The Gap Function: Evaluating Integer Programming Models over Multiple Right-hand Sides

Temitayo Ajayi *[1], Christopher Thomas †[2], and Andrew Schaefer ‡[1]

[1]Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005
[2]Bazean Corporation

Abstract

For an integer programming model with fixed data, the linear programming relaxation gap is considered one of the most important measures of model quality. There is no consensus, however, on appropriate measures of model quality that account for data variation. In particular, when the right-hand side is not known exactly, one must assess a model based on its behavior over many right-hand sides. Gap functions are the linear programming relaxation gaps parametrized by the right-hand side. Despite drawing research interest in the early days of integer programming (Gomory 1965), the properties and applications of these functions have been little studied. In this paper, we construct measures of integer programming model quality over sets of right-hand sides based on the absolute and relative gap functions. In particular, we formulate optimization problems to compute the expectation and extrema of gap functions over finite discrete sets and bounded hyper-rectangles. These optimization problems are linear programs (albeit of an exponentially large size) that contain at most one special ordered set constraint. These measures for integer programming models, along with their associated formulations, provide a framework for determining a model's quality over a range of right-hand sides.

Keywords: Gap function, linear programming relaxation, philosophy of modeling, superadditive duality

1 Introduction

The linear programming relaxation is foundational to most practical approaches for solving integer programs. Nemhauser and Wolsey (1988) state that "most integer programming

*temitayo.ajayi@rice.edu
†cthomas@bazean.com
‡andrew.schaefer@rice.edu

algorithms require an upper bound on the value of the objective function, and the efficiency of the algorithm is very dependent on the sharpness of the bound.” In practice, this upper bound is almost always the objective value of the linear programming relaxation (for maximization problems). In this sense, the linear programming relaxation gap is considered one of the most important measures of the quality of an integer programming model.

Most evaluative metrics for integer programming models concern only a single instance of the problem and may not consider how the model performs when the data vary. However, often it is necessary to evaluate a model under multiple parameters. For instance, the right-hand side of the constraints may not be known exactly. Alternatively, one may analyze the sensitivity of a model to changes in resources, or determine which right-hand sides, if any, produce a perfect formulation. In any of these cases, it is necessary to assess a model’s quality based on how it performs over a set of right-hand sides. This necessity is supported by the prevalence of linear programming sensitivity analysis (e.g., Higle and Wallace (2003)), and the lack of a thorough, complementary study in integer programming.

Gap functions represent the difference (absolute or relative) between the value function of the linear programming relaxation and the value function of the integer program. Studying gap functions over a set of right-hand sides was one of the first questions to arise in integer programming: Gomory (1965) proves that the absolute gap function is periodic within particular subsets of right-hand sides. However, the gap function has been little studied since. Blair and Jeroslow (1979) prove that the absolute gap function is bounded under certain conditions. Hoşten and Sturmfels (2007) characterize the maximum gap over parametrized right-hand sides by using monomial ideals. Aliev et al. (2017) provide bounds on the absolute gap function for knapsack problems. Eisenbrand et al. (2013) show that it is NP-hard to test if the absolute gap function is bounded by a given constant.

Related research includes the study of the proximity of solutions, either when the right-hand side is changed or when the integrality constraints are relaxed. Blair and Jeroslow (1977) and Mangasarian and Shiau (1987) bound the distance between solutions of two linear programs with distinct right-hand sides. Cook et al. (1986) and Paat et al. (2018) bound the distance from an optimal solution of the linear programming relaxation to the nearest optimal solution of the integer program and mixed-integer program, respectively. These results can be used directly to bound, but not necessarily compute, the expectation and supremum of the gap functions. Granot and Skorin-Kapov (1990) and Werman and Magagnosc (1991) extend Cook et al. (1986) to classes of quadratic and convex integer programs, respectively. Stein (2016a,b) study the feasibility and optimality error of the continuous relaxations of mixed-integer programs using grid relaxations.

The existing literature primarily focuses on worst-case performance, but there is little study of best-case or average performance. Evaluating a family of integer programs is important because it is useful to know how a model responds to changes in the data. Because the gap (absolute or relative) at a single right-hand side is a good measure of model quality, we propose that gap functions are a suitable extension for evaluation over multiple right-hand sides. Further, computing metrics based on gap functions enables the comparison of two models, even when one model does not dominate the other model (see Example 2).

In this paper, we propose evaluative metrics, based on gap functions, for pure integer programs parametrized over a set of right-hand sides. Notably, our metrics measure how the model behaves throughout the set of right-hand sides and address more than just worst-case performance. These other metrics may have benefits in areas such as stochastic programming. The remainder of the paper is organized as follows: Section 2 reviews preliminaries, defines gap functions, and provides some properties of gap functions. Sections 3-6 formulate optimization problems whose optimal objective values are the expectation, supremum, and infimum of the absolute and relative gap functions, and Table 1 provides a summary of evaluative metrics addressed in this paper. The optimization problems that we present are linear programs, albeit of exponentially large size, with at most one special-ordered-set constraint. The combination of these optimization problems with their associated metrics provides a framework for evaluating integer programming models over many right-hand sides.

Gap Function	RHS Set	Expectation	Supremum	Infimum
Absolute	Discrete Set	Section 3.1	Section 3.2	Section 3.3
	Hyper-rectangle	Section 4.1	Section 4.2	Section 4.3
Relative	Discrete Set	Section 5.1	Section 5.2	Section 5.3
	Hyper-rectangle	——	Section 6.1	Section 6.2

Table 1: A summary of gap function problems and metrics addressed in Sections 3-6. Note that we do not consider the expectation of the relative gap function over a hyper-rectangle.

The proofs of the results are in the electronic companion, unless otherwise stated.

2 Preliminaries and Properties of Gap Functions

Consider the following motivational examples.

Example.

Consider $\beta \in [0,1]^2$ and:

$$\max \{c^T x \mid x_1 \leq \beta_1, x_2 \leq \beta_2, x \in \mathbb{Z}^2_+\}, \tag{IP1}$$

$$\max \{c^T x \mid .6x_1 \leq \beta_1, .6x_2 \leq \beta_2, x \in \mathbb{Z}^2_+\}. \tag{IP2}$$

(IP1) is a tighter formulation than (IP2) for all $\beta \in [0,1]^2$. In fact, when β is integral, removing the integrality constraint for (IP1) (the relaxed feasible region) yields the convex hull of (IP1) and (IP2). In contrast, the convex hull of both (IP1) and (IP2) is a strict subset of the relaxed feasible region of (IP2). In cases such as this, one model (IP1) is preferred over another model (IP2) for all β in some right-hand side set ($\beta \in [0,1]^2$) with respect to the absolute gap function. $\square$

Determining which model is preferred is not always straightforward.

Example.

Consider $\beta \in [0,1]^2$ and:

$$\max \{c^T x \mid .7x_1 \leq \beta_1, x_2 \leq \beta_2, x \in \mathbb{Z}^2_+\}, \tag{IP3}$$

$$\max \{c^T x \mid x_1 \leq \beta_1, .8x_2 \leq \beta_2, x \in \mathbb{Z}^2_+\}. \tag{IP4}$$

Figure 1 illustrates how the feasible regions and the tightness of the relaxations of Example 2 change with respect to the right-hand side. The feasible regions of (IP3) and (IP4) are the same for any integral right-hand side β. Observe that the relaxation of (IP3) yields the convex hull when $\beta_1 = 0$ and the relaxation of (IP4) yields a strict superset of the convex hull. On the other hand, the relaxation of (IP4) yields the convex hull when $\beta_2 = 0$ and the relaxation of (IP3) yields a strict superset of the convex hull. Thus, in general, it is not obvious how to evaluate models over a set of right-hand sides.

Moreover, it may not be sufficient to consider a single metric to compare models. Let $\widehat{\mathcal{B}} = \{(0,0)^T, (1,0)^T, (0,1)^T, (1,1)^T\}$, and $c = (1,1)^T$. Define $\Gamma_3 : \widehat{\mathcal{B}} \to \mathbb{R}$ to be the absolute gap function for (IP3) and define Γ_4 similarly for (IP4) (see Definition 2.1 for details on the absolute gap function definition). Observe that $\sup\limits_{\beta \in \widehat{\mathcal{B}}} \Gamma_3(\beta) - \sup\limits_{\beta \in \widehat{\mathcal{B}}} \Gamma_4(\beta) = 3/7 - 1/4 > 0$. By the supremum metric, (IP4) is preferable to (IP3).

Let $\hat{\xi}$ be a discrete random variable with event space $\widehat{\mathcal{B}}$ and $\mathbf{P}\{\hat{\xi} = \beta\} = \hat{\mu}(\beta)$. Consider the probability distributions $\hat{\mu}$ given by: $\hat{\mu}((1,0)^T) = p, \hat{\mu}((0,1)^T) = q, \hat{\mu}((1,1)^T) = r, \hat{\mu}((0,0)^T) = 1-(p+q+r)$, where $p, q, r \geq 0$ and $p+q+r \leq 1$. Suppose $p = 1/10, q = 1/2$, and $r = 1/5$. Then the expectation of the difference of the absolute gap functions is $\mathbb{E}_{\hat{\xi}}[\Gamma_3(\beta) - \Gamma_4(\beta)] = -13/280$. By the expectation metric with this distribution, (IP3) is preferable to (IP4). Furthermore, for any p, q, r as described above, (IP3) is preferable to (IP4) if and only if $(3/7)(p+r) - (1/4)(q+r) < 0$. $\square$

Example 2 demonstrates that it can be difficult to assess whether one integer programming formulation is better than another over multiple right-hand sides. In addition, it may be important to consider multiple metrics during this evaluation, as they can yield different

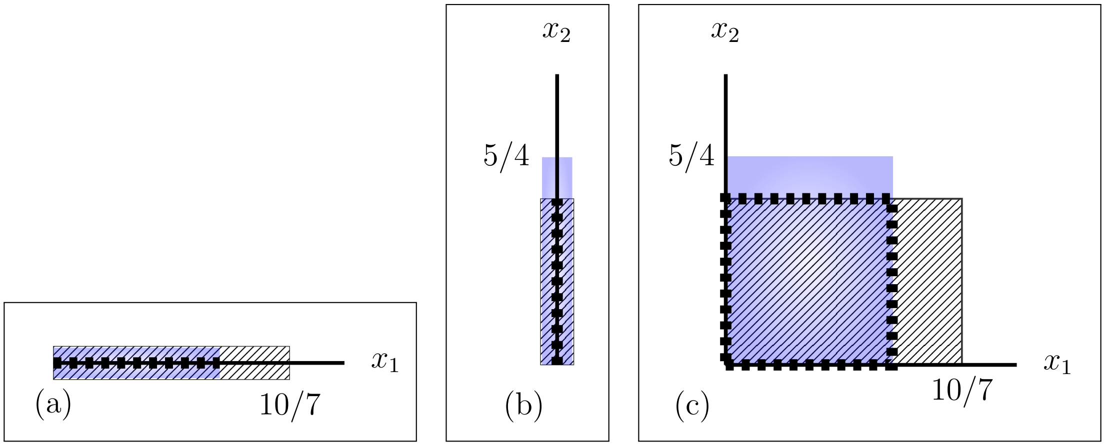

Figure 1: The feasible regions for the relaxations of IP3(β) (hatched) and IP4(β) (shaded) from Example 2. The polyhedra for $\beta = (1, 0)$ (a), $(0, 1)$ (b), and $(1, 1)$ (c) are shown. The feasible region when $\beta = (0, 0)$ is the origin (not shown), and the feasible regions for (a) and (b) are line segments along axes. The dashed line indicates the convex hull of the integral feasible points for both IP3 and IP4. Which formulation is tighter depends on the right-hand side.

information. The remainder of this paper focuses on providing evaluative metrics for models over multiple right-hand sides.

2.1 Preliminaries

Linear and integer programs are defined by the following data: $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$. In general, it is assumed implicitly that the data in such optimization problems are rational, and we continue this practice. Moreover, we make the following assumption.

A1. *The problem data, A, b, and c, are integral.*

Let a_i and a^j denote the i^{th} row and j^{th} column of A, respectively, and a_{ij} denote the j^{th} entry in the i^{th} row of A. Let $\tilde{b} \in \mathbb{Z}^m$ such that $\tilde{b} \leq b$. Define the bounded hyper-rectangle $\mathcal{B}(\tilde{b}, b) = \otimes_{i=1}^m [\tilde{b}_i, b_i]$, the Cartesian product of the intervals $[\tilde{b}_1, b_1], ..., [\tilde{b}_n, b_n]$, and let $\widehat{\mathcal{B}}(\tilde{b}, b) = \mathcal{B}(\tilde{b}, b) \cap \mathbb{Z}^m$; thus, $\widehat{\mathcal{B}}(\tilde{b}, b)$ contains only integral vectors. In the following parametrized optimization problems, the right-hand side parameter, β, is a vector in $\mathcal{B}(\tilde{b}, b)$.

Let $P(\beta) = \{x \in \mathbb{R}^n_+ \mid Ax \leq \beta\}$, and define the *parametrized linear program* LP(β) as follows:

$$z_{LP}(\beta) = \max\{c^T x \mid x \in P(\beta)\}, \text{ and } \text{opt}_{LP}(\beta) = \arg\max\{c^T x \mid x \in P(\beta)\}.$$

The *linear programming value function* is $z_{LP} : \mathcal{B}(\tilde{b}, b) \to \mathbb{R} \cup \{-\infty, \infty\}$. If LP($\beta$) is unbounded for some $\beta \in \mathcal{B}(\tilde{b}, b)$, define $z_{LP}(\beta) = \infty$. If LP(β) is infeasible, define $z_{LP}(\beta) = -\infty$.

Let $Q = \{v \in \mathbb{R}^m_+ \mid A^T v \geq c\}$, which has a finite set $\mathcal{K}$ of extreme points. Moreover, because $Q \subseteq \mathbb{R}^m_+$, $|\mathcal{K}| \geq 1$ if $Q \neq \emptyset$. Denote these extreme points as $\{v^k\}_{k\in\mathcal{K}}$.

The dual of LP(β), denoted DLP(β), is defined as follows:

$$z_{DLP}(\beta) = \min\{\beta^T v \mid v \in Q\}, \text{ and } \text{opt}_{DLP}(\beta) = \arg\min\{\beta^T v \mid v \in Q\}.$$

The *dual value function* is $z_{DLP} : \mathcal{B}(\tilde{b}, b) \to \mathbb{R} \cup \{-\infty, \infty\}$. If DLP($\beta$) is unbounded for some $\beta \in \mathcal{B}(\tilde{b}, b)$, define $z_{DLP}(\beta) = -\infty$. If DLP(β) is infeasible, define $z_{DLP}(\beta) = \infty$. DLP(β) is a weak and strong dual to LP(β), i.e., for any $x \in P(\beta)$ and any $v \in Q$, $c^T x \leq \beta^T v$, and $z_{DLP}(\beta) = z_{LP}(\beta)$ when both problems are feasible.

Let $S(\beta) = P(\beta) \cap \mathbb{Z}^n$. Define the *parametrized integer program* IP(β) as follows:

$$z_{IP}(\beta) = \max\{c^T x \mid x \in S(\beta)\}, \text{ and } \text{opt}_{IP}(\beta) = \arg\max\{c^T x \mid x \in S(\beta)\}.$$

The *integer programming value function* is $z_{IP} : \mathcal{B}(\tilde{b}, b) \to \mathbb{R} \cup \{-\infty, \infty\}$. If IP($\beta$) is infeasible for some $\beta \in \mathcal{B}(\tilde{b}, b)$, define $z_{IP}(\beta) = -\infty$. If IP(β) is unbounded, define $z_{IP}(\beta) = \infty$. The *linear programming relaxation* of IP(β) is LP(β).

One can construct gap functions for IP using value functions. Gap functions represent the difference, absolute or relative, between the linear programming relaxation value function and the integer programming value function. Gap functions measure the linear programming relaxation's accuracy for approximating the associated integer program, and as such, are indicators of model quality over many right-hand sides.

