

A Further details about plots

The version of the *facility location* problem we study involves a set of locations J and a set of clients C . Facilities are to be constructed at some subset of the locations, and the clients in C are served by these facilities. Each location $j \in J$ has a cost f_j of being the site of a facility, and a cost $s_{c,j}$ of serving client $c \in C$. Finally, each location j has a capacity κ_j which is a limit on the number of clients j can serve. The goal of the facility location problem is to arrive at a feasible set of locations for facilities and a feasible assignment of clients to these locations that minimizes the overall cost incurred.

The facility location problem can be formulated as the following 0, 1 IP:

$$\begin{aligned}
 & \text{minimize} && \sum_{j \in J} f_j x_j + \sum_{j \in J} \sum_{c \in C} s_{c,j} y_{c,j} \\
 & \text{subject to} && \sum_{j \in J} y_{c,j} = 1 && \forall c \in C \\
 & && \sum_{c \in C} y_{c,j} \leq \kappa_j x_j && \forall j \in J \\
 & && y_{c,j} \in \{0, 1\} && \forall c \in C, j \in J \\
 & && x_j \in \{0, 1\} && \forall j \in J
 \end{aligned}$$

We consider the following two distributions over facility location IPs.

First distribution Facility location IPs are generated by perturbing the costs and capacities of a base facility location IP. We generated the base IP with 40 locations and 40 clients by choosing the location costs and client-location costs uniformly at random from $[0, 100]$ and the capacities uniformly at random from $\{0, \dots, 39\}$. To sample from the distribution, we perturb this base IP by adding independent Gaussian noise with mean 0 and standard deviation 10 to the cost of each location, the cost of each client-location pair, and the capacity of each location.

Second distribution Facility location IPs are generated by placing 80 evenly-spaced locations along the line segment connecting the points $(0, 1/2)$ and $(1, 1/2)$ in the Cartesian plane. The location costs are all uniformly set to 1. Then, 80 clients are placed uniformly at random in the unit square $[0, 1]^2$. The cost $s_{c,j}$ of serving client c from location j is the distance between j and c . Location capacities are chosen uniformly at random from $\{0, \dots, 43\}$.

In our experiments, we add five cuts at the root of the B&C tree. These five cuts come from the set of Chvátal-Gomory and Gomory mixed integer cuts derived from the optimal simplex tableau of the LP relaxation. The five cuts added are chosen to maximize a weighting of cutting-plane scores:

$$\mu \cdot \text{score}_1 + (1 - \mu) \cdot \text{score}_2. \quad (4)$$

score_1 is the *parallelism* of a cut, which intuitively measures the angle formed by the objective vector and the normal vector of the cutting plane—promoting cutting planes that are nearly parallel with the objective direction. score_2 is the *efficacy*, or depth, of a cut, which measures the perpendicular distance from the LP optimum to the cut—promoting cutting planes that are “deeper”, as measured with respect to the LP optimum. More details about these scoring rules can be found in Balcan et al. [10] and references therein. Given an IP, for each $\mu \in [0, 1]$ (discretized at steps of 0.01) we choose the five cuts among the set of Chvátal-Gomory and Gomory mixed integer cuts that maximize (4). Figures 1a and 1b display the average tree size over 1000 samples drawn from the respective distribution for each value of μ used to choose cuts at the root. We ran our experiments using the C API of IBM ILOG CPLEX 20.1.0, with default cut generation disabled, and a 64-core machine with 512 GB of RAM.

B Omitted results and proofs from Section 3

B.1 Example in two dimensions

Consider the LP

$$\max\{x + y : x \leq 1, y \geq 0, y \leq x\}.$$

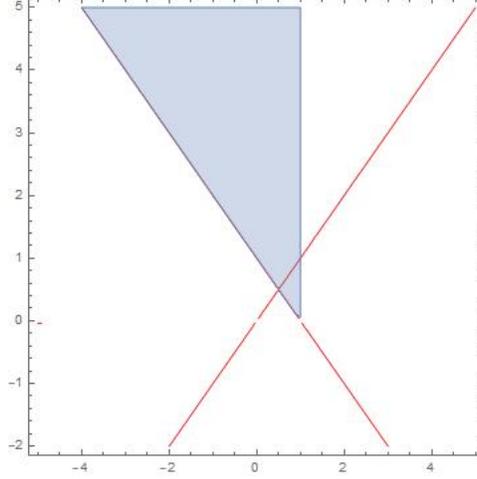


Figure 3: Decomposition of the parameter space: the blue region contains the set of (α_1, α_2) such that the constraint intersects the feasible region at $x = 1$ and $x = y$. The red lines consist of all (α_1, α_2) such that the objective value is equal at these intersection points. The red lines partition the blue region into two components: one where the new optimum is achieved at the intersection of h and $x = y$, and one where the new optimum is achieved at the intersection of h and $x = 1$.

The optimum is at $(x^*, y^*) = (1, 1)$. Consider adding an additional constraint $\alpha_1 x + \alpha_2 y \leq 1$. Let h denote the hyperplane $\alpha_1 x + \alpha_2 y = 1$. We derive a description of the set of parameters (α_1, α_2) such that h intersects the hyperplanes $x = 1$ and $y = x$. The intersection of h and $x = 1$ is given by

$$(x, y) = \left(1, \frac{1 - \alpha_1}{\alpha_2}\right),$$

which exists if and only if $\alpha_2 \neq 0$. This intersection point is in the LP feasible region if and only if $0 \leq \frac{1 - \alpha_1}{\alpha_2} \leq 1$ (which additionally ensures that $\alpha_2 \neq 0$). Similarly, h intersects $y = x$ at

$$(x, y) = \left(\frac{1}{\alpha_1 + \alpha_2}, \frac{1}{\alpha_1 + \alpha_2}\right),$$

which exists if and only if $\alpha_1 + \alpha_2 \neq 0$. This intersection point is in the LP feasible region if and only if $0 \leq \frac{1}{\alpha_1 + \alpha_2} \leq 1$. Now, we put down an ‘‘indifference’’ curve in (α_1, α_2) -space that represents the set of (α_1, α_2) such that the value of the objective achieved at the two aforementioned intersection points is equal. This surface is given by

$$\frac{2}{\alpha_1 + \alpha_2} = 1 + \frac{1 - \alpha_1}{\alpha_2}.$$

Since $\alpha_1 + \alpha_2 \neq 0$ and $\alpha_2 \neq 0$ (for the relevant α_1, α_2 in consideration), this is equivalent to $\alpha_1^2 - \alpha_2^2 - \alpha_1 + \alpha_2 = 0$, which is a degree-2 curve in α_1, α_2 . The left-hand-side can be factored to write this as $(\alpha_1 - \alpha_2)(\alpha_1 + \alpha_2 - 1) = 0$. Therefore, this curve is given by the two lines $\alpha_1 = \alpha_2$ and $\alpha_1 + \alpha_2 = 1$. Figure 3 illustrates the resulting partition of (α_1, α_2) -space.

It turns out that when $n = 2$ the indifference curve can always be factored into a product of linear terms. Let the objective of the LP be (c_1, c_2) , and let $s_1 x + s_2 y = u_1$ and $t_1 x + t_2 y = v_1$ be two intersecting edges of the LP feasible region. Let $\alpha_1 x + \alpha_2 y = \beta$ be an additional constraint. The intersection points of this constraint with the two lines, if they exist, are given by

$$\left(\frac{s_2 \beta - u \alpha_2}{s_2 \alpha_1 - s_1 \alpha_2}, \frac{s_1 \beta - u \alpha_1}{s_1 \alpha_2 - s_2 \alpha_1}\right) \text{ and } \left(\frac{t_2 \beta - v \alpha_2}{t_2 \alpha_1 - t_1 \alpha_2}, \frac{t_1 \beta - v \alpha_1}{t_1 \alpha_2 - t_2 \alpha_1}\right).$$

The indifference surface is thus given by

$$c_1 \frac{s_2 \beta - u \alpha_2}{s_2 \alpha_1 - s_1 \alpha_2} + c_2 \frac{s_1 \beta - u \alpha_1}{s_1 \alpha_2 - s_2 \alpha_1} = c_1 \frac{t_2 \beta - v \alpha_2}{t_2 \alpha_1 - t_1 \alpha_2} + c_2 \frac{t_1 \beta - v \alpha_1}{t_1 \alpha_2 - t_2 \alpha_1}.$$

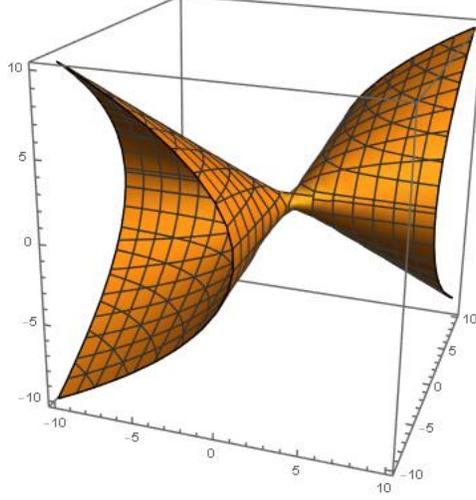


Figure 4: Indifference surface for two edges of the feasible region of an LP in three variables.

For α_1, α_2 such that $s_2\alpha_1 - s_1\alpha_2 \neq 0$ and $t_2\alpha_1 - t_1\alpha_2 \neq 0$, clearing denominators and some manipulation yields

$$(c_1\alpha_2 - c_2\alpha_1)((ut_1 - vs_1)\alpha_2 - (ut_2 - vs_2)\alpha_1 + (s_2t_2 - t_1s_2)\beta) = 0.$$

This curve consists of the two planes $c_1\alpha_2 - c_2\alpha_1 = 0$ and $(ut_1 - vs_1)\alpha_2 - (ut_2 - vs_2)\alpha_1 + (s_2t_2 - t_1s_2)\beta = 0$.

This is however not true if $n > 2$. For example, consider an LP in three variables x, y, z with the constraints $x + y \leq 1, x + z \leq 1, x \leq 1, z \leq 1$. Writing out the indifference surface (assuming the objective is $\mathbf{c} = (1, 1, 1)^T$) for the vertex on the intersection of $\{x + y = 1, x = 1\}$ and the vertex on $\{x + z = 1, z = 1\}$ yields

$$\alpha_1\alpha_2 - \alpha_2\beta - \alpha_3^2 + \alpha_3\beta = 0.$$

Setting $\beta = 1$, we can plot the resulting surface in $\alpha_1, \alpha_2, \alpha_3$ (Figure 4).

