Polyhedral Analysis of Symmetric Multilinear Polynomials over Box Constraints

Yibo Xu*   Warren Adams†   Akshay Gupte‡

April 29, 2021

Abstract

It is well-known that the convex and concave envelope of a multilinear polynomial over a box are polyhedral functions. Exponential-sized extended and projected formulations for these envelopes are also known. We consider the convexification question for multilinear polynomials that are symmetric with respect to permutations of variables. Such a permutation-invariant structure naturally implies a quadratic-sized extended formulation for the envelopes through the use of disjunctive programming. The optimization and separation problems are answered directly without using this extension. The problem symmetry allows the optimization and separation problems to be answered directly without using any extension. It also implies that permuting the coefficients of a core set of facets generates all the facets. We provide some necessary conditions and some sufficient conditions for a valid inequality to be a core facet. These conditions are applied to obtain envelopes for two classes: symmetric supermodular functions and multilinear monomials with reflection symmetry, thereby yielding alternate proofs to the literature. Furthermore, we use constructs from the reformulation-linearization-technique to completely characterize the set of points lying on each facet.

Keywords. Convex Hull, Permutation Invariance, Kuhn's Triangulation, Submodular and Supermodular, Reformulation-Linearization Technique

AMS 2020 subject classification. 90C23, 90C57, 05E05, 52B15

1 Introduction

A polynomial is multilinear if every monomial is square-free in the sense that it is a product of a subset of variables raised to the power one. Multilinear polynomials have degree n and they become linear functions when $n-1$ variables are fixed (hence the name). A multilinear polynomial can be expressed as $\sum_{J \subseteq N} c_J \prod_{j \in J} x_j$ for some $c \in \mathbb{R}^{2^n}$. We are interested in symmetric multilinear polynomials (SMPs) in this paper, by which we mean the polynomial

$$m(x) = \sum_{i=2}^{n} c_i \sum_{\substack{J \subseteq N \\ |J|=i}} \prod_{j \in J} x_j, \tag{1a}$$

---

*Department of Mathematical Sciences, Rensselaer Polytechnic Institute, USA. Email: `xuy24@rpi.edu`.

†Engineering and Information Science Branch (RTA), Air Force Office of Scientific Research, USA. Email: `warren.adams.2@us.af.mil`.

‡School of Mathematics, University of Edinburgh, UK. Email: `akshay.gupte@ed.ac.uk`.

where we have dropped the constant and linear terms since they are inconsequential to us from the point of view of convexification. The symmetry of m refers to the fact that for every $x \in \mathbb{R}^n$ and $\bar{x}$ equal to a permutation of x, we have $m(x) = m(\bar{x})$. Our goal is to study the convex hull of the graph of an SMP over a symmetric box in $\mathbb{R}^n$. Namely, we study $\operatorname{conv} G$ which is the convex hull of the set

$$G = \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R} : x \in X, \ y = m(x) \right\}, \tag{1b}$$

where X is a box imposing the same lower and upper bounds (finite $\ell < u$) on all variables,

$$X = \left\{ x \in \mathbb{R}^n : \ell \leqslant x_j \leqslant u, \ j \in N \right\}. \tag{1c}$$

Throughout this paper, we use $N := \{1, \ldots, n\}$. Although coordinatewise scaling and translation does not break the symmetry of $m(x)$ and can reduce the box X to be the unit hypercube $[0, 1]^n$, we work with arbitrary ℓ and u to avoid this affine transformation which can be cumbersome to perform.

The graph $G_p = \{(x, y) \in B \times \mathbb{R} : y = p(x)\}$ of a general multilinear polynomial $p(x) = \sum_{J \subseteq N} c_J \prod_{j \in J} x_j$ over an arbitrary box $B \subset \mathbb{R}^n$ appears not only as a substructure in some important applications but also when optimizing a polynomial over binary variables [DK16], which is equivalent to pseudo-Boolean optimization [BH02], and as an intermediate set when performing factorable reformulations of general polynomial optimization problems [BR08; Bao+15; DKS20]. Hence, convexification of G_p has been the subject of many studies in the literature. It is known that this convex hull is a polytope and exponential-sized extended formulations are available [Rik97; She97; BM14] but an explicit description in the (x, y)-space is not known in general. Earlier studies focused on convexifying multilinear monomials [RS01; Ben04; MF04; CLL10], motivated by the classical linearization of a monomial $x_1 x_2 \ldots x_n$ over $[0, 1]^n$ [GW74] which defines the convex hull for the graph of the monomial [Cra93] and leads to the standard linearization for G_p. Separation over G_p is NP-hard, and so different classes of valid inequalities and cutting plane procedures have been developed for use in global optimization algorithms [Bao+15; FKL15; CR17; DKS20]. There have also been many recent studies on describing G_p or generalizations of it in the monomial space, which is obtained by adding a new variable for each monomial, under different assumptions on the structure of the polynomial [LNL12; DK16; DK18a; DK18b; FFM18; BCR19; CDG20; FFM20; Gup+20].

Symmetry has not been exploited in the rich body of literature on convexifying multilinear polynomials. The main objective of this paper is to initiate a systematic polyhedral analysis of $\operatorname{conv} G$ by exploiting symmetry in this set. Our focus is on the minimal inequality description of this full-dimensional polytope in the original (x, y)-space, as opposed to the many studies in the literature about convexifying general multilinear polynomials in the monomial space. Note that there will be exponentially many monomials in $m(x)$ when $c_{\frac{n}{k}} \neq 0$ for some constant $k \geqslant 1$, and so a reformulation to the monomial space will not always be tractable. The symmetric structure we assume is interesting not only because it enables a thorough analysis of the convex hull and hence adds to the convexification literature, but also because it arises in many applications of combinatorial optimization [Ant+16; KTR19; BCR20; DMW20] and with regards to chromatic number of graphs and other areas of combinatorics [cf. Sta95; Eag+20].

1.1 Our Contributions

There are exponentially many facets, but symmetry of the function m and of the set X means that we only need to focus on a certain subset of inequalities, which we call core inequalities, and

all other inequalities are generated as permutations of coefficients in core inequalities. In theory, all these inequalities can be obtained by projecting an extended formulation having $O(n^2)$ many variables and $O(n)$ many constraints, which is much smaller in size than the exponential-sized extensions for general multilinear polynomials. However, projecting this extension is a tedious task due to the combinatorial explosion that generally occurs with the projection operation. Instead, we give several necessary conditions and some sufficient conditions for a core valid inequality to be facet-defining, and these are more tractable to verify than the conditions that come from the use of polarity since the latter require enumeration of extreme points of a polyhedron. They can be applied to certify whether a given description of $\mathsf{conv}\, G$ is minimal or not. In that regard, we use our conditions to obtain explicit listing of all the facet-defining inequalities for two families of SMPs. The first family is that of supermodular SMPs and the second family is that of monomials whose lower and upper bounds are reflections of each other $(-\ell = u > 0)$.

The Reformulation-Linearization Technique (RLT) is known to convexify the graph of a general multilinear polynomial over an arbitrary box. We show that RLT also implies that such a polynomial is nonnegative on a box if and only if it is nonnegative at every vertex of the box. This consequence enables us to characterize the set of points on G that lie on a facet of the convex hull of G. The sets, by derivation, turn out to be different unions of d-dimensional $(0 \leqslant d \leqslant n - 1)$ faces of $\mathsf{conv}\, G$.

The questions of optimization and separation are also answered for $\mathsf{conv}\, G$ for any SMP without using the quadratic-sized extension. A linear function can be optimized over $\mathsf{conv}\, G$ in $O(n \log n)$ time (assuming the value of function m at a vertex of X can be computed in $O(1)$ time) without using the extended formulation. Thus, a point can be separated from $\mathsf{conv}\, G$ in polynomial time via the ellipsoid method. There is also a direct separation algorithm that runs in $O(nt + n \log n)$ time where t is the number of core inequalities, and hence has polynomial time complexity when t is bounded by a polynomial in input size.

1.1.1   Related Results in Literature

Although we recognize that our convex hull descriptions for the two special families have been established before in literature, our necessary conditions for core facets of $\mathsf{conv}\, G$ yield alternate proofs for them. In this context, the following results are known. A set function is supermodular if and only if a certain extension of it from the vertices of a box to the entire box is a concave function [Lov83, Proposition 4.1]; see [Iwa08, Theorem 4] for a simple proof. The projection of this extension generates the concave envelope of a supermodular function, and this envelope is described in the x-space by so-called polymatroid inequalities; see [AN08, Theorem 1], [Von10, Lemma 4], [TRX13, Theorem 3.3] for some proofs of this well-known fact. These polymatroid inequalities can be derived from polarity and the results on optimizing over the polymatroid polyhedron [Edm70]. There is one inequality for each n-permutation, allowing for repetitions, and this immediately relates to there being a single core facet of the envelope and all other facets being permutations of it. The convex envelope of a supermodular function is not known in general and is NP-hard to separate over. For the symmetric structure that is considered in this paper, this envelope was first described by [TRX13, Theorem 4.6] using their approach of strategically computing an exponential number of subdivisions of X and by using the extreme points of these subdivisions to calculate the facets. A similarity between this proof and ours is that they both rely on the Kuhn's triangulation [Kuh60] of a box. The convex hull of a monomial with $-\ell = u$, was first established by the authors in [AGX19, Theorem 4.1]. However, it was done so without using the symmetry in the monomial to recognize the core facets.

3

Since the original submission of this paper, convexification of permutation-invariant sets, such as the graph G, has been studied by [KTR19] who give a general framework for obtaining an extended formulation with $O(n^2)$ many variables and inequalities for the convex hulls of such sets. This framework is based on first convexifying a strategically-defined subset of the region of interest, and then obtaining the convex hull for all permutations of each point within the convexified set. In the realm of convex hulls for SMPs, [KTR19] and this paper both utilize information relative to a specific simplex that generates the Kuhn's triangulation of a box, and this similarity is not entirely surprising because of symmetry in the problem.

1.2 Organisation of the Paper

Section 2 presents an extended formulation for $\mathsf{conv}\, G$, gives a polar description of this convex hull, introduces the concept of core inequalities from which all valid linear inequalities can be derived upto permutations the coefficients, and answers the questions of optimization and separation. Section 3 provides various conditions for a core valid equality to be a core facet. Section 4 gives the RLT theory for general multilinear polynomials and derives consequences of it on nonnegativity of the polynomial over a box. This RLT machinery enables us to characterize the set of points in the graph at which a core facet is satisfied exactly. Section 5 considers the case of $m(x)$ being a supermodular function and Section 6 considers the case of $m(x)$ being a monomial with the variable domain being a box that allows reflections across the origin. Finally, Section 7 provides a summary of the paper and highlights some outstanding open questions for future research. The Appendix gives alternate arguments for deriving basic properties of $\mathsf{conv}\, G$ using the RLT, and also has illustrative examples for our main results.

2 Preliminaries

The convex hull of G is a full-dimensional set because G is the surface of a nonlinear function taken over an n-dimensional box, and this set is a polytope whose vertices are in bijection to the vertices of the box X since this is known for general multilinear functions [Rik97; She97]. Throughout this paper, we will study inequalities of the form

$$\beta_0 + \sum_{j=1}^n \beta_j x_j + \beta' y \geqslant 0. \tag{2}$$

Trivial facets of $\mathsf{conv}\, G$ (also referred to as vertical facets in the literature) are the facets generated by valid β with $\beta' = 0$. For $n \geqslant 3$, trivial facets are precisely the bounds on the variables, and for $n = 2$ there are no trivial facets [BST09, Theorem 2.4 and Remark 2.1]. Wlog and upto scaling, we can assume that every nontrivial valid inequality has $\beta' = \pm 1$. Since $\mathsf{conv}\, G$ is the intersection of the epigraph of the convex envelope of m and hypograph of the concave envelope of m, a facet-defining inequality (facet) with $\beta' = 1$ (resp. $\beta' = -1$) represents a nontrivial facet of the epigraph (resp. hypograph). Since $\mathsf{conv}\, G$ is full-dimensional, for every nontrivial facet of $\mathsf{conv}\, G$ there exists a unique nontrivial valid inequality. We will characterize nontrivial facets using the concept of core inequalities.

Certain notation will be useful throughout our study. Our polyhedral analysis will rely on the simplex

$$\mathcal{S} \equiv \{x \in \mathbb{R}^n : u \geqslant x_1 \geqslant \ldots \geqslant x_n \geqslant \ell\}, \tag{3}$$

whose $n + 1$ extreme points are

$$x^k \equiv \left(\underbrace{u, \ldots, u}_{k}, \underbrace{\ell, \ldots, \ell}_{n-k} \right), \quad k = 0, \ldots, n. \tag{4}$$

The simplex $\mathcal{S}$ is the one that has been used in combinatorial geometry to yield the Kuhn's triangulation of a box [Kuh60]. It is also known to be useful for convexifying general submodular/supermodular functions [Lov83; TRX13] and we will see this also in §5. The value of the multilinear function at each x^k can be computed by substituting it into equation (1a). In the special case of $\ell = 0$ and $u = 1$, this becomes the combinatorial formula $m(x^k) = c_k + \sum_{j=2}^{k-1} \binom{k}{j} c_j$. For notational convenience, we let $\mathcal{L}_k(\beta_0, \beta) = \beta_0 + \sum_{j=1}^{n} \beta_j x_j^k$ be that value obtained by inserting extreme point x^k of $\mathcal{S}$ into the expression $\beta_0 + \sum_{j=1}^{n} \beta_j x_j$ of (2), where x_j^k denotes entry j of x^k, so that

$$\mathcal{L}_k(\beta_0, \beta) = \beta_0 + u \left(\sum_{j=1}^{k} \beta_j \right) + \ell \left(\sum_{j=k+1}^{n} \beta_j \right), \qquad \forall \, k \in \{0, \ldots, n\}, \tag{5}$$

with $\sum_{j=1}^{0} \beta_j = 0$ and $\sum_{j=n+1}^{n} \beta_j = 0$ in $\mathcal{L}_0(\beta_0, \beta)$ and $\mathcal{L}_n(\beta_0, \beta)$, respectively.

We begin by noting two implicit descriptions of $\operatorname{conv} G$, one is an extended formulation that projects onto this convex hull and another is a polarity result that gives a characterization of all the facets. Then we introduce core inequalities as those inequalities having a nondecreasing order on the coefficients and permutations of which generate all the valid inequalities. Lastly, we give algorithms for optimizing and separating over $\operatorname{conv} G$.

2.1 Implicit Descriptions of the Convex Hull

A straightforward application of disjunctive programming along with using the well-known result that the envelopes of a multilinear function are generated by the extreme points of the box gives us a $O(n^2)$-sized extended formulation for $\operatorname{conv} G$.

Proposition 2.1. *The following polyhedron projects onto $\operatorname{conv} G$,*

$$\left\{ (x, \{w^k\}_{k=0}^n, y, v, \lambda) : \ x_j = \sum_{k=0}^{n} w_j^k, \ \forall j \in N, \ y = \sum_{k=0}^{n} v_k, \ \sum_{k=0}^{n} \lambda_k = 1, \right.$$

$$\sum_{j=1}^{n} w_j^k = (ku + (n-k)\ell)\lambda_k, \ v_k = m(x^k)\lambda_k, \ k = 0, \ldots, n$$

$$\left. \ell \lambda_k \leqslant w_j^k \leqslant u \lambda_k \ \forall j, k, \ v \in \mathbb{R}^{n+1}, \ \lambda \in \mathbb{R}_+^{n+1} \right\}$$

Proof. It is well-known [Rik97; She97] that envelopes of a multilinear function over a box have the vertex extendability property meaning that for any multilinear function $p(x)$ and box $B \subset \mathbb{R}^n$, the convex hull of $\{(x, y) \in B \times \mathbb{R} : y = p(x)\}$ is equal to the convex hull of $\{(x, y) \in \operatorname{vert} B \times \mathbb{R} : y = p(x)\}$. Therefore, we have $\operatorname{conv} G$ being equal to $\operatorname{conv} \{(x, y) \in \{\ell, u\}^n \times \mathbb{R} : y = m(x)\}$. The vertex set $\{\ell, u\}^n$ can be partitioned into $n + 1$ sets with each set for $k \in \{0, \ldots, n\}$ corresponding to points having k entries of u, thereby giving us

$$\operatorname{vert}(\operatorname{conv} G) = \bigcup_{k=0}^{n} \bigcup_{\sigma \in \mathfrak{S}_n} \left\{ (\sigma \cdot x^k, m(\sigma \cdot x^k)) \right\} = \bigcup_{k=0}^{n} \left\{ Q_k \times \{m(x^k)\} \right\}, \tag{6}$$

where $\mathfrak{S}_n$ is the symmetric group of n elements, $\sigma \cdot v = (v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)})$ for a vector $v \in \mathbb{R}^n$, $Q_k := \cup_{\sigma \in \mathfrak{S}_n} \{\sigma \cdot x^k\}$, and the second equality is due to symmetry of $m(x)$. Since $\operatorname{conv} G$ is the convex hull of its vertices, we get $\operatorname{conv} G = \operatorname{conv}\left(\bigcup_{k=0}^n \left\{\operatorname{conv} Q_k \times \{m(x^k)\}\right\}\right)$. Since Q_k is the set of all points that have k entries of u and $n-k$ entries of ℓ, we have $\operatorname{conv} Q_k = \{x \in [\ell, u]^n \colon \sum_{j=1}^n x_j = ku + (n-k)\ell\}$. Set $P_k := \operatorname{conv} Q_k \times \{m(x^k)\} = \{(x,y) \in [\ell, u]^n \times \mathbb{R} \colon \sum_{j=1}^n x_j = ku + (n-k)\ell, \ y = m(x^k)\}$, which is a polyhedron. Since $\operatorname{conv} G = \operatorname{conv}(\cup_{k=0}^n P_k)$, applying disjunctive programming [Bal79] yields the desired extended formulation. $\square$

Another way of implicitly describing the convex hull is to characterise all its facets using the extreme points of the polar of the convex hull. This is known from [She97, Theorem 2] and [BST09, Theorem 2.4] for general multilinear functions over arbitrary boxes, and hence we state the below result for an SMP without proof.

Proposition 2.2. *An inequality* (2) *with $\beta' = 1$ (resp. $\beta' = -1$) is a facet of $\operatorname{conv} G$ if and only if (β_0, β) is an extreme point of the polyhedron $E := \{(\beta_0, \beta) \colon \beta_0 + \beta^\top \bar{x} \geqslant -m(\bar{x}), \ \bar{x} \in \{\ell, u\}^n\}$ (resp. $H := \{(\beta_0, \beta) \colon \beta_0 + \beta^\top \bar{x} \geqslant m(\bar{x}), \ \bar{x} \in \{\ell, u\}^n\}$).*

The challenge with using this polarity result is that it requires enumeration of extreme points of the set E and H, which are exponentially many in general. The next section observes that it suffices to focus attention on only a subset of inequalities since all other inequalities are obtained via permutations.

2.2 Core Inequalities

We define *core inequalities* as being those inequalities (2) that have $\beta' = \pm 1$ and $\beta_1 \leqslant \ldots \leqslant \beta_n$. The restriction that $\beta' = \pm 1$ for nonzero β' is nonrestrictive, as it follows from scaling. The restriction that $\beta_1 \leqslant \ldots \leqslant \beta_n$ follows from the problem symmetry, as an inequality will be valid (a facet) for $\operatorname{conv} G$ if and only if every symmetric copy obtained by permuting the entries in $\beta^\top$ is also valid (a facet). A *valid core inequality* is a core inequality that is valid for $\operatorname{conv} G$. A valid core inequality that is also a facet for $\operatorname{conv} G$ is a *core facet*.

Symmetry implies that the validity of a core inequality (2) can be checked in terms of only the $(n+1)$ extreme points of the simplex $\mathcal{S}$.

