we treat the conjugate as a function of the y variables and convexify it. Intuitively, Theorem 1 relates the process of lifting to the extension form of Hahn–Banach Theorem; see Theorem 1 in Sect. 5.12 of [22].

Observe that $f(x, y) \geq \langle \alpha, x\rangle + g(\alpha, y) \geq \langle \alpha, x\rangle + \langle \nu, y\rangle$ as long as $\langle \nu, y\rangle$ underestimates $g(\alpha, y)$. Since $\langle \nu, y\rangle$ is linear, it also underestimates $\mathrm{cl\,conv}\,(g(\alpha, \cdot))\,(y)$. To simplify notation, we denote $\mathrm{cl\,conv}\,(g(\alpha, \cdot))\,(y)$ by $\bar{g}_\alpha(y)$ henceforth.

Proposition 2 *Assume that* $\bar{g}_\alpha(0) = 0$ *and* $\partial \bar{g}_\alpha(0) \neq \emptyset$. *Then, the largest closed, positively homogeneous, convex function that underestimates* $\bar{g}_\alpha(y)$ *is*

$$h_\alpha(y) = \sup\{\langle \nu, y\rangle \mid \upsilon \in \partial \bar{g}_\alpha(0)\}.$$

Proof By definition, $\bar{\nu} \in \partial \bar{g}_\alpha(0)$ if and only if $\bar{g}_\alpha(y) \geq \langle \bar{\nu}, y \rangle$. Clearly, $h_\alpha(y)$, being a support function, is a closed, positively homogeneous convex function. Now consider any other such function $k_\alpha(y)$ that also underestimates $\bar{g}_\alpha(y)$ and is the largest. Since $\partial \bar{g}_\alpha(0) \neq \emptyset$, $k_\alpha(0) \geq 0$. But $k_\alpha(0) \leq \bar{g}_\alpha(0)$. Therefore, $k_\alpha(0) = 0$. In other words, $k_\alpha(y)$ is proper and can be expressed as a supremum of linear functions. Any linear function $\langle \nu, y \rangle$ that underestimates $k_\alpha(y)$ also underestimates $\bar{g}_\alpha(y)$ and therefore $\nu \in \partial \bar{g}_\alpha(0)$. In other words, $h_\alpha(y) \geq \langle \nu, y \rangle$. Therefore, $h_\alpha(y) \geq k_\alpha(y)$. □

We are interested in lifting inequalities of the form

$$f(x, 0) \geq \langle \bar{\alpha}, x \rangle + g(\bar{\alpha}, 0), \tag{4}$$

into inequalities of the form

$$f(x, y) \geq \langle \bar{\alpha}, x \rangle + \bar{g}_{\bar{\alpha}}(0) + \langle \nu, y \rangle, \tag{5}$$

where $\nu \in \partial \bar{g}_{\bar{\alpha}}(0)$ (a convex set). Inequality (5) is valid since it follows from Theorem 1 that $f(x, y) \geq \langle \bar{\alpha}, x \rangle + \bar{g}_{\bar{\alpha}}(y)$ and because $\nu \in \partial \bar{g}_{\bar{\alpha}}(0)$ implies that $\bar{g}_{\bar{\alpha}}(y) \geq \bar{g}_{\bar{\alpha}}(0) + \langle \nu, y \rangle$. In fact, one can obtain the nonlinear inequality

$$f(x, y) \geq \langle \bar{\alpha}, x \rangle + \bar{g}_{\bar{\alpha}}(0) + h_{\bar{\alpha}}(y), \tag{6}$$

where $h_{\bar{\alpha}}(y)$ is the largest closed positively homogeneous convex underestimator of $\bar{g}_{\bar{\alpha}}(y) - \bar{g}_{\bar{\alpha}}(0)$ obtained in Proposition 2. Since $\bar{g}_{\bar{\alpha}}(y)$ is the largest closed convex underestimator, $h_{\bar{\alpha}}(y)$ is also the largest closed positively homogeneous convex underestimator of $g(\bar{\alpha}, y) - \bar{g}_{\bar{\alpha}}(0)$. Observe that in the above development, we assumed that $\partial \bar{g}_{\bar{\alpha}}(0)$ is not empty, otherwise even a linear lifted inequality cannot be obtained. We are particularly interested in the cases when $\bar{g}_\alpha(0)$ equals $g(\alpha, 0)$ for all α. This will happen, for example, if 0 is an extreme point of $\text{proj}_y \text{dom}(f)$, $\text{proj}_x \text{dom}(f)$ is bounded, f is lower-semicontinuous (lsc) and has an affine minorant. In this case, for each $(\alpha, y) \in \mathbb{R}^{p+n}$ and $\delta \in \mathbb{R}$, there is a neighborhood of (α, y) such that $\{(\alpha, y, x) \mid \langle \alpha, x \rangle - f(x, y) \leq -\delta\}$ is compact and therefore its projection on the (α, y)-space is compact. In other words, $g(\alpha, y)$ is lsc. Since 0 is an extreme point of $\text{proj}_y \text{dom}(f)$, we know that $g(\alpha, 0) = \text{conv}\,(g(\alpha, \cdot))\,(0)$. Since $g(\alpha, \cdot)$ is closed and has an affine minorant (because f has such a minorant and $\text{proj}_x \text{dom}(f)$ is bounded), it follows that $\bar{g}_\alpha(0) = g(\alpha, 0)$.

Proposition 3 *Let* $f(x, y) \colon \mathbb{R}^{p+n} \to \mathbb{R}$, $g(\bar{\alpha}, y) = -\sup_x \{\langle \bar{\alpha}, x \rangle - f(x, y)\}$ *and* $\bar{g}_{\bar{\alpha}}(y) = \text{cl}\,\text{conv}\,(g(\bar{\alpha}, \cdot))\,(y)$. *If* $f(x, y) \geq \langle \bar{\alpha}, x \rangle + \langle \nu, y \rangle - \delta$ *where* δ *is such that* $f(x, 0) \geq \langle \bar{\alpha}, x \rangle - \delta + \epsilon$ *is not valid for any* $\epsilon > 0$, *then* $\nu \in \partial \bar{g}_{\bar{\alpha}}(0)$ *and* $\bar{g}_{\bar{\alpha}}(0) = g(\bar{\alpha}, 0)$.

Proof Clearly, $g(\bar{\alpha}, y) = \inf_x \{f(x, y) - \langle \bar{\alpha}, x \rangle\} \geq \langle \nu, y \rangle - \delta = \langle \nu, y \rangle + g(\bar{\alpha}, 0)$, where the last equality follows since $\delta = -\inf_x \{f(x, 0) - \langle \bar{\alpha}, x \rangle\} = -g(\bar{\alpha}, 0)$. Since, $\langle \nu, y \rangle + g(\bar{\alpha}, 0)$ is an affine function of y, $\bar{g}_{\bar{\alpha}}(y) \geq \langle \nu, y \rangle + g(\bar{\alpha}, 0)$. Setting $y = 0$, we obtain $\bar{g}_{\bar{\alpha}}(0) \geq g(\bar{\alpha}, 0)$. But $\bar{g}_{\bar{\alpha}}(\cdot)$ underestimates $g(\bar{\alpha}, \cdot)$. Therefore, $\bar{g}_{\bar{\alpha}}(0) = g(\bar{\alpha}, 0)$ and $\bar{g}_{\bar{\alpha}}(y) \geq \langle \nu, y \rangle + \bar{g}_{\bar{\alpha}}(0)$, which proves that $\nu \in \partial \bar{g}_{\bar{\alpha}}(0)$. □

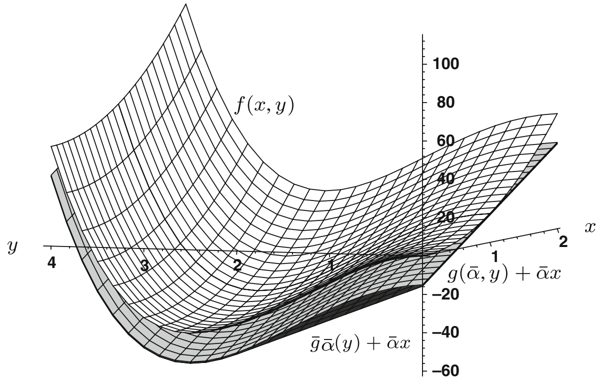

Fig. 1 Illustrating Theorem 1 and Proposition 2 on Example 4

In the above discussion, we have shown that the process of lifting inequalities is equivalent to finding a vector v in the subdifferential of $\bar{g}_{\bar{\alpha}}(y)$. On the other hand, the tightest underestimator is the one given in (6). This underestimator is constructed by identifying a vector $\bar{\alpha}$ such that $\bar{g}_{\bar{\alpha}}(0)$ equals $g(\bar{\alpha}, 0)$ and then constructing the largest positively homogeneous, closed convex function underestimating $\bar{g}_{\bar{\alpha}}(y) - g(\bar{\alpha}, 0)$. Henceforth, we further restrict our attention to the case where $\text{proj}_y\, \text{dom}(f) \subseteq \mathbb{R}^n_+$. If $0 \in \text{proj}_y\, \text{dom}(f)$, then 0 must be an extreme point of $\text{proj}_y\, \text{dom}(f)$.

Example 4 Let $f(x, y) = y^4 - 12y^2 - 3y + 15x^2$, where $x \in \mathbb{R}$, $y \in \mathbb{R}_+$. Let $\bar{\alpha} = 30$, and assume that the inequality we are lifting is $f(x, 0) \geq 30x - 15$. Then, $g(\bar{\alpha}, y) = y^4 - 12y^2 - 3y - 15$,

$$\bar{g}_{\bar{\alpha}}(y) = \begin{cases} -19y - 15 & \text{if y} < 2 \\ y^4 - 12y^2 - 3y - 15 & \text{otherwise,} \end{cases}$$

and $h_{\bar{\alpha}}(y) = -19y$. Observe that $h_{\bar{\alpha}}(y)$ is the largest positively homogeneous underestimator of $\bar{g}_{\bar{\alpha}}(y) + 15$. Here, as required in Proposition 2, we shift $\bar{g}_{\alpha}(y)$ by 15 so that it passes through the origin. Inequality (6) reduces to $f(x, y) \geq 30x - 19y - 15$, which can also be inferred from Fig. 1. □

The notion of sequence-independence has been central to the successful development of lifted inequalities for integer programs. We approach the notion of sequence-independence from a geometric perspective that is more general than the one used in the integer programming literature, but yields the more traditional definitions when the perturbation function is superadditive.

The lifted inequality (6) uses the largest closed positively homogeneous convex underestimator of $\bar{g}_{\bar{\alpha}}(y) - \bar{g}_{\bar{\alpha}}(0)$. Since y may have many components, we are interested in investigating when this underestimator can be obtained by considering subsets of the y-variables independently. Towards this end, consider the partitioning $y = (y_1, y_2)$ of the y-variables and define $\gamma(y_1, y_2) = g(\bar{\alpha}, y) - \bar{g}_{\bar{\alpha}}(0)$. To introduce the definition of sequence-independence, we assume that $\gamma(y_1, y_2)$ has a linear minorant and that

$\gamma(0, 0) = 0$. We define $\gamma_r(y_1, y_2)$ as the restriction of γ to the subspaces $(y_1, 0)$ and $(0, y_2)$, i.e.,

$$\gamma_r(y_1, y_2) = \begin{cases} \gamma(y_1, y_2) & \text{if } y_1 = 0 \text{ or } y_2 = 0 \\ \infty & \text{otherwise.} \end{cases}$$

Recall that $\hat{\gamma}(y_1, y_2)$ denotes the largest positive homogeneous convex underestimator of $\gamma(y_1, y_2)$.

Definition 5 Lifting of $\gamma(y_1, y_2)$ from $(0, 0)$ is said to be sequence-independent of y_1 and y_2 if $\hat{\gamma}(y_1, y_2) = \hat{\gamma}_r(y_1, y_2)$.

The definition above is inspired by convex extensions as defined in Tawarmalani and Sahinidis [38] where the convex envelope of a function is constructed by restricting attention to a subset of the domain of the function.

Theorem 6 *The lifting of* $\gamma(y_1, y_2)$ *from* $(0, 0)$ *is sequence-independent of* y_1 *and* y_2 *if and only if* $\hat{\gamma}(y_1, y_2) = \hat{\gamma}(y_1, 0) + \hat{\gamma}(0, y_2)$.

Proof Clearly, $\hat{\gamma}_r(y_1, 0)$ equals $\hat{\gamma}(y_1, 0)$ and is the largest positively homogeneous convex underestimator of $\gamma(y_1, 0)$. This is because $(y_1, 0)$ can only be expressed as convex combination of points of the same form. By symmetry, $\hat{\gamma}_r(0, y_2)$ is the largest positively homogeneous convex underestimator of $\gamma(0, y_2)$. Then, by the definition of $\hat{\gamma}_r$,

$$\hat{\gamma}_r(y_1, y_2) = \inf_{\lambda} \left\{ \lambda \hat{\gamma}_r(y_a, 0) + (1 - \lambda)\hat{\gamma}_r(0, y_b) \mid y_1 = \lambda y_a, y_2 = (1 - \lambda) y_b \right\}.$$

However, it follows from the positive homogeneity of $\hat{\gamma}_r$ and $\hat{\gamma}_r(0, 0) = 0$ that $\hat{\gamma}_r(y_1, y_2) = \hat{\gamma}_r(y_1, 0) + \hat{\gamma}_r(0, y_2)$. In other words, the lifting of $\gamma(y_1, y_2)$ from $(0, 0)$ is sequence-independent of y_1 and y_2 if

$$\hat{\gamma}(y_1, y_2) = \hat{\gamma}(y_1, 0) + \hat{\gamma}(0, y_2). \tag{7}$$

On the other hand, if (7) holds, then $\hat{\gamma}_r(y_1, y_2) \geq \hat{\gamma}(y_1, y_2) = \hat{\gamma}(y_1, 0) + \hat{\gamma}(0, y_2) = \hat{\gamma}_r(y_1, 0) + \hat{\gamma}_r(0, y_2) = \hat{\gamma}_r(y_1, y_2)$, where the first inequality follows from the fact that γ_r is a restriction of γ. Therefore, the equality holds throughout. □

Because the recession directions of the epigraph of cl $\hat{\gamma}_r(y_1, 0)$ are not in the opposite direction of those of the epigraph of cl $\hat{\gamma}_r(0, y_2)$ (since $\hat{\gamma}_r$ has a linear minorant), it follows that cl $\hat{\gamma}_r(y_1, 0) +$ cl $\hat{\gamma}_r(0, y_2)$ is closed. Therefore, if the lifting of $\gamma(y_1, y_2)$ is sequence-independent, then

$$\text{cl}\, \hat{\gamma}(y_1, y_2) = \text{cl}\, \hat{\gamma}(y_1, 0) + \text{cl}\, \hat{\gamma}(0, y_2). \tag{8}$$

It follows from Proposition 3 and the preceding discussion that the lifted inequality (6) is formed by constructing the largest closed positively convex homogeneous underestimator of $\bar{g}_{\bar{\alpha}}(y) - \bar{g}_{\bar{\alpha}}(0)$. Let $y = (y_1, y_2)$ and recall that $\gamma(y_1, y_2) = g(\bar{\alpha}, y) - \bar{g}_{\bar{\alpha}}(0)$.

If lifting of γ from $(0, 0)$ is sequence-independent of y_1 and y_2, then, using (8), we can rewrite (6) as

$$f(x, y) \geq \langle \bar{\alpha}, x \rangle + \bar{g}_{\bar{\alpha}}(0) + \text{cl}\,\hat{\gamma}(y_1, 0) + \text{cl}\,\hat{\gamma}(0, y_2). \tag{9}$$

The form of (9) justifies the sequence-independence terminology since each part of the lifting can be obtained independently from the other.

The sequence-independence condition in Theorem 6 may be difficult to verify directly. We now show that a relaxation of the notion of superadditivity already yields sequence-independence. As is standard, a function $\gamma(y_1, y_2)$ is superadditive if $\gamma(y_1 + y_1', y_2 + y_2') \geq \gamma(y_1, y_2) + \gamma(y_1', y_2')$.

Corollary 7 *If $\gamma(y_1, y_2) \geq \hat{\gamma}(y_1, 0) + \hat{\gamma}(0, y_2)$, then the lifting of $\gamma(y_1, y_2)$ from $(0, 0)$ is sequence-independent of y_1 and y_2. More generally, if $\gamma(y_1, \dots, y_m) \geq \sum_{i=1}^{m} \hat{\gamma}(0, \dots, 0, y_i, 0, \dots, 0)$, then the lifting of $\gamma(y_1, \dots, y_m)$ is sequence-independent of $y_1, \dots, y_m$, i.e., $\hat{\gamma}(y_1, \dots, y_m) = \sum_{i=1}^{m} \hat{\gamma}(0, \dots, 0, y_i, 0, \dots, 0)$.*

Proof Clearly, any convex positively homogeneous underestimator $h(y_1, y_2)$ of $\gamma(y_1, y_2)$ must satisfy $h(y_1, y_2) \leq h(y_1, 0) + h(0, y_2) \leq \hat{\gamma}(y_1, 0) + \hat{\gamma}(0, y_2)$, because h is convex and $\hat{\gamma}$ is the largest convex positively homogeneous underestimator. Since $\hat{\gamma}(y_1, 0) + \hat{\gamma}(0, y_2)$ is a convex positively homogeneous function that underestimates γ (by assumption), it must be the largest such function. In other words, $\hat{\gamma}(y_1, y_2) = \hat{\gamma}(y_1, 0) + \hat{\gamma}(0, y_2)$, which by Theorem 6, implies sequence-independence. The general result follows via induction on m. □

If $\gamma(y_1, y_2) \geq \gamma(y_1, 0) + \gamma(0, y_2)$, the condition in the Corollary 7 is certainly satisfied since $\hat{\gamma}(y_1, 0)$ (respectively, $\hat{\gamma}(0, y_2)$) underestimates $\gamma(y_1, 0)$ (respectively, $\gamma(0, y_2)$). Note that in Corollary 7, we do not require superadditivity of γ component-wise. This is therefore different, and less restrictive than the way sequence-independence is typically defined in integer programming via the superadditivity of the perturbation function; see Wolsey [43].

