Evaluating Mixed-Integer Programming Models over Multiple Right-hand Sides

Rachael M. Alfant[a], Temitayo Ajayi[b], Andrew J. Schaefer[a,*]

[a] *Department of Computational Applied Mathematics and Operations Research, Rice University, Houston, TX 77005*
[b] *Nature Source Improved Plants, Ithaca, NY 14850*

Abstract

A critical measure of model quality for a mixed-integer program (MIP) is the difference, or gap, between its optimal objective value and that of its linear programming relaxation. In some cases, the right-hand side is not known exactly; however, there is no consensus metric for evaluating a MIP model when considering multiple right-hand sides. In this paper, we provide model formulations for the expectation and extrema of absolute and relative MIP gap functions over finite discrete sets.

Keywords: Mixed-integer programming, superadditive duality, value function

---

1. Introduction

Given a (maximization) mixed-integer program (MIP), the gap is the difference between the optimal objective value of its linear programming (LP) relaxation and that of the MIP. The MIP gap is a critical measure of model quality for MIPs with fixed data. Some theoretical implications include improving solution algorithms, such as branch and bound [22]. Practical implications include the interpretation of the dual objective (price) function, which tells us how much extra resources are worth [31]. In practice, the right-hand side may not be known exactly or it may vary. Thus, evaluative metrics must be developed in order to assess a MIP model's quality over multiple right-hand sides. Such metrics may have applications in sensitivity analysis (e.g., [14, 31]) and stochastic programming (e.g., [20, 26, 29]).

Value functions and superadditive duality play central roles in this paper, and have various applications in optimization. [18] extends integer programming (IP) duality theory (see [16] for an extensive survey on IP duality) to MIPs by examining the group problem. [28] characterize MIP value functions and present a cutting-plane algorithm for their construction. [6] provide properties

[*] Corresponding author
Email address: andrew.schaefer@rice.edu (Andrew J. Schaefer)

of the MIP value function and use superadditivity to conduct sensitivity analyses on the optimal value. [7] bound the MIP gap as the right-hand side varies, and [5] identifies a class of computable formulas that precisely characterize value functions of MIPs.

Although our paper is, to our knowledge, the first to study MIP gap functions using superadditivity, there is an existing body of literature on IP gap functions and superadditivity. Most relevant, [1] optimize IP gap functions over multiple right-hand sides, whereas our paper optimizes MIP gap functions over multiple right-hand sides. There are a number of non-trivial differences between the optimization problems presented in [1] and our paper that make our models more complex, including our focus on MIPs versus IPs and the subsequent inclusion of dual variables for both the IP and LP embedded in the MIP. Thus, we present a novel framework by which to evaluate the quality of a MIP model over multiple right-hand sides. Furthermore, this paper presents a novel proof of strong duality for the MIP superadditive dual proposed in [21], in addition to a novel proof of the periodicity of absolute MIP gap functions.

2. Preliminaries

Let $\mathbf{A} \in \mathbb{Z}_{+}^{m\times n}$, $\mathbf{G} \in \mathbb{Q}_{+}^{m\times p}$, $\mathbf{b} \in \mathbb{R}_{+}^{m}$, $\mathbf{c} \in \mathbb{R}_{++}^{n}$, and $\mathbf{h} \in \mathbb{R}_{++}^{p}$. Let $\mathbf{a}_j$ be the j^{th} column of $\mathbf{A}$ and $\mathbf{g}_k$ the k^{th} column of $\mathbf{G}$. Consider the MIP problem:

$$z_{MIP}(\mathbf{b}) \;=\; \max_{\mathbf{x}\in\mathbb{Z}_{+}^{n},\ \mathbf{y}\in\mathbb{R}_{+}^{p}} \{\mathbf{c}^{\top}\mathbf{x} + \mathbf{h}^{\top}\mathbf{y} \mid \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{y} \le \mathbf{b}\}. \tag{1}$$

Let $z_{LPR}(\mathbf{b})$ be the optimal objective value of the LP relaxation of (1) with right-hand side $\mathbf{b}$. In this paper, we study MIP gaps over multiple right-hand sides. Thus, define $\mathcal{B}[\mathbf{0},\mathbf{b}] := \prod_{i=0}^{m}[0,b_i]$, i.e., the Cartesian product of the intervals $[0,b_1],\ldots,[0,b_m]$, and $\widehat{\mathcal{B}}[\mathbf{0},\mathbf{b}] := \mathcal{B}[\mathbf{0},\mathbf{b}]\cap\mathbb{Z}_{+}^{m}$. We assume the right-hand side parameter $\widehat{\boldsymbol{\beta}}$ is in $\mathcal{B}[\mathbf{0},\mathbf{b}]$ and $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0},\mathbf{b}]$ is such that $\boldsymbol{\beta} \le \widehat{\boldsymbol{\beta}}$. We formally define MIP gap functions as follows.

Definition 2.1. Given a set of right-hand sides, $\mathcal{B}[\mathbf{0},\mathbf{b}]$, the *absolute gap function* for MIPs is defined as: $\boldsymbol{\Gamma} : \mathcal{B}[\mathbf{0},\mathbf{b}] \to \mathbb{R}_{+}\cup\{\infty\}$, $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}) := z_{LPR}(\widehat{\boldsymbol{\beta}}) - z_{MIP}(\widehat{\boldsymbol{\beta}})$. Given a set of right-hand sides, $\mathcal{B}^{+}[\mathbf{0},\mathbf{b}] := \{\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0},\mathbf{b}] \mid z_{MIP}(\widehat{\boldsymbol{\beta}}) > 0\}$, the *relative gap function* for MIPs is defined as: $\boldsymbol{\gamma} : \mathcal{B}^{+}[\mathbf{0},\mathbf{b}] \to \mathbb{R}_{+}$, $\boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}}) := \frac{z_{MIP}(\widehat{\boldsymbol{\beta}})}{z_{LPR}(\widehat{\boldsymbol{\beta}})}$.

An absolute gap that is close to zero indicates that the LP relaxation provides a high-quality approximation for the optimal objective value of the corresponding MIP. In addition, because (1) is a maximization optimization problem, $z_{LPR}(\widehat{\boldsymbol{\beta}})$ is an upper bound for $z_{MIP}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0},\mathbf{b}]$.

Thus, $\mathbf{\Gamma}(\widehat{\boldsymbol{\beta}}) \geq 0$. A relative gap that is close to 1 indicates that the LP relaxation provides a high-quality approximation for the optimal objective value of the corresponding MIP. The domain of $\boldsymbol{\gamma}$ is restricted to $\mathcal{B}^+[\mathbf{0}, \mathbf{b}]$ in order to avoid division by zero. Thus, $\boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}}) \in [0, 1]$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{B}^+[\mathbf{0}, \mathbf{b}]$.

2.1. IP Value Functions and Duality

Our approach to gap functions for MIPs is very closely related to MIP value functions and MIP duality. Thus, to study gap functions for MIPs, we first characterize superadditive duality for pure IPs and define value functions of pure IPs and LPs. For any $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$ and $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$, the parametrized LP, LP$(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$, with value function z_{LP}, is defined as:

$$z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) := \max_{\mathbf{y} \in \mathbb{R}^p_+} \{\mathbf{h}^\top \mathbf{y} \mid \mathbf{G}\mathbf{y} \leq \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\}. \tag{2}$$

The dual of LP$(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$, LPD$(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$, with value function z_{LPD}, is defined as follows:

$$z_{LPD}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) := \min_{\boldsymbol{\pi} \in \mathbb{R}^m_+} \{\boldsymbol{\pi}^\top (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \mid \boldsymbol{\pi}^\top \mathbf{G} \geq \mathbf{h}^\top\}. \tag{3}$$

Let $\mathcal{Y} := \{\boldsymbol{\pi} \in \mathbb{R}^m_+ \mid \boldsymbol{\pi}^\top \mathbf{G} \geq \mathbf{h}^\top\}$. Denote $\mathbf{\Omega} := \{\boldsymbol{\pi}^r \mid r \in \mathcal{R}\}$ the set of extreme points of $\mathcal{Y}$. Because the primal (2) is feasible, (3) is always bounded. Thus, $\min_{\boldsymbol{\pi} \in \mathbf{\Omega}} \boldsymbol{\pi}^\top (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = z_{LPD}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$. Furthermore, because there exist a finite number of constraints and variables for LPD$(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$, the set of extreme points of $\mathcal{Y}$ is finite, i.e., $|\mathcal{R}| < +\infty$, as a consequence of Weyl's Theorem [10]. We use the variable, $\boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}}$, to model the value of z_{LPD} over $\widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$.

Assumption 2.1. $\mathbf{A}$ and $\mathbf{G}$ have no zero columns.

Assumption 2.1 implies finite optima for LP$(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$; weak duality then yields finite optima for LPD$(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$. For $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$, the parametrized IP, IP$(\boldsymbol{\beta})$, with value function z_{IP}, is:

$$z_{IP}(\boldsymbol{\beta}) := \max_{\mathbf{x} \in \mathbb{Z}^n_+} \{\mathbf{c}^\top \mathbf{x} \mid \mathbf{A}\mathbf{x} \leq \boldsymbol{\beta}\}.$$

Remark 2.1. For all $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$, $z_{IP}(\boldsymbol{\beta}) < +\infty$ and IP$(\boldsymbol{\beta})$ is feasible. In addition, z_{IP} is superadditive, i.e., for any $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$ with $\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$, $z_{IP}(\boldsymbol{\beta}_1) + z_{IP}(\boldsymbol{\beta}_2) \leq z_{IP}(\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2)$.

