Locally ideal formulations for piecewise linear functions with indicator variables

Srikrishna Sridhar a, Jeff Linderoth b, James Luedtke b,∗

a School of Computer Sciences, University of Wisconsin-Madison, 1209 Dayton Street, Madison, WI 53706, United States
b Department of Industrial and Systems Engineering, University of Wisconsin-Madison, 1513 University Avenue, Madison, WI 53706, United States

ARTICLE INFO

Article history:
Received 12 April 2013
Received in revised form 19 August 2013
Accepted 20 August 2013
Available online 29 August 2013

Keywords:
Piecewise linear function
Mixed-integer programming
Indicator variables

1. Introduction

Optimization problems involving piecewise linear functions (PLFs) appear in a wide range of applications. PLFs are frequently used to approximate nonlinear functions and to model cost functions involving economies of scale and fixed charges. Problems involving non-convex PLFs are commonly formulated as mixed integer programming (MIP) problems [4,17,8,2,23].

Consider a univariate PLF $f : [B_0, B_1] → \mathbb{R}$ with its domain $[B_0, B_1]$ divided into an increasing sequence of breakpoints $B_0, B_1, \ldots, B_n$. For simplicity, we assume that $f(\cdot)$ is continuous, $B_0 = 0$ and $f(0) = 0$. Our results can be extended to the case when $f(\cdot)$ is lower semi-continuous, $B_0 \neq 0$, and $f(B_0) \neq 0$. The function $f(\cdot)$ can be written as

$$f(x) := mx + c_i, \quad x \in [B_{i-1}, B_i] \forall i \in \{1, \ldots, n\}$$  \hspace{1cm} (1)

where $m_i \in \mathbb{R}$, $c_i \in \mathbb{R}$ and $B_0 < B_1 < \cdots < B_n$.

In this paper, we present MIP formulations for PLFs where setting a binary indicator variable to zero forces the argument of the function $f(\cdot)$ to zero which in turn forces the function to take a zero value. In other words,

$$z = 0 \Rightarrow x = 0, f(x) = 0.$$  \hspace{1cm} (2)

The goal of this work is to present a theoretical and computational comparison of MIP formulations that enforce the logical conditions in (2). Specifically, we examine properties of different formulations of the three variable set

$$X := \bigcup_{i=1}^{n} \{ (x, y, z) : x \in [B_{i-1}, B_i], y = mx + c_i, z = 1 \} \cup \{ (0, 0, 0) \}.$$  \hspace{1cm} (3)

In some applications, notably those where the PLF appears in a minimization objective, the relevant set to study has the variable $y$ constrained to lie in the epigraph of a function. We denote $X^\circ$ as the set where the equality relationship $y = mx + c_i$ in (3) is replaced with $y \geq mx + c_i$.

Methods for modeling PLFs include specially ordered sets of type II (SOS2) [4], the incremental model, or delta method (Delta) [17], the multiple choice model (MCM) [13], the convex combination (CC) model [8], the disaggregated convex combination model (DCC) [19], and approaches that require only logarithmically many binary variables [24]. Table 1 lists several applications in the literature that have modeled PLFs using these well-known methods in conjunction with variable upper bound constraints of the form

$$x \leq B_n z$$  \hspace{1cm} (4)

to enforce the logical on–off condition (2).

In this work, we propose a simple modeling artifice for PLFs that also enforces the logical condition (2), and we demonstrate its desirable theoretical and computational properties. We start by describing the idea using SOS2 to model a PLF as

$$x = \sum_{i=0}^{n} \lambda_i B_i, \quad y = \sum_{i=0}^{n} \lambda_i F_i \quad 1 = \sum_{i=0}^{n} \lambda_i$$  \hspace{1cm} (5)

$$\lambda := \left\{ \lambda_i \in \mathbb{R}_+ : \forall i \in \{0, \ldots, n\} \right\}$$  \hspace{1cm} (6)