Example 8 Consider $\gamma(y_1, y_2) = y_1^2 + 16\sqrt{y_1} + y_2^2 + 31.25\sqrt{y_2} + y_1 y_2$, where $(y_1, y_2) \in \mathbb{R}_+^2$. It can be easily verified that $\gamma(y_1, y_2) \geq \gamma(y_1, 0) + \gamma(0, y_2)$ and $\gamma(0, 0) = 0$. Then, $\gamma(y_1, y_2) \geq \hat{\gamma}(y_1, y_2) = 12y_1 + 18.75y_2 = \hat{\gamma}(y_1, 0) + \hat{\gamma}(0, y_2)$, where $\hat{\gamma}(y_1, y_2)$ is the highest positively homogeneous underestimator of $\gamma(y_1, y_2)$ as shown in Fig. 2, which depicts two views of $\gamma(y_1, y_2)$ and $\hat{\gamma}(y_1, y_2)$. □

It follows from Theorem 23.8 in [32] that the subdifferential of $\hat{\gamma}(y_1, \dots, y_m)$ is the sum of the subdifferentials of $\hat{\gamma}(0, \dots, 0, y_i, 0, \dots, 0)$. In other words, if $\hat{\gamma}(0, \dots, 0, y_i, 0, \dots, 0)$ is expressible as a supremum of t_i inequalities, then $\prod_{i=1}^{m} t_i$ underestimating inequalities for $f(x, y)$ are obtained. This phenomenon will be demonstrated later in Theorem 30 where an exponential family of facets for the 0–1 mixed-integer bilinear knapsack set is derived from lifting. Next, we describe how the lifting results presented here relate to lifting techniques in mixed-integer programming.

Example 9 Consider a pure integer programming problem whose feasible set is $S = \{Ax + By \leq d, x \in \mathbb{Z}_+^p, y \in \mathbb{Z}_+^n\}$. Let $\alpha x \leq \delta$ be a valid and tight inequality

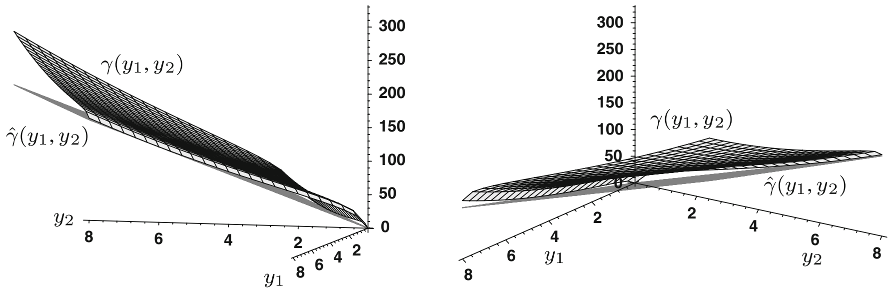

Fig. 2 Illustration of Corollary 7 using Example 8

for $\{Ax \leq d, x \in \mathbb{Z}^p_+\}$. We define $f(x, y)$ to be the indicator function of S. Let $P(w) = -\max\{\alpha x - \delta \mid Ax \leq d - w, x \in \mathbb{Z}^p_+\}$ define the perturbation problem in x-space. If y_is are re-introduced into the inequality, then by the definition of $P(w)$, $\alpha x - \delta \leq -P\left(\sum_{i=1}^n B_i y_i\right)$ where B_i is the ith column of B. If $P(\cdot)$ is superadditive, then $\alpha x - \delta \leq -\sum_{i=1}^n P(B_i)y_i$. Comparing the definition of $g(\alpha, y)$ in Theorem 1 and that of $P(w)$ above, it follows that $g(\alpha, y) = P(By) - \delta$. Further, since $P(0) = 0$, $g(\alpha, 0) = -\delta$. Then, as follows, the superadditivity of $P(\cdot)$ implies the superadditivity of $\gamma(y) = g(\alpha, y) - g(\alpha, 0)$:

$$\begin{aligned} g(\alpha, y_1 + y_2) + \delta &= P(B(y_1 + y_2)) = P(By_1 + By_2) \geq P(By_1) + P(By_2) \\ &= (g(\alpha, y_1) + \delta) + (g(\alpha, y_2) + \delta), \end{aligned}$$

which in turn implies sequence-independence (see Corollary 7). □

Example 10 Consider a single constraint mixed-integer knapsack set $S = \{ax + by \leq d, x \in \mathbb{Z}^p_+, y \in \mathbb{R}^n_+\}$. Let $\alpha x \leq \delta$ be a valid and tight inequality for $\{ax \leq d, x \in \mathbb{Z}^p_+\}$. Define the perturbation problem as $P(w) = -\max\{\alpha x - \delta \mid ax \leq d - w, x \in \mathbb{Z}^p_+\}$. Let $p = \inf_{\lambda < 0} \lambda P\left(\frac{1}{\lambda}\right)$. Clearly, pw underestimates $P(w)$. Further, if $b_i < 0$ and λ^k is a sequence such that $\lambda^k P\left(\frac{1}{\lambda^k}\right) \to p$, then we can show that $\hat{\gamma}(0, \ldots, 0, y_i, 0, \ldots, 0) = pb_i y_i$ by choosing $y_i^k = \frac{1}{b_i \lambda^k}$. Sequence-independence of the lifting coefficients then follows as:

$$\gamma(y) = P(by) = P\left(\sum_{i=1}^n b_i y_i\right) \geq p \sum_{i=1}^n b_i y_i = \sum_{i=1}^n \hat{\gamma}(0, \ldots, 0, y_i, 0, \ldots, 0).$$

We conclude that the lifting coefficients are pb_i, a result obtained in a different manner in [23]. □

Example 11 Here, we consider the mixed-integer knapsack from Example 10, with the added condition that each y_i is bounded from above. We study the interesting case with $b < 0$. This was the set studied in Richard et al. [30,31]. Without loss of generality, one may assume that $y_i \leq 1$ by scaling b_i. We show that lifted inequalities for the mixed-integer knapsack are in one-to-one correspondence with the inequalities obtained from piecewise linear concave underestimators $G(w)$ of $P(w)$ over

$\left[\sum_{i=1}^{n} b_i, 0\right)$ that are such that $G(0) = 0$ and the breakpoints are limited to occur at $\sum_{i=1}^{j} b_i$ for $j \in \{1, \ldots, n-1\}$ with appropriate reordering of variables. Let $G(w)$ be one such underestimator. We first show that a valid inequality for S can be constructed using $G(w)$. Here, $\gamma(y) = -\max_x \left\{\alpha x - \delta \mid ax \leq d - by, x \in \mathbb{Z}_+^p\right\}$. Since $\gamma(y)$ is constant whenever by is constant, it follows that the convex envelope of $\gamma(y)$ can be constructed by limiting attention to points where all but one of the y variables are at their bounds, i.e., 0 or 1 (see Corollary 5 in [38]). Let p_j be the slope of $G(w)$ between $\sum_{i=1}^{j+1} b_i$ and $\sum_{i=1}^{j} b_i$. It can be easily argued that $\sum_{j=1}^{n} b_j p_j y_j$ underestimates $\gamma(y)$ along each of the edges of the hypercube $[0, 1]^n$. This is because any function $G'(w)$ obtained by interchanging segments of $G(w)$ only underestimates $G(w)$, which in turn underestimates $P(w)$. Another way to see this by rewriting $G(w)$ as the value function of an LP as follows:

$$P(w) \geq G(w) = \max\left\{\sum_{i=1}^{n} b_i p_i y_i \;\middle|\; \sum_{i=1}^{n} b_i y_i = w,\ \forall i, 0 \leq y_i \leq 1\right\} \geq \sum_{i=1}^{n} b_i p_i y_i.$$

Now, we discuss the converse. Any inequality of the form $\alpha x + ty \leq \delta$ that is valid for S must be such that $ty \leq \gamma(y) = P(by)$ (see Proposition 3). Then, define $G(w) = \max\{ty \mid by = w, 0 \leq y \leq 1\}$. Clearly, $G(w)$ is piecewise-linear concave because it is the value function of a linear program. Further, $G(0) = 0$ and $G(w)$ underestimates $P(w)$. It remains to show that the breakpoints of $G(w)$ correspond to $\sum_{i=1}^{j} b_{\pi(i)}$ for some reordering $\pi(i)$ of variables. We define $\pi(i)$ as an order for which the ratios t_i/b_i do not increase. Then, the breakpoints of $G(w)$ can be verified to occur at $\sum_{i=1}^{j} b_{\pi(i)}$. The computational effort in obtaining such inequalities can be reduced by exploiting the following fact. Let w' be the smallest minimizer in $\operatorname{argmin}\left\{P(w)/w \;\middle|\; w \in \left[\sum_{i=1}^{n} b_i, 0\right)\right\}$, assuming that the minimum is in fact attained. Denote the corresponding optimal value by p. Then, since $G(w)$ is lower-semicontinuous and concave, $G(0) = P(0) = 0$, and $G(w') \leq P(w') = pw'$, it follows that the subgradient of G at w' is no less than p. Therefore, to construct $G(w)$ it suffices to limit our attention to $[w', 0]$. The above discussion provides a short proof of Theorems 24 and 29 in [30] and generalizes Theorems 7 and 24 in [31]. □

The argument of Example 11 remains valid even when we consider nonlinear sets of the form $\left\{(x, y) \mid \phi(x, by) \leq 0, y \in \mathbb{R}_+^n\right\}$ by defining the perturbation function as $P(w) = -\sup\{\alpha x - \delta \mid \phi(x, w) \leq 0\}$. In this case, concave underestimators of $P(w)$ with breakpoints at $\sum_{i=1}^{j} b_j$ are in one-to-one correspondence with the extended underestimating inequalities of the form $\alpha x + ty \leq \delta$.

Lifting in integer programming has an interesting property. If the seed inequality is facet-defining for the restriction of the problem and if each time the restriction on a variable is lifted in such a way that a point outside of the restriction satisfies the lifted inequality at equality, then the resulting inequality is facet-defining. In the context of nonlinear programs, we do not expect such a property to hold. In fact, no facet-defining inequality may exist (we may not be able to start with a facet-defining inequality for the restriction) and, even if it does, we may not be able to lift the inequality in a manner that adds the additional point as specified above. However,

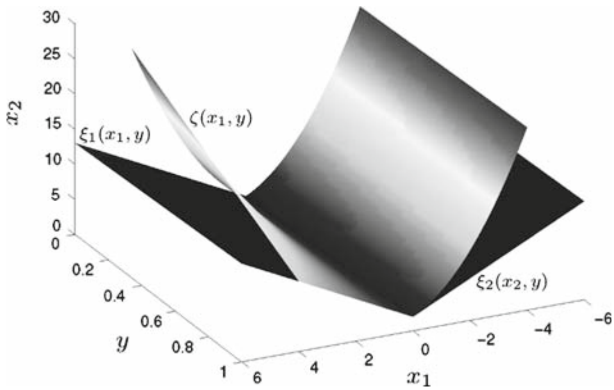

Fig. 3 Maximal faces are not lifted to maximal faces

it seems intuitive at first that such a property could hold if we used maximal proper faces instead of facets [32]. Let $W \subseteq \mathbb{R}^n \times \mathbb{R}^p_+$. Let W' be the restriction of W, $\{(x, y) \mid (x, y) \in W, y = 0\}$ and $S' = \text{conv}(W') = \text{conv}(W) \cap \{y \mid y = 0\}$. Let $\alpha x \leq 1$ correspond to a maximal face of W'. We are interested in lifting this inequality into an inequality $\alpha x + \beta y \leq 1$. Let T be a set of points in $S \backslash S'$ satisfying $\alpha x + \beta y = 1$. Further, assume that T is maximal in the sense that there does not exist T' satisfying $T \subset T' \subseteq S \backslash S'$, $\dim(T) < \dim(T')$ and a valid inequality $\alpha x + \beta' y \leq 1$ for S that is tight at every point in T'. Unfortunately, contrary to integer programming, we illustrate in Example 12 that such a lifted inequality may not define a maximal face of $\text{conv}(W)$.

Example 12 Consider the following set

$$\overline{W} = \left\{ (x_1, x_2, y) \;\middle|\; \begin{array}{l} x_2 \geq \zeta(x_1, y) = \max\left\{(1-y)x_1^2 + 2x_1 + 1, (x_1 - 1)^2\right\} + y \\ 0 \leq y \leq 1 \end{array} \right\}$$

that is represented in Fig. 3. It can be easily verified that $x_2 \geq 1 - 2x_1$ and $x_2 \geq 1 + 2x_1$ define the same maximal face of $\overline{W}_0 = \overline{W} \cap \{y \mid y = 0\}$. We lift $x_2 \geq 1 - 2x_1$ by constructing the maximal inequality $x_2 \geq 1 - 2x_1 + \beta_1 y$ such that $\beta_1 \leq \min\left\{\frac{x_2 + 2x_1 - 1}{y} \;\middle|\; (x_1, x_2, y) \in \overline{W} \backslash \overline{W}_0\right\} = 1$. Similarly, we lift $x_2 \geq 1 + 2x_1$ by constructing the maximal inequality $x_2 \geq 1 + 2x_1 + \beta_2 y$ such that $\beta_2 \leq \min\left\{\frac{x_2 - 2x_1 - 1}{y} \;\middle|\; (x_1, x_2, y) \in \overline{W} \backslash \overline{W}_0\right\} = 1$. The inequality $x_2 \geq \xi_2(x_1, y) = 1 - 2x_1 + y$ is satisfied at equality when $x_1 = 0$, and even though the coefficient of y is the maximum possible, the resulting inequality does not correspond to a maximal face. This is easily verified by considering the other inequality $x_2 \geq \xi_1(x_1, y) = 1 + 2x_1 + y$ that we obtained above. This inequality is satisfied at equality when $x_1 = 0$ or when $x_1 \geq 0$ and $y = 1$. The face defined by the second inequality contains the face defined by the first inequality. It can be verified that it also describes a maximal face. □

As in Examples 9–11, we are often interested in constrained optimization problems. Indicator functions of the constraint set allow us to naturally translate the lifting techniques developed here for such cases. In the remainder of this section, we prove a result that is practically useful in lifting inequalities for nonlinear programs, as will be illustrated in Sect. 3. Let $x, y_1, \ldots, y_m$ be vectors of variables, where $x \in \mathbb{R}^n$, $y_i \in \mathbb{R}_+^{n_i}$ for all i. Also, consider $\kappa(y_1, \ldots, y_m) \colon \mathbb{R}^{\sum_{i=1}^m n_i} \mapsto \mathbb{R}^k$ and $\phi(x, w) \colon \mathbb{R}^{n+k} \mapsto \mathbb{R}^q$. Define

$$S = \{(x, y_1, \ldots, y_m) \mid \phi\,(x, \kappa(y_1, \ldots, y_m)) \le 0\}.$$

Further assume that we are interested in lifting an inequality with a fixed slope α in the space of x variables. Then,

$$g(\alpha, y_1, \ldots, y_m) = -\sup_x \left\{\langle \alpha, x\rangle \;\middle|\; \phi\,(x, \kappa(y_1, \ldots, y_m)) \le 0\right\}$$

is the negative of the support function (conjugate of the indicator function) of S when $y_1, \ldots, y_m$ are fixed. Let $g(\alpha, 0, \ldots, 0) = -\delta$, or in other words, let $\alpha x \le \delta$ be the tightest inequality with slope α when $y_1, \ldots, y_m$ are each fixed at 0. Here, $\gamma(y_1, \ldots, y_m) = g(\alpha, y_1, \ldots, y_m) + \delta$. The perturbation function $P(w) \colon \mathbb{R}^k \mapsto \mathbb{R}$ is defined as

$$P(w) = \delta - \sup_x \left\{\langle \alpha, x\rangle \;\middle|\; \phi\,(x, \kappa(0, \ldots, 0) + w) \le 0\right\}.$$

Let $r(y_1, \ldots, y_m) = \kappa(y_1, \ldots, y_m) - \kappa(0, \ldots, 0)$. Then,

$$\gamma(y_1, \ldots, y_m) = P\,(r(y_1, \ldots, y_m)).$$

By definition, $\delta - \langle \alpha, x\rangle \ge \gamma(y_1, \ldots, y_m)$. If γ is superadditive, then $\delta - \langle \alpha, x\rangle \ge \sum_{i=1}^m \gamma(0, \ldots, 0, y_i, 0 \ldots, 0)$. Therefore, if $\nu_i y_i \le \gamma(0, \ldots, 0, y_i, 0, \ldots, 0)$ then it follows that $\langle \alpha, x\rangle + \sum_{i=1}^m \nu_i y_i \le \delta$ is valid for S. We generalize the above observation in the following result.

Theorem 13 *If there exist $h_1 \colon \mathbb{R}^{mk} \mapsto \mathbb{R}^k$ and $h_2 \colon \mathbb{R}^m \mapsto \mathbb{R}$ such that*

(A1) $r(y_1, \ldots, y_m) \ge h_1\,(r(y_1, 0, \ldots, 0), \ldots, r(0, \ldots, 0, y_m))$,
(A2) $P\,(h_1\,(r(y_1, 0, \ldots, 0), \ldots, r(0, \ldots, 0, y_m)))$
$\ge h_2\,(P\,(r(y_1, 0, \ldots, 0)), \ldots, P\,(r(0, \ldots, y_m)))$,
(A3) *$P(\cdot)$ is nondecreasing,*

where h_2 is convex and isotone (i.e., $h_2(a) \le h_2(b)$ whenever $a \le b$), and for all i, ${\gamma'}_i(y_i)$ is a convex underestimator for $\gamma(0, \ldots, 0, y_i, 0, \ldots, 0)$, then the set:

$$\overline{S} = \left\{(x, y_1, \ldots, y_m) \mid \delta - \alpha x \ge h_2({\gamma'}_1(y_1), \ldots, {\gamma'}_m(y_m))\right\}$$

is convex and outer-approximates S.