Definition. *Given a right-hand side set D, the absolute gap function for integer programs, $\Gamma : D \to \mathbb{R} \cup \{\infty\}$, is defined by $\Gamma(\beta) = z_{LP}(\beta) - z_{IP}(\beta)$.*

Unless otherwise stated, the domain of the absolute gap function is $\mathcal{B}(\tilde{b}, b)$. Because $S(\beta) \subseteq P(\beta)$, for all $\beta \in \mathcal{B}(\tilde{b}, b)$, $\Gamma(\beta)$ is nonnegative for all $\beta \in \mathcal{B}(\tilde{b}, b)$.

Let $\mathcal{B}^+(\tilde{b}, b) = \{\beta \in \mathcal{B}(\tilde{b}, b) \mid z_{IP}(\beta) \geq 0, z_{LP}(\beta) > 0\}$, and let $\widehat{\mathcal{B}}^+(\tilde{b}, b) = \widehat{\mathcal{B}}(\tilde{b}, b) \cap \mathcal{B}^+(\tilde{b}, b)$. The set $\mathcal{B}^+(\tilde{b}, b)$ is the domain of the relative gap function, defined below.

Definition. *The relative gap function for integer programs, $\gamma : \mathcal{B}^+(\tilde{b}, b) \to \mathbb{R}$, is defined by $\gamma(\beta) = z_{IP}(\beta)/z_{LP}(\beta)$.*

We restrict the domain of the relative gap function for two reasons. First, we avoid division by zero. Second, if $z_{IP}(\beta) < 0$ and $z_{LP}(\beta) < 0$ for some $\beta \in \mathcal{B}(\tilde{b}, b)$, it is possible that $z_{IP}(\beta)/z_{LP}(\beta) > 1$, which decreases the interpretability of this ratio. As a consequence of the domain, $\gamma(\beta) \in [0, 1]$ for all $\beta \in \mathcal{B}^+(\tilde{b}, b)$. The relative gap function also has the additional property that it evaluates models in a way that is invariant to scaling transformations of the objective function.

To solve mathematical programs with integer variables, Beale and Tomlin (1970) introduce Special Ordered Sets of Type 1 (SOS1), ordered sets in which at most one variable can be nonzero. SOS1 constraints typically replace constraints involving binary variables in which the sum of the binary variables is at most 1. Given a finite set $\mathcal{I}$, let $\text{SOS1}(\{y(i)\}_{i \in \mathcal{I}})$ denote a special-ordered-set constraint of type 1 on the decision variable vector y. That is, $|\{i \in \mathcal{I} \mid y(i) > 0\}| \leq 1$. Special Ordered Sets have computational advantages (Beale and Tomlin 1970, Williams 2013), over using binary variables. In a branching algorithm, special ordered sets can be exploited algorithmically; multiple variables are set to zero simultaneously in each new branch. This can decrease the depth of the branch-and-bound tree, for example.

Finally, we briefly mention the complexity of gap function problems. Complexity theory relevant to gap functions includes early results on the complexity of integer programming (for reference, see Garey and Johnson (1979)) and more specific results, such as those in Eisenbrand et al. (2013). It is NP-hard to determine if a constant bounds the absolute gap function over a given set of right-hand sides, and an analogous result holds for the relative gap function. These results follow from the well-known facts that it is NP-hard to test if a constant bounds the value function z_{IP} over a given set of right-hand sides and that linear programs can be solved in polynomial time (Khachiyan 1980).

2.2 Parametrized Superadditive Duality

A function $\phi : D \to \mathbb{R}$ is *superadditive* if $\phi(\beta_1)+\phi(\beta_2) \le \phi(\beta_1+\beta_2)$, for all $\beta_1, \beta_2, \beta_1+\beta_2 \in D$. It is well known that z_{LP} and z_{IP} are superadditive functions over the sets of right-hand sides for which the optimization problems are feasible (Jeroslow 1979, Johnson 1980, Tind and Wolsey 1981). Other early results in superadditive duality can be found in Johnson (1979) and Wolsey (1981). Recall the *superadditive dual* to IP(β), which we denote by SDP(β) (Jeroslow 1979, Wolsey 1981):

$$\begin{aligned} z_{SDP}(\beta) = \min_{\phi}\ & \phi(\beta) \\ \text{s.t.}\ & \phi(a^j) \ge c_j, \text{ for all } j \in \{1, ..., n\}, \\ & \phi(0) = 0, \\ & \phi : \mathbb{R}^m \to \mathbb{R} \text{ is superadditive and nondecreasing.} \end{aligned} \tag{1}$$

The *superadditive dual value function* is $z_{SDP} : \mathcal{B}(\tilde{b}, b) \to \mathbb{R} \cup \{-\infty, \infty\}$. If SDP($\beta$) is unbounded, we define $z_{SDP}(\beta) = -\infty$. If SDP(β) is infeasible, we define $z_{SDP}(\beta) = \infty$. The following are standard results in superadditive duality theory.

Theorem 1. (Jeroslow 1979, Wolsey 1981) *(Weak Duality) For any $x \in S(\beta)$ and any ϕ feasible for SDP(β), $c^T x \le \phi(\beta)$.*

Corollary 1. (Jeroslow 1979, Wolsey 1981) *Let z_{IP} and z_{SDP} be defined as above. Then $z_{IP}(\beta) \leq z_{SDP}(\beta)$ for all β. Further, $z_{IP}(\beta) = \infty$ implies SDP(β) is infeasible, and $z_{SDP}(\beta) = -\infty$ implies IP(β) is infeasible.*

Theorem 2. (Jeroslow 1979, Wolsey 1981) *(Strong Duality)*

1. *If either IP(β) or SDP(β) has a finite optimal value, then any optimal solution x^* to IP(β) and any optimal solution ϕ^* to SDP(β) satisfy $c^T x^* = \phi^*(\beta)$.*

2. *If IP(β) is infeasible, either SDP(β) is infeasible or $z_{SDP}(\beta) = -\infty$.*

3. *If SDP(β) is infeasible, either IP(β) is infeasible or $z_{IP}(\beta) = \infty$.*

Notice that feasible solutions of (1) are real-valued functions, which z_{IP} may not be. An *extension* of a function $f : D \to \mathbb{R} \cup \{-\infty, \infty\}$ is a function $g : D' \to \mathbb{R}$ such that $D' \supseteq D$ and $g(\beta) = f(\beta)$, for all $\beta \in D$ such that $f(\beta) \in \mathbb{R}$.

Proposition 1. *(Blair and Jeroslow 1977) The value function z_{IP} has an extension $\phi^* : \mathbb{R}^m \to \mathbb{R}$ such that ϕ^* is also the value function of a mixed-integer program.*

An important implication of Proposition 1 is that the value function of an integer program can be extended to a real-valued function that maintains its superadditive and nondecreasing properties.

Corollary 2. *For any β such that SDP(β) has an optimal solution, there exists an optimal solution ϕ^* of SDP(β) such that $\phi^*(\tilde{\beta}) = z_{IP}(\tilde{\beta})$ for all $\tilde{\beta}$ satisfying $S(\tilde{\beta}) \neq \emptyset$.*

We make the following assumption to avoid arithmetic with infinity when both IP(β) and LP(β) are unbounded for some $\beta \in \mathcal{B}(\tilde{b}, b)$.

A2. *The right-hand side b is nonnegative, $z_{IP}(b) < \infty$, and LP($\tilde{b}$) is feasible.*

Meyer (1974) shows that if A and β are rational and IP(β) is feasible, then $z_{IP}(\beta) = \infty$ if and only if $z_{LP}(\beta) = \infty$. IP(b) is feasible (the zero vector is a solution), and the data are rational (in fact, integral); thus, **A2** implies that $z_{LP}(b) < \infty$. The monotonicity of z_{IP} implies that $z_{IP}(\beta) \leq z_{IP}(b)$, and the monotonicity of z_{LP} implies that $z_{LP}(\beta) \geq z_{LP}(\tilde{b}) > -\infty$, for all $\beta \in \mathcal{B}(\tilde{b}, b)$. Hence, z_{IP} is bounded from above, and z_{LP} is bounded from above and below.

When $A \in \mathbb{Z}_+^{m\times n}$ and $\beta \in \mathbb{Z}_+^m$, the superadditive dual can be formulated as a linear program (Wolsey 1981). Consider the following set of linear constraints:

$$\phi(a^j) \geq c_j, \text{ for all } j \in \{1, ..., n\} \text{ with } a^j \leq \beta, \tag{2a}$$

$$\phi(\beta_1) + \phi(\beta_2) \leq \phi(\beta_1 + \beta_2), \text{ for all } \beta_1, \beta_2, \beta_1 + \beta_2 \in \widehat{\mathcal{B}}(0, \beta), \tag{2b}$$

$$\phi(\tilde{\beta}) \geq 0, \text{ for all } \tilde{\beta} \in \widehat{\mathcal{B}}(0, \beta), \tag{2c}$$

$$\phi(0) = 0. \tag{2d}$$

Note that ϕ represents a vector, not a function, that is indexed by the elements of $\widehat{\mathcal{B}}(0, \beta)$, and we denote the elements of the vector by $\phi(\beta')$, where $\beta' \in \widehat{\mathcal{B}}(0, \beta)$. Define the polyhedron $\Phi(\beta) = \{\phi \in \mathbb{R}^{|\widehat{\mathcal{B}}(0,\beta)|} \mid (2a) - (2d)\}$. Similarly, the *truncated value function* z_{IP} is a vector whose indices correspond to right-hand sides and the element of the vector corresponding to the right-hand side β is the value function evaluated at β.

Proposition 2. *Suppose $A \in \mathbb{Z}_+^{m\times n}$ and $b \in \mathbb{Z}_+^m$. The truncated value function $z_{IP} \in \mathbb{R}^{|\widehat{\mathcal{B}}(0,b)|}$ is in the polyhedron $\Phi(b)$.*

We use the polyhedron $\Phi(\beta)$ in the following linear programming formulation:

$$z_{SDP}(\beta) = \min\{\phi(\beta) \mid \phi \in \Phi(\beta)\}.$$

We denote this formulation by SDP1(β).

Theorem 3. (Wolsey 1981) *Suppose $A \in \mathbb{Z}_+^{m\times n}$ and $b \in \mathbb{Z}_+^m$. Then SDP1(β) is a strong dual of IP(β), for all $\beta \in \widehat{\mathcal{B}}(0, b)$.*

Notice that the feasible region $\Phi(\beta)$ of SDP1(β) changes with respect to the right-hand side parameter. We thus parametrize the superadditive dual over $\mathcal{B}(0, b)$ with the fixed feasible region $\Phi(b)$, which is useful for formulations in Sections 3-6 in which the value of $\phi(\beta)$ for each $\beta \in \widehat{\mathcal{B}}(0, b)$ is important. Denote the fixed feasible region formulations by SDP2(β):

$$\min\{\phi(\beta) \mid \phi \in \Phi(b)\}.$$

Note that SDP1(b) and SDP2(b) are the same optimization problem.

Theorem 4. *Suppose $A \in \mathbb{Z}_+^{m\times n}$ and $b \in \mathbb{Z}_+^m$. For each $\beta \in \widehat{\mathcal{B}}(0, b)$, SDP2($\beta$) is a strong dual to IP(β).*

Remark 1. *Theorem 4 implies that, given $\beta \in \widehat{\mathcal{B}}(0, b)$, if $\phi^* \in \Phi(b)$ is an optimal solution to SDP2(β), then $\phi^*(\beta) = z_{IP}(\beta)$. However, it does not claim that $\phi^*(\beta') = z_{IP}(\beta')$ for all $\beta' \in \widehat{\mathcal{B}}(0, b)$ (see Example 2.2). Theorem 8 in Section 2.5 establishes $z_{IP}(\beta') = \phi^*(\beta')$ for a subset of $\widehat{\mathcal{B}}(0, b)$.*

Corollary 3. *Suppose $A \in \mathbb{Z}_+^n$ and $b \in \mathbb{Z}_+^m$. For any $\phi \in \Phi(b), \phi(\beta) \geq z_{IP}(\beta)$, for all $\beta \in \widehat{\mathcal{B}}(0, b)$.*

Example.

Let $A = [2\ 1]$ and $c = [3\ 1]^T$. Consider the following family of integer programs:

$$\max\ \{3x_1 + x_2 \mid 2x_1 + x_2 \leq \beta, x \in \mathbb{Z}_+^2\},$$

where $\beta \in \widehat{\mathcal{B}}(0, 4)$.

Consider $\beta = 2$. Let $\phi^*(0) = 0, \phi^*(1) = 1, \phi^*(2) = 3, \phi^*(3) = 5$, and $\phi^*(4) = 6$. It can be seen that ϕ^* is an optimal solution to SDP2(β). In addition, $\phi^*(2) = 3 = z_{IP}(2)$. However, the value function does not coincide with ϕ^* everywhere, as $\phi^*(3) = 5 > z_{IP}(3) = 4$. □

A key point of the new superadditive dual formulations SDP2 is that the feasible region is the same, regardless of the right-hand side parameter. We exploit this property in Sections 3-6.

2.3 Gap Functions and Level Sets

A *level set* of the integer programming value function z_{IP} is a set of the form $\{\beta \in \mathcal{B}(\tilde{b}, b) \mid z_{IP}(\beta) = \alpha\}$, for some $\alpha \in \mathbb{Z}$. Trapp et al. (2013) define a vector $\beta \in \mathbb{Z}^m$ as "level-set minimal" for IP(b) if $z_{IP}(\beta - e_i) < z_{IP}(\beta)$ for all $i \in \{1, ..., m\}$ such that $\beta - e_i \in \mathcal{B}(\tilde{b}, b)$, where e_i is the i^{th} standard basis vector of $\mathbb{R}^m$. Our modified definition of level-set-minimal vectors incorporates the boundary of the right-hand side set, and we also provide a definition for a natural analog, level-set-maximal vectors.

Definition. *For the integer programming value function z_{IP} and right-hand side set $\mathcal{B}(\tilde{b}, b)$,*

1. *$\beta \in \mathcal{B}(\tilde{b}, b)$ is level-set minimal if $z_{IP}(\beta - e_i) < z_{IP}(\beta)$ for all $i \in \{1, ..., m\}$ such that $\beta - e_i \in \mathcal{B}(\tilde{b}, b)$. Let D^- denote the level-set-minimal vectors.*

2. *$\beta \in \mathcal{B}(\tilde{b}, b)$ is level-set maximal if $z_{IP}(\beta + e_i) > z_{IP}(\beta)$ for all $i \in \{1, ..., m\}$ such that $\beta + e_i \in \mathcal{B}(\tilde{b}, b)$. Let D^+ denote the level-set-maximal vectors.*

Trapp et al. (2013) note that level-set-minimal vectors are a generalization of minimal tenders, used by Kong et al. (2006). Trapp et al. (2013) also state it may "be computationally advantageous to optimize over" the set of level-set-minimal vectors. Specifically, Trapp et al. (2013) use a value function reformulation of a two-stage stochastic program and show that an optimal solution exists over the sets of level-set-minimal vectors. This leads to substantially less memory usage. The trade-off is that it is an NP-complete problem to verify that a vector is not level-set minimal (Trapp et al. 2013). Still, Trapp et al. (2013) provide an algorithm for generating a nontrivial superset of the level-set-minimal vectors.

We show that every level set of z_{IP} (over the bounded hyper-rectangle $\mathcal{B}(\tilde{b}, b)$) has a level-set-minimal vector and a level-set-maximal vector. Trapp et al. (2013) prove this statement for level-set-minimal vectors under a different definition without bounded hyper-rectangles. Because $\mathcal{B}(\tilde{b}, b)$ is bounded, as long as the set is nonempty, there exist non-empty level sets. Hence D^- and D^+ are nonempty.

Proposition 3.

(I) For any $\beta \in \mathcal{B}(\tilde{b}, b)$*, there exists a* $\tilde{\beta} \in D^+$ *with* $\tilde{\beta} \geq \beta$ *such that* $z_{IP}(\tilde{\beta}) = z_{IP}(\beta)$*.*

(II) For any $\beta \in \mathcal{B}(\tilde{b}, b)$*, there exists a* $\tilde{\beta} \in D^-$ *with* $\tilde{\beta} \leq \beta$ *such that* $z_{IP}(\tilde{\beta}) = z_{IP}(\beta)$*.*

We can use Proposition 3 to express the infimum and supremum of Γ (over $\mathcal{B}(\tilde{b}, b)$) and γ (over $\mathcal{B}^+(\tilde{b}, b)$).

