Submodular Maximization and Its generalization through Intersection Cuts

Liding Xu[1][0000−0002−0286−1109] and Leo Liberti[1][0000−0003−3139−6821]

LIX CNRS, École Polytechnique, Institut Polytechnique de Paris, Palaiseau 91128, France liding.xu@polytechnique.edu, liberti@lix.polytechnique.fr

Abstract. We study a mixed-integer set $\mathcal{S} := \{(x,t) \in \{0,1\}^n \times \mathbb{R} : f(x) \geq t\}$ arising in the submodular maximization problem, where f is a submodular function defined over $\{0,1\}^n$. We use the intersection cut paradigm to generate valid inequalities to tighten a polyhedral outer approximation of $\mathcal{S}$. We construct a continuous extension F of f, which is convex and defined over the entire $\mathbb{R}^n$. We show that the epigraph epi(F) of F is an $\mathcal{S}$-free set, and characterize maximal $\mathcal{S}$-free sets including epi(F). Using epi(F), which contains a vertex of the polyhedral outer approximation under some conditions, we use a discrete Newton algorithm to compute an intersection cut. Additionally, our results are generalized to the submodular-supermodular maximization problem, with application to Boolean multilinear programming. We evaluate our techniques on max cut and pseudo Boolean maximization problems within a MIP solver.

Keywords: Outer approximation · Polyhedra · Intersection cuts.

1 Introduction

In this work, we consider the submodular maximization problem:

$$\max_{t \in \mathbb{R}} t \quad \text{s.t.} \quad f(x) \geq t, \quad x \in \{0,1\}^n \cap \mathcal{X}. \tag{1}$$

where $f : \{0,1\}^n \to \mathbb{R}$ is a submodular function and $\mathcal{X} \subseteq \mathbb{R}^n$ is a set describing additional constraints. We study valid inequalities for the mixed-integer set $\text{hypo}_{\{0,1\}^n}(f) := \{(x,t) \in \{0,1\}^n \times \mathbb{R} : f(x) \geq t\}$, which is the hypograph of f over the binary points $\{0,1\}^n$. For arbitrary submodular functions, a class of exponentially many linear valid inequalities [18] exists. With the binary condition $x \in \{0,1\}^n$, these valid inequalities define a Mixed Binary Linear Programming problem equivalent to (1), but, in general, there is no polynomial time algorithm to find the most violated inequality [2]. For a class of special submodular functions, lifting procedures [2,20] can strengthen those general valid inequalities into facet-defining inequalities. Using such strengthened valid inequalities, a Benders-like exact approach based on the branch-and-cut algorithm [2] can provide global dual bounds for primal solutions, and achieve a finite convergence rate.

Many submodular maximization problems, such as MAX CUT with positive edge weights [19], have natural Mixed Binary Linear/Nonlinear Programming formulations. These formulations can be solved via general-purpose global optimization solvers, most of which rely on polyhedral outer approximations [7]. Intersection cuts can strengthen a polyhedral outer approximation of a non-convex set $\mathcal{S}$ that is considered hard to optimize over. The construction of intersection cuts [8] requires two key ingredients: a corner polyhedron relaxation of $\mathcal{S}$, and an $\mathcal{S}$-free set, which is a convex set that does not contain any interior point of $\mathcal{S}$. (Inclusion-wise) maximal $\mathcal{S}$-free sets generate strong intersection cuts that are not dominated by other intersection cuts. In the submodular maximization problem setting, we have $\mathcal{S} = \text{hypo}_{\{0,1\}^n}(f)$. Thereby, our main goals are the construction of $\text{hypo}_{\{0,1\}^n}(f)$-free (in short, hypograph-free) sets, the characterization of maximal hypograph-free sets, and computing intersection cuts. To the best of our knowledge, intersection cuts have not been applied directly to submodular maximization problems.

