On the Value Function of a Mixed Integer Linear Optimization Problem and an Algorithm for its Construction

Ted K. Ralphs and Anahita Hassanzadeh

Department of Industrial and Systems Engineering, Lehigh University, USA

COR@L Technical Report 14T-004

On the Value Function of a Mixed Integer Linear Optimization Problem and an Algorithm for its Construction

TED K. RALPHS*[1] AND ANAHITA HASSANZADEH†[1]

[1]Department of Industrial and Systems Engineering, Lehigh University, USA

August 3, 2014

Abstract

This paper addresses the *value function* of a general mixed integer linear optimization problem (MILP). The value function describes the change in optimal objective value as the right-hand side is varied and understanding its structure is central to solving a variety of important classes of optimization problems. We propose a discrete representation of the MILP value function and describe a cutting plane algorithm for its construction. We show that this algorithm is finite when the set of right-hand sides over which the value function of the associated pure integer optimization problem is finite is bounded. We explore the structural properties of the MILP value function and provide a simplification of the *Jeroslow Formula* obtained by applying our results.

1 Introduction

Understanding and exploiting the structure of the value function of an optimization problem is a critical element of solution methods for a variety of important classes of multi-stage and multi-level optimization problems. Previous findings on the value function of a *pure integer linear optimization problem* (PILP) have resulted in finite algorithms for constructing it, which have in turn enabled the development of solution methods for two-stage stochastic pure integer optimization problems (Schultz et al., 1998; Kong et al., 2006) and certain special cases of bilevel optimization problems (Bard, 1991, 1998; S DeNegre, 2011; Dempe et al., 2012). Studies of the value function of a general *mixed integer linear optimization problem* (MILP), however, have not yet led to algorithmic advances. Algorithms for construction have only been proposed in certain special cases (Güzelsoy and Ralphs, 2006), but no practical characterization is known. By "practical," we mean a characterization and associated representation that is suitable for computation and which we can use to formulate problems in which the value function of a MILP is embedded, e.g., multi-stage stochastic programs.

In this paper, we extend previous results by demonstrating that the MILP value function has an underlying *discrete structure* similar to the PILP value function, even in the general case. This

*E-mail: `ted@lehigh.edu`

†E-mail: `anahita.hassanzadeh@gmail.com`

discrete structure emerges from separating the function into discrete and continuous parts, which in turn enables a representation of the function in terms of two discrete sets. We show that this representation can be constructed and propose an algorithm for doing so. Both the representation and the algorithm are finite under the assumption that the set of right-hand sides over which the related PILP value function is finite is bounded.

We consider the value function associated with a nominal MILP instance defined as follows. The variables of the instance are indexed on the set $N = \{1, \ldots, n\}$, with $I = \{1, \ldots, r\}$ denoting the index set for the integer variables and $C = \{r+1, \ldots, n\}$ denoting the index set for the continuous variables. For any $D \subseteq N$ and a vector y indexed on N, we denote by y_D the sub-vector consisting of the corresponding components of y. Similarly, for a matrix M, we denote by M_D the sub-matrix constructed by columns of M that correspond to indices in D. Then, the nominal instance we consider throughout the paper is given by

$$z_{IP} = \inf_{(x,y)\in X} c_I^\top x + c_C^\top y, \tag{MILP}$$

where $(c_I^\top, c_C^\top)^\top \in \mathbb{R}^n$ is the objective function vector and $X = \{(x,y) \in \mathbb{Z}_+^I \times \mathbb{R}_+^C : A_I x + A_C y = d\}$ is the feasible region, described by $A_I \in \mathbb{Q}^{m\times r}$, $A_C \in \mathbb{Q}^{m\times(n-r)}$, and $d \in \mathbb{R}^m$.

A pair of vectors $(x, y) \in X$ is called a *feasible solution* and $c_I^\top x + c_C^\top y$ is its associated *solution value*. For such a solution, the vector x is referred to as the *integer part* and the vector y as the *continuous part*. Any $(x^*, y^*) \in X$ such that $c_I^\top x^* + c_C^\top y^* = z_{IP}$ is called an *optimal solution*. Throughout the paper, we assume $rank(A, d) = rank(A) = m$.

The value function is a function $z : \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$ that describes the change in the optimal solution value of a MILP as the right-hand side is varied. In the case of (MILP), we have

$$z(b) = \inf_{(x,y)\in S(b)} c_I^\top x + c_C^\top y \quad \forall b \in B, \tag{MVF}$$

where for $b \in \mathbb{R}^m$, $S(b) = \{(x,y) \in \mathbb{Z}_+^r \times \mathbb{R}_+^{n-r} : A_I x + A_C y = \hat{b}\}$ and $S_I(b) = \{x \in \mathbb{Z}_+^r : A_I x = b\}$. We let $B = \{b \in \mathbb{R}^m \ : S(b) \neq \emptyset\}$, $B_I = \{b \in \mathbb{R}^m \ : S_I(b) \neq \emptyset\}$ and $S_I = \cup_{b\in B} S_I(b)$. By convention, we let $z(b) = \infty$ for $b \in \mathbb{R}^m \setminus B$ and $z(b) = -\infty$ when the infimum in (MVF) is not attained for any $x \in S(b)$ for $b \in B$. To simplify the presentation, we assume that $z(0) = 0$, since otherwise $z(b) = -\infty$ for all $b \in \mathbb{R}^m$ with $S(b) \neq \emptyset$ (Nemhauser and Wolsey, 1988).

To illustrate the basic concepts, we now present a brief example that we refer to throughout the paper.

Example 1. Consider the MILP value function defined by

$$\begin{aligned} z(b) = \inf\ & 3x_1 + \frac{7}{2}x_2 + 3x_3 + 6x_4 + 7x_5 \\ \text{s.t. } & 6x_1 + 5x_2 - 4x_3 + 2x_4 - 7x_5 = b \\ & x_1, x_2, x_3 \in \mathbb{Z}_+,\ x_4, x_5 \in \mathbb{R}_+. \end{aligned} \tag{Ex.1}$$

Figure 1 shows this non-convex, non-concave piecewise polyhedral function. ■

Although the MILP with which the value function in Example 1 is associated has only a single constraint, the structure of the function is already quite complex. Nevertheless, the function does have an obvious regularity to it. The gradient of the function is always one of only two values,

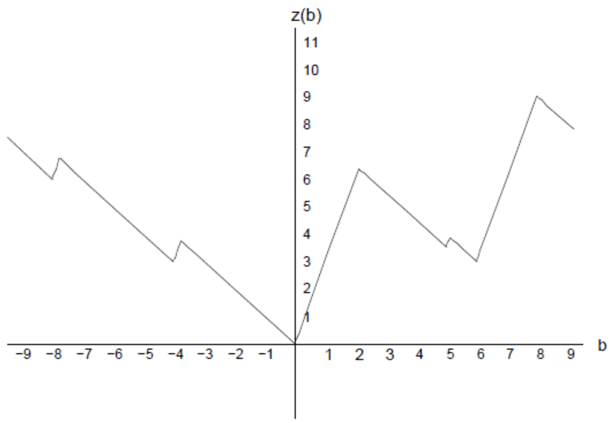


Figure 1: MILP Value Function of (Ex.1).

which means that its epigraph is the union of a set of polyhedral cones that are identical, aside from the location of their extreme points. This allows the function to be represented simply by describing this single polyhedral cone and a discrete set of points at which we place translations of it.

In the remainder of the paper, we formalize the basic idea illustrated in Example 1 and show that it can be generalized to obtain a similar result that holds for all MILP value functions. More specifically, in Section 3, we review some basic properties of the LP value function and define the *continuous restriction*, an LP whose value function yields the aforementioned polyhedral cone. In Sections 4, we present the first main result in Theorem 1, which characterizes the minimal discrete set of points that yields a full description of the value function. This discrete set generalizes the *minimal tenders* of Kong et al. (2006) used to represent the PILP value function. In Section 5, we exploit this representation to uncover further properties of the value function, culminating in Theorem 2, which shows that there is a one-to-one correspondence between the discrete set of Theorem 1 and the regions over which the value function is convex (the *local stability regions*). In Section 6, we discuss the relationship of our representation with the well-known *Jeroslow Formula*, showing in Theorem 3 that the original formula can be simplified using the discrete representation of Theorem 1. Finally, in Section 7 we demonstrate how to put the representation into computational practice by presenting a cutting plane algorithm for constructing a discrete representation such as the one in Theorem 1 (though possibly not provably minimal). Before getting to the main results, we next summarize related work.

2 Related Work

Much of the recent work on the value function has addressed the pure integer case, since the PILP value function has some desirable properties that enable more practical results. Blair and Jeroslow (1982) first showed that the value function of a PILP is a *Gomory function* that can be derived by taking the maximum of finitely many *subadditive functions*. Conti and Traverso (1991) then proposed using *reduced Gröbner basis* methods to solve PILPs. Subsequently, in the context of stochastic optimization, Schultz et al. (1998) used the so-called *Buchberger Algorithm* to compute the reduced Gröbner basis for solving sequences of integer optimization problems which differ only in their right-hand sides. In the same paper, the authors recognized that over certain regions of the right-hand side space, the pure integer value function remains constant. This property turned out to be quite significant, resulting in algorithms for two-stage stochastic optimization (Ahmed et al., 2004; Kong et al., 2006). In the same vein, Kong et al. (2006) proposed using the properties of a pure integer optimization problem in two algorithms for constructing the PILP value function when the set of right-hand sides is finite.

The complex structure of the MILP value function makes the extension of results in linear and pure integer optimization to the general case a challenge. In particular, with the introduction of continuous variables, we no longer have countability of the set of right-hand sides for which $S(b) \neq \emptyset$ (or finiteness in the case of a bounded $S(b)$), which is a central property in the PILP case. The MILP value function also ostensibly lacks certain properties used in previous algorithms for eliminating parts of the domain from consideration in the pure integer case. For the mixed integer case, Bank et al. (1983) studied the MILP value function in the context of parametric optimization and provided theoretical results on regions of the right-hand side set over which the function is continuous. In a series of papers, Blair and Jeroslow (1977, 1979) studied the properties of the MILP value function and showed that it is a piecewise polyhedral function. Blair and Jeroslow (1984) identified a subclass of Gomory functions called *Chvátal functions* to which the general MILP value function belongs. However, a closed form representation was not achieved until a decade later in a subsequent work of Blair (1995). The so-called *Jeroslow Formula* represents the MILP value function as collection of Gomory functions with linear correction terms. This characterization is related to ours and we discuss this relationship in Section 6.

3 The Continuous Restriction

To understand the MILP value function, it is important to first understand the structure of the value function of a linear optimization problem. In particular, we are interested in the structure of the value function of the LP arising from (MILP) by fixing the values of the integer variables. We call this problem the *continuous restriction* (CR) w.r.t a given $\hat{x} \in S_I$. Its value function is given by

$$\begin{aligned} \bar{z}(b;\hat{x}) = c_I^\top \hat{x} + \inf\ & c_C^\top y \\ \text{s.t. } & A_C y = b - A_I \hat{x} \\ & y \in \mathbb{R}_+^{n-r}. \end{aligned} \tag{CR}$$

For a given $\hat{x} \in S_I$, we let $S(b, \hat{x}) = \{y \in \mathbb{R}_+^{n-r} : A_C y = b - A_I \hat{x}\}$. As before, we let $\bar{z}(b;\hat{x}) = \infty$ if $S(b, \hat{x}) = \emptyset$ for a given $b \in B$ and $\bar{z}(b;\hat{x}) = -\infty$ if the function value is unbounded. As we will show formally in Proposition 4, it is evident that for any $\hat{x} \in S_I$, $\bar{z}(\cdot;\hat{x})$ bounds the value function from above, which is the reason for the notation.

When $\hat{x} = 0$ in (CR), the resulting function is in fact the value function of a general LP, since A_C is itself an arbitrary matrix. In the remainder of the section, we consider this important special case and define

$$\begin{aligned} z_C(b) = \inf\ & c_C^\top y \\ \text{s.t. } & A_C y = b \\ & y \in \mathbb{R}_+^{n-r}. \end{aligned} \tag{LVF}$$

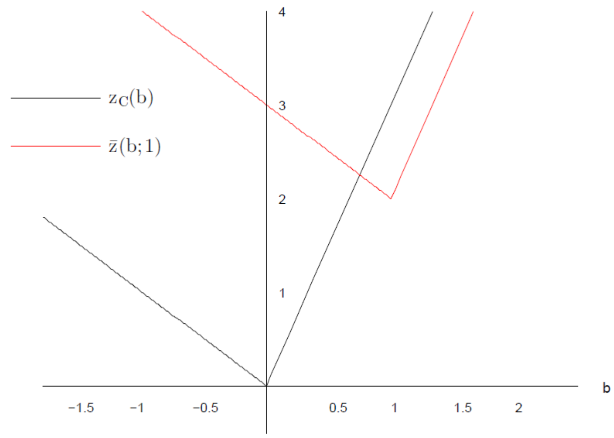


Figure 2: The value function of the continuous restriction of (Ex.2) and a translation.

We let $\mathcal{K}$ be the polyhedral cone that is the positive linear span of A_C, i.e., $\mathcal{K} = \{\lambda_1 A^{r+1} + \ldots + \lambda_{n-r} A^n : \lambda_1, \ldots, \lambda_{n-r} \geq 0\}$. As we discuss later, this cone is the set of right-hand sides for which z_C is finite and plays an important role in the structure of both the LP and MILP value functions. The following example illustrates the continuous restriction associated with a given MILP.

Example 2. Consider the MILP

$$\begin{aligned} \inf\ & 2x_1 + 6x_2 + 7x_3 + 5x_4 \\ \text{s.t. } & x_1 + 2x_2 - 7x_3 + x_4 = b \\ & x_1 \in \mathbb{Z}_+, x_2, x_3, x_4 \in \mathbb{R}_+. \end{aligned} \tag{Ex.2}$$

The value functions of the continuous restriction w.r.t. $x_1 = 0$ and $x_1 = 1$ are plotted in Figure 2. ■

Note that in the example just given, $\bar{z}(\cdot; 1)$ is simply a translation of z_C. As we will explore in more detail later, this is true in general, so that for $\hat{x} \in S_I$, we have

$$\bar{z}(b; \hat{x}) = c_I^\top \hat{x} + z_C(b - A_I \hat{x})\ \forall b \in B.$$

Thus, the following results can easily be generalized to the continuous restriction functions w.r.t. points other than the origin.