is SOS2.
In this formulation, the function $f(\cdot)$ and its argument $x$ are expressed as convex combinations of breakpoints $B := \{B_0, \ldots, B_n\}$ and their corresponding function evaluations $\{f_0, \ldots, f_n\}$ where $F_i := f(B_i) = m_i B_i + c_i$. The formulation introduces a non-negative set of variables $\lambda \in \mathbb{R}_+^{n+1}$ that satisfy the SOS2 property—most two of the variables can be positive, and if two variables are positive then they must be consecutive in the ordered set. Most modern general purpose MIP solvers enforce the SOS2 condition algorithmically by branching [4].

Using variable upper bound constraints (4) to enforce the logical condition (2) has two problems. First, the use of “big M” constraints may considerably weaken the LP relaxation of the MIP formulation. Second, the model introduces an additional constraint $x \leq B_0 z$.

We propose the following simple strengthening that replaces $x \leq B_0 z$ and $\sum_{i=0}^{n} \lambda_i = 1$ with

$$\sum_{i=0}^{n} \lambda_i = z. \quad (6)$$

Setting the binary variable $z = 0$ in (6) forces $\lambda_i = 0 \forall i \in \{0, \ldots, n\}$, which in turn forces the function to take a zero value. If the binary variable $z = 1$, then $\sum_{i=0}^{n} \lambda_i = 1$, which reduces to (5). We show in Section 2.1 that a formulation using (6) has the desirable property of being locally ideal, while one that uses $x \leq B_0 z$ does not.

In Section 2, we also show how to strengthen MIP formulations of $X$ that use the incremental model, the multiple choice model, the convex combination model, the disaggregated convex combination model, and logarithmic models to model the PLF. Therefore, this formulation strengthening technique could be directly applied to all of the applications listed in Table 1. In all cases, we show that our model retains the desirable theoretical property of the underlying PLF modeling method, either idealness or sharpness, but using a variable upper bound constraint $x \leq B_0 z$ destroys the property. Borghetti et al. [5] created a formulation of $X$ that employed the strengthening techniques we describe. They used the convex combination method to model the PLFs which does not have the locally ideal property [23]. In the case that the PLFs are convex, we describe a connection between the formulation strengthening techniques we describe and the perspective reformulation [11]. The Delta, MCM, CC, and DCC MIP formulations can all be extended to model multivariate piecewise linear functions [23]. When the multivariate piecewise linear functions are combined with the structure (2), our strengthened formulation can be similarly applied. We omit the details to simplify exposition. We conclude with a computational study on a practical application to illustrate the benefits of the new formulations. In our experiments, we observed that our formulation computes optimal solutions on average 40 times faster.

### 2. Properties of MIP formulations

Padberg and Rijal [21] define a locally ideal MIP formulation as one where the vertices of its corresponding LP relaxation satisfy all required integrality conditions. Extending this definition, Croxton et al. [7] and Keha et al. [14] define a locally ideal SOS2 formulation as one whose LP relaxation has extreme points that all satisfy the SOS2 property. As shown by Vielma et al. [23], all commonly used MIP formulations of PLFs, except for the original convex combination (CC) model, are known to be locally ideal. In this section, we demonstrate the theoretical strength of the proposed formulations for $X$ that include the logical condition (2).

#### 2.1. SOS2 model

We consider the following two SOS2-based formulations for $X$:

$$S_1 := \{(x, y, \lambda, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{n+1} \times \{0, 1\} :$$

$$x = \sum_{i=0}^{n} B_i \lambda_i, \quad y = \sum_{i=0}^{n} F_i \lambda_i, \quad 1 = \sum_{i=0}^{n} \lambda_i,$$

$$x \leq B_0 z, \quad \lambda \text{ is SOS2} \}$$

and

$$S_2 := \{(x, y, \lambda, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{n+1} \times \{0, 1\} :$$