Theorem 5. *Level-set-minimal and level-set-maximal vectors are sufficient to bound the gap function. That is,*

(I) $\inf_{\beta \in \mathcal{B}(\tilde{b}, b)} \Gamma(\beta) = \inf_{\beta \in D^-} \Gamma(\beta)$*.*

(II) $\sup_{\beta \in \mathcal{B}(\tilde{b}, b)} \Gamma(\beta) = \sup_{\beta \in D^+} \Gamma(\beta)$*.*

(III) $\inf_{\beta \in \mathcal{B}^+(\tilde{b}, b)} \gamma(\beta) = \inf_{\beta \in D^+ \cap \mathcal{B}^+(\tilde{b}, b)} \gamma(\beta)$*.*

(IV) $\sup_{\beta \in \mathcal{B}^+(\tilde{b}, b)} \gamma(\beta) = \sup_{\beta \in D^- \cap \mathcal{B}^+(\tilde{b}, b)} \gamma(\beta)$*.*

2.4 The Absolute Gap Function is the Minimum of Finitely Many Gomory Functions

Gomory functions are used throughout the integer programming literature (e.g., Blair and Jeroslow (1982, 1985), Williams (1996)).

Definition. *(Blair and Jeroslow 1982) The class* $\mathcal{C}^m$ *of* m*-dimensional Gomory functions is the smallest class* $\mathcal{C}$ *of functions with domain* $\mathbb{Q}^m$*, range* $\mathbb{R}$*, and these properties:*

(i) If there exists $\lambda \in \mathbb{Q}^m$ *such that* $f(\beta) = \lambda^T \beta$*, then* $f \in \mathcal{C}$*;*

(ii) If $f, g \in \mathcal{C}$ *and* $s, t \in \mathbb{Q}_+$*, then* $sf + tg \in \mathcal{C}$*;*

(iii) If $f \in \mathcal{C}$*, then* $\lceil f \rceil \in \mathcal{C}$*, where* $\lceil f \rceil$ *is the function defined by the condition* $\lceil f \rceil(v) = \lceil f(v) \rceil$*;*

(iv) If $f, g \in \mathcal{C}$, then $\max\{f, g\} \in \mathcal{C}$.

When integer programs are posed as minimization problems, Blair and Jeroslow (1982) show that every integer programming value function is a Gomory function.

Theorem 6. *The absolute gap function defined over rational vectors, $\Gamma : \mathcal{B}(\tilde{b}, b) \cap \mathbb{Q}^m \to \mathbb{R}$, is the minimum of finitely many Gomory functions.*

2.5 Periodicity of Gap Functions

For each feasible basis of the linear programming relaxation, Gomory (1965) considers the cone of right-hand sides for which the basis is optimal, and then shows that the absolute gap function, within a subset of the interior of this cone, is periodic with respect to the columns of A. We use superadditive duality (Proposition 5 and Theorem 7) to reproduce this result. First, Definition 2.5 provides a way to describe parts of level sets of the linear programming value function.

Definition. *For any $\beta \in \mathcal{B}(\tilde{b}, b), x, \hat{x} \in \mathbb{R}^n_+$ such that $x \leq \hat{x}$, define $\Lambda(\beta, x, \hat{x}) = \{\lambda \in \mathbb{R}^m_+ \mid Ax \leq \lambda \leq \beta - A(\hat{x} - x)\}$.*

Proposition 4. *For any $\beta \in \mathcal{B}(\tilde{b}, b)$, if $\hat{x} \in opt_{LP}(\beta)$, then for all x such that $0 \leq x \leq \hat{x}$, $x \in opt_{LP}(\tilde{\beta})$, for all $\tilde{\beta} \in \Lambda(\beta, x, \hat{x})$.*

Corollary 4. *Let $\beta, x, \hat{x}$ satisfy the conditions in Proposition 4. Suppose further that $x \in \mathbb{Z}^n_+$. Then for any $\tilde{\beta} \in \Lambda(\beta, x, \hat{x})$, $\Gamma(\tilde{\beta}) = 0$.*

Proposition 5. *Let $\beta \in \mathcal{B}(\tilde{b}, b)$. For all $j \in \{1, ..., n\}$ such that there exists $x^I \in opt_{IP}(\beta)$ and $x^L \in opt_{LP}(\beta)$ with $x^I_j \geq 1$ and $x^L_j \geq 1$, we have $z_{LP}(\beta - a^j) = z_{LP}(\beta) - c_j$, and $z_{IP}(\beta - a^j) = z_{IP}(\beta) - c_j$.*

Theorem 7. *Under the conditions of Proposition 5, $\Gamma(\beta - a^j) = \Gamma(\beta)$. If, in addition, $z_{LP}(\beta) > c_j$ and $z_{IP}(\beta) \geq c_j$, then $\gamma(\beta - a^j) = (z_{IP}(\beta) - c_j)/(z_{LP}(\beta) - c_j)$.*

While Theorem 7 refers to the periodicity of the gap functions, Theorem 8 shows that optimal solutions to the superadditive dual problem SDP2 also have a periodic property.

Theorem 8. *Let $\beta \in \widehat{\mathcal{B}}(0,b)$ and suppose $A \in \mathbb{Z}_+^{m\times n}$ and $b \in \mathbb{Z}_+^m$. For all $j \in \{1,...,n\}$ such that there exists $x \in opt_{IP}(\beta)$ with $x_j \geq 1$, any optimal solution ϕ^* of SDP2(β) satisfies $\phi^*(\beta - a^j) = z_{IP}(\beta - a^j)$.*

3 Absolute Gap Problems over a Discrete Set

In this section, we formulate optimization problems that compute the expectation, maximum, and minimum of the absolute gap function Γ over a discrete set of right-hand sides. The expectation provides global information for approximation quality weighted by a probability distribution. The maximum is a worst-case approximation bound. The minimum determines which formulations (if any) are perfect—integer programming models whose optimal objective value is the optimal objective value of its linear programming relaxation. These three measures of the strength of parametrized integer programming models can be computed through formulations with either linear constraints alone, or linear constraints and a single special-ordered-set constraint.

For the objective values of formulations throughout the remainder of this paper, we use the following nomenclature: The first letter indicates the quality measure (expectation, supremum, or infimum). The second letter denotes the gap function (absolute or relative). The third letter stands for the set of right-hand sides (discrete set or hyper-rectangle). For example, δ_{IAD} denotes the optimal objective value of the infimum absolute gap problem over a discrete set of right-hand sides. Table 1 lists all problems covered in this paper.

A3. *The problem data are nonnegative: $A \in \mathbb{Z}_+^{m\times n}, b \in \mathbb{Z}_+^m$, and $c \in \mathbb{Z}_+^n$. Also, $\tilde{b} = 0$.*

A3 holds for the rest of this paper. A consequence of **A3** is that, by Theorem 4, SDP2(β) is a strong dual to IP(β) over the sets of right-hand sides, $\widehat{\mathcal{B}}(0,b)$ and $\mathcal{B}(0,b)$, which are subsets of $\mathbb{R}_+^m$. We use supremum and infimum for notational consistency (see

Section 4); however, the maximum and minimum absolute gap over $\widehat{\mathcal{B}}(0,b)$ exist because $\widehat{\mathcal{B}}(0,b)$ is a finite set, and $0 \leq |\Gamma(\beta)| < \infty$ for all $\beta \in \widehat{\mathcal{B}}(0,b)$, which follows from **A2** and $\widehat{\mathcal{B}}(0,b) \subset \mathbb{R}^m_+$.

3.1 Expectation of the Absolute Gap Function over a Discrete Set

Let $\xi^{(1)}$ be a discrete random variable with event space $\widehat{\mathcal{B}}(0,b)$ and $\mathbf{P}\{\xi^{(1)} = \beta\} = \mu^{(1)}(\beta)$. The superscript (1) denotes the size of the discretization in $\widehat{\mathcal{B}}(0,b)$ (see Section 4.1). The expected absolute gap over $\widehat{\mathcal{B}}(0,b)$ is given by

$$\mathbb{E}_{\xi^{(1)}}[\Gamma(\xi^{(1)})] = \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(1)}(\beta)\Gamma(\beta).$$

Note that $\mathbb{E}_{\xi^{(1)}}[\Gamma(\xi^{(1)})]$ is not the expectation of the absolute gap function over $\mathcal{B}(0,b)$ (see Section 4.1). The expectation of the absolute gap function provides information about the integer programming model's quality over the set of right-hand sides. Namely, one gains an idea of the approximation quality of linear programming relaxations when solving many integer programs parametrized by right-hand sides in $\widehat{\mathcal{B}}(0,b)$. Furthermore, right-hand sides that are weighted more heavily contribute more to this metric. This metric may be of particular interest, in general, for stochastic programming in the case when the recourse matrix is fixed and bounds on the first-stage variables and stochastic right-hand sides of the second stage produce bounded right-hand side sets for the second-stage problem. In such a setting, the expected absolute gap may indicate that a relaxation of the second-stage problem (if it is an integer program) is a good approximation.

The following superadditive-dual-based formulation (3) computes $\mathbb{E}_{\xi^{(1)}}[\Gamma(\xi^{(1)})]$:

$$\delta_{EAD} = \max \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(1)}(\beta)\psi(\beta)$$

$$\text{s.t. } \psi(\beta) \leq \beta^T v^k - \phi(\beta), \text{ for all } k \in \mathcal{K}, \beta \in \widehat{\mathcal{B}}(0,b), \tag{3a}$$

$$\psi \in \mathbb{R}^{|\widehat{\mathcal{B}}(0,b)|}, \tag{3b}$$

$$\phi \in \Phi(b). \tag{3c}$$

Proposition 6. *The optimal objective value of* (3) *is* $\mathbb{E}_{\xi^{(1)}}[\Gamma(\xi^{(1)})]$.

Corollary 5. *Suppose* $\mu^{(1)}(\beta) > 0$ *for all* $\beta \in \widehat{\mathcal{B}}(0,b)$. *Then, any feasible solution* $(\bar{\phi}, \bar{\psi})$ *for which* $\bar{\phi} \neq z_{IP}$ *is not optimal for* (3).

The necessary condition of Corollary 5 is a consequence of the set $\Phi(b)$. A key point is that because $\phi \in \Phi(b)$, as opposed to $\Phi(\beta)$ for some $\beta \leq b, \beta \neq b$, every component of ϕ is constrained. When $\mu^{(1)}(\beta) > 0$, for all $\beta \in \widehat{\mathcal{B}}(0,b)$, it is optimal to minimize each component $\phi(\beta)$ in the same way that $\phi'(\beta)$ is minimized for an optimal solution ϕ' of SDP2(β).

The variance of the absolute gap function, denoted by σ_Γ^2 indicates how much the model's linear programming relaxation approximation quality changes over the discrete set. The variance over the discrete set $\widehat{\mathcal{B}}(0,b)$ can be calculated from optimal solutions of (3).

Corollary 6. *Let* $(\tilde{\phi}, \tilde{\psi})$ *be an optimal solution to* (3) *and* δ_{EAD} *the optimal objective value. Then the variance of the absolute gap over* $\widehat{\mathcal{B}}(0,b)$ *can be calculated as follows:*

$$\sigma_\Gamma^2 = \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(1)}(\beta)(\tilde{\psi}(\beta))^2 - \delta_{EAD}^2.$$

3.2 Supremum of the Absolute Gap Function over a Discrete Set

The expected absolute gap is one measure of the quality of an integer programming model over a set of right-hand sides. Because the absolute gap function is nonnegative, the expectation of the absolute gap function is equal to a weighted ℓ_1 norm of the absolute gap function when all of the probabilities are positive. In many cases, however, it is useful to consider other metrics. In particular, the supremum norm is also an important characteristic of a function. For the absolute gap function, the supremum can be used to determine the worst-case performance of the linear programming relaxation as an approximation for the integer program.

The supremum of the absolute gap function over the discrete set $\widehat{\mathcal{B}}(0,b)$ is:

$$\Delta_{SAD} = \sup_{\beta \in \widehat{\mathcal{B}}(0,b)} \Gamma(\beta) = \max_{\beta \in \widehat{\mathcal{B}}(0,b)} \Gamma(\beta).$$

Formulation (4) computes the supremum of the absolute gap function over $\widehat{\mathcal{B}}(0,b)$.

$$\begin{aligned} \delta_{SAD} = \max \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} & \psi(\beta) \\ \text{s.t. } & \psi(\beta) \leq \beta^T v^k - \phi(\beta), \text{ for all } \beta \in \widehat{\mathcal{B}}(0,b), k \in \mathcal{K}, && \text{(4a)} \\ & \text{SOS1}\left(\{\psi(\beta)\}_{\beta \in \widehat{\mathcal{B}}(0,b)}\right), && \text{(4b)} \\ & \psi \in \mathbb{R}^{|\widehat{\mathcal{B}}(0,b)|}, && \text{(4c)} \\ & \phi \in \Phi(b). && \text{(4d)} \end{aligned}$$

Proposition 7. *The optimal objective value of* (4) *is* Δ_{SAD}.

3.3 Infimum of the Absolute Gap Function over a Discrete Set

We also consider the other extremum of the absolute gap function. The infimum of the absolute gap function represents the linear programming relaxation's best-case approximation over the set of right-hand sides. By computing the infimum of the absolute gap function, one can determine if a perfect formulation exists within the family of integer programs. In particular, we exclude $\{0\}$ from consideration, as trivially, $\Gamma(0) = 0 = \min_{\beta \in \widehat{\mathcal{B}}(0,b)} \Gamma(\beta)$.

The infimum of the absolute gap function over $\widehat{\mathcal{B}}(0,b) \backslash \{0\}$ is given by:

$$\Delta_{IAD} = \inf_{\beta \in \widehat{\mathcal{B}}(0,b) \backslash \{0\}} \Gamma(\beta) = \min_{\beta \in \widehat{\mathcal{B}}(0,b) \backslash \{0\}} \Gamma(\beta).$$

Formulation (5) computes the infimum of the absolute gap function over $\widehat{\mathcal{B}}(0,b)\backslash\{0\}$.

$$
\begin{aligned}
\delta_{IAD} = \max\ & \psi \\
\text{s.t. }\ & \psi \leq \beta^T v^k - \phi(\beta), \text{ for all } \beta \in \widehat{\mathcal{B}}(0,b)\backslash\{0\}, k \in \mathcal{K}, && \text{(5a)} \\
& \psi \in \mathbb{R}, && \text{(5b)} \\
& \phi \in \Phi(b). && \text{(5c)}
\end{aligned}
$$

Proposition 8. *The optimal objective value of* (5) *is* Δ_{IAD}.

4 Absolute Gap Problems over a Hyper-rectangle

Gap functions can also evaluate integer programming models over non-discrete sets, such as a hyper-rectangle. In this section, we provide formulations to: estimate $\mathbb{E}_\xi(\Gamma(\xi))$, the expectation of Γ over $\mathcal{B}(0,b)$, compute $\Delta_{SAH} = \sup\{\Gamma(\beta) \mid \beta \in \mathcal{B}(0,b)\}$, and compute $\Delta_{IAH} = \inf\{\Gamma(\beta) \mid \beta \in \mathcal{B}(0,b)\backslash\mathcal{B}(0,1)\}$, where $\mathcal{B}(0,1)$ is the unit hyper-cube anchored at the vector of zeros and the vector of ones in $\mathbb{R}^m$.

Similar to Section 3.3, we do not consider $\mathcal{B}(0,1)$ when computing the infimum, as trivially, $\Gamma(0) = 0$. Note that the absolute gap function does not necessarily achieve its supremum, even on the compact set $\mathcal{B}(0,b)$, because z_{IP} is generally not continuous. This indicates that a brute-force sampling approach may only provide bounds on the absolute gap function's supremum. We show that when the set of right-hand sides is $\mathcal{B}(0,b)$, we can compute the supremum of the absolute gap function by considering only integral points. We also show that the infimum is always achieved at an integral point. The nomenclature in this section's optimization problems follows that of Section 3.