We shall now more formally analyze the structure of z_C. We first present a representation due to Blair and Jeroslow (1977), who characterized the LP value function in terms of its epigraph. Let $\mathcal{L} = \text{epi}(z_C)$.

Proposition 1 *(Blair and Jeroslow, 1977) The value function of z_C is a convex polyhedral function and its epigraph $\mathcal{L}$ is the convex cone*

$$\textit{cone}\{(A^{r+1}, c_{r+1}), (A^{r+2}, c_{r+2}), \ldots, (A^n, c_n), (0, 1)\}.$$

The above description of the LP value function in terms of a cone is not computationally convenient for reasons that will become clear. We can derive a more direct characterization of the LP value function by considering the structure of the dual of (LVF) for a fixed right-hand side $\hat{b} \in \mathbb{R}^m$. In particular, this dual problem is

$$\sup_{\nu \in S_D} \hat{b}^\top \nu, \tag{3.1}$$

where $S_D = \{\nu \in \mathbb{R}^m : A_C^\top \nu \leq c_C\}$. Note that our earlier assumption that $z(0) = 0$ implies $S_D \neq \emptyset$. Let $\{\nu^i\}_{i \in K}$ be the set of extreme points of S_D, indexed by set K. When S_D is unbounded, let its set of extreme directions $\{d^j\}_{j \in L}$ be indexed by set L. From strong duality, we have that $z_C(\hat{b}) = \sup_{\nu \in S_D} \hat{b}^\top \nu$ when $S_D \neq \emptyset$. If the LP with right-hand side $\hat{b}$ has a finite optimum, then

$$z_C(\hat{b}) = \sup_{\nu \in S_D} \hat{b}^\top \nu = \sup_{i \in K} \hat{b}^\top \nu^i. \tag{3.2}$$

Otherwise, for some $j \in L$, we have $\hat{b}^\top d^j > 0$ and $z_C(\hat{b}) = +\infty$. We can therefore obtain a representation of the cone $\mathcal{L}$ as

$$\{(b, z) \in \mathbb{R}^{m+1} : b^\top \nu^i \leq z,\ b^\top d^j \leq 0, i \in K, j \in L\}.$$

Let $\mathcal{E}$ be the set of index sets of the nonsingular square sub-matrices of A_C corresponding to dual feasible bases. That is, $E \in \mathcal{E}$ if and only if $\exists i \in K$ such that $A_E^\top \nu^i = c_E$. Abusing notation slightly, we denote this (unique) ν^i by ν_E in order to be consistent with the literature. The cone $\mathcal{L}$ has an extreme point if and only if there exist $m+1$ linearly independent vectors in the set $\{(\nu^i,\ -1) : i \in K\} \cup \{(d^j,\ 0) : j \in L\}$. It is easy to show that in this case, the origin is the single extreme point of $\mathcal{L}$ and all dual extreme points are optimal at the origin, i.e., $\nu_E^\top 0 = c_E^\top A_E^{-1} 0 = z_C(0) = 0$ for all $E \in \mathcal{E}$. Conversely, when $\mathcal{L}$ has an extreme point, it must be the single point at which all the inequalities in the description of $\mathcal{L}$ are binding.

The convexity of $z_C(b)$ follows from the representation (3.2), since $z_C(b)$ is the maximum of a finite number of affine functions and is hence a convex polyhedral function (Bazaraa et al., 1990; Blair and Jeroslow, 1977). With respect to differentiability, consider a right-hand side $b \in B$ for which the optimal solution to the corresponding LP is non-degenerate. Let the (unique) optimal basis and optimal dual solution be A_E and ν_E, respectively, for some $E \in \mathcal{E}$. As a result of the unchanged reduced costs, under a small enough perturbation in b, A_E and ν_E remain the optimal basis and dual solution to the new problem. Hence, the function is affine in a neighborhood of b and differentiability of the LP value function at b follows. On the other hand, whenever the value function is non-differentiable, the problem has multiple optimal dual solutions and every optimal basic solution to the primal problem is degenerate. These observations result in the following characterization of the differentiability of the LP value function.

Proposition 2 *(Bazaraa et al., 1990) If z_C is differentiable at $\hat{b} \in K$, then the gradient of z_C at $\hat{b}$ is the unique $\nu \in S_D$ such that $z_C(\hat{b}) = \hat{b}^\top \nu$. If $\hat{b} \in \mathrm{int}(\mathcal{K})$ is a point of non-differentiability of z_C, then there exist $\nu^1, \nu^2, \ldots, \nu^s \in S_D$ with $s > 1$ such that $z_C(\hat{b}) = \hat{b}^\top \nu^1 = \hat{b}^\top \nu^2 = \ldots = \hat{b}^\top \nu^s$ and every optimal basic solution to the associated LP with right-hand side $\hat{b}$ is degenerate.*

Example 3. In (Ex.2), we have

$$z_C(b) = \sup\{\nu b : -1 \leq \nu \leq 3, \nu \in \mathbb{R}\} = \begin{cases} 3b & \text{if } b \geq 0 \\ -b & \text{if } b < 0 \end{cases}$$

Then, $\mathcal{E} = \{\{1\}, \{2\}, \{3\}\}$ with $A_{\{1\}} = 2$, $A_{\{2\}} = -7$, and $A_{\{3\}} = 1$. The corresponding basic feasible solutions to the dual problem are 3, -1, and 5 respectively. If the value function is differentiable at $\hat{b} \in \mathbb{R}$, then its gradient at $\hat{b}$ is either -1 or 3. These extreme points describe the facets of the convex cone $\mathcal{L} = cone\{(2, 6), (-7, 7), (1, 5), (0, 1)\} = \{(b, z) \in \mathbb{R}^2 : z \geq 3b,\ z \geq -b\}$. Note that we can conclude that fixing x_1 to 0 in (Ex.2) does not affect its value function. Finally, note that $\mathcal{K} = \mathbb{R}$, i.e., $z_C(b) < \infty$ for all $b \in \mathbb{R}$. ■

We have so far examined the LP value function arising from restricting the integer variables to a fixed value and discussed that such a value function inherits the structure of a general LP value function. The LP value function, though it arises from a continuous optimization problem, has a discrete representation in terms of the extreme points and extreme directions of its dual. In the next section, we study the effect of the addition of integer variables.

4 A Characterization of the MILP Value Function

The goal of this section is to derive a discrete representation of a general MILP value function building from the results of the previous section. We observe that the MILP value function is the minimum of a countable number of translations of z_C and thus retains the same local structure as that of the continuous restriction (CR). By characterizing the set of points at which these translations occur, we arrive at Theorem 1, our discrete characterization.

From the MILP value function (Ex.1) and its continuous restriction w.r.t $\hat{x} = 0$, plotted respectively in Figures 1 and 2, we can observe that when integer variables are added to the continuous restriction, many desirable properties of the LP value function, such as convexity and continuity, may be lost. The value function in this particular example remains continuous, but as a result of the added integer variables, the function becomes piecewise linear and additional points of non-differentiability are introduced. In general, however, even continuity may be lost in some cases. Let us consider another example.

Example 4. Consider

$$
\begin{aligned}
z(b) = \inf\ & x_1 - \frac{3}{4}x_2 + \frac{3}{4}x_3 + \frac{5}{2}x_4 \\
\text{s.t. } & \frac{5}{4}x_1 - x_2 + \frac{1}{2}x_3 + \frac{1}{3}x_4 = b \\
& x_1, x_2 \in \mathbb{Z}_+, x_3, x_4 \in \mathbb{R}_+.
\end{aligned} \tag{Ex.4}
$$

Figure 3 shows this value function. As in (Ex.1), the value function is piecewise linear; however, in this case, it is also discontinuous. More specifically, it is a lower semi-continuous function. The next result formalizes these properties. ■

Proposition 3 *(Nemhauser and Wolsey, 1988; Bank et al., 1983) The MILP value function* (MVF) *is lower semi-continuous, subadditive, and piecewise polyhedral over B.*

Characterizing a piecewise polyhedral function amounts to determining its points of discontinuity and non-differentiability. In the case of the MILP value function, these points are determined by properties of the continuous restriction, which has already been introduced, and a second problem,

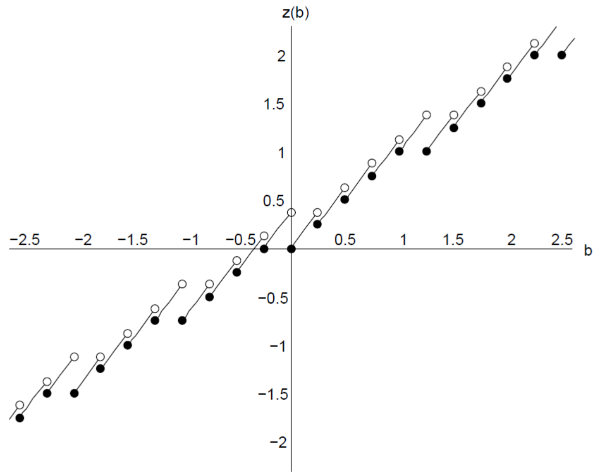


Figure 3: Value Function (Ex.4).

called the *integer restriction*, obtained by fixing the continuous variables to zero. This problem is defined as follows.

$$\begin{aligned} z_I(b) = \ & \inf c_I^\top x \\ & s.t.\ A_I x = b \\ & x \in \mathbb{Z}^r_+. \end{aligned} \tag{IR}$$

The role of the integer restriction in characterizing the value function will become clear shortly, but we first need to introduce some additional concepts.

Recalling that the continuous restriction for any $\hat{x} \in S_I$ can be expressed as $\bar{z}(b;\hat{x}) = c_I^\top \hat{x} + z_C(b - A_I\hat{x})$, we obtain the following representation of (MVF) in terms of the continuous restriction:

$$z(b) = \inf_{x \in S_I} c_I^\top x + z_C(b - A_I x) = \inf_{x \in S_I} \bar{z}(b;x) = \inf_{\hat{b} \in B_I} z(\hat{b}) + z_C(b - \hat{b}) \ \ \forall b \in B. \tag{4.1}$$

This shows that the MILP value function can be represented as a countable collection of value functions of continuous restriction functions arising from translations of the LP value function z_C. Describing the value function consists essentially of characterizing the minimal set of points at which such translations must be located to yield the entire function. The points at which translations may potentially be located can be thought of as corresponding to vectors $x \in S_I$, as in the first two equation above, though more than one member of S_I may specify the same location. Equivalently, we can also consider describing the function simply by specifying its value at points in B_I, as in the third equation above, which makes the correspondence one-to-one. Despite being finite under the assumption that B_I is finite, this characterization is nevertheless still quite impractical, as both S_I and B_I may be very large. As one might guess, it is not necessary to consider all members of B_I in order to obtain a complete representation. Later in this section, we characterize the subset of B_I necessary to guarantee a complete description. This characterization provides a key insight that leads eventually to our algorithm for construction.

Before moving on, we provide some examples that illustrate how the structure of z_C influences the structure of (MVF). First, we examine the significance of the domain of z_C in the structure and the continuity of the MILP value function with the following example.

Example 5. Consider again the value function (Ex.4). Its continuous restriction w.r.t $\hat{x} = 0$ is

$$\begin{aligned} z_C(b) = \inf \ & \frac{3}{4}x_1 + \frac{5}{2}x_2 \\ \text{s.t.} \ & \frac{1}{2}x_1 + \frac{1}{3}x_2 = b \\ & x_1, x_2 \in \mathbb{R}_+. \end{aligned}$$

Equivalently,

$$z_C(b) = \sup\{\nu b : \nu \leq \frac{3}{2}, \nu \in \mathbb{R}\}. \tag{4.2}$$

Here, the positive linear span of $\{\frac{1}{2}, \frac{1}{3}\}$ is $\mathcal{K} = \mathbb{R}_+$. We also have $z_C(b) = \frac{3}{2}b$ for all $b \in \mathcal{K}$. The gradient of $z_C(b)$ at any $b \in \mathbb{R}_+ \backslash \{0\}$ is $\frac{3}{2}$, which is the extreme point of the feasible region of (4.2). Note that for $b \in \mathbb{R}_-$, $z_C(b) = +\infty$ because the continuous restriction w.r.t the origin is infeasible whenever $b \in \mathbb{R}_-$ and its corresponding dual problem is therefore unbounded. However, in the modification of this problem in (Ex.4), we have $B = \mathbb{R}$, while $\mathcal{K}$ remains $\mathbb{R}_+$. This is because the additional integer variables result in translations of $\mathcal{K}$ into $\mathbb{R}_-$. These translations result in the discontinuity of the value function observed in (Ex.4). ■

The next result shows that the continuous restriction with respect to any fixed $\hat{x} \in S_I$ bounds the value function from above, as it is a restriction of the value function by definition.

Proposition 4 *For any $\hat{x} \in S_I$, $\bar{z}(\cdot; \hat{x})$ bounds z from above.*

Proof. For $\hat{x} \in S_I$ we have

$$\bar{z}(b; \hat{x}) = c_I^\top \hat{x} + z_C(b - A\hat{x}) \geq \inf_{x \in S_I} c_I^\top x + z_C(b - A_I x) = z(b). \qquad \square$$

The second result shows that the continuous restriction with respect to the origin coincides with the value function z over the intersection of $\mathcal{K}$ and some open ball centered at the origin. We denote an open ball with radius $\epsilon > 0$ centered at a point d by $\mathcal{N}_\epsilon(d)$.