$$x = \sum_{i=0}^{n} B_i \lambda_i, \quad y = \sum_{i=0}^{n} F_i \lambda_i, \quad z = \sum_{i=0}^{n} \lambda_i, \quad \lambda \text{ is SOS2} \}$$

where $S_1$ is a standard SOS2 model for PLFs that uses the constraint (4), while formulation $S_2$ uses the constraint (6) to model the logical condition (2). One can easily show that both $S_1$ and $S_2$ are valid formulations of $X$. In other words, for either $T = S_1$ or $T = S_2$,

$$X = \{(x, y, z) : \exists \lambda \in \mathbb{R}_+^{n+1} \text{ s.t. } x, y, z, \lambda \in T \}.$$ 

We use the standard definition of the linear programming (LP) relaxation of a model as the relaxation obtained by replacing integrality restrictions on variables with simple bound restrictions and by removing adjacency requirements for SOS2 variables. We now prove that the formulation $S_2$ is locally ideal while $S_1$ is not.

**Theorem 1.** Formulation $S_2$ is locally ideal.

**Proof.** The LP relaxation of $S_2$ has $n + 4$ variables, three equality constraints

$$x = \sum_{i=0}^{n} B_i \lambda_i, \quad y = \sum_{i=0}^{n} F_i \lambda_i, \quad z = \sum_{i=0}^{n} \lambda_i,$$

and $n + 2$ inequality constraints, $z \leq 1$ and $\lambda_i \geq 0 \forall i \in \{0, 1, \ldots, n\}$. Extreme points of the LP relaxation of $S_2$ have $n + 4$ binding constraints, which forces at least $n$ variables from $\lambda \in \mathbb{R}_+^{n+1}$ to be exactly equal to zero. Thus, the extreme points of the LP relaxation of $S_2$ are

$$\{(x = B_i, \quad y = F_i, \quad \lambda = B_i \bar{e}_i, \quad z = 1) \forall i \in \{1, \ldots, n\} \} \cup \{(x = 0, y = 0, \lambda = \bar{e}, z = 0)\} \quad (7)$$

where $\bar{e}_i$ are the $n$ dimensional unit vectors. All points in (7) have $z \in \{0, 1\}$ and satisfy the SOS2 properties for the $\lambda$ variables. Hence, $S_2$ is locally ideal. $\square$

A point $(x, y, z, \lambda)$ can be an extreme point of the set

$$P_T := \{(x, y, \lambda, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{n+1} \times \{0, 1\} :$$

$$x = \sum_{i=0}^{n} B_i \lambda_i, \quad y \geq \sum_{i=0}^{n} F_i \lambda_i, \quad z = \sum_{i=0}^{n} \lambda_i \}$$

only if $y = \sum_{i=0}^{n} F_i \lambda_i$. Therefore, the proof of Theorem 1 also establishes that expressing the logical condition (2) using (6) also...
results in a locally ideal formulation of $X^\infty$. Similar logic applies in our subsequent proofs of the local idealness of other formulations of $X$ (Theorems 4 and 6). In each case, our proposed modeling of the logical condition (2) also yields a locally ideal formulation of $X^\infty$.

**Theorem 2.** Formulation $S_1$ is not locally ideal.

**Proof.** Consider an instance with $n = 3$, $B = \{0, 1, \frac{1}{3}, \frac{3}{4}\}$, and $F = \{0, 4, 2, 3\}$. The point $\{x = \frac{1}{3}, y = 4, z = 0, 1\}$ is feasible to the LP relaxation of $S_1$, but not feasible for the LP relaxation of $S_2$. Since the projection of the LP relaxation of $S_2$ is a subset of the projection of the LP relaxation of $S_1$, $S_1$ cannot be locally ideal. □