4.1 Approximating the Expectation of the Absolute Gap Function over a Hyper-rectangle

We address how to approximate the expectation of the absolute gap function over the hyper-rectangle $\mathcal{B}(0,b)$. This is an inherently different problem from the other problems considered in this paper as it requires information about the absolute gap function over a non-discrete set. Let ξ be a continuous random variable with event space $\mathcal{B}(0,b)$ and $\mathbf{P}\{\xi=\beta\}=\mu(\beta)$. The expectation of the absolute gap function over $\mathcal{B}(0,b)$ is

$$\mathbb{E}_{\xi}[\Gamma(\xi)] = \int_{\beta\in\mathcal{B}(0,b)} \mu(\beta)\Gamma(\beta).$$

We provide a framework that approximates the expectation of the gap function arbitrarily well. Let $p\in\mathbb{N}$, and let $\widehat{\mathcal{B}}_p(0,b)=\{\beta\in\mathcal{B}(0,b)\mid \beta_i=\alpha_i/p, \alpha_i\in\mathbb{Z}^m_+\}$. Thus, $\widehat{\mathcal{B}}_p(0,b)\supseteq\widehat{\mathcal{B}}(0,b)$. Similar to Section 3.1, we let $\xi^{(p)}$ be a discrete random variable with event space $\widehat{\mathcal{B}}_p(0,b)$ and $\mathbf{P}\{\xi^{(p)}=\beta\}=\mu^{(p)}(\beta)$. Here, p denotes the discretization size of $\widehat{\mathcal{B}}_p(0,b)$. The probability mass function $\mu^{(p)}$ can be viewed as an approximation of the probability density function μ. The expected absolute gap over $\widehat{\mathcal{B}}_p(0,b)$ is given by

$$\mathbb{E}_{\xi^{(p)}}[\Gamma(\xi^{(p)})] = \sum_{\beta\in\widehat{\mathcal{B}}_p(0,b)} \mu^{(p)}(\beta)\Gamma(\beta).$$

The following superadditive-dual-based formulation (6) exactly computes the expectation of the absolute gap over $\widehat{\mathcal{B}}_p(0,b)$:

$$\delta_{EAH} = \max \sum_{\beta\in\widehat{\mathcal{B}}_p(0,b)} \mu^{(p)}(\beta)\psi(\beta)$$

$$\text{s.t. } \psi(\beta)\le\beta^T v^k-\phi(\lfloor\beta\rfloor), \text{ for all } k\in\mathcal{K},\beta\in\widehat{\mathcal{B}}_p(0,b), \tag{6a}$$

$$\psi\in\mathbb{R}^{|\widehat{\mathcal{B}}_p(0,b)|}, \tag{6b}$$

$$\phi\in\Phi(b). \tag{6c}$$

Proposition 9. *The optimal objective value of* (6) *is* $\delta_{EAH} = \mathbb{E}_{\xi^{(p)}}[\Gamma(\xi^{(p)})]$.

Corollary 7. *Suppose* $\mu^{(p)}(\beta) > 0$ *for all* $\beta \in \widehat{\mathcal{B}}_p(0, b)$. *Then, any feasible solution* $(\bar{\phi}, \bar{\psi})$ *in which* $\bar{\phi} \neq z_{IP}$ *is not optimal for* (6).

Corollary 8. *Let* $(\tilde{\phi}, \tilde{\psi})$ *be an optimal solution to* (6) *and* δ_{EAH} *the optimal objective value. Then the variance of the absolute gap over* $\widehat{\mathcal{B}}_p(0, b)$ *can be calculated as follows:*

$$\sigma_\Gamma^2 = \sum_{\beta \in \widehat{\mathcal{B}}_p(0,b)} \mu^{(p)}(\beta)(\tilde{\psi}(\beta))^2 - \delta_{EAH}^2.$$

Next, we show how to approximate the absolute gap function by extending part of optimal solutions of (6) to functions over $\mathcal{B}(0, b)$, and we derive a guarantee of the approximation quality.

Proposition 10. *(Mangasarian and Shiau 1987) The linear programming value function* z_{LP} *is Lipschitz continuous over* $\mathcal{B}(0, b)$.

We use Proposition 10 to make approximation guarantees for the expectation over the hyper-rectangle. Let $M = \min\{\bar{M} \in \mathbb{R}_{++} \mid \bar{M}||\beta^1 - \beta^2||_\infty \geq |z_{LP}(\beta^1) - z_{LP}(\beta^2)|, \text{ for all } \beta^1, \beta^2 \in \mathcal{B}(0, b)\}$, the best Lipschitz constant of z_{LP} over $\mathcal{B}(0, b)$. Define the *p-floor* function $\lfloor \cdot \rfloor^{(p)} : \mathcal{B}(0, b) \to \mathbb{R}$ by $(\lfloor \beta \rfloor^{(p)})_i = \lfloor \alpha_i \rfloor / p$, where $\beta_i = \alpha_i / p$.

Theorem 9. *Suppose that for all* $\beta \in \widehat{\mathcal{B}}_p(0, b), \mu^{(p)}(\beta) > 0$. *Let* $(\tilde{\phi}, \tilde{\psi})$ *be an optimal solution to* (6), *and define the extension* $\tilde{\psi}_p : \mathcal{B}(0, b) \to \mathbb{R}$ *by* $\tilde{\psi}_p(\beta) = \tilde{\psi}_p(\lfloor \beta \rfloor^{(p)})$. *Then the following results hold:*

$$||\Gamma - \tilde{\psi}_p||_\infty \leq M/p, \tag{7a}$$

$$\mathbb{E}_\xi|\Gamma(\xi) - \tilde{\psi}_p(\xi)| = \mathbb{E}_\xi(\Gamma(\xi) - \tilde{\psi}_p(\xi)) \leq M/p. \tag{7b}$$

We show that if the discrete probability mass function $\mu^{(p)}$ is an approximation of the probability density function μ, then the difference between $\mathbb{E}_\xi[\Gamma(\xi)]$ and objective value δ_{EAH}

is bounded. Cover $\mathcal{B}(0,b)$ by $T^{(p)} = \prod_{i=1}^{m}(pb_i)$ hyper-cubes, denoted $\mathcal{B}^t = \bigotimes_{i=1}^{m}[b_i^t, b_i^t + 1]$, with vertices on the integral points, where $b^t \in \widehat{\mathcal{B}}_p(0,b)$ and $b_i^t < b_i$ for all $i \in \{1, ..., m\}$.

Theorem 10. *Suppose the probability mass function $\mu^{(p)}$ approximates the probability density function μ as follows: $\mu^{(p)}(b^t) = \int_{\mathcal{B}^t} \mu(\beta)d\beta$, for all $t \in \{1, ..., T^{(p)}\}$. Then, the difference between the expectation of the absolute gap function over $\mathcal{B}(0,b)$ and the optimal objective value of (6) is bounded. Specifically, $|\mathbb{E}_\xi[\Gamma(\xi)] - \delta_{EAH}| \leq M/p$.*

Theorems 9 and 10 state that (6) produces a solution that can be used to approximate the expectation of the absolute gap function when at least one of two conditions are satisfied. The first is that the probability mass function $\mu^{(p)}$ is nonzero over all of $\widehat{\mathcal{B}}_{(p)}(0,b)$. The second condition is that $\mu^{(p)}$ approximates the probability mass function μ. We remark that a guarantee of the approximation of the variance is also possible, using the results of Theorem 9, for instance.

4.2 Supremum of the Absolute Gap Function over a Hyper-rectangle

Cover $\mathcal{B}(0,b)$ by $L = \Pi_{i=1}^{m} b_i$ unit hyper-cubes, denoted $\mathcal{B}^l = \bigotimes_{i=1}^{m}[b_i^l, b_i^l + 1]$, with vertices on the integral points, where $b^l \in \widehat{\mathcal{B}}(0,b)$ and $b_i^l < b_i$ for all $i \in \{1, ..., m\}$. Notice that

$$\sup\{z_{LP}(\beta) - z_{IP}(\beta) \mid \beta \in \mathcal{B}(0,b)\} = \max_{l \in \{1,...,L\}} \sup\{z_{LP}(\beta) - z_{IP}(\beta) \mid \beta \in \mathcal{B}^l\}.$$

Proceed by considering each unit hyper-cube, $\mathcal{B}^l, l \in \{1, ..., L\}$ independently. For each $\beta \in \mathcal{B}(0,b)$, let $r(\beta) = \beta + \sum_{\{i \mid \beta_i \leq b_i - 1\}} e_i$. Let $d^l = r(b^l)$ be the "top-right" corner of the unit hyper-cube $\mathcal{B}^l$. The "bottom-left" corner of $\mathcal{B}^l$ is b^l.

Proposition 11. *For any unit hyper-cube $\mathcal{B}^l$, $\sup_{\beta \in \mathcal{B}^l} \Gamma(\beta) = z_{LP}(d^l) - z_{IP}(b^l)$.*

Corollary 9. $\Delta_{SAH} = \sup_{\beta \in \mathcal{B}(0,b)} \Gamma(\beta) = \sup_{l \in \{1,...,L\}} z_{LP}(d^l) - z_{IP}(b^l)$.

Formulation (8) computes the supremum of the absolute gap function over $\mathcal{B}(0,b)$, and it provides the model's worst-case linear programming relaxation approximation over the indicated set of right-hand sides.

$$
\begin{aligned}
\delta_{SAH} = \max \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} & \psi(\beta) \\
\text{s.t. } & \psi(\beta) \leq r(\beta)^T v^k - \phi(\beta), \text{ for all } \beta \in \widehat{\mathcal{B}}(0,b), k \in \mathcal{K}, && \text{(8a)} \\
& \text{SOS1}\left(\{\psi(\beta)\}_{\beta \in \widehat{\mathcal{B}}(0,b)}\right), && \text{(8b)} \\
& \psi \in \mathbb{R}^{|\widehat{\mathcal{B}}(0,b)|}, && \text{(8c)} \\
& \phi \in \Phi(b). && \text{(8d)}
\end{aligned}
$$

Proposition 12. *The optimal solution of* (8) *is* Δ_{SAH}.

There is a strong connection between models (4) and (8). For each $l \in \{1, ..., L\}$ and each $k \in \mathcal{K}$, $(\beta^l)^T v^k - \phi(\beta^l)$ is an upper bound on $\Gamma(\beta^l)$. In the same light, $r(\beta^l)^T v^k - \phi(\beta^l)$ is an upper bound on $\sup_{\beta \in \mathcal{B}^l} \Gamma(\beta)$. This modified upper bound now allows the supremum of the absolute gap function to be computed given the results of Proposition 11 and Corollary 9.

4.3 Infimum of the Absolute Gap Function over a Hyper-rectangle

Unlike the supremum, the infimum of the absolute gap function over $\mathcal{B}(0,b) \backslash \mathcal{B}(0,1)$ is always achieved.

Proposition 13. *The infimum of the absolute gap function over* $\mathcal{B}(0,b) \backslash \mathcal{B}(0,1)$ *can be computed by considering only the integral points in* $\mathcal{B}(0,b)$. *Specifically,* $\inf_{\beta \in \mathcal{B}(0,b) \backslash \mathcal{B}(0,1)} \Gamma(\beta) = \min_{\beta \in \widehat{\mathcal{B}}(0,b) \backslash \{0\}} \Gamma(\beta)$. *Further,* (5) *can be used to compute* $\Delta_{IAH} = \inf_{\beta \in \mathcal{B}(0,b) \backslash \mathcal{B}(0,1)} \Gamma(\beta)$.

Proposition 13 states that the infimum of the absolute gap function is achieved at integral points, i.e., if a perfect formulation exists within the hyper-rectangle of right-hand sides, then a perfect formulation exists with integer data. More generally, assuming the set $\mathcal{B}(0,b) \backslash \mathcal{B}(0,1)$ is nonempty, there exists an integral nonempty subset of the right-hand sides for which a model's linear programming relaxation is tightest. Together, Proposition 12 and

Proposition 13 demonstrate that, on both the discrete set $\widehat{\mathcal{B}}(0,b)$ and hyper-rectangle $\mathcal{B}(0,b)$, one can formulate the worst-case and best-case absolute gap problems as linear programs with at most one special-ordered-set constraint.

5 Relative Gap Problems over a Discrete Set

The relative gap function provides another measure of integer programming model quality. The relative gap function evaluates models in a way that is invariant to scaling transformations of the objective function. The relative gap is unbounded or undefined at the origin. Hence, when solving relative gap problems, the domain of the relative gap function is either $\widehat{\mathcal{B}}^+(0,b)$ or $\mathcal{B}^+(0,b)$. Note that $\gamma(\beta) \in [0,1]$, for all $\beta \in \mathcal{B}^+(0,b)$. In addition, $\beta^T v^k > 0$ for all extreme points $v^k, k \in \mathcal{K}$ because $z_{LP}(\beta) > 0$, for all $\beta \in \mathcal{B}^+(0,b)$. The nomenclature in the optimization problems of this section follows that of Section 3.

5.1 Expectation of the Relative Gap Function over a Discrete Set

The expectation of the relative gap function over the discrete set $\widehat{\mathcal{B}}^+(0,b)$ evaluates integer programming models by giving the linear programming relaxations' long-run average approximation error for the integer programs.

Let $\xi^{(1)}$ be a discrete random variable with event space $\widehat{\mathcal{B}}^+(0,b)$ and $\mathbf{P}\{\xi^{(1)} = \beta\} = \mu^{(1)}(\beta)$. The superscript again denotes the discretization size of $\widehat{\mathcal{B}}^+(0,b)$. The expected relative gap over $\widehat{\mathcal{B}}^+(0,b)$ is given by:

$$\mathbb{E}_{\xi^{(1)}}[\gamma(\xi^{(1)})] = \sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \mu^{(1)}(\beta)\gamma(\beta).$$

The linear program (9) computes $\mathbb{E}_{\xi^{(1)}}[\gamma(\xi^{(1)})]$.

$$\delta_{ERD} = \min \sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \mu^{(1)}(\beta)\psi(\beta) \tag{9a}$$

$$\text{s.t. } \psi(\beta)\beta^T v^k \geq \phi(\beta), \text{ for all } \beta \in \widehat{\mathcal{B}}^+(0,b), k \in \mathcal{K}, \tag{9b}$$

$$\psi \in \mathbb{R}^{|\widehat{\mathcal{B}}^+(0,b)|}, \tag{9c}$$

$$\phi \in \Phi(b). \tag{9d}$$

Proposition 14. *The optimal objective value of* (9) *is* $\mathbb{E}_{\xi^{(1)}}(\gamma(\xi^{(1)}))$.

Corollary 10. *Suppose* $\mu^{(1)}(\beta) > 0$ *for all* $\beta \in \widehat{\mathcal{B}}^+(0,b)$. *Then, any feasible solution* $(\bar{\phi}, \bar{\psi})$ *in which* $\bar{\phi} \neq z_{IP}$ *is not optimal for* (9).

The variance of the relative gap function, denoted by σ_γ^2, indicates the extent to which this approximation error varies over $\widehat{\mathcal{B}}^+(0,b)$.

Corollary 11. *Let* $(\tilde{\phi}, \tilde{\psi})$ *be an optimal solution to* (9) *and* δ_{ERD} *the optimal objective value. Then the variance of the relative gap over* $\widehat{\mathcal{B}}(0,b)$ *can be calculated as follows:*

$$\sigma_\gamma^2 = \sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \mu^{(1)}(\beta)(\tilde{\psi}(\beta))^2 - \delta_{ERD}^2.$$

5.2 Supremum of the Relative Gap Function over a Discrete Set

The supremum of the relative gap function represents the integer programming model's best-case approximation error (by the linear programming relaxation) over the set of right-hand sides. The supremum of the relative gap function over $\widehat{\mathcal{B}}^+(0,b)$ is defined as

$$\Delta_{SRD} = \max_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \gamma(\beta).$$

Consider the following optimization problem (10):

$$\delta_{SRD} = \min \psi$$

$$\text{s.t. } \psi \cdot \beta^T v^k \geq \phi(\beta), \text{ for all } \beta \in \widehat{\mathcal{B}}^+(0,b), k \in \mathcal{K}, \tag{10a}$$

$$\psi \in \mathbb{R}, \tag{10b}$$

$$\phi \in \Phi(b). \tag{10c}$$

Proposition 15. *The optimal objective value of* (10) *is* Δ_{SRD}.