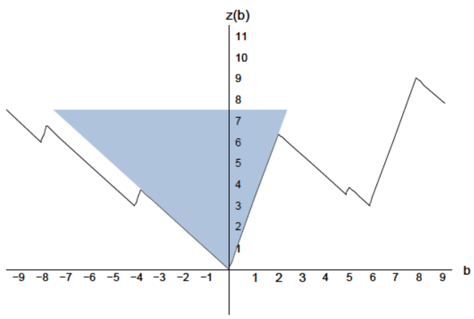
Proposition 5 *There exists $\epsilon > 0$ such that $z(b) = z_C(b)$ for all $b \in \mathcal{N}_\epsilon(0) \cap \mathcal{K}$.*

Proof. At the origin, we have $z(0) = 0$ with a corresponding optimal solution to the MILP being $(x_I^*, x_C^*) = (0, 0)$. For a given $\hat{b} \in \mathbb{R}$, as long as there exists an optimal solution $\hat{x}$ to the MILP with right-hand side b such that $\hat{x} = 0$, we must have $z(\hat{b}) = z_C(\hat{b})$. Therefore, assume to the contrary. Then for every $\epsilon > 0$, $\exists \tilde{b} \in \mathcal{N}_\epsilon(0) \cap \mathcal{K}, \tilde{b} \neq 0$ such that $z_C(\tilde{b}) > z(\tilde{b})$. Consider an arbitrary $\epsilon > 0$ and an arbitrary $\tilde{b} \in \mathcal{N}_\epsilon(0) \cap \mathcal{K}, \tilde{b} \neq 0$ such that $z_C(\tilde{b}) > z(\tilde{b})$. Then if $\tilde{x}$ is a corresponding optimal solution to the MILP with right-hand side $\tilde{b}$, we must have $\hat{x} \neq 0$. Let E and $\hat{E}$ denote the set of column indices of sub-matrices of A_C corresponding to optimal bases of the continuous restrictions at 0 and $\hat{x}$, respectively (note that both must exist).
Case i. $E = \hat{E}$. We have

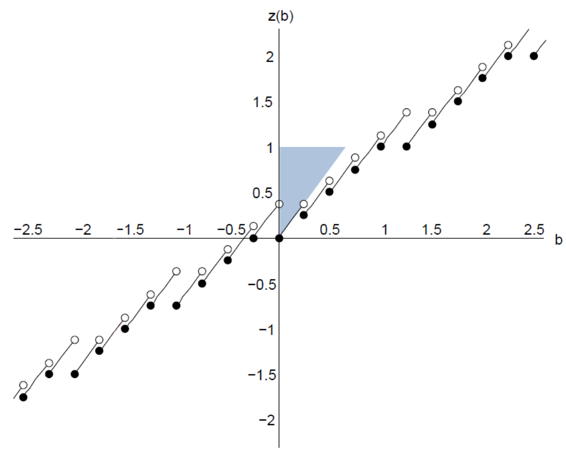
$$\begin{aligned} z_C(\hat{b}) > z(\hat{b}) &\Rightarrow c_E^\top A_E^{-1} \hat{b} > c_I^\top \hat{x} + c_{\hat{E}}^\top A_{\hat{E}}^{-1} \hat{b} - c_{\hat{E}}^\top A_{\hat{E}}^{-1} A_I \hat{x} \\ &\Rightarrow 0 > c_I^\top \hat{x} - c_{\hat{E}}^\top A_{\hat{E}}^{-1} A_I \hat{x}. \end{aligned}$$

However, the last inequality implies that at the origin, $(\hat{x}, A_{\hat{E}}^{-1} A_I \hat{x})$ provides an improved solution so that $z(0) < 0$, which is a contradiction.
Case ii. $A_E \neq A_{\hat{E}}$. We have $z_C(\hat{b}) = c_E^\top A_E^{-1} \hat{b} > c_{\hat{E}}^\top A_{\hat{E}}^{-1} \hat{b}$, which is a contradiction of the fact that z_C is the value function of the continuous restriction at 0.■

Example 6. Figure 4a shows that the epigraph of the value function of (Ex.1) coincides with the cone $\text{epi}(z_C) = \mathit{cone}\{(2,6), (-7,7), (0,1)\}$ on $\mathcal{N}_{2.125}(0)$. Similarly, Figure 4b demonstrates that the epigraph of the discontinuous value function (Ex.4) coincides with $\text{epi}(z_C) = \mathit{cone}\{(\frac{1}{2}, \frac{3}{4}), (\frac{1}{3}, \frac{5}{2}), (0,1)\}$ on $\mathcal{N}_{0.25}(0) \cap \mathcal{K} = [0, 0.25) \subseteq \mathbb{R}^+$. ■



(a)



(b)

Figure 4: The MILP value function and the epigraph of the (CR) value function at the origin.

The characterization of the value function we proposed in (4.1) is finite as long as the set S_I is finite. However, there are cases where the set

$$B_I = \{b \in B : S_I(b) \neq \emptyset\}$$

is finite, while S_I remains infinite. Clearly in such cases, there is a finite representation of the value function that (4.1) does not provide. We can address this issue by representing the value function in terms of the set B_I rather than the set S_I, but even then, the representation is not minimal, as not all members of B_I are necessary to the description. We next study the properties of the minimal subset of B_I that can fully characterize the value function of a MILP.

From the previous examples, we can observe that when the MILP has only a single constraint and the value function is thus piecewise linear, the points necessary to describe the function are the *lower break points*. To generalize the notion of lower break points to higher dimension, we need some additional machinery.

In Figure 4, the lower break points are also local minima of the MILP value function and one may be tempted to conjecture that knowledge of the local minima is enough to characterize the value function. Unfortunately, it is easy to find cases for which the value function has no local minima and yet still has the nonconvex structure characteristic of a general MILP value function. Consider the following example.

Example 7. Consider

$$
\begin{aligned}
z(b) = \inf\ & -2x_1 + 6x_2 - 7x_3 \\
\text{s.t. } & x_1 - 2x_2 + 7x_3 = b \\
& x_1 \in \mathbb{Z}_+, x_2, x_3 \in \mathbb{R}_+.
\end{aligned}
\tag{Ex.6}
$$

As illustrated in Figure 5, the extreme point of the epigraph of the continuous restriction of the problem does not correspond to a local minimum. In fact the value function does not have any local minima. ■

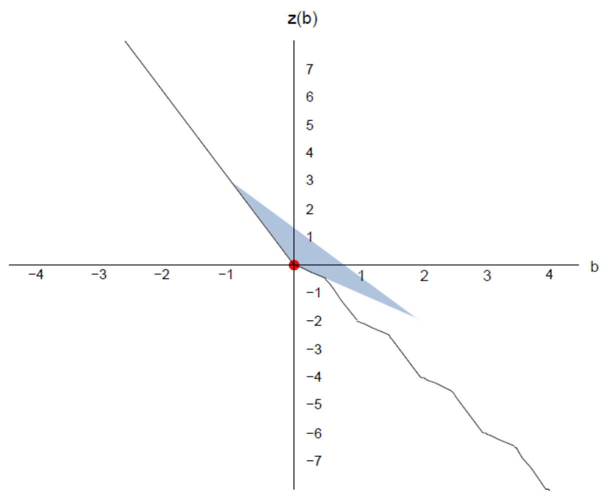


Figure 5: MILP value function (Ex.6) with no local minimum.

In the previous examples, the epigraph of z_C was also always a pointed cone. As a result, the MILP value function had lower break points that corresponded to the extreme points of $\text{epi}(\bar{z}(\cdot; x))$ for certain $x \in S_I$. However, the cone $\text{epi}(z_C)$ may not have an extreme point in general. When it fails to have one, in the single-dimensional case, the MILP value function will be linear and will have no break points. Consider the following example.

Example 8. Consider

$$
\begin{aligned}
z(b) = \inf\ & 2x_1 + 6x_2 - 7x_3 \\
\text{s.t. } & x_1 - 6x_2 + 7x_3 = b \\
& x_1 \in \mathbb{Z}_+, x_2, x_3 \in \mathbb{R}_+.
\end{aligned}
\tag{Ex.7}
$$

In this example, the value function (Ex.7) coincides with the value function of the continuous restriction w.r.t the origin. This function is plotted in Figure 6. ■

In this last example, the epigraph of the value function contains a line that passes through the origin. This property can be generalized to any dimension. If $\text{epi}(z_C)$ is not a pointed cone, then

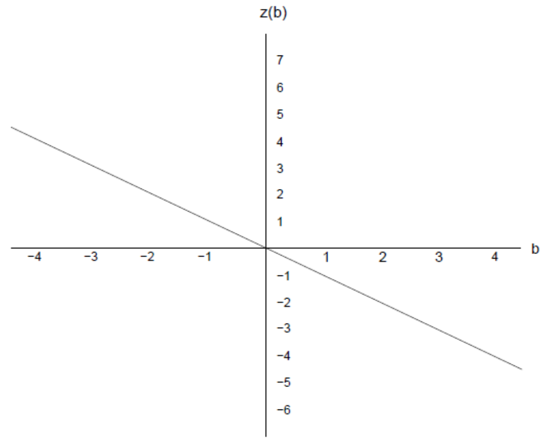


Figure 6: Linear and convex MILP and CR value functions to (Ex.7).

for any given $\hat{x} \in S_I$, the boundary of the epigraph of $\bar{z}(\cdot;\hat{x})$ contains a line that passes through $(A_I\hat{x}, \bar{z}(A_I\hat{x};\hat{x}))$. The boundary of the resulting MILP value function therefore contains parallel lines that result from translations of $\bar{z}$. Clearly, to characterize such a value function, one would need to have, for each such line, a point $\hat{b}$ such that $(\hat{b}, z(\hat{b}))$ is on both the line and the value function of the continuous restriction, z_C. This case, in which $\mathrm{epi}(z_C)$ is not a pointed cone, is, however, an edge case and its consideration would complicate the presentation substantially. For the remainder of this section, we therefore assume the more common case in which $\mathrm{epi}(z_C)$ is a pointed cone.

To generalize the set of lower break points to higher dimension, we introduce the notion of *points of strict local convexity* of the MILP value function. We denote the set of these points by B_{SLC}.

Definition 1 *A point $\hat{b} \in B_I$ is a* point of strict local convexity *of the function $f : \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$ if for some $\epsilon > 0$ and $g \in \mathbb{R}^m$, we have*

$$f(b) > f(\hat{b}) + g^\top (b - \hat{b}) \text{ for all } b \in \mathcal{N}_\epsilon(\hat{b}),\ b \neq \hat{b}.$$

This definition requires the existence of a hyperplane that is tangent to the function f at the point $\hat{b} \in B_I$, while lying strictly below f in some neighborhood of $\hat{b}$. For the continuous restriction with respect to $\hat{x} \in S_I$, this can happen only at the extreme point of the epigraph of the function, if such a point exists. Note that at such a point, we must have $\bar{z}(\hat{b};\hat{x}) = c_I^\top \hat{x}$. Furthermore, if $\hat{x} \in \arg\inf_{x \in S_I} \bar{z}(\hat{b};x)$, then we will also have $\bar{z}(\hat{b};\hat{x}) = z_I(\hat{b})$.

Proposition 6 *For a given $\hat{x} \in S_I$, $\hat{b} \in A_I\hat{x} + \mathcal{K}$ is a point of strict local convexity of $\bar{z}(\cdot;\hat{x})$ if and only if $(\hat{b}, \bar{z}(\hat{b};\hat{x}))$ is the extreme point of* $\mathrm{epi}(\bar{z}(\cdot;\hat{x}))$.

Proof. Let $\hat{x} \in S_I$ and $\hat{b} \in A_I\hat{x} + \mathcal{K}$ be given as in the statement of the theorem. We use the following property in the proof. Let the function H_t be defined by

$$H_t(b) = \begin{cases} c_I^\top \hat{x} + (b - A_I\hat{x})^\top \eta^t & \text{for } b \in \mathcal{K}, \\ +\infty & \text{otherwise,} \end{cases}$$

where $\eta^t \in \{\nu^i\}_{i\in K} \cup \{d^j\}_{j\in L}$. Then, we have

$$\bar{z}(b;\hat{x}) = \sup_{t\in K\cup L} H_t(b)$$

Moreover,

$$\partial\bar{z}(\hat{b};\hat{x}) = \mathrm{conv}(\{\nabla H_1,\ldots,\nabla H_p\}_{p\in P}) = \mathrm{conv}(\{\eta^1,\ldots,\eta^p\}_{p\in P}) \neq \emptyset,$$

where $P \subseteq K \cup L$ and $|P| > 1$ and finally, we have that

$$\bar{z}(\hat{b};\hat{x}) = H_1(\hat{b}) = \cdots = H_p(\hat{b}) \text{ for } p \in P. \tag{4.3}$$

($\Rightarrow$) Let ϵ and g be the radius of the ball and a corresponding subgradient showing the strict local convexity of $\bar{z}(\cdot;\hat{x})$ at $\hat{b}$. If $\bar{z}(\cdot;\hat{x})$ is differentiable at $\hat{b}$, then $\exists\nu \in \mathbb{R}^m$ such that $\partial\bar{z}(\cdot;\hat{x}) = \{\nu\}$, and therefore $g = \nu$. Then we trivially have that $\hat{b}$ cannot be a point of strict local convexity of $\bar{z}(\cdot;\hat{x})$, as there always exists ϵ' with $0 < \epsilon' < \epsilon$ such that on $\mathcal{N}_{\epsilon'}(\hat{b})$, we have $\bar{z}(b;\hat{x}) = \bar{z}(\hat{b};\hat{x}) + \nu^\top(b - \hat{b})$. Therefore, $\bar{z}(\cdot;\hat{x})$ cannot be differentiable at $\hat{b}$.

Since $\bar{z}(\cdot;\hat{x})$ is not differentiable at $\hat{b}$, there are $H_1,\ldots,H_p, p \in P$, as defined above. In the case that $p > m$, from the discussion in Section 3, $\hat{b}$ has to be the extreme point of $\mathrm{epi}(\bar{z}(\cdot;\hat{x}))$. Next, we show that $\hat{b}$ cannot be the extreme point of $\mathrm{epi}(\bar{z}(\cdot;\hat{x}))$ if $p \leq m$.