Intersection cuts were originally devised in the continuous setting, as they were used to approximate the set $\mathcal{S}$ taken as the hypograph of a convex function over a polytope [22] arising in Nonlinear Programming. There is a unique maximal $\mathcal{S}$-free set: the epigraph of that convex function. Later on, intersection cuts were used in the discrete setting [5], where lattice-free sets (e.g. , splits, triangles, and spheres [8,15]) are introduced for Integer Programming problems. The submodular maximization problem plays an intermediate role between these settings. On the one hand, considering the mixed-integer nature of $\text{hypo}_{\{0,1\}^n}(f)$, our first result shows that any maximal $\{0,1\}^n$-free set $\mathcal{C}$ can be lifted into a set $\mathcal{C} \times \mathbb{R}$ which is also maximally hypograph-free. On the other hand, as a discrete analogue to convex functions, the submodular function f has a convex (thus continuous) extension over the box $[0,1]^n$, called Lovász extension [16]. We further extend the Lovász extension to a convex function F over the entire n-dimensional Euclidean space $\mathbb{R}^n$. We show that the epigraph $\text{epi}(\mathsf{F})$ of F is a hypograph-free set, but unlike the continuous setting, $\text{epi}(\mathsf{F})$ is not maximally hypograph-free. However, we can characterize maximal hypograph-free sets including $\text{epi}(\mathsf{F})$. A nice property of $\text{epi}(\mathsf{F})$ is that it contains a vertex of the LP relaxation to the submodular maximization problem under some conditions. We reduce the intersection cut separation problem to solving univariate equations, which we achieve by a discrete Newton-like algorithm [13].

A function is *submodular-supermodular* (SS) if it is the difference of two submodular functions (call them f_1, f_2). We extend our results to generate intersection cuts for the set $\{(x,t) \in \{0,1\}^n \times \mathbb{R} : f_1(x) - f_2(x) \geq \ell t\}$ with $\ell \in \{0,1\}$. We show that a Boolean multilinear function is an SS function, and we derive intersection cuts for Boolean multilinear programming problems. We implement intersection cuts using $\text{epi}(\mathsf{F})$ and the lifted $\{0,1\}^n$-free sets within the `SCIP` solver [7] and test them on MAX CUT and PSEUDO BOOLEAN MAXIMIZATION problems. We show that intersection cuts can strengthen the LP relaxation of these problems.

In this paper we state and prove our main theoretical results Thm. 3, Prop. 6 and Prop. 7, but only state without proof the preliminary results leading to the main one, due to the page limit of this conference paper. The interested reader may find all proofs in the online appendix.

2 Preliminary Notion and Literature Review

To simplify the notation, we denote $\mathcal{N} := [n] := \{1, \cdots, n\}$, $\mathcal{B} := \{0,1\}^n$, $\bar{\mathcal{B}} := [0,1]^n$. We assume that $\mathcal{N}$ is equipped with the natural number order. We denote by $\mathbf{1}$ the all-one vector. For $S \subseteq \mathcal{N}$, we denote by $\mathsf{supp}(S) \in \mathcal{B}$ the characteristic vector of S. For vectors a, b, we let (a, b) be their concatenation, and extend this notation naturally to the case where b is a scalar. Given a set $\mathcal{D} \subseteq \mathbb{R}^n$ and a function $g : \mathcal{D} \to \mathbb{R}$, we adopt the usual notation $\mathrm{epi}_{\mathcal{D}}(g), \mathrm{gra}_{\mathcal{D}}(g), \mathrm{hypo}_{\mathcal{D}}(g)$ to denote the epigraph, graph and hypograph of g over $\mathcal{D}$, respectively. For example, $\mathrm{gra}_{\mathcal{D}}(g) := \{(x,t) \in \mathcal{D} \times \mathbb{R} : g(x) = t\}$. When $\mathcal{D}$ is omitted in the subscript, it is assumed to be $\mathbb{R}^n$. For any set $\mathcal{S}$, we denote by $\mathrm{bd}(\mathcal{S})$, $\mathrm{ext}(\mathcal{S})$, $\mathrm{int}(\mathcal{S})$ its boundary, extreme points, interior, respectively. When $\mathcal{S}$ is not full-dimensional, $\mathrm{relint}(\mathcal{S}), \mathrm{relbd}(\mathcal{S})$ denote its relative interior and relative boundary.