Proof

$$
\begin{aligned}
\gamma(y_1,\ldots,y_m) &= P\left(r(y_1\ldots,y_m)\right)\\
&\geq P\left(h_1\left(r(y_1,0,\ldots,0),\ldots,r(0,\ldots,0,y_m)\right)\right)\\
&\geq h_2\left(P\left(r(y_1,0,\ldots,0)\right),\ldots,P\left(r(0,\ldots,0,y_m)\right)\right)\\
&= h_2\left(\gamma(y_1,0,\ldots,0),\ldots,\gamma(0,\ldots,0,y_m)\right)
\end{aligned}
$$

where the first inequality follows because of Assumptions (A1) and (A3), and the second inequality follows from Assumption (A2). Since h_2 and γ'_i are convex, the following set

$$
A=\left\{(x,y_1,\ldots,y_m,z_1,\ldots,z_m)\;\middle|\;\begin{array}{l}\delta-\alpha x\geq h_2(z_1,\ldots,z_m)\\ z_i\geq\gamma'_i(y_i),\quad i=1,\ldots,m\end{array}\right\}
$$

is convex. Since $\gamma'_i(y_i)$ underestimates $\gamma(0,\ldots,0,y_i,0,\ldots,0)$, $\text{proj}_{(x,y)}(A)$ is a convex outer-approximation of S. If $(x,y_1,\ldots,y_m,z_1,\ldots,z_m)$ is feasible to A, then so is

$$
\left(x,y_1,\ldots,y_m,\gamma'_1(y_1),\ldots,\gamma'_m(y_m)\right),
$$

because h_2 is isotone. It follows that $\text{proj}_{(x,y)}(A)=\overline{S}$. □

Theorem 13 provides a recipe for outer-approximation by developing convex underestimators for γ restricted to the coordinate axes. It encompasses sequence-dependent as well as sequence-independent lifting procedures. For sequence-dependent lifting, we set m equal to 1 and note that h_1 and h_2 can be taken to be identity operators. For sequence-independent lifting, note that assumption (A3) is automatically satisfied when $\phi(x,w)$ is nondecreasing in w and also note that if r and P are superadditive over their relevant domains, then h_1 and h_2 can be chosen to be summation operators. In fact, Theorem 13 generalizes superadditive lifting in integer programming because, in this case, (A1) and (A3) follow easily from the fact that the defining constraints are linear and integrality is preserved when projecting perpendicular to coordinate axes and because (A2) is the familiar superadditivity of the perturbation function.

At the beginning of the section, we assumed that the restricted set is obtained by fixing y at 0. As might be apparent, the constructions in this section are affinely invariant. Therefore, it is straightforward to translate them to fixing y at $\bar{y}$. In that case, instead of positively homogenous underestimators, we develop conic underestimators of $g_{\bar{\alpha}}(y)$ with apex at $(\bar{y},g_{\bar{\alpha}}(\bar{y}))$. Similarly, we define $\gamma(y_1,\ldots,y_m)=\kappa(y_1,\ldots,y_m)-\kappa(\bar{y}_1,\ldots,\bar{y}_m)$ and replace the zeros in Theorem 13 by the corresponding $\bar{y}_i$s.

3 Application to nonlinear knapsack sets

In this section, we illustrate how the general lifting theory developed in Sect. 2 can be applied to specific problems and in particular to bilinear mixed-integer knapsack sets.

After presenting general lifting results in Sect. 3.1, we derive two families of strong inequalities for bilinear mixed-integer knapsack sets in Sect. 3.2. We will use these inequalities in Sect. 4 to illustrate the fact that, even when the set studied can be linearized, nonlinear lifting yields inequalities that cannot be easily obtained from traditional integer programming cutting plane techniques.

3.1 Sequence-dependent and sequence-independent lifting for nonlinear knapsack sets

In this section, we describe lifting tools for bilinear mixed-integer knapsack sets and derive conditions under which lifting is simple to perform. Because the tools we describe are common to all nonlinear knapsack sets in which binary and continuous variables are paired, we present the lifting results in this more general setting. In particular, we consider nonlinear knapsack sets of the form

$$K = \left\{ (x, y) \in \{0,1\}^n \times [0,1]^n \;\middle|\; \sum_{j=1}^{n} \rho_j(x_j, y_j) \le d \right\}$$

where $n \in \mathbb{Z}_+$, $\rho_j: \{0,1\} \times [0,1] \to \mathbb{R}$, and $d \in \mathbb{R}$. The set K is interesting for several reasons. First, it is an extension of various mixed-integer knapsack sets that have been previously studied. For example, it generalizes the linear mixed-integer knapsack set (choose $\rho_j(x_j, y_j) = a_j x_j + b_j y_j$) and the single node fixed charge flow model (choose $\rho_j(x_j, y_j) = a_j y_j$ if $y_j \le x_j$, $\rho_j(x_j, y_j) = \infty$ otherwise). The set K also generalizes many nonlinear knapsack sets. In particular, it generalizes bilinear knapsack sets (choose $\rho_j(x_j, y_j) = a_j x_j y_j$) that will be studied in more detail in Sect. 3.2.

Next we describe lifting tools to derive valid inequalities for $PK = \text{conv}(K)$. Given $T \subseteq N = \{1, \ldots, n\}$, we define

$$K(T) = \{(x, y) \in K \mid x_j = y_j = 0, \forall j \in T\}$$

and define $PK(T) = \text{conv}\,(K(T))$. The set $K(T)$ is the restriction of K obtained by fixing all the pairs of variables (x_j, y_j) for $j \in T$ to 0. Note that since ρ_j is arbitrary, we can always transform ρ_j in such a way that any extreme point of the unit-hypercube is mapped to the origin. Therefore, fixing x_j and y_j to 0 is not a restrictive assumption.

Assume now that $T = \{t, \ldots, n\}$, $t \ge 2$ and, for $i \in T$, we define $T_i = \{i + 1, \ldots, n\}$. We show in Proposition 14 how a valid inequality for PK can be derived starting from the following seed inequality

$$\sum_{j=1}^{t-1} \alpha_j x_j + \sum_{j=1}^{t-1} \beta_j y_j \le \delta \tag{10}$$

which is assumed to be valid for $K(T)$. The lifting we perform is sequential in that variable pairs (x_j, y_j) are lifted one at a time. Therefore, assuming that variables

(x_j, y_j) have already been lifted for $j = t, \ldots, i-1$, we are interested in studying the lifting of variables (x_i, y_i) in the inequality

$$\sum_{j=1}^{i-1} \alpha_j x_j + \sum_{j=1}^{i-1} \beta_j y_j \le \delta \tag{11}$$

which may be assumed to be valid for $PK(T_{i-1})$ by induction.

Inequality (11) corresponds to $\bar{\alpha}x \le \delta$ in Sect. 2 and $f(x_1, y_1, \ldots, x_i, y_i)$ is the indicator function of $K(T_i)$. As (11) is lifted from $K(T_{i-1})$ to be valid for $K(T_i)$,

$$g(\bar{\alpha}, x_i, y_i) = -\max\left\{\sum_{j=1}^{i} \alpha_j x_j + \sum_{j=1}^{i} \beta_j y_j \;\middle|\; (x, y) \in K(T_i)\right\}.$$

By definition, $\bar{\alpha}x + g(\bar{\alpha}, x_i, y_i) \le 0$. Further, by validity of (11), $-g(\bar{\alpha}, 0, 0) \le \delta$. Therefore, if $\alpha_i x_i + \beta_i y_i \le g(\bar{\alpha}, x_i, y_i) - g(\bar{\alpha}, 0, 0)$, then $\bar{\alpha}x + \alpha_i x_i + \beta_i y_i \le \delta$ is a valid inequality. The linear underestimation of $g(\bar{\alpha}, x_i, y_i) - g(\bar{\alpha}, 0, 0)$ is a special case of the cone underestimator of Proposition 2 from which the validity of the more general (5) was derived. The validity also follows from the even more general convex underestimation of $g(\bar{\alpha}, x_i, y_i)$ in Theorem 1 and the following discussion.

The main idea underlying the upcoming Proposition 14 is to translate the above inequality in terms of a perturbation function, as is typical in integer programming. Towards this end, we associate the following perturbation function with (11):

$$\begin{aligned} P^i(w) = \delta - \max &\left(\sum_{j=1}^{i} \alpha_j x_j + \sum_{j=1}^{i} \beta_j y_j\right) \\ &\text{s.t.} \sum_{j=1}^{i} \rho_j(x_j, y_j) \le d - \sum_{j=i+1}^{n} \rho_j(0, 0) - w \\ &x_j \in \{0, 1\}, y_j \in [0, 1]\, j = 1, \ldots, i. \end{aligned} \tag{12}$$

Observe that $g(\bar{\alpha}, x_i, y_i) - g(\bar{\alpha}, 0, 0) = P^{i-1}(\rho_i(x_i, y_i) - \rho_i(0, 0)) - \delta$. To provide a self-contained treatment in this section, we also present direct proofs using techniques prevalent in the integer programming literature. The reader is referred to Louveaux and Wolsey [21] for a survey that gives a unified presentation of such lifting techniques, and also considers superadditive lifting over sets of variables that are separable.

Proposition 14 *Assume that* (10) *is valid for* $PK(T)$. *Also assume that, for* $i \in T$, *there exist* α_i *and* β_i *that satisfy*

$$\alpha_i \phi_1 + \beta_i \phi_2 \le P^{i-1}(\rho_i(\phi_1, \phi_2) - \rho_i(0, 0)) \;\; \forall \phi_1 \in \{0, 1\}, \phi_2 \in [0, 1]. \tag{13}$$

Then, for $i \in T$, the inequality

$$\sum_{j=1}^{i} \alpha_j x_j + \sum_{j=1}^{i} \beta_j y_j \le \delta \tag{14}$$

is valid for $PK(T_i)$.

Proof We prove the result by induction. The base case is assumed to be true. Assume now that $\sum_{j=1}^{i-1} \alpha_j x_j + \sum_{j=1}^{i-1} \beta_j y_j \le \delta$ is valid for $PK(T_{i-1})$ and that there exist α_i and β_i satisfying (13). From the definition of the perturbation function $\delta - \sum_{j=1}^{i-1} \alpha_j x_j - \sum_{j=1}^{i-1} \beta_j y_j \ge P^{i-1}(w)$, as long as (x, y) satisfy (12). Then, if $x_i = \phi_1$ and $y_i = \phi_2$ (13), proves the validity of (14). □

The lifting method that is underlying Proposition 14 is systematic and constructive. If the seed inequality defines a facet of $PK(T)$ and the coefficients (α_i, β_i) for $i \in T$ are chosen in such a way that (14) is satisfied at equality by two new affinely independent points $(\phi_1, \phi_2) \in \{0, 1\} \times [0, 1]$ with $(\phi_1, \phi_2) \neq (0, 0)$, then the resulting inequality is facet-defining for PK.

If performed exactly, lifting generates strong inequalities for the problem considered. However, it is limited in that (1) the computation of each single function $P^i(w)$ might be difficult and/or computationally prohibitive, and (2) all the functions $P^i(w)$ must be computed. These limitations can be alleviated if the perturbation function $P^{t-1}(w)$ is well-structured. The following proposition shows that if $P^{t-1}(w)$ is superadditive then the perturbation function does not change after pairs of variables are lifted. This result follows directly from Corollary 7. It can also be derived as a consequence of Theorem 13. In order to see this, define h_1 and h_2 as summation operators. Let $\kappa(x_t, y_t, \ldots, x_n, y_n) = \sum_{j=t}^{n} \rho(x_j, y_j)$. Then, $r(x_t, y_t, \ldots, x_n, y_n) = \sum_{j=t}^{n} \left(\rho_j(x_j, y_j) - \rho_j(0, 0)\right)$ and $P(w)$ of Sect. 2 corresponds to $P^{t-1}(w)$ defined above. Assumption (A1) is satisfied since h_1 is a summation operator and (A3) is satisfied because the perturbation relaxes the feasible region in the definition of $P^{t-1}(\cdot)$. Superadditivity of $P^{t-1}(w)$ is precisely Assumption (A2). Therefore, $\sum_{j=1}^{n} \left(\alpha_j x_j + \beta_j y_j\right) \le \delta$ is valid for PK as long as $\alpha_j x_j + \beta_j y_j$ underestimates $P^{t-1}\left(\rho_j(x_j, y_j) - \rho_j(0, 0)\right)$. We also include below a direct algebraic proof along the lines of similar proofs in the integer programming literature.

Proposition 15 *Assume that (10) is valid for $PK(T)$, $P^{t-1}(w)$ is superadditive, i.e., $P^{t-1}(w_1) + P^{t-1}(w_2) \le P^{t-1}(w_1 + w_2)$, $\forall w_1, w_2 \in \mathbb{R}$, and that there exist α_i and β_i for $i \in T$ that satisfy*

$$\alpha_i \phi_1 + \beta_i \phi_2 \le P^{t-1}(\rho_i(\phi_1, \phi_2) - \rho_i(0, 0)) \quad \forall \phi_1 \in \{0, 1\}, \phi_2 \in [0, 1]. \tag{15}$$

Then, for $i \in T$, the inequality

$$\sum_{j=1}^{i} \alpha_j x_j + \sum_{j=1}^{i} \beta_j y_j \le \delta$$

is valid for $PK(T_i)$.

Proof We show by induction that $P^i = P^{t-1}$ for $i = t-1, \dots, n$. The rest of the result then follows from Proposition 14. The base case is straightforward. We assume now by induction that $P^{i-1} = P^{i-2} = \dots = P^{t-1}$ and prove that $P^i = P^{t-1}$. First note that $P^i(w) \le P^{i-1}(w) \le P^{t-1}(w)$. Further, using dynamic programming arguments, it can easily be seen that

$$P^i(w) = \inf_{\phi_1 \in \{0,1\} \phi_2 \in [0,1]} \Big\{ P^{i-1}\left(w + \rho_i(\phi_1, \phi_2) - \rho_i(0,0)\right) - \alpha_i \phi_1 - \beta_i \phi_2 \Big\}.$$

Using the fact that $P^{i-1} = P^{t-1}$ and that P^{t-1} is superadditive, we conclude that

$$\begin{aligned} P^i(w) &= \inf_{\phi_1 \in \{0,1\}, \phi_2 \in [0,1]} \Big\{ P^{t-1}\left(w + \rho_i(\phi_1, \phi_2) - \rho_i(0,0)\right) - \alpha_i \phi_1 - \beta_i \phi_2 \Big\} \\ &\ge P^{t-1}(w) + \inf_{\phi_1 \in \{0,1\}, \phi_2 \in [0,1]} \Big\{ P^{t-1}\left(\rho_i(\phi_1, \phi_2) - \rho_i(0,0)\right) - \alpha_i \phi_1 - \beta_i \phi_2 \Big\}. \end{aligned}$$

Finally, note that the conditions (15) governing the choice of α_i and β_i imply $P^i(w) \ge P^{t-1}(w)$. We therefore conclude that $P^i(w) = P^{t-1}(w)$. □

Observe that the superadditivity of the perturbation function is not typically needed over its complete domain. Indeed, the above argument holds even if the perturbation function is only superadditive over all realizable values of $\sum_{i=t}^{n} (\rho_i(x_i, y_i) - \rho_i(0,0))$ in the feasible region. Proposition 15 extends sequence-independent lifting to particular forms of nonlinear mixed-integer programs. It can be used even if the functions ρ_j have different forms for each j. Finally, provided that the lifting coefficients α_j and β_j can be derived quickly from the expression of P^{t-1}, Proposition 15 provides an efficient way to generate valid inequalities for MINLPs that contain *PK* as a substructure.

Proposition 15 has applications to linear and nonlinear programs. As an example, in the lifting of flow cover inequalities for the single node flow model, the lifting function is traditionally chosen as a one-dimensional function even though the integer program has $n+1$ inequalities. Proposition 15 shows that a one-dimensional perturbation function is natural in this context since the model can be reformulated as a nonlinear knapsack when $\rho_1, \dots, \rho_n$ are appropriately defined. We expand on this observation in the following example.

Example 16 We consider the single node flow model without inflows, which was initially studied by Padberg et al. [28], and Van Roy and Wolsey [39]. We can model this set as a nonlinear knapsack set of the form *PK* by setting $d > 0$ and by defining $\rho_j(x_j, y_j)$, for all $j \in N$, as follows:

$$\rho_j(x_j, y_j) = \begin{cases} m_j y_j & \text{if} \quad y_j \le x_j, \\ \infty & \text{otherwise,} \end{cases}$$

where we assume that $m_j > 0$. A flow cover F is a subset of N that is such that $\sum_{j \in F} m_j = d + \lambda$ for some $\lambda > 0$. We assume that $m_1 \ge \lambda$ and without loss of generality that $F = \{1, \ldots, p\}$. We define $PK(S_0, S_1) = \{(x, y) \in PK \mid x_j = y_j = 0 \; \forall j \in S_0 \text{ and } x_j = y_j = 1 \; \forall j \in S_1\}$. The defining inequality of $PK(N \backslash F, F \backslash \{1\})$ is $\rho_1(x_1, y_1) \leq m_1 - \lambda$. The convex hull of this two-dimensional set is polyhedral and its only nontrivial facet-defining inequality is

$$m_1 y_1 - (m_1 - \lambda)x_1 \leq 0. \tag{16}$$

We will use (16) as the seed inequality for lifting. It can easily be verified that the perturbation function associated with (16) is

$$P^1(w) = \begin{cases} -\lambda & \text{if } -\infty < w \leq -\lambda \\ w & \text{if } -\lambda < w \leq 0 \\ 0 & \text{if } 0 < w \leq m_1 - \lambda \\ \infty & \text{if } m_1 - \lambda < w \end{cases} \tag{17}$$

and that $P^1(w)$ is superadditive for $w \leq 0$. We now lift the variables (x_i, y_i) for $i \in F \backslash \{1\}$. Note that because these variables were fixed at 1 rather than 0, Proposition 15, suitably adapted, states that the lifting coefficients (α_i, β_i) of (x_i, y_i) satisfy

$$\begin{aligned} \beta_i &= \sup_{0 \leq \phi_2 < 1} \frac{-P^1(\rho_i(1, \phi_2) - \rho_i(1, 1))}{1 - \phi_2} \\ \alpha_i &= \sup_{0 \leq \phi_2 \leq 1} \Big(-P^1(\rho_i(0, \phi_2) - \rho_i(1, 1)) - \beta_i(1 - \phi_2)\Big) \end{aligned}$$

for $i \in F \backslash \{1\}$. Using (17), it is easily seen that $\alpha_i = -(m_i - \lambda)^+$ and $\beta_i = m_i$ for $i = 2, \ldots, p$. Therefore, we obtain that

$$\sum_{i \in F} -(m_i - \lambda)^+ x_i + \sum_{i \in F} m_i y_i \leq \sum_{i \in F \backslash \{1\}} m_i + \sum_{i \in F \backslash \{1\}} -(m_i - \lambda)^+ \tag{18}$$

is valid for $PK(N \backslash F, \emptyset)$. Note that because $\rho_i(x_i, y_i) \geq \rho_i(0, 0)$ for $i \in N \backslash F$, (18) is also valid for PK. Inequality (18) is the well-known flow cover inequality, which plays a central role in the study of single node flow models. We refer the reader to Gu et al. [15], Atamtürk [1], Shebalov and Klabjan [33], and Louveaux and Wolsey [21] among others for a discussion of other valid and facet-defining inequalities for fixed-charge flow problems. □

3.2 Two families of strong lifted inequalities for bilinear knapsack sets

In this section, we study mixed-integer bilinear knapsack sets obtained by choosing $\rho_j(x_j, y_j) = a_j x_j y_j$ in the definition of K. More precisely, we study

$$B = \left\{ (x, y) \in \{0,1\}^n \times [0,1]^n \;\middle|\; \sum_{j=1}^{n} a_j x_j y_j \le d \right\}$$

under the assumption that $n \in \mathbb{Z}_+$, $a_j > 0, \forall j = 1, \dots, n$, and $d > 0$. Although we use the set B primarily to illustrate the strength of nonlinear lifting, we note that B occurs as a relaxation of outsourcing problems where a quantity y_j of a certain product does not use the capacity of an available resource if it has been outsourced ($x_j = 0$). The set B also occurs in compact linearizations of 0–1 quadratic programs recently proposed by Chaovalitwongse et al. [10]. In this paper, the authors reformulate $x^t Qx$ as $x^t y$ with $y = Qx$ where bounds for y are derived by bounding Qx. They then linearize the term $x^t y$ using the fact that x is binary. We note that this transformation introduces only a linear number of additional variables and constraints. Further, observe that this scheme transforms $x^t Qx \le 0$ into the defining constraint of B whenever Q is a non-negative matrix.