Lemma 2.3. *A core inequality* (2) *is a valid core inequality if and only if*

$$\mathcal{L}_k(\beta_0, \beta) + \beta' m(x^k) \geqslant 0, \qquad \forall \, k \in \{0, \ldots, n\}. \tag{7}$$

Proof. The only if direction is trivial and so we consider the if direction. It is sufficient to show that the $(n+1)$ inequalities of (7) imply that (2) is nonnegative for all 2^n extreme points of $\operatorname{conv} G$. Toward this end, consider any $k \in \{0, \ldots, n\}$, and note that each of the $\binom{n}{k}$ extreme points (x,y) of $\operatorname{conv} G$ with x having k entries of value u and $(n-k)$ entries of value ℓ also has $y = m(x^k)$. In addition, each such extreme point yields a value for $\left(\beta_0 + \sum_{j=1}^n \beta_j x_j\right)$ that is at least as large as $\mathcal{L}_k(\beta_0, \beta)$, so that (2) is nonnegative for all $\binom{n}{k}$ such extreme points. As the result holds true for every such k, it holds true for all the extreme points of $\operatorname{conv} G$. This completes the proof. $\square$

6

2.3 Optimization and Separation

The question of optimizing a linear function over $\operatorname{conv} G$ can be solved via a sorting algorithm without using the quadratic-sized extended formulation of Proposition 2.1.

Proposition 2.4. *For any* $(\alpha, \alpha') \in \mathbb{R}^n \times \mathbb{R}$,

$$
\max \left\{ \alpha^\top x + \alpha' y : (x, y) \in \operatorname{conv} G \right\} = \max_{k=0,\dots,n} \left\{ u \sum_{j=1}^{k} \alpha_{\sigma(j)} + \ell \sum_{j=k+1}^{n} \alpha_{\sigma(j)} + \alpha' m(x^k) \right\},
$$

where the permutation σ *is such that* $\alpha_{\sigma(1)} \geqslant \alpha_{\sigma(2)} \geqslant \cdots \geqslant \alpha_{\sigma(n)}$.

Proof. The linear function has an optimum at a vertex of $\operatorname{conv} G$, and so the left-hand side is equivalent to maximising $\alpha^\top x + \alpha' y$ over $\operatorname{vert}(\operatorname{conv} G)$. Our claim follows after using $\operatorname{vert}(\operatorname{conv} G) = \cup_{k=0}^{n} \{Q_k \times \{m(x^k)\}\}$ from (6) and observing that $\max\{\alpha^\top x : x \in Q_k\} = u \sum_{j=1}^{k} \alpha_{\sigma(j)} + \ell \sum_{j=k+1}^{n} \alpha_{\sigma(j)}$. $\square$

Hence, assuming computation of each $m(x^k)$ takes $O(1)$ time, linear optimization over $\operatorname{conv} G$ can be solved in $O(n \log n)$ time by sorting the vector α.

Due to the well-known equivalence of complexity of optimization and separation over a polyhedron, it follows that a given point can be separated from the convex hull of G in polynomial time. However, this connection invokes the ellipsoid algorithm whose complexity is a high degree polynomial in the input encoding. The question of separation can be answered directly if there are polynomially many core facets. The approach is similar to that used for Proposition 2.4 but here the point $(\bar{x}, \bar{y})$ to be separated has $\bar{x}$ sorted in nonincreasing order.

Proposition 2.5. *If* $\operatorname{conv} G$ *has* t *core facets, then a point can be separated from* $\operatorname{conv} G$ *in* $O(nt + n \log n)$ *time.*

Proof. Let $(\bar{x}, \bar{y}) \in X \times \mathbb{R}$ be a given point. Take $\sigma \in \mathfrak{S}_n$ such that $\bar{x}_{\sigma(1)} \leqslant \bar{x}_{\sigma(2)} \leqslant \cdots \leqslant \bar{x}_{\sigma(n)}$. Choose any core facet (2). We have $\beta_1 \leqslant \cdots \leqslant \beta_n$ and need to check if this facet or any permuted facet is violated, i.e., whether there exists a permutation $\sigma \in \mathfrak{S}_n$ such that $\beta_0 + \sum_{j=1}^{n} \beta_{\sigma(j)} \bar{x}_j + \beta' \bar{y} < 0$. This is equivalent to checking whether $-\beta_0 - \beta' \bar{y} > \min\{\bar{x}^\top v : v \in P_n(\beta)\}$, where $P_n(\beta) = \operatorname{conv}\{(\beta_{\sigma(1)}, \beta_{\sigma(2)}, \dots, \beta_{\sigma(n)}) : \sigma \in \mathfrak{S}_n\}$ is the generalized permutahedron [Bow72] with respect to β. Linear optimization over the permutahedron can be done via the sorting algorithm since the permutahedron is the base polymatroid polyhedron corresponding to a certain submodular function [Rad52] and Edmonds [Edm70] established that the sorting algorithm works for optimizing over any base polymatroid. In particular, we have that our minimum is attained at $v^* = \sigma^{-1} \cdot \hat{\beta}$, where $\hat{\beta} := (\beta_n, \beta_{n-1}, \dots, \beta_1)$ is the reverse sorting of β, so that $v_j^* = \beta_{n+1-\sigma^{-1}(j)}$ for all $j \in N$. Thus, checking separation of a single core facet and its permuted copies is equivalent to checking whether $-\beta_0 - \beta' \bar{y} > \sum_{j=1}^{n} \beta_{n+1-\sigma^{-1}(j)} \bar{x}_j$, which runs in $O(n)$ time. The overall complexity then becomes $O(nt + n \log n)$ due to the sorting of $\bar{x}$ initially and enumeration over t core facets. $\square$

As seen in the above proof, the overall complexity is composed of nt additions and multiplications and a sorting step which requires $O(n \log n)$ comparisons.

3   Conditions for Core Facets

A challenge in determining whether a valid core inequality is a core facet is the identification of a maximum number of affinely independent points within $\operatorname{conv} G$ that satisfies the inequality exactly. Of course, we can restrict attention to only those points that are extreme to $\operatorname{conv} G$. The following proposition shows that the set of extreme points of $\operatorname{conv} G$ which satisfy a core inequality exactly can be completely characterized in terms of the extreme points of the simplex $\mathcal{S}$.

Lemma 3.1. *Given an extreme point $(\tilde{x}, m(\tilde{x}))$ of $\operatorname{conv} G$ with $\tilde{x}$ not an extreme point of $\mathcal{S}$, let r be the smallest index j such that $x_j = \ell$, and s be the largest index j such that $x_j = u$. Further let p be the number of entries of $\tilde{x}$ having value u (so that $r \leqslant p \leqslant s - 1$). Then $(x, y) = (\tilde{x}, m(\tilde{x}))$ satisfies a valid core inequality (2) exactly if and only if $(x, y) = (x^p, m(x^p))$ satisfies the inequality exactly, and $\beta_r = \ldots = \beta_s$.*

Proof. We have $\beta_0 + \sum_{j=1}^{n} \beta_j \tilde{x}_j + \beta' m(\tilde{x}) \geqslant \mathcal{L}_p(\beta_0, \beta) + \beta' m(x^p) \geqslant 0$, where the first inequality is due to the nondecreasing values of β_j and the symmetry of $m(x)$, and the second inequality is due to the validity of (2). The first inequality is satisfied exactly if and only if $\beta_r = \ldots = \beta_s$, while the second inequality is satisfied exactly if and only if (2) is satisfied exactly at $(x^p, m(x^p))$. $\square$

A consequence is that given a valid core inequality, we can identify the subset of 2^n extreme points of $\operatorname{conv} G$ that satisfy the inequality exactly by considering only the $(n+1)$ extreme points of $\mathcal{S}$. The next few results build upon this consequence to provide characteristics of core facets in terms of the extreme points of $\mathcal{S}$.

Proposition 3.2. *A valid core inequality (2) is a core facet only if it is satisfied exactly at $(x^k, m(x^k))$ for at least two extreme points x^k of the simplex $\mathcal{S}$, with at least one extreme point not being x^0 or x^n.*

Proof. Suppose that a valid core inequality (2) is satisfied exactly at fewer than two such points $(x^k, m(x^k))$, or at only the points $(x^0, m(x^0))$ and $(x^n, m(x^n))$. Since the dimension of $\operatorname{conv} G$ is $n + 1$, it is sufficient to show that the inequality is satisfied exactly at no more than n affinely independent extreme points of $\operatorname{conv} G$. Three cases arise. First, if the inequality is satisfied exactly at no such $(x^k, m(x^k))$, then (2) is not satisfied exactly at any of the 2^n extreme points of $\operatorname{conv} G$ by Lemma 3.1. Second, if the inequality is satisfied exactly at precisely one such $(x^{\tilde{k}}, m(x^{\tilde{k}}))$ then, by Lemma 3.1, the only extreme points $(x, m(x))$ of $\operatorname{conv} G$ that can possibly satisfy (2) exactly are the $\binom{n}{k}$ points with x having $\tilde{k}$ entries of value u and $(n - \tilde{k})$ entries of value ℓ. However, as the two linearly independent hyperplanes $\sum_{j=1}^{n} x_j = \tilde{k} u + (n - \tilde{k})\ell$ and $y = m(x^{\tilde{k}})$ both pass through these $\binom{n}{k}$ points, there exist at most n such affinely independent points. Finally, if the inequality is satisfied exactly at only $(x^0, m(x^0))$ and $(x^n, m(x^n))$, then Lemma 3.1 gives us that these are the only extreme points of $\operatorname{conv} G$ that satisfy (2) exactly. $\square$

The necessary condition in Proposition 3.2 to be a core facet is also sufficient when all the coefficients are equal.

Proposition 3.3. *A valid core inequality (2) with $\beta_1 = \ldots = \beta_n$ is a core facet if and only if it is satisfied exactly at $(x^k, m(x^k))$ for at least two extreme points x^k of the simplex $\mathcal{S}$, with at least one extreme point not being x^0 or x^n.*

Proof. The only if direction follows directly from Proposition 3.2, and so we consider the if direction. Suppose the inequality is satisfied exactly at two such points $(x^p, m(x^p))$ and $(x^q, m(x^q))$, with x^p not being x^0 or x^n. Then Lemma 3.1 gives us that the $\binom{n}{p}$ extreme points $(x, m(x))$ of $\mathsf{conv}\, G$ with x having p entries of value u and $(n-p)$ entries of value ℓ satisfy the inequality exactly. The proof is to show that there exist n affinely independent points from amongst this set of $\binom{n}{p}$ points. Then these n points, together with $(x^q, m(x^q))$, will form an affinely independent set of $(n+1)$ points because each of the first n points satisfies the equation $\sum_{j=1}^{n} x_j = pu + (n-p)\ell$, but $(x^q, m(x^q))$ does not. In fact, it is sufficient to show that the extreme points of X associated with these $\binom{n}{p}$ points are affinely independent.

Since the affine independence of a collection of points is unaffected when the same value is subtracted from every entry of each point, and when each point is multiplied by a nonzero scalar, the affine independence of the associated $\binom{n}{p}$ extreme points of X remains unchanged when every u is replaced with 1 and every ℓ is replaced with 0. Consider the $n \times \binom{n}{p}$ matrix A defined so that each column corresponds to one such point, upon application of these operations. If $p = 1$, then A is a permutation matrix, and the $\binom{n}{p}$ points are affinely independent. Otherwise, A has the two properties that: each row contains $\kappa_1 \equiv \binom{n-1}{p-1}$ entries of value 1, and every pair of two distinct rows contains $\kappa_2 \equiv \binom{n-2}{p-2}$ common entries of value 1. Hence, AA^T is the $n \times n$ matrix having κ_1 along the main diagonal and κ_2 elsewhere. Since $\mathrm{rank}(AA^T) = n$, because the common row sum allows us to subtract κ_2 from every entry of AA^T to obtain a lower-triangular matrix with $(\kappa_1 - \kappa_2) \neq 0$ along the diagonal, then $\mathrm{rank}(A) = n$, and the proof is complete. $\square$

The below proposition gives further conditions, in terms of the extreme points x^k of $\mathcal{S}$, for a valid core inequality to be a core facet.

Proposition 3.4. *Given any two valid core inequalities of the form*

$$\bar{\beta}_0 + \sum_{j=1}^{n} \bar{\beta}_j x_j + \beta' y \geqslant 0 \ \text{and} \ \hat{\beta}_0 + \sum_{j=1}^{n} \hat{\beta}_j x_j + \beta' y \geqslant 0, \tag{8}$$

if

$$\mathcal{L}_k(\bar{\beta}_0, \bar{\beta}) \leqslant \mathcal{L}_k(\hat{\beta}_0, \hat{\beta}) \ \forall \ k \in \{0, \dots, n\}, \tag{9}$$

with strict inequality holding for at least one k in (9), then the right inequality of (8) is not a facet.

Proof. It is sufficient to show that every extreme point $(x, m(x))$ of $\mathsf{conv}\, G$ that satisfies the right inequality of (8) exactly also satisfies the left inequality of (8) exactly. This statement holds for the $(n + 1)$ extreme points $(x^k, m(x^k)), k \in \{0, \dots, n\}$, by (9), and so we arbitrarily select any one of the remaining $2^n - (n + 1)$ extreme points, say $(\tilde{x}, m(\tilde{x}))$, and suppose that the right inequality holds exactly at this point. Some $p \in \{1, \dots, n - 1\}$ entries of $\tilde{x}$ have value u and $(n - p)$ entries have value ℓ, with the first p entries not all equal to u. The proof reduces to showing that

$$0 = \mathcal{L}_p(\hat{\beta}_0, \hat{\beta}) + \beta' m(x^p) = \mathcal{L}_p(\bar{\beta}_0, \bar{\beta}) + \beta' m(x^p) = \bar{\beta}_0 + \sum_{j=1}^{n} \bar{\beta}_j \tilde{x}_j + \beta' m(\tilde{x}). \tag{10}$$

Each equality of (10) is considered separately.

- Since, by assumption, $(\tilde{x}, m(\tilde{x}))$ satisfies the right inequality of (8) exactly, Lemma 3.1 gives us that the first equality of (10) holds true, and also that $\hat{\beta}^* = \hat{\beta}_r = \ldots = \hat{\beta}_s$ for some scalar $\hat{\beta}^*$, where r and s are, respectively, the indices of the first and last entries of $\tilde{x}$ which differ from x^p.

- The second equality of (10) holds true by (9) with $k = p$, the first equality of (10), and the validity of the left inequality of (8) at $(x^p, m(x^p))$.

- Since as noted above, $\hat{\beta}^* = \hat{\beta}_r = \ldots = \hat{\beta}_s$ for some scalar $\hat{\beta}^*$, we have that

$$\mathcal{L}_p(\hat{\beta}_0, \hat{\beta}) + \hat{\beta}^*(u - \ell)(s - p) = \mathcal{L}_s(\hat{\beta}_0, \hat{\beta}) \geqslant \mathcal{L}_s(\bar{\beta}_0, \bar{\beta}) = \mathcal{L}_p(\bar{\beta}_0, \bar{\beta}) + (u - \ell)(\sum_{j=p+1}^{s} \bar{\beta}_j) \quad (11)$$

and

$$\mathcal{L}_p(\hat{\beta}_0, \hat{\beta}) - \hat{\beta}^*(u - \ell)(p - r + 1) = \mathcal{L}_{r-1}(\hat{\beta}_0, \hat{\beta}) \geqslant \mathcal{L}_{r-1}(\bar{\beta}_0, \bar{\beta}) = \mathcal{L}_p(\bar{\beta}_0, \bar{\beta}) - (u - \ell)(\sum_{j=r}^{p} \bar{\beta}_j),$$
$$(12)$$

where, within (11) and (12), the equalities follow from (5), and the inequalities are due to (9). Combine these expressions and invoke $\mathcal{L}_p(\hat{\beta}_0, \hat{\beta}) = \mathcal{L}_p(\bar{\beta}_0, \bar{\beta})$ from the second equality of (10) to obtain

$$\hat{\beta}^* \leqslant \left(\frac{1}{p - r + 1} \right) \left(\sum_{j=r}^{p} \bar{\beta}_j \right) \leqslant \left(\frac{1}{s - p} \right) \left(\sum_{j=p+1}^{s} \bar{\beta}_j \right) \leqslant \hat{\beta}^*,$$

where the three inequalities follow from (12), the nondecreasing property of $\bar{\beta}$, and (11), respectively. Again by the nondecreasing property of $\bar{\beta}$, we have that $\hat{\beta}^* = \bar{\beta}_r = \ldots = \bar{\beta}_s$, giving the third equality of (10) by Lemma 3.1. $\square$

This leads us to the following necessary and sufficient condition for a valid inequality with distinct coefficients to be a core facet.

Corollary 3.5. *A valid inequality $\bar{\beta}_0 + \sum_{j=1}^{n} \bar{\beta}_j x_j + \beta' y \geqslant 0$ with $\bar{\beta}_1 < \ldots < \bar{\beta}_n$ is a core facet if and only if it is satisfied exactly at $(x^k, m(x^k))$ for $k \in \{0, \ldots, n\}$. In this case, no other core facet can exist with the given β'.*

Proof. The if direction follows from the $n + 1$ points $(x^k, m(x^k))$ for $k \in \{0, \ldots, n\}$ being affinely independent, and so we consider the only if direction. Lemma 3.1 gives us that the only extreme points to $\mathbf{conv}\, G$ that can possibly satisfy the given inequality exactly are $(x^k, m(x^k))$ for $k \in \{0, \ldots, n\}$. Since the inequality is a facet, these $(n + 1)$ affinely independent points must satisfy the inequality exactly, giving $\mathcal{L}_k(\bar{\beta}_0, \bar{\beta}) + \beta' m(x^k) = 0$ for $k \in \{0, \ldots, n\}$. No other valid inequality $\hat{\beta}_0 + \sum_{j=1}^{n} \hat{\beta}_j x_j + \beta' y \geqslant 0$ for $\mathbf{conv}\, G$ with the given β' can then be a facet by Proposition 3.4. $\square$

Another set of necessary conditions for being a core facet is obtained below.

Proposition 3.6. *Given any two valid core inequalities of the form (8), suppose that the following conditions hold.*

1. *$\bar{\beta}_u = \bar{\beta}_v$ for each $(u, v), u < v$, having $\hat{\beta}_u = \hat{\beta}_v$.*
2. *$\{k : \mathcal{L}_k(\hat{\beta}_0, \hat{\beta}) + \beta' m(x^k) = 0\} \subset \{k : \mathcal{L}_k(\bar{\beta}_0, \bar{\beta}) + \beta' m(x^k) = 0\}$.*

10

Then the right inequality of (8) *is not a facet.*

Proof. As with the proof of Proposition 3.4, it is sufficient to show that every extreme point of $\mathsf{conv}\,G$ that satisfies the right inequality of (8) exactly also satisfies the left inequality of (8) exactly. This statement holds for the $(n+1)$ extreme points $(x^k, m(x^k)), k \in \{0, \ldots, n\}$, by Condition 2 above, and so we arbitrarily select any one of the remaining $2^n - (n+1)$ extreme points, say $(\tilde{x}, m(\tilde{x}))$, and suppose that the right inequality holds exactly at this point. Some $p \in \{1, \ldots, n-1\}$ entries of $\tilde{x}$ have value u and $(n-p)$ entries have value ℓ, with the first p entries not all equal to u. By Lemma 3.1, we have that $\hat{\beta}_r = \ldots = \hat{\beta}_s$ where r and s are, respectively, the indices of the first and last entries of $\tilde{x}$ which differ from x^p. The first condition above then gives $\bar{\beta}_r = \ldots = \bar{\beta}_s$, so that by again invoking Lemma 3.1, we have the left inequality of (8) is satisfied exactly at $(\tilde{x}, m(\tilde{x}))$. $\qquad\square$

The below result identifies instances in which select coefficients of a core facet must have equal values. We do not need to consider $r = n$ or $s = 0$ because such cases are not possible by Proposition 3.2.