B.2 Linear programming sensitivity for multiple constraints

Lemma B.1. *Let $(\mathbf{c}, A, \mathbf{b})$ be an LP and let M denote the set of its m constraints. Let \mathbf{x}_{LP}^* and z_{LP}^* denote the optimal solution and its objective value, respectively. For $F \subseteq M$, let $A_F \in \mathbb{R}^{|F| \times n}$ and $\mathbf{b}_F \in \mathbb{R}^{|F|}$ denote the restrictions of A and \mathbf{b} to F . For $k \leq n$, $\alpha_1, \dots, \alpha_k \in \mathbb{R}^n$, $\beta_1, \dots, \beta_k \in \mathbb{R}$, and $F \subseteq M$ with $|F| = n - k$, let $A_{F, \alpha_1, \dots, \alpha_k} \in \mathbb{R}^{n \times n}$ denote the matrix obtained by adding row vectors $\alpha_1, \dots, \alpha_k$ to A_F and let $A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k} \in \mathbb{R}^{n \times n}$ be the matrix $A_{F, \alpha_1, \dots, \alpha_k} \in \mathbb{R}^{n \times n}$ with the i th column replaced by $[\mathbf{b}_F \ \beta_1 \ \dots \ \beta_k]^T$. There is a set of at most K hyperplanes, $nK^n m^n$ degree- K polynomial hypersurfaces, and $nK^n m^{2n}$ degree- $2K$ polynomial hypersurfaces partitioning $\mathbb{R}^{K(n+1)}$ into connected components such that for each component C , one of the following holds: either (1) $\mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K) = \mathbf{x}_{\text{LP}}^*$, or (2) there is a subset of cuts indexed by $\ell_1, \dots, \ell_k \in [K]$ and a set of constraints $F \subseteq M$ with $|F| = n - k$ such that*

$$\mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K) = \left(\frac{\det(A_{F, \alpha_{\ell_1}, \beta_{\ell_1}, \dots, \alpha_{\ell_k}, \beta_{\ell_k}})}{\det(A_{F, \alpha_{\ell_1}, \dots, \alpha_{\ell_k}})}, \dots, \frac{\det(A_{F, \alpha_{\ell_1}, \beta_{\ell_1}, \dots, \alpha_{\ell_k}, \beta_{\ell_k}})}{\det(A_{F, \alpha_{\ell_1}, \dots, \alpha_{\ell_k}})} \right),$$

for all $(\alpha_1, \beta_1, \dots, \alpha_K, \beta_K) \in C$.

Proof. First, if none of $\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K$ separate \mathbf{x}_{LP}^* , then $\mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K) = \mathbf{x}_{\text{LP}}^*$ and $z_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K) = z_{\text{LP}}^*$. The set of all such cuts is given by the intersection of halfspaces in $\mathbb{R}^{K(n+1)}$ given by

$$\bigcap_{j=1}^K \left\{ (\alpha_1, \beta_1, \dots, \alpha_k, \beta_k) \in \mathbb{R}^{K(n+1)} : \alpha_j^T \mathbf{x}_{\text{LP}}^* \leq \beta_j \right\}. \quad (5)$$

All other vectors of K cuts contain at least one cut that separates \mathbf{x}_{LP}^* , and those cuts therefore pass through $\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$. The new LP optimum is thus achieved at a vertex created by the cuts that separate \mathbf{x}_{LP}^* . As in the proof of Theorem 3.1, we consider all possible new vertices formed by our set of K cuts. In the case of a single cut, these new vertices necessarily were on edges of \mathcal{P} , but now they may lie on higher dimensional faces.

Consider a subset of $k \leq n$ cuts that separate \mathbf{x}_{LP}^* . Without loss of generality, denote these cuts by $\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_k^T \mathbf{x} \leq \beta_k$. We now establish conditions for these k cuts to “jointly” form a new vertex of \mathcal{P} . Any vertex created by these cuts must lie on a face f of \mathcal{P} with $\dim(f) = k$ (in the case that $k = n$, the relevant face f with $\dim(f) = n$ is \mathcal{P} itself). Letting M denote the set of m constraints that define \mathcal{P} , each dimension- k face f of \mathcal{P} can be identified with a (potentially empty) subset $F \subset M$ of size $n - k$ such that f is precisely the set of all points \mathbf{x} such that

$$\begin{aligned} \alpha_i^T \mathbf{x} &= b_i & \forall i \in F \\ \alpha_i^T \mathbf{x} &\leq b_i & \forall i \in M \setminus F, \end{aligned}$$

where α_i is the i th row of A . Let $A_F \in \mathbb{R}^{(n-k) \times n}$ denote the restriction of A to only the rows in F , and let $\mathbf{b}_F \in \mathbb{R}^{n-k}$ denote the entries of \mathbf{b} corresponding to the constraints in F . Consider removing the inequality constraints defining the face. The intersection of the cuts $\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_k^T \mathbf{x} \leq \beta_k$ and this unbounded surface (if it exists) is precisely the solution to the system of n linear equations

$$\begin{aligned} A_F \mathbf{x} &= \mathbf{b}_F \\ \alpha_1^T \mathbf{x} &= \beta_1 \\ &\vdots \\ \alpha_k^T \mathbf{x} &= \beta_k. \end{aligned}$$

Let $A_{F, \alpha_1, \dots, \alpha_k} \in \mathbb{R}^{n \times n}$ denote the matrix obtained by adding row vectors $\alpha_1, \dots, \alpha_k$ to A_F , and let $A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k} \in \mathbb{R}^{n \times n}$ denote the matrix $A_{F, \alpha_1, \dots, \alpha_k}$ where the i th column is replaced by

$$\begin{bmatrix} \mathbf{b}_F \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \in \mathbb{R}^n.$$

By Cramer’s rule, the solution to this system is given by

$$\mathbf{x} = \left(\frac{\det(A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k}^1)}{\det(A_{F, \alpha_1, \dots, \alpha_k})}, \dots, \frac{\det(A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k}^n)}{\det(A_{F, \alpha_1, \dots, \alpha_k})} \right),$$

and the value of the objective at this point is

$$\mathbf{c}^T \mathbf{x} = \sum_{i=1}^n c_i \cdot \frac{\det(A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k}^i)}{\det(A_{F, \alpha_1, \dots, \alpha_k})}.$$

Now, to ensure that the unique intersection point \mathbf{x} (1) exists and (2) actually lies on f (or simply lies in \mathcal{P} , in the case that $F = \emptyset$), we stipulate that it satisfies the inequality constraints in $M \setminus F$. That is,

$$\sum_{j=1}^n a_{ij} \frac{\det(A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k}^j)}{\det(A_{F, \alpha_1, \dots, \alpha_k})} \leq b_i \quad (6)$$

for every $i \in M \setminus F$. If $\alpha_1, \beta_1, \dots, \alpha_k, \beta_k$ satisfies any of these constraints, it must be that $\det(A_{F, \alpha_1, \dots, \alpha_k}) \neq 0$, which guarantees that $A_F \mathbf{x} = \mathbf{b}_F, \alpha_1^T \mathbf{x} = \beta_1, \dots, \alpha_k^T \mathbf{x} = \beta_k$ indeed has a unique solution. Now, $\det(A_{F, \alpha_1, \dots, \alpha_k})$ is a polynomial in $\alpha_1, \dots, \alpha_k$ of degree $\leq k$, since it is multilinear in each coefficient of each α_ℓ , $\ell = 1, \dots, k$. Similarly, $\det(A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k}^i)$ is a polynomial in $\alpha_1, \beta_1, \dots, \alpha_k, \beta_k$ of degree $\leq k$, again because it is multilinear in each cut parameter. Hence, the boundary each constraint of the form given by Equation 6 is a polynomial of degree at most k .

The collection of these polynomials for every k , every subset of $\{\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K\}$ of size k , and every face of \mathcal{P} of dimension k , along with the hyperplanes determining separation constraints (Equation 5), partition $\mathbb{R}^{K(n+1)}$ into connected components such that for all $(\alpha_1, \beta_1, \dots, \alpha_K, \beta_K)$ within a given connected component, there is a fixed subset of K and a fixed set of faces of \mathcal{P} such that the cuts with indices in that subset intersect every face in the set at a common vertex.

Now, consider a single connected component, denoted by C . Let f_1, \dots, f_ℓ denote the faces intersected by vectors of cuts in C , and let (without loss of generality) $1, \dots, k$ denote the subset of cuts that intersect these faces. Let $F_1, \dots, F_\ell \subset M$ denote the sets of constraints that are binding at each of these faces, respectively. For each pair f_p, f_q , consider the surface

$$\sum_{i=1}^n c_i \cdot \frac{\det(A_{F_p, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k}^i)}{\det(A_{F_p, \alpha_1, \dots, \alpha_k})} = \sum_{i=1}^n c_i \cdot \frac{\det(A_{F_q, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k}^i)}{\det(A_{F_q, \alpha_1, \dots, \alpha_k})},$$

which can be equivalently written as

$$\sum_{i=1}^n c_i \cdot \det(A_{F_p, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k}^i) \det(A_{F_q, \alpha_1, \dots, \alpha_k}) = \sum_{i=1}^n c_i \cdot \det(A_{F_q, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k}^i) \det(A_{F_p, \alpha_1, \dots, \alpha_k}). \quad (7)$$

This is a degree- $2k$ polynomial hypersurface in $(\alpha_1, \beta_1, \dots, \alpha_K, \beta_K) \in \mathbb{R}^{K(n+1)}$. This hypersurface is precisely the set of all cut vectors for which the LP objective achieved at the vertex on face f_p is equal to the LP objective value achieved at the vertex on face f_q . The collection of these surfaces for each p, q partitions C into further connected components. Within each of these connected components, the face containing the vertex that maximizes the objective is invariant, and the subset of cuts passing through that vertex is invariant. If $F \subseteq M$ is the set of binding constraints representing this face, and $\ell_1, \dots, \ell_k \in [K]$ represent the subset of cuts intersecting this face, $\mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K)$ and $z_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K)$ have the closed forms:

$$\mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K) = \left(\frac{\det(A_{F, \alpha_{\ell_1}, \beta_{\ell_1}, \dots, \alpha_{\ell_k}, \beta_{\ell_k}}^1)}{\det(A_{F, \alpha_{\ell_1}, \dots, \alpha_{\ell_k}})} , \dots , \frac{\det(A_{F, \alpha_{\ell_1}, \beta_{\ell_1}, \dots, \alpha_{\ell_k}, \beta_{\ell_k}}^n)}{\det(A_{F, \alpha_{\ell_1}, \dots, \alpha_{\ell_k}})} \right),$$

and

$$z_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K) = \sum_{i=1}^n c_i \cdot \frac{\det(A_{F, \alpha_{\ell_1}, \beta_{\ell_1}, \dots, \alpha_{\ell_k}, \beta_{\ell_k}}^i)}{\det(A_{F, \alpha_{\ell_1}, \dots, \alpha_{\ell_k}})}.$$

for all $(\alpha_1, \beta_1, \dots, \alpha_K, \beta_K)$ within this component. We now count the number of surfaces used to obtain our decomposition. First, we added K hyperplanes encoding separation constraints for each of the K cuts (Equation 5). Then, for every subset $S \subseteq K$ of size $\leq n$, and for every face F of \mathcal{P} with $\dim(F) = |S|$, we first considered at most $|M \setminus F| \leq m$ degree- $\leq K$ polynomial hypersurfaces representing decision boundaries for when cuts in S intersected that face (Equation 6). The number of k -dimensional faces of \mathcal{P} is at most $\binom{m}{n-k} \leq m^{n-k} \leq m^{n-1}$, so the total number of these hypersurfaces is at most $(\binom{K}{0} + \dots + \binom{K}{n})m^n \leq nK^n m^n$. Finally, we considered a degree- $2K$ polynomial hypersurface for every subset of cuts and every pair of faces with degree equal to the size of the subset, of which there are at most $nK^n \binom{m}{2} \leq nK^n m^{2n}$. \square

C Omitted results and proofs from Section 4

We first require the following lemma which bounds the number of relevant subsets of $\mathcal{BC} := \{\mathbf{x}[i] \leq \ell, \mathbf{x}[i] \geq \ell\}_{0 \leq \ell \leq \tau, i \in [n]}$ that define a possible node expanded by B&C. \mathcal{BC} is a set of size $2n(\tau + 1)$ so naïvely there are at most $2^{2n(\tau+1)}$ subsets of branching constraints. The following observation allows us to greatly reduce the number of sets we consider.