Because we assume the data are nonnegative, Remark 2.1 is a direct result of Assumption 2.1.

Definition 2.2. [8] *Chvátal functions* are a recursively defined class of functions constructed using sums, nonnegative multiples, and floors of linear functions. *Gomory functions* are similar, but also include minimums of linear functions.

There are various formulations for the dual of IP($\boldsymbol{\beta}$). We use the superadditive dual, denoted $SIP(\boldsymbol{\beta})$, as it is a strong dual to IP for all $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$ [1, 17, 31]. Furthermore, it is particularly adaptable for developing measures of model quality over multiple right-hand sides. The formulation is as follows [1, 31]:

$$\begin{aligned} z_{SIP}(\boldsymbol{\beta}) := \min \quad & \boldsymbol{\phi}(\boldsymbol{\beta}) && \text{(4a)} \\ \text{s.t.} \quad & \boldsymbol{\phi}(\mathbf{a}_j) \geq c_j \quad \forall\, j \in 1, 2, \ldots, n, && \text{(4b)} \\ & \boldsymbol{\phi} \text{ nondecreasing and superadditive,} && \text{(4c)} \\ & \boldsymbol{\phi}(\mathbf{0}) = 0, && \text{(4d)} \\ & \boldsymbol{\phi}(\boldsymbol{\beta}_1) \in \mathbb{R} \quad \forall\, \boldsymbol{\beta}_1 \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]. && \text{(4e)} \end{aligned}$$

Note that we use the variable, $\boldsymbol{\phi}(\boldsymbol{\beta})$, to model the value of z_{SIP} over $\widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$. Denote $\boldsymbol{\Phi}(\widehat{\boldsymbol{\beta}}) := \{\boldsymbol{\phi} \in \mathbb{R}^{|\widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]|} \mid \text{(4b)} - \text{(4e)}\}$.

[1] use superadditive duality to model IP gap functions over multiple right-hand sides. We provide an analogous framework for MIPs: in particular, we use superadditive duality to model MIP gap functions over multiple (discrete) right-hand sides. For the MIP extension, we must account for the continuous variables by using dual extreme points, which significantly complicates the superadditive dual formulation of MIP, as discussed in the following section.

2.2. MIP Value Functions and Superadditive Duality

Let $\mathcal{S}(\widehat{\boldsymbol{\beta}}) := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{Z}^n_+ \times \mathbb{R}^p_+ \mid \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{y} \leq \widehat{\boldsymbol{\beta}}\}$. We define the parametrized mixed-integer program, MIP($\widehat{\boldsymbol{\beta}}$), with value function z_{MIP}, as:

$$z_{MIP}(\widehat{\boldsymbol{\beta}}) := \max_{\mathbf{x}, \mathbf{y}} \{\mathbf{c}^\top \mathbf{x} + \mathbf{h}^\top \mathbf{y} \mid (\mathbf{x}, \mathbf{y}) \in \mathcal{S}(\widehat{\boldsymbol{\beta}})\}. \tag{5}$$

As with IP value functions, z_{MIP} is superadditive [6]. Duality for MIPs is more complex than that of IPs and LPs because we must account for both the integer and continuous variables in the MIP formulation. We do this by computing the gap function over $\widehat{\boldsymbol{\beta}}$ (the right-hand side corresponding to the MIP), while simultaneously solving for the optimal portion of $\widehat{\boldsymbol{\beta}}$ to allocate to the IP problem (with right-hand side $\boldsymbol{\beta}$) versus the LP problem (with right-hand side $\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}$) embedded in the MIP. One possible approach is to use a formulation similar to (4), with an additional constraint containing a directional derivative that accounts for the continuous variables. However, having a constraint containing a directional derivative may present additional modeling complications. For this reason, we instead formulate the superadditive dual of MIP by exploiting the structure of the MIP value function presented in Proposition 2.1.

Proposition 2.1. [21] *For any $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$, $z_{MIP}(\widehat{\boldsymbol{\beta}}) := \max_{\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]} \{z_{IP}(\boldsymbol{\beta}) + z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})\}$.*

Proposition 2.1 decomposes the value function, z_{MIP}, into its integer and continuous value functions. We exploit this property to construct a superadditive dual formulation to MIP that does not require the use of directional derivatives. We compute the gap function over a (potentially unknown) right-hand side parameter, $\widehat{\boldsymbol{\beta}}$, while also solving for the optimal portion of $\widehat{\boldsymbol{\beta}}$ to allocate to the IP problem versus the LP problem embedded in the MIP; as such, we propose an alternative MIP dual formulation, SDMIP($\widehat{\boldsymbol{\beta}}$):

$$z_{SDMIP}(\widehat{\boldsymbol{\beta}}) := \max_{\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]} \{z_{SIP}(\boldsymbol{\beta}) + z_{LPD}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})\}.$$

Recall $\boldsymbol{\Phi}(\widehat{\boldsymbol{\beta}}) := \{\boldsymbol{\phi} \in \mathbb{R}^{|\widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]|} \mid (4b) - (4e)\}$, i.e., the feasible region of SIP($\widehat{\boldsymbol{\beta}}$). [1] use SIP($\boldsymbol{\beta}$) to model IP gap functions parametrized over a set of right-hand sides as LPs (albeit of exponentially large size) with at most one SOS1 constraint. In this paper, we use the MIP superadditive dual formulation presented in Proposition 2.2 to model MIP gap functions over multiple (discrete) right-hand sides as (exponentially large) LPs with at most one SOS1 constraint. Proposition 2.2 extends the dual formulation presented in [31] to MIPs.

Proposition 2.2. [21] *Let $\mathbf{A} \in \mathbb{Z}_{+}^{m \times n}$, $\mathbf{G} \in \mathbb{Q}_{+}^{m \times p}$, and $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$. SDMIP($\widehat{\boldsymbol{\beta}}$) is equivalent to:*

$$\begin{aligned}
z_{SDMIP}(\widehat{\boldsymbol{\beta}}) \;:=\; \min_{\boldsymbol{\phi}, \boldsymbol{\pi}} \quad & \boldsymbol{\phi}(\boldsymbol{\beta}^{'}) + \boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{'}}^{\top}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{'}) && \text{(6a)}\\
s.t. \quad & \boldsymbol{\phi}(\boldsymbol{\beta}^{'}) + \boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{'}}^{\top}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{'}) \geq \boldsymbol{\phi}(\boldsymbol{\beta}) + \boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}}^{\top}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \quad \forall\, \boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}], && \text{(6b)}\\
& \boldsymbol{\phi} \in \boldsymbol{\Phi}(\widehat{\boldsymbol{\beta}}), && \text{(6c)}\\
& \boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}} \in \boldsymbol{\Omega} \quad \forall\, \boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]. && \text{(6d)}
\end{aligned}$$

Vector $\boldsymbol{\phi}$ is indexed by $\boldsymbol{\beta}$; vector $\boldsymbol{\pi}$ is indexed by $(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ and dot-producted with $(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ in (6a) and (6b). Formulation (6) avoids various modeling complications presented by the use of directional derivatives, but at the expense of a larger LP. We present what are, to our knowledge, novel proofs showing that SDMIP($\widehat{\boldsymbol{\beta}}$) is both a weak and strong dual to MIP($\widehat{\boldsymbol{\beta}}$) for all $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$.

Theorem 2.1. *Let $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$, and let $\boldsymbol{\beta}^* \in \arg\max_{\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]} z_{IP}(\boldsymbol{\beta}) + z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$. Then, for $\boldsymbol{\phi} \in \boldsymbol{\Phi}(\widehat{\boldsymbol{\beta}})$ and $\boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*} \in \boldsymbol{\Omega}$, we have that $\boldsymbol{\phi}(\boldsymbol{\beta}^*) + \boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*}^{\top}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \geq z_{MIP}(\widehat{\boldsymbol{\beta}})$.*

Proof. Let $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$. Choose $\boldsymbol{\beta}^* \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$ such that $z_{MIP}(\widehat{\boldsymbol{\beta}}) = z_{IP}(\boldsymbol{\beta}^*) + z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)$. Let $\boldsymbol{\phi} \in \boldsymbol{\Phi}(\widehat{\boldsymbol{\beta}})$ and $\boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*} \in \boldsymbol{\Omega}$ be such that $(\boldsymbol{\phi}, \boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*})$ is feasible for SDMIP($\widehat{\boldsymbol{\beta}}$). Then, $z_{SDMIP}(\widehat{\boldsymbol{\beta}}) \leq$

$\phi(\beta^*)+\pi^\top_{\widehat{\beta}-\beta^*}(\widehat{\beta}-\beta^*)$. By IP weak duality [1, 31], $z_{IP}(\beta^*) \le \phi(\beta^*)$. By LP weak duality, $z_{LP}(\widehat{\beta}-\beta^*) \le \pi^\top_{\widehat{\beta}-\beta^*}(\widehat{\beta}-\beta^*)$. Thus, $z_{MIP}(\widehat{\beta}) = z_{IP}(\beta^*) + z_{LP}(\widehat{\beta}-\beta^*) \le \phi(\beta^*) + \pi^\top_{\widehat{\beta}-\beta^*}(\widehat{\beta}-\beta^*)$. □

Theorem 2.2. *Let $\widehat{\beta} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$. Then, SDMIP($\widehat{\beta}$) is a strong dual to MIP($\widehat{\beta}$).*