Proposition 3.7. *Given a core facet* (2)*, let*

$$r = \min\left\{k \in \{0, \ldots, n-1\} \colon \mathcal{L}_k(\beta_0, \beta) + \beta' m(x^k) = 0\right\},$$
$$s = \max\left\{k \in \{1, \ldots, n\} \colon \mathcal{L}_k(\beta_0, \beta) + \beta' m(x^k) = 0\right\}.$$

Then $\beta_1 = \ldots = \beta_{r+1}$ and $\beta_s = \ldots = \beta_n$.

Proof. We consider separately the conclusions that $\beta_1 = \ldots = \beta_{r+1}$ and $\beta_s = \ldots = \beta_n$.

- If $r = 0$, the first conclusion follows trivially. Otherwise, $r \geqslant 1$ and, by contradiction, define $p \leqslant r$ so that $\beta_p < \beta_{p+1} = \beta_{r+1}$. It is sufficient to show that there exists an $\varepsilon > 0$ so that the inequality

$$(\beta_0 - pu\varepsilon) + \sum_{j=1}^{p}(\beta_j + \varepsilon)x_j + \sum_{j=p+1}^{n}\beta_j x_j + \beta' y \geqslant 0, \tag{13}$$

with $(\beta_p + \varepsilon) \leqslant \beta_{p+1}$, is valid for $\mathsf{conv}\,G$. Then every extreme point $(x, m(x))$ to $\mathsf{conv}\,G$ with $x_j = u$ for $j \in \{1, \ldots, p\}$ will have the left side of (13) equal to the left side of (2), while every extreme point $(x, m(x))$ to $\mathsf{conv}\,G$ with $x_j = \ell$ for at least one $j \in \{1, \ldots, p\}$ will have the left side of (13) strictly less than the left side of (2). Relative to the extreme points of the simplex $\mathcal{S}$, the left sides of (13) and (2) will both take value $\mathcal{L}_k(\beta_0, \beta) + \beta' m(x^k)$ at $(x^k, m(x^k))$ for each $k \in \{p, \ldots, n\}$, but the left side of (13) will be $\mathcal{L}_k(\beta_0, \beta) + \beta' m(x^k) - (p-k)(u-\ell)\varepsilon$ at $(x^k, m(x^k))$ for each $x^k, k \in \{0, \ldots, p-1\}$, which is $(p-k)(u-\ell)\varepsilon > 0$ less than the left side of (2). Define $\varepsilon = \min\{\varepsilon', (\beta_{p+1} - \beta_p)\}$, where $\varepsilon' = \min_{k \in \{0, \ldots, p-1\}}\left\{\frac{\mathcal{L}_k(\beta_0, \beta) + \beta' m(x^k)}{(p-k)(u-\ell)}\right\}$. Then $\varepsilon > 0$ and inequality (13) is valid for $\mathsf{conv}\,G$ by Lemma 2.3, as it is valid at all $(x^k, m(x^k)), k \in \{0, \ldots, n\}$.

- If $s = n$, the second conclusion follows trivially. Otherwise $s \leqslant (n-1)$ and, by contradiction, define $p \geqslant s$ so that $\beta_s = \beta_p < \beta_{p+1}$. It is sufficient to show that there exists an $\varepsilon > 0$ so that the inequality

$$\beta_0 + (n-p)\ell\varepsilon + \sum_{j=1}^{p}\beta_j x_j + \sum_{j=p+1}^{n}(\beta_j - \varepsilon)x_j + \beta' y \geqslant 0, \tag{14}$$

11

with $\beta_p \leqslant (\beta_{p+1} - \varepsilon)$, is valid for $\text{conv}\,G$. Then every extreme point $(x, m(x))$ to $\text{conv}\,G$ with $x_j = \ell$ for $j \in \{p+1, \ldots, n\}$ will have the left side of (14) equal to the left side of (2), while every extreme point $(x, m(x))$ to $\text{conv}\,G$ with $x_j = u$ for at least one $j \in \{p+1, \ldots, n\}$ will have the left side of (14) strictly less than the left side of (2). Relative to the extreme points of the simplex $\mathcal{S}$, the left sides of (14) and (2) will both take value $\mathcal{L}_k(\beta_0, \beta) + \beta' m(x^k)$ at $(x^k, m(x^k))$ for each $k \in \{1, \ldots, p\}$, but the left side of (14) will be $\mathcal{L}_k(\beta_0, \beta) + \beta' m(x^k) - (k-p)(u-\ell)\varepsilon$ at $(x^k, m(x^k))$ for each $x^k, k \in \{p+1, \ldots, n\}$, which is $(k-p)(u-\ell)\varepsilon > 0$ less than the left side of (2). Define $\varepsilon = \min\{\varepsilon', (\beta_{p+1} - \beta_p)\}$, where $\varepsilon' = \min_{k \in \{p+1, \ldots, n\}} \left\{ \frac{\mathcal{L}_k(\beta_0, \beta) + \beta' m(x^k)}{(k-p)(u-\ell)} \right\}$. Then $\varepsilon > 0$ and inequality (14) is valid for $\text{conv}\,G$ by Lemma 2.3, as it is valid at all $(x^k, m(x^k)), k \in \{0, \ldots, n\}$. $\qquad\square$

4 RLT for Multilinear Functions

In this section, we present the Reformulation-Linearization Technique (RLT) as it pertains to the set G. We provide a brief description of select aspects of the general RLT process that are relevant to this study, emphasizing some key properties. Then we review the mathematical details of the RLT in terms of Kronecker products of matrices. This RLT machinery enables us to characterize exactness of the core facets for $\text{conv}\,G$.

4.1 Main Ideas

The RLT is a general methodology for reformulating mixed-integer linear and polynomial programs for the purpose of obtaining tight linear programming relaxations. While there is a rich body of literature on the topic [SA90; SA94; SA99] we focus attention here on a box-constrained region of n variables x_j, where each x_j is restricted to lie between variable bounds L_j and U_j. The RLT gives a hierarchy of successively tighter polyhedral relaxations, but we consider only the highest level n which affords the convex hull representations.

Specifically, consider the set

$$X' \equiv \left\{ x \in \mathbb{R}^n : L_j \leqslant x_j \leqslant U_j \ \forall\, j \in N \right\} \tag{15}$$

having each $L_j < U_j$, which is a generalization of the set X of (1c) that reduces to (1c) when $L_j = \ell$ and $U_j = u$ for all $j \in N$. The RLT process that we apply to (15) consists of the two steps of *reformulation* and *linearization*. The *reformulation* step computes products of the expressions $(x_j - L_j)$ and $(U_j - x_j)$, taken n at a time, such that one such expression is chosen for each j. In this manner, 2^n multilinear polynomial functions of degree n emerge. To elaborate, define the 2^n functions $F_K(x)$ so that

$$F_K(x) = \prod_{j \in K} (x_j - L_j) \prod_{j \notin K} (U_j - x_j), \quad K \subseteq N. \tag{16}$$

Then we have 2^n multilinear polynomial functions of the form

$$F_K(x), \quad K \subseteq N. \tag{17}$$

Each of these functions is nonnegative for all $x \in X'$, and the RLT enforces this nonnegativity to obtain the 2^n multilinear polynomial inequalities

$$F_K(x) \geqslant 0, \quad K \subseteq N, \tag{18}$$

12

that are satisfied for all $x \in X'$.

The *linearization* step then substitutes a continuous variable w_J for each of the $2^n - (n+1)$ distinct product terms $\prod_{j \in J} x_j$ with $J \subseteq N$ and $|J| \geqslant 2$ that are found in (18). Denote the linearized form of each function $F_K(x)$ that is obtained via such substitutions as $F_K(x)$. The RLT gives the following *polyhedral* set:

$$P \equiv \left\{ (x, w) \in \mathbb{R}^n \times \mathbb{R}^{2^n - (n+1)} : F_K(x) \geqslant 0, \ K \subseteq N \right\}. \tag{19}$$

We adopt the notation that $\{\bullet\}_L$ is the linearized form of the vector $\bullet$ that is obtained by substituting $w_J = \prod_{j \in J} x_j$ for all $J \subseteq N$ with $|J| \geqslant 2$ throughout $\bullet$. In this manner, $F_K(x) = \{F_K(x)\}_L$ for all $K \subseteq N$. Two properties of the set P are as follows.

Proposition 4.1. *The set P is a polytope having exactly 2^n extreme points and at each such extreme point (x, w) we have*

1. $x_j \in \{L_j, U_j\}$ *for all $j \in N$,*
2. $w_J = \prod_{j \in J} x_j$ *for all $J \subseteq N$ with $|J| \geqslant 2$.*

Consequently, we have $P = \mathsf{conv}\, T$ where

$$T \equiv \left\{ (x, w) \in \prod_{j \in N} [L_j, U_j] \times \mathbb{R}^{2^n - (n+1)} : w_J = \prod_{j \in J} x_j, \ J \subseteq N \text{ with } |J| \geqslant 2 \right\}. \tag{20}$$

These properties of P were originally established by Sherali and Adams [SA90] when P has $L_j = 0$ and $U_j = 1$ for all $j \in N$; see also [FKL15, Theorem 1]. They were more generally proven in [AS05; Hen11] using Kronecker products of matrices for richer families of polytopes that subsume P. These same type Kronecker products are also found in the generalization of the RLT constructs to handle general integer variables [AS05; Ada11]. For completeness, we present a tailored version of the arguments in the next subsection. In addition to establishing the two stated properties of P, this Kronecker product representation of (17) through (19) motivates two consequences that will be used later to characterize exactness of facet-defining inequalities. These apply to any multilinear polynomial and therefore encompass the functions $m(x)$.

Corollary 4.2. *For every $p(x) = \sum_{J \subseteq N} \alpha_j \prod_{j \in J} x_j$, there exists a unique $\pi \in \mathbb{R}^{2^n}$ so that $p(x) = \sum_{K \subseteq N} \pi_K F_K(x)$. In particular, $\pi_K = \frac{p(\hat{x})}{F_K(\hat{x})}$, where $\hat{x}_j = U_j$ for all $j \in K$ and $\hat{x}_j = L_j$ for all $j \notin K$.*

Corollary 4.3. *For every $p(x) = \sum_{J \subseteq N} \alpha_j \prod_{j \in J} x_j$, we have $p(x) \geqslant 0$ for all $x \in X'$ if and only if $p(x) \geqslant 0$ for every extreme point $x \in X'$.*

Proposition 4.1 can be extended to include the variable y and restriction $y = m(x)$ of (1b), so that the projection operation from the resulting higher-variable space onto the original (x, y) space can, in theory, be used to compute $\mathsf{conv}\, G$. These properties were already noted by Sherali [She97] using a different approach, we give our arguments in Appendix A for sake of completeness.

4.2 Use of Kronecker Products

Proof of Proposition 4.1. We prove Proposition 4.1 by expressing (17) and (18) as Kronecker products of matrices, and then using a known result of Kronecker products. The consequence

$P = \mathbf{conv}\, T$ follows since every extreme point of the polytope P is in T, and since $T \subseteq P$ by construction.

Recall that the Kronecker product of an $m_1 \times n_1$ matrix A with an $m_2 \times n_2$ matrix B, denoted by $A \otimes B$, is that $m_1 m_2 \times n_1 n_2$ matrix $A \otimes B = \begin{bmatrix} a_{11} B & \dots & a_{1n_1} B \\ \vdots & \ddots & \vdots \\ a_{m_1 1} B & \dots & a_{m_1 n_1} B \end{bmatrix}$, where a_{ij} is the $(i,j)^{th}$ entry of A. A result of Kronecker products is that, given any collection of r pairs of matrices A_i and B_i for $i = 1, ..., r$ such that the products $A_i B_i$ are defined, we have that

$$A_1 B_1 \otimes A_2 B_2 \otimes \dots \otimes A_r B_r = (A_1 \otimes A_2 \otimes \dots \otimes A_r)(B_1 \otimes B_2 \otimes \dots \otimes B_r). \tag{21}$$

As a consequence, if each of the matrices A_i is invertible with inverse A_i^{-1}, then by setting $B_i = A_i^{-1}$ for all i within (21), we obtain that

$$(A_1 \otimes A_2 \otimes \dots \otimes A_r)^{-1} = A_1^{-1} \otimes A_2^{-1} \otimes \dots A_r^{-1}, \tag{22}$$

since the left side of (21) becomes the suitably-sized identity matrix.

Relative to the functions $F_K(x)$ of (17) and (18), and the functions $F_K(x)$ and set P of (19), define for each variable x_j, the two matrices $\begin{bmatrix} U_j & -1 \\ -L_j & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ x_j \end{bmatrix}$ so that the functions $(U_j - x_j)$ and $(x_j - L_j)$ can be represented in matrix form as the two entries of the vector $\begin{bmatrix} U_j & -1 \\ -L_j & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_j \end{bmatrix}$. Then the 2^n functions of (17) can be collectively expressed using Kronecker products as

$$\left(\begin{bmatrix} U_1 & -1 \\ -L_1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \end{bmatrix} \right) \otimes \dots \otimes \left(\begin{bmatrix} U_n & -1 \\ -L_n & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_n \end{bmatrix} \right),$$

and the 2^n inequalities of (18) can be similarly expressed as

$$\left(\begin{bmatrix} U_1 & -1 \\ -L_1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \end{bmatrix} \right) \otimes \dots \otimes \left(\begin{bmatrix} U_n & -1 \\ -L_n & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_n \end{bmatrix} \right) \geqslant \mathbf{0}, \tag{23}$$

where $\mathbf{0}$ is the 2^n-dimensional column vector of all zeroes. Identity (21) allows us to rewrite (23) as

$$\left(\begin{bmatrix} U_1 & -1 \\ -L_1 & 1 \end{bmatrix} \otimes \dots \otimes \begin{bmatrix} U_n & -1 \\ -L_n & 1 \end{bmatrix} \right) \left(\begin{bmatrix} 1 \\ x_1 \end{bmatrix} \otimes \dots \otimes \begin{bmatrix} 1 \\ x_n \end{bmatrix} \right) \geqslant \mathbf{0}. \tag{24}$$

The set P of (19) can then be expressed in terms of (24) as

$$P = \left\{ (x, w) \in \mathbb{R}^n \times \mathbb{R}^{2^n - (n+1)} : \right.$$
$$\left. \left(\begin{bmatrix} U_1 & -1 \\ -L_1 & 1 \end{bmatrix} \otimes \dots \otimes \begin{bmatrix} U_n & -1 \\ -L_n & 1 \end{bmatrix} \right) \left\{ \begin{bmatrix} 1 \\ x_1 \end{bmatrix} \otimes \dots \otimes \begin{bmatrix} 1 \\ x_n \end{bmatrix} \right\}_L \geqslant \mathbf{0} \right\}, \tag{25}$$

which is equivalent to

$$P = \left\{ (x, w) \in \mathbb{R}^n \times \mathbb{R}^{2^n - (n+1)} : \right.$$
$$\left(\begin{bmatrix} U_1 & -1 \\ -L_1 & 1 \end{bmatrix} \otimes \dots \otimes \begin{bmatrix} U_n & -1 \\ -L_n & 1 \end{bmatrix} \right) \left\{ \begin{bmatrix} 1 \\ x_1 \end{bmatrix} \otimes \dots \otimes \begin{bmatrix} 1 \\ x_n \end{bmatrix} \right\}_L = \lambda$$
$$\left. \text{for some nonnegative } \lambda \in \mathbb{R}^{2^n} \right\}. \tag{26}$$

For each j, the inverse of the matrix $\begin{bmatrix} U_j & -1 \\ -L_j & 1 \end{bmatrix}$ is $\frac{1}{d_j}\begin{bmatrix} 1 & 1 \\ L_j & U_j \end{bmatrix}$, where $d_j = (U_j - L_j)$ denotes the difference between U_j and L_j. Then we apply (22) to (26) to rewrite the set P as

$$
P = \left\{ (x, w) \in \mathbb{R}^n \times \mathbb{R}^{2^n - (n+1)} : \right.
$$
$$
\left\{ \begin{bmatrix} 1 \\ x_1 \end{bmatrix} \otimes \ldots \otimes \begin{bmatrix} 1 \\ x_n \end{bmatrix} \right\}_L = \left(\frac{1}{d_1}\begin{bmatrix} 1 & 1 \\ L_1 & U_1 \end{bmatrix} \otimes \ldots \otimes \frac{1}{d_n}\begin{bmatrix} 1 & 1 \\ L_n & U_n \end{bmatrix} \right) \lambda
$$
$$
\left. \text{for some nonnegative } \lambda \in \mathbb{R}^{2^n} \right\}.
\tag{27}
$$

The two claimed properties of P become apparent in light of (27). The first equation of (27) is equivalent via scaling to $D_n = \sum_{j=1}^{2^n} \lambda_j$, with $D_n \equiv \prod_{j=1}^n d_j$, thereby establishing P as a polytope having exactly 2^n extreme points. These points can be numbered so that extreme point j has λ given by $\lambda_j = D_n$ and $\lambda_i = 0$ for $i \neq j$, and (x, w) given by column j of the matrix $\begin{bmatrix} 1 & 1 \\ L_1 & U_1 \end{bmatrix} \otimes \ldots \otimes \begin{bmatrix} 1 & 1 \\ L_n & U_n \end{bmatrix}$, less the first row. Each of the 2^n columns of this last matrix consists of a distinct realization of the vector $\begin{bmatrix} 1 \\ x_1 \end{bmatrix} \otimes \ldots \otimes \begin{bmatrix} 1 \\ x_n \end{bmatrix}$, less the first entry, evaluated at some x having each x_j fixed at either its lower bound L_j or its upper bound U_j. $\qquad\square$

Proof of Corollary 4.2. Invertibility of the matrix $\left(\begin{bmatrix} U_1 & -1 \\ -L_1 & 1 \end{bmatrix} \otimes \ldots \otimes \begin{bmatrix} U_n & -1 \\ -L_n & 1 \end{bmatrix} \right)$ from (24) establishes the 2^n functions $F_K(x)$ of (17) as a basis for the vector space consisting of all multilinear polynomials having degree at most n. This space clearly has dimension 2^n, and the invertibility demonstrates the functions to be linearly independent. As a result, associated with every multilinear polynomial $\sum_{J \subseteq N} \alpha_j \prod_{j \in J} x_j$, there exists a unique $\pi \in \mathbb{R}^{2^n}$ so that $p(x) = \sum_{K \subseteq N} \pi_K F_K(x)$ holds. This identity can be readily solved in terms of π_K due to the structure of the functions $F_K(x)$ to obtain the claimed expression for π. $\qquad\square$

Proof of Corollary 4.3. The multipliers π_K of Corollary 4.2 can be used to establish, in terms of the extreme points of the set X', the nonnegativity of $\sum_{J \subseteq N} \alpha_J \prod_{j \in J} x_j$ over X', and the collection of points at which the polynomial vanishes. Clearly, such a polynomial is nonnegative over X' if and only if $\pi_K \geqslant 0$ for all K. Corollary 4.2 uniquely defines each such multiplier to be the positive scalar $\frac{p(\hat{x})}{F_K(\hat{x})}$ evaluated at a distinct extreme point $\hat{x}$. $\qquad\square$

4.3   Exactness of Core Inequalities

In this section, we exploit the RLT properties to characterise the set of points $(x, y) \in G$ at which each valid core inequality is satisfied exactly. We consider only core inequalities since other inequalities can be handled via permutations of the variables x_j to coincide with the permutations of the coefficients β_j. We begin with three propositions and a theorem. The first proposition provides a necessary and sufficient condition for a point $\tilde{x} \in X$ to satisfy a function $F_K(x)$ of (16) exactly, and the second gives three equivalent conditions for exactness to hold. The third proposition uses the RLT results of Observations 1 and 2 as a theoretical bridge to translate the conditions of the second proposition to that of satisfying a valid inequality for $\text{conv}\,G$ exactly. The theorem uses this third proposition to give a necessary and sufficient condition for a point $(\tilde{x}, \tilde{y}) \in G$ to satisfy a valid core inequality exactly. We subsequently present two special cases of the theorem as corollaries. Consistent with our earlier work, we then separately address the core facets when the SMP $m(x)$ is supermodular over the extreme points of X of (1c), and when $m(x)$ is a monomial having $-\ell = u > 0$. The two corollaries serve

to simplify this analysis. Throughout, the inequalities $F_K(x) \geqslant 0$ of (18) are assumed to have $L_j = \ell$ and $U_j = u$ for all $j \in N$ within (16) as in X of (1c).