The study of combinatorial structures of submodular functions is due to Edmonds [12], and we refer to [19] for common concepts and definitions. A complete description of the convex hull of the epigraph of a submodular function f (i.e., $\mathrm{conv}(\mathrm{epi}_{\mathcal{B}}(f))$) is given by [3]. Submodular functions are a subclass of discrete convex functions, and we refer to [17] for more details about discrete convex analysis. As already mentioned, intersection cuts generate valid inequalities for hard sets to optimize over.

Assume that we are solving the optimization problem $\min_{z \in \mathcal{S}} cz$. Given a polyhedral outer approximation $\mathcal{P}$ of a non-convex set $\mathcal{S}$, an LP relaxation is $\min_{z \in \mathcal{P}} cz$. Then, an optimal *relaxation point* $\tilde{z}$ is a vertex of $\mathcal{P}$. An intersection cut is a particular type of cut that separates $\tilde{z}$ from $\mathcal{S}$.

Definition 1. *Given $\mathcal{S} \in \mathbb{R}^p$, a closed set $\mathcal{C}$ is called (convex) $\mathcal{S}$-free, if $\mathcal{C}$ is convex and $\mathrm{int}(\mathcal{C}) \cap \mathcal{S} = \varnothing$.*

The construction of intersection cuts [8] needs two components: a corner polyhedron relaxation $\mathcal{R}$ of $\mathcal{S}$ with apex $\tilde{z}$ ($\mathcal{R}$ can be extracted from $\mathcal{P}$), and an $\mathcal{S}$-free set $\mathcal{C}$ containing $\tilde{z}$ in its interior. Then, an intersection cut separates $\tilde{z}$ from $\mathrm{conv}\,(\mathcal{R} \setminus \mathrm{int}(\mathcal{C}))$ (a set which, we note, contains $\mathcal{S}$) as follows. The corner polyhedron $\mathcal{R}$ is given in the half-space and ray representation as $\mathcal{R} := \{z \in \mathbb{R}^p : A(z - \tilde{z}) \leq 0\} = \{z \in \mathbb{R}^p : \exists \eta \in \mathbb{R}^p_+ \; z = \tilde{z} + \sum_{j=1}^p \eta_j r^j\}$, where A is a p-by-p invertible matrix, r^j is the j-th column of $-A^{-1}$ and an extreme ray of $\mathcal{R}$. Define the *step length* $\eta_j^* := \sup_{\eta_j \geq 0}\{\eta_j : \tilde{z} + \eta_j r^j \in \mathcal{C}\}$. The point $\tilde{z}$ is separated by an intersection cut $\sum_{j=1}^p \frac{1}{\eta_j^*} A_j (z - \tilde{z}) \leq -1$. Let $\mathcal{C}, \mathcal{C}^*$ be two $\mathcal{S}$-free sets such that $\mathcal{C} \subseteq \mathcal{C}^*$, then the intersection cut derived from $\mathcal{C}^*$ dominates the intersection cut derived from $\mathcal{C}$. This makes maximal $\mathcal{S}$-free sets relevant in the study.

3 Hypograph-free Sets for Submodular Functions

In this section, we construct hypograph-free sets for submodular functions. W.l.o.g., we assume in the sequel that for any submodular function f, $f(0)=0$ holds (by a translation of a constant). First, we show that one can lift a maximal $\mathcal{B}$-free set into a maximal $\mathrm{hypo}_{\mathcal{B}}(f)$-free set.

Theorem 1. *Let $f:\mathcal{B}\to\mathbb{R}$ be an arbitrary function, and let $\mathcal{K}$ be a maximal $\mathcal{B}$-free set in $\mathbb{R}^n$. Then, $\mathcal{C}:=\mathcal{K}\times\mathbb{R}$ is a maximal $\mathrm{hypo}_{\mathcal{B}}(f)$-free set.*

The construction of set $\mathcal{C}$ does not rely on any structure of f, as it just lifts a $\mathcal{B}$-free set. We next construct $\mathrm{hypo}_{\mathcal{B}}(f)$-free sets using the submodularity of f, for both theoretical and computational interests.