Next, we study the geometric structure of B in more detail. Although the set B is defined using a nonlinear inequality, its convex hull, PB, is a polyhedron since B is expressible as a union of a finite number of polytopes.

Proposition 17 *PB is a full-dimensional polyhedron.* □

Even though the set PB has a simple defining inequality, it has a very rich polyhedral structure. We now illustrate the variety of facet-defining inequalities of PB on an example. The description of the convex hull of feasible solutions to this problem was obtained using PORTA; see Christof and Löbel [11]. Its complete linear description is given in the Appendix.

Example 18 Consider the bilinear mixed-integer knapsack set PB defined by

$$19x_1y_1 + 17x_2y_2 + 15x_3y_3 + 10x_4y_4 \le 20.$$

The linear description of this polytope contains 64 inequalities. We list a subset of nine inequalities below

$$7x_2 + 7x_4 + 17y_2 + 10y_4 \le 34 \quad (19)$$

$$7x_2 + 5x_3 + 7x_4 + 17y_2 + 15y_3 + 10y_4 \le 49 \quad (20)$$

$$63x_1 + 63x_2 + 63x_4 + 133y_1 + 153y_2 + 90y_4 \le 439 \quad (21)$$

$$63x_1 + 63x_2 + 45x_3 + 63x_4 + 133y_1 + 153y_2 + 135y_3 + 90y_4 \le 574 \quad (22)$$

$$63x_1 + 63x_2 + 63x_3 + 63x_4 + 133y_1 + 153y_2 + 189y_3 + 126y_4 \le 664 \quad (23)$$

$$19x_1 + 17x_2 + 15x_3 + 19y_1 + 17y_2 + 15y_3 \le 71 \quad (24)$$

$$19x_1 + 17x_2 + 15x_3 + 10x_4 + 19y_1 + 17y_2 + 15y_3 + 10y_4 \le 81 \quad (25)$$

$$x_1 \le 1 \quad (26)$$

$$x_1 \ge 0. \quad (27)$$

□

Among the above inequalities, (24)–(27) are easy to derive. In particular, it is easily proven that bound inequalities $x_i \le 1$, $x_i \ge 0$ and $y_i \ge 0$ are always facet-defining for PB and that $y_i \le 1$ is facet-defining for PB whenever $a_i \le d$. Also, given any set $F \subseteq N$, we can use the underestimation $x_i y_i \ge x_i + y_i - 1$ for $i \in F$ and $x_i y_i \ge 0$ for $i \in N \backslash F$ to obtain the valid inequality

$$\sum_{i \in F} a_i x_i + \sum_{i \in F} a_i y_i \le d + \sum_{i \in F} a_i. \tag{28}$$

Inequality (28) may be facet-defining for PB, as illustrated by (24)–(25).

Next, we obtain facet-defining inequalities for PB via lifting. The first family is obtained using sequence-dependent lifting tools. It uses clique inequalities as seeds and illustrates that the coefficients of the continuous variables in the facet-defining inequalities of PB can be much more diverse than those encountered in mixed-integer linear knapsack sets. The second family is obtained from covers using sequence-independent tools. Because B reduces to a 0–1 knapsack set when the continuous variables are fixed at one and because cover inequalities have been shown to be both theoretically [6] and practically [12] important for 0–1 knapsack sets, we believe that these inequalities are important in the description of PB. Further, these inequalities will help us illustrate two important facts about nonlinear lifting. First, because we show in Sect. 4 that lifted clique and lifted cover inequalities are difficult to obtain from a reasonably tight integer programming reformulation of B, these inequalities illustrate the potential of using lifting in nonlinear settings, even when linear reformulations exist. Second, because lifted cliques form a subfamily of lifted cover inequalities, these inequalities also illustrate the fact that, although sequential lifting is a more general tool than sequence-independent lifting, aggressively searching for situations where sequence-independence conditions hold, yields many inequalities whose overall structure is not transparent when performing sequential lifting, and the geometric intuitions of Sect. 2 help in conducting this search.

To simplify the description of the subsequent lifting procedures, we introduce the following notation. For $N_0, N_1 \subseteq N$ such that $N_0 \cap N_1 = \emptyset$ and $\tilde{N}_0, \tilde{N}_1 \subseteq N$ such that $\tilde{N}_0 \cap \tilde{N}_1 = \emptyset$, we define

$$PB(N_0, N_1, \tilde{N}_0, \tilde{N}_1) = \left\{ (x, y) \in B \;\middle|\; \begin{array}{l} x_j = 0 \text{ for } j \in N_0, x_j = 1 \text{ for } j \in N_1 \\ y_j = 0 \text{ for } j \in \tilde{N}_0, y_j = 1 \text{ for } j \in \tilde{N}_1 \end{array} \right\}.$$

3.2.1 Lifted clique inequalities

The notion of a clique inequality was introduced for node packing problems in [27]. We restate the definition for PB as follows:

Definition 19 Let $K \subseteq N$. We say that K is a clique for B if $a_i + a_j > d$ for all $i, j \in K$ such that $i \neq j$. □

For a clique K, it is easy to verify that the inequality

$$\sum_{j\in K} x_j \leq 1 \tag{29}$$

is valid for $PB(N\backslash K, \emptyset, N\backslash K, K)$. Further, it is facet-defining if $a_j \leq d$ for all $j \in K$. We will now obtain a family of facet-defining inequalities for PB by lifting the seed inequality (29) first with respect to the continuous variables y_j fixed at 1 and then with respect to the pairs of variables (x_j, y_j) fixed at $(0, 0)$. We note that the lifting coefficients of the variables y_j can be obtained from a simple modification to Proposition 14 while the lifting coefficients of pairs of variables, (x_j, y_j), are obtained directly from Proposition 14. We assume without loss of generality that $K = \{1, \dots, k\}$ and that $a_1 \geq \dots \geq a_k$.

Proposition 20 *Let K be a clique for B and assume that $a_i \leq d$ for $i \in K$. The inequality*

$$\sum_{j\in K} x_j + \sum_{j\in K} \frac{a_j}{a_j - d + \min\{a_i \mid i \in K, i \neq j\}}(y_j - 1) \leq 1 \tag{30}$$

is facet-defining for PB.

Proof We prove by induction on j that the lifting coefficient of y_j, denoted by β_j, equals $\frac{a_j}{a_j - d + \min\{a_i \mid i\in K, i\neq j\}}$ for $j = 1, \dots, k-1$. A simple variation of Proposition 14 that fixes y_j at 1, instead of 0, shows that the lifting coefficient is:

$$\begin{aligned} \hat{\beta}_j = \max\ & \frac{\sum_{i\in K} x_i + \sum_{i=1}^{j-1} \beta_i (y_i - 1) - 1}{1 - y_j} \\ \text{s.t. } & (x, y) \in B_j \\ & y_j < 1, \end{aligned} \tag{P}$$

where $B_j = \{(x, y) \in \{0, 1\}^k \times [0, 1]^j \mid \sum_{i=1}^{j} a_i x_i y_i + \sum_{i=j+1}^{k} a_i x_i \leq d\}$. Assume that we have proven that $\hat{\beta}_l = \beta_l$ for $l = 1, \dots, j - 1$. We next prove that the result holds for j. It can easily be seen that the convex hull of B_j is a polytope and therefore has a finite number of extreme points. Consider now the problem

$$\begin{aligned} \beta'_j = \max\ & \frac{\sum_{i\in K} x_i + \sum_{i=1}^{j-1} \beta_i (y_i - 1) - 1}{1 - y_j} \\ \text{s.t.} & (x, y) \in V(\text{conv}(B_j)) \\ & 0 < y_j < 1 \\ & x_j = 1 \\ & y_i = 1 \qquad \forall i = 1, \dots, j - 1 \end{aligned} \tag{R}$$

where $V(\text{conv}(B_j))$ represents the set of extreme points of $\text{conv}(B_j)$.

We first show that $\beta'_j = \beta_j$ and then we show that $\hat{\beta}_j = \beta'_j$. We claim that $\beta'_j \geq \beta_j$. Note that when the x variables are fixed, B_j reduces to a traditional continuous knapsack set. Therefore, if $(x, y) \in V(\text{conv}(B_j))$, then each $(x_i, y_i) \in \{(0,0), (1,0), (0,1), (1,1), (1, f)\}$, for $i \in \{1, \ldots, j\}$, where $0 \leq f \leq 1$. Further at most one y_i is fractional. Let $l' = \max\{l | l \neq j\}$ and consider the solution $(x_i, y_i) = (0, 1)$ for $i = 1, \ldots, \min\{j-1, l'-1\}$, $(x_j, y_j) = (1, 1 - \frac{1}{\beta_j})$, $x_i = 0$ for $i = j+1, \ldots, k-1$ and $x_{l'} = 1$. This solution, henceforth referred to as p^j, is feasible to (R) (as well as (P)) and has an objective value of β_j. Therefore, $\beta'_j \geq \beta_j$. Now, we show that $\beta'_j \leq \beta_j$. We first claim that $\sum_{i \in K} x_i \leq 2$ in the optimal solution to (R). Because K is a clique and $y_i = 1$ for $i < j$, x_i must be 0 for all but one of the variables x_i for $i \in K \backslash \{j\}$. Therefore, the numerator in the objective of (R) cannot be larger than 1. Further, since $y_j \leq 1 - \frac{1}{\beta_j}$, the denominator cannot be smaller than $\frac{1}{\beta_j}$. Therefore, $\beta'_j \leq \beta_j$. This completes our proof of $\beta'_j = \beta_j$.

Now, we claim that $\beta'_j = \hat{\beta}_j$. Clearly, $\beta'_j \leq \hat{\beta}_j$ since the feasible region of (R) is a subset of that of (P). Assume now, so as to derive a contradiction, that $\hat{\beta}_j > \beta'_j$. Let $(x', y') \in B_j$ be a point that most violates the following inequality:

$$\eta(x, y) = \sum_{i \in K} x_i + \sum_{i=1}^{j-1} \beta_i (y_i - 1) + \beta'_j (y_j - 1) - 1 \leq 0. \tag{31}$$

Such a point exists since $\hat{\beta}_j > \beta'_j$, $\eta(x, y)$ is continuous and B_j is compact. By linearity of $\eta(x, y)$, we can pick (x', y') from $V(\text{conv}(B_j))$. We assume, without loss of generality that $y'_j > 0$. Otherwise, define $x''_i = x'_i$ for $i \neq j$ and $y''_i = y'_i$ for $i < j$ and $(x''_j, y''_j) = (0, 1)$. Then, $(x'', y'') \in B_j$. However, $\eta(x'', y'') \geq \eta(x', y')$ since $\beta'_j \geq \beta_j \geq 1$, and, therefore, (x'', y'') is a solution of the desired type. Now, we show that $y'_j < 1$. This follows from the induction hypothesis, since $\eta(x, y) \leq 0$ is valid for B_j when $y_j = 1$. Since, for $i < j$, $\beta_i \geq 1$, we may assume that $(x'_i, y'_i) \in \{(0, 1), (1, 1)\}$ for $i = 1, \ldots, j-1$. We have thus shown that (x', y') is feasible to (R), which is a contradiction to the assumption that β'_j is the optimal value of (R). Therefore, $\beta'_j = \hat{\beta}_j$. Combining the two steps, $\hat{\beta}_j = \beta'_j = \beta_j$.

We now prove that the lifting coefficients for (x_j, y_j), $j \in N \backslash K$ can be chosen to be zero. Consider (x, y) feasible to B. Let (x', y') be such that $x'_j = x_j$ and $y'_j = y_j$ for $j \in K$ and $(x'_j, y'_j) = 0$ for $j \in N \backslash K$. Then, (x', y') is feasible to $PB(N \backslash K, \emptyset, N \backslash K, \emptyset)$, and therefore satisfies (30). Since (x, y) matches (x', y') on the support of (30), (x, y) satisfies it as well.

Now we show that (30) is facet-defining for PB. Let $q^j = (e_j, \sum_{i=1}^k e_i)$, for $j = 1, \ldots, k$, where e_i is the ith unit vector in $\mathbb{R}^n$. These points are tight for the seed inequality. When the restriction $y_j = 1$ for $j \in K$ is relaxed, we showed earlier that p^j is tight for (30). When the restriction $x_j = 0$ is relaxed for $j \in N \backslash K$, the point $r_1^j = q^1 + (e_j, 0)$ is tight for (30). Similarly, when the restriction $y_j = 0$ is relaxed for $j \in N \backslash K$, the point $r_2^j = q^1 + (0, e_j)$ is tight for (30). Therefore, the above $2n$ affinely independent points show that (30) is facet-defining for PB. □

We now illustrate the use of Proposition 20 on the bilinear knapsack set presented in Example 18.

Example 21 Consider the set *PB* introduced in Example 18. Because $K = \{1, 2, 3, 4\}$ is a clique and $a_i \leq d$ for $i \in K$, the inequality

$$x_1 + x_2 + x_3 + x_4 \leq 1$$

is facet-defining for $PB(\emptyset, \emptyset, \emptyset, N)$. This clique inequality can be lifted with respect to the continuous variables into the following inequality:

$$x_1 + x_2 + x_3 + x_4 + \frac{19}{9}y_1 + \frac{17}{7}y_2 + \frac{15}{5}y_3 + \frac{10}{5}y_4 \leq 1 + \frac{19}{9} + \frac{17}{7} + \frac{15}{5} + \frac{10}{5}$$

or equivalently

$$63x_1 + 63x_2 + 63x_3 + 63x_4 + 133y_1 + 153y_2 + 189y_3 + 126y_4 \leq 664.$$

This inequality is facet-defining for *PB* and corresponds to (23) in Example 18. □

In Sect. 3.2.2, we generalize the lifted clique inequality by exploiting the sequence-independent lifting theory.

3.2.2 Lifted cover inequalities

The notion of a cover inequality was introduced for the 0–1 knapsack polytope by Wolsey [41], Balas [3] and Hammer et al. [17]. We apply their definition to *PB* as follows:

Definition 22 A set $C \subseteq N$ is said to be a cover for B if $\sum_{j \in C} a_j = d + \mu$ where $\mu > 0$. Furthermore, μ is said to be the excess of the cover. □

Assume that a cover C is known for B. We assume without loss of generality that $C = \{1, \ldots, p\}$ and that $a_1 \geq a_2 \geq \ldots \geq a_p$. To obtain lifted cover inequalities, we first fix the variables (x_j, y_j) for $j \in C \backslash \{1\}$ to $(1, 1)$ and the variables (x_j, y_j) for $j \in N \backslash C$ to $(0, 1)$. The defining inequality of $PB(N \backslash C, C \backslash \{1\}, \emptyset, N \backslash \{1\})$ is

$$a_1 x_1 y_1 \leq a_1 - \mu. \tag{32}$$

Assuming that $a_1 > \mu$, it is easy to verify that

$$\mu x_1 + a_1 y_1 \leq a_1 \tag{33}$$

is the only nontrivial facet-defining inequality of $PB(N \backslash C, C \backslash \{1\}, \emptyset, N \backslash \{1\})$. To obtain a strong inequality for *PB*, we lift the seed inequality (33) in two steps. First, we lift it with respect to the variables (x_j, y_j) for $j \in C \backslash \{1\}$. Second, we lift it with respect to the variables (x_j, y_j) for $j \in N\backslash C$. In order to perform these liftings, we first derive a closed-form expression for the perturbation function

$$P^1(w) = a_1 - \max_{x_1 \in \{0,1\}, y_1 \in [0,1]} \left\{ \mu x_1 + a_1 y_1 \;\middle|\; a_1 x_1 y_1 \le a_1 - \mu - w \right\} \tag{34}$$

associated with (33) and then verify that it is superadditive.

Proposition 23 *The perturbation function of* (33) *is*

$$P^1(w) = \begin{cases} -\mu & \text{if } w \le -\mu \\ w & \text{if } -\mu < w \le 0 \\ 0 & \text{if } 0 < w \le a_1 - \mu \\ \infty & \text{if } a_1 - \mu < w. \end{cases}$$

Moreover, $P^1(w)$ is superadditive for $w \le 0$.