When $1 < p \leq m$, equation (4.3) must still hold. Let

$$R = \{(b, \bar{z}(b;\hat{x})) \in (A_I\hat{x} + \mathcal{K}) \times \mathbb{R} : \bar{z}(b;\hat{x}) = H_1(b) = \cdots = H_p(b) \text{ for } p \in P\}$$

Then there exists $\tilde{b} \in N_\epsilon(\hat{b})$ such that $(\tilde{b}, z(\tilde{b})) \in R$ and $\tilde{b} \neq \hat{b}$. We have

$$\bar{z}(\tilde{b};\hat{x}) - \bar{z}(\hat{b};\hat{x}) = (\tilde{b} - \hat{b})^\top\eta^t, t \in P.$$

Then we can conclude that for $g \in \partial\bar{z}(\hat{b};\hat{x}) = \mathrm{conv}(\{\eta^1,\ldots,\eta^p\})$, the function $\bar{z}(\hat{b};\hat{x}) + g^\top(\tilde{b}-\hat{b})$ also coincides with $\bar{z}(\tilde{b};\hat{x})$ as follows. Choose $0 \leq \lambda^t \leq 1, t \in P$ such that $g = \sum_{t\in P}\lambda^t\eta^t, \sum_{t\in P}\lambda^t = 1$. From the equations

$$\lambda^t(\bar{z}(\tilde{b};\hat{x}) - \bar{z}(\hat{b};\hat{x})) = \lambda^t(\tilde{b} - \hat{b})^\top\eta^t, t \in P$$

we have a contradiction to $\hat{b}$ being the point of strict local convexity of $\bar{z}(\cdot;\hat{x})$, since

$$\bar{z}(\tilde{b};\hat{x}) - \bar{z}(\hat{b};\hat{x}) = \sum_{t=1}^{p}\lambda^t(\tilde{b} - \hat{b})^\top\eta^t = g^\top(\tilde{b} - \hat{b}).$$

($\Leftarrow$) Since $(\hat{b}, z(\hat{b}))$ is the extreme point of $\mathrm{epi}(\bar{z}(\cdot;\hat{x}))$, then $\partial\bar{z}(\hat{b};\hat{x}) = \mathrm{conv}(\{\eta^1,\ldots,\eta^p\})$, where $p \in P$ and we must have that $|P| > m$. Choose $g \in \mathrm{int}(\mathrm{conv}(\{\eta^1,\ldots,\eta^p\}))$. For an arbitrary $\tilde{b} \in A_I\hat{x} + \mathcal{K}$, $\tilde{b} \neq \hat{b}$ there exists $\tilde{\eta} \in \{\eta^1,\ldots,\eta^p\}$ such that $\bar{z}(\tilde{b};\hat{x}) = (\tilde{b} - A_I\hat{x})^\top\tilde{\eta}$. Then, from the monotonicity of the subgradient of a convex function we have $(\tilde{\eta}^\top - g^\top)(\tilde{b} - \hat{b}) > 0$. Therefore,

$$\bar{z}(\tilde{b};\hat{x}) = \bar{z}(\hat{b};\hat{x}) + \tilde{\eta}^\top(\tilde{b} - \hat{b}) > \bar{z}(\hat{b};\hat{x}) + g^\top(b - \hat{b}) \quad \forall\tilde{b} \in A_I\hat{x} + \mathcal{K},\ \tilde{b} \neq \hat{b}. \tag{4.4}$$

That is, $\hat{b}$ is a point of strict local convexity of $\bar{z}(\cdot;\hat{x})$. ■

Example 9. Consider the MILP in Example 4. The blue shaded region in Figure 7 is $\mathrm{epi}(\bar{z}(\cdot;1))$. The point $(1,-2)$ is the extreme point of the cone $\mathrm{epi}(\bar{z}(b;1))$ and $\hat{b} = 1$ is a point of strict local convexity of the value function. ■

Next, we discuss the points of strict local convexity of the MILP value function.

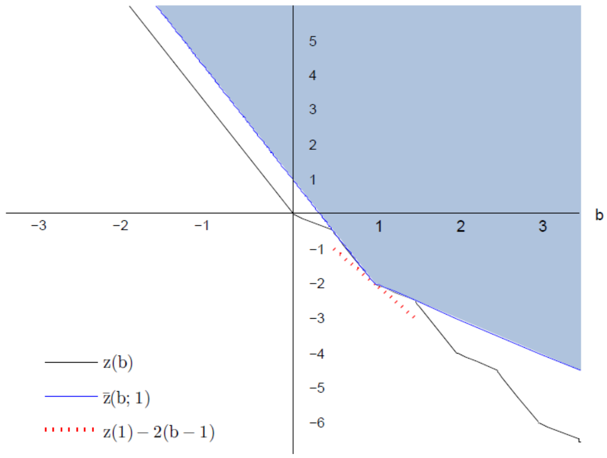


Figure 7: The value function of the MILP in (Ex.6)

Proposition 7 *If $\hat{b}$ is a point of strict local convexity, then there exists $\hat{x} \in S_I$ such that*

- $\hat{b} = A_I\hat{x}$;
- $(\hat{b}, \bar{z}(\hat{b};\hat{x}))$ *is the extreme point of* $\mathrm{epi}(\bar{z}(\cdot;\hat{x}))$; *and*
- $\bar{z}(\hat{b};\hat{x}) = c_I^\top \hat{x} = z_I(\hat{b}) = z(\hat{b})$.

Proof. Let $\hat{b}$ be a point of strict local convexity. If there exists $\hat{x} \in S_I(\hat{b})$ such that $\hat{x} \in \arg\inf_{x \in S_I} \bar{z}(\hat{b}; x)$, then we have that $c_I^\top x = z(\hat{b})$. The remainder of the statement is trivial in this case. Consider the case where such x does not exist. That is, for any $(x, y) \in S(\hat{b})$ such that $c_I^\top x + c_C^\top y = z(\hat{b})$, we have $y > 0$. Let one such point be $(\hat{x}, \hat{y})$. Consider $\epsilon > 0$ used to show $\hat{b}$ is a point of strict local convexity. If $\bar{z}(\cdot;\hat{x})$ coincides with z on $\mathcal{N}_\epsilon(\hat{b})$, then from from Proposition 6 it follows that $\hat{b}$ cannot be a point of strict local convexity. On the other hand, if z is constructed by multiple translations of $\bar{z}$ over $\mathcal{N}_\epsilon(\hat{b})$, since it attains the minimum of these functions, there cannot be a supporting hyperplane to z at $\hat{b}$, therefore $\hat{b}$ cannot be a point of strict local convexity. ■

We note that the reverse direction of Proposition 7 does not hold. In particular, it is possible that for some $\hat{x} \in S_I$ we have $z(A_I\hat{x}) = c_I^\top \hat{x}$, but that $A_I\hat{x}$ is not a point of strict local convexity. For instance, in example 1, for $\hat{x} = (1, 0, 1)$ we have that $A_I\hat{x} = 2$ and that $\bar{z}(2;\hat{x}) = z(2) = 6$. Nevertheless, 2 is not a point of strict local convexity.

Points of strict local convexity may lie on the boundary of B_I. The next example illustrates a case where this happens.

Example 10. Consider the MILP value function

$$\begin{aligned} z(b) = \inf \;& -x_1 + 3x_2 \\ \text{s.t. } & x_1 - 3x_2 = b \\ & x_1 \in \mathbb{Z}_+,\; x_2 \in \mathbb{R}_+. \end{aligned} \tag{Ex.9}$$

shown in Figure 8. If we artificially impose the additional restriction that $b \in [0, 2]$ for the purposes

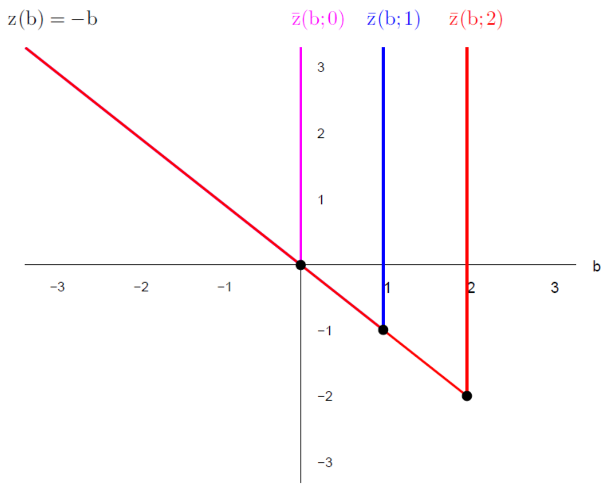


Figure 8: MILP Value Function of (Ex.9) with $B_I = [0, 2]$.

of illustration, it is clear that there is no point of strict local convexity in the interior of B_I, although epi(z_C) is a pointed cone. ■

Let us further examine the phenomena illustrated by the previous example. For a given $\hat{x} \in S_I$, let $\hat{b} = A_I \hat{x}$. We know that the single extreme point of epi$(\bar{z}(\cdot; \hat{x}))$ is $(\hat{b}, \bar{z}(\hat{b}; \hat{x}))$ and that there must therefore be $m + 1$ (2 in this example) facets of epi$(\bar{z}(\cdot; \hat{x}))$ whose intersection is this single extreme point. Now, if $\hat{b}$ is not a point of strict local convexity, then on any $\mathcal{N}_\epsilon(\hat{b})$ with $\epsilon > 0$, at most m facets of epi$(\bar{z}(\cdot; \hat{x}))$ coincide with the facets of the epigraph of the value function. This means that there exists a direction in which z is affine in the neighborhood of $(\hat{b}, \bar{z}(\hat{b}; \hat{x}))$; that is, $\hat{b}$ cannot be a point of strict local convexity of z. Given that the set B_I is assumed to be bounded, the value function must contain a point $(\tilde{b}, z(\tilde{b}))$ such that $\tilde{b} \in \text{bd}(conv(B_I)) \cap B_I$ along a line in this direction. Let $\bar{bd}(B_I) = \text{bd}(conv(B_I)) \cap B_I$. Since epi($z_C$) is pointed, then $\tilde{b}$ has to be a point of strict local convexity of z. This latter point is the one needed to describe the value function—the epigraph of the associated continuous restriction associated with $\tilde{b}$ contains the continuous restriction w.r.t. $\hat{x}$, which means that $A\hat{x}$ is not contained in the minimal set of points at which we need to know the value function.

We are now almost ready to formally state our main result. So far, we have discussed certain properties of the points of strict local convexity and showed that such points can belong to the interior or boundary of B_I. Our goal is to show that the set B_{SLC}, which was previously defined to the set of all points of strict local convexity, is precisely the minimal subset of B_I, denoted in the following result by B_{min}, needed to characterize the full value function. Let us now formally define B_{min} to be a minimal subset of B_I such that

$$z(b) = \inf_{\hat{b} \in B_{min}} z(\hat{b}) + z_C(b - \hat{b}) \quad \forall b \in B. \tag{4.5}$$

Then we have the following result.

Proposition 8 $B_{min} = B_{SLC}$.

Proof. First, we show that if $\hat{b} \in B \backslash B_{SLC}$, then it is not in the set B_{min}. If $\hat{b} \in B \backslash B_I$, then from (4.1) it follows that $\hat{b}$ is not necessary to describe the value function, then $\hat{b} \notin B_{min}$. Consider

$\hat{b} \in B_I \backslash B_{SLC}$. Let $\hat{x} \in S_I(\hat{b})$ such that $c_I^\top \hat{x} = z(\hat{b})$. Since epi($z_C$) is assumed to be pointed, we have $z(\hat{b}) = \min\{\bar{z}(\hat{b}; x_1), \bar{z}(\hat{b}; x_2),$
$\ldots, \bar{z}(\hat{b}; x_k)\}$, where $k > 1$ and $x_1, \ldots, x_k \in S_I$. Then, for some $l = 1, \ldots, k$ and $x_l \neq \hat{x}$ we have $\min\{\bar{z}(\hat{b}; \hat{x}), \bar{z}(\hat{b}; x_l)\} = \bar{z}(\hat{b}; x_l)$ and it follows that $\hat{b} \notin B_{min}$. Therefore, if $\hat{b} \in B_{min}$, then $\hat{b} \in B_{SLC}$.

We next show that if $\hat{b} \in B_{SLC}$, then $\hat{b} \in B_{min}$. Let us denote by $S'(\hat{b})$ the set of points $x \in S_I$ such that $(A_I x, c_I^\top x)$ coincides with the value function at $\hat{b}$. If $\hat{b} \notin B_{min}$, then all the points in $S'(\hat{b})$ can be eliminated from the description of the value function in (4.1). That is, we have

$$z(\hat{b}) = \inf_{x \in S_I \backslash S'(\hat{b})} \bar{z}(\hat{b}; x).$$

Therefore, for any pair $(x, y) \in S(\hat{b})$ that is an optimal solution to the MILP with right-hand side fixed at $\hat{b}$, we have $y > 0$. This, however, contradicts with Proposition 7 and we have that $\hat{b}$ cannot be a point of strict local convexity of z. ■

Because it will be convenient to think of the value function as being described by a subset of S_I, rather than as a subset of B_I, we now express our main result in those terms. From Proposition 8, it follows that there is a subset of S_{min} of S_I that can be used to represent the value function, as shown in the following theorem. Note, however, that while B_{min} is unique, S_{min} is not.