An interesting consequence of Theorem 1 is that when the PLF is convex, the application of the reformulation technique we suggest to the set $X^\infty$ is equivalent to the perspective reformulation [11], a preprocessing technique for (convex) mixed integer nonlinear programs that have the logical indicator structure (2). If $f(\cdot)$ is convex, then $h_1 > h_2 > \ldots > h_n$, and the perspective reformulation of $X^\infty$ is

$$P = \{(x, y, z) \in \mathbb{R}^3 \times \{0, 1\} : y \geq h_1 x + c z \forall i \in \{1, \ldots, n\}, 0 \leq x \leq b_0,\}$$

where $h_i := (F_i - F_{i-1})/(B_i - B_{i-1})$ and $c_i := (F_{i-1} - b_0 (F_i - F_{i-1})/(B_i - B_{i-1}))$. Günlük and Linderoth [10] show that if $f(\cdot)$ is convex, then $P = \text{conv}(X^\infty)$. The formulation $S_2$ is locally ideal, so $P_2^\infty$ must also be a formulation that is similarly strong.

**Corollary 3.** $\text{Proj}_{x,y}(P_2^\infty) = P = \text{conv}(X^\infty)$

**2.2. Incremental model**

The incremental model introduces a set of non-negative variables $\delta = \{\delta_1, \ldots, \delta_n\}$ to model the portion of each interval “filled” by the variable $x$. The interval $i+1$ can be filled if $\delta_{i+1} > 0$ only if the interval $i$ is already filled ($\delta_i = 1$). Unlike the S02 model, the incremental model specifically requires the introduction of binary variables $b \in \{0, 1\}^{n+1}$ to enforce the necessary ordering conditions. To model the on-off logical condition (2), the incremental model can be augmented with a variable upper bound constraint $x \leq b_0 z$, resulting in a formulation

$$\Delta_1 := \{(x, y, \delta, z, b) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \{0, 1\} \times \{0, 1\}^{n+1} :$$

$$x = \sum_{i=1}^n [B_i - B_{i-1}] \delta_i, \quad y = \sum_{i=1}^n [F_i - F_{i-1}] \delta_i, \quad x \leq b_0 z,$$

$$\delta_1 \leq 1, \quad 0 \leq \delta_n, \quad \delta_{i+1} \leq b_i \leq \delta_i \forall i \in \{1, \ldots, n-1\} \}.$$ 

Alternatively, the on-off condition can be enforced by replacing the constraint $\delta_1 \leq 1$ with $\delta_1 \leq z$, yielding the formulation

$$\Delta_2 := \{(x, y, \delta, z, b) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \{0, 1\} \times \{0, 1\}^{n+1} :$$

$$x = \sum_{i=1}^n [B_i - B_{i-1}] \delta_i, \quad y = \sum_{i=1}^n [F_i - F_{i-1}] \delta_i,$$

$$\delta_1 \leq z, \quad 0 \leq \delta_n, \quad \delta_{i+1} \leq b_i \leq \delta_i \forall i \in \{1, \ldots, n-1\} \}.$$ 

Incremental models that use $\delta_1 \leq z$ are locally ideal, while those that use $x \leq b_0 z$ are not.

**Theorem 4.** Formulation $\Delta_2$ is locally ideal.

**Proof.** The matrix for the constraint system in $\Delta_2$, ignoring the equality constraints defining $x$ and $y$, is

$$-\delta_1 + z \geq 0, \quad b_i - b_{i-1} \geq 0 \quad \forall i \in \{1, \ldots, n-1\}, \quad -\delta_{i+1} + b_i \geq 0 \quad \forall i \in \{1, \ldots, n-1\},$$

$$\delta_0 \geq 0,$$

which is a network matrix, and hence is totally unimodular. Thus all extreme points of the LP relaxation of $\Delta_2$ naturally satisfy the requisite integrality properties. □