Therefore, the best-case relative gap problem over the discrete set $\widehat{\mathcal{B}}^{+}(0,b)$ can be formulated as a linear program. The analog of (10) for the absolute gap function is (5), the best-case absolute gap problem.

5.3 Infimum of the Relative Gap Function over a Discrete Set

To evaluate an integer programming model by worst-case linear programming relaxations, one can use the infimum of the relative gap function. The infimum of the relative gap function over $\widehat{\mathcal{B}}^{+}(0,b)$ is given by

$$\Delta_{IRD} = \min_{\beta \in \widehat{\mathcal{B}}^{+}(0,b)} \gamma(\beta).$$

Formulation (11) can be used to compute the infimum of the relative gap function over $\widehat{\mathcal{B}}^{+}(0,b)$.

$$\delta_{IRD} = \max \sum_{\beta \in \widehat{\mathcal{B}}^{+}(0,b)} \psi(\beta)$$

$$\text{s.t. } (1-\psi(\beta))\beta^T v^k \geq \phi(\beta), \text{ for all } \beta \in \widehat{\mathcal{B}}(0,b), k \in \mathcal{K}, \tag{11a}$$

$$\text{SOS1}\left(\{\psi(\beta)\}_{\beta \in \widehat{\mathcal{B}}^{+}(0,b)}\right), \tag{11b}$$

$$\psi \in \mathbb{R}^{|\widehat{\mathcal{B}}^{+}(0,b)|}, \tag{11c}$$

$$\phi \in \Phi(b). \tag{11d}$$

Proposition 16. *The optimal objective value of* (11) *is* $1 - \Delta_{IRD}$.

The counterpart to (11) for the absolute gap function is (4). Notice that for all $\beta \in \widehat{\mathcal{B}}(0,b)$, the constraint $\psi(\beta) \leq 1$ is implicit because $\phi(\beta) \geq 0$ and $\beta^T v^k \geq 0$.

6 Relative Gap Problems over a Hyper-rectangle

We extend our results from $\widehat{\mathcal{B}}^+(0,b)$ to $\mathcal{B}^+(0,b)$ for the worst- and best-case relative gap problems to evaluate integer programming models on non-discrete sets of right-hand sides. That is, we wish to compute $\Delta_{IRH} = \inf\{\gamma(\beta) \mid \beta \in \mathcal{B}^+(0,b)\}$ and $\Delta_{SRH} = \sup\{\gamma(\beta) \mid \beta \in \mathcal{B}^+(0,b)\}$. We show that when the set of right-hand sides is $\mathcal{B}^+(0,b)$, the infimum can be computed using only integral points, although it is generally not achieved, whereas the supremum is achieved at an integral point. We remark that we do not provide a framework to approximate the expectation of the relative gap function over a hyper-rectangle—the Lipschitzian argument from the absolute gap function case does not immediately translate to the relative gap function case—and we leave it as future research.

6.1 Supremum of the Relative Gap Function over a Hyper-rectangle

The supremum of the relative gap function over $\mathcal{B}(0,b)^+$ is always achieved at an integral right-hand side.

Proposition 17. *The supremum of the relative gap function over $\mathcal{B}(0,b)^+$ can be computed by only considering the integral points. Specifically,* $\sup_{\beta \in \mathcal{B}^+(0,b)} \gamma(\beta) = \max_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \gamma(\beta)$. *Further,* (10) *can be used to compute* $\Delta_{SRH} = \sup_{\beta \in \mathcal{B}^+(0,b)} \gamma(\beta)$.

6.2 Infimum of the Relative Gap Function over a Hyper-rectangle

Cover $\mathcal{B}^+(0,b)$ by L unit hyper-cubes such that each hyper-cube of the cover is a subset of $\mathcal{B}^+(0,b)$, and recall the notation from Section 4. Notice that

$$\inf\{z_{IP}(\beta)/z_{LP}(\beta) \mid \beta \in \mathcal{B}^+(0,b)\} = \min_{l \in \{1,\dots,L\}} \inf\{z_{IP}(\beta)/z_{LP}(\beta) \mid \beta \in \mathcal{B}^l\}.$$

Thus, we proceed by considering each unit hyper-cube, $\mathcal{B}^l, l \in \{1, \dots, L\}$ independently.

Proposition 18. *For any unit hyper-cube $\mathcal{B}^l$,* $\inf_{\beta}\{\gamma(\beta) \mid \beta \in \mathcal{B}^l\} = z_{IP}(b^l)/z_{LP}(d^l)$.

Corollary 12. $\Delta_{IRH} = \inf_{\beta \in \mathcal{B}^+(0,b)} \gamma(\beta) = \inf_{l \in \{1,\dots,L\}} z_{IP}(b^l)/z_{LP}(d^l)$.

Formulation (12) computes the infimum of the relative gap function over $\mathcal{B}^+(0, b)$.

$$
\begin{aligned}
\delta_{IRH} = \max \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} & \psi(\beta) \\
\text{s.t. } & (1 - \psi(\beta)) \cdot r(\beta)^T v^k \geq \phi(\beta), \text{ for all } \beta \in \widehat{\mathcal{B}}^+(0, b), k \in \mathcal{K}, && \text{(12a)} \\
& \text{SOS1}\left(\{\psi(\beta)\}_{\beta \in \widehat{\mathcal{B}}^+(0,b)}\right), && \text{(12b)} \\
& \psi \in \mathbb{R}^{|\widehat{\mathcal{B}}^+(0,b)|}, && \text{(12c)} \\
& \phi \in \Phi(b). && \text{(12d)}
\end{aligned}
$$

Proposition 19. $\delta_{IRH} = 1 - \Delta_{IRH}$.

7 Conclusion

This paper provides a linear-programming framework to evaluate integer programming models using linear programming relaxation gap functions. Through the expectation, supremum, and infimum of the absolute and relative gap functions, models can be evaluated on their performance as the problem data change, which is vital when the right-hand side is uncertain. Our formulations also reveal structure in the gap functions. For instance, our formulations show that one can compute metrics over hyper-rectangles using information at discrete sets; a brute-force sampling approach may not guarantee such results. In addition to creating a framework for evaluating integer programming models over many right-hand sides, we also present properties of the absolute and relative gap functions. Further, we introduce level-set-maximal vectors and demonstrate how they can be used to bound the gap functions.

There are various natural extensions of this work that can lead to future research. We only address integer programs; however, gap problems may be useful for mixed-integer programs and stochastic integer programs. Although we consider the expectation of the absolute gap function, we do not explore special properties of the probability distribution. One may also wish to consider different sets of right-hand sides, in addition to hyper-rectangles. We assume minimal structural knowledge; additional results may arise with assumptions about the constraint matrix and objective function.

References

Aliev, I., M. Henk, T. Oertel. 2017. Integrality gaps of integer knapsack problems. F. Eisenbrand, J. Koenemann, eds., *Integer Programming and Combinatorial Optimization*. Springer International Publishing, 25–38.

Beale, E. M. L., J. A. Tomlin. 1970. Special facilities in a general mathematical programming system for non-convex problems using ordered sets of variables. *Proceeding of the Fifth International Conference on Operations Research*, vol. 69. 447–454.

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

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

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

Blair, C. E., R. G. Jeroslow. 1985. Constructive characterizations of the value function of a mixed-integer program II. *Discrete Mathematics* **10**(3) 227–240.

Cook, W., A. M. H. Gerards, A. Schrijver, É. Tardos. 1986. Sensitivity theorems in integer programming. *Mathematical Programming* **34**(3) 251–264.

Eisenbrand, F., N. Hähnle, D. Pálvölgyi, G. Shmonin. 2013. Testing additive integrality gaps. *Mathematical Programming* **141**(1) 257–271.

Garey, M. R., D. S. Johnson. 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman.

Gomory, R. E. 1965. On the relation between integer and noninteger solutions to linear programs. *Proceedings of the National Academy of Science* **53**(2) 260–265.

Granot, F., J. Skorin-Kapov. 1990. Some proximity and sensitivity results in quadratic integer programming. *Mathematical Programming* **47**(1) 259–268.

Higle, J. L., S. W. Wallace. 2003. Sensitivity analysis and uncertainty in linear programming. *Interfaces* **33**(4) 53–60.

Hoşten, S., B. Sturmfels. 2007. Computing the integer programming gap. *Combinatorica* **27**(3) 367–382.

Jeroslow, R. G. 1979. Minimal inequalities. *Mathematical Programming* **17**(1) 1–15.

Johnson, E. L. 1979. On the group problem and a subadditive approach to integer programming. *Annals of Discrete Mathematics* **5** 97–112.

Johnson, E. L. 1980. Subadditive lifting methods for partitioning and knapsack problems. *Journal of Algorithms* **1**(1) 75 – 96.

Khachiyan, L. G. 1980. Polynomial algorithms in linear programming. *USSR Computational Mathematics and Mathematical Physics* **20**(1) 53 – 72.

Kong, N., A. J. Schaefer, B. Hunsaker. 2006. Two-stage integer programs with stochastic right-hand sides: a superadditive dual approach. *Mathematical Programming* **108**(2-3) 275–296.

Mangasarian, O. L., T. H. Shiau. 1987. Lipschitz continuity of solutions of linear inequalities, programs, and complementarity problems. *SIAM Journal on Control and Optimization* **25**(3) 583–595.

Meyer, R. R. 1974. On the existence of optimal solutions to integer and mixed-integer programming problems. *Mathematical Programming* **7**(1) 223–235.

Nemhauser, G. L., L. A. Wolsey. 1988. *Integer and Combinatorial Optimization*. Wiley-Interscience, John Wiley & Sons.

Paat, J., R. Weismantel, S. Weltge. 2018. Distances between optimal solutions of mixed-integer programs. *Mathematical Programming* doi:10.1007/s10107-018-1323-z.

Stein, O. 2016a. Error bounds for mixed integer linear optimization problems. *Mathematical Programming* **156**(1) 101–123.

Stein, O. 2016b. Error bounds for mixed integer nonlinear optimization problems. *Optimization Letters* **10**(6) 1153–1168.

Tind, J., L. A. Wolsey. 1981. An elementary survey of general duality theory in mathematical programming. *Mathematical Programming* **21**(1) 241–261.

Trapp, A. C., O. A. Prokopyev, A. J. Schaefer. 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.

Werman, M., D. Magagnosc. 1991. The relationship between integer and real solutions of constrained convex programming. *Mathematical Programming* **51**(1) 133–135.

Williams, H. P. 1996. Constructing the value function for an integer linear programme over a cone. *Computational Optimization and Applications* **6**(1) 15–26.

Williams, H. P. 2013. *Model Building in Mathematical Programming*. 5th ed. John Wiley & Sons.

Wolsey, L. A. 1981. Integer programming duality: Price functions and sensitivity analysis. *Mathematical Programming* **20**(1) 173–195.

A Section 2 Results

Corollary 2. *For any β such that SDP(β) has an optimal solution, there exists an optimal solution ϕ^* of SDP(β) such that $\phi^*(\tilde{\beta}) = z_{IP}(\tilde{\beta})$ for all $\tilde{\beta}$ satisfying $S(\tilde{\beta}) \neq \emptyset$.*

Proof: Let ϕ^* be an extension of z_{IP} satisfying the claim of Proposition 1. Because the value function of a mixed-integer program is superadditive and nondecreasing (Jeroslow 1979), ϕ^* is superadditive and nondecreasing. The function ϕ^* is an extension of z_{IP}, so $\phi^*(a^j) \geq c_j$ for all $j \in \{1, ..., n\}$, which proves feasibility. Optimality is immediate. $\square$

Proposition 2. *Suppose $A \in \mathbb{Z}_+^{m\times n}$ and $b \in \mathbb{Z}_+^m$. The truncated value function $z_{IP} \in \mathbb{R}^{|\widehat{\mathcal{B}}(0,b)|}$ is in the polyhedron $\Phi(b)$.*

One can prove Proposition 2 directly by verifying z_{IP} satisfies all of the constraints that define $\Phi(b)$.

Theorem 4. *Suppose $A \in \mathbb{Z}_+^{m\times n}$ and $b \in \mathbb{Z}_+^m$. For each $\beta \in \widehat{\mathcal{B}}(0,b)$, SDP2($\beta$) is a strong dual to IP(β).*

Proof: We embed the feasible region $\Phi(\beta)$ in $\mathbb{R}^{|\widehat{\mathcal{B}}(0,b)|}$. Let $\Phi_{emb}(\beta) = \{\phi \in \mathbb{R}^{|\widehat{\mathcal{B}}(0,b)|} \mid \phi$ satisfies (2a)–(2d) with respect to $\beta\}$. Observe that $\Phi(b) \subseteq \Phi_{emb}(\beta)$, for all $\beta \in \widehat{\mathcal{B}}(0,b)$, and for any $\phi \in \Phi(\beta)$, there exists $\phi_{emb} \in \Phi_{emb}(\beta)$ such that $\phi_{emb}(\beta') = \phi(\beta')$, for all $\beta' \in \widehat{\mathcal{B}}(0,\beta)$. Furthermore, because A and β are nonnegative, IP(β) is feasible, and by **A2**, IP(β) is bounded.

Define $\phi^* \in \mathbb{R}^{|\widehat{\mathcal{B}}(0,b)|}$ such that $\phi^*(\beta') = z_{IP}(\beta')$, for all $\beta' \in \widehat{\mathcal{B}}(0,b)$. Then $\phi^* \in \Phi(b)$, which implies that $\phi^* \in \Phi_{emb}(\beta)$. Define the truncated vector $\tilde{\phi} \in \mathbb{R}^{|\widehat{\mathcal{B}}(0,\beta)|}$ such that $\tilde{\phi}(\beta') = \phi^*(\beta')$, for all $\beta' \in \widehat{\mathcal{B}}(0,\beta)$. Then $\tilde{\phi}(\beta) = \phi^*(\beta) = z_{IP}(\beta)$ and $\tilde{\phi}$ satisfies (2a)-(2d) with respect to β, which implies that $\tilde{\phi} \in \Phi(\beta)$. Hence, $\tilde{\phi}$ is an optimal solution for SDP2(β). $\square$

Corollary 3. *Suppose $A \in \mathbb{Z}_+^n$ and $b \in \mathbb{Z}_+^m$. For any $\phi \in \Phi(b), \phi(\beta) \geq z_{IP}(\beta)$, for all $\beta \in \widehat{\mathcal{B}}(0,b)$.*

The proof of Corollary 3 follows from Theorem 4 and is omitted.

Proposition 3.

(I) For any $\beta \in \mathcal{B}(\tilde{b}, b)$, there exists a $\tilde{\beta} \in D^+$ with $\tilde{\beta} \geq \beta$ such that $z_{IP}(\tilde{\beta}) = z_{IP}(\beta)$.

(II) For any $\beta \in \mathcal{B}(\tilde{b}, b)$, there exists a $\tilde{\beta} \in D^-$ with $\tilde{\beta} \le \beta$ such that $z_{IP}(\tilde{\beta}) = z_{IP}(\beta)$.

Proof: We prove only (I) as (II) is analogous. If $\beta \in D^+$, the proof is trivial. Suppose $\beta \in \mathcal{B}(\tilde{b}, b)$ is not level-set maximal. Let $s^0 = 0$ and $\beta^0 = \beta$. Perform the following improvement procedure: for each $i \in \{1, ..., m\}$, let $s^i = \max\{s \in \mathbb{Z}_+ \mid \beta^{i-1} + s \cdot e_i \in \mathcal{B}(\tilde{b}, b), z_{IP}(\beta^{i-1} + s \cdot e_i) = z_{IP}(\beta)\}$ and $\beta^i = \beta^{i-1} + s^i \cdot e_i$. At the end of each improvement procedure (which consists of a finite number of iterations), $z_{IP}(\beta^m) = z_{IP}(\beta^0) = z_{IP}(\beta)$, and either $\beta^m = \beta^0$ or $\beta^m \neq \beta^0$.