Proof When $w > a_1 - \mu$, it is easily seen that the feasible region of (34) is empty and therefore $P^1(w) = \infty$. When $w \le a_1 - \mu$, it can be verified that an optimal solution to (34) is either obtained by setting $x_1^* = 0$, $y_1^* = 1$ or by setting $x_1^* = 1$ and $y_1^* = \min\{\frac{a_1 - \mu - w}{a_1}, 1\}$. Therefore, we obtain $P^1(w) = a_1 - \max\{a_1, \mu + \min\{a_1 - \mu - w, a_1\}\} = \min\{0, \max\{w, -\mu\}\}$. The proof that $P^1(w)$ is superadditive over $(-\infty, 0]$ is straightforward. □

Because $P^1(w)$ is superadditive over $(-\infty, 0]$, the result of Proposition 15 can be used to obtain the following family of facet-defining inequalities for $PB(N\backslash C, \emptyset, \emptyset, N\backslash C)$.

Relating the current setup to that of Sect. 2,

$$\gamma(1, 1, \ldots, x_i, y_i, \ldots, 1, 1) = \begin{cases} -\min\{a_i, \mu\} & \text{if } x_i = 0 \\ -\min\{a_i - a_i y_i, \mu\} & \text{if } x_i = 1. \end{cases}$$

It can be easily verified that

$$\min\{a_i, \mu\}(x_i - 1) + a_i(y_i - 1) \le \gamma(1, 1, \ldots, x_i, y_i, \ldots, 1, 1).$$

Since the seed inequality may be written as $\min\{a_1, \mu\}(x_1 - 1) + a_1(y_1 - 1) \le -\mu$, it follows from Theorem 13 that (35) is valid for $PB(N\backslash C, \emptyset, \emptyset, N\backslash C)$. A direct proof is presented next.

Proposition 24 *Let $C \subseteq N$ be a cover for B satisfying $a_1 > \mu$. Then*

$$\sum_{j\in C} \min\{a_j, \mu\} x_j + \sum_{j \in C} a_j y_j \le \sum_{j\in C} \min\{a_j, \mu\} + \sum_{j\in C} a_j - \mu \tag{35}$$

is a facet-defining inequality of $PB(N\backslash C, \emptyset, \emptyset, N\backslash C)$.

Proof Note first that, because the variables (x_j, y_j) for $j \in C\backslash\{1\}$ were fixed at (1,1), the results of Proposition 15 must be reformulated. It can be easily established that, because $P^1(w)$ is superadditive over $(-\infty, 0]$, we can obtain valid lifting coefficients for (x_j, y_j) if we choose (α_j, β_j) to satisfy the conditions

$$\beta_j \geq \sup_{0\leq\phi<1} \frac{-P^1(a_j\phi - a_j)}{1-\phi} \tag{36}$$

$$\alpha_j + \inf_{0\leq\phi\leq 1} \beta_j(1-\phi) \geq -P^1(-a_j). \tag{37}$$

First, we compute β_j. Because $a_j > 0$, there are only two cases. In the first case, $\mu > a_j \geq 0$. From Proposition 23, $P^1(a_j\phi - a_j) = a_j\phi - a_j$. Therefore, $\beta_j \geq a_j$. In the second case, $a_j \geq \mu$. From Proposition 23, $P^1(a_j\phi - a_j) = a_j\phi - a_j$ when $a_j\phi - a_j > -\mu$ and $P^1(a_j\phi - a_j) = -\mu$ otherwise. When $a_j\phi - a_j > -\mu$, we conclude that $\sup_{0\leq\phi<1} \frac{-P^1(a_j\phi-a_j)}{1-\phi} \leq a_j$. When $a_j\phi - a_j \leq -\mu$, we conclude that $\sup_{0\leq\phi<1} \left\{ \frac{\mu}{1-\phi} \middle| a_j\phi - a_j \leq -\mu \right\} \leq a_j$. Because the supremum is achieved when $a_j\phi - a_j = -\mu$, we find that $\beta_j \geq a_j$. We now compute α_j. Because $\beta_j > 0$, we deduce from (37) that $\alpha_j \geq -P^1(-a_j)$. Therefore, the best lifting coefficients are $\alpha_j = \min\{a_j, \mu\}$ and $\beta_j = a_j$. By Proposition 23, (35) is valid for $PB(N\backslash C, \emptyset, \emptyset, N\backslash C)$. Finally observe that (36) and (37) are satisfied at equality for our choice of α_j and β_j and that the supremum and infimum are achieved in (36) and (37). We conclude that (35) is facet-defining for $PB(N\backslash C, \emptyset, \emptyset, N\backslash C)$. □

We next illustrate the use of Proposition 24 on an example.

Example 25 Consider the bilinear mixed-integer knapsack set *PB* defined by

$$19x_1y_1 + 17x_2y_2 + 15x_3y_3 + 10x_4y_4 + 6x_5y_5 + 2x_6y_6 \leq 62.$$

The set $C = \{1, 2, 3, 4, 5, 6\}$ forms a cover whose excess $\mu = 7$. Note that this cover satisfies the assumptions of Proposition 24. It follows that

$$\begin{aligned} 7x_1 &+ 7x_2 + 7x_3 + 7x_4 + 6x_5 + 2x_6 \\ &+ 19y_1 + 17y_2 + 15y_3 + 10y_4 + 6y_5 + 2y_6 \leq 98 \end{aligned} \tag{38}$$

is facet-defining for *PB*. □

We observed in Example 18 that valid inequalities of the form (28) could easily be obtained for any $F \subseteq N$ using simple linearization arguments. Proposition 24 shows that this linearization procedure can be improved in the presence of 0–1 variables. In fact, (35) can be seen as a strengthening of (28) obtained by reducing the coefficients of integer variables and the constraint right-hand-side by the same amount for all integer variables with a reasonably large coefficient.

We now lift the remaining pairs of fixed variables in (35), i.e., (x_j, y_j) for $j \in N \backslash C$. These variables are fixed at (0, 1). To perform this lifting, we first compute in Proposition 26 a closed-form expression for the perturbation function associated with (35). This function is defined as

$$\begin{aligned} P^p(w) = \min & \sum_{j \in C} \min\{a_j, \mu\}(1 - x_j) + \sum_{j \in C} a_j(1 - y_j) - \mu \\ s.t. & \sum_{j \in C} a_j x_j y_j \leq \sum_{j \in C} a_j - \mu - w \\ & x_j \in \{0, 1\} \forall j \in C \\ & y_j \in [0, 1] \forall j \in C. \end{aligned} \tag{39}$$

In Proposition 28, we show that $P^p(w)$ is superadditive over $[0, d]$ and use this result to derive the desired lifting coefficients in Theorem 30.

We next derive an analytical form for the perturbation function (39). Towards this end, let $q \in C$ be the index for which $a_q > \mu \geq a_{q+1}$. Further, let $A_0 = 0$, $A_l = \sum_{j=1}^{l} a_j$ for $l = 1, \ldots, p$. Note that $A_p = \sum_{j \in C} a_j = d + \mu$.

Proposition 26 *For $w \geq 0$, the perturbation function associated with (35) is given by*

$$P^p(w) = \begin{cases} (i-1)\mu & if\ A_{i-1} \leq w \leq A_i - \mu \\ i\mu + w - A_i & if\ A_i - \mu \leq w \leq A_i \\ q\mu + w - A_q & if\ A_q \leq w \leq A_p - \mu \\ +\infty & otherwise, \end{cases}$$

where $i = 1, \ldots, q$.

Proof Let $w \in \mathbb{R}$. We make the following four observations about optimal solutions to (39). First, we observe that there is an optimal solution (x^*, y^*) in which no more than one of the continuous variables y^* is fractional and that if y_k^* is fractional then $x_k^* = 1$. This observation holds because, once the 0–1 variables are fixed, the problem reduces to a continuous knapsack problem. Second, we note that, because the objective coefficient of x_j is greater than or equal to that of y_j in (39), a solution with $(x_j^*, y_j^*) = (1, 0)$ for some $j \in \{1, \ldots, p\}$ is no better than the solution obtained by changing (x_j^*, y_j^*) to (0, 1). We conclude that there is an optimal solution in which the variables (x_j^*, y_j^*) take the values (0, 1), (1, 1), or $(1, f)$. Third, we observe that any solution where $(x_j^*, y_j^*; x_k^*, y_k^*) = (1, 1; 1, f)$ for $j < k$ is no better than the solution obtained by changing $(x_j^*, y_j^*; x_k^*, y_k^*)$ to $(1, f'; 1, 1)$ because $a_j \geq a_k$. Fourth, we note that any solution where $(x_j^*, y_j^*; x_k^*, y_k^*) = (1, f; 0, 1)$ for $j < k$ is no better than the solution obtained by changing $(x_j^*, y_j^*; x_k^*, y_k^*)$ to $(0, 1; 1, f')$ when either $a_k \geq \mu$ or $a_j f \leq a_k$. Finally, $(x_j^*, y_j^*, x_k^*, y_k^*) = (1, f'; 1, 1)$ is as good as $(x_j^*, y_j^*, x_k^*, y_k^*) = (1, f; 0, 1)$ when $a_k < \mu$ and $a_j f > a_k$. We conclude that, when $w \in [0, d]$, the following is an optimal solution for (39):

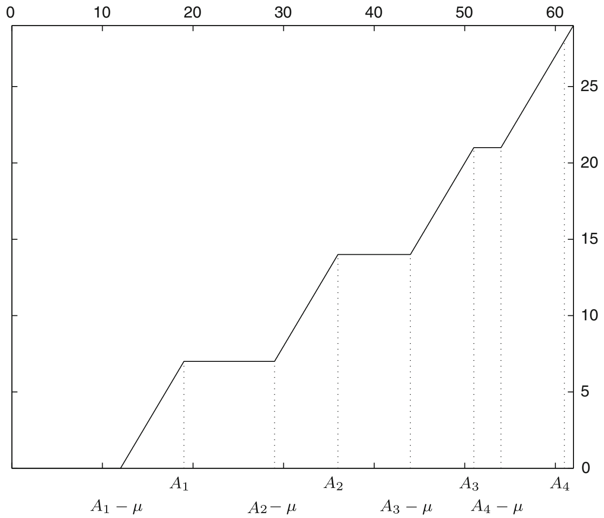

Fig. 4 Perturbation function of (38)

$$(x_j^*, y_j^*) = \begin{cases} (0, 1) & \text{if } j = 1, \ldots, k-1 \\ (0, 1) & \text{if } j = k \text{ and } \min\{a_k, \mu\} < a_k - d(w) \\ (1, \frac{d(w)}{a_k}) & \text{if } j = k \text{ and } \min\{a_k, \mu\} \geq a_k - d(w) \\ (1, 1) & \text{if } j = k+1, \ldots, p \end{cases}$$

where $k = \max \left\{ l \mid \sum_{j=l}^{p} a_j > d - w \right\}$ and $d(w) = d - w - \sum_{j=k+1}^{p} a_j$. □

Note that, in the description of the function $P^p(w)$, the interval $[A_q, A_p - \mu]$ may be empty if $\sum_{i=q+1}^{p} a_i \leq \mu$. In Fig. 4, we illustrate the function $P^p(w)$ derived in Proposition 26 for the cover inequality (38) presented in Example 25. In particular, we observe that $P^p(w)$ is a piecewise linear, nondecreasing, and continuous function that has only two slopes. We also observe that the heights of the plateaus of the function are integer multiples of μ.

This observation combined with the fact that the intervals $[0, A_1]$, $[A_1, A_2]$, $[A_2, A_3], \ldots, [A_{q-1}, A_q]$ are nondecreasing in length implies that the corresponding function $P^p(w)$ is superadditive. We use the following result of Richard [29] to prove the superadditivity of $P^p(w)$.

Proposition 27 *Let $C_0, \ldots, C_r$ be integers. Assume that $C_0 = 0$ and that $C_i \geq C_{i-1} + \lambda$ for some positive integer λ. Then the function*

$$\mathcal{P}(x) = \begin{cases} i & \text{if } C_i \leq x \leq C_{i+1} - \lambda \quad \text{for } i = 0, \ldots, r-1 \\ i + \frac{x - C_i}{\lambda} & \text{if } C_i - \lambda \leq x \leq C_i \quad \text{for } i = 1, \ldots, r \\ r + \frac{x - C_r}{\lambda} & \text{if } C_r \leq x \end{cases}$$

is superadditive if and only if $C_i + C_j \geq C_{i+j}$ for $0 \leq i \leq j \leq r$ with $i + j \leq r$. □

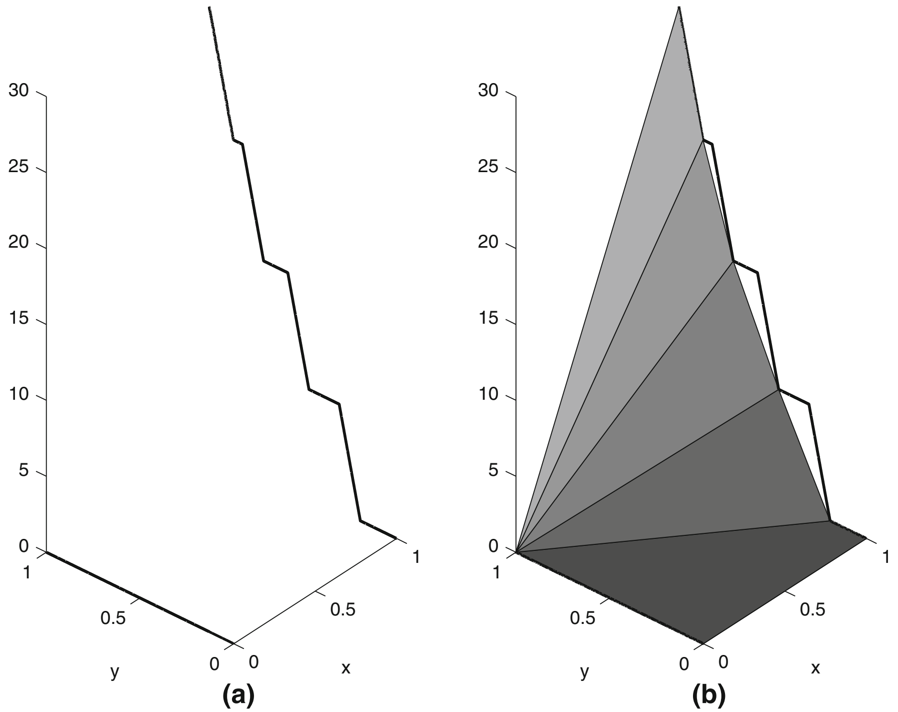

Fig. 5 Deriving lifting coefficients for (38)

Using the result of Proposition 27, we next prove that the perturbation function of the lifted cover inequality (35) is superadditive.

Proposition 28 *The perturbation function* $P^p(w)$ *associated with* (35) *is superadditive over* $[0, A_p - \mu]$.

Proof Note that, if we choose $r = p$ and $C_i = A_i$ for $i = 1, \ldots, p$, $\mathcal{P}(w) = \frac{P^p(w)}{\mu}$ for $w \leq A_p - \mu$. To prove that $P^p(w)$ is superadditive, it is therefore sufficient to prove that $A_i + A_j \geq A_{i+j}$ for $0 \leq i \leq j \leq r$. These conditions are trivially satisfied because A_i is the sum of the largest i coefficients of the cover. □

We now use the result of Propositions 15, 26, and 28 to obtain facet-defining inequalities for PB. A graphical representation of the lifting problem is given in Fig. 5. In Fig. 5a, we plot the amount of slack in constraint (35) for every feasible value of (x_j, y_j). When $x_j = 0$, the amount of slack is 0 since $P^p(a_j x_j y_j) = P^p(0) = 0$. When $x_j = 1$, the amount of slack is given by $P^p(a_j x_j y_j) = P^p(a_j y_j)$. Because adding $\alpha_j x_j + \beta_j y_j \leq \beta_j$ to the seed inequality must produce a valid inequality, the coefficients α_j and β_j must be selected in such a way that the plane they define lower-approximates the function in Fig. 5a. Further, if we want the inequality generated to be facet-defining, the underestimating plane must touch the function of Fig. 5a in at least two new affinely independent points.

We observe from Fig. 5b that various lifting coefficients, satisfying the above properties, can be derived from the initial cover inequality (35). Further, these coefficients can be obtained easily from a piecewise linear convex underestimator $p(w)$ of the function $P^p(w)$. Because $p(w)$ plays an important role in the derivation of lifting coefficients, we first express it in closed-form in the following lemma.

Lemma 29 *Let* $a_j \in [0, d]$ *and define* l_j *to be the only integer such that* $A_{l_j} - \mu \leq a_j < A_{l_j+1} - \mu$. *Let* $W_i^j = A_i - \mu$ *for* $i = 1, \ldots, l_j$ *and* $W_{l_j+1}^j = a_j$. *Define* $p_{0j}(w) = 0$ *and* $p_{ij}(w) = P^p(W_i^j) + \frac{P^p(W_{i+1}^j) - P^p(W_i^j)}{a_{i+1}}(w - W_i^j)$, *for* $i = 1, \ldots, l_j$. *Then, the function*

$$p(w) := \max\left\{p_{ij}(w) \mid i \in \{0, \ldots, l_j\}\right\} \tag{40}$$

is a convex underestimator of the perturbation function $P^p(w)$ *associated with* (35) *over* $[0, a_j]$.

Proof The fact that $p(w)$ is convex is easily established since $p(w)$ is defined as the maximum of a finite number of linear functions. It is also obvious that 0 underestimates $P^p(w)$. Note that the slope of $p_{ij}(w)$ is no more than that of $p_{i'j}(w)$ whenever $i < i'$. Therefore, it can be easily verified that the maximum in (40) is attained for $i = k$ when $w \in [W_k^j, W_{k+1}^j]$. Also, by definition, $p_{kj}(W_k^j) = P^p(W_k^j)$ and $p_{kj}(W_{k+1}^j) = P^p(W_{k+1}^j)$. Then, $p_{kj}(w) \leq P^p(w)$ because $P^p(w)$ is concave when $w \in [W_k^j, W_{k+1}^j]$. □

Next, we derive the lifting coefficients for the variables (x_j, y_j) for $j \in N \backslash C$, which are indexed from $p+1$ to n. We now relate the current setup to the notation of Sect. 2. Then, $\bar{\alpha} = \left(\min\{a_1, \mu\}, a_1, \ldots, \min\{a_p, \mu\}, a_p\right)$, $\delta = \sum_{j \in C}\left(\min\{a_j, \mu\} + a_j\right) - \mu$, $r(x_{p+1}, y_{p+1}, \ldots, x_n, y_n) = \sum_{j=p+1}^n a_j x_j y_j$, and

$$\gamma(x_{p+1}, y_{p+1}, \ldots, x_n, y_n) = P^p\left(r(x_{p+1}, y_{p+1}, \ldots, x_n, y_n)\right).$$

Let h_1 and h_2 (see Theorem 13) be summation operators. Assumption (A1) and (A3) are trivially satisfied as in the discussion before Proposition 15. Proposition 28 proves Assumption (A2), i.e.,

$$\gamma(x_{p+1}, y_{p+1}, \ldots, x_n, y_n) \geq \sum_{j=p+1}^{n} \gamma(0, 0, \ldots, x_j, y_j, \ldots, 0, 0).$$

Moreover, Proposition 29 derives a convex underestimator (actually the convex envelope; see Fig. 5) for $\gamma(0, 0, \ldots, x_j, y_j, \ldots, 0, 0)$. This proves the validity of exponentially many inequalities for B (the content of Theorem 30), a direct proof of which is provided next.