Proof. Let $\widehat{\beta} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$. Choose $\beta^* \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\beta}]$ such that $z_{MIP}(\widehat{\beta}) = z_{IP}(\beta^*)+z_{LP}(\widehat{\beta}-\beta^*)$. Let $\mathbf{x}^* \in \text{opt}_{IP}(\beta^*)$ and $\mathbf{y}^* \in \text{opt}_{LP}(\widehat{\beta}-\beta^*)$. By IP strong duality [1], $z_{IP}(\beta^*) = \mathbf{c}^\top\mathbf{x}^* = \phi^*(\beta^*) = z_{SIP}(\beta^*)$ for some $\phi^* \in \mathbf{\Phi}(\widehat{\beta})$. Also, by LP duality, $z_{LP}(\widehat{\beta}-\beta^*) = \mathbf{h}^\top\mathbf{y}^* = {\pi^*}^\top_{\widehat{\beta}-\beta^*}(\widehat{\beta}-\beta^*) = z_{LPD}(\widehat{\beta}-\beta^*)$ for ${\pi^*}_{\widehat{\beta}-\beta^*} \in \mathbf{\Omega}$ where $\pi^*_{\widehat{\beta}-\beta^*} \in \arg\min_{\pi\in\mathbf{\Omega}} \pi^\top(\widehat{\beta}-\beta^*)$. Now, consider the tuple $(\phi^*, {\pi^*}_{\widehat{\beta}-\beta^*})$. Note that this tuple is feasible for SDMIP($\widehat{\beta}$): ϕ^* satisfies (6c), ${\pi^*}_{\widehat{\beta}-\beta^*}$ satisfies (6d), and because $\beta^* \in \arg\max_{\beta\in\widehat{\mathcal{B}}[\mathbf{0},\widehat{\beta}]}\{z_{IP}(\beta) + z_{LP}(\widehat{\beta}-\beta)\}$, $\phi^*(\beta^*) + {\pi^*}^\top_{\widehat{\beta}-\beta^*}(\widehat{\beta}-\beta^*) \ge \phi^*(\beta) + {\pi^*}^\top_{\widehat{\beta}-\beta}(\widehat{\beta}-\beta)$ for all $\beta \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\beta}]$, thus satisfying (6b). Furthermore: $\phi^*(\beta^*) + {\pi^*}^\top_{\widehat{\beta}-\beta^*}(\widehat{\beta}-\beta^*) = z_{SIP}(\beta^*) + z_{LPD}(\widehat{\beta}-\beta^*) = z_{IP}(\beta^*)+z_{LP}(\widehat{\beta}-\beta^*) = z_{MIP}(\widehat{\beta})$. By Theorem 2.1, $\phi^*(\beta^*) + {\pi^*}^\top_{\widehat{\beta}-\beta^*}(\widehat{\beta}-\beta^*) = z_{MIP}(\widehat{\beta}) \le \phi(\beta^*) + \pi^\top_{\widehat{\beta}-\beta^*}(\widehat{\beta}-\beta^*)$ for all $\phi \in \mathbf{\Phi}(\widehat{\beta})$ and $\pi_{\widehat{\beta}-\beta^*} \in \mathbf{\Omega}$. Thus, $z_{SDMIP}(\widehat{\beta}) = \phi^*(\beta^*) + {\pi^*}^\top_{\widehat{\beta}-\beta^*}(\widehat{\beta}-\beta^*)$, and $z_{MIP}(\widehat{\beta}) = z_{SDMIP}(\widehat{\beta})$. □

Now, consider the LP relaxation of (5): $z_{LPR}(\widehat{\beta}) := \max_{\mathbf{x}\in\mathbb{R}^n_+,\ \mathbf{y}\in\mathbb{R}^p_+}\{\mathbf{c}^\top\mathbf{x} + \mathbf{h}^\top\mathbf{y} \mid \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{y} \le \widehat{\beta}\}$. As with IP and MIP value functions, z_{LPR} is superadditive.

Remark 2.2. For all $\widehat{\beta} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$, $z_{LPR}(\widehat{\beta}) < +\infty$ and LPR($\widehat{\beta}$) is feasible.

Because we assume the data are nonnegative, Remark 2.2 follows from Assumption 2.1. Let $\mathcal{Q} := \{\mathbf{u} \in \mathbb{R}^m_+ \mid \mathbf{A}^\top\mathbf{u} \ge \mathbf{c},\ \mathbf{G}^\top\mathbf{u} \ge \mathbf{h}\}$, and let $\{\mathbf{u}^q \mid q \in \kappa\}$ be the set of extreme points of $\mathcal{Q}$. The dual of $z_{LPR}(\widehat{\beta})$ may be formulated as follows:

$$z_{DLPR}(\widehat{\beta}) \ = \ \min_{q\in\kappa} \widehat{\beta}^\top\mathbf{u}^q.$$

Because $\mathbf{c}$ and $\mathbf{h}$ are strictly positive, as a result of Assumption 2.1, Remark 2.2, and Weyl's Theorem [10], DLPR($\widehat{\beta}$) is feasible and $z_{DLPR}(\widehat{\beta}) \ge 0$ for all $\widehat{\beta} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$. Furthermore, there are a finite number of extreme points, i.e., $|\kappa| < +\infty$.

Remark 2.3. There always exists an extreme point of $\mathcal{Q}$ that is an optimal solution to DLPR($\widehat{\beta}$).

Remark 2.3 allows for one to encode the objective function of DLPR($\widehat{\beta}$) as a function of the extreme points of $\mathcal{Q}$. We exploit this in Sections 4 and 5, where we optimize the expectation and extrema of absolute and relative MIP gap functions over finite discrete sets.

3. Properties of Absolute MIP Gap Functions

Absolute MIP gap functions have a number of properties that are unrelated to superadditivity, but are interesting nonetheless - particularly because these properties may lead to algorithmic innovations in the computation of absolute MIP gap functions. We begin by relating Gomory functions to absolute MIP gap functions (see [1] for a proof of the IP case).

Proposition 3.1. *The absolute MIP gap function defined over* $\mathcal{B}[\mathbf{0}, \mathbf{b}]$ *is the minimum of finitely many Gomory functions.*

The proof of Proposition 3.1 (and other results with omitted proofs) is in the E.C. The remainder of this section presents results for absolute MIP gap function periodicity. [13] proves that the absolute gap function for IPs is periodic with respect to the columns of the constraint matrix. [1] use superadditivity and IP complementary slackness to reproduce this result for absolute IP gap functions. We provide a generalization of these results that apply to absolute MIP gap functions.

Proposition 3.2. *Let* $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$ *and* $(\mathbf{x}^*, \mathbf{y}^*) \in opt_{MIP}(\widehat{\boldsymbol{\beta}})$*. Given* $\eta \in \mathbb{N}$*, let* $J_\eta \subseteq \{1, \ldots, n\}$ *denote the set of indices such that* $x_j^* \geq \eta$ *for* $j \in J_\eta$*. Denote* $K := \{1, \ldots, p\}$*, and let* $\lambda^* = \min_{k \in K}\{y_k^*\}$*. Then, for any* $j \in J_\eta$*,* $k \in K$*, and* $\lambda \in [0, \lambda^*]$*,* $z_{MIP}(\widehat{\boldsymbol{\beta}} - \eta \mathbf{a}_j - \lambda \mathbf{g}_k) = z_{MIP}(\widehat{\boldsymbol{\beta}}) - \eta c_j - \lambda h_k$*.*

Proposition 3.2 also applies to z_{LPR}. Note that Proposition 3.2 is a complementary slackness condition: if $J_\eta = \emptyset$, Proposition 3.2 implies LP complementary slackness, and if $K = \emptyset$, Proposition 3.2 implies IP complementary slackness [25]. We use this to prove the following theorem on the periodicity of absolute MIP gap functions.

Theorem 3.1. *Let* $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$*,* $(\tilde{\mathbf{x}}^M, \tilde{\mathbf{y}}^M) \in opt_{MIP}(\widehat{\boldsymbol{\beta}})$*, and* $(\tilde{\mathbf{x}}^L, \tilde{\mathbf{y}}^L) \in opt_{LPR}(\widehat{\boldsymbol{\beta}})$*. Given* $\eta \in \mathbb{N}$*, let* $J_\eta \subseteq \{1, \ldots, n\}$ *denote the set of indices such that* $\tilde{x}_j^M, \tilde{x}_j^L \geq \eta$ *for* $j \in J_\eta$*. Denote* $K := \{1, \ldots, p\}$*, and let* $\lambda^* = \min\{\tilde{y}_1^M, \ldots, \tilde{y}_p^M, \tilde{y}_1^L, \ldots, \tilde{y}_p^L\}$*. Then, for* $j \in J_\eta$*,* $k \in K$*, and* $\lambda \in [0, \lambda^*]$*, we have that* $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}} - \eta \mathbf{a}_j - \lambda \mathbf{g}_k) = \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$*. If, in addition,* $z_{LPR}(\widehat{\boldsymbol{\beta}}) > \eta c_j + \lambda h_k$*, then* $\boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}} - \eta \mathbf{a}_j - \lambda \mathbf{g}_k) = \frac{z_{MIP}(\widehat{\boldsymbol{\beta}}) - \eta c_j - \lambda h_k}{z_{LPR}(\widehat{\boldsymbol{\beta}}) - \eta c_j - \lambda h_k}$*.*