Theorem 1 (Discrete Representation) *Let S_{min} be any minimal subset of S_I such that for any $b \in B_{min}$, $\exists x \in S_{min}$ such that $A_I x = b$ and $c_I^\top x = z(b)$. Then for $b \in B$, we have*

$$z(b) = \inf_{x \in S_I} \bar{z}(b; x) = \inf_{x \in S_{min}} \bar{z}(b; x). \tag{4.6}$$

Proof. The proof follows from Proposition 8, noting that a point $\hat{x} \in S_I$ such that $c_I^\top x > z(A_I x)$ cannot be necessary to describe the value function. ■

Example 11. We apply the theorem to (Ex.1). In this example, over $b \in [-9, 9]$, we have that $B_{min} = \{-8, -4, 0, 5, 6, 10\}$ and $S_{min} = \{[0; 0; 2], [0; 0; 1], [0; 0; 0], [0; 1; 0], [1; 0; 0], [0; 2; 0]\}$. Clearly, the knowledge of the latter set is enough to represent the value function. ■

Theorem 1 provides a minimal subset of S_I required to describe the value function. We discuss in Section 7 that constructing a minimal such subset exactly may be difficult. Alternatively, we propose an algorithm to approximate S_{min} (with a superset that is thus still guaranteed to yield the full value function). This has proven empirically to be a close approximation. Before further addressing the practical matter of how to generate the representation, we discuss some theoretical properties of the value function that arise from our result so far in the next two sections. The reader interested in the computational aspects of constructing the value function can safely skip to Section 7 for the proposed algorithm, as that algorithm does not depend on the results in the following two sections.

5 Local Stability Regions

In this section, we demonstrate that certain structural properties of the value function, such as regions of convexity and points of non-differentiability and discontinuity, can also be characterized in the context of our representation. We show that there is a one-to-one correspondence between regions over which the value function is convex and continuous—the so-called *local stability sets*—and the set B_{min}. We also provide results on the relationships between this set and the sets of non-differentiability and discontinuity of the value function.

We start this section by introducing notation for the sets of right-hand sides with particular properties.

Definition 2

- $B_{LS}(\hat{b}) = \{b \in B : z(b) = z(\hat{b}) + z_C(b - \hat{b})\}$ *is the* local stability set *w.r.t* $\hat{b} \in B$*;*
- $B_{ES}(\hat{b}) = \text{bd}(B_{LS}(\hat{b}))$ *is the* local boundary set *w.r.t* $\hat{b} \in B$*;*
- $B_{ES} = \cup_{b \in B_{min}} B_{ES}(b)$ *is the* boundary set*;*
- $B_{ND} = \{b \in B : z \textit{ is not differentiable at } b\}$ *is the* non-differentiability set*; and*
- $B_{DC} = \{b \in B : z \textit{ is discontinuous at } b\}$ *is the* discontinuity set.

Example 12. To illustrate the above definitions, consider the value function in Example 1. Let $\hat{b} = 3$. Over the interval $[-9, 9]$ we have that the function $z(3) + z_C(b - 3)$ coincides with z at $b \in B_{LS}(\hat{b}) = [2.125, 3]$. Then, $B_{ES}(\hat{b}) = \{2.125, 3\}$. The minimal set is $B_{min} = \{-8, -4, 0, 5, 6\}$. The boundary set consists of the union of the local boundary sets w.r.t. minimal points; i.e.,

$$\begin{aligned} B_{ES} =& \{\{-9, -7.75\} \cup \{-7.75, -3.75\} \cup \{-3.75, 2.125\} \cup \{2, 125, 5.125\} \cup \{5.125, 8\}\} \\ =& \{-9, -7.75, -3.75, 2.125, 5.125, 8\}. \end{aligned}$$

The non-differentiablity set is

$$B_{ND} = \{-9, -8, -7.75, -4, -3.75, 0, 2.125, 5, 5.125, 6, 8, 9\}.$$

Finally, $B_{DC} = \emptyset$. ■

The main result of this section is Theorem 2. The goal is to show that the value function is convex and continuous over the the local stability sets associated with the members of B_{min}. Furthermore, in this theorem we demonstrate the relationship between the set B_{min}, the boundary set, B_{ES}, and the sets of point of non-differentiability and discontinuity of the value function. We next state the theorem.

Theorem 2

i. Let $\hat{b} \in B$.

- *There exists* $x^* \in S_{min}$ *such that for any* $\tilde{b} \in \text{int}(B_{LS}(\hat{b}))$*, there exists* $y \in \mathbb{R}^{n-r}_+$ *such that* (x^*, y) *is an optimal solution to the MILP with right-hand side* $\tilde{b}$.

– z *is continuous and convex over* $\text{int}(B_{LS}(\hat{b}))$.

ii. $\hat{b} \in B_{ES}$ *if and only if for any* $\epsilon > 0$, $\nexists x^* \in S_I$ *such that* $z(b) = c_I^\top x^* + z_C(b - A_I x^*)$ *for all* $b \in \mathcal{N}_\epsilon(\hat{b})$.

iii. Let $\hat{b} \in B_{min}$. *Then,* $\text{int}(B_{LS}(\hat{b}))$ *is the maximal set of right-hand sides containing* $\hat{b}$ *over which the value function is convex and continuous.*

iv. For the general MILP value function, we have $B_{min} \subseteq B_{ND}$ *and* $B_{ES} \subseteq B_{ND}$. *Furthermore, if the MILP value function is discontinuous, we have* $B_{min} \subseteq B_{DC} \subseteq B_{ES} \subseteq B_{ND}$.

Proof. We build to the proof of the theorem, which constitute the remainder of this section, by proving lemmas 1–8, The first and second parts of the theorem follow from lemma 1 and lemma 2. The third part of the theorem is shown in lemma 3. The last part follows from lemmas 5–8.

In the first lemma, we show properties of the function on differentiable regions within local stability sets.

Lemma 1 *Let* $\hat{b} \in B$. *Then there exists* $x^* \in S_I$ *such that for any* $\tilde{b} \in \text{int}(B_{LS}(\hat{b}))$, *there exists* $y \in \mathbb{R}_+^{n-r}$ *such that* (x^*, y) *is an optimal solution to the MILP with right-hand side* $\tilde{b}$. *Furthermore,* z *is continuous and convex over* $\text{int}(B_{LS}(\hat{b}))$

Proof. From Theorem 1, for any $\hat{b} \in B$ there exists $x^* \in S_{min}$ such that $\text{int}(B_{LS}(\hat{b})) = \text{int}(\{b \in B : z(b) = c_I^\top x^* + z_C(b - A_I x^*)\})$. Therefore, for any $\tilde{b} \in \text{int}(B_{LS}(\hat{b}))$, $z(\tilde{b}) = c_I^\top x^* + c_C^\top y^*$ where $y^* = \text{argmin}\{c_C^\top y : A_C y = \tilde{b} - A_I x^*, y \in \mathbb{R}_+^{n-r}\}$. The convexity and continuity of z on $\text{int}(B_{LS}(\hat{b}))$ follows trivially. ■

Corollary 1 *If* z *is differentiable over* $\mathcal{N} \subseteq B$, *then there exist* $x^* \in S_I$ *and* $E \in \mathcal{E}$ *such that* $z(b) = c_I^\top x^* + \nu_E^\top (b - A_I x^*)$ *for all* $b \in \mathcal{N}$.

Proof. Let an arbitrary $\hat{b} \in \mathcal{N}$ be given. By Theorem 1, we know that there exists $\hat{x}^* \in S_{min}$ such that $z(\hat{b}) = \bar{z}(\hat{b}; \hat{x})$ and $A_I x^* \in B_{min}$. Then, we have $z(\hat{b}) = c_I^\top x^* + \nu_E^\top (\hat{b} - A_I x^*)$ with $E \in \mathcal{E}$ and there exists (x^*, x_E, x_N), an optimal solution to the given MILP with right-hand side $\hat{b}$, where x_E and x_N correspond to the basic and non-basic variables in the corresponding solution to the continuous restriction w.r.t. x^*. It follows that the vector $(x^*, x_E + A_E^{-1}(b - \hat{b}), x_N)$ is a feasible solution for any $b \in \mathcal{N}$.

Now, let another arbitrary point $\tilde{b} \in \mathcal{N}$ be given. We show that $(x^*, x_E + A_E^{-1}(\tilde{b} - \hat{b}), x_N)$ must be an optimal solution for right-hand side $\tilde{b}$. Since $\hat{b} \in \mathcal{N}, \tilde{b} \in \mathcal{N}$ and z is differentiable over $\mathcal{N}$, then ν_E is the unique optimal dual solution to the continuous restriction by Proposition 2 and we have

$$
\begin{aligned}
z(\tilde{b}) &= c_I^\top x^* + c_E^\top (x_E + A_E^{-1}(\tilde{b} - \hat{b})) + c_N^\top x_N \\
&= z(\hat{b}) + \nu_E^\top (\tilde{b} - \hat{b}) = c_I^\top x^* + \nu_E^\top (\hat{b} - A_I x^*) + \nu_E^\top (\tilde{b} - \hat{b}) \\
&= c_I^\top x^* + \nu_E^\top (\tilde{b} - A_I x^*) = \bar{z}(\tilde{b}; x^*).
\end{aligned}
$$

Since $\tilde{b}$ and $\hat{b}$ were arbitrary points in $\mathcal{N}$, the result holds for all such pairs and this ends the proof. ■

It follows from the previous result that if the value function is differentiable over $\mathcal{N} \subseteq B$, then its gradient at every right-hand side in $\mathcal{N}$ is a unique optimal dual solution to the continuous restriction problem w.r.t. some $x^* \in S_I$. This generalizes Proposition 2 on the gradient of the function at a differentiable point of the LP value function to the mixed integer case. As an example, in (Ex.4) the gradient of z at any differentiable point is $\nu = \frac{3}{2}$. Next, we show the second part of Theorem 2 in the following result.

Lemma 2 *$\hat{b} \in B_{ES}$ if and only if for any $\epsilon > 0$, $\nexists x^* \in S_I$ such that $z(b) = c_I^\top x^* + z_C(b - A_I x^*)$ for all $b \in \mathcal{N}_\epsilon(\hat{b})$.*

Proof. ($\Rightarrow$) Let $\epsilon > 0$ be given and assume $\exists x^* \in S_I$ such that $z(b) = c_I^\top x^* + z_C(b - A_I x^*)$ for all $b \in \mathcal{N}_\epsilon(\hat{b})$. Now, let $\tilde{b} \in B_{min}$ be such that $\hat{b} \in B_{ES}(\tilde{b})$. Then for all $b \in \mathcal{N}_\epsilon(\hat{b})$, we have $z(b) = \bar{z}(b; x^*) = z(\tilde{b}) + z_C(b - \tilde{b})$. That is, $\hat{b} \in \text{int}(B_{LS}(\tilde{b}))$.

($\Leftarrow$) Let $\tilde{b} \in B_{min}$ be such that $b \in \text{int}(B_{LS}(\tilde{b}))$. Then from Lemma 1, there exists $x^* \in S_I$ optimal for all $b \in B_{LS}(\tilde{b})$. ■

Next, we arrive at showing the third part of Theorem 2. This is shown in the following result.

Lemma 3 *Let $\hat{b} \in B_{min}$. Then, $\text{int}(B_{LS}(\hat{b}))$ is the maximal set of right-hand sides containing $\hat{b}$ over which the value function is convex and continuous.*

Proof. Assume the contrary that $B_{LS}(\hat{b})$ with $\hat{b} \in B_{min}$ is not the maximal set. Then, there exists $\tilde{b}$ in the boundary set w.r.t $\hat{b}$, $B_{ES}(\hat{b})$, and $\epsilon > 0$ such that the value function is continuous and convex at $\mathcal{N}_\epsilon(\tilde{b})$. From Theorem 1 and Lemma 2 we have

$$z(b) = \min_{x^i \in S_{min}} \{c_I^\top x^i + (b - A_I x^i)^\top \nu^i\}, b \in \mathcal{N}_\epsilon(\tilde{b}), \tag{5.1}$$

where ν^i is the optimal dual solution to $z_C(\tilde{b} - A_I x^i)$ and the set $x^i \in S_{min}$ contains two or more distinct members. Then, z is concave over $\mathcal{N}_\epsilon(\tilde{b})$ unless all the polyhedral functions in (5.1) are the same. But then $\mathcal{N}_\epsilon(\tilde{b})$ is a subset of $B_{LS}(\hat{b})$. ■

So far, in Theorem 2 we have demonstrated that over the local stability set w.r.t a minimal point, the integer part of the solution to the MILP remains constant and the value function of the MILP is a translation of the continuous restriction value function. This can be viewed as a generalization of a similar result that the value function of a PILP with inequality constraints is constant over its local stability sets ($z_C(b) = 0$ for $b \in \mathbb{R}^m$). These regions are characterized by Schultz et al. (1998). In this case, the members of B_{min} generalize the notion of *minimal tenders* discussed in (Trapp et al., 2013).

Before showing the forth and last part of Theorem 2, we need another lemma on the necessary conditions for the continuity of the value function.

Lemma 4 *If $z_C(b) < \infty$ for all $b \in B$, then z is continuous over B.*

Proof. If $z_C(\hat{b}) < \infty$ for some $\hat{b} \in B$, then the continuous restriction w.r.t. the origin and its dual are both feasible and have optimal solution values equal to $z_C(\hat{b})$. Therefore, $\bar{z}$ is finite and continuous on B. It can be proved by induction that the minimum of countably many continuous functions defined on B is continuous on B. The continuity of z follows by the representation in (4.1). ■

We now proceed to show the last part of the theorem. Lemmas 5–8 address the relationships between the discontinuity set of the value function with the minimal set of right-hand sides, the boundary set, and the set of non-differentiability points. Combining the following lemmas, the proof of the theorem is complete.

Lemma 5 $B_{min} \subseteq B_{ND}$.

Proof. Assume the value function is differentiable at some $\hat{b} \in B_{min}$. Let $\nabla z(\hat{b}) = g$. Then, there exists some $\epsilon > 0$, such that $z(b) = z(\hat{b}) + g^\top(b - \hat{b})$ for all $b \in \mathcal{N}_\epsilon(\hat{b})$. But then, from the definition of a point of strict local convexity, $\hat{b}$ cannot be in B_{SLC} and therefore, $\hat{b} \notin B_{min}$. ■

Earlier we showed that the discontinuities of the MILP value function may only happen when it no longer attains its minimum over some translated $\bar{z}$ and a switch to another translation is required. This is used next to show the relationship between the discontinuity and boundary sets.

Lemma 6 $B_{DC} \subseteq B_{ES}$.