Suppose $\beta^m \neq \beta^0$. Then $\beta^m \ge \beta^0$ and $\beta_i^m \ge \beta_i^0 + 1$ for some $i \in \{1, ..., m\}$. Because $\mathcal{B}(\tilde{b}, b)$ is a finite set, this case can only occur finitely often. Suppose $\beta^m = \beta^0$. By definition, β^m is level-set maximal because there does not exist $i \in \{1, ..., m\}$ such that $z_{IP}(\beta^m + e_i) = z_{IP}(\beta^m)$. Hence, we conclude that after finitely many applications of the improvement procedure, $\beta^m = \beta^0$ and β^m is level-set maximal with $z_{IP}(\beta) = z_{IP}(\beta^m)$ and $\beta^m \ge \beta$. □

THEOREM 5. *Level-set-minimal and level-set-maximal vectors are sufficient to bound the gap function. That is,*

(I) $\inf_{\beta \in \mathcal{B}(\tilde{b}, b)} \Gamma(\beta) = \inf_{\beta \in D^-} \Gamma(\beta).$

(II) $\sup_{\beta \in \mathcal{B}(\tilde{b}, b)} \Gamma(\beta) = \sup_{\beta \in D^+} \Gamma(\beta).$

(III) $\inf_{\beta \in \mathcal{B}^+(\tilde{b}, b)} \gamma(\beta) = \inf_{\beta \in D^+ \cap \mathcal{B}^+(\tilde{b}, b)} \gamma(\beta).$

(IV) $\sup_{\beta \in \mathcal{B}^+(\tilde{b}, b)} \gamma(\beta) = \sup_{\beta \in D^- \cap \mathcal{B}^+(\tilde{b}, b)} \gamma(\beta).$

Proof: We prove only (I) as the other proofs are similar. Suppose $\tilde{\beta} \in \mathcal{B}(\tilde{b}, b) \backslash D^-$. By Proposition 3, there exists $\bar{\beta} \in D^-$ such that $z_{IP}(\bar{\beta}) = z_{IP}(\tilde{\beta})$ and $\bar{\beta} \le \tilde{\beta}$. By the monotonicity of z_{LP}, $\Gamma(\tilde{\beta}) \ge \Gamma(\bar{\beta}) \ge \inf_{\beta \in D^-} \Gamma(\beta)$. Hence, $\inf_{\beta \in \mathcal{B}(\tilde{b}, b)} \Gamma(\beta) \ge \inf_{\beta \in D^-} \Gamma(\beta)$. Finally, we have that $D^- \subseteq \mathcal{B}(\tilde{b}, b)$, which implies $\inf_{\beta \in D^-} \Gamma(\beta) \ge \inf_{\beta \in \mathcal{B}(\tilde{b}, b)} \Gamma(\beta)$ and proves the equality. □

THEOREM 6. *The absolute gap function defined over rational vectors, $\Gamma : \mathcal{B}(\tilde{b}, b) \cap \mathbb{Q}^m \to \mathbb{R}$, is the minimum of finitely many Gomory functions.*

Proof: It is known that $-z_{IP}(\beta) = \min\{-c^T x \mid Ax \le \beta, x \in \mathbb{Z}^n_+\}$ is a Gomory function when defined over rational vectors (Blair and Jeroslow 1982). Also, $z_{LP}(\beta) = \min\{(v^k)^T\beta \mid k \in \mathcal{K}\}$. For each $k \in \mathcal{K}$, let $L_k : \mathcal{B}(\tilde{b}, b) \cap \mathbb{Q}^m \to \mathbb{R}$, where $L_k(\beta) = (v^k)^T\beta - z_{IP}(\beta)$. Notice that for each $k \in \mathcal{K}, L_k$ is a Gomory function, as it is the sum of two Gomory functions. Further, $\Gamma(\beta) = \min\{(v^k)^T\beta \mid k \in \mathcal{K}\} - z_{IP}(\beta) = \min\{L_k(\beta) \mid k \in \mathcal{K}\}$, for all $\beta \in \mathcal{B}(\tilde{b}, b) \cap \mathbb{Q}^m$. □

PROPOSITION 4. *For any $\beta \in \mathcal{B}(\tilde{b}, b)$, if $\hat{x} \in opt_{LP}(\beta)$, then for all x such that $0 \le x \le \hat{x}$, $x \in opt_{LP}(\tilde{\beta})$, for all $\tilde{\beta} \in \Lambda(\beta, x, \hat{x})$.*

Proof: From **A2**, $|z_{LP}(\beta')| < \infty$, for all $\beta' \in \mathcal{B}(\tilde{\beta}, \beta)$. Hence, $\text{opt}_{LP}(\beta')$ and $\text{opt}_{DLP}(\beta')$ are nonempty, for all $\beta' \in \mathcal{B}(\tilde{\beta}, \beta)$. Let $v \in \text{opt}_{DLP}(\beta)$ and $\hat{x} \in \text{opt}_{LP}(\beta)$. Consider $x \in \mathbb{R}^n$ such that $0 \le x \le \hat{x}$. For any $\beta' \in \mathbb{R}^m$, let $M^<_{\beta'} = \{i \in \{1, ..., m\} \mid a_i^T\hat{x} < \beta'_i\}$. By linear programming complementary slackness:

(I) $v_i = 0$ for all $i \in M^<_\beta$.

(II) $\hat{x}_j > 0 \Rightarrow (a^j)^T v = c_j$.

We show that v also satisfies these two conditions when x replaces $\hat{x}$ and $\beta - A(\hat{x} - x)$ replaces β in (I) and (II). For condition (II), notice $\{j : x_j > 0\} \subseteq \{j : \hat{x}_j > 0\}$, since $x_j \le \hat{x}_j$. This implies $(a^j)^T v = c_j$, for all j such that $x_j > 0$, which proves condition (II). Furthermore, $M^<_{\beta - A(\hat{x}-x)} = M^<_\beta$, since $a_i^T\hat{x} < \beta_i$ if and only if $a_i^T x < \beta_i - a_i^T(\hat{x} - x)$. This implies condition (I) holds, so $v \in \text{opt}_{DLP}(\beta - A(\hat{x} - x))$. By assumption, $v \in \text{opt}_{DLP}(\beta)$, thus, $\beta^T v = c^T\hat{x}$. Hence, complementary slackness implies that $0 = (\beta_i - a_i^T\hat{x})v_i = (\beta_i - a_i^T(\hat{x}-x) - a_i^T x)v_i$. The vector x is feasible for $\text{LP}(\beta - A(\hat{x}-x))$. To see this, first note that $x \in \mathbb{R}^n_+$. Also $\beta - A\hat{x} \ge 0$, which implies $Ax \le Ax + \beta - A\hat{x} = \beta - A(\hat{x} - x)$. By complementary slackness (primal solution x, dual solution v), $x \in \text{opt}_{LP}(\beta - A(\hat{x} - x))$. By monotonicity, $x \in \text{opt}_{LP}(\tilde{\beta})$, for any $\tilde{\beta} \in \Lambda(\beta, x, \hat{x})$. □

COROLLARY 4. *Let $\beta, x, \hat{x}$ satisfy the conditions in Proposition 4. Suppose further that $x \in \mathbb{Z}^n_+$. Then for any $\tilde{\beta} \in \Lambda(\beta, x, \hat{x})$, $\Gamma(\tilde{\beta}) = 0$.*

Proof: By Proposition 4, $x \in \text{opt}_{LP}(\tilde{\beta})$. Now, $x \in \mathbb{Z}^n_+$, thus $x \in S(\tilde{\beta})$ and $c^T x \le z_{IP}(\tilde{\beta})$.

Therefore, $c^T x = z_{LP}(\tilde{\beta}) \geq z_{IP}(\tilde{\beta}) \geq c^T x$. Thus, $z_{LP}(\tilde{\beta}) = z_{IP}(\tilde{\beta})$, and $\Gamma(\tilde{\beta}) = 0$. $\square$

Lemma 1. (Nemhauser and Wolsey 1988) *Suppose $\hat{x} \in opt_{IP}(\beta)$, for some $\beta \in \mathcal{B}(\tilde{b}, b)$ and ϕ^* is an optimal solution to SDP(β). Then $\phi^*(Ax) = c^T x$ and $\phi^*(Ax) + \phi^*(\beta - Ax) = \phi^*(\beta)$, for all $x \in \mathbb{Z}^n_+$ such that $x \leq \hat{x}$.*

Proposition 5. *Let $\beta \in \mathcal{B}(\tilde{b}, b)$. For all $j \in \{1, ..., n\}$ such that there exists $x^I \in opt_{IP}(\beta)$ and $x^L \in opt_{LP}(\beta)$ with $x^I_j \geq 1$ and $x^L_j \geq 1$, we have $z_{LP}(\beta - a^j) = z_{LP}(\beta) - c_j$, and $z_{IP}(\beta - a^j) = z_{IP}(\beta) - c_j$.*

Proof: Let e_j be the j^{th} standard basis vector in $\mathbb{R}^n$, which implies $e_j \in \mathbb{Z}^n_+$. By hypothesis, $e_j \leq x^I$ and $e_j \leq x^L$. Recall that z_{IP} can be extended to an optimal solution ϕ^* of SDP(β) (Corollary 2) so that $\phi^*(\beta) = z_{IP}(\beta)$, for all β such that $S(\beta) \neq \emptyset$. By Lemma 1, $z_{IP}(\beta - Ae_j) = z_{IP}(\beta - a^j) = \phi^*(\beta - a^j) = \phi^*(\beta) - \phi^*(a^j) = z_{IP}(\beta) - z_{IP}(a^j)$.

Observe that $\beta - A(x^L - e_j) = a^j + \beta - Ax^L \geq a^j$. Thus, $a^j \in \Lambda(\beta, e_j, x^L)$. By Proposition 4, $e_j \in \text{opt}_{LP}(a^j)$. This implies that $z_{LP}(a^j) = c_j$. Because $e_j \in S(a^j), z_{IP}(a^j) = c_j$. Therefore, $z_{IP}(\beta - a^j) = z_{IP}(\beta) - c_j$.

By the superadditivity of z_{LP}, $z_{LP}(\beta - a^j) \leq z_{LP}(\beta) - z_{LP}(a^j) = z_{LP}(\beta) - c_j$. Observe that $x^L - e_j \in P(\beta - a^j)$. Therefore, $z_{LP}(\beta - a^j) \geq c^T(x^L - e_j) = z_{LP}(\beta) - c_j$. This implies $z_{LP}(\beta - a^j) = z_{LP}(\beta) - c_j$. $\square$

Theorem 7. *Under the conditions of Proposition 5, $\Gamma(\beta - a^j) = \Gamma(\beta)$. If, in addition, $z_{LP}(\beta) > c_j$ and $z_{IP}(\beta) \geq c_j$, then $\gamma(\beta - a^j) = (z_{IP}(\beta) - c_j)/(z_{LP}(\beta) - c_j)$.*

We omit the proof of Theorem 7 as it follows from Proposition 5.

Theorem 8. *Let $\beta \in \widehat{\mathcal{B}}(0, b)$ and suppose $A \in \mathbb{Z}^{m \times n}_+$ and $b \in \mathbb{Z}^m_+$. Suppose there exists an $x \in opt_{IP}(\beta)$ such that $x_j \geq 1$ for some $j \in \{1, ..., n\}$. Then for any optimal solution ϕ^* to SDP2(β), $\phi^*(\beta - a^j) = z_{IP}(\beta - a^j)$.*

Proof: By an argument similar to that shown in the proof of Proposition 5, $z_{IP}(\beta - a^j) = z_{IP}(\beta) - z_{IP}(a^j)$. Because $(\beta - a^j) + a^j = \beta \in \widehat{\mathcal{B}}(0, b)$ and ϕ^* is an optimal solution to

SDP2(β),

$$\begin{aligned}\phi^*(\beta - a^j) + \phi^*(a^j) &\leq \phi^*(\beta) \\ &= z_{IP}(\beta) \\ \iff \phi^*(\beta - a^j) &\leq z_{IP}(\beta) - \phi^*(a^j).\end{aligned}$$

The first inequality comes from the superadditive constraints in SDP2(β), and the equality is due to SDP2(β) being a strong dual (Theorem 4). By Corollary 3, $\phi^*(a^j) \geq z_{IP}(a^j)$, which implies that

$$\begin{aligned}\phi^*(\beta - a^j) &\leq z_{IP}(\beta) - z_{IP}(a^j) \\ &= z_{IP}(\beta - a^j).\end{aligned}$$

By Corollary 3, $\phi^*(\beta - a^j) \geq z_{IP}(\beta - a^j)$, which proves the desired equality. $\square$

B Section 3 Results

PROPOSITION 6. *The optimal objective value of* (3) *is* $\mathbb{E}_{\xi^{(1)}}[\Gamma(\xi^{(1)})]$.

Proof: Let $(\tilde{\phi}, \tilde{\psi}) = (z_{IP}, \Gamma)$, the truncated value function and the truncated gap function. Observe that $\tilde{\phi} \in \Phi(b)$ (Proposition 2). Consider $\beta \in \widehat{\mathcal{B}}(0, b)$. Then $\tilde{\psi}(\beta) = \Gamma(\beta) = z_{LP}(\beta) - z_{IP}(\beta) \leq \beta^T v^k - \tilde{\phi}(\beta)$, for all $k \in \mathcal{K}$. Hence the solution is feasible.

Let $(\bar{\phi}, \bar{\psi})$ be a feasible solution for (3). Because $\bar{\phi}$ satisfies (3c), by Corollary 3, $\bar{\phi}(\beta) \geq z_{IP}(\beta)$, for all $\beta \in \widehat{\mathcal{B}}(0, b)$. By (3a), for each $\beta \in \widehat{\mathcal{B}}(0, b)$, $\bar{\psi}(\beta) \leq \beta^T v^k - \bar{\phi}(\beta) \leq \beta^T v^k - z_{IP}(\beta)$, for all $k \in \mathcal{K}$. This implies that $\bar{\psi}(\beta) \leq z_{LP}(\beta) - z_{IP}(\beta) = \tilde{\psi}(\beta)$ and $\sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(1)}(\beta)\bar{\psi}(\beta) \leq \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(1)}(\beta)\tilde{\psi}(\beta)$.

Thus, $(\tilde{\phi}, \tilde{\psi})$ is an optimal solution to (3), and the optimal objective value is $\mathbb{E}_{\xi^{(1)}}[\Gamma(\xi^{(1)})]$.

$\square$

COROLLARY 5. *Suppose $\mu^{(1)}(\beta) > 0$ for all $\beta \in \widehat{\mathcal{B}}(0,b)$. Then, any feasible solution $(\bar{\phi}, \bar{\psi})$ in which $\bar{\phi} \neq z_{IP}$ is not optimal for* (3).

One can prove Corollary 5 by noticing that every component of an optimal solution ϕ^* of (3) must be as small as possible. We omit a formal proof.