Proof. Let $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$, $(\tilde{\mathbf{x}}^M, \tilde{\mathbf{y}}^M) \in \text{opt}_{MIP}(\widehat{\boldsymbol{\beta}})$, and $(\tilde{\mathbf{x}}^L, \tilde{\mathbf{y}}^L) \in \text{opt}_{LPR}(\widehat{\boldsymbol{\beta}})$. Note that the relative gap function result follows directly from Proposition 3.2. By definition, $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}} - \eta \mathbf{a}_j - \lambda \mathbf{g}_k) = z_{LPR}(\widehat{\boldsymbol{\beta}} - \eta \mathbf{a}_j - \lambda \mathbf{g}_k) - z_{MIP}(\widehat{\boldsymbol{\beta}} - \eta \mathbf{a}_j - \lambda \mathbf{g}_k)$. By hypothesis, we consider a pair of indices, (j, k), with $j \in J_\eta$ and $k \in K$, such that $\tilde{x}_j^M, \tilde{x}_j^L \geq \eta$ and $0 \leq \lambda^* \leq \tilde{y}_k^M, \tilde{y}_k^L$. Then, by Proposition 3.2, $z_{LPR}(\widehat{\boldsymbol{\beta}} - \eta \mathbf{a}_j - \lambda \mathbf{g}_k) = z_{LPR}(\widehat{\boldsymbol{\beta}}) - \eta c_j - \lambda h_k$ and $z_{MIP}(\widehat{\boldsymbol{\beta}} - \eta \mathbf{a}_j - \lambda \mathbf{g}_k) = z_{MIP}(\widehat{\boldsymbol{\beta}}) - \eta c_j - \lambda h_k$

for all $\lambda \in [0, \lambda^*]$. Thus, $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}} - \eta \mathbf{a}_j - \lambda \mathbf{g}_k) = z_{LPR}(\widehat{\boldsymbol{\beta}}) - \eta c_j - \lambda h_k - (z_{MIP}(\widehat{\boldsymbol{\beta}}) - \eta c_j - \lambda h_k) = z_{LPR}(\widehat{\boldsymbol{\beta}}) - z_{MIP}(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$. □

4. Absolute Gap Functions over a Discrete Set

In this section, we present formulations for optimizing the expectation, infimum, and supremum of the absolute gap function, $\boldsymbol{\Gamma}$, over finite discrete sets. Following the notation of [1], each formulation is associated with three letters: the first letter indicates the quality measure (expectation, infimum, or supremum), the second letter designates the gap function (absolute or relative), and the third letter, D, indicates that the gap is measured over a discrete set. For all of the absolute and relative gap function formulations, it is important to note that as a consequence of Remark 2.3, the optimal objective value of DLPR($\widehat{\boldsymbol{\beta}}$) can be written solely in terms of the extreme points of the feasible region $\mathcal{Q}$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$.

Note that while the formulations presented in this section bear similarities to those presented in [1], there are a number of non-trivial differences, including the domain over which the formulations are defined, and the inclusion of the dual variables for both the IP and LP embedded in the MIP. Furthermore, unlike the formulations presented in [1], there are two right-hand sides for us to consider: the right-hand side corresponding to the MIP, $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$, and the portion of $\widehat{\boldsymbol{\beta}}$ allocated to the IP embedded in the MIP, $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$.

For each formulation in this section, let $\mathcal{D}$ be a finite, discrete subset of $\mathcal{B}[\mathbf{0}, \mathbf{b}]$. The expectation of the absolute gap function can be used to determine the expected performance of the LP relaxation as an approximation for the MIP, with a gap close to zero indicating a high-quality approximation for the MIP in expectation. The infimum can be used to determine the best-case performance, with a gap of zero indicating a perfect formulation for at least one right-hand side in $\mathcal{D}$. Finally, the supremum can be used to determine the worst-case performance, with a gap close to zero indicating a consistently high-quality approximation for the MIP.

4.1. Expectation of the Absolute Gap Function over a Discrete Set

Denote $\boldsymbol{\xi}$ a discrete random variable with event space $\mathcal{D}$. Let $\mathbb{P}\{\boldsymbol{\xi} = \widehat{\boldsymbol{\beta}}\} = \mu(\widehat{\boldsymbol{\beta}})$. The expectation of the absolute gap function over $\mathcal{D}$ is: $\mathbb{E}_{\boldsymbol{\xi}}[\boldsymbol{\Gamma}(\boldsymbol{\xi})] := \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{D}} \mu(\widehat{\boldsymbol{\beta}}) \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$. Consider the formulation:

$$\delta_{EAD} = \max \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{D}} \mu(\widehat{\boldsymbol{\beta}}) \boldsymbol{\psi}(\widehat{\boldsymbol{\beta}}) \tag{7a}$$

$$\text{s.t.}\quad \boldsymbol{\psi}(\widehat{\boldsymbol{\beta}}) \leq \widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q - (\boldsymbol{\phi}(\boldsymbol{\beta}) + \boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})) \quad \forall\, q \in \boldsymbol{\kappa},\ \boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{D}, \tag{7b}$$

$$\boldsymbol{\phi} \in \boldsymbol{\Phi}(\mathbf{b}), \tag{7c}$$

$$\boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}} \in \boldsymbol{\Omega} \quad \forall\, \boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{D}, \tag{7d}$$

$$\boldsymbol{\psi} \in \mathbb{R}_+^{|\mathcal{D}|}. \tag{7e}$$

Theorem 4.1. *The optimal objective value of (7) is* $\delta_{EAD} = \mathbb{E}_{\boldsymbol{\xi}}[\boldsymbol{\Gamma}(\boldsymbol{\xi})]$.

Proof. Let $\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{D}$, and let $\boldsymbol{\beta}^* \in \arg \max_{\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]} z_{IP}(\boldsymbol{\beta}) + z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$. For each $\widehat{\boldsymbol{\beta}} \in \mathcal{D}$, let $\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) = z_{IP}(\boldsymbol{\beta}^*)$ and $\tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) = z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)$ such that $\tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*} \in \boldsymbol{\Omega}$ (note that this is guaranteed to exist by Remark 2.1). We show that the triple $(\tilde{\boldsymbol{\phi}}, \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}, \tilde{\boldsymbol{\psi}})$ is feasible for (7). Note that by IP strong duality [1], $\tilde{\boldsymbol{\phi}}$ satisfies (6c). Furthermore, by strong duality, $z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) = z_{LPD}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)$. So, $z_{LPD}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) = \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)$ where $\tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*} \in \boldsymbol{\Omega}$. Thus, $\tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}$ satisfies (6d). In addition, by construction, $\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) + \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \geq \tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}) + \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})\ \forall \boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$, thus, satisfying (6b). By strong duality, $z_{LPR}(\widehat{\boldsymbol{\beta}}) = z_{DLPR}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{D}$. Thus,

$$\begin{aligned}\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}) = z_{LPR}(\widehat{\boldsymbol{\beta}}) - z_{MIP}(\widehat{\boldsymbol{\beta}}) &= \min_{q \in \boldsymbol{\kappa}} \widehat{\boldsymbol{\beta}}^\top u^q - (\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) + \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)) \\ &\leq \widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q - (\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) + \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)),\ \forall\, q \in \boldsymbol{\kappa}, \widehat{\boldsymbol{\beta}} \in \mathcal{D}.\end{aligned}$$

Therefore, the triple $(\tilde{\boldsymbol{\phi}}, \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}, \tilde{\boldsymbol{\psi}})$ is feasible for (7).

Suppose $(\boldsymbol{\phi}^*, \boldsymbol{\pi}^*_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}, \boldsymbol{\psi}^*)$ is feasible for (7). By Theorem 2.1, $\boldsymbol{\phi}^*(\boldsymbol{\beta}^*) + {\boldsymbol{\pi}^*}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \geq z_{MIP}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{D}$. By feasibility, $\boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}) \leq \widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q - (\boldsymbol{\phi}^*(\boldsymbol{\beta}) + {\boldsymbol{\pi}^*}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}))$ for all $q \in \boldsymbol{\kappa}$, $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$, and $\widehat{\boldsymbol{\beta}} \in \mathcal{D}$. It follows that $\boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}) \leq z_{LPR}(\widehat{\boldsymbol{\beta}}) - z_{MIP}(\widehat{\boldsymbol{\beta}}) = \tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}})$. Hence, $\sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{D}} \mu(\widehat{\boldsymbol{\beta}}) \boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}) \leq \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{D}} \mu(\widehat{\boldsymbol{\beta}}) \tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) = \mathbb{E}_{\boldsymbol{\xi}}[\boldsymbol{\Gamma}(\boldsymbol{\xi})]$, i.e., the optimal objective value of (7) is $\mathbb{E}_{\boldsymbol{\xi}}[\boldsymbol{\Gamma}(\boldsymbol{\xi})]$. □

4.2. Infimum of the Absolute Gap Function over a Discrete Set

Because, trivially, $\boldsymbol{\Gamma}(\mathbf{0}) = 0 = \min_{\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]} \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$, we exclude $\{\mathbf{0}\}$ from consideration. Denote $\mathcal{D}^+ = \mathcal{D} \setminus \{\mathbf{0}\}$. The infimum of the absolute gap function over $\mathcal{D}^+$ is: $\Delta_{IAD} := \inf_{\widehat{\boldsymbol{\beta}} \in \mathcal{D}^+} \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}) = \min_{\widehat{\boldsymbol{\beta}} \in \mathcal{D}^+} \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$. Consider the formulation:

$$\delta_{IAD} = \max\ \boldsymbol{\psi} \tag{8a}$$

$$\text{s.t.}\quad \boldsymbol{\psi} \leq \widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q - (\boldsymbol{\phi}(\boldsymbol{\beta}) + \boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})) \quad \forall\, q \in \boldsymbol{\kappa},\ \boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}] \setminus \{\mathbf{0}\},\ \widehat{\boldsymbol{\beta}} \in \mathcal{D}^+, \tag{8b}$$

$$\boldsymbol{\phi} \in \boldsymbol{\Phi}(\mathbf{b}), \tag{8c}$$

$$\boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}} \in \boldsymbol{\Omega} \quad \forall\, \boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}] \setminus \{\mathbf{0}\},\ \widehat{\boldsymbol{\beta}} \in \mathcal{D}^+, \tag{8d}$$

$$\boldsymbol{\psi} \in \mathbb{R}_+. \tag{8e}$$

Theorem 4.2. *The optimal objective value of (8) is Δ_{IAD}. That is, $\delta_{IAD} = \Delta_{IAD}$.*