Proof. Assume to the contrary that there exists $\tilde{b} \in B_{DC}$ but $\tilde{b} \in \text{int}(B_{LS}(\hat{b}))$ for some $\hat{b} \in B_{min}$. Then from Theorem 2, there exists $\epsilon > 0$ such that $z(b) = z(\hat{b}) + z_C(b - \hat{b})$ for all $b \in \mathcal{N}_\epsilon(\tilde{b})$. Therefore, z can only be continuous on $\mathcal{N}_\epsilon(\tilde{b})$, which is a contradiction.■

Lemma 7 *If the value function is discontinuous, then* $B_{min} \subseteq B_{DC}$.

Proof. Since $B_{DC} \neq \emptyset$, from Lemma 4 we have $\mathcal{K} \neq \mathbb{R}^m$. Then, for any $\hat{b} \in B$ we have $\hat{b} \in B_{ES}(b)$; that is, any right-hand side lies on the boundary of its local stability set. Consider $\hat{b} \in B_{min}$. If z is continuous at $\hat{b}$ then there exists $\epsilon > 0$ and $\tilde{b} \in B_{min}$ such that $\tilde{b} \neq \hat{b}$ and for any $b_1 \in \mathcal{N}_\epsilon(\hat{b}) \backslash B_{ES}(\hat{b})$ we have $z(b_1) = z(\tilde{b}) + z_C(b_1 - \tilde{b})$. Consider $b_2 \in \mathcal{N}_\epsilon(\hat{b}) \cap B_{ES}(\hat{b})$. If $z(\tilde{b}) + z_C(b_2 - \tilde{b}) < z(b_2)$, then z cannot be the value function at b_2. On the other hand, if $z(\tilde{b}) + z_C(b_2 - \tilde{b})$ lies above or on $z(b_2)$, it can be easily shown that there cannot exist a supporting hyperplane of z at $\hat{b}$ that lies strictly below z on an arbitrarily small neighborhood of $\hat{b}$. Then $\hat{b}$ cannot be in B_{min}. ■

The next result shows that if $\hat{b}$ belongs to the boundary set w.r.t a minimal point, then z is non-differentiable at $\hat{b}$.

Lemma 8 $B_{ES} \subseteq B_{ND}$.

Proof. Assume there exist some $\tilde{b} \in B_{ES}(\hat{b}), \hat{b} \in B_{min}$ such that z is differentiable at $\tilde{b}$. Then there exists $\epsilon > 0$ and $E \in \mathcal{E}$ such that for all $b \in N_\epsilon(\tilde{b})$ we have

$$\begin{aligned} z(b) &= z(\tilde{b}) + \nu_E^\top(b - \tilde{b}) \\ &= z(\hat{b}) + \nu_E^\top(\tilde{b} - \hat{b}) + \nu_E^\top(b - \tilde{b}) \\ &= z(\hat{b}) + \nu_E^\top(b - \hat{b}) = z(\hat{b}) + z_C(b - \hat{b}). \end{aligned} \tag{5.2}$$

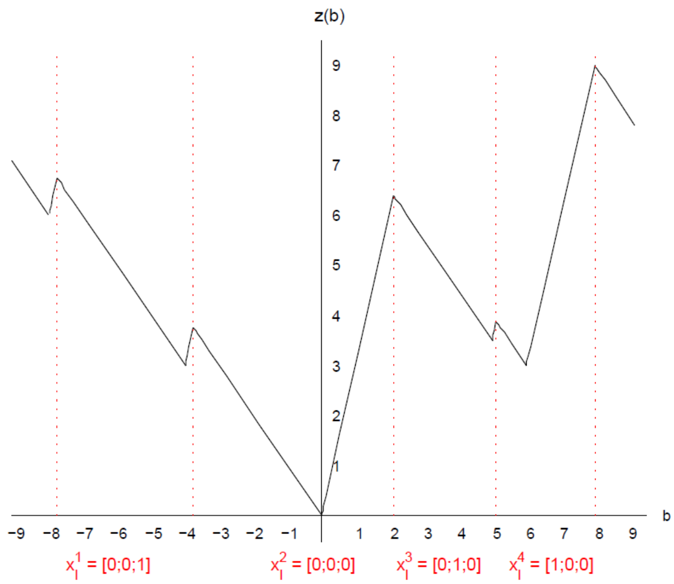


Figure 9: Local stability sets and corresponding integer part of solution in (Ex.1).

But this contradicts the third part of Theorem 2. ■

We finish this section by applying Theorem 2 to the continuous value function in Example 1 and the discontinuous value function in Example 4.

Example 13. Consider the value function (Ex.1). Figure 9 shows the optimal integer parts $x^1, \dots, x^4$ of solutions to the corresponding MILP over the local stability sets $B_{LS}(-4), B_{LS}(0), B_{LS}(5)$ and $B_{LS}(6)$, respectively. One can observe that both the minimal set and the boundary set of the value function are subsets of its set of non-differentiability points.

Similarly, in Example 4, $x^1 = [1\ 2]^\top, x^2 = [2\ 3]^\top, x^3 = [3\ 4]^\top, x^4 = [0\ 0]^\top, x^5 = [1\ 1]^\top, x^6 = [2\ 2]^\top, x^7 = [3\ 3]^\top$ are respectively the integer parts of the solutions for right-hand sides in the local stability sets $B_{LS}(-0.75), B_{LS}(-0.5), \dots, B_{LS}(0.5), B_{LS}(0.75)$. In this case, the value function is discontinuous on the points that belong to the minimal set and we have $B_{min} \cup B_{ES} = B_{ES} = B_{DC} = B_{ND}$. ■

Remark 1 *If z is continuous over B, then $S_D \neq \emptyset$. This follows from the fact that if $S_D = \emptyset$, then $z(0) = z_C(0) = -\infty$ which contradicts $z(0) = 0$. Therefore, we have that $z_C(b) > -\infty$ for all $b \in \mathbb{R}^m$. However, we may still have $z_C(b) = \infty$ for some $b \in \mathbb{R}^m$. The following is an example.*

Example 14. The value function defined by (Ex.12) below is continuous on $\mathbb{R}$, although $\mathcal{K} = \mathbb{R}_+$.

$$\begin{aligned} z(b) = \ & \inf x_1 - x_2 \\ \text{s.t.}\ & -x_1 + x_2 = b \\ & x_1 \in \mathbb{Z}_+,\ x_2 \in \mathbb{R}_+. \quad ■ \end{aligned} \tag{Ex.12}$$

6 A Simplified Jeroslow Formula

The representation we have just described is related (though not so obviously) to a closed form representation of the MILP value function identified by Blair (1995), which he called the *Jeroslow Formula*. In this formula, the value function is obtained by taking the minimum of $|\mathcal{E}|$ functions, each consisting of a PILP value function and a linear term. In this section, we study the connection between our representation of the MILP value function and the representation in the Jeroslow Formula and provide a simpler representation of it.

Let us denote by $\lfloor \cdot \rfloor$ the component-wise floor function. For $E \in \mathcal{E}$, we define

$$\lfloor b \rfloor_E = A_E \left\lfloor A_E^{-1} b \right\rfloor \ \forall b \in B, \ T_E = \{b \in B : A_E^{-1} b \in \mathbb{Z}^m\}, \text{ and } T = \bigcap_{E \in \mathcal{E}} T_E.$$

Consider a given $\hat{b} \in \mathcal{K}$. Let $E \in \mathcal{E}$ be such that $\hat{x}_E = A_E^{-1}\hat{b}$ is the corresponding solution to the continuous restriction w.r.t. the origin. If $\hat{b} \in T_E$, it follows that $\hat{b} = A_E \hat{x}_E$ is an integer linear combination of vectors in feasible basis A_E. Hence, the same is true for any member of T.

Now consider the continuous restriction w.r.t to a given $\hat{x} \in S_I$. Then we have more generally that the corresponding solution to the continuous restriction w.r.t. $\hat{x}$ at a given $\hat{b} \in \mathcal{K} + A_I \hat{x}$ is

$$\tilde{x}_E = A_E^{-1}(\hat{b} - A_I \hat{x}),$$

where $E \in \mathcal{E}$. In this case, when $\hat{b} \in T$, we can no longer guarantee that $\tilde{x}_E \in \mathbb{Z}^m$. By an appropriate scaling, however, we can ensure this property, and this is one of the key steps in deriving the Jeroslow formula. Since all matrices are assumed to be rational, there exists $M \in \mathbb{Z}_+$ such that $M A_E^{-1} A_j \in \mathbb{Z}^m$ for all $E \in \mathcal{E}$ and all $j \in I$, with A_j denoting the j^{th} column of A. Then, since $A_E^{-1} b$ is integral for any $b \in T$ and $E \in \mathcal{E}$, we have that the value function of the following PILP is equal to the value function of the original MILP for all $b \in T$.

Proposition 9 *(Blair, 1995) There exists $M \in \mathbb{Z}_+$ such that $z(b) = z_M(b)$ for all $b \in T$, where*

$$\begin{aligned} z_M(b) = \inf \ & c_I^\top x + \frac{1}{M} c_C^\top y \\ \text{s.t. } & A_I x + \frac{1}{M} A_C y = b \\ & (x, y) \in \mathbb{Z}_+^r \times \mathbb{Z}_+^{n-r}. \end{aligned} \tag{6.1}$$

Proof. Let $M \in \mathbb{Z}_+$ such that $M A_E^{-1} A_j$ is a vector of integers for all $E \in \mathcal{E}$ and $j \in I$. Scaling A_C and c_C by $\frac{1}{M}$ in (MVF) guarantees that $A_j \in T$ for all $j \in I$. Therefore, $M A_E^{-1}(b - A_I x) \in \mathbb{Z}^m$ for all $x \in \mathbb{Z}^r$ and $E \in \mathcal{E}$. It follows that the solution value to z and z_M is equal for any $b \in T$. ■

We illustrate the scaling procedure in the following example.

Example 15. Consider Example 1. In (Ex.1) we have $A_I^j \in \{6, 5, -4\}$ for $j = 1, \ldots, 3$ and $\mathcal{E} =$

$\{\{1\},\{2\}\}$ with $A_{\{1\}} = 2$ and $A_{\{2\}} = -7$. We choose $M = 14$ so that $MA_E^{-1}A_I^j \in \mathbb{Z}$ for all $E \in \mathcal{E}$. The corresponding scaled PILP problem is

$$\begin{aligned} z_M(b) = \inf\ & 3x_1 + \frac{7}{2}x_2 + 3x_3 + \frac{3}{7}x_4 + \frac{1}{2}x_5 \\ \text{s.t. } & 6x_1 + 5x_2 - 4x_3 + \frac{1}{7}x_4 - \frac{1}{2}x_5 = b \\ & x_1, x_2, x_3, x_4, x_5 \in \mathbb{Z}_+. \end{aligned} \tag{6.2}$$

Figure 10a demonstrates the value function (6.2) for $b \in [-9, 9]$. From the figure we can see that z and z_M coincide on intervals of length $\frac{1}{7}$ where $A_{\{1\}} = 2$ is optimal, while the two functions coincide at intervals of length $\frac{1}{2}$ where $A_{\{2\}} = -7$ is optimal.

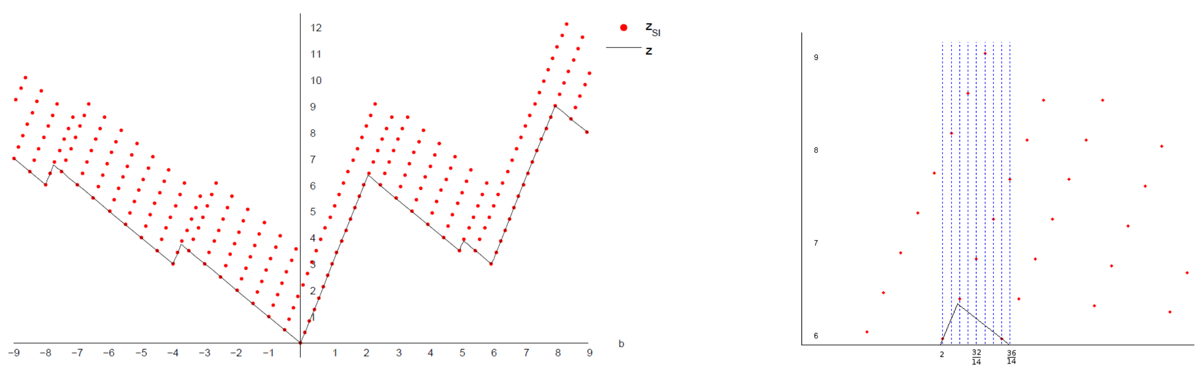
Let us have a closer look at the interval $[2, 2.6]$ illustrated in Figure 10b. The set of feasible right-hand sides for the scaled PILP (6.2) in this interval is $\{2, 2\frac{1}{14}, 2\frac{2}{14}, 2\frac{3}{14}, , \ldots, 2\frac{8}{14}\}$. Among these points, the value function coincides with (6.2) at 2 and 2.5. We have

$$z(b) = z_C(b) = \nu_{\{1\}}b = 3b, \text{ for } b \in [2, 2.125]$$

and

$$z(b) = \bar{z}(b; [0,1,0]^\top) = \frac{7}{2} + \nu_{\{2\}}(b-5) = \frac{7}{2} - (b-5), \text{ for } b \in [2.125, 2.6].$$

Since $T_{\{1\}} = \{b : MA_{\{1\}}^{-1}b = 7b \in \mathbb{Z}\}$, we have $T_{\{1\}} \cap [2, 2.6] = \{b : b = \frac{i}{7}, i = 14, \ldots, 18\}$. Similarly, $T_{\{2\}} \cap [2, 2.6] = \{b : b = \frac{i}{2}, i = 4, 5\}$.



(a) (b)

Figure 10: The scaled PILP value function (6.2).