Lemma C.1. *Fix an IP (c, A, b) . Define an equivalence relation on pairs of branching-constraint sets $\sigma_1, \sigma_2 \subseteq \mathcal{BC}$, by $\sigma_1 \sim \sigma_2 \iff \mathbf{x}_{\text{LP}}^*(\alpha^T \mathbf{x} \leq \beta, \sigma_1) = \mathbf{x}_{\text{LP}}^*(\alpha^T \mathbf{x} \leq \beta, \sigma_2)$ for all possible cutting planes $\alpha^T \mathbf{x} \leq \beta$. The number of equivalence classes of \sim is at most τ^{3n} .*

Proof of Lemma C.1. Consider as an example $\sigma_1 = \{\mathbf{x}[1] \leq 1, \mathbf{x}[1] \leq 5\}$ and $\sigma_2 = \{\mathbf{x}[1] \leq 1\}$. We have $\mathbf{x}_{\text{LP}}^*(\alpha^T \mathbf{x} \leq \beta, \sigma_1) = \mathbf{x}_{\text{LP}}^*(\alpha^T \mathbf{x} \leq \beta, \sigma_2)$ for any cut $\alpha^T \mathbf{x} \leq \beta$, because the constraint

$\mathbf{x}[1] \leq 5$ is redundant in σ_1 . More generally, any $\sigma \subseteq \mathcal{BC}$ can be reduced by preserving only the tightest \leq constraint and tightest \geq constraint without affecting the resulting LP optimal solutions. The number of such unique reduced sets is at most $((\tau + 2)^2)^n < \tau^{3n}$ (for each variable, there are $\tau + 2$ possibilities for the tightest \leq constraint: no constraint or one of $\mathbf{x}[i] \leq 0, \dots, \mathbf{x}[i] \leq \tau$, and similarly $\tau + 2$ possibilities for the \geq constraint). \square

Proof of Lemma 4.1. We carry out the same reasoning in the proof of Theorem 3.1 for each reduced σ . The number of edges of $\mathcal{P}(\sigma)$ is at most $\binom{m+|\sigma|}{n-1} \leq (m+|\sigma|)^{n-1}$. For each edge E , we considered at most $|(M \cup \sigma) \setminus E| \leq m + |\sigma|$ hyperplanes, for a total of at most $(m + |\sigma|)^n$ halfspaces. Then, we had a degree-2 polynomial hypersurface for every pair of edges, of which there are at most $\binom{m+|\sigma|}{2} \leq (m + |\sigma|)^2$. Summing over all reduced σ (of which there are at most τ^{3n}), combined with the fact that if σ is reduced then $|\sigma| \leq 2n$, we get a total of at most $(m + 2n)^n \tau^{3n}$ hyperplanes and at most $(m + 2n)^{2n} \tau^{3n}$ degree-2 hypersurfaces, as desired. \square

Let $\mathcal{V} \subseteq \mathbb{R}^{n+1}$ denote the set of all valid cuts for the input IP (c, A, b) . The set \mathcal{V} is a polyhedron since it can be expressed as

$$\mathcal{V} = \bigcap_{\bar{\mathbf{x}} \in \mathcal{P}_1} \{(\boldsymbol{\alpha}, \beta) \in \mathbb{R}^{n+1} : \boldsymbol{\alpha}^T \bar{\mathbf{x}} \leq \beta\},$$

and \mathcal{P}_1 is finite as \mathcal{P} is bounded. For cuts outside \mathcal{V} , we assume the B&C tree takes some special form denoting an invalid cut. Our goal now is to decompose \mathcal{V} into connected components such that $\mathbf{1}[\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) \in \mathbb{Z}^n]$ is invariant for all $(\boldsymbol{\alpha}, \beta)$ in each component.

Proof of Lemma 4.3. Fix a connected component C in the decomposition that includes the facets defining \mathcal{V} and the surfaces obtained in Lemma 4.2. For all $\sigma \in \mathcal{BC}$, $\mathbf{x}_1 \in \mathcal{P}_1$, and $i = 1, \dots, n$, consider the surface

$$\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] = \mathbf{x}_1[i]. \quad (8)$$

This surface is a hyperplane, since by Lemma 4.1, either $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] = \mathbf{x}_{\text{LP}}^*(\sigma)[i]$ or $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] = \frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^*)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})}$, where $E \subseteq M \cup \sigma$ is the subset of constraints corresponding to σ and C . Clearly, within any connected component of C induced by these hyperplanes, for every σ and $\mathbf{x}_1 \in \mathcal{P}_1$, $\mathbf{1}[\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) = \mathbf{x}_1]$ is invariant. Finally, if $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) \in \mathbb{Z}^n$ for some cut $\boldsymbol{\alpha}^T \mathbf{x} \leq \beta$ within a given connected component, $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) = \mathbf{x}_1$ for some $\mathbf{x}_1 \in \mathcal{P}_{\text{IH}}(\sigma) \subseteq \mathcal{P}_1$, which means that $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) = \mathbf{x}_1 \in \mathbb{Z}^n$ for all cuts $\boldsymbol{\alpha}^T \mathbf{x} \leq \beta$ in that connected component.

We now count the number of hyperplanes given by Equation 8. For each σ , there are $\binom{m+|\sigma|}{n-1} \leq (m + 2n)^{n-1}$ binding edge constraints $E \subseteq M \cup \sigma$ defining the formula of Lemma 4.1, and we have $n|\mathcal{P}_1|$ hyperplanes for each E . Since $\tau = \max_{\mathbf{x} \in \mathcal{P}_1} \|\mathbf{x}\|_\infty$, $|\mathcal{P}_1| \leq \tau^n$. So the total number of hyperplanes given by Equation 8 is at most $\tau^{3n}(m + 2n)^{n-1}n\tau^n \leq (m + 2n)^n \tau^{4n}$. The number of facets defining \mathcal{V} is at most $|\mathcal{P}_{\text{IH}}| \leq |\mathcal{P}_1| \leq \tau^n$. Adding these to the counts obtained in Lemma 4.2 yields the final tallies in the lemma statement. \square

Proof of Theorem 4.4. Fix a connected component C in the decomposition induced by the set of hyperplanes and degree-2 hypersurfaces established in Lemma 4.3. Let

$$Q_1, \dots, Q_{i_1}, I_1, Q_{i_1+1}, \dots, Q_{i_2}, I_2, Q_{i_2+1}, \dots \quad (9)$$

denote the nodes of the tree branch-and-cut creates, in order of exploration, under the assumption that a node is pruned if and only if either the LP at that node is infeasible or the LP optimal solution is integral (so the ‘‘bounding’’ of branch-and-bound is suppressed). Here, a node is identified by the list σ of branching constraints added to the input IP. Nodes labeled by Q are either infeasible or have fractional LP optimal solutions. Nodes labeled by I have integral LP optimal solutions and are candidates for the incumbent integral solution at the point they are encountered. (The nodes are functions of $\boldsymbol{\alpha}$ and β , as are the indices i_1, i_2, \dots) By Lemma 4.3 and the observation following it, this ordered list of nodes is invariant over all $(\boldsymbol{\alpha}, \beta) \in C$.

Now, given an node index ℓ , let $I(\ell)$ denote the incumbent node with the highest objective value encountered up until the ℓ th node searched by B&C, and let $z(I(\ell))$ denote its objective value. For each node Q_ℓ , let σ_ℓ denote the branching constraints added to arrive at node Q_ℓ . The hyperplane

$$z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_\ell) = z(I(\ell)) \quad (10)$$

(which is a hyperplane due to Lemma 4.1) partitions C into two subregions. In one subregion, $z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_\ell) \leq z(I(\ell))$, that is, the objective value of the LP optimal solution is no greater than the objective value of the current incumbent integer solution, and so the subtree rooted at Q_ℓ is pruned. In the other subregion, $z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_\ell) > z(I(\ell))$, and Q_ℓ is branched on further. Therefore, within each connected component of C induced by all hyperplanes given by Equation 10 for all ℓ , the set of node within the list (9) that are pruned is invariant. Combined with the surfaces established in Lemma 4.3, these hyperplanes partition \mathbb{R}^{n+1} into connected components such that as $(\boldsymbol{\alpha}, \beta)$ varies within a given component, the tree built by branch-and-cut is invariant.

Finally, we count the total number of surfaces inducing this partition. Unlike the counting stages of the previous lemmas, we will first have to count the number of connected components induced by the surfaces established in Lemma 4.3. This is because the ordered list of nodes explored by branch-and-cut (9) can be different across each component, and the hyperplanes given by Equation 10 depend on this list. From Lemma 4.3 we have $3(m+2n)^n \tau^{4n}$ hyperplanes, $3(m+2n)^{3n} \tau^{4n}$ degree-2 polynomial hypersurfaces, and $(m+2n)^{6n} \tau^{4n}$ degree-5 polynomial hypersurfaces. To determine the connected components of \mathbb{R}^{n+1} induced by the zero sets of these polynomials, it suffices to consider the zero set of the product of all polynomials defining these surfaces. Denote this product polynomial by p . The degree of the product polynomial is the sum of the degrees of $3(m+2n)^n \tau^{4n}$ degree-1 polynomials, $3(m+2n)^{3n} \tau^{4n}$ degree-2 polynomials, and $(m+2n)^{6n} \tau^{4n}$ degree-5 polynomials, which is at most $3(m+2n)^n \tau^{4n} + 2 \cdot 3(m+2n)^{3n} \tau^{4n} + 5 \cdot (m+2n)^{6n} \tau^{4n} < 14(m+2n)^{3n} \tau^{4n}$. By Warren's theorem, the number of connected components of $\mathbb{R}^{n+1} \setminus \{(\boldsymbol{\alpha}, \beta) : p(\boldsymbol{\alpha}, \beta) = 0\}$ is $O((14(m+2n)^{3n} \tau^{4n})^{n-1})$, and by the Milnor-Thom theorem, the number of connected components of $\{(\boldsymbol{\alpha}, \beta) : p(\boldsymbol{\alpha}, \beta) = 0\}$ is $O((14(m+2n)^{3n} \tau^{4n})^{n-1})$ as well. So, the number of connected components induced by the surfaces in Lemma 4.3 is $O(14^n (m+2n)^{3n^2} \tau^{4n^2})$. For every connected component C in Lemma 4.3, the closed form of $z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_\ell)$ is already determined due to Lemma 4.1, and so the number of hyperplanes given by Equation 10 is at most the number of possible $\sigma \subseteq \mathcal{BC}$, which is at most τ^{3n} . So across all connected components C , the total number of hyperplanes given by Equation 10 is $O(14^n (m+2n)^{3n^2} \tau^{5n^2})$. Finally, adding this to the surface-counts established in Lemma 4.3 yields the lemma statement. \square

C.1 Product scoring rule for variable selection

Let σ be the set of branching constraints added thus far. The product scoring rule branches on the variable $i \in [n]$ that maximizes:

$$\max\{z_{\text{LP}}^*(\sigma) - z_{\text{LP}}^*(x_i \leq \lfloor x_{\text{LP}}^*(\sigma)[i] \rfloor, \sigma), \gamma\} \cdot \max\{z_{\text{LP}}^*(\sigma) - z_{\text{LP}}^*(x_i \geq \lceil x_{\text{LP}}^*(\sigma)[i] \rceil, \sigma), \gamma\},$$

where $\gamma = 10^{-6}$.