Lemma 4.4. *Given any $K \subseteq N$ and any $\tilde{x} \in X$ of (1c), partition N into N_1, N_2, and N_3 so that $N_1 \equiv \{j : \tilde{x}_j = \ell\}$, $N_2 \equiv \{j : \tilde{x}_j = u\}$, and $N_3 \equiv \{j : \ell < \tilde{x}_j < u\}$. Then $\tilde{x}$ satisfies $F_K(x) = 0$ if and only if $\big([K \cap N_1] \cup [(N \setminus K) \cap N_2]\big) \neq \emptyset$.*

Proof. Trivial by the definition of the polynomial $F_K(x)$ in (16). $\square$

Proposition 4.5. *Given any $K \subseteq N$ and any $\tilde{x} \in X$ of (1c), partition N into N_1, N_2, and N_3 so that $N_1 \equiv \{j : \tilde{x}_j = \ell\}$, $N_2 \equiv \{j : \tilde{x}_j = u\}$, and $N_3 \equiv \{j : \ell < \tilde{x}_j < u\}$. The following statements are equivalent:*

1. *$\tilde{x}$ satisfies $F_K(x) = 0$,*

2. *$\hat{x}$ satisfies $F_K(x) = 0$ at every $\hat{x} \in X$ having $\hat{x}_j = \ell \ \forall \ j \in N_1$ and $\hat{x}_j = u \ \forall \ j \in N_2$,*

3. *$\hat{x}$ satisfies $F_K(x) = 0$ at every extreme point $\hat{x}$ of X having $\hat{x}_j = \ell \ \forall \ j \in N_1$ and $\hat{x}_j = u \ \forall \ j \in N_2$.*

Proof. Given any $K \subseteq N$ and any $\tilde{x} \in X$ with N partitioned into N_1, N_2, and N_3 as described, the proof is to show that $1 \to 2 \to 3 \to 1$. The case $2 \to 3$ is trivial and so we consider the remaining two cases. For convenience, we define the set $\mathcal{N}_K \subseteq N$ as $\mathcal{N}_K \equiv \big([K \cap N_1] \cup [(N \setminus K) \cap N_2]\big)$.
$(1 \to 2)$ Lemma 4.4 gives us that $\tilde{x}$ satisfies $F_K(x) = 0$ only if $\mathcal{N}_K \neq \emptyset$. But $\mathcal{N}_K$ remains nonempty for every $\hat{x} \in X$ having $\hat{x}_j = \ell \ \forall \ j \in N_1$ and $\hat{x}_j = u \ \forall \ j \in N_2$ so that the "if" condition of Lemma 4.4 establishes the result.
$(3 \to 1)$ Consider that extreme point $\hat{x}$ of X having $\hat{x}_j = \ell \ \forall \ j \in N_1$, $\hat{x}_j = u \ \forall \ j \in N_2$, $\hat{x}_j = \ell \ \forall \ j \in (N_3 - K)$, and $\hat{x}_j = u \ \forall \ j \in (K \cap N_3)$. Then

$$F_K(x) = 0 = (u - \ell)^{(n - |\mathcal{N}_K|)} \left(\prod_{j \in (K \cap N_1)} (\hat{x}_j - \ell) \prod_{j \in (N \setminus K) \cap N_2} (u - \hat{x}_j) \right)$$

when evaluated at $\hat{x}$, implying that $\mathcal{N}_K \neq \emptyset$. Then the "if" condition of Lemma 4.4 establishes the result. $\square$

The result below mirrors that of Proposition 4.5, but invokes Observations 1 and 2 to extend the results from the functions $F_K(x)$ to the inequalities (2).

Proposition 4.6. *Given any inequality (2) that is valid for $\mathrm{conv}\, G$ and any $\tilde{x} \in X$ of (1c), partition N into N_1, N_2, and N_3 so that $N_1 \equiv \{j : \tilde{x}_j = \ell\}$, $N_2 \equiv \{j : \tilde{x}_j = u\}$, and $N_3 \equiv \{j : \ell < \tilde{x}_j < u\}$. The following statements are equivalent:*

1. *$(x, y) = (\tilde{x}, m(\tilde{x}))$ satisfies (2) exactly,*

2. *$(x, y) = (\hat{x}, m(\hat{x}))$ satisfies (2) exactly at every $\hat{x} \in X$ having $\hat{x}_j = \ell \ \forall \ j \in N_1$ and $\hat{x}_j = u \ \forall \ j \in N_2$,*

3. *$(x, y) = (\hat{x}, m(\hat{x}))$ satisfies (2) exactly at every extreme point $\hat{x}$ of X having $\hat{x}_j = \ell \ \forall \ j \in N_1$ and $\hat{x}_j = u \ \forall \ j \in N_2$.*

Proof. Given any $K \subseteq N$ and any $\tilde{x} \in X$ with N partitioned into N_1, N_2, and N_3 as described, the proof is to show that $1 \to 2 \to 3 \to 1$. The case $2 \to 3$ is trivial and so we consider the remaining two cases. To begin, note that Observations 1 and 2 combine to show, given any

inequality (2) that is valid for $\text{conv}\,G$, there exists a unique, nonnegative $\pi \in \mathbb{R}^{2^n}$ satisfying $\beta_0 + \sum_{j=1}^{n} \beta_j x_j + \beta' m(x) = \prod_{K \subseteq N} \pi_K F_K(x)$ so that, given any $\bar{x} \in X$,

$$\beta_0 + \sum_{j=1}^{n} \beta_j \bar{x}_j + \beta' m(\bar{x}) = 0 \text{ if and only if every } \pi_K > 0 \text{ has } F_K(x) = 0 \text{ at } \bar{x}. \qquad (28)$$

$(1 \to 2)$ Given $(x, y) = (\tilde{x}, m(\tilde{x}))$ satisfies (2) exactly, the "only if" direction of (28) with $\bar{x} = \tilde{x}$ gives us that every $\pi_K > 0$ has $F_K(x) = 0$ at $\tilde{x}$. Implication $1 \to 2$ of Proposition 4.5 then gives us that every $\pi_K > 0$ has $F_K(x) = 0$ at every $\hat{x} \in X$ having $\hat{x}_j = \ell \; \forall \; j \in N_1$ and $\hat{x}_j = u \; \forall \; j \in N_2$. Then the "if" direction of (28) with each such $\hat{x}$ substituted for $\bar{x}$ gives the result.

$(3 \to 1)$ Given $(x, y) = (\hat{x}, m(\hat{x}))$ satisfies (2) exactly at every extreme point $\hat{x}$ of X having $\hat{x}_j = \ell \; \forall \; j \in N_1$ and $\hat{x}_j = u \; \forall \; j \in N_2$, the "only if" direction of (28) with each such $\hat{x}$ substituted for $\bar{x}$ gives us that every $\pi_K > 0$ has $F_K(x) = 0$ at every such $\hat{x}$. Implication $3 \to 1$ of Proposition 4.5 then gives us that every $\pi_K > 0$ has $F_K(x) = 0$ at $\tilde{x}$. Then the "if" direction of (28) with $\bar{x} = \tilde{x}$ gives the result. $\qquad \square$

We invoke Lemma 3.1 and Proposition 4.6 to establish a theorem and two corollaries. The theorem gives, in terms of the extreme points of the simplex $\mathcal{S}$ of (3), a necessary and sufficient condition for a valid core inequality (2) to be satisfied exactly at a point $(\tilde{x}, \tilde{y}) \in G$. This theorem is a generalization of Lemma 3.1 in that, by restricting $n = (p + b)$ within the theorem, the stated point $(\tilde{x}, \tilde{y}) \in G$ must be the extreme point $(\tilde{x}, m(\tilde{x})) \in G$ of $\text{conv}\,G$ found within Lemma 3.1. The corollaries are special cases of the theorem when the valid core inequality (2): is satisfied exactly at $(x^j, m(x^j))$ for all $j \in \{0, \ldots, n\}$, and has all coefficients β_j equal to the same scalar, say $\bar{\beta}$, respectively.

Theorem 4.7. *Given a point $(\tilde{x}, \tilde{y}) \in G$ with $\tilde{x}$ not an extreme point of the simplex $\mathcal{S}$ of (3), let r be the smallest index j such that $\tilde{x}_j < u$ and s be the largest index j such that $\tilde{x}_j > \ell$. Further let p be the number of entries of $\tilde{x}$ having value u and b be the number of entries of $\tilde{x}$ having value ℓ. Then $(\tilde{x}, \tilde{y})$ satisfies a valid core inequality (2) exactly if and only if $(x, y) = (x^j, m(x^j))$ satisfies (2) exactly for all $j \in \{p, \ldots, n - b\}$, and $\beta_r = \ldots = \beta_s$.*

Proof. If $n = (p + b)$ so that all entries of $\tilde{x}$ take either value ℓ or u, then $(\tilde{x}, \tilde{y}) \in G$ is the extreme point $(\tilde{x}, m(\tilde{x}))$ of $\text{conv}\,G$, so the result is Lemma 3.1. Otherwise, $n \geqslant (p + b + 1)$, and we adopt the notation of Proposition 4.6 that $N_1 \equiv \{j : \tilde{x}_j = \ell\}$, $N_2 \equiv \{j : \tilde{x}_j = u\}$, and $N_3 \equiv \{j : \ell < \tilde{x}_j < u\}$, so that $|N_1| = b$, $|N_2| = p$, and $|N_3| = n - (p + b)$. For any chosen $k \in \{0, \ldots, n - (p + b)\}$, define a new point $\hat{x}$ in terms of $\tilde{x}$ by setting k values of $\hat{x}_j$ with $j \in N_3$ to the value u and the remaining $(n - (p + b) - k)$ values of $\hat{x}_j$ with $j \in N_3$ to the value ℓ, and by setting $\hat{x}_j = \tilde{x}_j$ for all $j \in (N_1 \cup N_2)$. Then, if $\hat{x} \neq x^{k+p}$, Lemma 3.1 gives us that $(x, y) = (\hat{x}, m(\hat{x}))$ satisfies (2) exactly if and only if $(x^{k+p}, m(x^{k+p}))$ satisfies (2) exactly, and $\beta_{\hat{r}} = \ldots = \beta_{\hat{s}}$, where $\hat{r}$ is the smallest index j such that $\hat{x}_j = \ell$ and $\hat{s}$ is the largest index j such that $\hat{x}_j = u$. Consequently, by considering all $k \in \{0, \ldots, n - (p + b)\}$ and all corresponding $\binom{n-(p+b)}{k}$ possible fixings of $\hat{x}_j$ for $j \in N_3$, $(x, y) = (\hat{x}, m(\hat{x}))$ will satisfy (2) exactly at every extreme point $\hat{x}$ of X having $\hat{x}_j = \ell$ for all $j \in N_1$ and $\hat{x}_j = u$ for all $j \in N_2$ if and only if $(x, y) = (x^j, m(x^j))$ satisfies (2) exactly for all $j \in \{p, \ldots, n - b\}$, and $\beta_r = \ldots = \beta_s$. The "only if" consequence that $\beta_r = \ldots = \beta_s$ is due to the following: if $r = s$, the result is trivial while if $r < s$, then any $\hat{x}$ defined as above so that $\hat{x}_r = \ell$ and $\hat{x}_s = u$, regardless of k, will yield $r = \hat{r}$ and $s = \hat{s}$. Then implication $3 \to 1$ of Proposition 4.6 gives us that $(\tilde{x}, \tilde{y})$ satisfies (2) exactly, as desired. $\qquad \square$

17

Theorem 4.7 simplifies when the valid core inequality (2) is satisfied exactly at $(x^j, m(x^j))$ for all $j \in \{0, \ldots, n\}$. In this case, the parameters p and b within the theorem are no longer needed, as they serve only to restrict the points $(x^j, m(x^j))$ that must satisfy the inequality exactly. The simplification is below.

Corollary 4.8. *Given a point $(\tilde{x}, \tilde{y}) \in G$ with $\tilde{x}$ not an extreme point of the simplex $\mathcal{S}$ of (3), let r be the smallest index j such that $\tilde{x}_j < u$ and s be the largest index j such that $\tilde{x}_j > \ell$. Given a valid core inequality (2) that is satisfied exactly at $(x^j, m(x^j))$ for all $j \in \{0, \ldots, n\}$, the point $(\tilde{x}, \tilde{y})$ satisfies this inequality exactly if and only if $\beta_r = \ldots = \beta_s$.*

Theorem 4.7 also simplifies when the valid core inequality (2) has, for some scalar $\bar{\beta}$, $\beta_j = \bar{\beta}$ for all $j \in N$. In this case, the parameters r and s within the theorem are no longer needed, as they serve only to restrict a subset of the coefficients β_j to equal. The simplification is stated formally below.

Corollary 4.9. *Given a point $(\tilde{x}, \tilde{y}) \in G$ with $\tilde{x}$ not an extreme point of the simplex $\mathcal{S}$ of (3), let p be the number of entries of $\tilde{x}$ having value u and b be the number of entries of $\tilde{x}$ having value ℓ. Then $(\tilde{x}, \tilde{y})$ satisfies a valid core inequality (2) having $\beta_j = \bar{\beta}$ for all $j \in N$ exactly if and only if $(x, y) = (x^j, m(x^j))$ satisfies (2) exactly for all $j \in \{p, \ldots, n - b\}$.*

5   Supermodular Functions

In this subsection, we describe $\text{conv} \, G$ for SMPs $m(x)$ that are *supermodular* over the extreme points of X. Such a function $m(x)$ over these 2^n points can be expressed as a set function $f(A)$ defined over $A \subseteq N$ so that $f(A) = m(x)$ when evaluated at that extreme point x having $x_j = u$ for $j \in A$ and $x_j = \ell$ for $j \notin A$. Recall that a set function f is defined to be supermodular over N if and only if

$$f(S \cup \{r\}) - f(S) \leqslant f(S \cup \{r, t\}) - f(S \cup \{t\}) \ \text{ for } r, t \in N, r \neq t, \text{ and } S \subseteq N \backslash \{r, t\}, \quad (29)$$

we have that the function $m(x)$ is supermodular over the 2^n extreme points of X if and only if

$$m(x^k) - m(x^{k-1}) \leqslant m(x^{k+1}) - m(x^k), \quad \forall \, k \in \{1, \ldots, n - 1\}. \tag{30}$$

This equivalence follows by considering, for each $k \in \{1, \ldots, n - 1\}$, all sets $S \subseteq N$ within (29) having $|S| = (k - 1)$, and all $\{r, t\} \in N \backslash S$, and by invoking the symmetry of $m(x)$ to obtain that $f(T) = m(x^k)$ for all $T \subseteq N$ with $|T| = k$. The function $m(x)$ is *strictly supermodular* over the 2^n extreme points of X if and only if the $(n - 1)$ inequalities of (30) are satisfied strictly. Henceforth, for brevity, we will refer to functions $m(x)$ that are (strictly) supermodular over the 2^n extreme points of X as being (strictly) supermodular.

The below theorem explicitly states all core facets for supermodular $m(x)$.

Theorem 5.1. *When $m(x)$ is supermodular, there exist at most $(n+1)$ core facets for $\text{conv} \, G$, and all such facets, subject to repetition, are*

$$\frac{u m(x^0) - \ell m(x^n)}{u - \ell} + \sum_{j=1}^{n} \left(\frac{m(x^j) - m(x^{j-1})}{u - \ell} \right) x_j - y \geqslant 0, \tag{31}$$

and

$$-m(x^k) - \left(\frac{m(x^k) - m(x^{k-1})}{u - \ell} \right) \left(\sum_{j=1}^{n} x_j - [ku + (n - k)\ell] \right) + y \geqslant 0, \quad \forall \, k \in N. \tag{32}$$

18

Proof. The $(n+1)$ inequalities of (31) and (32) are valid for $\operatorname{conv} G$ by Lemma 2.3, as they are readily verified to hold for all $(x^k, m(x^k)), k \in \{0, \dots, n\}$. Inequality (31) is a facet because these same $(n+1)$ affinely independent points satisfy it exactly, and it is a core facet because the β_j are nondecreasing by (30). In addition, no other core facet can exist with $\beta' = -1$ by Proposition 3.4 because the functions $\mathcal{L}_k(\beta_0, \beta)$ from (5), with (β_0, β) defined in terms of (31), have $\mathcal{L}_k(\beta_0, \beta) = m(x^k)$ for all $k \in \{0, \dots, n\}$. Relative to (32), for each $k \in N$, the corresponding inequality is a core facet by Proposition 3.3, as it is satisfied exactly at the two points $(x^{k-1}, m(x^{k-1}))$ and $(x^k, m(x^k))$. To show that no other core facet can exist with $\beta' = 1$ and complete the proof, it is sufficient to show that every core facet $\beta_0 + \sum_{j=1}^n \beta_j x_j + y \geqslant 0$ which is satisfied exactly at some $(x^r, m(x^r))$ and $(x^s, m(x^s))$ with $r < s$, is also satisfied exactly at $(x^{r+1}, m(x^{r+1}))$. Then induction has the core facet being satisfied exactly at $(x^p, m(x^p))$ for all $p \in \{r, \dots, s\}$ so that, for each $p \in \{r+1, \dots, s\}$, by inserting $(x^{p-1}, m(x^{p-1}))$ and $(x^p, m(x^p))$ into the facet and subtracting the first expression from the second, we obtain $\beta_p(u - \ell) = -\left(m(x^p) - m(x^{p-1}) \right)$. As the β_p are nondecreasing and the values $-\left(m(x^p) - m(x^{p-1}) \right)$ are nonincreasing, it will then follow that $\beta_p(u - \ell) = -\left(m(x^{r+1}) - m(x^r) \right)$ for all $p \in \{r+1, \dots, s\}$. Hence, by selecting $r = \operatorname{argmin}_{k \in \{0,\dots,n-1\}} \left\{ \mathcal{L}_k(\beta_0, \beta) + m(x^k) = 0 \right\}$ and $s = \operatorname{argmax}_{k \in \{1,\dots,n\}} \left\{ \mathcal{L}_k(\beta_0, \beta) + m(x^k) = 0 \right\}$, Proposition 3.7 will give us that $\beta_j = -\left(\frac{m(x^{r+1}) - m(x^r)}{u - \ell} \right)$ for all $j \in N$. Then the chosen core facet must be inequality (32) with $k = (r+1)$.