**Theorem 5.** Formulation $\Delta_1$ is not locally ideal.

**Proof.** Consider an instance with $n = 3$, $B = \{0, 1, \frac{1}{3}, \frac{3}{4}\}$ and $F = \{0, 4, 2, 3\}$. The point $\{x = \frac{1}{3}, y = 4, z = (1, 0, 2, 1)， b = (0, 0)\}$ is feasible to the LP relaxation of $\Delta_1$, but not feasible to the LP relaxation of $\Delta_2$. □

**2.3. Multiple choice model**

In the multiple choice model, a non-negative set of variables $w := \{w_1, \ldots, w_n\}$ and an additional set of binary variables $b := \{b_1, \ldots, b_n\}$ are introduced, with the logical implication that $w_i = x$ if $x$ is in the $i$th interval, and $w_i = 0$ otherwise. Using a variable upper bound constraint to enforce the logical condition (2) with the multiple choice model gives the following formulation of $X$:

$$M_1 := \{(x, y, w, z, b) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \{0, 1\} \times \{0, 1\}^{n+1} :$$

$$w_i = x, \quad y = \sum_{i=1}^n (m_i w_i + c_i b_i), \quad x \leq b_0 z,$$

$$b_0 \leq 1, \quad b_{i+1} \leq w_i \leq b_i \forall i \in \{1, \ldots, n\} \}.$$ 

Instead, the on-off condition can be formulated by replacing the constraints $\sum_{i=1}^n b_i = 1$ with $\sum_{i=1}^n b_i = z$, yielding a formulation

$$M_2 := \{(x, y, w, z, b) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \{0, 1\} \times \{0, 1\}^{n+1} :$$

$$w_i = x, \quad y = \sum_{i=1}^n (m_i w_i + c_i b_i),$$

$$b_0 \leq z, \quad b_{i+1} \leq w_i \leq b_i \forall i \in \{1, \ldots, n\} \}.$$ 

**Theorem 6.** Formulation $M_2$ is locally ideal.

**Proof.** Following Balas [3], we write an extended formulation for the convex hull of the union of the $n+1$ polytopes $X_0 = \{(0, 0, 0)\}$, $X_i = \{(x, y, z) : B_i \leq x \leq B_{i+1}, y = m_i x + c_i, z = 0\} \forall i \in \{1, \ldots, n\}$ as the convex hull of the union of the $n+1$ polytopes $X_0$, $X_i$, $X_{i+1}$ such that for which there exist vectors $w := (w_0, \ldots, w_n)$, $v := (v_0, \ldots, v_n)$, $u := (u_0, \ldots, u_n)$, $b := (b_0, \ldots, b_n)$ such that the following inequality system is satisfied:

$$x = \sum_{i=0}^n u_i, \quad y = \sum_{i=0}^n v_i, \quad z = \sum_{i=0}^n u_i, \quad 1 = \sum_{i=0}^n b_i,$$

$$u_0 = 0, \quad v_0 = 0, \quad u_n = 0, \quad b_i \geq 0 \forall i \in \{0, \ldots, n\},$$

$$b_{i+1} \leq w_i \leq b_i \forall i \in \{1, \ldots, n\},$$

$$u_i = b_i \forall i \in \{1, \ldots, n\}.$$
We can eliminate $b_0$, $u$, and $v$ from this system to obtain
\[
X = \sum_{i=1}^{n} w_i, \quad Y = \sum_{i=1}^{n} (m_i w_i + c(b_i)),
\]
\[
z = \sum_{j=1}^{n} b_j, \quad z \leq 1
\]
\[
b_i \geq 0, \quad B_i \preceq u_i \preceq B_i \quad \forall i \in \{1, \ldots, n\},
\]
which is equivalent to the LP relaxation of $M_2$. □

**Theorem 7.** Formulation $M_1$ is not locally ideal.

**Proof.** Consider an instance with $n = 3$, $B = \{(0, 1, 2, 1)\}$, and
\[
F = \{0, 4, 2, 3\).
\]
The point $x = \frac{1}{3}, y = 4, w = (0, 1, 0), z = \frac{1}{3}, b = (0, 1, 0)\}$ is feasible to the linear programming relaxation of $M_1$, but not feasible for $M_2$. □