Lemma C.2. *There is a set of at most $3(m+2n)^n \tau^{3n}$ hyperplanes and $(m+2n)^{2n} \tau^{3n}$ degree-2 polynomial hypersurfaces partitioning \mathbb{R}^{n+1} into connected components such that for any connected component C and any σ , the set of branching constraints $\{x_i \leq \lfloor x_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] \rfloor, x_i \geq \lceil x_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] \rceil \mid i \in [n]\}$ is invariant across all $(\boldsymbol{\alpha}, \beta) \in C$.*

Proof. Fix a connected component C in the decomposition established in Lemma 4.1. By Lemma 4.1, for each σ , either $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) = \mathbf{x}_{\text{LP}}^*(\sigma)$ or there exists $E \subseteq M \cup \sigma$ such that $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] = \frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})}$ for all $(\boldsymbol{\alpha}, \beta) \in C$. Fix a variable $i \in [n]$, which corresponds to two branching constraints

$$x_i \leq \lfloor \mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] \rfloor \quad \text{and} \quad x_i \geq \lceil \mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] \rceil. \quad (11)$$

If C is a component where $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) = \mathbf{x}_{\text{LP}}^*(\sigma)$, then these two branching constraints are trivially invariant over $(\boldsymbol{\alpha}, \beta) \in C$. Otherwise, in order to further decompose C such that the right-hand-sides of these constraints are invariant for every σ , we add the two decision boundaries given by

$$k \leq \frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})} \leq k + 1$$

for every i, σ , and every integer $k = 0, \dots, \tau - 1$, where $\tau = \max_{\mathbf{x} \in \mathcal{P} \cap \mathbb{Z}^n} \|\mathbf{x}\|_\infty$. This ensures that within every connected component of C induced by these boundaries (hyperplanes),

$$\lfloor \mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] \rfloor = \left\lfloor \frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})} \right\rfloor \text{ and } \lceil \mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] \rceil = \left\lceil \frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})} \right\rceil$$

are invariant, so the branching constraints from Equation (11) are invariant. For a fixed σ , there are two hyperplanes for every $E \subseteq M \cup \sigma$ corresponding to an edge of $\mathcal{P}(\sigma)$ and $i = 1, \dots, n$, for a total of at most $2n \binom{m+|\sigma|}{n-1} \leq 2n(m+|\sigma|)^{n-1}$ hyperplanes. Summing over all reduced σ , we get a total of $2n(m+2n)^{n-1} \tau^{3n} < 2(m+2n)^n \tau^{3n}$ hyperplanes. Adding these hyperplanes to the set of hyperplanes established in Lemma 4.1 yields the lemma statement. \square

Proof of Lemma 4.2. Fix a connected component C in the decomposition established in Lemma C.2. We know that for each set of branching constraints σ :

- By Lemma 4.1, either $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) = \mathbf{x}_{\text{LP}}^*(\sigma)$ or there exists $E \subseteq M \cup \sigma$ such that $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] = \frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})}$ for all $(\boldsymbol{\alpha}, \beta) \in C$ and all $i \in [n]$, and
- The set of branching constraints $\{x_i \leq \lfloor \mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] \rfloor, x_i \geq \lceil \mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[i] \rceil \mid i \in [n]\}$ is invariant across all $(\boldsymbol{\alpha}, \beta) \in C$.

Suppose that σ is the list of branching constraints added so far. For any variable $k \in [n]$, let

$$\sigma_k^- = (x_k \leq \lfloor \mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[k] \rfloor, \sigma) \text{ and } \sigma_k^+ = (x_k \geq \lceil \mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma)[k] \rceil, \sigma).$$

So long as $(\boldsymbol{\alpha}, \beta) \in C$, σ_k^- and σ_k^+ are fixed. With this notation, we can write the product scoring rule as

$$\max\{z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) - z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_k^-), \gamma\} \cdot \max\{z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) - z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_k^+), \gamma\},$$

where $\gamma = 10^{-6}$.

By Lemma 4.1, we know that across all $(\boldsymbol{\alpha}, \beta) \in C$, either $z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_k^+) = z_{\text{LP}}^*(\sigma_k^+)$ or there exists $E_k^+ \subseteq M \cup \sigma_k^+$ such that

$$z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_k^+) = \sum_{i=1}^n c_i \cdot \frac{\det(A_{E_k^+, \boldsymbol{\alpha}, \beta, \sigma_k^+}^i)}{\det(A_{E_k^+, \boldsymbol{\alpha}, \sigma_k^+})},$$

and similarly for σ_k^- , defined according to some edge set $E_k^- \subseteq M \cup \sigma_k^-$. Therefore, for each $k \in [n]$, there is a single degree-2 polynomial hypersurface partitioning C into connected components such that within each connected component, either

$$z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) - z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_k^-) \geq \gamma \tag{12}$$

or vice versa, and similarly for σ_k^+ . In particular, the former hypersurface will have one of four forms:

1. $z_{\text{LP}}^*(\sigma) - z_{\text{LP}}^*(\sigma_k^-) \geq \gamma$, which is uniformly satisfied or not satisfied across all $(\boldsymbol{\alpha}, \beta) \in C$,
2. $z_{\text{LP}}^*(\sigma) - \sum_{i=1}^n c_i \cdot \frac{\det(A_{E_k^-, \boldsymbol{\alpha}, \beta, \sigma_k^-}^i)}{\det(A_{E_k^-, \boldsymbol{\alpha}, \sigma_k^-})} \geq \gamma$, which is a hyperplane,
3. $\sum_{i=1}^n c_i \cdot \frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})} - z_{\text{LP}}^*(\sigma_k^-) \geq \gamma$, which is a hyperplane, or
4. $\sum_{i=1}^n c_i \left(\frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})} - \frac{\det(A_{E_k^-, \boldsymbol{\alpha}, \beta, \sigma_k^-}^i)}{\det(A_{E_k^-, \boldsymbol{\alpha}, \sigma_k^-})} \right) \geq \gamma$, which is a degree-2 polynomial hypersurface.

Simply said, these are all degree-2 polynomial hypersurfaces.

Within any region induced by these hypersurfaces, the comparison between any two variables x_k and x_j will have the form

$$\begin{aligned} & \max\{z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) - z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_k^-), \gamma\} \cdot \max\{z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) - z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_k^+), \gamma\} \\ & \geq \max\{z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) - z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_j^-), \gamma\} \cdot \max\{z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma) - z_{\text{LP}}^*(\boldsymbol{\alpha}^T \mathbf{x} \leq \beta, \sigma_j^+), \gamma\} \end{aligned}$$

which at its most complex will equal

$$\begin{aligned} & \sum_{i=1}^n c_i \left(\frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})} - \frac{\det(A_{E_k^-, \boldsymbol{\alpha}, \beta, \sigma_k^-}^i)}{\det(A_{E_k^-, \boldsymbol{\alpha}, \sigma_k^-})} \right) \cdot \sum_{i=1}^n c_i \left(\frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})} - \frac{\det(A_{E_k^+, \boldsymbol{\alpha}, \beta, \sigma_k^+}^i)}{\det(A_{E_k^+, \boldsymbol{\alpha}, \sigma_k^+})} \right) \\ & \geq \sum_{i=1}^n c_i \left(\frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})} - \frac{\det(A_{E_j^-, \boldsymbol{\alpha}, \beta, \sigma_j^-}^i)}{\det(A_{E_j^-, \boldsymbol{\alpha}, \sigma_j^-})} \right) \cdot \sum_{i=1}^n c_i \left(\frac{\det(A_{E, \boldsymbol{\alpha}, \beta, \sigma}^i)}{\det(A_{E, \boldsymbol{\alpha}, \sigma})} - \frac{\det(A_{E_j^+, \boldsymbol{\alpha}, \beta, \sigma_j^+}^i)}{\det(A_{E_j^+, \boldsymbol{\alpha}, \sigma_j^+})} \right). \end{aligned} \quad (13)$$

This inequality can be written as a degree-5 polynomial hypersurface. In any region induced by these hypersurfaces, the variable that branch-and-cut branches on will be fixed.

We now count the total number of hypersurfaces. First, we count the number of degree-2 polynomial hypersurfaces from Equation (12): there is a hypersurface defined by each variable x_k , set of branching constraints σ , cutoff $t \in [\tau]$ such that $\sigma_k^- = (x_k \leq t, \sigma)$, set $E \subseteq M \cup \sigma$ corresponding to an edge of $\mathcal{P}(\sigma)$, and set $E_k^- \subseteq M \cup \sigma_k^-$ (and similarly for σ_k^+ and E_k^+). For a fixed σ , this amounts to $2n\tau \binom{m+|\sigma|}{n-1} \binom{m+|\sigma|+1}{n-1} \leq 2n\tau(m+|\sigma|+1)^{2(n-1)}$ hypersurfaces. Summing over all τ^{3n} reduced σ , we have $2n\tau^{3n+1}(m+2n+1)^{2(n-1)}$ degree-2 polynomial hypersurfaces.

Next, we count the number of degree-5 polynomial hypersurfaces from Equation (13): there is a hypersurface defined by each pair of variables x_k, x_j , set of branching constraints σ , cutoffs $t_k, t_j \in [\tau]$ such that $\sigma_k^- = (x_k \leq t_k, \sigma)$ and $\sigma_j^- = (x_j \leq t_j, \sigma)$, and sets $E, E_k^-, E_k^+, E_j^-, E_j^+$ corresponding to edges of $\mathcal{P}(\sigma), \mathcal{P}(\sigma_k^-), \mathcal{P}(\sigma_k^+), \mathcal{P}(\sigma_j^-), \mathcal{P}(\sigma_j^+)$. For a fixed σ , this amounts to $n^2\tau^2 \binom{m+|\sigma|}{n-1} \binom{m+|\sigma|+1}{n-1}^4 \leq n^2\tau^2(m+|\sigma|+1)^{5(n-1)}$ hypersurfaces. Summing over all τ^{3n} reduced σ , we have $n^2\tau^{3n+2}(m+2n+1)^{5(n-1)}$ degree-5 polynomial hypersurfaces.