Theorem 30 *Let C be a cover for B such that $a_1 \geq \mu$. Then,*

$$\sum_{j\in C}\min\{a_j,\mu\}x_j + \sum_{j\in C} a_j y_j + \sum_{j\in N\backslash C}\alpha_j x_j + \sum_{j\in N\backslash C}\beta_j y_j \leq \sum_{j\in C}\min\{a_j,\mu\} + \sum_{j\in C} a_j + \sum_{j\in N\backslash C}\beta_j - \mu$$

is a facet-defining inequality for PB if

$$(\alpha_j,\beta_j) \in (0,0) \cup \bigcup_{i=1}^{l_j}\left(\frac{P^p(W^j_{i+1}) - P^p(W^j_i)}{a_{i+1}}(a_j - W^j_i) + P^p(W^j_i),\ a_j\frac{P^p(W^j_{i+1}) - P^p(W^j_i)}{a_{i+1}}\right).$$

for $j \in N\backslash C$, where W^j_i and l_j are as defined in Lemma 29.

Proof By superadditivity of $P^p(w)$ for $w \geq 0$, and Proposition 15 adapted suitably for y_j, $j \in N\backslash C$, fixed at 1 instead of 0, it follows that a lifted inequality is valid if the lifting coefficients (α_j,β_j) are chosen such that

$$\beta_j \geq \frac{-P^p(0)}{1-\phi} \quad \text{for all } 0 \leq \phi < 1 \tag{41}$$

$$\alpha_j + \beta_j(\phi - 1) \leq P^p(a_j\phi) \quad \text{for all } 0 \leq \phi \leq 1. \tag{42}$$

We next derive pairs of lifting coefficients (α_j,β_j) for the variables (x_j,y_j) that satisfy (41) and (42). Furthermore, we prove that each of the above lifting coefficients preserves the facet-defining character of the seed inequality during lifting by providing two points $\hat{\phi}$ and $\tilde{\phi}$ for which inequalities (41) and/or (42) are satisfied at equality.

First consider the coefficients $(\alpha_j,\beta_j) = (0,0)$. It is easy to verify that they satisfy (41) and (42) since $P^p(w)$ is nonnegative and $P^p(0) = 0$. Also it can easily be verified that equality holds in (42) for $\hat{\phi} = 0$ and $\tilde{\phi} = \min\left\{\frac{A_1-\mu}{a_j}, 1\right\}$.

Second, let

$$\alpha_j = \frac{P^p(W^j_{i+1}) - P^p(W^j_i)}{a_{i+1}}(a_j - W^j_i) + P^p(W^j_i),$$

$$\beta_j = a_j\frac{P^p(W^j_{i+1}) - P^p(W^j_i)}{a_{i+1}}.$$

Because $\beta_j > 0$, (α_j,β_j) satisfies (41). We must now show that (α_j,β_j) satisfies (42). It follows from Lemma 29 that

$$P^p(a_j\phi) \geq P^p(W^j_i) + \frac{P^p(W^j_{i+1}) - P^p(W^j_i)}{a_{i+1}}(a_j\phi - W^j_i)$$

$$
\begin{aligned}
&= P^p(W_i^j) + \frac{P^p(W_{i+1}^j) - P^p(W_i^j)}{a_{i+1}}(a_j - W_i^j) \\
&\quad + a_j \frac{P^p(W_{i+1}^j) - P^p(W_i^j)}{a_{i+1}}(\phi - 1) \\
&= \alpha_j + \beta_j(\phi - 1).
\end{aligned}
$$

Further, it can easily be verified that equality holds throughout for the points $\hat{\phi} = \frac{W_i^j}{a_j}$ and $\tilde{\phi} = \frac{W_{i+1}^j}{a_j}$. □

Note that, in Theorem 30, the perturbation function remains unchanged over $[0, d]$ after each pair of variable (x_j, y_j) is lifted for $j \in N\backslash C$. Note also that since we identified different choices of lifting coefficients (α_j, β_j) for each pair of variables (x_j, y_j), a single cover inequality yields an exponential number of lifted cover inequalities. Further, all of these inequalities are facet-defining for PB since the choice of (α_j, β_j) was made in such a way that two new tight affinely independent points were added to the inequality during the lifting of each pair of variables. It can also be observed from Fig. 5b that, had we fixed the variables (x_j, y_j) for $j \in N\backslash C$ to $(0, 0)$ instead of $(0, 1)$, we would have obtained only $(0, 0)$ as the lifting coefficients. It was therefore crucial to fix (x_j, y_j) to $(0, 1)$ for $j \in N\backslash C$ in our lifting procedure to generate an exponential number of inequalities. Incidentally, Theorem 30 provides a concrete example of the observation we made after Example 8 that it may be possible to derive an exponential number of lifted inequalities from sequence-independent lifting by choosing subgradients over subsets of variables independently.

Using the result of Theorem 30, we can explain many inequalities of the linear description of the mixed-integer bilinear knapsack set presented in Example 18. We derive some of them next.

Example 31 For the bilinear knapsack set described in Example 18, it can easily be verified that $C = \{2, 4\}$ is a cover whose excess is 7. We conclude from Proposition 24 that the cover inequality

$$
7x_2 + 7x_4 + 17y_2 + 10y_4 \leq 34 \tag{43}
$$

is facet-defining for $PB(T, \emptyset, \emptyset, T)$ where $T = \{1, 3\}$. We now determine the lifting coefficients of the pairs of variables (x_1, y_1) and (x_3, y_3) using the result of Theorem 30. Note first that the perturbation function $P^2(w)$ is given by

$$
P^2(w) = \begin{cases} 0 & \text{if } w \leq 10 \\ w - 10 & \text{if } 10 < w \leq 17 \\ 7 & \text{if } 17 < w \leq 20. \end{cases}
$$

Applying Theorem 30, we obtain the following four inequalities

$$7x_2+7x_4+17y_2+10y_4+\begin{cases} & \leq 34 \\ 7x_1 \quad\quad + \frac{133}{9}y_1 & \leq \frac{439}{9} \\ \quad\quad 5x_3 \quad\quad\quad + 15y_3 & \leq 49 \\ 7x_1 + 5x_3 + \frac{133}{9}y_1 + 15y_3 & \leq \frac{574}{9} \end{cases} \tag{44}$$

which are all facet-defining for PB. In fact, they are described as (19)–(22) in the linear description of PB. □

The exponential family of inequalities we presented for PB can be generated efficiently because the perturbation function $P^p(w)$ is naturally superadditive. In other applications, however, the perturbation function of an inequality of interest may not be superadditive. In such a case, it is still possible to derive valid inequalities efficiently if a superadditive lower approximation of the perturbation can be found. The inequalities obtained with this procedure will typically not be the strongest possible. However, for IPs, the idea of replacing perturbation functions with superadditive lower approximations has proven to be a very successful approach to generating strong inequalities. We refer to Gu et al. [16], Atamtürk [2], and Shebalov and Klabjan [33] among others for examples. We leave the investigation of approximate superadditive lifting in nonlinear settings for future research.

The lifted cover inequalities described in Theorem 30 and illustrated in Example 31 support the claim that the coefficients of continuous variables in the facets of the single constraint bilinear mixed-integer knapsack set can be extremely different from those of the defining knapsack constraint. This is in sharp contrast with the linear case where coefficients of the continuous variables typically have the ratio property discussed in Richard et al. [30], i.e., the ratio of the coefficients of continuous variables in a facet-defining inequality over their coefficients in the defining knapsack constraint is either 0 or a constant θ. This is reminiscent of the conclusion we drew after deriving lifted cliques inequalities. In fact, this is not surprising, since it can be shown that the family of lifted covers subsumes the family of lifted cliques.

Corollary 32 *Lifted clique inequalities of Proposition 20 are lifted covers with* $C = \{k-1, k\}$. □

Corollary 32 illustrates an interesting point about lifting. Although sequential lifting is a more general tool than sequence-independent lifting, sequence-independent lifting often reveals many inequalities whose structure is difficult to detect when variables are lifted sequentially.

Finally, we mention that although the set B we considered is nonlinear, we managed to describe a large family of facet-defining inequalities for its convex hull using nonlinear lifting. An interesting question is whether these inequalities can be obtained easily from a linearization of the set B. We will show in the following section that the lifted cover inequalities and lifted cliques inequalities are difficult to obtain from the natural linearization of the set B.

4 Nonlinear lifting and strength of inequalities

In the previous section, we derived strong inequalities for *PB* using nonlinear lifting. In this section, we show that these inequalities are difficult to obtain from the standard linearization of B. Towards this end, we first introduce an aggregation-tightening procedure that generalizes many cut-generation techniques for integer programs including lifting covers and mixed-integer rounding. We derive necessary and sufficient conditions that characterize when an inequality can be obtained using the aggregation-tightening procedure. We envision that this construction will also be useful in other contexts for proving strength of inequalities. Second, we specialize the construction to show that lifted clique and cover inequalities of Sect. 3.2 cannot be obtained using a single round of the aggregation-tightening procedure on a mixed-integer linear reformulation of the bilinear knapsack set.

A variety of techniques for generating cuts in integer programming proceed by first aggregating the constraints to generate a single valid inequality which is then strengthened using integrality restrictions on the variables. Examples of such a procedure include Chvátal–Gomory cuts and lifted cover inequalities. More abstractly, the above procedure derives a cut by tightening a valid inequality for a relaxation (in the above, the linear programming relaxation) using another relaxation (the integer lattice). We formalize this abstract notion of generating inequalities in the following procedure. Let S be a possibly nonconvex set. Let A_1 and A_2 be two sets that contain S. We define *aggregation-tightening* as the following two-step procedure that generates valid inequalities for S. First, we determine a hyperplane H such that A_1 belongs to one of the closed halfspaces of H, say H_+. Second, we find an inequality valid for $A_2 \cap H_+$. We denote the set of inequalities that can be obtained in this manner as $I(A_1, A_2)$.

We say an inequality $\alpha x \le \delta$ in $I(A_1, A_2)$ strongly separates a set A from S if there exists an $\varepsilon > 0$ such that $\alpha x \ge \delta + \varepsilon$ for all $x \in A$. We say two sets A and B can be separated strongly if there exists an α such that $\inf\{\alpha x \mid x \in A\} > \sup\{\alpha x \mid x \in B\}$.

Theorem 33 *Let A be a compact/polyhedral subset of $A_2 \backslash S$. Then, A can be strongly separated from S by an inequality in $I(A_1, A_2)$ if and only if A can be strongly separated from A_1 by a hyperplane. Further, if at least one of A or A_1 is compact, then A and A_1 can be strongly separated if and only if* $\operatorname{cl}\operatorname{conv}(A) \cap \operatorname{cl}\operatorname{conv}(A_1) = \emptyset$.

Proof ($\Leftarrow$) Suppose A can be separated strongly from A_1. In particular, there is (α, δ) and $\varepsilon > 0$ such that $\alpha x \ge \delta + \varepsilon$ for all $x \in A$ and $\alpha x \le \delta$ for all $x \in A_1$. Then, $\alpha x \le \delta$ is clearly valid for $x \in A_2 \cap \{x \mid \alpha x \le \delta\}$. Therefore, $\alpha x \le \delta$ is in $I(A_1, A_2)$ and strongly separates S and A. ($\Rightarrow$) Let P be the hyperplane corresponding to the inequality in $I(A_1, A_2)$ that strongly separates A and S. Assume by contradiction that there does not exist an inequality strongly separating A_1 and A. Let $\alpha x \le \delta$ be the inequality valid for A_1 used in the derivation of P. Now, $\inf\{\alpha x \mid x \in A\} \le \delta$, otherwise A and A_1 can be strongly separated. If A is compact or polyhedral, then $\{x \mid x \in A, \alpha x \le \delta\}$ is non-empty and contains a point, say x'. Then, $x' \in A_2 \cap \{x \mid \alpha x \le \delta\}$. Now, x' must satisfy the inequality corresponding to P and, therefore, A is not strongly separated from S by this inequality.

We now argue that $\text{cl}\,\text{conv}(A) \cap \text{cl}\,\text{conv}(A_1) = \emptyset$ is equivalent to strong separation of A and A_1. A hyperplane strongly separates A and A_1 if and only if it strongly separates $\text{conv}(A)$ and $\text{conv}(A_1)$, since half-spaces are convex. Then, from Theorem 11.4 in [32], it follows that A and A_1 can be strongly separated if and only if $0 \notin \text{cl}(\text{cl}\,\text{conv}(A) - \text{cl}\,\text{conv}(A_1))$. But by Corollary 9.1.1 in [32] and compactness of A or A_1, it follows that $\text{cl}(\text{cl}\,\text{conv}(A) - \text{cl}\,\text{conv}(A_1)) = \text{cl}\,\text{conv}(A) - \text{cl}\,\text{conv}(A_1)$. Or, in other words A and A_1 can be strongly separated if and only if $\text{cl}\,\text{conv}(A) \cap \text{cl}\,\text{conv}(A_1) = \emptyset$. □

Definition 34 Consider an inequality $\alpha x \leq \delta$ that is valid for all $x \in S$. We say that a subset A of $A_2 \backslash S$ provides a non-inclusion certificate of $\alpha x \leq \delta$ in $I(A_1, A_2)$ if there exists an $\varepsilon > 0$ such that $\alpha x \geq \delta + \varepsilon$ for all $x \in A$, A cannot be strongly separated from A_1 by a hyperplane, and at least one of A or A_1 is compact.

The goal of this section is to establish that lifted covers and cliques are not obtained easily for *PB* using integer programming cut-generation techniques. Towards this end, we first linearize B using standard techniques. Second, we show that the resulting linearization is the best possible if one does not use the integrality of the x variables. Finally, we show that aggregation-tightening procedure applied to the linearized problem does not generate the lifted cover and lifted clique inequalities by constructing a non-inclusion certificate for an appropriately defined relaxation of the set.

Consider the bilinear mixed-integer knapsack set B defined in Sect. 3.2. In Proposition 17, we argued that $PB = \text{conv}(B)$ is a polyhedron. In fact, one can easily reformulate B into the following mixed-integer linear set via standard linearization techniques [14]:

$$LB = \left\{ (x, y) \in \{0, 1\}^n \times [0, 1]^n \;\middle|\; \sum_{j \in J} a_j (x_j + y_j - 1) \leq d,\ \forall J \subseteq N \right\}.$$

The correctness of the reformulation follows easily by verifying the equivalence of the bilinear constraint with the linear constraint corresponding to the index set $J = \{j \mid x_j = 1\}$. Now, consider the following relaxation of B

$$RB = \left\{ (x, y) \in [0, 1]^n \times [0, 1]^n \;\middle|\; \sum_{j=1}^{n} a_j x_j y_j \leq d \right\},$$

obtained by ignoring the integrality requirements on the x variables. The defining inequalities of *LB* are valid for *RB* as well. This follows since $(x - 1)(y - 1) \geq 0 \Rightarrow xy \geq x + y - 1$; see McCormick [24]. Therefore, it is natural to first investigate whether additional constraints can be derived without enforcing integrality on the x variables. Let

$$RLB = \left\{ (x, y) \in [0, 1]^n \times [0, 1]^n \;\middle|\; \sum_{j \in J} a_j (x_j + y_j - 1) \leq d,\ \forall J \subseteq N \right\},$$

be the relaxation of *LB* obtained when the integrality restrictions on the x variables are ignored. *RLB* will serve the purpose of A_1 in our aggregation-tightening procedure. We prove next that *RLB* is the best possible linearization of *RB*. Observe that this result does not directly follow from the well-known fact that $\max_{J\subseteq N}\sum_{j\in J} a_j(x_j+y_j-1)$ is the convex envelope of $\sum_{j=1}^n a_j x_j y_j$ over the unit hypercube since in general $\text{conv}\{x \mid f(x)\le d\} \neq \{x \mid \text{conv}\, f(x)\le d\}$.

Proposition 35 $RLB = \text{conv}(RB)$.

Proof Because *RLB* is a convex relaxation of *RB*, we conclude that $RLB \supseteq \text{conv}(RB)$. To prove the opposite inclusion, i.e. $RLB \subseteq \text{conv}(RB)$, we show that every vertex of *RLB* belongs to *RB*. Consider any vertex (x^*, y^*) of *RLB*. There exist coefficients b_i and c_i for $i \in N$ such that (x^*, y^*) is the only optimal solution of the problem

$$z^* = \max\left\{\sum_{i\in N}(b_i x_i + c_i y_i)\;\middle|\;(x,y)\in RLB\right\}.$$

First, note that we may assume that $b_i \ge c_i$. Otherwise, we interchange y_i and x_i. Second, define $N^{++} = \{i \in N \mid c_i \ge 0\}$, $N^{--} = \{i \in N \mid b_i \le 0\}$ and $N^{+-} = N\backslash(N^{++}\cup N^{--})$. We assume that $N^{++} = \{1,\ldots,p\}$, $N^{+-} = \{p+1,\ldots,q\}$ and $N^{--} = \{q+1,\ldots,n\}$. Within each one of these sets, we assume without loss of generality that the subscripts are ordered in such a way that $\frac{c_i}{a_i}$ forms a nonincreasing sequence. Finally, we define $j = \max\left\{r \in N^{++} \mid \sum_{i=1}^r a_i \le d\right\}$.