4.3. Supremum of the Absolute Gap Function over a Discrete Set

Let $\text{SOS1}(\{\mathbf{w}(\widehat{\boldsymbol{\beta}})\}_{\widehat{\boldsymbol{\beta}}\in\mathcal{D}})$ denote a Special Ordered Set constraint of Type 1 on the decision variable $\mathbf{w} \in \mathbb{R}_+^{|\mathcal{D}|}$, so that $|\{\widehat{\boldsymbol{\beta}} \in \mathcal{D} \mid \mathbf{w}(\widehat{\boldsymbol{\beta}}) > 0\}| \leq 1$ [3]. The supremum of the absolute gap function over $\mathcal{D}$ is: $\Delta_{SAD} := \sup_{\widehat{\boldsymbol{\beta}}\in\mathcal{D}} \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}) = \max_{\widehat{\boldsymbol{\beta}}\in\mathcal{D}} \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$. Consider the formulation:

$$\delta_{SAD} = \max \sum_{\widehat{\boldsymbol{\beta}}\in\mathcal{D}} \boldsymbol{\psi}(\widehat{\boldsymbol{\beta}}) \tag{9a}$$

$$\text{s.t. } \boldsymbol{\psi}(\widehat{\boldsymbol{\beta}}) \leq \widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q - (\boldsymbol{\phi}(\boldsymbol{\beta}) + \boldsymbol{\pi}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})) \quad \forall\, q \in \boldsymbol{\kappa},\ \boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{D}, \tag{9b}$$

$$\text{SOS1}(\{\boldsymbol{\psi}(\widehat{\boldsymbol{\beta}})\}_{\widehat{\boldsymbol{\beta}}\in\mathcal{D}}), \tag{9c}$$

$$\boldsymbol{\phi} \in \boldsymbol{\Phi}(\mathbf{b}), \tag{9d}$$

$$\boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}} \in \boldsymbol{\Omega} \quad \forall\, \boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{D}, \tag{9e}$$

$$\boldsymbol{\psi} \in \mathbb{R}_+^{|\mathcal{D}|}. \tag{9f}$$

Theorem 4.3. *The optimal objective value of (9) is Δ_{SAD}. That is, $\delta_{SAD} = \Delta_{SAD}$.*

5. Relative Gap Functions over a Discrete Set

In this section, we optimize the expectation, infimum, and supremum of the relative gap function, $\boldsymbol{\gamma}$, over finite discrete sets. As with Section 4, the formulations presented in this section bear similarities to those presented in [1]. However, there are a number of non-trivial differences, including: the domain over which the formulations are defined, the inclusion of the dual variables for both the IP and LP embedded in the MIP, and the consideration of the right-hand side corresponding to the MIP, $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}]$, as well as the portion of $\widehat{\boldsymbol{\beta}}$ allocated to the IP embedded in the MIP, $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$.

Recall $\mathcal{B}^+[\mathbf{0}, \mathbf{b}] = \{\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0}, \mathbf{b}] \mid z_{MIP}(\widehat{\boldsymbol{\beta}}) > 0\}$. We maintain the same notation from Section 4, along with the following sets: let $\mathcal{S}^+$ be a finite subset of $\mathcal{B}^+[\mathbf{0}, \mathbf{b}]$, and let $\widehat{\mathcal{B}}^+[\mathbf{0}, \mathbf{b}] = \mathcal{B}^+[\mathbf{0}, \mathbf{b}] \cap \mathbb{Z}_+^m$. The expectation of the relative gap function can be used to determine the expected performance of the LP relaxation as an approximation for the MIP, with a gap close to 1 indicating a high-quality approximation for the MIP in expectation. The infimum can be used to determine the worst-case performance, with a gap close to 1 indicating a consistently high-quality approximation for the MIP.

The supremum can be used to determine the best-case performance, with a gap of 1 indicating a perfect formulation for at least one right-hand side in $\mathcal{S}^+$.

5.1. Expectation of the Relative Gap Function over a Discrete Set

Denote $\boldsymbol{\xi}$ a discrete random variable with event space $\mathcal{S}^+$. Note that $\mathbb{P}\{\boldsymbol{\xi} = \widehat{\boldsymbol{\beta}}\} = \mu(\widehat{\boldsymbol{\beta}})$. The expectation of the relative gap function over $\mathcal{S}^+$ is: $\mathbb{E}_{\boldsymbol{\xi}}[\boldsymbol{\gamma}(\boldsymbol{\xi})] := \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+} \mu(\widehat{\boldsymbol{\beta}})\boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}})$. Consider the formulation:

$$\delta_{ERD} = \min \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+} \mu(\widehat{\boldsymbol{\beta}})\boldsymbol{\psi}(\widehat{\boldsymbol{\beta}}) \tag{10a}$$

$$\text{s.t.} \quad \boldsymbol{\psi}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q \geq \boldsymbol{\phi}(\boldsymbol{\beta}) + \boldsymbol{\pi}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \quad \forall\, q \in \boldsymbol{\kappa},\ \boldsymbol{\beta} \in \widehat{\mathcal{B}}^+[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{S}^+, \tag{10b}$$

$$\boldsymbol{\phi} \in \boldsymbol{\Phi}(\mathbf{b}), \tag{10c}$$

$$\boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}} \in \boldsymbol{\Omega} \quad \forall\, \boldsymbol{\beta} \in \widehat{\mathcal{B}}^+[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{S}^+, \tag{10d}$$

$$\boldsymbol{\psi} \in \mathbb{R}_+^{|\mathcal{S}^+|}. \tag{10e}$$

Theorem 5.1. *The optimal objective value of (10) is* $\delta_{ERD} = \mathbb{E}_{\boldsymbol{\xi}}[\boldsymbol{\gamma}(\boldsymbol{\xi})]$.

Proof. Let $\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+$, and let $\boldsymbol{\beta}^* \in \arg\max_{\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]} z_{IP}(\boldsymbol{\beta}) + z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$. For each $\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+$, let $\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) = z_{IP}(\boldsymbol{\beta}^*)$ and $\tilde{\boldsymbol{\pi}}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) = z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)$ such that $\tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*} \in \boldsymbol{\Omega}$. By arguments similar to those in the proof of Theorem 4.1, the triple satisfies (6b)-(6d).

By strong duality, $z_{LPR}(\widehat{\boldsymbol{\beta}}) = z_{DLPR}(\widehat{\boldsymbol{\beta}}) = \min_{q \in \boldsymbol{\kappa}} \widehat{\boldsymbol{\beta}}^\top u^q$. In addition, $z_{LPR}(\widehat{\boldsymbol{\beta}}) > 0$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+$. Thus, for all $q \in \boldsymbol{\kappa}$, $\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+$, and $\boldsymbol{\beta} \in \widehat{\mathcal{B}}^+[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$:

$$\begin{aligned} \tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q = \boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q \geq \boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}}) z_{LPR}(\widehat{\boldsymbol{\beta}}) &= z_{MIP}(\widehat{\boldsymbol{\beta}}) = z_{IP}(\boldsymbol{\beta}^*) + z_{LP}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \\ &= \tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) + \tilde{\boldsymbol{\pi}}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \geq \tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}) + \tilde{\boldsymbol{\pi}}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}). \end{aligned}$$

Hence, the triple $(\tilde{\boldsymbol{\phi}}, \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}, \tilde{\boldsymbol{\psi}})$ is feasible for (10).

Suppose $(\boldsymbol{\phi}^*, \boldsymbol{\pi}^*_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}, \boldsymbol{\psi}^*)$ is feasible for (10). By Theorem 2.1, $\boldsymbol{\phi}^*(\boldsymbol{\beta}^*) + \boldsymbol{\pi}^{*\top}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \geq z_{MIP}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+$. By feasibility, $\boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}) \geq \frac{\boldsymbol{\phi}^*(\boldsymbol{\beta}) + \boldsymbol{\pi}^{*\top}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{\widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q}$ $\forall q \in \boldsymbol{\kappa}$, $\boldsymbol{\beta} \in \widehat{\mathcal{B}}^+[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$, and $\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+$. Hence, $\boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}) \geq \frac{\boldsymbol{\phi}^*(\boldsymbol{\beta}^*) + \boldsymbol{\pi}^{*\top}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*)}{z_{LPR}(\widehat{\boldsymbol{\beta}})} \geq \frac{z_{MIP}(\widehat{\boldsymbol{\beta}})}{z_{LPR}(\widehat{\boldsymbol{\beta}})} = \boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}}) = \tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}})$. Thus, $\mathbb{E}_{\boldsymbol{\xi}}[\boldsymbol{\Gamma}(\boldsymbol{\xi})] = \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+} \mu(\widehat{\boldsymbol{\beta}})\boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}}) = \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+} \mu(\widehat{\boldsymbol{\beta}})\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) \leq \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+} \mu(\widehat{\boldsymbol{\beta}})\boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}})$, i.e., $\delta_{ERD} = \mathbb{E}_{\boldsymbol{\xi}}[\boldsymbol{\Gamma}(\boldsymbol{\xi})]$. □