We look at some polyhedra associated with f [3,19]: its *extended polymatroid* defined as $EPM_f := \{s\in\mathbb{R}^n : \forall x\in\mathcal{B}\ sx\le f(x)\}$; the convex hull of the epigraph f over $\mathcal{B}$ defined as $Q_f := \mathrm{conv}(\mathrm{epi}_{\mathcal{B}}(f))$. Recall that $\mathrm{ext}(EPM_f)$ are the vertices of EPM_f, and we can further define a polyhedron

$$EE_f := \{(x,t)\in\mathbb{R}^{n+1} : \forall s\in\mathrm{ext}(EPM_f)\ sx\le t\}. \tag{2}$$

In fact, EE_f includes Q_f, because of the following lemma:

Lemma 1 ([3]). $Q_f = EE_f\cap(\bar{\mathcal{B}}\times\mathbb{R})$.

Therefore, $x\in\bar{\mathcal{B}}$ just defines trivial facets of Q_f, and non-trivial facets of Q_f are defined by $sx\le t$, where s is a vertex of EPM_f.

These polyhedra in turn give rise to some functions associated with f. A convex function g is a *convex underestimating function* of f over $\mathcal{B}$, if for all $x\in\mathcal{B}$, $g(x)\le f(x)$. The *convex envelope* $\mathrm{env}_{\mathcal{B}}(f)$ is thus the maximal convex underestimating function of f over $\mathcal{B}$. By Lemma 1, Q_f is the epigraph of $\mathrm{env}_{\mathcal{B}}(f)$, so $\mathrm{env}_{\mathcal{B}}(f):\bar{\mathcal{B}}\to\mathbb{R}: x\to\max_{s\in\mathrm{ext}(EPM_f)} sx$ (note that the domain is $\bar{\mathcal{B}}$). We remark that $\mathrm{env}_{\mathcal{B}}(f)$ is equivalent to the Lovász continuous extension of f [3,16]. We define the envelope of f extended to $\mathbb{R}^n$ as $\mathsf{F}:\mathbb{R}^n\to\mathbb{R}: x\to\max_{s\in\mathrm{ext}(EPM_f)} sx$. We note that F simply enlarges the domain of $\mathrm{env}_{\mathcal{B}}(f)$ from $\bar{\mathcal{B}}$ to $\mathbb{R}^n$. This extension is algebraically simple, but analytical properties of $\mathsf{F}(x)$ outside $\bar{\mathcal{B}}$ will be studied in further detail. We find that EE_f is the epigraph of F, i.e., $EE_f=\mathrm{epi}(\mathsf{F})$, so F is a convex function. Since every facet $sx\le t$ of EE_f is in one-to-one correspondence to a linear underestimator function sx of F, we call EE_f the *extended envelope epigraph*.

We look at combinatorial structures of the facets of EE_f. Recall that a permutation π on $\mathcal{N}$ is a bijective map from $\mathcal{N}$ to itself. We denote by Π the set of permutations on $\mathcal{N}$, and $\pi(i)\in\mathcal{N}$ is the map of an element $i\in\mathcal{N}$ under this permutation. We define the following sets and vectors related to permutations.

Definition 2. *Given a permutation $\pi\in\Pi$ and $i\in\mathcal{N}\cup\{0\}$, define $\mathcal{N}_i:=\{1,\cdots,i\}$ ($\mathcal{N}_0:=\varnothing$), $\pi(\mathcal{N}_i):=\{\pi(1),\cdots,\pi(i)\}$ ($\pi(\mathcal{N}_0):=\varnothing$), and $v^i(\pi):=\mathsf{supp}(\pi(\mathcal{N}_i))$. Define the map $\sigma:\Pi\to\mathbb{R}^n$ satisfying $\sigma(\pi)_{\pi(i)}=f(v^i(\pi))-f(v^{i-1}(\pi))$ for all $\pi\in\Pi$ and for all $i\in\mathcal{N}$.*

The set of vertices $\text{ext}(EPM_f)$ is the image of Π under the map σ.