It is easily verified that the point (x^*, y^*, w^*) where $w_i^* = \max\{x_i^* + y_i^* - 1, 0\}$ for $i \in N$ is an optimal solution to

$$\begin{aligned}
\tilde{z} = \max &\sum_{i\in N}(b_i x_i + c_i y_i)\\
s.t. &\sum_{i\in N} a_i w_i \le d \quad (\pi_0)\\
&x_i + y_i - 1 \le w_i \quad \forall i\in N \quad (\pi_i)\\
&x_i \le 1 \quad \forall i\in N \quad (\rho_i)\\
&y_i \le 1 \quad \forall i\in N \quad (\sigma_i)\\
&x_i, y_i, w_i \ge 0 \quad \forall i\in N
\end{aligned} \tag{45}$$

and that $z^* = \tilde{z}$. In the above linear program, the dual variables for each of the constraints are listed in parenthesis. We now derive an upper bound on z^* by considering the following dual feasible solution

$$\pi_i = \begin{cases} \frac{c_j}{a_j} & \text{if } i = 0\\ \frac{c_j}{a_j}a_i & \text{if } i\in\{1,\ldots,j\}\\ c_i & \text{if } i\in\{j+1,\ldots,q\}\\ 0 & \text{otherwise,}\end{cases}$$

$$\rho_i = \begin{cases} b_i - \frac{c_j}{a_j} a_i & \text{if } i \in \{1, \ldots, j\} \\ b_i - c_i & \text{if } i \in \{j+1, \ldots, q\} \\ 0 & \text{otherwise,} \end{cases}$$

$$\sigma_i = \begin{cases} c_i - \frac{c_j}{a_j} a_i & \text{if } i \in \{1, \ldots, j\} \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, the ordering of $\frac{c_i}{a_i}$ helps here since we know that for $i = 1, \ldots, j$, $\frac{c_j}{a_j} a_i \le c_i \le b_i$. Therefore, the coefficients in the objective are underestimated. We will use $x \le 1$ and $y \le 1$ for the remaining part. Further, for $i = j+1, \ldots, n$ we know that $c_i \le \frac{c_j}{a_j} a_i$. Therefore, the coefficients of w_i in (45) are underestimated. We use $w \ge 0$ for the remaining part. Adding the constraints and appropriately using $w \ge 0$, $x \le 1$ and $y \le 1$, we obtain the following upper bound

$$z^{\text{UB}} = \sum_{i=1}^{j-1} (b_i + c_i) + \sum_{i=j}^{q} b_i + c_j \frac{\left(d - \sum_{i=1}^{j-1} a_i\right)}{a_j}.$$

Now consider the following problem

$$\max \left\{ \sum_{i=1}^{n} (b_i x_i + c_i y_i) \;\middle|\; (x, y) \in RB \right\}. \tag{46}$$

The following solution

$$\left.\begin{array}{ll} (x_i, y_i) = (1, 1) & \text{for } \ i = 1, \ldots, j-1 \\ (x_j, y_j) = \left(1, \left(d - \sum_{i=1}^{j-1} a_i\right) / a_j\right) & \\ (x_i, y_i) = (1, 0) & \text{for } \ i = j+1, \ldots, q \\ (x_i, y_i) = (0, 0) & \text{for } \ i = q+1, \ldots, n. \end{array}\right\} \tag{47}$$

is feasible for (46). Note, incidentally, that this solution is optimal for (46). This is because, assuming $b_i \ge c_i$, it can be argued that $x_i \ge y_i$ for an optimal solution. Further, $(x_i + \epsilon)(y_i - \epsilon) = x_i y_i - \epsilon(x_i - y_i) - \epsilon^2 < x_i y_i$ yields that $(x_i + \epsilon, y_i - \epsilon)$ is feasible with an improved objective function value. We can iteratively apply the above argument until either x_i is 1 or y_i is 0. But, if y_i is 0, then x_i can be made 1. Now, since x_i's are binary, we are left with a continuous knapsack problem in y variables. Therefore, only one component of y is fractional. The remaining follows from the non-increasing order of $\frac{c_i}{a_i}$ in N^{++}.

The objective value associated with (47) matches z^{UB}, which is larger or equal to z^*. However, since (x^*, y^*) is feasible for RB, it is feasible for RLB and because its value is larger or equal to z^* then it is optimal for RLB, i.e., the vertex (x^*, y^*) of RLB belongs to RB. This completes the proof. □

The above argument can be generalized in a straightforward manner to find the convex hull of sets defined via a multilinear inequality as long as each variable appears in only one term. Although the defining inequality of *RB* is nonlinear, Proposition 35 shows that conv(*RB*) is a polytope. In fact, *RLB* is a tractable reformulation for *RB* since it can be expressed as a projection of polynomially many linear inequalities by introducing z_j for $x_j y_j$ and then linearizing as in the proof of Proposition 35. Therefore, it is reasonable to say that the difficulty of optimizing linear functions over B arises primarily from the integrality of the x variables and that *LB* is a reasonable reformulation for deriving inequalities for B using integer programming techniques. Following this scheme, a valid inequality for *LB* (or equivalently B) is obtained by tightening a valid inequality for *RLB* using integrality of the x variables and the bounds on the x and y variables. One may then conjecture that the lifted inequalities of Sect. 3.2 can be found via this reformulation using lifting or other cutting plane techniques in mixed-integer programming. In the remaining part of this section, we show that such a procedure will not yield the desired outcome. This result demonstrates that lifting techniques become even more powerful when generalized to nonlinear programs not only by offering new capabilities for nonlinear problems, but also by exposing new inequalities for integer programming problems with alternate nonlinear formulations. From here onwards, A_2 will refer to $\{0,1\}^n \times [0,1]^n$ and is intended to capture the effects of tightening inequalities valid for *RLB*.

Proposition 36 *Not all lifted covers and lifted clique inequalities are included in the inequalities $I(RLB, A_2)$, where $A_2 = \{0,1\}^n \times [0,1]^n$.*

Proof Consider

$$B = \left\{ (x, y) \in \{0,1\}^3 \times [0,1]^3 \;\middle|\; \sum_{j=1}^{3} a_j x_j y_j \le a_2 + a_3 - \mu \right\}$$

where (i) $a_1 > a_2 \ge a_3 > \mu > 0$, (ii) $a_1 < a_2 + a_3 - \mu$. The set $C = \{2, 3\}$ is a cover and therefore, from Theorem 30, the lifted cover inequality

$$\eta(x,y) = x_1 + x_2 + x_3 + \frac{a_1}{\mu + a_1 - a_2}(y_1 - 1) + \frac{a_2}{\mu}(y_2 - 1) + \frac{a_3}{\mu}(y_3 - 1) - 1 \le 0 \tag{48}$$

is facet-defining for *PB*. The above inequality is also a lifted clique inequality since our assumptions imply that $\{1, 2, 3\}$ is a clique. Now, consider the following points:

1. $p = (x_1 = 1, x_2 = 0, x_3 = 1, y_1 = \frac{a_2 - \mu}{a_1} + \epsilon, y_2 = 1, y_3 = 1)$
2. $q = (x_1 = 0, x_2 = 1, x_3 = 1, y_1 = \frac{a_2 - \mu}{a_1} + \epsilon, y_2 = 1, y_3 = 1)$

First note that $\eta(p) = \eta(q) = \frac{\epsilon a_1}{\mu + a_1 - a_2} > 0$ as long as $\epsilon > 0$. Define $e_I(x, y) = \sum_{i \in I} a_i(x_i + y_i - 1) - a_2 - a_3 + \mu$ where $I \subseteq \{1, 2, 3\}$. Then

$$RLB = \{(x, y) \mid e_I(x, y) \le 0, I \subseteq N, 0 \le x_i \le 1, 0 \le y_i \le 1\}.$$

Now, define $h(x, y)$ to be the vector of violations of constraints $e_I(x, y) \leq 0$ for $I \subseteq \{1, 2, 3\}$, i.e.,

$$h(x, y) = [e_{\{1\}}(x, y), e_{\{2\}}(x, y), e_{\{3\}}(x, y), e_{\{1,2\}}(x, y), e_{\{1,3\}}(x, y), e_{\{2,3\}}(x, y), e_{\{1,2,3\}}(x, y)].$$

Then, by direct calculation, we obtain that

$$\begin{aligned} h(p) &= [a_1\epsilon - a_3, \mu - a_2 - a_3, \mu - a_2, a_1\epsilon - a_3, a_1\epsilon, \mu - a_2, \epsilon a_1] \\ h(q) &= [a_1(\epsilon - 1) - a_3, \mu - a_3, \mu - a_2, \\ &\quad a_1(\epsilon - 1) + (a_2 - a_3), a_1(\epsilon - 1), \mu, a_2 + a_1(\epsilon - 1)]\,. \end{aligned}$$

Consider $r = \lambda p + (1 - \lambda)q$ and set $\epsilon = (1 - \lambda)\frac{a_1 - a_2}{a_1}$. Then,

$$h(r) = [-a_2(1 - \lambda) - a_3, -\lambda a_2 - (a_3 - \mu), \mu - a_2, -a_3, -a_2(1 - \lambda), -\lambda a_2 + \mu, 0]\,.$$

Note that $0 \leq r \leq 1$ whenever $0 \leq \lambda \leq 1$. Therefore, as long as $1 > \lambda \geq \frac{\mu}{a_2}$, r is feasible to A_2 and ϵ is greater than zero (i.e., r is infeasible to the lifted cover). Since p and q are feasible to *RLB*, by Theorem 33, $p \cup q$ provides a certificate of the non-inclusion of the lifted clique and lifted cover inequality (48) in $I(RLB, A_2)$. □

The aggregation-tightening procedure includes as special cases many inequalities for integer programming. Clearly $I(RLB, A_2)$ subsumes lifted cover inequalities generated from a single row relaxation of LB and rank-one fractional Gomory cuts for LB. Theorem 30 and our discussion in the last paragraph give evidence that the lifted cover inequalities, which were obtained in closed form using superadditive lifting for the nonlinear formulation of the bilinear knapsack set, are extremely hard to obtain using standard integer programming cut-generation procedures. This is a clear motivation to further develop cutting plane procedures for nonlinear programming problems. As a side note, the proofs of Proposition 36 and Theorem 39 provide an example of a six-dimensional mixed-integer polytope for which the intersection of the convex hull of all knapsack constraints (even those obtained via aggregation of inequalities) is not sufficient to obtain the convex hull of the integer program. Next, we use the construction of Proposition 36 to show that the lifted cover inequality (48) cannot even be obtained as a rank-one split cut of LB.

Proposition 37 *Not all lifted covers and lifted clique inequalities are rank-one split cuts for LB.*

Proof Consider the set B in the proof of Proposition 36 and the corresponding lifted cover inequality (48). Define the set

$$V = \left\{ (x, y) \in [0, 1]^3 \times [0, 1]^3 \;\middle|\; 1 < \sum_{i=1}^{3} x_i \leq \frac{d}{a_1}, y_i = 1\, \forall i \right\}.$$

Observe that $V \subseteq RB \subseteq RLB$ (see Proposition 35). Further, all the points in V violate (48) and since $d > a_1$, $V \neq \emptyset$. We assume for a proof by contradiction that the lifted cover inequality is a rank-one split cut. The corresponding split disjunction is induced by $(\pi, \pi_0) \in \mathbb{Z}^4$, and the split cut is derived as a valid inequality for $RLB_1 \cup RLB_2$ where $RLB_1 = RLB \cap \{(x, y) \mid \pi x \leq \pi_0\}$ and $RLB_2 = RLB \cap \{(x, y) \mid \pi x \geq \pi_0 + 1\}$. Since the points in V do not satisfy (48),

$$(RLB_1 \cup RLB_2) \cap V = \emptyset, \tag{49}$$

otherwise, we find a contradiction to our assumption.

Consider the points e_1, e_2 and e_3 in the space of the x variables. We assume without loss of generality, and by invoking the pigeon-hole principle, that two of these points, say e_i and e_j, $i \neq j$, satisfy $\pi x \leq \pi_0$. Assume further that $\pi e_k \geq \pi_0 + 1$, where $k \neq i$ and $k \neq j$. If $\pi(e_i + e_j) \leq \pi_0$, then for $\lambda \in [0, 1]$,

$$(1 - \lambda)\pi e_i + \lambda\pi(e_i + e_j) = \pi e_i + \lambda\pi e_j = \pi(e_i + \lambda e_j) \leq \pi_0.$$

For λ sufficiently small, $(e_i + \lambda e_j, 1) \in V$, yielding a contradiction to (49). If $\pi(e_i + e_j) \geq \pi_0 + 1$, then for $\lambda \in [0, 1]$,

$$(1 - \lambda)\pi e_k + \lambda\pi(e_i + e_j) = \pi\left((1 - \lambda)e_k + \lambda e_i + \lambda e_j\right) \geq \pi_0 + 1.$$

For λ sufficiently small, $((1-\lambda)e_k + \lambda e_i + \lambda e_j, 1) \in V$, again yielding a contradiction to (49). Therefore, we may assume that $\pi e_k \leq \pi_0$. We claim that $\pi e_k = \pi e_i = \pi e_j = \pi_0$. Otherwise, if $\pi e_t < \pi_0$ for $t \in \{i, j, k\}$, then $\pi(e_t + \lambda e_r) < \pi_0$ for $r \neq t$ and a sufficiently small $\lambda > 0$, which is a contradiction to (49). It follows that $\pi = \pi_0 1$, where 1 is a vector of all ones. However, for $\lambda > 0$ and sufficiently small, $e_1 + \lambda e_2$ does not satisfy $\pi x \leq \pi_0$. Therefore, $\pi_0(1 + \lambda) > \pi_0$, or, in other words, $\pi_0 > 0$. Consider the points p, q, and r defined in proof of Proposition 36. Because $\pi p = \pi q = 2\pi_0 \geq \pi_0 + 1$, r belongs to RLB_2 and violates (48). □

Remark 38 We reinterpret the proof of Proposition 37 to allow for more general two-term disjunctions. Define $L = [0, 1]^3 \cap \left\{x \mid \sum_{i=1}^{3} x_i \leq 1\right\}$, $R(k) = [0, 1]^3 \cap \left\{x \mid \sum_{i=1}^{3} x_i \geq k\right\}$, $T_1 = [0, 1]^3 \cap \{x \mid \pi x \leq \pi_0\}$, and $T_2 = [0, 1]^3 \cap \left\{x \mid \pi x \geq \pi_0 + 1\right\}$. More generally, T_1 and T_2 are any convex subsets of $[0, 1]^3$ such that $T_1 \cup T_2 \supseteq \{0, 1\}^3$. Since L and $R\left(\frac{d}{a_1}\right)$ are separated by a non-zero distance, if $T_1 \cap L \neq \emptyset$ and $T_1 \cap R\left(\frac{d}{a_1}\right) \neq \emptyset$, then either T_1 is not connected or RLB_1 contains a point in V, contradicting respectively the convexity of T_1 or (49). Therefore, without loss of generality, we assume that $T_1 \subseteq L$ and, similarly, $T_2 \subseteq R\left(\frac{d}{a_1}\right)$. However, $T_1 \supseteq L$ since T_1 contains e_1, e_2, e_3, and 0. Similarly, T_2 must contain the remaining corner points of $[0, 1]^3$ and, therefore, $R(2) \subseteq T_2$. Consider the points p, q and r defined in proof of Proposition 36. Since p and q satisfy $\sum_{i=1}^{3} x_i \geq 2$, it follows that $r \in RLB \cap \{(x, y) \mid x \in R(2)\} \subseteq RLB \cap \{(x, y) \mid x \in T_2\} = RLB_2$. Moreover, r violates (48), yielding the desired contradiction. □

Even though Proposition 36 shows that a single inequality in $I(RLB, A_2)$ is not capable of dominating the lifted cover inequality of Theorem 30, it does not immediately show that the lifted cover inequality cannot be obtained by aggregating inequalities in $I(RLB, A_2)$. Next, we show strong separation of the lifted cover inequality from the inequalities obtained by aggregation of inequalities in $I(RLB, A_2)$.

Theorem 39 *Lifted cover and lifted clique inequalities are not implied by the intersection of all the inequalities in $I(RLB, A_2)$, where $A_2 = \{0, 1\}^n \times [0, 1]^n$.*

Proof Let B, p and q be as in the proof of Proposition 36. Let I_1 be the set of facet-defining inequalities for $\text{conv}(B) \cup p$ and I_2 be the set of facet-defining inequalities for $\text{conv}(B) \cup q$. Clearly, inequalities in $I(A_1, A_2)$ can be obtained from conic combinations of inequalities in $I_1 \cup I_2$. Further, we rescale the inequalities of the type $\alpha x \le \delta$ in $I_1 \cup I_2$ to satisfy $\|(\alpha, \delta)\| = 1$. By Theorem 30 and Proposition 20, we know that the lifted cover/lifted clique inequality is facet-defining. Then, let $\{u_1, \ldots, u_{n+1}\}$ be a set of $n + 1$ affinely independent points feasible to B that satisfy the lifted cover/lifted clique inequality at equality. Define:

$$z = \max_{(\alpha,\delta)} \left\{ v(\alpha, \delta) \;\middle|\; \|(\alpha, \delta)\| = 1, (\alpha, \delta) \in I_1 \cup I_2 \right\},$$

where $v(\alpha, \delta) = \sum_{i=1}^{n+1} (\alpha u_i - \delta)$. First, observe that $v(\alpha, \delta) \le 0$, since $\alpha x \le \delta$ is satisfied for all u_i. In fact, $v(\alpha, \delta) < 0$, otherwise $\alpha u_i - \delta = 0$ for all u_i. Then, (α, δ) is uniquely determined (up to scaling) and must, therefore, correspond to the lifted cover inequality. This yields the desired contradiction since the lifted cover cuts off p and q. Now, since $\text{conv}(B)$ is a full-dimensional polytope by Proposition 17 and $I_1 \cup I_2$ is finite, it follows that $z < 0$. Let $(\alpha^1, \delta^1), \ldots, (\alpha^t, \delta^t)$ be the inequalities in $I_1 \cup I_2$. Define

$$\mathcal{V} = \begin{pmatrix} \alpha^1 & \delta^1 \\ \vdots & \vdots \\ \alpha_t & \delta_t \end{pmatrix}, w = (w_1, \ldots, w_t), e = (1, \ldots, 1)^T, \text{ and } Y = \begin{pmatrix} u_1 & \ldots & u_{n+1} \\ -1 & \ldots & -1 \end{pmatrix}.$$

where $\sum_{i=1}^{t} w_i = 1$. By our previous argument $\mathcal{V}Ye \le ze$. Therefore, $w\mathcal{V}Ye \le zwe = z < 0$. On the other hand, for the lifted cover inequality $\gamma x \le \varsigma$, $[\gamma \;\, \varsigma]Ye = 0$. Therefore, Ye separates $w\mathcal{V}$ (the set of inequalities obtained by aggregating inequalities in $I(RLB, A_2)$) from $[\gamma, \varsigma]$. □

5 Conclusion

In this paper, we have extended mixed-integer programming lifting techniques to nonlinear mixed-integer programming. In particular, we have interpreted lifting geometrically and have generalized the concept of sequence-independent lifting to nonlinear programs. For mixed-integer bilinear knapsack sets, we have used these techniques to obtain an exponential family of inequalities that are not easily obtained from linear integer programming reformulations. In the process, we have developed a unified approach for demonstrating strength of inequalities with respect to typical elementary closures studied in integer programming. Although originally designed for the generation of linear cuts, the theory we have developed can be used to generate nonlinear cuts and convex hulls in the space of original variables. Future research will concentrate on the computational aspects of generating lifted cuts on a larger selection of nonlinear problems. The lifting theory has the potential of yielding new inequalities for various models that have been previously studied, such as complementarity and quadratic problems.