To show that the core facet is satisfied exactly at $(x^{r+1}, m(x^{r+1}))$, we restrict attention to $s \geqslant (r+2)$ since the result is trivial for $s = (r+1)$. We have that

$$
\begin{aligned}
m(x^{r+1}) \quad \leqslant \quad & m(x^r) + \frac{m(x^s) - m(x^r)}{s - r} \\
= \quad & -\beta_0 - u \left(\sum_{j=1}^r \beta_j \right) - \ell \left(\sum_{j=r+1}^n \beta_j \right) - \left(\frac{u - \ell}{s - r} \right) \left(\sum_{j=r+1}^s \beta_j \right) \\
= \quad & -\beta_0 - u \left(\sum_{j=1}^{r+1} \beta_j \right) - \ell \left(\sum_{j=r+2}^n \beta_j \right) - \left(\frac{u - \ell}{s - r} \right) \left(\sum_{j=r+1}^s (\beta_j - \beta_{r+1}) \right) \\
\leqslant \quad & -\beta_0 - u \left(\sum_{j=1}^{r+1} \beta_j \right) - \ell \left(\sum_{j=r+2}^n \beta_j \right) \\
\leqslant \quad & m(x^{r+1}),
\end{aligned}
$$

where the first inequality follows from summing the $(s - r)$ inequalities $m(x^{r+1}) - m(x^r) \leqslant m(x^{k+1}) - m(x^k)$ for $k \in \{r, \dots, s-1\}$ that are implied by (30) for $k > r$, the first equality follows from the facet holding exactly at $(x^r, m(x^r))$ and $(x^s, m(x^s))$, the second equality is algebra, the second inequality follows from the nondecreasing values of β_j, and the final inequality follows from the feasibility of $(x^{r+1}, m(x^{r+1}))$ to $\operatorname{conv} G$. Then $m(x^{r+1}) = -\beta_0 - u \left(\sum_{j=1}^{r+1} \beta_j \right) - \ell \left(\sum_{j=r+2}^n \beta_j \right)$ so that the core facet is satisfied exactly at $(x^{r+1}, m(x^{r+1}))$. The proof is complete. $\qquad\square$

Theorem 5.1 states that there exist *at most* $(n+1)$ core facets because the inequalities of (32) can repeat. Repetition will occur whenever $\left(m(x^p) - m(x^{p-1}) \right) = \left(m(x^q) - m(x^{q-1}) \right)$ for distinct $p, q \in N$. Notably, if $m(x)$ is strictly supermodular, then repetitions will not occur so that (32) will contain n distinct facets.

19

Subject to permutations of β, inequalities (31) define the concave envelope of $m(x)$. These inequalities are a special case of the polymatroid inequalities that are known for general supermodular functions [Lov83; TRX13]. Polymatroid inequalities are known to be separated easily in $O(n \log n)$ time using a sorting algorithm since Edmonds [Edm70] showed how to optimize a linear function over the polymatroid polyhedron corresponding to a submodular function. Inequalities (32) have equal coefficients on the variables and hence they do not need to be permuted and define the convex envelope of $m(x)$. Our proof for these inequalities is an alternative proof to that of Tawarmalani et al. [TRX13, Theorem 4.6]. For a submodular function $m(x)$ (meaning that the inequality in (30) is switched to $\geqslant$), inequalities (31), subject to permutations of β, define the convex envelope, and inequalities (32) define the concave envelope subject to the following modifications: all occurrences of $m(x^j)$, $j \in \{0, \ldots, n\}$, and y are negated.

Now we use Corollarys 4.8 and 4.9 to identify, for each of the $(n+1)$ core facets (31) and (32) of Theorem 5.1, the set of all points $(x, y) \in G$ that satisfies it exactly. A key ingredient of Theorem 5.1 that is useful in our upcoming proof is that every core facet $\beta_0 + \sum_{j=1}^{n} \beta_j x_j + y \geqslant 0$ of the form (32) which is satisfied exactly at some $(x^r, m(x^r))$ and $(x^s, m(x^s))$ with $r < s$, is also satisfied exactly at $(x^p, m(x^p))$ for all $p \in \{r, \ldots, s\}$. We then have that the set of points $(x^j, m(x^j))$ which satisfies the facet exactly must be consecutive.

Theorem 5.2. *Let $m(x)$ be supermodular and $(x, y) \in G$.*

1. *(x, y) satisfies (31) exactly if and only if $x = x^k$ for some k, or $\beta_r = \cdots = \beta_s$ where $r = \min\{j : x_j < u\}$ and $s = \max\{j : x_j > \ell\}$.*

2. *(x, y) satisfies (32) exactly if and only if $(x^j, m(x^j))$ satisfies (32) exactly for $j = p, \ldots, n - q$, where $p = |\{j : x_j = u\}|$ and $q = |\{j : x_j = \ell\}|$.*

3. *(x, y) satisfies (32) exactly if and only if $|\{j : x_j = u\}| \geqslant v$ and $|\{j : x_j = \ell\}| \geqslant n - w$, where v and w are the smallest and largest, respectively, values of j such that the point $(x^j, m(x^j))$ satisfies (32) exactly.*

Proof. (1) As noted in the proof of Theorem 5.1, the points $(x^k, m(x^k))$ satisfy (31) exactly for all $k \in \{0, \ldots, n\}$. Then the result trivially holds true when x is an extreme point of $\mathcal{S}$, and it holds true when x is not an extreme point of $\mathcal{S}$ by Corollary 4.8.

(2) The result trivially holds true when x is an extreme point of $\mathcal{S}$, and it holds true when x is not an extreme point of $\mathcal{S}$ by Corollary 4.9 with $\bar{\beta} = -\left(\frac{m(x^k) - m(x^{k-1})}{u - \ell}\right)$.

(3) This is a restatement of above using Remark 6 of Theorem 5.1 which states that the set of extreme points of $\mathcal{S}$ satisfying the facet exactly must be consecutive. $\square$

More refined characterizations for exactness hold when the function is strictly supermodular.

Corollary 5.3. *Let $m(x)$ be strictly supermodular and $(x, y) \in G$.*

1. *(x, y) satisfies (31) exactly if and only if x is a convex combination of two consecutive extreme points x^j and x^{j+1} of $\mathcal{S}$.*

2. *(x, y) satisfies (32) exactly if and only if $|\{j : x_j = u_j\}| \geqslant k - 1$ and $|\{j : x_j = \ell\}| \geqslant n - k$.*

Proof. (1) Follows directly from Theorem 5.2 since the coefficients $\beta_j = \left(\frac{m(x^j) - m(x^{j-1})}{u - \ell}\right)$ for all $j \in N$ of (31) are distinct, as every inequality within (30) is satisfied strictly for strictly supermodular $m(x)$.

(2) Consider any $k \in N$. As noted in the proof of Theorem 5.1 and readily verified, $(x, y) = (x^{k-1}, m(x^{k-1}))$ and $(x, y) = (x^k, m(x^k))$ each satisfy the core facet exactly. Thus, it is sufficient

to show that supermodular $m(x)$ enforces $r = (k-1)$ and $s = k$ in Theorem 5.2. By contradiction, suppose that $s \geqslant (r+2)$ so that each of $(x^r, m(x^r))$, $(x^{r+1}, m(x^{r+1}))$, and $(x^{r+2}, m(x^{r+2}))$ satisfy the facet exactly. As in the proof of Theorem 5.1, for each $p \in \{r+1, r+2\}$, by inserting $(x^{p-1}, m(x^{p-1}))$ and $(x^p, m(x^p))$ into the facet and subtracting the second expression from the first, we obtain

$$\frac{-\big(m(x^{r+1}) - m(x^r)\big)}{u - \ell} = \beta_{r+1} = \beta_{r+2} = \frac{-\big(m(x^{r+2}) - m(x^{r+1})\big)}{u - \ell},$$

where $\beta_{r+1} = \beta_{r+2}$ is due to the form of (32). The strict inequality of (30) for supermodular $m(x)$ with $k = (r+1)$ yields the contradiction that $\beta_{r+1} > \beta_{r+2}$. $\qquad \square$

We finish this section by remarking on supermodularity of an SMP, giving us some important families of functions $m(x)$ so that Theorem 5.1 characterizes their convex hull. First we provide an alternate characterization to (30) by expressing these requirements in terms of the coefficients c_j of the function m.

Proposition 5.4. $m(x)$ is supermodular over X if and only if $\sum_{d=2}^{n} c_d (u - \ell)^2 \vartheta_{k,d} \geqslant 0$ for $k = 1, \ldots, n-1$, where $\vartheta_{2,2} = 1$ and

$$\vartheta_{k,d} = \sum_{\substack{J \subseteq N \setminus \{k, k+1\} \\ |J| = d-2}} \prod_{j \in J} x_j^{k-1}, \qquad d = 3, \ldots, n, \ k = 1, \ldots, n-1,$$

Proof. For each $k \in \{1, \ldots, n-1\}$, the difference $m(x^{k+1}) - m(x^k) - \big(m(x^k) - m(x^{k-1})\big)$ is computable in terms of only those expressions $\prod_{j \in J} x_j$ within $m(x)$ that contain both $k \in J$ and $\{k+1\} \in J$. For $d = 2$, this difference for the $\binom{n}{d}$ expressions in $m(x)$ of degree d is given by $c_2 (u - \ell)^2$, while for each $d \in \{3, \ldots, n\}$, it is given by

$$c_d (u - \ell)^2 \left(\sum_{\substack{J \subseteq N - \{k, k+1\} \\ |J| = d-2}} \left(\prod_{j \in J} x_j^{k-1} \right) \right). \tag{33}$$

Then (30) is equivalent to having the sum of these expressions from 2 to n being nonnegative for all $k \in \{1, \ldots, n-1\}$. $\qquad \square$

A multilinear monomial $x_1 x_2 \cdots x_n$ is supermodular over the nonnegative orthant [cf. TRX13]. Since supermodularity is preserved under taking nonnegative combinations of functions, it follows that an SMP with $c_j \geqslant 0$ for all j is supermodular over $\mathbb{R}_+^n$. This fact for nonnegative valued SMPs also follows from our characterization.

Corollary 5.5. $m(x)$ is supermodular over X if $\ell \geqslant 0$ and $c_j \geqslant 0$ for all $j = 2, \ldots, n$.

Proof. $\ell \geqslant 0$ implies that $x_j^k \geqslant 0$ for all j, k, which implies $\vartheta_{k,d} \geqslant 0$ for all k, d. The summation in Proposition 5.4 is nonnegative because $c_j \geqslant 0$ for all j. $\qquad \square$

Another consequence is that symmetric quadratic polynomials are always either submodular or supermodular, regardless of the box in $\mathbb{R}^n$.

Corollary 5.6. The symmetric quadratic polynomial $\tau \sum_{i \neq j} x_i x_j$ is either submodular or supermodular over X for any $\tau \neq 0$.

Proof. A symmetric quadratic polynomial is $m(x)$ with $c_3 = \cdots = c_n = 0$, and $c_2 = \tau$. Since all monomials are of the same degree 2, we can assume wlog that $\ell \geqslant 0$ because otherwise we can negate all the variables and consider the reflected box $X' = \{x' : -u_j \leqslant x'_j \leqslant -\ell_j\}$. Applying Corollary 5.5 to $m(x)$ if $\tau > 0$ or to $-m(x)$ if $\tau < 0$ yields the desired claim. $\qquad \square$

When considering the unit hypercube in $\mathbb{R}^n$, supermodularity is attained through nonnegativity of a partial sum. Denote $\binom{0}{0} = 1$.

Corollary 5.7. *$m(x)$ is supermodular over $[0, 1]^n$ if and only if*

$$\sum_{d=2}^{k+1} \binom{k-1}{d-2} c_d \geqslant 0, \quad \forall \, k \in \{1, \ldots, n-1\}. \tag{34}$$

Proof. Follows immediately by simplifying the sum in (33). $\qquad \square$

More general families of SMPs are encompassed by (34) including, for example, those having $c_d \geqslant 0$ for $d \in \{2, \ldots, n-1\}$ and $c_n \geqslant -\sum_{d=2}^{n-1} \binom{n-2}{d-2} c_d$, which reduces to $c_2 \geqslant \max\{0, (2-n)c_3\}$ for cubic functions.

6 Monomials with Reflection Symmetry

We consider $m(x)$ to be a monomial $m(x) = c_n \prod_{j=1}^{n} x_j$ over the box X. The coefficient c_n can be taken to be 1 since we can scale y to y/c_n. Three cases arise depending on the location of the box in $\mathbb{R}^n$: (1) $\ell u = 0$, (2) $\ell u > 0$, and (3) $\ell u < 0$. After performing appropriate scalings, the first two cases are equivalent to that of $\ell = 0, u = 1$ and $\ell = 1, u = r$ for some fixed $r > 1$, respectively. These two cases fall within the class of supermodular SMPs (cf. Corollary 5.5), and so Theorem 5.1 gives their convex hull. Indeed, it is easily verified that the resulting description of the convex hull matches the known envelopes for $\prod_{j=1}^{n} x_j$ over $[0, 1]^n$ [Cra93] and over $[1, r]^n$ for some fixed $r > 1$ (see [AGX19, Proposition 4.1] which is a direct consequence of results from [Ben04, Theorem 1] and [TRX13, Theorem 4.6]). The third case $\ell u < 0$ is equivalent (upto scaling) to taking $\ell = -1$ and $u = r$ for some $r > 0$ so that $X = [-1, r]^n$, and it is readily verified that a monomial is neither submodular nor supermodular over $[-1, r]^n$. Note that since we are considering monomials, scaling means that if each variable x_j has different lower and upper bounds ℓ_j and u_j, then as long as $u_j = -\ell_j$ we can reduce to the case $X = [-1, r]^n$. The convex hull for the specific subcase having $r = 1$ was first established by the authors in [AGX19, Theorem 4.1]. However, it was done so without using the symmetry of the monomial and hence did not recognize the core facets. Our main goal in this section is to independently describe this convex hull by identifying the core facets and their properties.

Theorem 6.1. *For $m(x) = \prod_{j=1}^{n} x_j$ and $-\ell = u = 1$, there exist precisely $n + 3$ core facets,*

and these are

$$1 - y \geqslant 0 \;\; and \;\; 1 + y \geqslant 0 \tag{35}$$

$$(n-1) + \sum_{j=1}^{n} x_j + (-1)^n y \geqslant 0, \tag{36}$$

$$(n-1) - \sum_{j=1}^{n} x_j + y \geqslant 0. \tag{37}$$

$$(n-1) - \left(\sum_{j=1}^{t+1} x_j - \sum_{j=t+2}^{n} x_j \right) + (-1)^{n-(t+1)} y \geqslant 0, \quad \forall\, t \in \{0, \ldots, n-2\}. \tag{38}$$

The general case of convexifying $\prod_{j=1}^{n} x_j$ over $[-1, r]^n$ for $r > 0, r \neq 1$ remains an open question. Before giving our proof for the above theorem, let us note the complexity of separation and comment on the structure of the proposed inequalities which includes identifying exactness of the core facets.

Corollary 6.2. *For $m(x) = \prod_{j=1}^{n} x_j$ and $-\ell = u = 1$, a point can be separated from $\operatorname{conv} G$ in $O(n \log n)$ time.*

Proof. Theorem 6.1 tells us that there are $t = n + 3$ core facets. Proposition 2.5 then implies that the complexity of separation is $O(n^2 + n \log n)$. In the proof of this result, the complexity $O(nt)$ came from having to take a summation of n terms while separating each of the t core facets. Since the summation in both (36) and (37) is the same for any permutation of variables and therefore can be stored and reused, we get that the overall complexity is $O(t + n \log n) = O(n \log n)$. $\square$

6.1 Structure of the inequalities

All facets for $\operatorname{conv} G$ are computable in terms of the core facets enumerated in Theorem 6.1 via permutations of β. Since each of the four core facets of (35), (36), and (37) have all β_j equal, no permutations of β will produce additional facets. However, for each $t \in \{0, \ldots, n-2\}$, the core facet (38) admits $\binom{n}{t+1}$ facets. Then (36), (37), and (38) motivate the 2^n inequalities

$$(n-1) - \left(\sum_{j \in J} x_j - \sum_{j \in N \setminus J} x_j \right) + (-1)^{n-|J|} y \geqslant 0, \quad \forall\, J \subseteq N, \tag{39}$$

where (39) with $J = \emptyset$ is (36), where (39) with $J = N$ is (37), and where (39) for each $J \subseteq N$ with $|J| \in \{1, \ldots, n-1\}$ are the $\binom{n}{|J|}$ facets obtained by permutations of β in (38) when $(t+1) = |J|$. The desired representation of $\operatorname{conv} G$ is then the two inequalities $-1 \leqslant y \leqslant 1$ of (35), the 2^n inequalities of (39) and, for the case in which $n \geqslant 3$, the $2n$ inequalities $-1 \leqslant x_j \leqslant 1 \;\forall\, j \in N$. We can combine and more succinctly write the $(2 + 2^n + 2n)$ inequalities of (35), (39), and the restrictions $-1 \leqslant x_j \leqslant 1$ for $j \in N$ for the cases having $n \geqslant 3$, by letting the variable x_{n+1} denote y and by letting $N' = \{1, \ldots, n+1\}$. Using this new definition of variables, (39) can be partitioned in terms of the exponents $(n - |J|)$ as

$$(n-1) - \left(\sum_{j \in J} x_j - \sum_{j \in (N \setminus J) \cup \{n+1\}} x_j \right) \geqslant 0, \quad \forall\, J \subseteq N \text{ with } n - |J| \text{ even,}$$

and

$$(n-1) - \left(\sum_{j \in J \cup \{n+1\}} x_j - \sum_{j \in N \setminus J} x_j \right) \geqslant 0, \quad \forall\, J \subseteq N \text{ with } n - |J| \text{ odd}.$$

Consequently, $-1 \leqslant x_j \leqslant 1$ for $j \in N$, (35), and (39) can be expressed as (40) and (41) below, where (40) encompasses $-1 \leqslant x_j \leqslant 1$ for $j \in N$ and (35), and (41) encompasses the above two families of inequalities.

$$-1 \leqslant x_j \leqslant 1 \qquad\qquad \forall\, j \in N' \qquad\qquad (40)$$

$$(n-1) - \left(\sum_{j \in J} x_j - \sum_{j \in N' \setminus J} x_j \right) \geqslant 0, \qquad \forall\, J \subseteq N' \text{ with } n + 1 - |J| \text{ odd} \qquad (41)$$

Of course, if $n = 2$, then (40) simplifies to $-1 \leqslant x_{n+1} \leqslant 1$.

Now we use Theorem 4.7 and Corollary 4.9 to identify, for each of the $(n+3)$ core facets (35), (36), (37), and (38) of Theorem 6.1, the set of all points $(x, y) \in G$ that satisfies it exactly.

Theorem 6.3. *For $m(x) = \prod_{j=1}^{n} x_j$ and $-\ell = u = 1$, $(\tilde{x}, \tilde{y}) \in G$ satisfies the core facet*

1. *(35) exactly if and only if $\tilde{x}$ contains*

    (a) *even number of entries of value -1 and the remaining entries of value 1 when $\beta' = -1$,*
    (b) *odd number of entries of value -1 and the remaining entries of value 1 when $\beta' = 1$,*

2. *(36) exactly if and only if $\tilde{x}$ contains at least $(n-1)$ entries of value -1,*
3. *(37) exactly if and only if $\tilde{x}$ contains at least $(n-1)$ entries of value 1,*
4. *(38) for $t \in \{0, \ldots, n-2\}$ if and only if $\tilde{x}$ differs from x^{t+1} in at most one entry.*

Proof. Each statement is considered separately.