### 2.4. Convex combination model

Another popular formulation for PLFs is the convex combination model, also known as the lambda method. The convex combination model uses continuous variables $\lambda \in \mathbb{R}^{n+1}$ and binary variables $b \in \{0, 1\}^n$. The continuous variables are used to express $x$ and $y$ in terms of the breakpoints $B$ and function values $F$. The binary variables are used to enforce the adjacency condition that $b_i = 1 \Rightarrow \lambda_i = 0$, $\forall i \not\in \{1, n\}$. Using a variable upper bound to model the logical on–off condition (2) in combination with the most commonly used convex combination model gives the following formulation of $X$:
\[
C_1 = \{ (x, y, z, b) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n+1} \times \{0, 1\} \times \{0, 1\}^n \ :
\]
\[
\begin{aligned}
X &= \sum_{i=0}^{n} \lambda_i B_i, \quad Y = \sum_{i=0}^{n} \lambda_i F_i, \quad x \preceq B_0 z, \\
\sum_{i=0}^{n} \lambda_i &= 1, \quad \sum_{i=0}^{n} b_i = 1, \\
\lambda_0 &\preceq b_1, \quad \lambda_n \preceq b_n, \quad \lambda_i \preceq b_i + b_{i+1} \forall i \in \{1, \ldots, n-1\}.
\end{aligned}
\]

Instead, the on–off condition can be directly imposed by replacing $\sum_{i=0}^{n} b_i = 1$ and $\sum_{i=0}^{n} \lambda_i = 1$ with the constraints $\sum_{i=0}^{n} b_i = 1$ and $\sum_{i=0}^{n} \lambda_i = z$. This gives the following formulation of $X$:
\[
C_2 = \{ (x, y, z, b) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n+1} \times \{0, 1\} \times \{0, 1\}^n \ :
\]
\[
\begin{aligned}
X &= \sum_{i=0}^{n} \lambda_i B_i, \quad Y = \sum_{i=0}^{n} \lambda_i F_i, \quad \sum_{i=0}^{n} \lambda_i = z, \quad \sum_{i=0}^{n} b_i = z, \\
\lambda_0 &\preceq b_1, \quad \lambda_n \preceq b_n, \quad \lambda_i \preceq b_i + b_{i+1} \forall i \in \{1, \ldots, n-1\}.
\end{aligned}
\]

It has been shown by Padberg [20] and Lee and Wilson [15] that the convex combination model that uses the constraints
\[
\lambda_0 \preceq b_1, \quad \lambda_n \preceq b_n, \quad \lambda_i \preceq b_i + b_{i+1} \forall i \in \{1, \ldots, n-1\} \tag{8}
\]

of model adjacency is not locally ideal. Padberg [20] gives the following improved formulation:
\[
\sum_{i=0}^{n} \lambda_i \leq \sum_{i=0}^{n} b_i, \quad \sum_{j=0}^{n-1} \lambda_j \leq \sum_{i=0}^{n} b_i \forall j = 1, \ldots, n,
\]
which does result in a locally ideal formulation of PLFs. However, in most presentations of the convex combination model in the literature [8,5,12,9] the non-ideal formulation (8) is used.

The convex combination model with constraints (8) does not result in a formulation that is locally ideal, however it does satisfy sharpness, a slightly weaker desirable property. An extended MIP formulation of a convex set is sharp if the extreme points of the projection of the LP relaxation of the formulation to the original space of variables ($x$, $y$, $z$ in this case) satisfy integrality [13]. Vielma et al. [23] showed that the convex combination model that uses adjacency constraint (8) is sharp. We now show that the formulation $C_2$ is sharp while $C_1$ is not sharp.