Adding these hypersurfaces to those from Lemma C.2, we get the lemma statement. \square

C.2 Extension to multiple cutting planes

We can similarly derive a multi-cut version of Lemma 4.1 that controls $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K, \sigma)$ for any set of branching constraints. We use the following notation. Let $(\mathbf{c}, A, \mathbf{b})$ be an LP and let M denote the set of its m constraints. For $F \subseteq M \cup \sigma$, let $A_{F, \sigma} \in \mathbb{R}^{|F| \times n}$ and $\mathbf{b}_{F, \sigma} \in \mathbb{R}^{|F|}$ denote the restrictions of A_σ and \mathbf{b}_σ to F . For $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k \in \mathbb{R}^n, \beta_1, \dots, \beta_k \in \mathbb{R}$, and $F \subseteq M \cup \sigma$ with $|F| = n - k$, let $A_{F, \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k, \sigma} \in \mathbb{R}^{n \times n}$ denote the matrix obtained by adding row vectors $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$ to $A_{F, \sigma}$ and let $A_{F, \boldsymbol{\alpha}_1, \beta_1, \dots, \boldsymbol{\alpha}_k, \beta_k, \sigma}^i \in \mathbb{R}^{n \times n}$ be the matrix $A_{F, \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k, \sigma} \in \mathbb{R}^{n \times n}$ with the i th column replaced by $[\mathbf{b}_{F, \sigma} \ \beta_1 \ \dots \ \beta_k]^T$.

Corollary C.3. *Fix an IP $(\mathbf{c}, A, \mathbf{b})$. There is a set of at most K hyperplanes, $nK^n(m+2n)\tau^{3n}$ degree- K polynomial hypersurfaces, and $nK^n(m+2n)^{2n}\tau^{3n}$ degree-2 K polynomial hypersurfaces partitioning $\mathbb{R}^{K(n+1)}$ into connected components such that for each component C and every $\sigma \subseteq \text{BC}$, one of the following holds: either (1) $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K, \sigma) = \mathbf{x}_{\text{LP}}^*(\sigma)$, or (2) there is a subset of cuts indexed by $\ell_1, \dots, \ell_k \in [K]$ and a set of constraints $F \subseteq M \cup \sigma$ with $|F| = n - k$ such that*

$$\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K, \sigma) = \left(\frac{\det(A_{F, \boldsymbol{\alpha}_{\ell_1}, \beta_{\ell_1}, \dots, \boldsymbol{\alpha}_{\ell_k}, \beta_{\ell_k}, \sigma}^1)}{\det(A_{F, \boldsymbol{\alpha}_{\ell_1}, \dots, \boldsymbol{\alpha}_{\ell_k}, \sigma})}, \dots, \frac{\det(A_{F, \boldsymbol{\alpha}_{\ell_1}, \beta_{\ell_1}, \dots, \boldsymbol{\alpha}_{\ell_k}, \beta_{\ell_k}, \sigma}^n)}{\det(A_{F, \boldsymbol{\alpha}_{\ell_1}, \dots, \boldsymbol{\alpha}_{\ell_k}, \sigma})} \right),$$

for all $(\boldsymbol{\alpha}_1, \beta_1, \dots, \boldsymbol{\alpha}_K, \beta_K) \in C$.

Proof. The exact same reasoning in the proof of Lemma B.1 applies. We still have K hyperplanes. Now, for each σ , for each subset $S \subseteq K$ with $|S| \leq n$, and for every face F of $\mathcal{P}(\sigma)$ with $\dim(F) = |S|$, we have at most m degree- K polynomial hypersurfaces. The number of k -dimensional faces of $\mathcal{P}(\sigma)$ is at most $\binom{m+|\sigma|}{n-k} \leq (m+2n)^{n-1}$, so the total number of these hypersurfaces is at most $nK^n(m+2n)^n\tau^{3n}$. Finally, for every σ , we considered a degree- $2K$ polynomial hypersurfaces for every subset of cuts and every pair of faces with degree equal to the size of the subset, of which there are at most $nK^n(m+2n)^{2n}\tau^{3n}$, as desired. \square

We now refine the decomposition obtained in Lemma 4.1 so that the branching constraints added at each step of branch-and-cut are invariant within a region. For ease of exposition, we assume that branch-and-cut uses a lexicographic variable selection policy. This means that the variable branched on at each node of the search tree is fixed and given by the lexicographic ordering x_1, \dots, x_n . Generalizing the argument to work for other policies, such as the product scoring rule, can be done as in the single-cut case.

Lemma C.4. *Suppose branch-and-cut uses a lexicographic variable selection policy. Then, there is a set of at most K hyperplanes, $3n^2K^n(m+2n)^n\tau^{3n}$ degree- K polynomial hypersurfaces, and $nK^n(m+2n)^{2n}\tau^{3n}$ degree- $2K$ polynomial hypersurfaces partitioning \mathbb{R}^{n+1} into connected components such that within each connected component, the branching constraints used at every step of branch-and-cut are invariant.*

Proof. Fix a connected component C in the decomposition established in Corollary C.3. Then, by Corollary C.3, for each σ , either $\mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K, \sigma) = \mathbf{x}_{\text{LP}}^*(\sigma)$ or there exists cuts (without loss of generality) labeled by indices $1, \dots, k \in [K]$ and there exists $F \subseteq M \cup \sigma$ such that

$$\mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K, \sigma)[i] = \frac{\det(A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k, \sigma}^i)}{\det(A_{F, \alpha_1, \dots, \alpha_k, \sigma})}$$

for all $(\alpha, \beta) \in C$ and all $i \in [n]$. Now, if we are at a stage in the branch-and-cut tree where σ is the list of branching constraints added so far, and the i th variable is being branched on next, the two constraints generated are

$$x_i \leq \lfloor \mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K, \sigma)[i] \rfloor \quad \text{and} \quad x_i \geq \lceil \mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K, \sigma)[i] \rceil,$$

respectively. If C is a component where $\mathbf{x}_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K, \sigma) = \mathbf{x}_{\text{LP}}^*(\sigma)$, then there is nothing more to do, since the branching constraints at that point are trivially invariant over $(\alpha_1, \beta_1, \dots, \alpha_K, \beta_K) \in C$. Otherwise, in order to further decompose C such that the right-hand-side of these constraints are invariant for every σ and every $i = 1, \dots, n$, we add the two decision boundaries given by

$$k \leq \frac{\det(A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k, \sigma}^i)}{\det(A_{F, \alpha_1, \dots, \alpha_k, \sigma})} \leq k+1$$

for every i, σ , and every integer $k = 0, \dots, \tau - 1$, where $\tau = \lceil \max_{\mathbf{x} \in \mathcal{P}} \|\mathbf{x}\|_\infty \rceil$. This ensures that within every connected component of C induced by these boundaries (degree- K polynomial hypersurfaces),

$$\lfloor \mathbf{x}_{\text{LP}}^*(\alpha^T \mathbf{x} \leq \beta, \sigma)[i] \rfloor = \left\lfloor \frac{\det(A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k, \sigma}^i)}{\det(A_{F, \alpha_1, \dots, \alpha_k, \sigma})} \right\rfloor$$

and

$$\lceil \mathbf{x}_{\text{LP}}^*(\alpha^T \mathbf{x} \leq \beta, \sigma)[i] \rceil = \left\lceil \frac{\det(A_{F, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k, \sigma}^i)}{\det(A_{F, \alpha_1, \dots, \alpha_k, \sigma})} \right\rceil$$

are invariant, so the branching constraints added by, for example, a lexicographic branching rule, are invariant. For a fixed σ , there are two hypersurfaces for every subset $S \subseteq [K]$, every $F \subseteq M \cup \sigma$ corresponding to a $|S|$ -dimensional face of $\mathcal{P}(\sigma)$, and every $i = 1, \dots, n$, for a total of at most $2n^2K^n \binom{m+|\sigma|}{|S|} \leq 2n^2K^n(m+2n)^n$. Summing over all reduced σ , we get a total of $2n^2K^n(m+2n)^n\tau^{3n}$ hypersurfaces. Adding these hypersurfaces to the set of hypersurfaces established in Corollary C.3 yields the lemma statement. \square

Now, as in the single-cut case, we consider the constraints that ensure that all cuts are valid. Let $\mathcal{V} \subseteq \mathbb{R}^{K(n+1)}$ denote the set of all vectors of valid K cuts. As before, \mathcal{V} is a polyhedron, since we may write

$$\mathcal{V} = \bigcap_{k=1}^K \bigcap_{\mathbf{x}_{\text{IH}} \in \mathcal{P}_{\text{IH}}} \left\{ (\boldsymbol{\alpha}_1, \beta_1, \dots, \boldsymbol{\alpha}_K, \beta_k) \in \mathbb{R}^{K(n+1)} : \boldsymbol{\alpha}_k^T \mathbf{x}_{\text{IH}} \leq \beta_k \right\}.$$

We now refine our decomposition further to control the integrality of the various LP solutions at each node of branch-and-cut.

Lemma C.5. *Given an IP (c, A, b) , there is a set of at most $2K\tau^n$ hyperplanes, $4n^2K^n(m+2n)^n\tau^{4n}$ degree- K polynomial hypersurfaces, and $nK^n(m+2n)^{2n}\tau^{3n}$ degree- $2K$ polynomial hypersurfaces partitioning $\mathbb{R}^{K(n+1)}$ into connected components such that for each component C , and each $\sigma \subseteq \mathcal{BC}$,*

$$\mathbf{1}[\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K, \sigma) \in \mathbb{Z}^n]$$

is invariant for all $(\boldsymbol{\alpha}_1, \beta_1, \dots, \boldsymbol{\alpha}_K, \beta_K) \in C$.

Proof. Fix a connected component C in the decomposition that includes the facets defining \mathcal{V} and the surfaces obtained in Lemma C.4. For all $\sigma \in \mathcal{BC}$, $\mathbf{x}_1 \in \mathcal{P}_1$, and $i = 1, \dots, n$, consider the surface

$$\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K, \sigma)[i] = \mathbf{x}_1[i]. \quad (14)$$

This surface is a polynomial hypersurface of degree at most K , due to Corollary C.3. Clearly, within any connected component of C induced by these hyperplanes, for every σ and $\mathbf{x}_1 \in \mathcal{P}_1$, $\mathbf{1}[\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K, \sigma) = \mathbf{x}_1]$ is invariant. Finally, if $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K, \sigma) \in \mathbb{Z}^n$ for some K cuts $\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K$ within a given connected component, $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K, \sigma) = \mathbf{x}_1$ for some $\mathbf{x}_1 \in \mathcal{P}_{\text{IH}}(\sigma) \subseteq \mathcal{P}_1$, which means that $\mathbf{x}_{\text{LP}}^*(\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K, \sigma) = \mathbf{x}_1 \in \mathbb{Z}^n$ for all vectors of K cuts $\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K$ in that connected component.