1. Follows directly from the definition of $m(x)$ when $-\ell = u = 1$.
2. Lemma 6.4 and its proof give us that (36) is satisfied exactly at $(\tilde{x}, \tilde{y}) = (x^k, m(x^k))$ for $k \in \{0, 1\}$, but at no other k. Then Corollary 4.9 with $\bar{\beta} = 1$ and $\tilde{x}$ having $p = 0$ and $b = (n-1)$ gives the result.
3. Lemma 6.4 and its proof give us that (37) is satisfied exactly at $(\tilde{x}, \tilde{y}) = (x^k, m(x^k))$ for $k \in \{n-1, n\}$, but at no other k. Then Corollary 4.9 with $\bar{\beta} = -1$ and $\tilde{x}$ having $p = (n-1)$ and $b = 0$ gives the result.
4. For each $t \in \{0, \ldots, n-2\}$, Lemma 6.6 and its proof give us that the associated inequality of (38) is satisfied exactly at $(\tilde{x}, \tilde{y}) = (x^k, m(x^k))$ for $k \in \{t, t+1, t+2\}$, but at no other k. Then Theorem 4.7 with $\tilde{x}$ having $p = t$ and $b = (n - t - 1)$ allows $\tilde{x}$, with suitable r and s, to differ from x^{t+1} in only a single entry j less than or equal to $(t+1)$. Similarly, Theorem 4.7 with $\tilde{x}$ having $p = (t+1)$ and $b = (n - t - 1)$ allows $\tilde{x}$, with suitable r and s, to differ from x^{t+1} in only a single entry j greater than or equal to $(t+1)$. Combining these outcomes gives the result. $\square$

6.2 Proof of the Convex Hull

We prove Theorem 6.1 using three main lemmas. Let us begin with two simple observations. First, there exist no facets (2) with $\beta' = 0$ if $n = 2$, and that there exist $2n$ such facets of the form $-1 \leqslant x_j \leqslant 1$ for $j \in N$ if $n \geqslant 3$. Second, the two inequalities of (35) are core facets. They are trivially valid, and they are core facets by Proposition 3.3 because each has $\beta_j = 0$ for all

$j \in N$, and each is satisfied exactly at two extreme points x^k of $\mathcal{S}$, with one point not being x^0 or x^n; equality occurs at x^{n-2} and x^n for the left inequality and occurs at x^{n-3} and x^{n-1} for the right inequality.

Now, to facilitate the derivation of the remaining core facets and thereby characterize $\mathsf{conv}\,G$, we define, for each $(p,q), p < q$, the value

$$D(p,q) = 2 \sum_{j=p+1}^{q} \beta_j + \beta' \left(m(x^q) - m(x^p) \right)$$

that is computed by subtracting the left side of (2) evaluated at $(x,y) = (x^p, m(x^p))$ from the left side evaluated at $(x^q, m(x^q))$, upon recalling that $-\ell = u = 1$.

The following result identifies two core facets in terms of the extreme points x^k of $\mathcal{S}$, and also eliminates, in terms of these same points, other inequalities from consideration.

Lemma 6.4. *The only core facets that are satisfied exactly at $(x^k, m(x^k))$ for two consecutive extreme points x^k of the simplex $\mathcal{S}$, and for no other such extreme points, are (36) and (37).*

Proof. Suppose that a core facet is satisfied exactly at $(x^t, m(x^t))$ and $(x^{t+1}, m(x^{t+1}))$ for extreme points x^t and x^{t+1} of the simplex $\mathcal{S}$, and for no other such extreme points. Then Proposition 3.7 with $r = t$ and $s = (t+1)$ states that there exists a $\bar{\beta}$ so that $\bar{\beta} = \beta_j$ for all $j \in N$. Three cases exist.

- $t \in \{1, \ldots, n-2\}$. We have $D(t-1,t) = 2\bar{\beta} + \beta' \left(m(x^{t+2}) - m(x^{t+1}) \right) < 0$ because $m(x^t) = m(x^{t+2})$ and $m(x^{t-1}) = m(x^{t+1})$, and we have $D(t+1,t+2) = 2\bar{\beta} + \beta' \left(m(x^{t+2}) - m(x^{t+1}) \right) > 0$. Combining, $0 < 2\bar{\beta} + \beta' \left(m(x^{t+2}) - m(x^{t+1}) \right) < 0$, which is not possible.

- $t = 0$. We have $D(0,1) = 2\bar{\beta} + \beta' \left(2(-1)^{n+1} \right) = 0$, with the first equality holding since $m(x^0) = -m(x^1) = (-1)^n$. Then $\bar{\beta} = \beta'(-1)^n$ and (2) evaluated at $(x,y) = (x^0, m(x^0))$ gives $\beta_0 = \beta' \left((-1)^n (n-1) \right)$ because $m(x^0) = (-1)^n$. Inequality (2) becomes

$$\beta' \left((-1)^n(n-1) + (-1)^n \sum_{j=1}^{n} x_j + y \right) \geqslant 0.$$

  This inequality is valid when $\beta' = (-1)^n$ by the "if" direction of Lemma 2.3 since it holds true at $(x,y) = (x^k, m(x^k))$ for $k \in \{0, \ldots, n\}$, but it is invalid when $\beta' = (-1)^{n+1}$ since it is then violated at $(x,y) = (x^k, m(x^k))$ for $k \in \{2, \ldots, n\}$. The inequality with $\beta' = (-1)^n$ is (36), and is a facet by the "if" direction of Proposition 3.3 since it holds exactly at $(x,y) = (x^k, m(x^k))$ for $k \in \{0,1\}$.

- $t = (n-1)$. We have $D(n-1,n) = 2\bar{\beta} + 2\beta' = 0$, with the first equality holding since $-m(x^{n-1}) = m(x^n) = 1$. Then $\bar{\beta} = -\beta'$ and (2) evaluated at $(x,y) = (x^{n-1}, m(x^{n-1}))$ gives $\beta_0 = \beta'(n-1)$ because $m(x^{n-1}) = -1$. Inequality (2) becomes

$$\beta' \left((n-1) - \sum_{j=1}^{n} x_j + y \right) \geqslant 0.$$

  This inequality is valid when $\beta' = 1$ by the "if" direction of Lemma 2.3 since it holds true at $(x,y) = (x^k, m(x^k))$ for $k \in \{0, \ldots, n\}$, but it is invalid when $\beta' = -1$ since it is then violated at $(x,y) = (x^k, m(x^k))$ for $k \in \{0, \ldots, n-2\}$. The inequality with $\beta' = 1$ is (37), and is a facet by the "if" direction of Proposition 3.3 since it holds exactly at $(x,y) = (x^k, m(x^k))$ for $k \in \{n-1, n\}$. $\quad\square$

The below result builds further by providing, in terms of the extreme points x^k of $\mathcal{S}$, necessary conditions for valid core inequalities that are not found in (35), (36), or (37) to be facets.

Lemma 6.5. *Every core facet for $\mathsf{conv}\, G$ which is not found in (35), (36), or (37) is satisfied exactly at $(x^k, m(x^k))$ for three consecutive extreme points x^k of the simplex $\mathcal{S}$, and for no other such extreme points.*

Proof. Proposition 3.2 gives us that a core facet is satisfied exactly at $(x^k, m(x^k))$ for at least two extreme points of the simplex $\mathcal{S}$. Let r and s be as in Proposition 3.7. Let $d = (s - r) \geqslant 1$ denote the difference between s and r. Lemma 6.4 exhausts the cases having $d = 1$, and so it is sufficient to show three results: a core facet having $d = 2$ that is not of the form (35) must be satisfied exactly at $(x, y) = (x^{r+1}, m(x^{r+1}))$, there exists no valid inequality having $d \geqslant 3$ and d odd, and a core facet having $d \geqslant 4$ and d even must be of the form (35).

- Suppose that a core facet $\beta_0 + \sum_{j=1}^{n} \beta_j x_j + \beta' y \geqslant 0$ has $d = 2$ and is not of the form (35). We have $\beta_1 = \ldots = \beta_{r+1}$ and $\beta_{r+2} = \ldots = \beta_n$ by Proposition 3.7, so that we cannot have $\beta_{r+1} = \beta_{r+2} = 0$ since the facet would be of the form (35). As a result, $D(r, r+2) = 0$ with $m(x^r) = m(x^{r+2})$ gives $0 < -\beta_{r+1} = \beta_{r+2} = \bar{\beta}$ for some $\bar{\beta} > 0$. Then $D(r+1, r+2) \leqslant 0$ gives $\bar{\beta} \leqslant \beta' m(x^{r+1})$ because $m(x^{r+2}) = -m(x^{r+1})$ and, since $0 < \bar{\beta}$ and $m(x^{r+1}) = (-1)^{n-(r+1)}$, we have $0 < \bar{\beta} \leqslant \beta' m(x^{r+1}) = \beta'(-1)^{n-(r+1)}$ so that $\beta' = (-1)^{n-(r+1)}$. Consequently, the core facet takes the form

$$\beta_0 - \bar{\beta} \left(\sum_{j=1}^{r+1} x_j - \sum_{j=r+2}^{n} x_j \right) + (-1)^{n-(r+1)} y \geqslant 0. \tag{42}$$

  We now use Proposition 3.6 to show that the core facet must be satisfied exactly at $(x, y) = (x^{r+1}, m(x^{r+1}))$ by setting $\beta_0 = (n-1)$ and $\bar{\beta} = 1$ within (42). The resulting inequality is valid for $\mathsf{conv}\, G$ and is satisfied exactly at $(x, y) = (x^k, m(x^k))$ for $k \in \{r, r+1, r+2\}$. The validity follows from Lemma 2.3 since, for any $r \in \{0, \ldots, n-2\}$, the left side is $2|k - (r+1)| - 1 + (-1)^{k-(r+1)}$ at $(x, y) = \left(x^k, m(x^k) \right)$ for each $k \in \{0, \ldots, n\}$ because $m(x^k) = (-1)^{n-k}$.

- By contradiction, suppose there exists a valid inequality $\beta_0 + \sum_{j=1}^{n} \beta_j x_j + \beta' y \geqslant 0$ having $d \geqslant 3$ and d odd. Then $D(r, s) = 0$ gives $\sum_{j=r+1}^{s} \beta_j = \beta' m(x^r)$ since $m(x^r) = -m(x^s)$. If $\beta' m(x^r) = 1$ so that $\beta' \left(m(x^{r+1}) - m(x^r) \right) = -2$, then $D(r, r+1) \geqslant 0$ gives $\beta_{r+1} \geqslant 1$, and nondecreasing β_j gives $\beta_{r+1} \leqslant \frac{1}{d}$. Thus, a contradiction. Similarly, if $\beta' m(x^r) = -1$ so that $\beta' \left(m(x^s) - m(x^{s-1}) \right) = 2$, then $D(s-1, s) \leqslant 0$ gives $\beta_s \leqslant -1$, and nondecreasing β_j gives $\beta_s \geqslant \frac{-1}{d}$. Again a contradiction.

- Suppose that a core facet $\beta_0 + \sum_{j=1}^{n} \beta_j x_j + \beta' y \geqslant 0$ has $d \geqslant 4$ and d even. The value $D(r, s) = 0$ gives $\sum_{j=r+1}^{s} \beta_j = 0$ since $m(x^r) = m(x^s)$. But $D(r, r+2) \geqslant 0$ gives $(\beta_{r+1} + \beta_{r+2}) \geqslant 0$ since $m(x^r) = m(x^{r+2})$, and the nondecreasing β_j gives $\beta_j = 0$ for $j \in \{r+1, \ldots, s\}$. Then $\beta_j = 0$ for $j \in N$ by Proposition 3.7, and the facet is of the form (35). □

A key component of the proof of Lemma 6.5 is the family of $(n-1)$ valid core inequalities (42) with $\beta_0 = (n-1)$ and $\bar{\beta} = 1$ that has the property that, for each $r \in \{0, \ldots, n-2\}$, the corresponding inequality is satisfied exactly at $(x, y) = (x^k, m(x^k))$ for $k \in \{r, r+1, r+2\}$. It turns out that each such inequality is a core facet, and that there exist no other core facets that are satisfied exactly at $(x^k, m(x^k))$ for three consecutive extreme points x^k of the simplex $\mathcal{S}$.

This result is established below. Here, we find it convenient to substitute the index t in (38) for r in (42).

Lemma 6.6. *There exist precisely $(n-1)$ core facets that are satisfied exactly at $(x^k, m(x^k))$ for three consecutive extreme points x^k of the simplex $\mathcal{S}$, and these inequalities are (38).*

Proof. The proof consists of two parts: the first part shows that inequalities (38) are the only candidate core facets that are satisfied exactly at three consecutive extreme points of the simplex $\mathcal{S}$, and the second part shows that each such inequality is a core facet.

Suppose, for some $t \in \{0, \ldots, n-2\}$, that $\beta_0 + \sum_{j=1}^{n} \beta_j x_j + \beta' y \geqslant 0$ is a core facet that is satisfied exactly at $(x, y) = (x^k, m(x^k))$ for $k \in \{t, t+1, t+2\}$. Then $d = 2$ in the proof of Lemma 6.5 since $d \geqslant 3$, d odd, was shown not possible, and $d \geqslant 4$, d even, was shown to yield a core facet of the form (35). Clearly, the facet is not of the form (35) since $D(t, t+1) = 0$ gives $\beta_{t+1} = \beta' m(x^t) \neq 0$. As a result, the first part of the proof of Lemma 6.5 gives us that the facet must be of the form (42). The two equations in two unknowns β_0 and $\bar{\beta}$ that are obtained by evaluating $(x, y) = (x^k, m(x^k))$ for $k \in \{t, t+1\}$ within (42) give $\beta_0 = (n-1)$ and $\bar{\beta} = 1$ as in (38), since $-m(x^t) = m(x^{t+1}) = (-1)^{n-(t+1)}$.

The proof of Lemma 6.5 showed inequalities (38) to be valid for $\mathsf{conv}\, G$ and, for each $t \in \{0, \ldots, n-2\}$, the associated inequality to be satisfied exactly at $(x, y) = (x^k, m(x^k))$ for $k \in \{t, t+1, t+2\}$. Thus, given such a t, it is sufficient to identify $(n+1)$ affinely independent points in $\mathsf{conv}\, G$ that satisfy the inequality exactly. Consider the n extreme points of $\mathsf{conv}\, G$, denoted by $(p^1, m(p^1)), \ldots, (p^n, m(p^n))$, so that p^i differs from x^{t+1} in only position i. Since $\beta_1 = \ldots = \beta_{t+1}$ and $(x, y) = (x^t, m(x^t))$ satisfies the inequality exactly, Lemma 3.1 gives us that $(x, y) = (p^i, m(p^i))$ for $i \in \{1, \ldots, t+1\}$ satisfies the inequality exactly. Similarly, since $\beta_{t+2} = \ldots = \beta_n$ and $(x, y) = (x^{t+1}, m(x^{t+1}))$ satisfies the inequality exactly, Lemma 3.1 gives us that $(x, y) = (p^i, m(p^i))$ for $i \in \{t+2, \ldots, n\}$ satisfies the inequality exactly. Subtract x^{t+1} from every such point to reduce p^i to $-2e^i$ for $i \leqslant t+1$ and to $2e^i$ for $i \geqslant t+2$, where e^i is the unit vector in $\mathbb{R}^n$ having a 1 in position i and 0 elsewhere. Hence, x^{t+1}, together with $p^1, \ldots, p^n$, is an affinely independent set of points, so that $(x^{t+1}, m(x^{t+1}))$, together with $(p^1, m(p^1)), \ldots, (p^n, m(p^n))$, is an affinely independent set of points. $\square$

***Proof of Theorem* 6.1.** Follows from Lemmas 6.4, 6.5, and 6.6. $\square$

7 Summary and Open Questions

This paper derives polyhedral results for the convex hull of symmetric multilinear polynomials (SMPs) taken over a box domain. Exponential-sized extended formulations of general multilinear polynomials are available via the reformulation-linearization-technique (RLT), but symmetry and disjunctive programming enable a quadratic-sized extended formulation. The goal of this paper is to obtain the convex hulls in the original variable spaces. Instead of adopting the tedious method of projecting the extended formulations, our approach is more elegant in the sense that we directly exploit the problem structure to define special core facets by which all facets can be characterized, and to then devise necessary and/or sufficient conditions on the coefficients of these facets. Whereas much of the theory is applicable to general SMPs over box constraints, we focus attention on two special problem classes: general supermodular (submodular) functions, and monomials having the variable lower bound equal to the negative of the upper bound. For each class, we use the necessary conditions to motivate families of core facets, and then prove that no other such facets can exist. Our derivations of these convex hulls provides alternate

proofs to those in literature. For both classes, we use RLT results to characterize for each facet the set of all points at which the inequality is satisfied exactly.

A direction of future research is the identification of convex hull forms for more general families of SMPs than the two types within this paper, and likewise for the identification of all points within the resulting graphs that satisfy each facet exactly. One open question in this regard is to generalize Theorem 6.1 to monomials taken over $[-1, r]^n$ for arbitrary $r > 0$. Even more broadly, we do not know an explicit minimal description for convex hulls of general SMPs in the original variable space. Projecting the extended formulation of Proposition 2.1 is an option, but this is likely to result in a combinatorial explosion as is usually the case when projecting from a higher-dimensional space. Another question that is open is whether similar to the two families of SMPs analysed in this paper, every SMP has a linear (or even polynomial) number of core facets. A positive answer to this question would imply a straightforward separation algorithm for the convex hull due to Proposition 2.5 without invoking the ellipsoid method and the optimization algorithm in Proposition 2.4.

Bibliography

[AGX19]   W. Adams, A. Gupte, and Y. Xu. "Error bounds for monomial convexification in polynomial optimization". In: *Mathematical Programming* 175 (2019), pp. 355–393.

[Ada11]   W. P. Adams. "Use of Lagrange Interpolating Polynomials in the RLT". In: *Wiley Encyclopedia of Operations Research and Management Science*, ed. by J. J. Cochran et al. John Wiley & Sons, Inc., 2011, pp. 5684–5694.

[AS05]   W. P. Adams and H. D. Sherali. "A hierarchy of relaxations leading to the convex hull representation for general discrete optimization problems". In: *Annals of Operations Research* 140.1 (2005), pp. 21–47.

[Ant+16]   M. Anthony, E. Boros, Y. Crama, and A. Gruber. "Quadratization of symmetric pseudo-Boolean functions". In: *Discrete Applied Mathematics* 203 (2016), pp. 1–12.

[AN08]   A. Atamtürk and V. Narayanan. "Polymatroids and mean-risk minimization in discrete optimization". In: *Operations Research Letters* 36.5 (2008), pp. 618–622.

[Bal79]   E. Balas. "Disjunctive Programming". In: *Discrete Optimization II*, ed. by P. Hammer et al. Vol. 5. Annals of Discrete Mathematics. Elsevier, 1979, pp. 3–51.

[BM14]   M. Ballerstein and D. Michaels. "Extended formulations for convex envelopes". In: *Journal of Global Optimization* 60.2 (2014), pp. 217–238.

[Bao+15]   X. Bao, A. Khajavirad, N. V. Sahinidis, and M. Tawarmalani. "Global optimization of nonconvex problems with multilinear intermediates". In: *Mathematical Programming Computation* 7.1 (2015), pp. 1–37.

[BST09]   X. Bao, N. V. Sahinidis, and M. Tawarmalani. "Multiterm polyhedral relaxations for nonconvex, quadratically constrained quadratic programs". In: *Optimization Methods and Software* 24.4-5 (2009), pp. 485–504.

[Ben04]   H. P. Benson. "Concave envelopes of monomial functions over rectangles". In: *Naval Research Logistics (NRL)* 51.4 (2004), pp. 467–476.

[BCR20]   E. Boros, Y. Crama, and E. Rodríguez-Heck. "Compact quadratizations for pseudo-Boolean functions". In: *Journal of Combinatorial Optimization* 39.3 (2020), pp. 687–707.

[BH02]   E. Boros and P. L. Hammer. "Pseudo-boolean optimization". In: *Discrete Applied Mathematics* 123.1-3 (2002), pp. 155–225.

[Bow72]   V. J. Bowman. "Permutation polyhedra". In: *SIAM Journal on Applied Mathematics* 22.4 (1972), pp. 580–589.

[BCR19]   C. Buchheim, Y. Crama, and E. Rodríguez-Heck. "Berge-acyclic multilinear 0–1 optimization problems". In: *European Journal of Operational Research* 273.1 (2019), pp. 102–107.

[BR08]   C. Buchheim and G. Rinaldi. "Efficient reduction of polynomial zero-one optimization to the quadratic case". In: *SIAM Journal on Optimization* 18.4 (2008), pp. 1398–1413.