Lemma 2 ([12]). $\sigma(\Pi) = \text{ext}(EPM_f)$.

So, every permutation π induces a vertex $\sigma(\pi)$ of $\text{ext}(EPM_f)$ through the map σ. This shows that every facet of EE_f (a non-trivial facet of Q_f) is given as $\sigma(\pi)x \leq t$, and every linear underestimator of F is given as $\sigma(\pi)x$.

Proposition 1. *Given a permutation $\pi \in \Pi$, for all $i \in \mathcal{N} \cup \{0\}$, the facet-defining inequality $\sigma(\pi)x \leq t$ is supported by $\big(v^i(\pi), f(v^i(\pi))\big)$, i.e., $\sigma(\pi)v^i(\pi) = f(v^i(\pi))$.*

Conversely to Prop. 1, given a point in the graph of f, we can construct all the facets supported by it.

Corollary 1. *For a point $v \in \mathcal{B}$, let ι be the number of ones in v. If a permutation $\pi \in \Pi$ satisfies that $v = v^\iota(\pi)$, then $(v, f(v))$ supports the facet-defining inequality $\sigma(\pi)x \leq t$ of EE_f.*

Since EE_f is the epigraph of F, the shape of EE_f is determined by F, so it suffices to look at F. From a convex analysis perspective, the nonsmooth polyhedral function F is the maximum of a set of linear functions, so it is convex and positive homogeneous of degree 1, and thus it is subdifferentiable [14]. Moreover, F has the following analytical properties.

Proposition 2. *For all $x', x \in \mathbb{R}^n$ and all $s \in \partial\mathsf{F}(x')$, $\mathsf{F}(x') = sx'$ and $\mathsf{F}(x) \geq sx$. Moreover, $\partial\mathsf{F}(x') = \text{conv}(\text{argmax}_{s\in\text{ext}(EPM_f)}\, sx')$.*

Given $\tilde{x} \in \mathbb{R}^n$, the evaluation of $\mathsf{F}(\tilde{x})$ is called the *extended polymatroid vertex maximization problem*, as by definition $\mathsf{F}(\tilde{x})$ equals

$$\max_{s\in\text{ext}(EPM_f)} s\tilde{x}. \tag{3}$$

By Prop. 2, an optimal solution s^* is a subgradient of F at $\tilde{x}$ (i.e., $s^* \in \partial\mathsf{F}(\tilde{x})$). By Lemma 2, $\max_{s\in\text{ext}(EPM_f)} s\tilde{x} = \max_{\pi\in\Pi} \sigma(\pi)\tilde{x}$, so Eq. 3 asks for a permutation π^* that maximizes $\sigma(\pi^*)\tilde{x}$. By Prop. 3, Eq. 3 can be solved by a simple *sorting algorithm*, i.e., π^* is a permutation such that $\tilde{x}_{\pi^*(1)} \geq \cdots \geq \tilde{x}_{\pi^*(n)}$.

Proposition 3. *The sorting algorithm finds an optimal solution to the extended polymatroid vertex maximization problem* (3)*.*

From Prop. 1 and Cor. 1, for all $x \in \mathcal{B}$, $(x, f(x))$ supports some facets of EE_f.

Theorem 2. $EE_f \cap \text{hypo}_{\mathcal{B}}(f) = \text{gra}_{\mathcal{B}}(f) \subseteq \text{bd}(EE_f)$.

As already mentioned, F is convex and $EE_f = \text{epi}(\mathsf{F})$, now we find that F is also a continuous extension of f, i.e., for all $x \in \mathcal{B}$, $\mathsf{F}(x) = f(x)$. As EE_f includes Q_f, F further extends $\text{env}_{\mathcal{B}}(f)$ (the Lovász continuous extension). We show that the extended envelope epigraph is a hypograph-free set.