Over the intervals for which A_{E^*} is the optimal dual basis for the corresponding continuous restriction, z and z_M coincide at $T_{E^*} = \{b \in B : b = \lfloor b \rfloor_{E^*} = kMA_{E^*}^{-1}, k \in \mathbb{Z}_+\}$. For instance, $z(2) = z_M(2)$ with $2 \in T_{\{1\}}$, but $z(2\frac{1}{7}) \neq z_M(2\frac{1}{7})$ with $2\frac{1}{7} \in T_{\{1\}}$. This is due to the fact that $A_{\{1\}}$ is the optimal dual basis at $z(2) = z_C(2)$ but not at $\bar{z}(2\frac{1}{7}; [0,1,0]^\top) = z(2\frac{1}{7})$. ■

Remark 2 *Note that the z and z_M may coincide at some right-hand side that is not in the set T, e.g., $b = 2.5 \in T_{\{2\}} \backslash T_{\{1\}}$.*

Blair and Jeroslow (1984) identified a class of functions called Gomory functions and showed that for any PILP, there exists a Gomory function whose value coincides with that of the value function of the PILP wherever it is finite. To extend this result to the MILP case, Blair (1995) proposed "rounding" any $b \in B$ to some $\lfloor b \rfloor_E$ with $E \in \mathcal{E}$ and evaluating the latter using a PILP. Note that (6.1) has to be modified to be used for this purpose, since it is not necessarily feasible for all $\lfloor b \rfloor_E$, $E \in \mathcal{E}$; i.e., it is possible to have $\tilde{x}_E = M(A_E^{-1} - A_E^{-1} A_I \hat{x}) < 0$ for $\hat{x} \in \mathbb{Z}_+^r$. To achieve feasibility for all $\lfloor b \rfloor_E$, Blair (1995) proposed the following modification of (6.1) and used it in the Jeroslow formula.

$$
\begin{aligned}
z_{JF}(t) = \ & \inf c_I^\top x + \frac{1}{M} c_C^\top y + z(-\frac{1}{M}\sum_{j \in C} A_j)y \\
& \text{s.t. } A_I x + \frac{1}{M} A_C y + (-\frac{1}{M}\sum_{j \in C} A_j)y = t \\
& \qquad y \in \mathbb{Z}_+^{n-r}, x \in \mathbb{Z}_+^r, y \in \mathbb{Z}_+.
\end{aligned}
\tag{6.3}
$$

Finally, he used linear terms of the form of $\nu_E^\top(b - \lfloor b \rfloor_E)$ to compensate for the "rounding" of b to $\lfloor b \rfloor_E$ with $E \in \mathcal{E}$. Together, he showed that for any MILP, there is a Gomory function G corresponding to the value function of the PILP (6.3) with

$$
z(b) = \inf_{E \in \mathcal{E}} \{ G(\lfloor b \rfloor_E) + \nu_E^\top (b - \lfloor b \rfloor_E) \}. \tag{6.4}
$$

The representation of the value function in (6.4) is known as the Jeroslow Formula.

Although it is a bit difficult to tease out, given the technical nature of the Jeroslow Formula, there is an underlying connection between it and our representation. In particular, the set T has a role similar to the role of B_{min} in our representation—it is a discrete subset of the domain of the value function of the original MILP over which the original value function agrees with the value function of a related PILP. This is the same property our set B_{min} has and it is what allows the value function to have a discrete representation. Furthermore, the correction terms in the Jeroslow Formula play a role similar to the value function of the continuous restriction in our representation.

The advantage our representation has over the Jeroslow Formula is that B_{min} is potentially a much smaller set and the value M in the Jeroslow formula would be difficult to calculate a priori. Furthermore, even if M could be obtained in some cases, evaluating the value function for a given $b \in B$ using the Jeroslow formula ostensibly requires the evaluation of a Gomory function for every $\lfloor b \rfloor_E$ for all $E \in \mathcal{E}$, including those feasible bases A_E that are not optimal at b. The number of evaluations required is equal to the size of $\bigcup_{E \in \mathcal{E}} T_E$. These drawbacks relegate the Jeroslow formula to purely theoretical purposes. On the surface, there does not seem to be any way to utilize it in practice. Nevertheless, it is possible to simplify the Jeroslow Formula, replacing T by B_{min} and eliminating the need to calculate M in the process. This leads to a more practicable variant of the original formula. First, we show formally that B_{min} is a subset of T.

Proposition 10 *$B_{min} \subseteq T$.*

Proof. Let $(\hat{x}, \hat{y})$ be an optimal solution to (MVF) at $\hat{b} \in B_{min}$. From Proposition 7, we have that $\hat{y} = 0$ in any optimal solution of the value function at $\hat{b}$. Then for all $E \in \mathcal{E}$ we have

$$
\left\lfloor \hat{b} \right\rfloor_E = \frac{1}{M} A_E \left\lfloor M A_E^{-1} A_I \hat{x} + M A_E^{-1} A_C \hat{y} \right\rfloor = A_I \hat{x} = \hat{b}.
$$

Then, $\left\lfloor \hat{b} \right\rfloor_E = \hat{b}$ for all $E \in \mathcal{E}$ and $\hat{b} \in T$. ■

Corollary 2 *If $\hat{b} \in B_{min}$, then $z(\hat{b}) = z_I(\hat{b}) = z_M(\hat{b}) = z_{JF}(\hat{b}) = G(\hat{b})$ where G is the PILP value function in (6.4).*

Proof. The first equality follows from Proposition 7. The second equality holds since $y = 0$ in any optimal solution to (MVF) at a right-hand side in B_{min}. $z_M(b)$ is equal to $z_{JF}(b)$ since for $b = \lfloor b \rfloor_E$ for all b when $b \in B_{min}$ and y can be fixed to zero in (6.3). The last equality holds for any $\lfloor b \rfloor_E$, $E \in \mathcal{E}$. ■

Theorem 3 (Simplified Jeroslow formula)

$$z(b) = \inf_{\hat{b} \in B_{min}, E \in \mathcal{E}} \{z_I(\hat{b}) - \nu_E^\top (b - \hat{b})\}. \tag{6.5}$$

Proof. We have

$$\begin{aligned} z(b) &= \inf_{E \in \mathcal{E}} \{G(\lfloor b \rfloor_E) + \nu_E^\top (b - \lfloor b \rfloor_E)\} \\ &= \inf_{\hat{b} \in T_E, E \in \mathcal{E}} \{G(\hat{b}) + \nu_E^\top (b - \hat{b})\} \\ &= \inf_{\hat{b} \in B_{min}} \{z_I(\hat{b}) + \sup_{E \in \mathcal{E}} \nu_E^\top (b - \hat{b})\} \\ &= \inf_{\hat{b} \in B_{min}, E \in \mathcal{E}} \{z_I(\hat{b}) - \nu_E^\top (b - \hat{b})\}. \end{aligned}$$

The first equation is the Jeroslow Formula. The second one is because $\lfloor b \rfloor_E \in T_E$ for any $E \in \mathcal{E}$ and $b \in B$. From Theorem 1, $z(b) = \inf\{\bar{z}(b; x) : A_I x \in B_{min}\}$, then the third equality holds. The last equation follows trivially. ■

The above result provides a variation of the Jeroslow formula where there is no need to find the value of M, or to evaluate the PILP value function z_{JF} for members of $\bigcup_{E \in \mathcal{E}} T_E$. Instead, we need to evaluate the simpler PILP value function z_I for the set $B_{min} \subseteq T \subseteq \bigcup_{E \in \mathcal{E}} T_E$. The difference in the size of B_{min} and $\bigcup_{E \in \mathcal{E}} T_E$ can be significant. We provide an illustrative example next.

Example 16. The value function of (6.2) for right-hand sides in $T_{\{1\}} \cup T_{\{2\}}$ is plotted in Figure 11 with filled blue circles. At a point $\hat{b} \in T_{\{1\}} \cup T_{\{2\}}$, we have $z_M(\hat{b}) = G(\hat{b})$ where G is the Gomory function corresponding to the PILP (6.2). The Jeroslow formula for the MILP value function over $[-9, 9]$ requires finding all such points. Alternatively, we can have a smaller representation by constructing the value function of the integer restriction of (Ex.1). i.e., $z_I(\hat{b}) = \inf\{3x_1 + \frac{7}{2}x_2 + 3x_3 : 6x_1 + 5x_2 - 4x_3 = \hat{b}, x_1, x_2, x_3 \in \mathbb{Z}^+\}$. This value function is plotted in Figure 12. However, the alternative formulation (6.5) requires finding $G(\hat{b}) = z_I(\hat{b})$ for $\hat{b} \in B_{min} = \{-8, -4, 0, 4, 5, 10\}$. ■

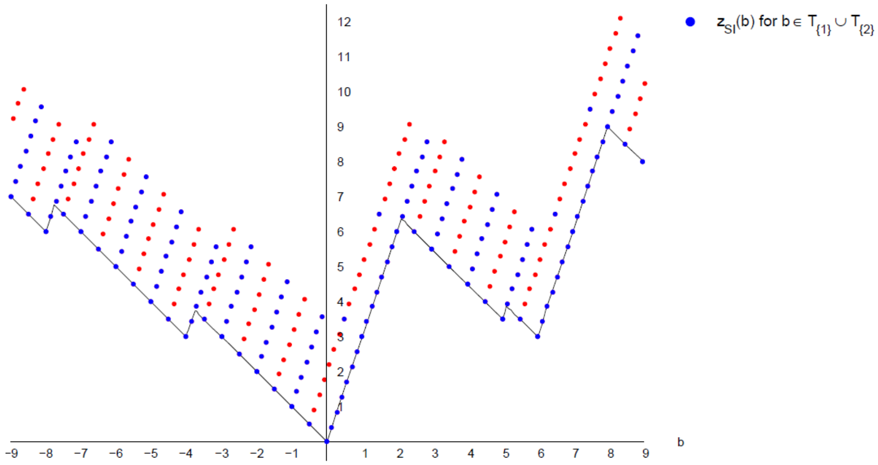


Figure 11: The value function of (6.2) for $b \in T_{\{1\}} \cup T_{\{2\}} \cup [-9, 9]$.

7 Finite Algorithm for Construction

In this final section, we discuss the use of our representation in computational practice, which is the ultimate goal of this work. To sum up what we have seen so far, we have shown that there exists a discrete set S_{min} (not necessarily unique) over which the value of z can be determined by solving instances of the integer restriction. Theorem 1 tells us that, in principle, if we knew $z(A_I x)$ for all $x \in S_{min}$, then $z(b)$ could be computed at any $b \in B$ by solving $|S_{min}|$ LPs.

Our discrete representation of the value function in Theorem 1 is equivalent to

$$z(b) = \inf_{x \in S_{min}} c_I^\top x + z_C(b - A_I x). \tag{7.1}$$

If $|S_{min}|$ is relatively small, this yields a practical representation. The most straightforward way to utilize our representation would then be to generate the set S_{min} a priori and to apply the above formula to evaluate $z(b)$ for $b \notin B_{min}$.

In general, obtaining an exact description of the set S_{min} seems to be difficult. One solution to this problem would be to instead generate the value function of the integer restriction first by the procedure of Kong et al. (2006), which is finite under our assumptions. We illustrate this hypothetical procedure in the following example.

Example 17. Consider constructing the value function defined by (Ex.1) for $b \in [-7, 7]$. The value function of the integer restriction z_I is plotted in Figure 12. Clearly, complete knowledge of z_I is unnecessary to describe the MILP value function, as this requires evaluation for each point in S_I, whereas we have already shown that evaluation of points in S_{min} is enough. In this example, over $b \in [-7, 7]$, we have that $S_{min} = \{0; 0; 1], [0; 0; 0], [0; 1; 0], [1; 0; 0]\}$. Therefore, four evaluations is enough, yet at least 15 are required for constructing the value function of the PILP. ■

Hence, this approach does not seem to be efficient. Instead, we anticipate overcoming this difficulty in two different ways, depending on the context in which the value function is needed. First, working with a subset of S_{min} still yields an upper approximation of z, which might be useful

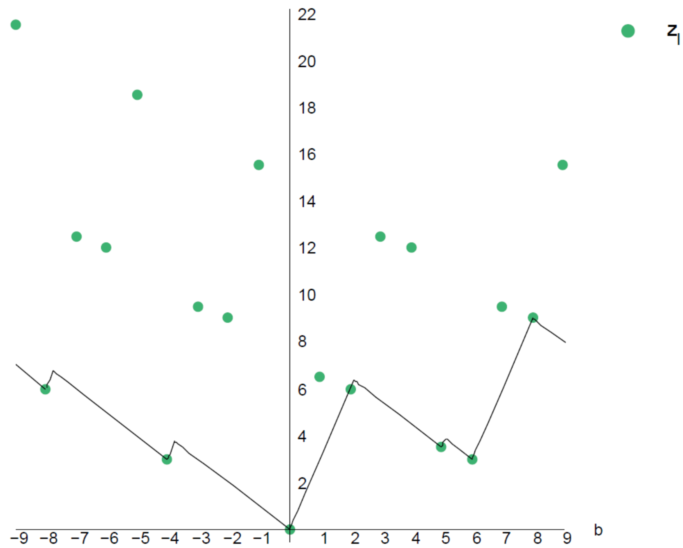


Figure 12: The value function of the integer restriction of (Ex.1) for $b \in [-9, 9]$.

in particular applications where an approximate solution will suffice. Second, we anticipate that in most cases, it would be possible to dynamically generate the set S_{min}, adding points only as necessary for improving the approximation in the part of the domain required for solution of a particular instance. This approach would be similar to that of using dynamic cut generation to solve fixed MILPs.

We demonstrate the potential of both such techniques here by describing a method for iteratively improving a given discrete approximation of the value function by dynamically generating improving members of S_I after the fashion of a cutting plane algorithm for MILP. At iteration k, we begin with an approximation arising from $S^k \subseteq S_I$ using the formula

$$\bar{z}(b) = \inf\{c_I^\top x + z_C(b - A_I x) : x \in S_I^k, z(A_I x) = c_I^\top x\} \tag{7.2}$$

and we generate the set S^{k+1} by determining the point at which the current approximation is maximally different from the true value function. This is akin to generation of the most violated valid inequality in the case of MILP. An important feature of the algorithm is that it produces a performance guarantee after each step, which bounds the maximum gap between the approximation and the true function value. In what follows, we denote the current upper bounding function by $\bar{z}$.