5.2. Infimum of the Relative Gap Function over a Discrete Set

The infimum of the relative gap function over $\mathcal{S}^+$ is: $\Delta_{IRD} := \min_{\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+} \boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}})$. Consider the formulation:

$$\delta_{IRD} = \max \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+} \boldsymbol{\psi}(\widehat{\boldsymbol{\beta}}) \tag{11a}$$

$$\text{s.t. } (1 - \boldsymbol{\psi}(\widehat{\boldsymbol{\beta}}))\widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q \geq \boldsymbol{\phi}(\boldsymbol{\beta}) + \boldsymbol{\pi}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \quad \forall\, q \in \boldsymbol{\kappa},\ \boldsymbol{\beta} \in \widehat{\mathcal{B}}^+[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{S}^+, \tag{11b}$$

$$\text{SOS1}(\{\boldsymbol{\psi}(\widehat{\boldsymbol{\beta}})\}_{\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+}), \tag{11c}$$

$$\boldsymbol{\phi} \in \boldsymbol{\Phi}(\mathbf{b}), \tag{11d}$$

$$\boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}} \in \boldsymbol{\Omega} \quad \forall\, \boldsymbol{\beta} \in \widehat{\mathcal{B}}^+[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{S}^+, \tag{11e}$$

$$\boldsymbol{\psi} \in \mathbb{R}_+^{|\mathcal{S}^+|}. \tag{11f}$$

Theorem 5.2. *The optimal objective value of (11) is* $1 - \Delta_{IRD}$. *That is,* $\delta_{IRD} = 1 - \Delta_{IRD}$.

5.3. Supremum of the Relative Gap Function over a Discrete Set

The supremum of the relative gap function over $\mathcal{S}^+$ is: $\Delta_{SRD} := \max_{\widehat{\boldsymbol{\beta}} \in \mathcal{S}^+} \boldsymbol{\gamma}(\widehat{\boldsymbol{\beta}})$. Consider the formulation:

$$\delta_{SRD} = \min \boldsymbol{\psi} \tag{12a}$$

$$\text{s.t. } \boldsymbol{\psi} \cdot \widehat{\boldsymbol{\beta}}^\top \mathbf{u}^q \geq \boldsymbol{\phi}(\boldsymbol{\beta}) + \boldsymbol{\pi}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \quad \forall\, q \in \boldsymbol{\kappa},\ \boldsymbol{\beta} \in \widehat{\mathcal{B}}^+[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{S}^+, \tag{12b}$$

$$\boldsymbol{\phi} \in \boldsymbol{\Phi}(\mathbf{b}), \tag{12c}$$

$$\boldsymbol{\pi}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}} \in \boldsymbol{\Omega} \quad \forall\, \boldsymbol{\beta} \in \widehat{\mathcal{B}}^+[\mathbf{0}, \widehat{\boldsymbol{\beta}}],\ \widehat{\boldsymbol{\beta}} \in \mathcal{S}^+, \tag{12d}$$

$$\boldsymbol{\psi} \in \mathbb{R}_+. \tag{12e}$$

Theorem 5.3. *The optimal objective value of of (12) is* Δ_{SRD}. *That is,* $\delta_{SRD} = \Delta_{SRD}$.

Acknowledgments

The authors thank the review team, as well as Seth Brown and Dr. Mustafa Can Camur of Rice University, for their helpful comments. This research was supported by National Science Foundation grant CMMI-1933373.

References

[1] T. Ajayi, C. Thomas, and A. J. Schaefer. The gap function: Evaluating integer programming models over multiple right-hand sides. *Oper. Res.*, 70(2):1259–1270, 2022.

[2] M. Baes, T. Oertel, and R. Weismantel. Duality for mixed-integer convex minimization. *Math. Program.*, 158(1):547–564, 2016.

[3] E. M. L. Beale and J. A. Tomlin. Special facilities in a general mathematical programming system for non-convex problems using ordered sets of variables. *Proc. Fifth Int. Conf. Oper. Res.* (J. Laurence, Ed.) Tavistock Publications, London, pages 447–454, 1970.

[4] C. E. Blair. Extensions of subadditive functions used in cutting-plane theory. Technical report, MSRR, No. 360, Carnegie Mellon University, Pittsburgh, PA, 1974.

[5] C. E. Blair. A closed-form representation of mixed-integer program value functions. *Math. Program.*, 71(2):127–136, 1995.

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

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

[8] C. E. Blair and R. G. Jeroslow. The value function of an integer program. *Math. Program.*, 23(1):237–273, 1982.

[9] C. E. Blair and R. G. Jeroslow. Constructive characterizations of the value-function of a mixed-integer program I. *Discret. Appl. Math.*, 9(3):217–233, 1984.

[10] A. Charnes and W. W. Cooper. The strong Minkowski-Farkas-Weyl theorem for vector spaces over ordered fields. *PNAS*, 44(9):914–916, 1958.

[11] W. Cook, A. M. H. Gerards, A. Schrijver, and É. Tardos. Sensitivity theorems in integer linear programming. *Math. Program.*, 34(3):251–264, 1986.

[12] R. E. Gomory. An algorithm for the mixed integer problem. Technical report, RM-2597, RAND Corp Santa Monica CA, 1960.

[13] R. E. Gomory. On the relation between integer and noninteger solutions to linear programs. *PNAS*, 53(2):260–265, 1965.

[14] M. Güzelsoy. *Dual methods in mixed integer linear programming.* PhD Thesis, Lehigh University, Bethlehem, PA, USA, 2009.

[15] M. Güzelsoy and T. K. Ralphs. Duality for mixed-integer linear programs. *IJOR*, 4(3):118–137, 2007.

[16] M. Güzelsoy and T. K. Ralphs. Integer programming duality. In *Encyclopedia of Oper. Res. Management Sci.*, pages 1–13. J. Cochran, Ed. Hoboken, NJ, USA: Wiley, 2010.

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

[18] E. L. Johnson. On the group problem for mixed integer programming. *Math. Prog. Study*, 2: 137–179, 1974.

[19] B. Kocuk and D. A. Morán R. On subadditive duality for conic mixed-integer programs. *SIAM J. Optim.*, 29(3):2320–2336, 2019.

[20] N. Kong, A. J. Schaefer, and B. Hunsaker. Two-stage integer programs with stochastic right-hand sides: A superadditive dual approach. *Math. Program.*, 108(2):275–296, 2006.

[21] J. B. Lamperski and A. J. Schaefer. A polyhedral characterization of the inverse-feasible region of a mixed-integer program. *Oper. Res. Lett.*, 43(6):575–578, 2015.

[22] P. Le Bodic and G. L. Nemhauser. An abstract model for branching and its application to mixed integer programming. *Math. Program.*, 166(1):369–405, 2017.

[23] D. C. Llewellyn and J. Ryan. A primal dual integer programming algorithm. *Discret. Appl. Math.*, 45(3):261–275, 1993.

[24] D. A. Morán R, S. S. Dey, and J. P. Vielma. A strong dual for conic mixed-integer programs. *SIAM J. Optim.*, 22(3):1136–1150, 2012.

[25] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization.* John Wiley & Sons, 1999.

[26] O. Y. Özaltın, O. A. Prokopyev, and A. J. Schaefer. Two-stage quadratic integer programs with stochastic right-hand sides. *Math. Program.*, 133(1):121–158, 2012.

[27] J. Paat, R. Weismantel, and S. Weltge. Distances between optimal solutions of mixed-integer programs. *Math. Program.*, 179(1):455–468, 2020.

[28] T. K. Ralphs and A. Hassanzadeh. On the value function of a mixed integer linear optimization problem and an algorithm for its construction. Technical report, COR@L, 14T–004, 2014.

[29] S. Sen. Algorithms for stochastic mixed-integer programming models. *Handbooks Oper. Res. Management Sci.*, 12:515–558, 2005.

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

[31] L. A. Wolsey. Integer programming duality: Price functions and sensitivity analysis. *Math. Program.*, 20(1):173–195, 1981.

E.C. Electronic Companion

E.C.1. Section 3 Results

Proposition 3.1. *The absolute MIP gap function defined over $\mathcal{B}[\mathbf{0},\mathbf{b}]$ is the minimum of finitely many Gomory functions.*

Proof. Let $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0},\mathbf{b}]$. Consider the negative of the value function:

$$-z_{MIP}(\widehat{\boldsymbol{\beta}}) = \min_{\mathbf{x}\in\mathbb{Z}^n_+,\ \mathbf{y}\in\mathbb{R}^p_+} \{-\mathbf{c}^\top\mathbf{x} - \mathbf{h}^\top\mathbf{y} \mid \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{y} \le \widehat{\boldsymbol{\beta}}\}.$$