[CLL10]   S. Cafieri, J. Lee, and L. Liberti. "On convex relaxations of quadrilinear terms". In: *Journal of Global Optimization* 47.4 (2010), pp. 661–685.

[CDG20]   R. Chen, S. Dash, and O. Günlük. "Cardinality Constrained Multilinear Sets". In: *Combinatorial Optimization: ISCO 2020*, ed. by M. Baïou et al. Vol. 12176. Lecture Notes in Computer Science. Springer, Cham, 2020, pp. 54–65.

[Cra93]   Y. Crama. "Concave extensions for nonlinear 0–1 maximization problems". In: *Mathematical Programming* 61.1-3 (1993), pp. 53–60.

[CR17]   Y. Crama and E. Rodríguez-Heck. "A class of valid inequalities for multilinear 0–1 optimization problems". In: *Discrete Optimization* 25 (2017), pp. 28–47.

[DK16]   A. Del Pia and A. Khajavirad. "A polyhedral study of binary polynomial programs". In: *Mathematics of Operations Research* (2016).

[DK18a]   A. Del Pia and A. Khajavirad. "The multilinear polytope for acyclic hypergraphs". In: *SIAM Journal on Optimization* 28.2 (2018), pp. 1049–1076.

[DK18b]   A. Del Pia and A. Khajavirad. *The running intersection relaxation of the multilinear polytope*. Preprint. 2018. Optimization-Online: http://www.optimization-online.org/DB_HTML/2018/05/6618.html.

[DKS20]   A. Del Pia, A. Khajavirad, and N. V. Sahinidis. "On the impact of running intersection inequalities for globally solving polynomial optimization problems". In: *Mathematical Programming Computation* 12.2 (2020), pp. 165–191.

[DMW20]   S. S. Dey, M. Molinaro, and G. Wang. *Solving row-sparse principal component analysis via convex integer programs*. Preprint. 2020. arXiv: 2010.11152 [math.OC].

[Eag+20] N. M. Eagles, A. M. Foley, A. Huang, E. Karangozishvili, and A. Yu. *H-chromatic symmetric functions*. Preprint. 2020. arXiv: 2011.06063 [math.CO].

[Edm70]   J. Edmonds. "Submodular functions, matroids, and certain polyhedra". In: *Combinatorial structures and their applications*, ed. by R. Guy et al. Proceedings 1969 Calgary Conference. Gordon and Breach, New York, 1970, pp. 69–87.

[FFM18]   A. Fischer, F. Fischer, and S. T. McCormick. "Matroid optimisation problems with nested non-linear monomials in the objective function". In: *Mathematical Programming* 169.2 (2018), pp. 417–446.

[FFM20]   A. Fischer, F. Fischer, and S. T. McCormick. "Matroid optimization problems with monotone monomials in the objective". In: *Discrete Applied Mathematics* In Press (2020).

[FKL15]   F. D. Fomeni, K. Kaparis, and A. N. Letchford. "Cutting planes for RLT relaxations of mixed 0–1 polynomial programs". In: *Mathematical Programming* 151.2 (2015), pp. 639–658.

[GW74]   F. Glover and E. Woolsey. "Converting the 0-1 Polynomial Programming Problem to a 0-1 Linear Program". In: *Operations Research* 22.1 (1974), pp. 180–182.

[Gup+20] A. Gupte, T. Kalinowski, F. Rigterink, and H. Waterer. "Extended formulations for convex hulls of some bilinear functions". In: *Discrete Optimization* 36 (2020), p. 100569.

[Hen11]   S. M. Henry. "Tight Polyhedral Representations of Discrete Sets using Projections, Simplices, and Base-2 Expansions". PhD thesis. Clemson, SC, USA: Department of Mathematical Sciences, Clemson University, 2011.

[Iwa08]   S. Iwata. "Submodular function minimization". In: *Mathematical Programming* 112.1 (2008), p. 45.

[KTR19]   J. Kim, M. Tawarmalani, and J.-P. P. Richard. *Convexification of permutation-invariant sets and applications*. Preprint. 2019. arXiv: 1910.02573 [math.OC].

[Kuh60]   H. W. Kuhn. "Some combinatorial Lemmas in topology". In: *IBM Journal of Research and Development* 4.5 (1960), pp. 518–524.

[Lov83]   L. Lovász. "Submodular functions and convexity". In: *Mathematical Programming: The State of the Art – Bonn 1982*, ed. by A. Bachem et al. XI$^{\text{th}}$ International Symposium on Mathematical Programming. Springer Berlin Heidelberg, 1983, pp. 235–257.

[LNL12]   J. Luedtke, M. Namazifar, and J. Linderoth. "Some results on the strength of relaxations of multilinear functions". In: *Mathematical Programming* 136.2 (2012), pp. 325–351.

[Mac98]   I. G. Macdonald. *Symmetric functions and Hall polynomials*. Second Edition. Oxford Classic Texts in the Physical Sciences. Oxford University Press, 1998.

[MF04]   C. Meyer and C. Floudas. "Trilinear monomials with mixed sign domains: Facets of the convex and concave envelopes". In: *Journal of Global Optimization* 29.2 (2004), pp. 125–155.

[Rad52]   R. Rado. "An inequality". In: *Journal of London Mathematical Society* 27 (1952), pp. 1–6.

[Rik97]   A. Rikun. "A convex envelope formula for multilinear functions". In: *Journal of Global Optimization* 10.4 (1997), pp. 425–437.

[RS01]   H. S. Ryoo and N. V. Sahinidis. "Analysis of bounds for multilinear functions". In: *Journal of Global Optimization* 19.4 (2001), pp. 403–424.

[She97]   H. D. Sherali. "Convex envelopes of multilinear functions over a unit hypercube and over special discrete sets". In: *Acta Mathematica Vietnamica* 22.1 (1997), pp. 245–270.

[SA90]   H. D. Sherali and W. P. Adams. "A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems". In: *SIAM Journal on Discrete Mathematics* 3.3 (1990), pp. 411–430.

[SA94]   H. D. Sherali and W. P. Adams. "A hierarchy of relaxations and convex hull characterizations for mixed-integer zero-one programming problems". In: *Discrete Applied Mathematics* 52.1 (1994), pp. 83–106.

[SA99]   H. D. Sherali and W. P. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Vol. 31. Nonconvex Optimization and its Applications. Kluwer Academic Publishers, 1999.

[Sta95]   R. P. Stanley. "A symmetric function generalization of the chromatic polynomial of a graph". In: *Advances in Mathematics* 111.1 (1995), pp. 166–194.

[TRX13]   M. Tawarmalani, J.-P. P. Richard, and C. Xiong. "Explicit convex and concave envelopes through polyhedral subdivisions". In: *Mathematical Programming* 138.1-2 (2013), pp. 531–577.

[Von10]   J. Vondrák. *Continuous extensions of submodular functions*. Polyhedral techniques in combinatorial optimization. Lecture Notes CS 369P. https://theory.stanford.edu/~jvondrak/CS369P/CS369P.html. Stanford University, 2010.

Appendix A   Properties of $\text{conv}\,G$ from RLT

The key observation is that the equivalence of (20) between the polyhedral sets P of (19) and $\text{conv}\,T$ of (20) continues to hold true with the inclusion of the restriction $y = m(x)$ within these two sets. Consider a generalization of the set G given by

$$G' \equiv \left\{(x,y) \in \mathbb{R}^n \times \mathbb{R} : x \in X',\ y = m(x)\right\}, \tag{A.1}$$

that is obtained by replacing $x \in X$ of (1c) with $x \in X'$ of (15). Let

$$P_y \equiv \left\{(x,w,y) \in \mathbb{R}^n \times \mathbb{R}^{2^n - (n+1)} \times \mathbb{R} : (x,w) \in P,\ y = \left\{m(x)\right\}_L\right\} \tag{A.2}$$

be the set P of (19) that is modified to include the additional variable y and additional restriction $y = \left\{m(x)\right\}_L$, and let

$$T_y \equiv \left\{(x,w,y) \in \mathbb{R}^n \times \mathbb{R}^{2^n - (n+1)} \times \mathbb{R} : (x,w) \in T,\ y = m(x)\right\}$$

be the set T of (20) that is modified to include the additional variable y and additional restriction $y = m(x)$. Then we have that

$$P_y = conv\left(T_y\right). \tag{A.3}$$

Equality (A.3) holds true from (20) because the extreme points of P and P_y correspond in a one-to-one fashion in such a manner that (x,w) is an extreme point of P if and only if (x,w,y) is an extreme point of P_y having $y = \left\{m(x)\right\}_L = m(x)$. Therefore, every extreme point of the polytope P_y is in T_y and $T_y \subseteq P_y$ by construction.

The three propositions below are consequences of (A.3), where $\text{Proj}_{(x,y)}(\bullet)$ denotes the projection of the set $\bullet$ onto the space of the variables (x,y).

Proposition A.1. $\text{conv}\,G' = \text{conv}\left(Proj_{(x,y)}(T_y)\right) = Proj_{(x,y)}\left(\text{conv}\,T_y\right) = Proj_{(x,y)}\left(P_y\right).$

Proof. The first equality follows because $G' = \text{Proj}_{(x,y)}(T_y)$, the second equality follows from interchanging the projection and convex hull operators, and the third equality follows from (A.3). $\qquad\square$

Proposition A.2. *conv* $\left(G'\right)$ *is a polytope with 2^n extreme points that have one-to-one correspondence with the extreme points of X' in such a manner that $y = m(x)$ at each such point x.*

Proof. conv $\left(G'\right)$ is a polytope with no more than 2^n extreme points since it is the projection of the polytope P_y having 2^n extreme points onto the (x,y) space, as stated in Proposition A.1. However, each of the 2^n points of *conv* $\left(G'\right)$ satisfying the property that every entry of x has either $x_j = L_j$ or $x_j = U_j$ for all $j \in N$, and that $y = m(x)$, is trivially an extreme point of *conv* $\left(G'\right)$. $\qquad\square$

Proposition A.2 also follows from results in Rikun [Rik97]. As opposed to using the projection from a higher-dimensional RLT space as in the above proof, Rikun shows that every $(x,y) \in G'$ which has some $x_j \in (L_j, U_j)$ can be expressed as a strict convex combination of two distinct points in G'.

The following result addresses the validity of a linear inequality for *conv* $\left(G'\right)$ in terms of the restrictions of the set P_y.

Corollary A.3. *A linear inequality $\beta_0 + \sum_{j=1}^n \beta_j x_j + \beta' y \geqslant 0$ is valid for $\mathrm{conv}\,(G')$ if and only if it can be uniquely expressed as a linear combination of the restrictions of P_y using nonnegative multipliers $\pi \in \mathbb{R}^{2^n}$ and the scalar β', so that*

$$\beta_0 + \sum_{j=1}^n \beta_j x_j + \beta' y = \sum_{K \subseteq N} \pi_K F_K(x) + \beta' \left(y - \{m(x)\}_L \right). \tag{A.4}$$

Proof. The existence of nonnegative multipliers π satisfying (A.4) follows from the result of Proposition A.1 that $\mathrm{conv}\,G' \mathrm{Proj}_{(x,y)}\left(P_y\right)$, as then a linear inequality will be valid for $\mathrm{conv}\,G'$ if and only if it is valid for P_y. The uniqueness follows from the invertibility of the matrix $\begin{bmatrix} U_1 & -1 \\ -L_1 & 1 \end{bmatrix} \otimes \dots \otimes \begin{bmatrix} U_n & -1 \\ -L_n & 1 \end{bmatrix}$ of (25). $\qquad\square$

Appendix B Illustrative Examples

B.1 Use of Kronecker Products

The example below illustrates the use of Kronecker products to establish the two stated properties of P, to demonstrate (20), and to provide insights into the two observations.

Example 1. Consider X' of (15) with $N = \{1, 2, 3\}$ in the $n = 3$ variables x_1, x_2, and x_3. The $2^3 = 8$ inequalities of (18) are expressed in (24) as

$$\begin{bmatrix} U_1U_2U_3 & -U_1U_2 & -U_1U_3 & U_1 & -U_2U_3 & U_2 & U_3 & -1 \\ -U_1U_2L_3 & U_1U_2 & U_1L_3 & -U_1 & U_2L_3 & -U_2 & -L_3 & 1 \\ -U_1L_2U_3 & U_1L_2 & U_1U_3 & -U_1 & L_2U_3 & -L_2 & -U_3 & 1 \\ U_1L_2L_3 & -U_1L_2 & -U_1L_3 & U_1 & -L_2L_3 & L_2 & L_3 & -1 \\ -L_1U_2U_3 & L_1U_2 & L_1U_3 & -L_1 & U_2U_3 & -U_2 & -U_3 & 1 \\ L_1U_2L_3 & -L_1U_2 & -L_1L_3 & L_1 & -U_2L_3 & U_2 & L_3 & -1 \\ L_1L_2U_3 & -L_1L_2 & -L_1U_3 & L_1 & -L_2U_3 & L_2 & U_3 & -1 \\ -L_1L_2L_3 & L_1L_2 & L_1L_3 & -L_1 & L_2L_3 & -L_2 & -L_3 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ x_3 \\ x_2 \\ x_2x_3 \\ x_1 \\ x_1x_3 \\ x_1x_2 \\ x_1x_2x_3 \end{pmatrix} \geqslant \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{B.5}$$

The equations of (27) in nonnegative variables $\boldsymbol{\lambda}$ take the form

$$\begin{pmatrix} 1 \\ x_3 \\ x_2 \\ w_{23} \\ x_1 \\ w_{13} \\ w_{12} \\ w_{123} \end{pmatrix} = \frac{1}{d_1d_2d_3} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ L_3 & U_3 & L_3 & U_3 & L_3 & U_3 & L_3 & U_3 \\ L_2 & L_2 & U_2 & U_2 & L_2 & L_2 & U_2 & U_2 \\ L_2L_3 & L_2U_3 & U_2L_3 & U_2U_3 & L_2L_3 & L_2U_3 & U_2L_3 & U_2U_3 \\ L_1 & L_1 & L_1 & L_1 & U_1 & U_1 & U_1 & U_1 \\ L_1L_3 & L_1U_3 & L_1L_3 & L_1U_3 & U_1L_3 & U_1U_3 & U_1L_3 & U_1U_3 \\ L_1L_2 & L_1L_2 & L_1U_2 & L_1U_2 & U_1L_2 & U_1L_2 & U_1U_2 & U_1U_2 \\ L_1L_2L_3 & L_1L_2U_3 & L_1U_2L_3 & L_1U_2U_3 & U_1L_2L_3 & U_1L_2U_3 & U_1U_2L_3 & U_1U_2U_3 \end{bmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \lambda_6 \\ \lambda_7 \\ \lambda_8 \end{pmatrix} \tag{B.6}$$

where the RLT linearization step sets $w_{12} = x_1x_2$, $w_{13} = x_1x_3$, $w_{23} = x_2x_3$, and $w_{123} = x_1x_2x_3$.

There are eight extreme points to this system, with extreme point j having $\lambda_j = d_1d_2d_3$ and $\lambda_i = 0$ for $i \neq j$. Then the eight extreme points to P of (19) and (25) are given by the eight columns of the above 8×8 matrix, less the first row. Consequently, we have that $P = conv(T)$ with

$$T = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ w_{12} \\ w_{13} \\ w_{23} \\ w_{123} \end{pmatrix} : L_j \leqslant x_j \leqslant U_j \; \forall \, j = 1, 2, 3, \; \begin{pmatrix} x_1x_2 \\ x_1x_3 \\ x_2x_3 \\ x_1x_2x_3 \end{pmatrix} = \begin{pmatrix} w_{12} \\ w_{13} \\ w_{23} \\ w_{123} \end{pmatrix} \right\},$$

as in (20).

Now consider any multilinear polynomial $\sum_{J \subseteq N} \alpha_J \prod_{j \in J} x_j$ as found in Corollarys 4.2 and 4.3, and the corresponding system in variables $\boldsymbol{\pi} \in \mathbb{R}^{2^n}$. Using obvious notation, denote (B.5) by $A\boldsymbol{v} \geqslant \boldsymbol{0}$ so that the (scaled) 8×8 matrix of (B.6) is A^{-1}. Observe that

the eight inequalities of (B.5) correspond, in order, to the functions $F(K)$ of (16) having $K = \emptyset, \{3\}, \{2\}, \{2,3\}, \{1\}, \{1,3\}, \{1,2\}, \{1,2,3\}$. (This order coincides with the variable indices of the products within the vector $\boldsymbol{v}$.) Accordingly define $\boldsymbol{\alpha}^T = (\alpha_\emptyset, \alpha_3, \alpha_2, \alpha_{23}, \alpha_1, \alpha_{13}, \alpha_{12}, \alpha_{123})$, and $\boldsymbol{\pi} \in \mathbb{R}^8$ by $\boldsymbol{\pi}^T = (\pi_\emptyset, \pi_3, \pi_2, \pi_{23}, \pi_1, \pi_{13}, \pi_{12}, \pi_{123})$ so that $\boldsymbol{\pi}^T = \boldsymbol{\alpha}^T A^{-1}$. Then we have

$$\sum_{J \subseteq N} \alpha_J \prod_{j \in J} x_j = \boldsymbol{\alpha}^T \boldsymbol{v} = (\boldsymbol{\alpha}^T A^{-1})(A\boldsymbol{v}) = \boldsymbol{\pi}^T(A\boldsymbol{v}) = \sum_{K \subseteq N} \pi_K F(K), \tag{B.7}$$

where the four equalities follow, from left to right, from the definitions of $\boldsymbol{\alpha}$ and $\boldsymbol{v}$, the multiplicative inverse of A, the definition of $\boldsymbol{\pi}$, and the stated equivalence between the vector $A\boldsymbol{v}$ and the functions $F(K)$ of (16). The computing of the multipliers π_K by Corollary 4.3 follows from (B.7) since each entry of the vector $\boldsymbol{\alpha}^T A^{-1}$ corresponds to a distinct $K \subseteq N$, and realizes value $\frac{1}{D_3}(\sum_{J \subseteq N} \alpha_J \prod_{j \in J} \hat{x}_j)$, where $\hat{x}_j = U_j$ for all $j \in K$ and $\hat{x}_j = L_j$ for all $j \notin K$. For each such K, this value is the multiplier π_K on the associated function $F(K)$. Corollary 4.3 gives us, provided it is nonnegative over the extreme points of X', that the polynomial $\sum_{J \subseteq N} \alpha_J \prod_{j \in J} x_j$ vanishes at a point $\hat{x} \in X'$ if and only if $\pi_K F(K) = 0$ for all $K \subseteq N$, where each $F(K)$ is evaluated at $\hat{x}$. $\diamond$

B.2 Convex Hull via Projection

The example below illustrates the computation of the set $conv\,(G')$ via the projection from the extended variable space $(x, \boldsymbol{w}, y)$ of P_y onto the space of the variables (x, y), as stated in Proposition A.1. For simplicity, the sets X' and G' of (15) and (A.1) are reduced to the sets X and G, respectively, by setting $L_j = \ell$ and $U_j = u$ for all $j \in N$. The example also demonstrates the result of Corollary 4.3 that allows for the identification of those points within X at which each facet is satisfied exactly. Example 2 builds upon Example 1, and will be later referenced.