COROLLARY 6. *Let $(\tilde{\phi}, \tilde{\psi})$ be an optimal solution to* (3) *and δ_{EAD} the optimal objective value. Then the variance of the absolute gap over $\widehat{\mathcal{B}}(0,b)$ can be calculated as follows:*

$$\sigma_\Gamma^2 = \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(1)}(\beta)(\tilde{\psi}(\beta))^2 - \delta_{EAD}^2.$$

Proof: For any $\beta \in \widehat{\mathcal{B}}(0,b)$ such that $\mu^{(1)}(\beta) > 0, \tilde{\phi}(\beta) = z_{IP}(\beta)$; otherwise, $(\tilde{\phi}, \tilde{\psi})$ is not an optimal solution. It follows that for all such β, $\tilde{\psi}(\beta) = \Gamma(\beta)$. Hence, $\sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(1)}(\beta)(\tilde{\psi}(\beta))^2 = \sum_{\beta : \mu^{(1)}(\beta) > 0} \mu^{(1)}(\beta)(\tilde{\psi}(\beta))^2 = \sum_{\beta : \mu^{(1)}(\beta) > 0} \mu^{(1)}(\beta)(\Gamma(\beta))^2 = \mathbb{E}_{\xi^{(1)}}\left[(\Gamma(\xi^{(1)}))^2\right]$. Therefore, $\sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(1)}(\beta)(\tilde{\psi}(\beta))^2 - \delta_{EAD}^2 = \mathbb{E}_{\xi^{(1)}}\left[(\Gamma(\xi^{(1)}))^2\right] - \mathbb{E}_{\xi^{(1)}}\left[\Gamma(\xi^{(1)})\right]^2 = \sigma_\Gamma^2$. □

PROPOSITION 7. *The optimal objective value of* (4) *is Δ_{SAD}.*

Proof: Let $\tilde{\beta} \in \arg\max\{\Gamma(\beta) \mid \beta \in \widehat{\mathcal{B}}(0,b)\}$. Let $\tilde{\phi} = z_{IP}$, $\tilde{\psi}(\tilde{\beta}) = \Gamma(\tilde{\beta})$, and $\tilde{\psi}(\beta) = 0$ for all $\beta \in \widehat{\mathcal{B}}(0,b) \backslash \tilde{\beta}$. It is immediate that $\tilde{\phi} \in \Phi(b)$ (Proposition 2). By construction, $\tilde{\psi}$ satisfies (4b). Now, $\tilde{\psi}(\tilde{\beta}) = \Gamma(\tilde{\beta})$, and $\tilde{\psi}(\beta) = 0 \leq \Gamma(\beta)$, for all $\beta \in \widehat{\mathcal{B}}(0,b) \backslash \tilde{\beta}$. Thus, $\tilde{\psi}(\beta) \leq \Gamma(\beta) = z_{LP}(\beta) - z_{IP}(\beta) \leq \beta^T v^k - \tilde{\phi}(\beta)$, for all $k \in \mathcal{K}$.

Let $(\bar{\phi}, \bar{\psi})$ be feasible for (4), where $\bar{\psi}(\beta) = 0$ for all $\beta \in \widehat{\mathcal{B}}(0,b) \backslash \bar{\beta}$, for some $\bar{\beta} \in \widehat{\mathcal{B}}(0,b)$. Because $\bar{\phi}$ satisfies (4d), by Corollary 3, $\bar{\phi}(\beta) \geq z_{IP}(\beta)$ for all $\beta \in \widehat{\mathcal{B}}(0,b)$. By (4a), $\bar{\psi}(\bar{\beta}) \leq z_{LP}(\bar{\beta}) - \bar{\phi}(\bar{\beta}) \leq z_{LP}(\bar{\beta}) - z_{IP}(\bar{\beta}) = \Gamma(\bar{\beta})$. Because $\tilde{\beta} \in \arg\max\{\Gamma(\beta) \mid \beta \in \widehat{\mathcal{B}}(0,b)\}, \bar{\psi}(\bar{\beta}) = \Gamma(\bar{\beta}) \leq \Gamma(\tilde{\beta}) = \tilde{\psi}(\tilde{\beta})$, which implies $\sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \bar{\psi}(\beta) \leq \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \tilde{\psi}(\beta) = \max\{\Gamma(\beta) \mid \beta \in \widehat{\mathcal{B}}(0,b)\} = \Delta_{SAD}$. □

PROPOSITION 8. *The optimal objective value of* (5) *is Δ_{IAD}.*

Proof: For each $\beta \in \widehat{\mathcal{B}}(0,b)$, let $\tilde{\phi}(\beta) = z_{IP}(\beta)$. Let $\tilde{\psi} = \Delta_{IAD}$. It is immediate that $\tilde{\phi} \in \Phi(b)$ (Proposition 2). Further, $\tilde{\psi} = \Delta_{IAD} \leq \Gamma(\beta) \leq \beta^T v^k - z_{IP}(\beta) = \beta^T v^k - \tilde{\phi}(\beta)$, for all $\beta \in \widehat{\mathcal{B}}(0,b) \backslash \{0\}$, $k \in \mathcal{K}$. Thus, $(\tilde{\phi}, \tilde{\psi})$ is feasible for (5).

Let $(\bar{\phi}, \bar{\psi})$ be feasible for (5). Recall that for all $\beta \in \widehat{\mathcal{B}}(0,b)$, by Corollary 3, $z_{IP}(\beta) \leq \bar{\phi}(\beta)$. By (5a), for all $\beta \in \widehat{\mathcal{B}}(0,b)\backslash\{0\}$ and $k \in \mathcal{K}$, $\bar{\psi} \leq \beta^T v^k - \bar{\phi}(\beta)$, which implies $\bar{\psi} \leq z_{LP}(\beta) - z_{IP}(\beta) = \Gamma(\beta)$. Hence, $\bar{\psi} \leq \Delta_{IAD} = \tilde{\psi}$. This proves that $(\tilde{\phi}, \tilde{\psi})$ is optimal for (5). □

C Section 4 Results

PROPOSITION 9. *The optimal objective value of* (6) *is* $\delta_{EAH} = \mathbb{E}_{\xi^{(p)}}[\Gamma(\xi^{(p)})]$.

Proof: The proof is similar to that of Proposition 6; however, we present a proof to illustrate the new details introduced by the use of $\widehat{\mathcal{B}}_p(0,b)$. Let $(\tilde{\phi}, \tilde{\psi}) = (z_{IP}, \Gamma)$, where $\tilde{\psi} \in \mathbb{R}^{|\widehat{\mathcal{B}}_p(0,b)|}$. Observe that $\tilde{\phi} \in \Phi(b)$ (Proposition 2). Consider $\beta \in \widehat{\mathcal{B}}_p(0,b)$. Then $\tilde{\psi}(\beta) = \Gamma(\beta) = z_{LP}(\beta) - z_{IP}(\beta) = z_{LP}(\beta) - z_{IP}(\lfloor\beta\rfloor) \leq \beta^T v^k - \tilde{\phi}(\lfloor\beta\rfloor)$, for all $k \in \mathcal{K}$. Hence the solution is feasible.

Let $(\bar{\phi}, \bar{\psi})$ be a feasible solution for (6). Because $\bar{\phi}$ satisfies (6c), by Corollary 3, $\bar{\phi}(\beta) \geq z_{IP}(\beta)$, for all $\beta \in \widehat{\mathcal{B}}(0,b)$. By (6a), for each $\beta \in \widehat{\mathcal{B}}_p(0,b)$, $\bar{\psi}(\beta) \leq \beta^T v^k - \bar{\phi}(\lfloor\beta\rfloor) \leq \beta^T v^k - z_{IP}(\beta)$, for all $k \in \mathcal{K}$. This implies that $\bar{\psi}(\beta) \leq z_{LP}(\beta) - z_{IP}(\beta) = \tilde{\psi}(\beta)$ and $\sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(p)}(\beta)\bar{\psi}(\beta) \leq \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \mu^{(p)}(\beta)\tilde{\psi}(\beta)$.

Thus, $(\tilde{\phi}, \tilde{\psi})$ is an optimal solution to (6), and the optimal objective value is $\mathbb{E}_{\xi^{(p)}}[\Gamma(\xi^{(p)})]$. □

COROLLARY 7. *Suppose* $\mu^{(p)}(\beta) > 0$ *for all* $\beta \in \widehat{\mathcal{B}}_p(0,b)$. *Then, any feasible solution* $(\bar{\phi}, \bar{\psi})$ *in which* $\bar{\phi} \neq z_{IP}$ *is not optimal for* (6).

We omit the proof of Corollary 7 due to reasoning similar to that of Corollary 5.

COROLLARY 8. *Let* $(\tilde{\phi}, \tilde{\psi})$ *be an optimal solution to* (6) *and* δ_{EAH} *the optimal objective value. Then the variance of the absolute gap over* $\widehat{\mathcal{B}}_p(0,b)$ *can be calculated as follows:*

$$\sigma_\Gamma^2 = \sum_{\beta \in \widehat{\mathcal{B}}_p(0,b)} \mu^{(p)}(\beta)(\tilde{\psi}(\beta))^2 - \delta_{EAH}^2.$$

The proof of Corollary 8 is similar to that of Corollary 6 and is omitted.

Lemma 2. *For any $\beta \in \mathcal{B}(0,b)$:*

(I) $||\beta - \lfloor\beta\rfloor^{(p)}||_\infty \leq 1/p$.

(II) $z_{IP}(\lfloor\beta\rfloor^{(p)}) = z_{IP}(\beta)$.

(III) $\Gamma(\beta) \geq \Gamma(\lfloor\beta\rfloor^{(p)})$.

Proof: Let $\beta \in \mathcal{B}(0,b)$. For any $i \in \{1,...,m\}, |\beta_i - (\lfloor\beta\rfloor^{(p)})_i| = \beta_i - (\lfloor\beta\rfloor^{(p)})_i = \frac{\alpha_i - \lfloor\alpha_i\rfloor}{p} \leq 1/p$, where $\beta = \alpha/p$. This proves (I). Also, $\lfloor\beta\rfloor = \lfloor(\lfloor\beta\rfloor^{(p)})\rfloor$, and (II) follows. The monotonicity of z_{LP} and (II) imply (III). $\square$

Proposition 10. *(Mangasarian and Shiau 1987) The linear programming value function z_{LP} is Lipschitz continuous over $\mathcal{B}(0,b)$.*

Proof: Mangasarian and Shiau (1987) show that over the set of right-hand sides for which the linear program is feasible, the optimal solutions follow a Lipschitzian relationship. That is, for any β^1, β^2 such that $z_{LP}(\beta^1)$ and $z_{LP}(\beta^2)$ are finite, there exist optimal solutions $x^1 \in \text{opt}_{LP}(\beta^1), x^2 \in \text{opt}_{LP}(\beta^2)$ such that $||x^1 - x^2||_S \leq M'||\beta^1 - \beta^2||_R$, where $||\cdot||_S$ is a norm on $\mathbb{R}^n$ and $||\cdot||_R$ is a norm on $\mathbb{R}^m$. LP(β) is feasible and bounded for all $\beta \in \mathcal{B}(0,b)$, and the objective function is linear. Hence z_{LP} is Lipschitz continuous over $\mathcal{B}(0,b)$. $\square$

Theorem 9. *Suppose that for all $\beta \in \widehat{\mathcal{B}}_p(0,b), \mu^{(p)}(\beta) > 0$. Let $(\tilde{\phi}, \tilde{\psi})$ be an optimal solution to* (6)*, and define the extension $\tilde{\psi}_p : \mathcal{B}(0,b) \to \mathbb{R}$ by $\tilde{\psi}_p(\beta) = \tilde{\psi}_p(\lfloor\beta\rfloor^{(p)})$. Then the following results hold:*

$$||\Gamma - \tilde{\psi}_p||_\infty \leq M/p, \tag{13a}$$

$$\mathbb{E}_\xi|\Gamma(\xi) - \tilde{\psi}_p(\xi)| = \mathbb{E}_\xi(\Gamma(\xi) - \tilde{\psi}_p(\xi)) \leq M/p. \tag{13b}$$

Proof: Let $\beta \in \mathcal{B}(0,b)$. By Lemma 2 ((II) and (III)) and Corollary 7, $|\Gamma(\beta) - \tilde{\psi}_p(\beta)| = |\Gamma(\beta) - \tilde{\psi}(\lfloor\beta\rfloor^{(p)})| = |\Gamma(\beta) - \Gamma(\lfloor\beta\rfloor^{(p)})| = \Gamma(\beta) - \Gamma(\lfloor\beta\rfloor^{(p)}) = z_{LP}(\beta) - z_{LP}(\lfloor\beta\rfloor^{(p)})$. By the definition of the Lipschitz constant M and Lemma 2 (I), $z_{LP}(\beta) - z_{LP}(\lfloor\beta\rfloor^{(p)}) \leq M||\beta - \lfloor\beta\rfloor^{(p)}||_\infty \leq M/p$. This proves (13a). (13b) follows from

$$\begin{aligned}
\mathbb{E}_\xi|\Gamma(\xi) - \tilde{\psi}_p(\xi)| &= \int_{\beta\in\mathcal{B}(0,b)} |\Gamma(\beta) - \tilde{\psi}_p(\beta)|\mu(\beta)d\beta \\
&= \int_{\beta\in\mathcal{B}(0,b)} (\Gamma(\beta) - \tilde{\psi}_p(\beta))\mu(\beta)d\beta \\
&\leq M/p \int_{\beta\in\mathcal{B}(0,b)} \mu(\beta)d\beta \\
&= M/p. \qquad \square
\end{aligned}$$

THEOREM 10. *Suppose the probability mass function $\mu^{(p)}$ approximates the probability density function μ as follows: $\mu^{(p)}(b^t) = \int_{\mathcal{B}^t} \mu(\beta)d\beta$, for all $t \in \{1, ..., T^{(p)}\}$. Then, the difference between the expectation of the absolute gap function over $\mathcal{B}(0,b)$ and the optimal objective value of (6) is bounded. Specifically, $|\mathbb{E}_\xi[\Gamma(\xi)] - \delta_{EAH}| \leq M/p$.*

Proof: Observe:

$$\begin{aligned}
|\mathbb{E}_\xi[\Gamma(\xi)] - \delta_{EAH}| &= \left| \int_{\mathcal{B}(0,b)} \mu(\beta)\Gamma(\beta)d\beta - \sum_{\beta\in\widehat{\mathcal{B}}_p} \mu^{(p)}(\beta)\Gamma(\beta) \right| \\
&= \left| \int_{\mathcal{B}(0,b)} \mu(\beta)\Gamma(\beta)d\beta - \sum_{t\in T^{(p)}} \mu^{(p)}(\beta)\Gamma(\beta) \right| \\
&= \left| \sum_{t\in T^{(p)}} \int_{\mathcal{B}^t} \mu(\beta)\Gamma(\beta)d\beta - \sum_{t\in T^{(p)}} \int_{\mathcal{B}^t} d\beta\Gamma(b^t) \right| \\
&= \left| \sum_{t\in T^{(p)}} \int_{\mathcal{B}^t} \mu(\beta)(\Gamma(\beta) - \Gamma(b^t))d\beta \right| \\
&\leq \sum_{t\in T^{(p)}} \int_{\mathcal{B}^t} \left|\mu(\beta)(\Gamma(\beta) - \Gamma(b^t))\right| d\beta \\
&\leq \sum_{t\in T^{(p)}} \int_{\mathcal{B}^t} \mu(\beta)M/pd\beta
\end{aligned}$$

$$= M/p \int\limits_{\mathcal{B}(0,b)} \mu(\beta) d\beta$$

$$= M/p. \qquad \square$$

Lemma 3. *Let $l \in \{1, ..., L\}$. For any $\epsilon > 0$, there exists $\beta \in (\mathcal{B}^l)^\circ$ such that $z_{LP}(d^l) - z_{LP}(\beta) < \epsilon$, where $(\mathcal{B}^l)^\circ$ denotes the interior of $\mathcal{B}^l$.*

Proof: Fix $\epsilon > 0$. Notice that if $z_{LP}(d^l) = 0$, then $z_{LP}(\beta) = 0$ for all $\beta \in (\mathcal{B}^l)^\circ$. Now assume $z_{LP}(d^l) \neq 0$, and let $\delta = \frac{1}{2} \min \left\{ \frac{\epsilon}{z_{LP}(d^l)}, \frac{1}{\max\limits_{i \in \{1,...,m\}} d_i^l} \right\}$. Define $\beta = (1-\delta)d^l$. Thus, $\beta \in (\mathcal{B}^l)^\circ$. Choose $x^d \in \mathrm{opt}_{LP}(d^l)$ and let $x^\beta = (1-\delta)x^d$. Then, $Ax^\beta = (1-\delta)Ax^d \leq (1-\delta)d^l = \beta$. Thus, $x^\beta \in P(\beta)$. It follows that $z_{LP}(\beta) \geq c^T x^\beta = (1-\delta)c^T x^d = z_{LP}(d^l) - \delta z_{LP}(d^l) \geq z_{LP}(d^l) - \frac{\epsilon}{2 z_{LP}(d^l)} z_{LP}(d^l) = z_{LP}(d^l) - \frac{\epsilon}{2} > z_{LP}(d^l) - \epsilon$. Further, $z_{LP}(d^l) \geq z_{LP}(\beta)$ because z_{LP} is increasing. $\square$

Proposition 11. *For any unit hyper-cube $\mathcal{B}^l$, $\sup\limits_{\beta \in \mathcal{B}^l} \Gamma(\beta) = z_{LP}(d^l) - z_{IP}(b^l)$.*

Proof: For any $\beta \in \mathcal{B}^l$, $z_{LP}(\beta) \leq z_{LP}(d^l)$ and $z_{IP}(b^l) \leq z_{IP}(\beta)$, which implies $z_{LP}(\beta) - z_{IP}(\beta) \leq z_{LP}(d^l) - z_{IP}(b^l)$. Fix $\epsilon > 0$. By Lemma 3, there exists $\beta^* \in (\mathcal{B}^l)^\circ$ such that $z_{LP}(d^l) - z_{LP}(\beta^*) < \epsilon$. Because $A \in \mathbb{Z}^{m \times n}$ and $\lfloor \beta^* \rfloor = b^l$, we have that $z_{IP}(\beta^*) = z_{IP}(b^l)$. Therefore, $[z_{LP}(d^l) - z_{IP}(b^l)] - [z_{LP}(\beta^*) - z_{IP}(\beta^*)] < \epsilon$. Therefore, $z_{LP}(d^l) - z_{IP}(b^l) = \sup\limits_{\beta \in \mathcal{B}^l} [z_{LP}(\beta) - z_{IP}(\beta)] = \sup\limits_{\beta \in \mathcal{B}^l} \Gamma(\beta)$. $\square$

Corollary 9. $\Delta_{SAH} = \sup\limits_{\beta \in \mathcal{B}(0,b)} \Gamma(\beta) = \sup\limits_{l \in \{1,...,L\}} z_{LP}(d^l) - z_{IP}(b^l)$.