We now count the number of hyperplanes given by Equation 14. For each σ , there are nK^n possible subsets of cut indices and at most $(m+2n)^{n-1}$ binding face constraints $F \subseteq M \cup \sigma$ defining the formula of Corollary C.3. For each subset-face pair, there are $n|\mathcal{P}_1| \leq n\tau^n$ degree- K polynomial hypersurfaces given by Equation 14. So the total number of such hypersurfaces over all σ is at most $\tau^{3n}n^2K^n(m+2n)^{n-1}\tau^n$. The number of facets defining \mathcal{V} is at most $K|\mathcal{P}_1| \leq K\tau^n$. Adding these to the counts obtained in Lemma C.4 yields the final tallies in the lemma statement. \square

At this point, as in the single-cut case, if the bounding aspect of branch-and-cut is suppressed, our decomposition yields connected components over which the branch-and-cut tree built is invariant. We now prove our main structural theorem for B&C as a function of multiple cutting planes at the root.

Theorem C.6. *Given an IP (c, A, b) , there is a set of at most $O(12^n n^{2n} K^{2n^2} (m+2n)^{2n^2} \tau^{5n^2})$ polynomial hypersurfaces of degree at most $2K$ partitioning $\mathbb{R}^{K(n+1)}$ into connected components such that the branch-and-cut tree built after adding the K cuts $\boldsymbol{\alpha}_1^T \mathbf{x} \leq \beta_1, \dots, \boldsymbol{\alpha}_K^T \mathbf{x} \leq \beta_K$ at the root is invariant over all $(\boldsymbol{\alpha}_1, \beta_1, \dots, \boldsymbol{\alpha}_K, \beta_K)$ within a given component. In particular, $f_{c,A,b}(\boldsymbol{\alpha}_1, \beta_1, \dots, \boldsymbol{\alpha}_K, \beta_K)$ is invariant over each connected component.*

Proof. Fix a connected component C in the decomposition induced by the set of hyperplanes, degree- K hypersurfaces, and degree- $2K$ hypersurfaces established in Lemma C.5. Let

$$Q_1, \dots, Q_{i_1}, I_1, Q_{i_1+1}, \dots, Q_{i_2}, I_2, Q_{i_2+1}, \dots \quad (15)$$

denote the nodes of the tree branch-and-cut creates, in order of exploration, under the assumption that a node is pruned if and only if either the LP at that node is infeasible or the LP optimal solution is integral (so the ‘‘bounding’’ of branch-and-bound is suppressed). Here, a node is identified by the list σ of branching constraints added to the input IP. Nodes labeled by Q are either infeasible or have fractional LP optimal solutions. Nodes labeled by I have integral LP optimal solutions and are candidates for the incumbent integral solution at the point they are encountered. (The nodes are functions of $\boldsymbol{\alpha}_1, \beta_1, \dots, \boldsymbol{\alpha}_K, \beta_K$, as are the indices i_1, i_2, \dots) By Lemma C.5, this ordered list of nodes is invariant for all $(\boldsymbol{\alpha}_1, \beta_1, \dots, \boldsymbol{\alpha}_K, \beta_K) \in C$.

Now, given an node index ℓ , let $I(\ell)$ denote the incumbent node with the highest objective value encountered up until the ℓ th node searched by B&C, and let $z(I(\ell))$ denote its objective value. For each node Q_ℓ , let σ_ℓ denote the branching constraints added to arrive at node Q_ℓ . The hyperplane

$$z_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_K^T \mathbf{x} \leq \beta_K, \sigma_\ell) = z(I(\ell)) \quad (16)$$

(which is a hyperplane due to Corollary C.3) partitions C into two subregions. In one subregion, $z_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_k^T \mathbf{x} \leq \beta_k, \sigma_\ell) \leq z(I(\ell))$, that is, the objective value of the LP optimal solution is no greater than the objective value of the current incumbent integer solution, and so the subtree rooted at Q_ℓ is pruned. In the other subregion, $z_{\text{LP}}^*(\alpha_1^T \mathbf{x} \leq \beta_1, \dots, \alpha_k^T \mathbf{x} \leq \beta_k, \sigma_\ell) > z(I(\ell))$, and Q_ℓ is branched on further. Therefore, within each connected component of C induced by all hyperplanes given by Equation 16 for all ℓ , the set of node within the list (15) that are pruned is invariant. Combined with the surfaces established in Lemma C.5, these hyperplanes partition $\mathbb{R}^{K(n+1)}$ into connected components such that as $(\alpha_1, \beta_1, \dots, \alpha_K, \beta_K)$ varies within a given component, the tree built by branch-and-cut is invariant.

Finally, we count the total number of surfaces inducing this partition. Unlike the counting stages of the previous lemmas, we will first have to count the number of connected components induced by the surfaces established in Lemma C.5. This is because the ordered list of nodes explored by branch-and-cut (15) can be different across each component, and the hyperplanes given by Equation 16 depend on this list. From Lemma C.5 we have $6n^2 K^n (m+2n)^{2n} \tau^{4n}$ polynomial hypersurfaces of degree $\leq 2K$. The set of all $(\alpha_1, \beta_1, \dots, \alpha_K, \beta_K) \in \mathbb{R}^{K(n+1)}$ such that $(\alpha_1, \beta_1, \dots, \alpha_K, \beta_K)$ lies on the boundary of any of these surfaces is precisely the zero set of the product of all polynomials defining these surfaces. Denote this product polynomial by p . The degree of the product polynomial is the sum of the degrees of $6n^2 K^n (m+2n)^{2n} \tau^{4n}$ polynomials of degree $\leq 2K$, which is at most $2K \cdot 6n^2 K^n (m+2n)^{2n} \tau^{4n} = 12n^2 K^{n+2} (m+2n)^{2n} \tau^{4n}$. By Warren's theorem, the number of connected components of $\mathbb{R}^{n+1} \setminus \{(\alpha, \beta) : p(\alpha, \beta) = 0\}$ is $O((12n^2 K^{n+2} (m+2n)^{2n} \tau^{4n})^{n-1})$, and by the Milnor-Thom theorem, the number of connected components of $\{(\alpha, \beta) : p(\alpha, \beta) = 0\}$ is $O((12n^2 K^{n+2} (m+2n)^{2n} \tau^{4n})^{n-1})$ as well. So, the number of connected components induced by the surfaces in Lemma C.5 is $O(12n^2 n^{2n} K^{2n^2} (m+2n)^{2n^2} \tau^{4n^2})$. For every connected component C in Lemma C.5, the closed form of $z_{\text{LP}}^*(\alpha^T \mathbf{x} \leq \beta, \sigma_\ell)$ is already determined due to Corollary C.3, and so the number of hyperplanes given by Equation 16 is at most the number of possible $\sigma \subseteq \mathcal{BC}$, which is at most τ^{3n} . So across all connected components C , the total number of hyperplanes given by Equation 16 is $O(12n^2 n^{2n} K^{2n^2} (m+2n)^{2n^2} \tau^{5n^2})$. Finally, adding this to the surface-counts established in Lemma C.5 yields the theorem statement. \square

D Omitted results from Section 5

Proof of Theorem 5.1. For a set \mathcal{X} , $\mathcal{X}^{<\mathbb{N}}$ denotes the set of finite sequences of elements from \mathcal{X} . There is a bijection between the set of IPs $(c, A, \mathbf{b}) \in \mathcal{I} := \mathbb{R}^n \times \mathbb{Z}^{m \times n} \times \mathbb{Z}^m$ and \mathbb{R} , so IPs can be uniquely represented as real numbers (and vice versa). Now, consider the set of all finite sequences of pairs of IPs and ± 1 labels of the form $((\mathbf{c}_1, A_1, \mathbf{b}_1), \varepsilon_1), \dots, ((\mathbf{c}_N, A_N, \mathbf{b}_N), \varepsilon_N)$, $\varepsilon_1, \dots, \varepsilon_N \in \{-1, 1\}$, that is, the set $(\mathcal{I} \times \{-1, 1\})^{<\mathbb{N}}$. There is a bijection between this set and $(\mathbb{R} \times \{-1, 1\})^{<\mathbb{N}}$, and in turn there is a bijection between $(\mathbb{R} \times \{-1, 1\})^{<\mathbb{N}}$ and \mathbb{R} . Hence, there exists a bijection between \mathcal{U} and $(\mathcal{I} \times \{-1, 1\})^{<\mathbb{N}}$. Fix such a bijection $\varphi : \mathcal{U} \rightarrow (\mathcal{I} \times \{-1, 1\})^{<\mathbb{N}}$, and let $\varphi^{-1} : (\mathcal{I} \times \{-1, 1\})^{<\mathbb{N}} \rightarrow \mathcal{U}$ denote the inverse of φ , which is well defined and also a bijection.

Let n be odd. For $c \in \mathbb{R}$, let $\text{IP}_c \in \mathcal{I}$ denote the IP

$$\begin{aligned} & \text{maximize} && c \\ & \text{subject to} && 2x_1 + \dots + 2x_n = n \\ & && \mathbf{x} \in \{0, 1\}^n. \end{aligned} \quad (17)$$

Since n is odd, IP_c is infeasible, independent of c . Jeroslow [32] showed that without the use of cutting planes or heuristics, branch-and-bound builds a tree of size $2^{(n-1)/2}$ before determining infeasibility and terminating. The objective c is irrelevant, but is important in generating distinct IPs with this property. Consider the cut $x_1 + \dots + x_n \leq \lfloor n/2 \rfloor$, which is a valid cut for IP_c (this is in fact a Chvátal-Gomory cut [10]). In particular, since n is odd, $x_1 + \dots + x_n \leq \lfloor n/2 \rfloor \implies x_1 + \dots + x_n \leq (n-1)/2 < n/2$, so the equality constraint of IP_c is violated by this cut. Thus, the feasible region of

the LP relaxation after adding this cut is empty, and branch-and-bound will terminate immediately at the root (building a tree of size 1). Denote this cut by $(\alpha^{(-1)}, \beta^{(-1)}) = (\mathbf{1}, \lfloor n/2 \rfloor)$. On the other hand, let $(\alpha^{(1)}, \beta^{(1)}) = (\mathbf{0}, 0)$ be the trivial cut $0 \leq 0$. Adding this cut to the IP constraints does not change the feasible region, so branch-and-bound will build a tree of size $2^{(n-1)/2}$.