**Theorem 8.** Formulation $C_2$ is sharp.

**Proof.** Suppose that $t = (x, y, z, b)$ is an extreme point of the linear programming relaxation of $C_2$ with $0 < z < 1$. For $e > 0$ define the points $t^+ = (x^+, y^+, z^+, b^+)$ and $t^- = (x^-, y^-, z^-, b^-)$ as
\[
b_i^+ = (1 + e) b_i, \quad \lambda_i^+ = (1 + e) \lambda_i, \quad \forall i \in \{1, \ldots, n\}
\]
\[
z^+ = (1 + e) z, \quad \lambda^+ = \sum_{i=0}^{n} \lambda_i^+ B_i, \quad y^+ = \sum_{i=0}^{n} \lambda_i^+ F_i
\]
\[
b_i^- = (1 - e) b_i, \quad \lambda_i^- = (1 - e) \lambda_i, \quad \forall i \in \{1, \ldots, n\}
\]
\[
z^- = (1 - e) z, \quad \lambda^-= \sum_{i=0}^{n} \lambda_i^- B_i, \quad y^- = \sum_{i=0}^{n} \lambda_i^- F_i.
\]

For some $e > 0$, the points $t^+, t^-$ are both feasible for the linear programming relaxation of $C_2$, and $t = 0.5(t^+ + t^-)$, so $t$ must not have been an extreme point. □

**Theorem 9.** Formulation $C_1$ is not sharp.

**Proof.** Consider an instance with $n = 3$, $B = \{(0, 1, 2, 1)\}$, and
\[
F = \{0, 4, 2, 3\}.
\]
One can verify that one of extreme points of the projection of the linear programming relaxation of $C_1$ is $\{x = \frac{1}{3}, y = 4, z = \frac{1}{3}\}$, which does not satisfy the required integrality constraint on $z$. □

### 2.5. Other formulations

The disaggregated convex combination model for PLFs uses two sets of non-negative variables $\lambda \in \mathbb{R}_+^n$ and $\mu \in \mathbb{R}_+^n$ and a set of binary variables $b \in \{0, 1\}^n$. The disaggregated convex combination model for a PLF is
\[
y = \sum_{i=1}^{n} (\lambda_i F_i + \mu_i F_{i-1}), \quad x = \sum_{i=1}^{n} (\lambda_i B_i + \mu_i B_{i-1})
\]
\[
\sum_{i=1}^{n} b_i = 1, \quad b_i = \lambda_i + \mu_i \forall i \in \{1, \ldots, n\}.
\]

This formulation can be extended to model $X$ by replacing the constraints $\sum_{i=1}^{n} b_i = 1$ with $\sum_{i=1}^{n} b_i = z$. The disaggregated convex combination model that uses these constraints is a locally ideal formulation of $X$.

Vielma and Nemhauser [24] modify the disaggregated convex combination model to use a logarithmic number of binary variables. Using notation defined in Vielma and Nemhauser [24], replacing $\sum_{i=1}^{n} \lambda_i = 1$ with $\sum_{i=1}^{n} \lambda_i = z$ is a valid locally ideal reformulation of model $X$. For the sake of brevity, we have omitted detailed discussions and proofs concerning the disaggregated convex combination models.
3. Computational results

In this section, we illustrate with numerical experiments the impact of using a locally ideal formulation (S₂) instead of a weaker model (S₁) that is not locally ideal.

3.1. Practical application

To make the numerical comparison, we consider an advertising budget allocation problem introduced by Zoltners and Sinha [25]. In this problem, a company is required to allocate an advertising budget D among a set \( K \) of advertising strategies for a set of \( P \) products. Let \( x_{jk} \) denote the amount of the advertising resource allocated to strategy \( k \in K \) for product \( j \in P \). The company incurs a fixed cost \( G_j \) for entering the market with product \( j \in P \) as well as a variable cost \( c_{jk} \) for each unit of the resource allocated to strategy \( k \in K \) of product \( j \in P \). The return on investment is evaluated by piecewise-linear functions \( y_{jk} = f_j(x_{jk}) \) which have the typical form shown in Fig. 1.