We omit the proof of Corollary 9 as it follows from Proposition 11.

Proposition 12. *The optimal solution of* (8) *is* Δ_{SAH}.

Proof: Let $\tilde{\beta} \in \arg\max\{z_{LP}(r(\beta)) - z_{IP}(\beta) \mid \beta \in \widehat{\mathcal{B}}(0,b)\}$ so that $z_{LP}(r(\tilde{\beta})) - z_{IP}(\tilde{\beta}) = \sup\{z_{LP}(d^\ell) - z_{IP}(b^\ell) = \Delta_{SAH}$, by Corollary 9. Let $\tilde{\phi} = z_{IP}$, $\tilde{\psi}(\tilde{\beta}) = z_{LP}(r(\tilde{\beta})) - z_{IP}(\tilde{\beta})$, and $\tilde{\psi}(\beta) = 0$, for all $\beta \in \widehat{\mathcal{B}}(0,b) \backslash \tilde{\beta}$. It is immediate that $\tilde{\phi} \in \Phi(b)$ (Proposition 2). By construction, $\tilde{\psi}$ satisfies (8b). For all $\beta \in \widehat{\mathcal{B}}(0,b) \backslash \tilde{\beta}$ and $k \in \mathcal{K}$, $\tilde{\psi}(\beta) = 0 \leq z_{LP}(r(\beta)) - z_{IP}(\beta) \leq r(\beta)^T v^k - \tilde{\phi}(\beta)$. Additionally, $\tilde{\psi}(\tilde{\beta}) = z_{LP}(r(\tilde{\beta})) - z_{IP}(\tilde{\beta}) \leq r(\tilde{\beta})^T v^k - \tilde{\phi}(\tilde{\beta})$, for all $k \in \mathcal{K}$. Hence, $(\tilde{\phi}, \tilde{\psi})$ is feasible for (8).

Let $(\bar{\phi}, \bar{\psi})$ be feasible for (8). Then $\bar{\psi}(\beta) = 0$ for all $\beta \in \widehat{\mathcal{B}}(0,b)\backslash\bar{\beta}$ for some $\bar{\beta} \in \widehat{\mathcal{B}}(0,b)$. Because $\bar{\phi} \in \Phi(b)$, by Corollary 3, $\bar{\phi}(\beta) \geq z_{IP}(\beta)$ for all $\beta \in \widehat{\mathcal{B}}(0,b)$. Let $k^* \in \arg\min\{r(\bar{\beta})^T v^k \mid k \in \mathcal{K}\}$. By (8a), $\bar{\psi}(\bar{\beta}) \leq r(\bar{\beta})^T v^{k^*} - \bar{\phi}(\bar{\beta}) \leq z_{LP}(r(\bar{\beta})) - z_{IP}(\bar{\beta}) \leq z_{LP}(r(\tilde{\beta})) - z_{IP}(\tilde{\beta}) = \tilde{\psi}(\tilde{\beta})$. Hence, $\sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \tilde{\psi}(\beta) \geq \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \bar{\psi}(\beta)$. This proves $(\tilde{\phi}, \tilde{\psi})$ is optimal for (8). Further, $\delta_{SAH} = \sum_{\beta \in \widehat{\mathcal{B}}(0,b)} \tilde{\psi}(\beta) = \tilde{\psi}(\tilde{\beta}) = \Delta_{SAH}$. □

Proposition 13. *The infimum of the absolute gap function over $\mathcal{B}(0,b)\backslash\mathcal{B}(0,1)$ can be computed by considering only the integral points in $\mathcal{B}(0,b)$. Specifically,* $\inf_{\beta \in \mathcal{B}(0,b)\backslash\mathcal{B}(0,1)} \Gamma(\beta) = \min_{\beta \in \widehat{\mathcal{B}}(0,b)\backslash\{0\}} \Gamma(\beta)$. *Further,* (5) *can be used to compute* $\Delta_{IAH} = \inf_{\beta \in \mathcal{B}(0,b)\backslash\mathcal{B}(0,1)} \Gamma(\beta)$.

Proof: Let $\beta \in \mathcal{B}(0,b)\backslash\mathcal{B}(0,1)$, then $\Gamma(\beta) = z_{LP}(\beta) - z_{IP}(\beta) = z_{LP}(\beta) - z_{IP}(\lfloor\beta\rfloor) \geq z_{LP}(\lfloor\beta\rfloor) - z_{IP}(\lfloor\beta\rfloor) = \Gamma(\lfloor\beta\rfloor)$. Because $\beta \in \mathcal{B}(0,b)\backslash\mathcal{B}(0,1), \lfloor\beta\rfloor \in \widehat{\mathcal{B}}(0,b)\backslash\{0\}$. In addition, for any $\beta' \in \widehat{\mathcal{B}}(0,1)$ and $\epsilon > 0$, using arguments similar to those in Lemma 3, from the continuity of z_{LP} and the fact that $z_{IP}(\tilde{\beta}) = z_{IP}(\beta')$, there exists $\tilde{\beta} \in (\mathcal{B}(\beta', \beta' + 1))^\circ \subset \mathcal{B}(0,b)\backslash\mathcal{B}(0,1)$ such that $\Gamma(\tilde{\beta}) - \Gamma(\beta') < \epsilon$. Hence, $\Delta_{IAH} = \inf_{\beta \in \mathcal{B}(0,b)\backslash\mathcal{B}(0,1)} \Gamma(\beta) = \min_{\beta \in \widehat{\mathcal{B}}(0,b)\backslash\{0\}} \Gamma(\beta) = \Delta_{IAD}$, by Proposition 8. □

D Section 5 Results

Proposition 14. *The optimal objective value of* (9) *is* $\mathbb{E}_{\xi^{(1)}}(\gamma(\xi^{(1)}))$.

Proof: Let $(\tilde{\phi}, \tilde{\psi}) = (z_{IP}, \gamma)$. It is immediate that $\tilde{\phi} \in \Phi(b)$ (Proposition 2). Also, for each $\beta \in \widehat{\mathcal{B}}^+(0,b)$, $\gamma(\beta) \cdot \beta^T v^k \geq \gamma(\beta) \cdot z_{LP}(\beta) = z_{IP}(\beta) = \tilde{\phi}(\beta)$, for all $k \in \mathcal{K}$. Hence, $(\tilde{\phi}, \tilde{\psi})$ is feasible for (9).

Let $(\bar{\phi}, \bar{\psi})$ be feasible for (9). Because $\bar{\phi} \in \Phi(b)$, by Corollary 3, $\bar{\phi}(\beta) \geq z_{IP}(\beta)$, for all $\beta \in \widehat{\mathcal{B}}^+(0,b)$. By (9b), $\bar{\psi}(\beta) \geq \frac{\bar{\phi}(\beta)}{\beta^T v^k}$, for all $k \in \mathcal{K}$, which implies $\bar{\psi}(\beta) \geq \frac{\bar{\phi}(\beta)}{z_{LP}(\beta)} \geq \gamma(\beta) = \tilde{\psi}(\beta)$. Therefore, $\sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \mu^{(1)}(\beta)\bar{\psi}(\beta) \geq \sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \mu^{(1)}(\beta)\tilde{\psi}(\beta)$, and $(\tilde{\phi}, \tilde{\psi})$ is optimal. Thus, the optimal objective value is $\sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \mu^{(1)}(\beta)\tilde{\psi}(\beta) = \sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \mu^{(1)}(\beta)\gamma(\beta) = \mathbb{E}_{\xi^{(1)}}(\gamma(\beta))$. □

Corollary 10. *Suppose $\mu^{(1)}(\beta) > 0$ for all $\beta \in \widehat{\mathcal{B}}^+(0,b)$. Then, any feasible solution $(\bar{\phi}, \bar{\psi})$ in which $\bar{\phi} \neq z_{IP}$ is not optimal for* (9).

We omit the proof of Corollary 10 due to reasoning similar to that of Corollary 5.

The variance of the relative gap function, denoted by σ_γ^2, indicates the extent to which this approximation error varies over $\widehat{\mathcal{B}}^+(0,b)$. The variance can be calculated from optimal solutions of (9).

Corollary 11. *Let $(\tilde{\phi}, \tilde{\psi})$ be an optimal solution to (9) and δ_{ERD} the optimal objective value. Then the variance of the relative gap over $\widehat{\mathcal{B}}(0,b)$ can be calculated as follows:*

$$\sigma_\gamma^2 = \sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \mu^{(1)}(\beta)(\tilde{\psi}(\beta))^2 - \delta_{ERD}^2.$$

The proof of Corollary 11 is similar to that of Corollary 6 and is omitted.

Proposition 15. *The optimal objective value of (10) is Δ_{SRD}.*

The proof is similar to that of Proposition 8 and is omitted.

Proposition 16. *The optimal objective value of (11) is $1 - \Delta_{IRD}$.*

The proof is similar to that of Proposition 7 and is omitted.

E Section 6 Results

Proposition 17. *The supremum of the relative gap function over $\mathcal{B}(0,b)^+$ can be computed by only considering the integral points. Specifically, $\sup_{\beta \in \mathcal{B}^+(0,b)} \gamma(\beta) = \max_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \gamma(\beta)$. Further, (10) can be used to compute $\Delta_{SRH} = \sup_{\beta \in \mathcal{B}^+(0,b)} \gamma(\beta)$.*

Proof: For any $\beta \in \mathcal{B}^+(0,b)$, $\gamma(\beta) = \frac{z_{IP}(\beta)}{z_{LP}(\beta)} = \frac{z_{IP}(\lfloor\beta\rfloor)}{z_{LP}(\beta)} \leq \frac{z_{IP}(\lfloor\beta\rfloor)}{z_{LP}(\lfloor\beta\rfloor)}$, and because $\beta \in \mathcal{B}^+(0,b)$, $\lfloor\beta\rfloor \in \widehat{\mathcal{B}}^+(0,b)$. Hence, $\Delta_{SRH} = \sup_{\beta \in \mathcal{B}^+(0,b)} \gamma(\beta) = \max_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \gamma(\beta) = \Delta_{SRD}$, by Proposition 12. □

Proposition 18. *For any unit hyper-cube $\mathcal{B}^l$, $\inf_\beta\{\gamma(\beta) \mid \beta \in \mathcal{B}^l\} = z_{IP}(b^l)/z_{LP}(d^l)$.*

Proof: For any $\beta \in \mathcal{B}^l$, $z_{LP}(\beta) \leq z_{LP}(d^l)$ and $z_{IP}(b^l) \leq z_{IP}(\beta)$, which implies $z_{IP}(\beta)/z_{LP}(\beta) \geq z_{IP}(b^l)/z_{LP}(d^l)$. Fix $\epsilon > 0$. By Lemma 3, there exists $\beta^* \in (\mathcal{B}^l)^\circ$ such that $z_{LP}(d^l) - z_{LP}(\beta^*) \leq$

$\frac{\epsilon z_{LP}(d^l) z_{LP}(b^l)}{z_{IP}(b^l)+1}$. Then,

$$\begin{aligned}
\frac{z_{IP}(\beta^*)}{z_{LP}(\beta^*)} - \frac{z_{IP}(b^l)}{z_{LP}(d^l)} &= \frac{z_{IP}(b^l)}{z_{LP}(d^l)} \cdot \frac{z_{LP}(d^l) - z_{LP}(\beta^*)}{z_{LP}(\beta^*)} \\
&\leq \frac{z_{IP}(b^l)}{z_{LP}(d^l)} \cdot \frac{z_{LP}(d^l) - z_{LP}(\beta^*)}{z_{LP}(b^l)} \\
&\leq \frac{z_{IP}(b^l)}{z_{LP}(d^l) z_{LP}(b^l)} \cdot \frac{z_{LP}(d^l) z_{LP}(b^l)}{z_{IP}(b^l)+1} \cdot \epsilon \\
&< \epsilon.
\end{aligned}$$

By definition of infimum, $z_{IP}(b^l)/z_{LP}(d^l) = \inf_{\beta \in \mathcal{B}^l} \gamma(\beta)$. $\square$

Corollary 12. $\Delta_{IRH} = \inf_{\beta \in \mathcal{B}^+(0,b)} \gamma(\beta) = \inf_{l \in \{1,\ldots,L\}} z_{IP}(b^l)/z_{LP}(d^l)$.

Corollary 12 follows from Proposition 18, and we omit the proof.

Proposition 19. $\delta_{IRH} = 1 - \Delta_{IRH}$.

Proof: Let $\tilde{\beta} \in \arg\min_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \{z_{IP}(\beta)/z_{LP}(r(\beta))\}$, $\tilde{\psi}(\beta) = 1 - \Delta_{IRH}$, if $\beta = \tilde{\beta}$ and 0 otherwise, and let $\tilde{\phi} = z_{IP}$. It is immediate that $\tilde{\phi} \in \Phi(b)$ (Proposition 2). Observe that $(1 - \tilde{\psi}(\tilde{\beta}))r(\tilde{\beta})^T v^k = \Delta_{IRH} r(\tilde{\beta})^T v^k \geq \Delta_{IRH} z_{LP}(r(\tilde{\beta})) = z_{IP}(\tilde{\beta}) \geq \tilde{\phi}(\tilde{\beta})$, for all $k \in \mathcal{K}$. Further, for all $\beta \in \widehat{\mathcal{B}}^+(0,b) \backslash \tilde{\beta}$, $(1 - \tilde{\psi}(\beta))r(\beta)^T v^k \geq z_{LP}(r(\beta)) \geq z_{IP}(\beta) = \tilde{\phi}(\beta)$, for all $k \in \mathcal{K}$. Thus, $(\tilde{\phi}, \tilde{\psi})$ is feasible.

Let $(\bar{\phi}, \bar{\psi})$ be feasible for (12). For all $\beta \in \widehat{\mathcal{B}}^+(0,b), k \in \mathcal{K}$, $(1 - \bar{\phi}(\beta))r(\beta)^T v^k \geq \bar{\phi}(\beta) \geq z_{IP}(\beta)$, because $\bar{\phi} \in \Phi(b)$. By (12a), $1 - \bar{\phi}(\beta) \geq \frac{\bar{\phi}(\beta)}{r(\beta)^T v^k} \geq \frac{z_{IP}(\beta)}{r(\beta)^T v^k}$, which implies $1 - \bar{\psi}(\beta) \geq \gamma(\beta) \geq \Delta_{IRH}$. Because $\bar{\psi}(\beta) > 0$ for at most one $\beta \in \widehat{\mathcal{B}}^+(0,b)$, $\sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \bar{\psi}(\beta) \leq 1 - \Delta_{IRH} = \sum_{\beta \in \widehat{\mathcal{B}}^+(0,b)} \tilde{\psi}(\beta)$. Hence, $(\tilde{\phi}, \tilde{\psi})$ is an optimal solution. $\square$