[9] prove that for any $\widehat{\boldsymbol{\beta}}$ such that MIP($\widehat{\boldsymbol{\beta}}$) is feasible, which includes $\mathcal{B}[\mathbf{0},\mathbf{b}]$, $-z_{MIP}$ is the minimum of finitely many Gomory functions. So, let $-z_{MIP}(\widehat{\boldsymbol{\beta}}) = \min\{G_1(\widehat{\boldsymbol{\beta}}),\ldots,G_L(\widehat{\boldsymbol{\beta}})\}$, where $\{G_i(\widehat{\boldsymbol{\beta}}) \mid i = 1,\ldots,L\}$ are Gomory functions. Recall that $z_{DLPR}(\widehat{\boldsymbol{\beta}}) = \min_{q\in\boldsymbol{\kappa}} \widehat{\boldsymbol{\beta}}^\top\mathbf{u}^q$. By strong duality, $z_{LPR}(\widehat{\boldsymbol{\beta}}) = z_{DLPR}(\widehat{\boldsymbol{\beta}})$. Thus, $z_{LPR}(\widehat{\boldsymbol{\beta}}) = \min_{q\in\boldsymbol{\kappa}} \widehat{\boldsymbol{\beta}}^\top\mathbf{u}^q$, where $|\boldsymbol{\kappa}| < +\infty$. Let $q(\widehat{\boldsymbol{\beta}})^* \in \boldsymbol{\kappa}$ be such that $q(\widehat{\boldsymbol{\beta}})^* \in \arg\min_{q\in\boldsymbol{\kappa}} \widehat{\boldsymbol{\beta}}^\top\mathbf{u}^q$. Recall that as a consequence of Assumption 2.1, Remark 2.2, Weyl's Theorem [10], and the assumption that $\mathbf{c},\mathbf{h} > \mathbf{0}$, $\widehat{\boldsymbol{\beta}}^\top\mathbf{u}^{q(\widehat{\boldsymbol{\beta}})^*} = z_{DLPR}(\widehat{\boldsymbol{\beta}}) \ge 0$. Thus, for $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0},\mathbf{b}]$: $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}) = z_{LPR}(\widehat{\boldsymbol{\beta}}) - z_{MIP}(\widehat{\boldsymbol{\beta}}) = \widehat{\boldsymbol{\beta}}^\top\mathbf{u}^{q(\widehat{\boldsymbol{\beta}})^*} + \min_{i=1,\ldots,L} G_i(\widehat{\boldsymbol{\beta}}) = \min_{i=1,\ldots,L}(\widehat{\boldsymbol{\beta}}^\top\mathbf{u}^{q(\widehat{\boldsymbol{\beta}})^*} + G_i(\widehat{\boldsymbol{\beta}}))$. Notice for each i, $\widehat{\boldsymbol{\beta}}^\top\mathbf{u}^{q(\widehat{\boldsymbol{\beta}})^*} + G_i(\widehat{\boldsymbol{\beta}})$ is a Gomory function, as it is the sum of two Gomory functions. Thus, $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$ is the minimum of finitely many Gomory functions for all $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0},\mathbf{b}]$. □

Proposition 3.2. *Let $\widehat{\boldsymbol{\beta}} \in \mathcal{B}[\mathbf{0},\mathbf{b}]$ and $(\mathbf{x}^*,\mathbf{y}^*) \in opt_{MIP}(\widehat{\boldsymbol{\beta}})$. Given $\eta \in \mathbb{N}$, let $J_\eta \subseteq \{1,\ldots,n\}$ denote the set of indices such that $x^*_j \ge \eta$ for $j \in J_\eta$. Denote $K := \{1,\ldots,p\}$, and let $\lambda^* = \min_{k\in K}\{y^*_k\}$. Then, for any $j \in J_\eta$, $k \in K$, and $\lambda \in [0,\lambda^*]$, $z_{MIP}(\widehat{\boldsymbol{\beta}} - \eta\mathbf{a}_j - \lambda\mathbf{g}_k) = z_{MIP}(\widehat{\boldsymbol{\beta}}) - \eta c_j - \lambda h_k$.*

Proof. Denote $\mathbf{e}_j$ the j^{th} unit vector in $\mathbb{R}^n_+$ and $\boldsymbol{\epsilon}_k$ the k^{th} unit vector in $\mathbb{R}^p_+$. Let $(\mathbf{x}^*,\mathbf{y}^*) \in \text{opt}_{MIP}(\widehat{\boldsymbol{\beta}})$, so $z_{MIP}(\widehat{\boldsymbol{\beta}}) = \mathbf{c}^\top\mathbf{x}^* + \mathbf{h}^\top\mathbf{y}^*$. Let $\lambda \in [0,\lambda^*]$ where $\lambda^* = \min_{k\in K}\{y^*_k\}$. Suppose $J_\eta \ne \emptyset$, and choose $\eta \in \mathbb{N}$ is such that $x^*_j \ge \eta$ for all $j \in J_\eta$. Note that $\mathbf{A}(\mathbf{x}^* - \eta\mathbf{e}_j) + \mathbf{G}(\mathbf{y}^* - \lambda\boldsymbol{\epsilon}_k) = \mathbf{A}\mathbf{x}^* + \mathbf{G}\mathbf{y}^* - \eta\mathbf{a}_j - \lambda\mathbf{g}_k \le \widehat{\boldsymbol{\beta}} - \eta\mathbf{a}_j - \lambda\mathbf{g}_k \le \widehat{\boldsymbol{\beta}}$, with $\mathbf{x}^* - \eta\mathbf{e}_j \in \mathbb{Z}^n_+$ and $\mathbf{y}^* - \lambda\boldsymbol{\epsilon}_k \in \mathbb{R}^p_+$. Therefore, $(\mathbf{x}^* - \eta\mathbf{e}_j, \mathbf{y}^* - \lambda\boldsymbol{\epsilon}_k)$ is feasible for MIP($\widehat{\boldsymbol{\beta}} - \eta\mathbf{a}_j - \lambda\mathbf{g}_k$) and MIP($\widehat{\boldsymbol{\beta}}$).

Now, suppose for the sake of contradiction that $(\mathbf{x}^* - \eta\mathbf{e}_j, \mathbf{y}^* - \lambda\boldsymbol{\epsilon}_k) \notin \text{opt}_{MIP}(\widehat{\boldsymbol{\beta}} - \eta\mathbf{a}_j - \lambda\mathbf{g}_k)$. Note that $\text{opt}_{MIP}(\widehat{\boldsymbol{\beta}} - \eta\mathbf{a}_j - \lambda\mathbf{g}_k) \ne \emptyset$ due to Assumption 2.1 and the data being nonnegative. Thus, let $(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) \in \text{opt}_{MIP}(\widehat{\boldsymbol{\beta}} - \eta\mathbf{a}_j - \lambda\mathbf{g}_k)$. Then, $\mathbf{A}\tilde{\mathbf{x}} + \mathbf{G}\tilde{\mathbf{y}} \le \widehat{\boldsymbol{\beta}} - \eta\mathbf{a}_j - \lambda\mathbf{g}_k \le \widehat{\boldsymbol{\beta}}$, with $\tilde{\mathbf{x}} \in \mathbb{Z}^n_+$ and $\tilde{\mathbf{y}} \in \mathbb{R}^p_+$, so $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is feasible for MIP($\widehat{\boldsymbol{\beta}}$). Also, $\mathbf{c}^\top\tilde{\mathbf{x}} + \mathbf{h}^\top\tilde{\mathbf{y}} > \mathbf{c}^\top(\mathbf{x}^* - \eta\mathbf{e}_j) + \mathbf{h}^\top(\mathbf{y}^* - \lambda\boldsymbol{\epsilon}_k)$. Now, let $\widehat{\mathbf{x}} = \tilde{\mathbf{x}} + \eta\mathbf{e}_j$ and $\widehat{\mathbf{y}} = \tilde{\mathbf{y}} + \lambda\boldsymbol{\epsilon}_k$. Note that $\mathbf{A}\widehat{\mathbf{x}} + \mathbf{G}\widehat{\mathbf{y}} = \mathbf{A}\tilde{\mathbf{x}} + \mathbf{G}\tilde{\mathbf{y}} + \eta\mathbf{a}_j + \lambda\mathbf{g}_k \le \widehat{\boldsymbol{\beta}} - \eta\mathbf{a}_j - \lambda\mathbf{g}_k + \eta\mathbf{a}_j + \lambda\mathbf{g}_k = \widehat{\boldsymbol{\beta}}$, with $\widehat{\mathbf{x}} \in \mathbb{Z}^n_+$ and $\widehat{\mathbf{y}} \in \mathbb{R}^p_+$. Thus, $(\widehat{\mathbf{x}},\widehat{\mathbf{y}})$ is feasible for MIP($\widehat{\boldsymbol{\beta}}$). Moreover,

$\mathbf{c}^\top\widehat{\mathbf{x}}+\mathbf{h}^\top\widehat{\mathbf{y}} = \mathbf{c}^\top\tilde{\mathbf{x}}+\mathbf{h}^\top\tilde{\mathbf{y}}+\eta c_j+\lambda h_k > \mathbf{c}^\top(\mathbf{x}^*-\eta\mathbf{e}_j)+\mathbf{h}^\top(\mathbf{y}^*-\lambda\boldsymbol{\epsilon}_k)+\eta c_j+\lambda h_k = \mathbf{c}^\top\mathbf{x}^*+\mathbf{h}^\top\mathbf{y}^*$, which contradicts the fact that $(\mathbf{x}^*,\mathbf{y}^*) \in \text{opt}_{MIP}(\widehat{\boldsymbol{\beta}})$. Thus, $(\mathbf{x}^*-\eta\mathbf{e}_j,\mathbf{y}^*-\lambda\boldsymbol{\epsilon}_k) \in \text{opt}_{MIP}(\widehat{\boldsymbol{\beta}}-\eta\mathbf{a}_j-\lambda\mathbf{g}_k)$. Therefore, $z_{MIP}(\widehat{\boldsymbol{\beta}}-\eta\mathbf{a}_j-\lambda\mathbf{g}_k) = \mathbf{c}^\top(\mathbf{x}^*-\eta\mathbf{e}_j)+\mathbf{h}^\top(\mathbf{y}^*-\lambda\boldsymbol{\epsilon}_k) = \mathbf{c}^\top\mathbf{x}^*+\mathbf{h}^\top\mathbf{y}^*-\eta c_j-\lambda h_k = z_{MIP}(\widehat{\boldsymbol{\beta}})-\eta c_j-\lambda h_k$. □