In addition to the initial assumption $z(0) = 0$, we also assume the set B_I is non-empty and bounded (while B can remain unbounded) to guarantee finite termination. Note, however, that it is possible to apply the algorithm even if this is not the case. It is a simple matter, for example, to generate the value function within a given box, even if B_I is an unbounded set. We note that we do not require the assumption $\mathcal{K} = \mathbb{R}^m$, although this is not a restrictive assumption in practice anyway, since A_C can always be modified to satisfy it (Kall and Mayer, 2010).

Algorithm

Initialize: Let $\bar{z}(b) = \infty$ for all $b \in B$, $\Gamma^0 = \infty$, $x^0 = 0$, $S^0 = \{x^0\}$, and $k = 0$.
while $\Gamma^k > 0$ **do**:

- Let $\bar{z}(b) = \min\{\bar{z}, \bar{z}(b; x^k)\}$ for all $b \in B$.
- $k \leftarrow k + 1$.
- Solve

$$\begin{aligned} \Gamma^k = \max\ & \bar{z}(b) - c_I^\top x \\ \text{s.t. } & A_I x = b \\ & x \in \mathbb{Z}_+^r. \end{aligned} \tag{SP}$$

  to obtain x^k.
- Set $S^k \leftarrow S^{k-1} \cup \{x^k\}$

end while
return $z(b) = \bar{z}(b)$ for all $b \in B$.

The key to this method is effective solution of (SP). We show how to formulate this problem as a mixed integer nonlinear program below. For practical computation, (SP) can be rewritten conceptually as

$$\begin{aligned} \Gamma^k = \max\ & \theta \\ \text{s.t. } & \theta \leq \bar{z}(b) - c_I^\top x \\ & A_I x = b \\ & x \in \mathbb{Z}_+^r. \end{aligned} \tag{7.3}$$

The upper approximating function $\bar{z}(b)$ is a non-convex and non-concave piecewise polyhedral function that is obtained by taking the minimum of a finite number of convex piecewise polyhedral functions $\bar{z}$. In particular, in iteration $k > 1$ of the algorithm we have $\bar{z}(b) = \min_{i=1,\ldots,k-1} \bar{z}(b; x^i)$. Therefore, the first constraint in (SP) can be reformulated as $k-1$ constraints, the right-hand side of each of which is a convex piecewise polyhedral function.

$$\theta + c_I^\top x \leq c_I^\top x^i + z_C(b - A_I x^i) \quad i = 1, \ldots, k-1. \tag{7.4}$$

Next, we can write z_C as

$$z_C(b - A_I x^i) = \sup\{(b - A_I x^i)^\top \nu^i : A_C^\top \nu^i \leq c_C, \nu^i \in \mathbb{R}^m\} \tag{7.5}$$

and reformulate each of $k-1$ constraints in (7.4) as

$$\begin{aligned} & \theta + c_I^\top x \leq c_I^\top x^i + (b - A_I x^i)^\top \nu^i \\ & A_C^\top \nu^i \leq c_C \\ & \nu^i \in \mathbb{R}^m \end{aligned} \tag{7.6}$$

for $i \in \{1, \ldots, k-1\}$. Together, then, in each iteration we solve

$$\begin{aligned} \Gamma^k = \ & \max \theta \\ \text{s.t. } & \theta + c_I^\top x \leq c_I^\top x^i + (A_I x - A_I x^i)^\top \nu^i \quad i = 1, \ldots, k-1 \\ & A_C^\top \nu^i \leq c_C \quad i = 1, \ldots, k-1 \\ & \nu^i \in \mathbb{R}^m \quad i = 1, \ldots, k-1 \\ & x \in \mathbb{Z}_+^r. \end{aligned} \tag{7.7}$$

Due to the first constraint, the resulting problem is a nonlinear optimization problem. Nevertheless, solvers do exist, e.g., Couenne (Belotti, 2009) that are capable of solving these problems. Assuming that there is a finite method to solve (7.7), we next show that the proposed algorithm terminates finitely and returns the correct value function.

Theorem 4 (Algorithm for Construction) *Under the assumptions that B_I is non-empty and bounded and (7.7) can be solved finitely, the algorithm terminates with the correct value function in finitely many steps.*

Proof. For any $x \in S_I$, $c_I^\top x \geq z(b)$ for all $b \in B$. From Proposition 7, we have that for $x \in S_{min} \subseteq S_I$, $c_I^\top x = z(A_I x)$. Therefore, for the solution of (7.7) at iteration k we have $x^k \in S_I$ and $c_I^\top x^k = z(A_I x^k)$. Since B_I is assumed to be bounded, then there is a finite number of such points that can be generated in the algorithm. That is, $\bar{z}$ can only be updated a finite number of times.

To see that at termination, $\bar{z}$ is the value function, first note that Proposition 4 implies that the initialization and the updates of the approximating function result in valid upper bounding functions. If in iteration k, the approximation $\bar{z}(b)$ is strictly above the value function at some $b \in B$, then $\Gamma^k > 0$ and there is some $x \in S_I$ for which $c_I x$ lies on the value function and below the approximation. The subproblem is guaranteed to find such a point, therefore, in each intermediate iteration we improve the approximation. When no such a point is found, the approximation is exact everywhere and we terminate with $\Gamma^k = 0$. ■

To illustrate, we apply the algorithm to two value functions: the first one is the function (Ex.1). The second value function is from the two-stage stochastic integer optimization literature and refers to the value function of the second-stage problem of the stochastic server location problem (SSLP) in (Ntaimo and Sen, 2005).

Example 18. Consider (Ex.1) where $x_1, x_2, x_3 \in \{1, \dots, 5\}$. Figure 13 plots Γ^k normalized by Γ^1, the initial gap reported with $\bar{z} = z_C$, versus the iteration number for problem (7.7) . When the algorithm is executed, over $b \in [-7, 7]$, the updates only occur for $\hat{x}$ such that $A_I \hat{x} \in \{-4, 5, 6\}$. This is because the remainder of the right-hand sides $A_I \hat{x}$ in $[-7, 7]$ correspond to $(A_I \hat{x}, c_I^\top \hat{x})$ (green circles in Figure 12) that lie either on or above z_C (and therefore below the following updated approximating functions). ■

The proposed algorithm can be applied to MILPs with inequality constraints by adding appropriate non-negativity restrictions to the dual variables ν in (7.7). We see an example next.

Example 19. Consider the second-stage problem of SSLP with 2 potential server locations and 3 potential clients. The first-stage variables and stochastic parameters are captured in the right-hand sided $b_1, \dots, b_5$. The resulting formulation is

$$\begin{aligned} z(b) = \min\ & 22y_{12} + 15y_{21} + 11y_{22} + 4y_{31} + 22y_{32} + 100R \\ \text{s.t. } & 15y_{21} + 4y_{31} - R \le b_1 \\ & 22y_{12} + 11y_{22} + 22y_{32} - R \le b_2 \\ & y_{11} + y_{12} = b_3 \\ & y_{21} + y_{22} = b_4 \\ & y_{31} + y_{32} = b_5 \\ & y_{ij} \in \mathbb{B}, i \in \{1,2,3\}, j \in \{1,2\}, R \in \mathbb{R}_+. \end{aligned} \tag{7.8}$$

The normalized gap Γ^k/Γ^1 versus the iteration number k is plotted in Figure 13. For this example, non-positivity constraints on the dual variables corresponding to the first two constraints are added to (7.7). ■

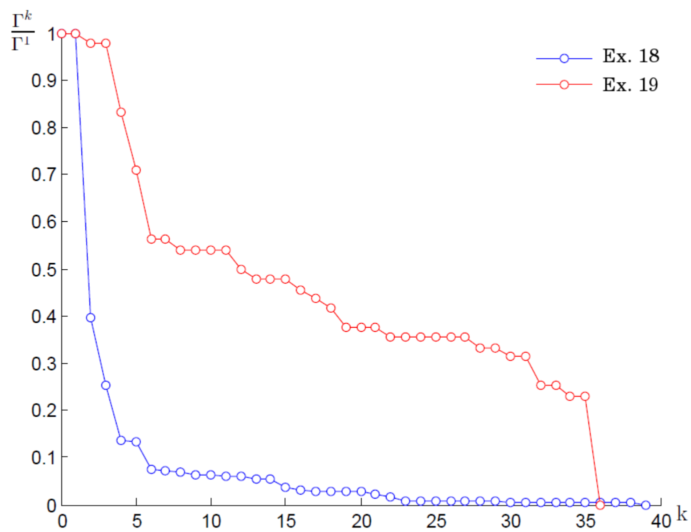


Figure 13: Normalized approximation gap vs. iteration number.

As one can observe in Figure 13, the quality of approximations improves significantly as the algorithm progresses. The upper-approximating functions $\bar{z}$ obtained from the intermediate iterations of the algorithm can be utilized within other solution methods that rely on bounding a MILP from above. Clearly, such piecewise approximating functions $\bar{z}$ are structurally simpler than the original MILP value function. Furthermore, as with SSLP, a common class of two-stage stochastic optimization problems considers stochasticity in the right-hand side. With a description of the value function of the second-stage problem, finding the solution to different second-stage problems reduces to evaluations of the value function at different right-hand sides. The proposed algorithm can therefore be incorporated into methods to solve stochastic optimization problems with a large number of scenarios.

8 Conclusion

In this work, we study the MILP value function, which is key to integer optimization sensitivity analysis and solution methods for various classes of optimization problems. The backbone of our work is to derive a discrete characterization of the MILP value function which can be utilized in developing algorithms for such problems. We identify a countable set of right-hand sides that describe the discrete structure of the value function and use this set to propose an algorithm for the construction of MILP value function. This algorithm is finite when the set of right-hand sides over which the value function of the associated pure integer problem is finite is bounded.

We further outline the connection between the MILP, PILP, and LP value functions. In particular, we show that the MILP value function arises from the combination of a PILP value function and a single LP value function. We address the relationship between our representation and the classic Jeroslow formula for the MILP value function. Finally, we study the continuity and convexity properties of the value function, as well as the relationships between several critical sets of the right-hand sides such as the set of discontinuity and non-differentiability points.

As a result of our work, we now have a method to dynamically generate points necessary to describe a MILP value function. A subset of such points can be used to derive functions that bound the value function from above, while the full collection of them is sufficient to have a complete characterization of the value function. The dynamic generation of these points can be integrated with iterative methods to solve stochastic integer and bilevel integer optimization problems. We show describe such a method for the case of two-stage stochastic programming with mixed integer recourse in (Hassanzadeh et al., 2014).

References

Ahmed, S., Tawarmalani, M., and Sahinidis, N. (2004). A finite branch-and-bound algorithm for two-stage stochastic integer programs. *Mathematical Programming*, 100(2):355–377.

Bank, B., Guddat, J., Klatte, D., Kummer, B., and Tammer, K. (1983). *Non-linear parametric optimization*. Birkhäuser verlag.

Bard, J. (1991). Some properties of the bilevel programming problem. *Journal of optimization theory and applications*, 68(2):371–378.

Bard, J. F. (1998). *Practical bilevel optimization: algorithms and applications*, volume 30. Springer.

Bazaraa, M., Jarvis, J., Sherali, H., and Bazaraa, M. (1990). *Linear programming and network flows*, volume 2. Wiley Online Library.

Belotti, P. (2009). Couenne: a users manual. *Technical Report*.

Blair, C. (1995). A closed-form representation of mixed-integer program value functions. *Mathematical Programming*, 71(2):127–136.

Blair, C. and Jeroslow, R. (1977). The value function of a mixed integer program: I. *Discrete Mathematics*, 19(2):121–138.

Blair, C. and Jeroslow, R. (1979). The value function of a mixed integer program: Ii. *Discrete Mathematics*, 25(1):7–19.

Blair, C. and Jeroslow, R. (1982). The value function of an integer program. *Mathematical Programming*, 23(1):237–273.

Blair, C. and Jeroslow, R. (1984). Constructive characterizations of the value-function of a mixed-integer program i. *Discrete Applied Mathematics*, 9(3):217–233.

Conti, P. and Traverso, C. (1991). Buchberger algorithm and integer programming. *Applied algebra, algebraic algorithms and error-correcting codes*, pages 130–139.

Dempe, S., Mordukhovich, B. S., and Zemkoho, A. B. (2012). Sensitivity analysis for two-level value functions with applications to bilevel programming. *SIAM Journal on Optimization*, 22(4):1309–1343.

Güzelsoy, M. and Ralphs, T. (2006). The value function of a mixed-integer linear program with a single constraint. *To be submitted*.

Hassanzadeh, A., , and Ralphs, T. K. (2014). A generalization of Benders' algorithm for two-stage stochastic optimization problems with mixed integer recourse. Technical report, COR@L Laboratory, Lehigh University.

Kall, P. and Mayer, J. (2010). *Stochastic linear programming: models, theory, and computation*. Springer Verlag.

Kong, N., Schaefer, A., and Hunsaker, B. (2006). Two-stage integer programs with stochastic right-hand sides: a superadditive dual approach. *Mathematical Programming*, 108(2):275–296.

Nemhauser, G. and Wolsey, L. (1988). *Integer and combinatorial optimization*, volume 18. Wiley New York.

Ntaimo, L. and Sen, S. (2005). The million-variable march for stochastic combinatorial optimization. *Journal of Global Optimization*, 32(3):385–400.

S DeNegre (2011). *Interdiction and Discrete Bilevel Linear Programming*. Phd, Lehigh University.

Schultz, R., Stougie, L., and Van Der Vlerk, M. (1998). Solving stochastic programs with integer recourse by enumeration: A framework using Gröbner basis. *Mathematical Programming*, 83(1):229–252.

Trapp, A. C., Prokopyev, O. A., and Schaefer, A. J. (2013). On a level-set characterization of the value function of an integer program and its application to stochastic programming. *Operations Research*, 61(2):498–511.