A MIP formulation for this problem is

\[
\text{max} \sum_{j \in P} \sum_{k \in K} y_{jk} \\
\sum_{j \in P} \sum_{k \in K} c_{jk} x_{jk} + \sum_{j \in P} G_j z_j \leq D \\
(x_{jk}, y_{jk}, z_j) \in X_{jk}, \forall j \in P, \forall k \in K \text{,} \quad (AP)
\]

where \( X_{jk} \) is meant to denote that each of the triplets of variables \( (x_{jk}, y_{jk}, z_j) \) must lie in a set \( X \) (defined in (3)) specific to the product/strategy pair. We denote by AP(S₁) the MIP formulation of AP that uses \( S_1 \) to model (3) and AP(S₂) as the MIP formulation of AP that uses the stronger formulation \( S_2 \).

3.2. Numerical results

We report tests conducted on 120 simulated instances of AP(X). We created 20 random instances for each of the six problem sizes \( (|P|, |K|, n) \in \{(50, 50, 10), (50, 100, 10), (100, 100, 10), (50, 50, 20), (50, 100, 20), (100, 100, 20)\} \). All instances were solved to 0.1% optimality using Gurobi 4.5.1 with default options on 2.66 GHz Intel Core2 Quad CPU Q9400 processor with 8GB RAM. For all instances, we compare the quality of the LP relaxation as the percentage gap between the root LP relaxation value of the MIP formulations AP(S₂) and AP(S₁) relative to the optimal solution for each instance. We also measure the CPU time taken (using a single thread) and number of nodes in the search tree. Table 2 shows the summary statistics of our experiment.

The results convincingly demonstrate the advantage of using the locally ideal formulation AP(S₂). The average root gap for AP(S₁) was 0.05%, while for AP(S₂) the average root gap was 19.6%. In fact, the best root gap for any instance of AP(S₁) was 17.1%. In terms of MIP solve times, AP(S₂) was solved in 16.8 s on average, 41.8 times faster than the 703 s average time to solve AP(S₁). In the worst case, Gurobi explored 1117 times more nodes on an instance modeled with AP(S₁) than with AP(S₂). Clearly, one should use the locally ideal model AP(S₂).

4. Concluding remarks

The multiple choice formulation can be seen as an application of disjunctive programming [3,22]. An interesting direction for future work would be to explore the application of our proposed formulation strengthening technique to MIP models of disjunctive constraints that have a structure similar to (2).

Acknowledgments

This work was supported through a contract from Argonne, a U.S. Department of Energy Office of Science laboratory. This work was supported by the Applied Mathematics activity, Advance Scientific Computing Research program within the DOE Office of Science.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.orl.2013.08.010.

References


Table 2

<table>
<thead>
<tr>
<th>Metric</th>
<th>Model</th>
<th>A.M</th>
<th>St. Dev</th>
<th>G.M</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP gap (%)</td>
<td>AP(S₁)</td>
<td>0.05</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>AP(S₂)</td>
<td>19.6</td>
<td>1.5</td>
<td>19.6</td>
</tr>
<tr>
<td>Time (s)</td>
<td>AP(S₁)</td>
<td>16.8</td>
<td>12.0</td>
<td>12.3</td>
</tr>
<tr>
<td></td>
<td>AP(S₂)</td>
<td>703.0</td>
<td>853.1</td>
<td>255.6</td>
</tr>
<tr>
<td>Nodes</td>
<td>AP(S₁)</td>
<td>402.9</td>
<td>312.4</td>
<td>314.5</td>
</tr>
<tr>
<td></td>
<td>AP(S₂)</td>
<td>312.9</td>
<td>312.4</td>
<td>314.5</td>
</tr>
</tbody>
</table>

Fig. 1. Sample curves modeling return on investment for five different product/strategy pairs.