Appendix

$$19x_1y_1 + 17x_2y_2 + 15x_3y_3 + 10x_4y_4 \leq 20$$

The linear description of the convex hull of the above set is as follows:

$$
\begin{array}{lllllllll}
-1x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 0\\
0x_1 & -\,1x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 0\\
0x_1 & +\,0x_2 & -\,1x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 0\\
0x_1 & +\,0x_2 & +\,0x_3 & -\,1x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 0\\
0x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & -\,1y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 0\\
0x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & -\,1y_2 & +\,0y_3 & +\,0y_4 & \leq 0\\
0x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & -\,1y_3 & +\,0y_4 & \leq 0\\
0x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & -\,1y_4 & \leq 0\\
0x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & +\,1y_4 & \leq 1\\
0x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & +\,1y_3 & +\,0y_4 & \leq 1\\
0x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & +\,1y_2 & +\,0y_3 & +\,0y_4 & \leq 1\\
0x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & +\,1y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 1\\
0x_1 & +\,0x_2 & +\,0x_3 & +\,1x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 1\\
0x_1 & +\,0x_2 & +\,1x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 1\\
0x_1 & +\,0x_2 & +\,1x_3 & +\,1x_4 & +\,0y_1 & +\,0y_2 & +\,3y_3 & +\,2y_4 & \leq 6\\
0x_1 & +\,1x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 1\\
0x_1 & +\,7x_2 & +\,0x_3 & +\,7x_4 & +\,0y_1 & +\,17y_2 & +\,0y_3 & +\,10y_4 & \leq 34\\
0x_1 & +\,7x_2 & +\,5x_3 & +\,7x_4 & +\,0y_1 & +\,17y_2 & +\,15y_3 & +\,10y_4 & \leq 49\\
0x_1 & +\,7x_2 & +\,7x_3 & +\,7x_4 & +\,0y_1 & +\,17y_2 & +\,21y_3 & +\,14y_4 & \leq 59\\
0x_1 & +\,12x_2 & +\,12x_3 & +\,0x_4 & +\,0y_1 & +\,17y_2 & +\,15y_3 & +\,0y_4 & \leq 44\\
0x_1 & +\,12x_2 & +\,12x_3 & +\,5x_4 & +\,0y_1 & +\,17y_2 & +\,15y_3 & +\,10y_4 & \leq 54\\
0x_1 & +\,12x_2 & +\,12x_3 & +\,7x_4 & +\,0y_1 & +\,17y_2 & +\,21y_3 & +\,14y_4 & \leq 64\\
0x_1 & +\,14x_2 & +\,12x_3 & +\,7x_4 & +\,0y_1 & +\,17y_2 & +\,15y_3 & +\,10y_4 & \leq 56\\
0x_1 & +\,17x_2 & +\,15x_3 & +\,10x_4 & +\,0y_1 & +\,17y_2 & +\,15y_3 & +\,10y_4 & \leq 62\\
1x_1 & +\,0x_2 & +\,0x_3 & +\,0x_4 & +\,0y_1 & +\,0y_2 & +\,0y_3 & +\,0y_4 & \leq 1\\
9x_1 & +\,0x_2 & +\,0x_3 & +\,9x_4 & +\,19y_1 & +\,0y_2 & +\,0y_3 & +\,10y_4 & \leq 38\\
9x_1 & +\,0x_2 & +\,5x_3 & +\,9x_4 & +\,19y_1 & +\,0y_2 & +\,15y_3 & +\,10y_4 & \leq 53\\
9x_1 & +\,0x_2 & +\,9x_3 & +\,9x_4 & +\,19y_1 & +\,0y_2 & +\,27y_3 & +\,18y_4 & \leq 73\\
9x_1 & +\,7x_2 & +\,0x_3 & +\,9x_4 & +\,19y_1 & +\,17y_2 & +\,0y_3 & +\,10y_4 & \leq 55\\
9x_1 & +\,7x_2 & +\,5x_3 & +\,9x_4 & +\,19y_1 & +\,17y_2 & +\,15y_3 & +\,10y_4 & \leq 70
\end{array}
$$

$$
\begin{array}{lllllllll}
14x_1 & + 0x_2 & + 14x_3 & + 0x_4 & + 19y_1 & + 0y_2 & + 15y_3 & + 0y_4 & \le 48 \\
14x_1 & + 0x_2 & + 14x_3 & + 5x_4 & + 19y_1 & + 0y_2 & + 15y_3 & + 10y_4 & \le 58 \\
14x_1 & + 0x_2 & + 14x_3 & + 9x_4 & + 19y_1 & + 0y_2 & + 27y_3 & + 18y_4 & \le 78 \\
14x_1 & + 12x_2 & + 14x_3 & + 0x_4 & + 19y_1 & + 17y_2 & + 15y_3 & + 0y_4 & \le 65 \\
14x_1 & + 12x_2 & + 14x_3 & + 5x_4 & + 19y_1 & + 17y_2 & + 15y_3 & + 10y_4 & \le 75 \\
14x_1 & + 12x_2 & + 14x_3 & + 9x_4 & + 19y_1 & + 17y_2 & + 27y_3 & + 18y_4 & \le 95 \\
16x_1 & + 16x_2 & + 0x_3 & + 0x_4 & + 19y_1 & + 17y_2 & + 0y_3 & + 0y_4 & \le 52 \\
16x_1 & + 16x_2 & + 0x_3 & + 7x_4 & + 19y_1 & + 17y_2 & + 0y_3 & + 10y_4 & \le 62 \\
16x_1 & + 16x_2 & + 12x_3 & + 0x_4 & + 19y_1 & + 17y_2 & + 15y_3 & + 0y_4 & \le 67 \\
16x_1 & + 16x_2 & + 12x_3 & + 7x_4 & + 19y_1 & + 17y_2 & + 15y_3 & + 10y_4 & \le 77 \\
18x_1 & + 0x_2 & + 14x_3 & + 9x_4 & + 19y_1 & + 0y_2 & + 15y_3 & + 10y_4 & \le 62 \\
18x_1 & + 16x_2 & + 0x_3 & + 9x_4 & + 19y_1 & + 17y_2 & + 0y_3 & + 10y_4 & \le 64 \\
18x_1 & + 16x_2 & + 14x_3 & + 0x_4 & + 19y_1 & + 17y_2 & + 15y_3 & + 0y_4 & \le 69 \\
18x_1 & + 16x_2 & + 14x_3 & + 9x_4 & + 19y_1 & + 17y_2 & + 15y_3 & + 10y_4 & \le 79 \\
19x_1 & + 0x_2 & + 15x_3 & + 10x_4 & + 19y_1 & + 0y_2 & + 15y_3 & + 10y_4 & \le 64 \\
19x_1 & + 17x_2 & + 0x_3 & + 10x_4 & + 19y_1 & + 17y_2 & + 0y_3 & + 10y_4 & \le 66 \\
19x_1 & + 17x_2 & + 15x_3 & + 0x_4 & + 19y_1 & + 17y_2 & + 15y_3 & + 0y_4 & \le 71 \\
19x_1 & + 17x_2 & + 15x_3 & + 10x_4 & + 19y_1 & + 17y_2 & + 15y_3 & + 10y_4 & \le 81 \\
63x_1 & + 63x_2 & + 0x_3 & + 63x_4 & + 133y_1 & + 153y_2 & + 0y_3 & + 90y_4 & \le 439 \\
63x_1 & + 63x_2 & + 45x_3 & + 63x_4 & + 133y_1 & + 153y_2 & + 135y_3 & + 90y_4 & \le 574 \\
63x_1 & + 63x_2 & + 63x_3 & + 63x_4 & + 133y_1 & + 153y_2 & + 189y_3 & + 126y_4 & \le 664 \\
84x_1 & + 84x_2 & + 84x_3 & + 0x_4 & + 114y_1 & + 119y_2 & + 105y_3 & + 0y_4 & \le 422 \\
84x_1 & + 84x_2 & + 84x_3 & + 35x_4 & + 114y_1 & + 119y_2 & + 105y_3 & + 70y_4 & \le 492 \\
84x_1 & + 84x_2 & + 84x_3 & + 49x_4 & + 114y_1 & + 119y_2 & + 147y_3 & + 98y_4 & \le 562 \\
96x_1 & + 96x_2 & + 84x_3 & + 0x_4 & + 114y_1 & + 119y_2 & + 105y_3 & + 0y_4 & \le 434 \\
96x_1 & + 96x_2 & + 84x_3 & + 42x_4 & + 114y_1 & + 119y_2 & + 105y_3 & + 60y_4 & \le 494 \\
96x_1 & + 96x_2 & + 84x_3 & + 47x_4 & + 114y_1 & + 119y_2 & + 105y_3 & + 70y_4 & \le 504 \\
96x_1 & + 96x_2 & + 84x_3 & + 49x_4 & + 114y_1 & + 119y_2 & + 111y_3 & + 74y_4 & \le 514 \\
98x_1 & + 98x_2 & + 98x_3 & + 63x_4 & + 133y_1 & + 153y_2 & + 189y_3 & + 126y_4 & \le 699 \\
112x_1 & + 112x_2 & + 0x_3 & + 63x_4 & + 133y_1 & + 153y_2 & + 0y_3 & + 90y_4 & \le 488 \\
112x_1 & + 112x_2 & + 84x_3 & + 63x_4 & + 133y_1 & + 153y_2 & + 105y_3 & + 90y_4 & \le 593 \\
112x_1 & + 112x_2 & + 94x_3 & + 63x_4 & + 133y_1 & + 153y_2 & + 135y_3 & + 90y_4 & \le 623 \\
112x_1 & + 112x_2 & + 96x_3 & + 56x_4 & + 133y_1 & + 136y_2 & + 120y_3 & + 80y_4 & \le 581 \\
112x_1 & + 112x_2 & + 98x_3 & + 63x_4 & + 133y_1 & + 153y_2 & + 147y_3 & + 98y_4 & \le 643
\end{array}
$$

References

1. Atamtürk, A.: Flow pack facets of the single node fixed-charge flow polytope. Oper. Res. Lett. **29**, 107–114 (2001)
2. Atamtürk, A.: On the facets of the mixed-integer knapsack polyhedron. Math. Program. **98**, 145–175 (2003)
3. Balas, E.: Facets of the knapsack polytope. Math. Program. **8**, 146–164 (1975)
4. Balas, E.: Disjunctive programming and hierarchy of relaxations for discrete optimization problems. SIAM J. Alg. Disc. Meth. **6**, 466–486 (1985)
5. Balas, E.: Disjunctive programming: properties of the convex hull of feasible points. Discrete Appl. Math. **89**(1-3), 3–44, (1998). Original manuscript was published as a technical report in (1974)
6. Balas, E., Zemel, E.: Lifting and complementing yields all the facets of positive zero-one programming polytopes. Math. Program. **9**, 13–24 (1984)

7. Bixby, R.E., Fenelon, M., Gu, Z., Rothberg, E., Wunderling, R.: MIP: theory and practice-closing the gap. In: Powell, M.J.D, Scholtes, S. (eds.) System Modelling and Optimization: Methods, Theory and Applications., pp. 19–49. Kluwer, Dordrecht (2000)
8. Bixby, R.E., Gu, Z., Rothberg, E., Wunderling, R.: Mixed integer programming: a progress report. In: Grötschel, M. (ed.) The Sharpest Cut: the Impact of Manfred Padberg and His Work, MPS/SIAM Series on Optimization, pp. 309–326. SIAM, Philadelphia (2004)
9. Ceria, S., Soares, J.: Convex programming for disjunctive convex optimization. Math. Program. **86**, 595–614 (1999)
10. Chaovalitwongse, W., Pardalos, P.M., Prokopyev, O.A.: A new linearization technique for multi-quadratic 0–1 programming problems. Oper. Res. Lett. **32**, 517–522 (2004)
11. Christof, T., Löbel, A.: PORTA : A POlyhedron Representation Transformation Algorithm (1997). http://www.zib.de/Optimization/Software/Porta/
12. Crowder, H.P., Johnson, E.L., Padberg, M.W.: Solving large-scale zero-one linear programming problems. Oper. Res. **31**, 803–834 (1983)
13. de Farias, I.R., Johnson, E.L., Nemhauser, G.L.: Facets of the complementarity knapsack polytope. Math. Oper. Res. **27**, 210–226 (2002)
14. Glover, F., Woolsey, E.: Converting a $0-1$ polynomial programming problem to a $0-1$ linear program. Oper. Res. **22**, 180–182 (1974)
15. Gu, Z., Nemhauser, G.L., Savelsbergh, M.W.P.: Lifted flow cover inequalities for mixed 0–1 integer programs. Math. Program. **85**, 439–468 (1999)
16. Gu, Z., Nemhauser, G.L., Savelsbergh, M.W.P.: Sequence independent lifting. J. Comb. Optim. **4**, 109–129 (2000)
17. Hammer, P.L., Johnson, E.L., Peled, U.N.: Facets of regular 0–1 polytopes. Math. Program. **8**, 179–206 (1975)
18. Hiriart-Urruty, J.-B., Lemaréchal, C.: Fundamentals of Convex Analysis. Springer, Heidelberg (2004)
19. Laundy, R., Perregaard, M., Tavares, G., Tipi, H., Vazacopoulos, A.: Solving hard mixed integer programming problems with Xpress-MP: A MIPLIB 2003 case study. Technical Report RRR 2-2007, Rutgers Center for Operations Research, Rutgers University, Piscataway (2007)
20. Lin, T.-C., Vandenbussche, D.: Boxed-constrained quadratic programs with fixed charge variables. J. Global Optim. (forthcoming)
21. Louveaux, Q., Wolsey, L.A.: Lifting, superadditivity, mixed integer rounding and single node flow sets revisited. Ann. Oper. Res. **153**, 47–77 (2007)
22. Luenberger, D.G.: Optimization by Vector Space Methods. Wiley, New York (1969)
23. Marchand, H., Wolsey, L.A.: The 0–1 knapsack problem with a single continuous variable. Math. Program. **85**, 15–33 (1999)
24. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: Part I—Convex underestimating problems. Math. Program. **10**, 147–175 (1976)
25. Nemhauser, G.L., Wolsey, L.A.: Integer and Combinatorial Optimization. Wiley Interscience Series in Discrete Mathematics and Optimization. Wiley, New York (1988)
26. Nemhauser, G.L., Wolsey, L.A.: A recursive procedure for generating all cuts for 0–1 mixed integer programs. Math. Program. **46**, 379–390 (1990)
27. Padberg, M.W.: On the facial structure of set packing polyhedra. Math. Program. **5**, 199–215 (1973)
28. Padberg, M.W., Van Roy, T.J., Wolsey, L.A.: Valid linear inequalities for fixed charge problems. Oper. Res. **32**, 842–861 (1984)
29. Richard, J.-P.P.: Lifted Inequalities for 0–1 Mixed Integer Programming. PhD thesis, Georgia Institute of Technology (2002)
30. Richard, J.-P.P., de Farias, I.R., Nemhauser, G.L.: Lifted inequalities for 0–1 mixed integer programming : Basic theory and algorithms. Math. Program. **98**, 89–113 (2003)
31. Richard, J.-P.P., de Farias, I.R., Nemhauser, G.L.: Lifted inequalities for 0–1 mixed integer programming : superlinear lifting. Math. Program. **98**, 115–143 (2003)
32. Rockafellar, R.T.: Convex Analysis. Princeton Mathematical Series. Princeton University Press, Princeton (1970)
33. Shebalov, S., Klabjan, D.: Sequence independent lifting for mixed integer programs with variable upper bounds. Math. Program. **105**, 523–561 (2006)
34. Sherali, H.D., Adams, W.P.: A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. SIAM J. Discrete Math. **3**, 411–430 (1990)

35. Sherali, H.D., Adams, W.P.: A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems, vol. 31 of Nonconvex Optimization and its Applications. Kluwer, Dordrecht (1999)
36. Sherali, H.D., Alameddine, A.: A new reformulation-linearization technique for bilinear programming problems. J. Global Optim. **2**, 379–410 (1992)
37. Stubbs, R., Mehrotra, S.: A branch and cut method for 0–1 mixed convex programming. Math. Program. **86**, 515–532 (1999)
38. Tawarmalani, M., Sahinidis, N.V.: Convex extensions and convex envelopes of l.s.c. functions. Math. Program. **93**, 247–263 (2002)
39. Van Roy, T.J., Wolsey, L.A.: Valid inequalities and separation for uncapacitated fixed charge networks. Oper. Res. Lett. **4**, 105–112 (1985)
40. Vandenbussche, D., Nemhauser, G.L.: A polyhedral study of nonconvex quadratic programs with box constraints. Math. Program. **102**, 531–557 (2005)
41. Wolsey, L.A.: Faces for a linear inequality in 0–1 variables. Math. Program. **8**, 165–178 (1975)
42. Wolsey, L.A.: Facets and strong valid inequalities for integer programs. Oper. Res. **24**, 367–372 (1976)
43. Wolsey, L.A.: Valid inequalities and supperadditivity for 0–1 integer programs. Math. Oper. Res. **2**, 66–77 (1977)