Example 2. Consider the set G with $N = \{1, 2, 3\}$ in the $n = 3$ variables x_1, x_2, x_3, and the variable y, with $y = m(x) = x_1 x_2 x_3,$. The set P_y of (A.2), whose projection onto the (x, y) variable space gives $conv\,G$ as stated in Proposition A.1, is expressed in matrix form below. The matrix partitioning is used to emphasize that the first eight restrictions are inequalities and the last restriction is equality.

$$\begin{bmatrix} u^3 & -u^2 & -u^2 & u & -u^2 & u & u & -1 & 0 \\ -\ell u^2 & u^2 & \ell u & -u & \ell u & -u & -\ell & 1 & 0 \\ -\ell u^2 & \ell u & u^2 & -u & \ell u & -\ell & -u & 1 & 0 \\ \ell^2 u & -\ell u & -\ell u & u & -\ell^2 & \ell & \ell & -1 & 0 \\ -\ell u^2 & \ell u & \ell u & -\ell & u^2 & -u & -u & 1 & 0 \\ \ell^2 u & -\ell u & -\ell^2 & \ell & -\ell u & u & \ell & -1 & 0 \\ \ell^2 u & -\ell^2 & -\ell u & \ell & -\ell u & \ell & u & -1 & 0 \\ -\ell^3 & \ell^2 & \ell^2 & -\ell & \ell^2 & -\ell & -\ell & 1 & 0 \end{bmatrix} \begin{pmatrix} 1 \\ x_3 \\ x_2 \\ w_{23} \\ x_1 \\ w_{13} \\ w_{12} \\ w_{123} \end{pmatrix} \geqslant \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{pmatrix} y \end{pmatrix} = \begin{pmatrix} 0 \end{pmatrix} \tag{B.8}$$

The eight inequalities are the linearized form of (B.5) when L_1, L_2, and L_3 are set to ℓ, and when U_1, U_2, and U_3 are set to u, while the equation is $-\{m(x)\}_L + y = 0$. Since we desire to project (B.8) onto the space of the variables (x_1, x_2, x_3, y), the projection cone takes the form

$$\begin{bmatrix} u & -u & -u & u & -\ell & \ell & \ell & -\ell & 0 \\ u & -u & -\ell & \ell & -u & u & \ell & -\ell & 0 \\ u & -\ell & -u & \ell & -u & \ell & u & -\ell & 0 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 \end{bmatrix} \begin{pmatrix} \pi_\emptyset \\ \pi_3 \\ \pi_2 \\ \pi_{23} \\ \pi_1 \\ \pi_{13} \\ \pi_{12} \\ \pi_{123} \\ \beta' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{B.9}$$

Table 1: Facet and Points in G where Satisfied Exactly when $\ell > 0$

Facet	Points in G where Satisfied Exactly		
$(\ell^2)x_1 + (\ell u)x_2 + (u^2)x_3 - y \geqslant \ell u(\ell + u)$	(x_1,ℓ,ℓ)	(u,x_2,ℓ)	(u,u,x_3)
$(\ell^2)x_1 + (u^2)x_2 + (\ell u)x_3 - y \geqslant \ell u(\ell + u)$	(x_1,ℓ,ℓ)	(u,x_2,u)	(u,ℓ,x_3)
$(\ell u)x_1 + (\ell^2)x_2 + (u^2)x_3 - y \geqslant \ell u(\ell + u)$	(x_1,u,ℓ)	(ℓ,x_2,ℓ)	(u,u,x_3)
$(u^2)x_1 + (\ell^2)x_2 + (\ell u)x_3 - y \geqslant \ell u(\ell + u)$	(x_1,u,u)	(ℓ,x_2,ℓ)	(ℓ,u,x_3)
$(\ell u)x_1 + (u^2)x_2 + (\ell^2)x_3 - y \geqslant \ell u(\ell + u)$	(x_1,ℓ,u)	(u,x_2,u)	(ℓ,ℓ,x_3)
$(u^2)x_1 + (\ell u)x_2 + (\ell^2)x_3 - y \geqslant \ell u(\ell + u)$	(x_1,u,u)	(ℓ,x_2,u)	(ℓ,ℓ,x_3)
$-(\ell^2)x_1 - (\ell^2)x_2 - (\ell^2)x_3 + y \geqslant -2\ell^3$	(x_1,ℓ,ℓ)	(ℓ,x_2,ℓ)	(ℓ,ℓ,x_3)
$-(\ell u)x_1 - (\ell u)x_2 - (\ell u)x_3 + y \geqslant -\ell u(\ell + u)$	Any $x_i = \ell$, any $x_j = u$, $i \neq j$.		
$-(u^2)x_1 - (u^2)x_2 - (u^2)x_3 + y \geqslant -2u^3$	(x_1,u,u)	(u,x_2,u)	(u,u,x_3)

where $\boldsymbol{\pi}$ are the nonnegative multipliers on the eight inequality restrictions of (B.8), and where β' is the multiplier on the equation. When $\ell > 0$, there are fifteen extreme directions to this cone, with six "trivial directions" resulting in the six facets $x_j \geqslant \ell$ and $-x_j \geqslant -u$ for $j \in \{1,2,3\}$. Specifically, setting $\beta' = 0$ for each j, the "trivial direction" having $\pi_K = 1$ if $j \in K$ and $\pi_K = 0$ otherwise gives $x_j - \ell \geqslant 0$, and the "trivial direction" having $\pi_K = 1$ if $j \in (N - K)$ and $\pi_K = 0$ otherwise gives $u - x_j \geqslant 0$, all inequalities scaled by $\frac{1}{(u-\ell)^2}$. (Clearly, for each $j \in \{1,2,3\}$, the inequality $x_j - \ell \geqslant 0$ is satisfied exactly at all points $(x_1,x_2,x_3,y) \in G$ having $x_j = \ell$, and the inequality $u - x_j \geqslant 0$ is satisfied exactly at all points $(x_1,x_2,x_3,y) \in G$ having $x_j = u$.) Each of the remaining nine directions is depicted as a column of the below matrix.

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ell & \ell+2u \\
\ell+u & \ell & \ell+u & \ell & 0 & 0 & 0 & 0 & u \\
\ell & \ell+u & 0 & 0 & \ell+u & \ell & 0 & 0 & u \\
\ell+u & \ell+u & u & 0 & u & 0 & \ell & 0 & 0 \\
0 & 0 & \ell & \ell+u & \ell & \ell+u & 0 & 0 & u \\
u & 0 & \ell+u & \ell+u & 0 & u & \ell & 0 & 0 \\
0 & u & 0 & u & \ell+u & \ell+u & \ell & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 2\ell+u & u & 0 \\
\ell-u & \ell-u & \ell-u & \ell-u & \ell-u & \ell-u & -\ell+u & -\ell+u & -\ell+u
\end{bmatrix}
$$

The set $\operatorname{conv} G$ is then defined in terms of the $x_j - \ell \geqslant 0$ and $u - x_j \geqslant 0$ restrictions for $j \in \{1,2,3\}$, together with the nine facets that are listed in the first column of the below table, upon dividing each inequality by $(u - \ell)$.

By letting $y = m(x)$ within each facet of Table 1, Corollary 4.3 allows us to identify, in terms of the positive multipliers $\boldsymbol{\pi}$ of (B.9), the points in G where each such inequality is satisfied exactly. Definition (16) gives us, for this example, that a function $F(K)$ vanishes at a point $\hat{x} \in X$ if and only if either $\hat{x}_j = \ell$ for some $j \in K$ or $\hat{x}_j = u$ for some $j \notin K$. The second column of Table 1, which then logically follows, lists the set of all points $(x_1,x_2,x_3,y) \in G$ where each facet is satisfied exactly, with the value $y = x_1x_2x_3$ suppressed for simplicity. The notation "x_j" found within this table indicates that the variable x_j can have $\ell \leqslant x_j \leqslant u$. Observe that, given any realization of (x_1,x_2,x_3) wherein two of the variables have values at either their lower or upper bounds, the first six facets enforce $y \leqslant x_1x_2x_3$ while the last three facets enforce $y \geqslant x_1x_2x_3$, ensuring that $y = x_1x_2x_3$.

For the case when $\ell = 0$, each of the first eight facets of Table 1 is twice listed by repetition, and the inequalities $x_j \geqslant 0$ for $j \in \{1,2,3\}$ are not facets. A statement of the five facets resulting from Table 1 and those points $(x_1,x_2,x_3,y) \in G$, with $y = x_1x_2x_3$ suppressed for simplicity,

Table 2: Facet and Points in G where Satisfied Exactly when $\ell = 0$

Facet	Points in G where Satisfied Exactly		
$(u^2)x_3 - y \geqslant 0$	$(x_1, x_2, 0)$	(u, u, x_3)	
$(u^2)x_2 - y \geqslant 0$	$(x_1, 0, x_3)$	(u, x_2, u)	
$(u^2)x_1 - y \geqslant 0$	$(0, x_2, x_3)$	(x_1, u, u)	
$y \geqslant 0$	Any $x_i = 0$		
$-(u^2)x_1 - (u^2)x_2 - (u^2)x_3 + y \geqslant -2u^3$	(x_1, u, u)	(u, x_2, u)	(u, u, x_3)

where each facet is satisfied exactly is given in Table 2. Here, the notation "x_j" indicates that the variable x_j can have $0 \leqslant x_j \leqslant u$. $\diamond$

B.3 Convex Hull for Supermodular Function

Example 3. Consider G in the $n = 3$ variables x_1, x_2, x_3, and the variable y, with $m(x) = x_1 x_2 x_3$ as in Example 2, and with $\ell \geqslant 0$. Inequalities (30) hold true because $(m(x^k) - m(x^{k-1}))$ takes values $\ell^2(u-\ell)$, $(\ell u)(u-\ell)$, and $u^2(u-\ell)$ when $k = 1$, 2, and 3, respectively. Then Theorem 5.1 is applicable, and core facet (31) is

$$-\ell u(\ell + u) + (\ell^2)x_1 + (\ell u)x_2 + (u^2)x_3 - y \geqslant 0, \tag{B.10}$$

and core facets (32) are

$$-\ell^2 u - \ell^2 \left(x_1 + x_2 + x_3 - [u + 2\ell]\right) + y \quad \geqslant \quad 0, \tag{B.11}$$
$$-\ell u^2 - \ell u \left(x_1 + x_2 + x_3 - [2u + \ell]\right) + y \quad \geqslant \quad 0, \tag{B.12}$$
$$-u^3 - u^2 \left(x_1 + x_2 + x_3 - [3u]\right) + y \quad \geqslant \quad 0. \tag{B.13}$$

For $\ell > 0$, $m(x)$ is strictly supermodular and facet (B.10), together with the additional five facets obtained by permuting the coefficients β, are the first six facets of Table 1. Facets (B.11)–(B.13) are the last three facets of Table 1. In addition, the six inequalities $\ell \leqslant x_j \leqslant u$ for $j \in \{1, 2, 3\}$ comprise the remaining facets of $\mathbf{conv}\,G$. For $\ell = 0$, $m(x)$ is supermodular (but not strictly supermodular because $(m(x^k) - m(x^{k-1}))$ takes values 0, 0, and u^3 when $k = 1$, 2, and 3, respectively), and three of the six facets obtained by permuting β in (B.10) are repetitive. The three resulting facets are the first three inequalities of Table 2. Also, and consistent with Remark 1, facets (B.11) and (B.12) are the same. Facets (B.11) and (B.13) are the last two inequalities of Table 2. In addition for $\ell = 0$, $x_j \geqslant 0$ for $j \in \{1, 2, 3\}$ are not facets.

Suppose that G in the $n = 3$ variables x_1, x_2, x_3, and the variable y, is changed to have $y = m(x) = (2u - \ell)(x_1 x_2 + x_1 x_3 + x_2 x_3) - x_1 x_2 x_3$, with arbitrary ℓ and u. Inequalities (30) hold true because the difference between the two sides takes values $2(u - \ell)^3$ and $(u - \ell)^3$ when $k = 1$ and $k = 2$, respectively, so that Theorem 5.1 is applicable with $m(x)$ strictly supermodular. Core facet (31) is

$$(\ell u)(4\ell - 5u) + (4\ell u - 3\ell^2)x_1 + \left(2u^2 - \ell^2\right)x_2 + \left(3u^2 - 2\ell u\right)x_3 - y \geqslant 0,$$

Table 3: Facet and Points in G where Satisfied Exactly

Facet	Points in G where Satisfied Exactly			
$40 - y \geqslant 0$	$(2,2,2)$	$(2,-2,-2)$	$(-2,2,-2)$	$(-2,-2,2)$
$40 + y \geqslant 0$	$(-2,-2,-2)$	$(-2,2,2)$	$(2,-2,2)$	$(2,2,-2)$
$80 + 20x_1 + 20x_2 + 20x_3 - y \geqslant 0$	$(x_1,-2,-2)$	$(-2,x_2,-2)$	$(-2,-2,x_3)$	
$80 - 20x_1 - 20x_2 - 20x_3 + y \geqslant 0$	$(x_1,2,2)$	$(2,x_2,2)$	$(2,2,x_3)$	
$80 - 20x_1 + 20x_2 + 20x_3 + y \geqslant 0$	$(x_1,-2,-2)$	$(2,x_2,-2)$	$(2,-2,x_3)$	
$80 + 20x_1 - 20x_2 + 20x_3 + y \geqslant 0$	$(x_1,2,-2)$	$(-2,x_2,-2)$	$(-2,2,x_3)$	
$80 + 20x_1 + 20x_2 - 20x_3 + y \geqslant 0$	$(x_1,-2,2)$	$(-2,x_2,2)$	$(-2,-2,x_3)$	
$80 - 20x_1 - 20x_2 + 20x_3 - y \geqslant 0$	$(x_1,2,-2)$	$(2,x_2,-2)$	$(2,2,x_3)$	
$80 - 20x_1 + 20x_2 - 20x_3 - y \geqslant 0$	$(x_1,-2,2)$	$(2,x_2,2)$	$(2,-2,x_3)$	
$80 + 20x_1 - 20x_2 - 20x_3 - y \geqslant 0$	$(x_1,2,2)$	$(-2,x_2,2)$	$(-2,2,x_3)$	

and core facets (32) are

$$
\begin{aligned}
-(4\ell u^2 - \ell^3 - \ell^2 u) - (4\ell u - 3\ell^2)\left(x_1 + x_2 + x_3 - [u + 2\ell]\right) + y &\geqslant 0, \\
-(2\ell u^2 - 2\ell^2 u + 2u^3) - (2u^2 - \ell^2)\left(x_1 + x_2 + x_3 - [2u + \ell]\right) + y &\geqslant 0, \\
-(5u^3 - 3\ell u^2) - (3u^2 - 2\ell u)\left(x_1 + x_2 + x_3 - [3u]\right) + y &\geqslant 0.
\end{aligned}
$$

There are exactly nine facets (2) with $\beta' = \pm 1$ for $\operatorname{conv} G$, including the four listed above and the five additional obtained by permuting the coefficients β in the first inequality. The six inequalities $\ell \leqslant x_j \leqslant u$ for $j \in \{1,2,3\}$ comprise the remaining facets of $\operatorname{conv} G$, as $\ell < u$. $\diamond$

B.4 Convex Hull for Monomial

Example 4. Consider G in the $n = 3$ variables x_1, x_2, x_3, and the variable y, with $m(x) = 5x_1 x_2 x_3$, and with $-\ell = u = 2$. Since $n \geqslant 3$, the convex hull representation is given by the (scaled) six inequalities $-2 \leqslant x_j \leqslant 2$ for $j \in \{1,2,3\}$ found in (40), together with the remaining (scaled) two inequalities of (40), $-40 \leqslant x_4 \leqslant 40$, and the eight inequalities of (41), with the last ten such inequalities summarized in Table 3 below. Here, y is used in lieu of x_4 for clarity. Each inequality $\beta_0 + \sum_{j=1}^n \beta_j x_j + \beta' y \geqslant 0$ has been scaled to $\beta_0 + \sum_{j=1}^n \left(\frac{\beta_j}{u}\right) x_j + \left(\frac{\beta'}{c_n u^n}\right) y \geqslant 0$ as discussed at the beginning of this section to handle the coefficient $c_n = 5$ found in $m(x)$ and the variable bounds $-\ell = u = 2$, and has then been rescaled to have integer coefficients. The first column of the table gives the facet and the second column, explained later in Example 6, gives the set of all points $(x_1, x_2, x_3, y) \in G$ where each facet is satisfied exactly. Consistent with Tables 1 and 2, the notation "x_j" indicates that the variable x_j can have $-2 \leqslant x_j \leqslant 2$, and the value $y = x_1 x_2 x_3$ is suppressed for simplicity. $\diamond$

B.5 Exactness for Supermodular Functions

Example 5. Consider G in the $n = 3$ variables x_1, x_2, x_3, and the variable y, with $m(x) = x_1 x_2 x_3$ as introduced in Example 2 and revisited in Example 3 with $\ell \geqslant 0$. As noted in Example 3, Theorem 5.1 is applicable, and (31) takes the form (B.10), while (32) takes the forms (B.11), (B.12), and (B.13). For $\ell > 0$, $m(x)$ was noted to be strictly supermodular. Then Corollary 5.3

gives us inequality (B.10) is satisfied exactly at only those points $(x, y) \in G$ where x is of the form (x_1, ℓ, ℓ), (u, x_2, ℓ), or (u, u, x_3), matching the first inequality of Table 1. The next five inequalities of Table 1, which follow from permutations of the coefficients of (B.10), have the set of points at which each is satisfied exactly obtained via the same permutations. Corollary 5.3 gives us that inequalities (B.11), (B.12), and (B.13) are satisfied exactly at only those points $(x, y) \in G$ where x has, respectively, at least two entries of value ℓ, at least one entry of value u and at least one entry of value ℓ, and at least two entries of value u. These results of Corollary 5.3 match the last three rows of Table 1. Also as noted in Example 3, the six inequalities $\ell \leqslant x_j \leqslant u$ for $j \in \{1, 2, 3\}$ are all facets of $\mathsf{conv}\, G$. For each of these last six facets, the set of points $(x, y) \in G$ which satisfy it exactly is obvious. For $\ell = 0$, $m(x)$ was noted to be supermodular, but not strictly supermodular. In this case, Theorem 5.2 gives us inequality (B.10) is satisfied exactly at only those points $(x, y) \in G$ having either $x_3 = 0$ or $x_1 = x_2 = u$, matching the first inequality of Table 2. The next two inequalities of Table 2, which follow from permutations of the coefficients of (B.10), have the set of points at which each is satisfied exactly obtained via the same permutations. Also as shown in Example 3, inequalities (32) are the last two inequalities of Table 2. Then Theorem 5.2 gives us that the $y \geqslant 0$ inequality of (B.11) and the $-(u^2)x_1 - (u^2)x_2 - (u^2)x_3 + y \geqslant -2u^3$ inequality of (B.13) are satisfied exactly at only those points $(x, y) \in G$ where x has, respectively, at least one entry of value 0, and at least two entries of value u, matching Table 2. $\diamond$

B.6 Exactness for Monomial

Example 6. Reconsider G in the $n = 3$ variables x_1, x_2, x_3, with $m(x) = 5x_1x_2x_3$, and with $-\ell = u = 2$, as found in Example 4 and Table 3. Recall that each facet within Table 3 has been suitably scaled to handle the coefficient $c_n = 5$ found in $m(x)$ and the variable bounds $-\ell = u = 2$. Therefore, consistent with the discussion at the beginning of this subsection, every point $(\tilde{x}, \tilde{y})$ identified in Theorem 6.3 as satisfying an inequality of the form (35), (36), (37), or (38) exactly for the case in which $c_n = 1$ and $-\ell = u = 1$ must be scaled to $(u\tilde{x}, c_n u^n \tilde{y}) = (2\tilde{x}, 40\tilde{y})$. The second column of Table 3 gives these scaled points for each such inequality, with the value $y = 5x_1x_2x_3$ suppressed for simplicity. Here, the two inequalities in the first block of constraints within Table 3 are scaled (35) for $\beta' = -1$ and $\beta' = 1$, respectively, the inequality in the second block is scaled (36), the inequality in the third block is scaled (37), and the inequalities in the fourth and fifth blocks are the three inequalities resulting from coefficient permutations of the scaled (38), with $t = 0$ and $t = 1$, respectively. $\diamond$