E.C.2. Section 4 Results

Theorem 4.2. *The optimal objective value of (8) is Δ_{IAD}. That is, $\delta_{IAD} = \Delta_{IAD}$.*

Proof. Let $\tilde{\boldsymbol{\psi}} = \Delta_{IAD}$, and let $\boldsymbol{\beta}^* \in \arg\max_{\boldsymbol{\beta}\in\widehat{\mathcal{B}}[\mathbf{0},\widehat{\boldsymbol{\beta}}]} z_{IP}(\boldsymbol{\beta}) + z_{LP}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})$. For each $\widehat{\boldsymbol{\beta}} \in \mathcal{D}^+$, let $\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) = z_{IP}(\boldsymbol{\beta}^*)$ and $\tilde{\boldsymbol{\pi}}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*) = z_{LP}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*)$ such that $\tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*} \in \boldsymbol{\Omega}$ (note that this is guaranteed to exist by Remark 2.1). We show that the triple $(\tilde{\boldsymbol{\phi}},\tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*},\tilde{\boldsymbol{\psi}})$ is feasible for (8). By arguments similar to those in the proof of Theorem 4.1, the triple satisfies (6b)-(6d), and $\tilde{\boldsymbol{\psi}} = \Delta_{IAD} = \min_{\widehat{\boldsymbol{\beta}}\in\mathcal{D}^+}\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}) \le z_{LPR}(\widehat{\boldsymbol{\beta}}) - z_{MIP}(\widehat{\boldsymbol{\beta}}) \le \widehat{\boldsymbol{\beta}}^\top\mathbf{u}^q - (\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) + \tilde{\boldsymbol{\pi}}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*))$ for all $q \in \boldsymbol{\kappa}$ and $\widehat{\boldsymbol{\beta}} \in \mathcal{D}^+$. Hence, the triple $(\tilde{\boldsymbol{\phi}},\tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*},\tilde{\boldsymbol{\psi}})$ is feasible for (8).

Now, suppose $(\boldsymbol{\phi}^*,\boldsymbol{\pi}^*_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}},\boldsymbol{\psi}^*)$ is feasible for (8). By Theorem 2.1, $\boldsymbol{\phi}^*(\boldsymbol{\beta}^*) + {\boldsymbol{\pi}^*}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*) \ge z_{MIP}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{D}^+$. Furthermore, by feasibility, $\boldsymbol{\psi}^* \le \widehat{\boldsymbol{\beta}}^\top\mathbf{u}^q - (\boldsymbol{\phi}^*(\boldsymbol{\beta}) + {\boldsymbol{\pi}^*}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}))$, for all $q \in \boldsymbol{\kappa}$, $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0},\widehat{\boldsymbol{\beta}}] \setminus \{\mathbf{0}\}$, and $\widehat{\boldsymbol{\beta}} \in \mathcal{D}^+$. It follows that $\boldsymbol{\psi}^* \le \Delta_{IAD} = \tilde{\boldsymbol{\psi}}$. Thus, $\tilde{\boldsymbol{\psi}} = \delta_{IAD}$, and the optimal objective value of (8) is Δ_{IAD}. □

Theorem 4.3. *The optimal objective value of (9) is Δ_{SAD}. That is, $\delta_{SAD} = \Delta_{SAD}$.*

Proof. Let $\widehat{\boldsymbol{\beta}}_{max} \in \arg\max_{\widehat{\boldsymbol{\beta}}\in\mathcal{D}}\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})\}$. Let $\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}_{max}) = \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}_{max})$ and $\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) = 0$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{D} \setminus \widehat{\boldsymbol{\beta}}_{max}$. Note that by construction, $\tilde{\boldsymbol{\psi}}$ satisfies (9c). Let $\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{D}$, and let $\boldsymbol{\beta}^* \in \arg\max_{\boldsymbol{\beta}\in\widehat{\mathcal{B}}[\mathbf{0},\widehat{\boldsymbol{\beta}}]} z_{IP}(\boldsymbol{\beta}) + z_{LP}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})$. For each $\widehat{\boldsymbol{\beta}} \in \mathcal{D}$, let $\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) = z_{IP}(\boldsymbol{\beta}^*)$ and $\tilde{\boldsymbol{\pi}}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*) = z_{LP}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*)$ such that $\tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*} \in \boldsymbol{\Omega}$ (note that this is guaranteed to exist by Remark 2.1). By arguments similar to those in the proof of Theorem 4.1, the triple satisfies (6b)-(6d).

Now, note that $\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) = 0 \le \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{D} \setminus \widehat{\boldsymbol{\beta}}_{max}$. Therefore for $\widehat{\boldsymbol{\beta}} \in \mathcal{D} \setminus \widehat{\boldsymbol{\beta}}_{max}$,

$$\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) \le \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}) = z_{LPR}(\widehat{\boldsymbol{\beta}}) - z_{MIP}(\widehat{\boldsymbol{\beta}}) \le \widehat{\boldsymbol{\beta}}^\top\mathbf{u}^q - (\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) + \tilde{\boldsymbol{\pi}}^\top_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*)),\ \forall\ q \in \boldsymbol{\kappa}, \widehat{\boldsymbol{\beta}} \in \mathcal{D}^+.$$

Recall that by strong duality, $z_{LPR}(\widehat{\boldsymbol{\beta}}_{max}) = z_{DLPR}(\widehat{\boldsymbol{\beta}}_{max})$. Thus, for $\widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}}_{max}$,

$$\begin{aligned}\tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}_{max}) = \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}_{max}) &= z_{LPR}(\widehat{\boldsymbol{\beta}}_{max}) - z_{MIP}(\widehat{\boldsymbol{\beta}}_{max}) \\ &= \min_{q\in\boldsymbol{\kappa}} \widehat{\boldsymbol{\beta}}^\top_{max}u^q - (\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) + \tilde{\boldsymbol{\pi}}^\top_{\widehat{\boldsymbol{\beta}}_{max}-\boldsymbol{\beta}^*}(\widehat{\boldsymbol{\beta}}_{max}-\boldsymbol{\beta}^*))\end{aligned}$$

$$\leq \widehat{\boldsymbol{\beta}}_{max}^{\top}\mathbf{u}^q - (\tilde{\boldsymbol{\phi}}(\boldsymbol{\beta}^*) + \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}_{max}-\boldsymbol{\beta}^*}^{\top}(\widehat{\boldsymbol{\beta}}_{max} - \boldsymbol{\beta}^*)),\ \forall\ q \in \boldsymbol{\kappa}.$$

Hence, the triple $(\tilde{\boldsymbol{\phi}}, \tilde{\boldsymbol{\pi}}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}, \tilde{\boldsymbol{\psi}})$ is feasible for (9).

Now, let $(\boldsymbol{\phi}^*, \boldsymbol{\pi}^*_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}, \boldsymbol{\psi}^*)$ be feasible for (9) such that there exists some $\widehat{\boldsymbol{\beta}}^* \in \mathcal{D}$ for which $\boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}) = 0$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{D} \setminus \widehat{\boldsymbol{\beta}}^*$. By Theorem 2.1, $\boldsymbol{\phi}^*(\boldsymbol{\beta}^*) + {\boldsymbol{\pi}^*}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^*}^{\top}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*) \geq z_{MIP}(\widehat{\boldsymbol{\beta}})$ for all $\widehat{\boldsymbol{\beta}} \in \mathcal{D}$. Furthermore, by feasibility, $\boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}) \leq \widehat{\boldsymbol{\beta}}^{\top}\mathbf{u}^q - (\boldsymbol{\phi}^*(\boldsymbol{\beta}) + {\boldsymbol{\pi}^*}_{\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}}^{\top}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}))$ for all $q \in \boldsymbol{\kappa}$, $\boldsymbol{\beta} \in \widehat{\mathcal{B}}[\mathbf{0}, \widehat{\boldsymbol{\beta}}]$, and $\widehat{\boldsymbol{\beta}} \in \mathcal{D}$. It follows that $\boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}^*) \leq z_{LPR}(\widehat{\boldsymbol{\beta}}^*) - z_{MIP}(\widehat{\boldsymbol{\beta}}^*) = \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}^*)$. Now, recall that $\widehat{\boldsymbol{\beta}}_{max} \in \arg\max_{\widehat{\boldsymbol{\beta}} \in \mathcal{D}}\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}})\}$. Then, $\boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}^*) \leq \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}^*) \leq \boldsymbol{\Gamma}(\widehat{\boldsymbol{\beta}}_{max}) = \tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}_{max})$. Thus, $\sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{D}} \boldsymbol{\psi}^*(\widehat{\boldsymbol{\beta}}) \leq \sum_{\widehat{\boldsymbol{\beta}} \in \mathcal{D}} \tilde{\boldsymbol{\psi}}(\widehat{\boldsymbol{\beta}}) = \Delta_{SAD}$, i.e., the optimal objective value of (9) is Δ_{SAD}. □

E.C.3. Section 5 Results

Theorem 5.2. *The optimal objective value of (11) is Δ_{IRD}. That is, $\delta_{IRD} = \Delta_{IRD}$.*

Proof. The proof follows similarly from Theorem 4.3 and is therefore omitted. □

Theorem 5.3. *The optimal objective value of (12) is Δ_{SRD}. That is, $\delta_{SRD} = \Delta_{SRD}$.*

Proof. The proof follows similarly from Theorem 4.2 and is therefore omitted. □