We now define $\alpha_{\mathbf{c}, A, \mathbf{b}}$ and $\beta_{\mathbf{c}, A, \mathbf{b}}$. Let

$$(\alpha_{\mathbf{c}, A, \mathbf{b}}(\mathbf{u}), \beta_{\mathbf{c}, A, \mathbf{b}}(\mathbf{u})) = \begin{cases} (\alpha^{(1)}, \beta^{(1)}) & \text{if } ((\mathbf{c}, A, \mathbf{b}), 1) \in \varphi(\mathbf{u}) \text{ and } ((\mathbf{c}, A, \mathbf{b}), -1) \notin \varphi(\mathbf{u}) \\ (\alpha^{(-1)}, \beta^{(-1)}) & \text{if } ((\mathbf{c}, A, \mathbf{b}), -1) \in \varphi(\mathbf{u}) \text{ and } ((\mathbf{c}, A, \mathbf{b}), 1) \notin \varphi(\mathbf{u}) \\ (\mathbf{0}, 0) & \text{otherwise} \end{cases}.$$

The choice to use $(\mathbf{0}, 0)$ in the case that either $((\mathbf{c}, A, \mathbf{b}), \varepsilon) \notin \varphi(\mathbf{u})$ for each $\varepsilon \in \{-1, 1\}$, or $((\mathbf{c}, A, \mathbf{b}), -1) \in \varphi(\mathbf{u})$ and $((\mathbf{c}, A, \mathbf{b}), 1) \in \varphi(\mathbf{u})$ is arbitrary and unimportant. Now, for any integer $N > 0$, constructing a set of N IPs and N thresholds that is shattered is almost immediate. Let $c_1, \dots, c_N \in \mathbb{R}$ be distinct reals, and let $1 < r_1, \dots, r_N < 2^{(n-1)/2}$. Then, the set $\{(\text{IP}_{c_1}, r_1), \dots, (\text{IP}_{c_N}, r_N)\}$ can be shattered. Indeed, given a sign pattern $(\varepsilon_1, \dots, \varepsilon_N) \in \{-1, 1\}^N$, let

$$\mathbf{u} = \varphi^{-1}((\text{IP}_{c_1}, \varepsilon_1), \dots, (\text{IP}_{c_N}, \varepsilon_N)).$$

Then, if $\varepsilon_i = 1$, $(\alpha_{\text{IP}_{c_i}}(\mathbf{u}), \beta_{\text{IP}_{c_i}}(\mathbf{u})) = (\alpha^{(1)}, \beta^{(1)})$, so $g_{\mathbf{u}}(\text{IP}_{c_i}) = 2^{(n-1)/2}$ and $\text{sign}(g_{\mathbf{u}}(\text{IP}_{c_i}) - r_i) = 1$. If $\varepsilon_i = -1$, $(\alpha_{\text{IP}_{c_i}}(\mathbf{u}), \beta_{\text{IP}_{c_i}}(\mathbf{u})) = (\alpha^{(-1)}, \beta^{(-1)})$, so $g_{\mathbf{u}}(\text{IP}_{c_i}) = 1$ and $\text{sign}(g_{\mathbf{u}}(\text{IP}_{c_i}) - r_i) = -1$. So for any N there is a set of IPs and thresholds that can be shattered, which yields the theorem statement. \square

Lemma D.1. *Consider the family of GMI cuts parameterized by $\mathbf{u} \in [-U, U]^m$. There is a set of at most $O(nU^2 \|A\|_1 \|\mathbf{b}\|_1)$ hyperplanes partitioning $[-U, U]^m$ into connected components such that $\lfloor \mathbf{u}^T \mathbf{a}_i \rfloor$, $\lfloor \mathbf{u}^T \mathbf{b} \rfloor$, and $\mathbf{1}[f_i \leq f_0]$ are invariant, for every i , within each component.*

Proof of Lemma D.1. We have $f_i = \mathbf{u}^T \mathbf{a}_i - \lfloor \mathbf{u}^T \mathbf{a}_i \rfloor$, $f_0 = \mathbf{u}^T \mathbf{b} - \lfloor \mathbf{u}^T \mathbf{b} \rfloor$, and since $\mathbf{u} \in [-U, U]^m$, $\lfloor \mathbf{u}^T \mathbf{a}_i \rfloor \in [-U \|\mathbf{a}_i\|_1, U \|\mathbf{a}_i\|_1]$ and $\lfloor \mathbf{u}^T \mathbf{b} \rfloor \in [-U \|\mathbf{b}\|_1, U \|\mathbf{b}\|_1]$. Now, for all i , $k_i \in [-U \|\mathbf{a}_i\|_1, U \|\mathbf{a}_i\|_1] \cap \mathbb{Z}$ and $k_0 \in [-U \|\mathbf{b}\|_1, U \|\mathbf{b}\|_1] \cap \mathbb{Z}$, put down the hyperplanes defining the two halfspaces

$$\lfloor \mathbf{u}^T \mathbf{a}_i \rfloor = k_i \iff k_i \leq \mathbf{u}^T \mathbf{a}_i < k_i + 1 \quad (18)$$

and the hyperplanes defining the two halfspaces

$$\lfloor \mathbf{u}^T \mathbf{b} \rfloor = k_0 \iff k_0 \leq \mathbf{u}^T \mathbf{b} < k_0 + 1. \quad (19)$$

In addition, consider the hyperplane

$$\mathbf{u}^T \mathbf{a}_i - k_i = \mathbf{u}^T \mathbf{b} - k_0 \quad (20)$$

for each i . Within any connected component of \mathbb{R}^m determined by these hyperplanes, $\lfloor \mathbf{u}^T \mathbf{a}_i \rfloor$ and $\lfloor \mathbf{u}^T \mathbf{b} \rfloor$ are constant. Furthermore, $\mathbf{1}[f_i \leq f_0]$ is invariant within each connected component, since if $\lfloor \mathbf{u}^T \mathbf{a}_i \rfloor = k_i$ and $\lfloor \mathbf{u}^T \mathbf{b} \rfloor = k_0$, $f_i \leq f_0 \iff \mathbf{u}^T \mathbf{a}_i - k_i \leq \mathbf{u}^T \mathbf{b} - k_0$, which is the hyperplane given by Equation 20. The total number of hyperplanes of type 18 is $O(nU \|A\|_1)$, the total number of hyperplanes of type 19 is $O(U \|\mathbf{b}\|_1)$, and the total number of hyperplanes of type 20 is $nU^2 \|A\|_1 \|\mathbf{b}\|_1$. Summing yields the lemma statement. \square

The next lemma allows us to transfer the polynomial partition of \mathbb{R}^{n+1} from Theorem 4.4 to a polynomial partition of $[-U, U]^m$, incurring only a factor 2 increase in degree.

Lemma D.2. *Let $p \in \mathbb{R}[y_1, \dots, y_{n+1}]$ be a polynomial of degree d . Let $D \subseteq [-U, U]^m$ be a connected component from Lemma D.1. Define $q : D \rightarrow \mathbb{R}$ by $q(\mathbf{u}) = p(\alpha(\mathbf{u}), \beta(\mathbf{u}))$. Then q is a polynomial in \mathbf{u} of degree $2d$.*

Proof. By Lemma D.1, there are integers k_0, k_i for $i \in [n]$ such that $\lfloor \mathbf{u}^T \mathbf{a}_i \rfloor = k_i$ and $\lfloor \mathbf{u}^T \mathbf{b} \rfloor = k_0$ for all $\mathbf{u} \in D$. Also, the set $S = \{i : f_i \leq f_0\}$ is fixed over all $\mathbf{u} \in D$.

A degree- d polynomial p in variables y_1, \dots, y_{n+1} can be written as $\sum_{T \subseteq [n+1], |T| \leq d} \lambda_T \prod_{i \in T} y_i$ for some coefficients $\lambda_T \in \mathbb{R}$, where $T \subseteq [n+1]$ means that T is a multiset of $[n+1]$. Evaluating at $(\boldsymbol{\alpha}(\mathbf{u}), \beta(\mathbf{u}))$, we get

$$\sum_{|T| \leq d} \lambda_T \prod_{\substack{i \in T \cap S \\ i \neq n+1}} f_i(1 - f_0) \prod_{\substack{i \in T \setminus S \\ i \neq n+1}} f_0(1 - f_i) \prod_{\substack{i \in T \\ i = n+1}} f_0(1 - f_0).$$

Now, $f_i = \mathbf{u}^T \mathbf{a}_i - k_i$ and $f_0 = \mathbf{u}^T \mathbf{b} - k_0$ are linear in \mathbf{u} . The sum is over all multisets of size at most d , so each monomial consists of the product of at most d degree-2 terms of the form $f_i(1 - f_0)$, $f_0(1 - f_i)$, or $f_0(1 - f_0)$. Thus, $\deg(q) \leq 2d$, as desired. \square

Proof of Lemma 5.2. Let $C \subseteq \mathbb{R}^{n+1}$ be a connected component in the partition established in Theorem 4.4, so C can be written as the intersection of at most $14^n(m+2n)^{3n^2} \tau^{5n^2}$ polynomial constraints of degree at most 5. Let $D \subseteq [-U, U]^m$ be a connected component in the partition established in Lemma D.1. By Lemma D.2, there are at most $14^n(m+2n)^{3n^2} \tau^{5n^2}$ polynomials of degree at most 10 partitioning D into connected components such that within each component, $\mathbf{1}[(\boldsymbol{\alpha}(\mathbf{u}), \beta(\mathbf{u})) \in C]$ is invariant. If we consider the overlay of these polynomial surfaces over all components C , we will get a partition of $[-U, U]^m$ such that for every C , $\mathbf{1}[(\boldsymbol{\alpha}(\mathbf{u}), \beta(\mathbf{u})) \in C]$ is invariant over each connected component of $[-U, U]^m$. Once we have this we are done, since all \mathbf{u} in the same connected component of $[-U, U]^m$ will be sent to the same connected component of \mathbb{R}^{n+1} by $(\boldsymbol{\alpha}(\mathbf{u}), \beta(\mathbf{u}))$, and thus by Theorem 4.4 the behavior of branch-and-cut will be invariant.

We now tally up the total number of surfaces. The number of connected components C was given by Warren's theorem and the Milnor-Thom theorem to be $O(14^{n(n+1)}(m+2n)^{3n^2(n+1)} \tau^{5n^2(n+1)})$, so the total number of degree-10 hypersurfaces is $14^n(m+2n)^{3n^2} \tau^{5n^2}$ times this quantity, which yields the lemma statement. \square

D.1 Multiple GMI cuts at the root

In this section we extend our results to allow for multiple GMI cuts at the root of the B&C tree. These cuts can be added simultaneously, sequentially, or in rounds. If GMI cuts $\mathbf{u}_1, \mathbf{u}_2$ are added simultaneously, both of them have the same dimension and are defined in the usual way. If GMI cuts $\mathbf{u}_1, \mathbf{u}_2$ are added sequentially, \mathbf{u}_2 has one more entry than \mathbf{u}_1 . This is because when cuts are added sequentially, the LP relaxation is re-solved after the addition of the first cut, and the second cut has a multiplier for all original constraints as well as for the first cut (this ensures that the second cut can be chosen in a more informed manner). If K cuts are made at the root, they can be added in sequential rounds of simultaneous cuts. In the following discussion, we focus on the case where all K cuts are added sequentially—the other cases can be viewed as instantiations of this. We refer the reader to the discussion in Balcan et al. [10] for more details.

To prove an analogous result for multiple GMI cuts (in sequence, that is, each successive GMI cut has one more parameter than the previous), we combine the reasoning used in the single-GMI-cut case with some technical observations in Balcan et al. [10].

Lemma D.3. *Consider the family of K sequential GMI cuts parameterized by $\mathbf{u}_1 \in [-U, U]^m$, $\mathbf{u}_2 \in [-U, U]^{m+1}, \dots, \mathbf{u}_K \in [-U, U]^{m+K-1}$. For any IP $(\mathbf{c}, A, \mathbf{b})$, there are at most*

$$O(nK(1+U)^{2K} \|A\|_1 \|\mathbf{b}\|_1)$$

degree- K polynomial hypersurfaces and

$$2^{O(n^2)} K^{O(n^3)} (m+2n)^{O(n^3)} \tau^{O(n^3)}$$

degree- $4K^2$ polynomial hypersurfaces partitioning $[-U, U]^m \times \dots \times [-U, U]^{m+K-1}$ connected components such that the B&C tree built after sequentially adding the GMI cuts defined by $\mathbf{u}_1, \dots, \mathbf{u}_K$ is invariant over all $(\mathbf{u}_1, \dots, \mathbf{u}_K)$ within a single component.

Proof. We start with the setup used by Balcan et al. [10] to prove similar results for sequential Chvátal-Gomory cuts. Let $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ be the columns of A . We define the following augmented

columns $\tilde{\mathbf{a}}_i^1 \in \mathbb{R}^m, \dots, \tilde{\mathbf{a}}_i^K \in \mathbb{R}^{m+K-1}$ for each $i \in [n]$, and the augmented constraint vectors $\tilde{\mathbf{b}}^1 \in \mathbb{R}^m, \dots, \tilde{\mathbf{b}}^K \in \mathbb{R}^{m+K-1}$ via the following recurrences:

$$\begin{aligned}\tilde{\mathbf{a}}_i^1 &= \mathbf{a}_i \\ \tilde{\mathbf{a}}_i^k &= \begin{bmatrix} \tilde{\mathbf{a}}_i^{k-1} \\ \mathbf{u}_{k-1}^T \tilde{\mathbf{a}}_i^{k-1} \end{bmatrix}\end{aligned}$$

and

$$\begin{aligned}\tilde{\mathbf{b}}^1 &= \mathbf{b} \\ \tilde{\mathbf{b}}^k &= \begin{bmatrix} \tilde{\mathbf{b}}^{k-1} \\ \mathbf{u}_{k-1}^T \tilde{\mathbf{b}}^{k-1} \end{bmatrix}\end{aligned}$$

for $k = 2, \dots, K$. In other words, $\tilde{\mathbf{a}}_i^k$ is the i th column of the constraint matrix of the IP and $\tilde{\mathbf{b}}^k$ is the constraint vector after applying cuts $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$. An identical induction argument to that of Balcan et al. [10] shows that for each $k \in [K]$,

$$[\mathbf{u}_k^T \tilde{\mathbf{a}}_i^k] \in \left[-(1+U)^k \|\mathbf{a}_i\|_1, (1+U)^k \|\mathbf{a}_i\|_1 \right]$$

and

$$[\mathbf{u}_k^T \tilde{\mathbf{b}}^k] \in \left[-(1+U)^k \|\mathbf{b}\|_1, (1+U)^k \|\mathbf{b}\|_1 \right].$$

Now, as in the single-GMI-cut setting, consider the surfaces

$$[\mathbf{u}_k^T \tilde{\mathbf{a}}_i^k] = \ell_i \iff \ell_i \leq \mathbf{u}_k^T \tilde{\mathbf{a}}_i^k < \ell_i + 1 \quad (21)$$

and

$$[\mathbf{u}_k^T \tilde{\mathbf{b}}^k] = \ell_0 \iff \ell_0 \leq \mathbf{u}_k^T \tilde{\mathbf{b}}^k < \ell_0 + 1 \quad (22)$$

for every i, k , and every integer $\ell_i \in [-(1+U)^k \|\mathbf{a}_i\|_1, (1+U)^k \|\mathbf{a}_i\|_1] \cap \mathbb{Z}$ and every integer $\ell_0 \in [-(1+U)^k \|\mathbf{b}\|_1, (1+U)^k \|\mathbf{b}\|_1] \cap \mathbb{Z}$. In addition, consider the surfaces

$$\mathbf{u}_k^T \tilde{\mathbf{a}}_i^k - \ell_i = \mathbf{u}_k^T \tilde{\mathbf{b}}^k - \ell_0 \quad (23)$$

for each i, k, ℓ_i, ℓ_0 . As observed by Balcan et al. [10], $\mathbf{u}_k^T \tilde{\mathbf{a}}_i^k$ is a polynomial in $\mathbf{u}_1[1], \dots, \mathbf{u}_1[m], \mathbf{u}_2[1], \dots, \mathbf{u}_2[m+1], \dots, \mathbf{u}_k[1], \dots, \mathbf{u}_k[m+k-1]$ of degree at most k (as is $\mathbf{u}_k^T \tilde{\mathbf{b}}^k$), so surfaces 21, 22, and 23 are all degree- K polynomial hypersurfaces for all i, k . Within any connected component of $[-U, U]^m \times \dots \times [-U, U]^{m+K-1}$ induced by these hypersurfaces, $[\mathbf{u}_k^T \tilde{\mathbf{a}}_i^k]$ and $[\mathbf{u}_k^T \tilde{\mathbf{b}}^k]$ are constant. Furthermore $\mathbf{1}[f_i^k \leq f_0^k]$ is invariant for every i, k , where $f_i^k = \mathbf{u}_k^T \tilde{\mathbf{a}}_i^k - [\mathbf{u}_k^T \tilde{\mathbf{a}}_i^k]$ and $f_0^k = \mathbf{u}_k^T \tilde{\mathbf{b}}^k - [\mathbf{u}_k^T \tilde{\mathbf{b}}^k]$.

Now, fix a connected component $D \subseteq [-U, U]^m \times \dots \times [-U, U]^{m+K-1}$ induced by the above hypersurfaces, and let $C \subseteq \mathbb{R}^{K(n+1)}$ be the intersection of q polynomial inequalities of degree at most d . Consider a single degree- d polynomial inequality in $K(n+1)$ variables $y_1, \dots, y_{K(n+1)}$, which can be written as

$$\sum_{\substack{T \subseteq [K(n+1)] \\ |T| \leq d}} \lambda_T \prod_{j \in T} y_j = \sum_{\substack{T_1, \dots, T_K \subseteq [n+1] \\ |T_1| + \dots + |T_K| \leq d}} \lambda_{T_1, \dots, T_K} \prod_{j_1 \in T_1} y_{j_1} \cdots \prod_{j_K \in T_K} y_{j_K} \leq \gamma.$$

Now, the sets S_1, \dots, S_K defined by $S_k = \{i : f_i^k \leq f_0^k\}$ are fixed within D , so we can write this as

$$\sum_{\substack{T_1, \dots, T_K \subseteq [n+1] \\ |T_1| + \dots + |T_K| \leq d}} \lambda_{T_1, \dots, T_K} \prod_{k=1}^K \left[\prod_{\substack{j \in T_k \cap S_k \\ j \neq n+1}} f_j^k (1 - f_0^k) \prod_{\substack{j \in T_k \setminus S_k \\ j \neq n+1}} f_0^k (1 - f_j^k) \prod_{\substack{j \in T_k \\ j = n+1}} f_0^k (1 - f_0^k) \right] \leq \gamma.$$

We have that f_j^k and f_0^k are degree- k polynomials in $\mathbf{u}_1, \dots, \mathbf{u}_k$. Since the sum is over all multisets T_1, \dots, T_K such that $|T_1| + \dots + |T_K| \leq d$, there are at most d terms across the products, each of the form $f_j^k (1 - f_0^k)^k, f_0^k (1 - f_j^k)^k$, or $f_0^k (1 - f_0^k)^k$. Therefore, the left-hand-side is a polynomial of

degree at most $2dK$, and if $C \subseteq \mathbb{R}^{K(n+1)}$ is the intersection of q polynomial inequalities each of degree at most d , the set

$$\{(\mathbf{u}_1, \dots, \mathbf{u}_K) \in D : (\boldsymbol{\alpha}(\mathbf{u}_1, \dots, \mathbf{u}_K), \beta(\mathbf{u}_1, \dots, \mathbf{u}_K)) \in C\} \subseteq [-U, U]^m \times \dots \times [-U, U]^{m+K-1}$$

can be expressed as the intersection of q degree- $2dK$ polynomial inequalities.

To finish, we run this process for every connected component $C \subseteq \mathbb{R}^{K(n+1)}$ in the partition established by Theorem C.6. This partition consists of $O(12^n n^{2n} K^{2n^2} (m+2n)^{2n^2} \tau^{5n^2})$ degree- $2K$ polynomials over $\mathbb{R}^{K(n+1)}$. By Warren's theorem and the Milnor-Thom theorem, these polynomials partition $\mathbb{R}^{K(n+1)}$ into $O(12^{n(n+1)} n^{2n(n+1)} K^{2n^2(n+1)} (m+2n)^{2n^2(n+1)} \tau^{5n^2(n+1)})$ connected components. Running the above argument for each of these connected components of $\mathbb{R}^{K(n+1)}$ yields a total of $O\left(12^{n(n+1)} n^{2n(n+1)} K^{2n^2(n+1)} (m+2n)^{2n^2(n+1)} \tau^{5n^2(n+1)}\right)$.

$O\left(12^n n^{2n} K^{2n^2} (m+2n)^{2n^2} \tau^{5n^2}\right) = 2^{O(n^2)} K^{O(n^3)} (m+2n)^{O(n^3)} \tau^{O(n^3)}$ polynomials of degree $4K^2$. Finally, we count the surfaces of the form (21), (22), and (23). The total number of degree- K polynomials of type 21 is at most $O(nK(1+U)^K \|A\|_1)$, the total number of degree- k polynomials of type 22 is $O(K(1+U)^K \|\mathbf{b}\|_1)$, and the total number of degree- K polynomials of type 23 is $O(nK(1+U)^{2K} \|A\|_1 \|\mathbf{b}\|_1)$. Summing these counts yields the desired number of surfaces in the lemma statement.

In any connected component of $[-U, U]^m$ determined by these surfaces, $\mathbf{1}[(\boldsymbol{\alpha}(\mathbf{u}), \beta(\mathbf{u})) \in C]$ is invariant for every connected component $C \subseteq \mathbb{R}^{K(n+1)}$ in the partition of $\mathbb{R}^{K(n+1)}$ established in Theorem C.6. This means that the tree built by branch-and-cut is invariant, which concludes the proof. \square

Finally, applying the main result of Balcan et al. [9] to Lemma D.3, we get the following pseudo-dimension bound for the class of K sequential GMI cuts at the root of the B&C tree.

Theorem D.4. *For $\mathbf{u}_1 \in [-U, U]^m, \mathbf{u}_2 \in [-U, U]^{m+1}, \dots, \mathbf{u}_K \in [-U, U]^{m+K-1}$, let $g_{\mathbf{u}_1, \dots, \mathbf{u}_K}(\mathbf{c}, A, \mathbf{b})$ denote the number of nodes in the tree B&C builds given the input $(\mathbf{c}, A, \mathbf{b})$ after sequentially applying the GMI cuts defined by $\mathbf{u}_1, \dots, \mathbf{u}_K$ at the root. The pseudo-dimension of the set of functions $\{g_{\mathbf{u}_1, \dots, \mathbf{u}_K} : (\mathbf{u}_1, \dots, \mathbf{u}_K) \in [-U, U]^m \times \dots \times [-U, U]^{m+K-1}\}$ on the domain of IPs with $\|A\|_1 \leq a$ and $\|\mathbf{b}\|_1 \leq b$ is*

$$O(mK^3 \log U + mn^3 K^2 \log(mnK\tau) + mK^2 \log(ab)).$$