Corollary 2. *EE_f, Q_f are $\mathrm{hypo}_{\mathcal{B}}(f)$-free sets.*

We show the following theorem on (maximal) hypograph-free sets.

Theorem 3. *Let $\mathcal{C}$ be a full-dimensional closed convex set in $\mathbb{R}^{n+1}$ including EE_f. Then, $\mathcal{C}$ is a $\mathrm{hypo}_{\mathcal{B}}(f)$-free set if and only if $\mathcal{C}$ is $\mathrm{gra}_{\mathcal{B}}(f)$-free. Moreover, $\mathcal{C}$ is a maximal $\mathrm{hypo}_{\mathcal{B}}(f)$-free set if and only if $\mathcal{C}$ is a polyhedron and there is at least one point of $\mathrm{gra}_{\mathcal{B}}(f)$ in the relative interior of each facet of $\mathcal{C}$.*

Proof. We note that by Thm. 2, $\mathrm{gra}_{\mathcal{B}}(f) \subseteq \mathrm{bd}(EE_f) \subseteq EE_f \subseteq \mathcal{C}$. Thereby, $\mathrm{gra}_{\mathcal{B}}(f) \cap \mathrm{int}(\mathcal{C}) = \varnothing$ (i.e., $\mathcal{C}$ is $\mathrm{gra}_{\mathcal{B}}(f)$-free) if and only if $\mathrm{gra}_{\mathcal{B}}(f) \subseteq \mathrm{bd}(\mathcal{C})$.

We consider $\mathcal{S}$-freeness first. We prove the forward direction. Assume that $\mathcal{C}$ is a $\mathrm{hypo}_{\mathcal{B}}(f)$-free set. Suppose, to aim at a contradiction, that there exists a point $(v, f(v)) \in \mathrm{int}(\mathcal{C}) \cap \mathrm{gra}_{\mathcal{B}}(f)$. Then, there exists a sufficiently small $\epsilon > 0$ such that $(v, f(v) - \epsilon) \in \mathrm{int}(\mathcal{C})$, but $(v, f(v) - \epsilon) \in \mathrm{hypo}_{\mathcal{B}}(f)$, which leads to a contradiction. We prove the reverse direction. Assume that $\mathcal{C}$ is $\mathrm{gra}_{\mathcal{B}}(f)$-free. Suppose, to aim at a contradiction, that there exists a point $(v, f(v)-\delta) \in \mathrm{int}(\mathcal{C})$ with $v \in \mathcal{B}$ and $\delta > 0$. As, for some $\epsilon > 0$, $(v, f(v) + \epsilon) \subseteq \mathrm{int}(EE_f) \subseteq \mathrm{int}(\mathcal{C})$, by convexity of $\mathcal{C}$, $(v, f(v)) \in \mathrm{int}(\mathcal{C})$, which leads to a contradiction. This implies that $\mathcal{C}$ is $\mathrm{hypo}_{\mathcal{B}}(f)$-free if and only if $\mathrm{gra}_{\mathcal{B}}(f)$-free (or $\mathrm{gra}_{\mathcal{B}}(f) \subseteq \mathrm{bd}(\mathcal{C})$).

We consider the maximality next. Due to [6], a full-dimensional lattice-free set is maximal if it is a polyhedron and there is at least one lattice point in the relative interior of each facet. As $\mathrm{gra}_{\mathcal{B}}(f)$ is a finite set, although it is not a subset of any lattice, the proof strategy is similar. Then, the result follows. □

The set Q_f is the convex hull of $\mathrm{epi}_{\mathcal{B}}(f)$, i.e., the minimal convex set including $\mathrm{epi}_{\mathcal{B}}(f)$. As we aim at obtaining an inclusion-wise maximal $\mathrm{hypo}_{\mathcal{B}}(f)$-free set, we can remove some facets from Q_f and thus enlarge this polyhedron. As already mentioned, after removing trivial facets of Q_f, the enlarged polyhedron is the extended envelope epigraph EE_f. To decide whether EE_f is maximally hypograph-free, we will give a geometrical characterization of the facets of EE_f.

Let $x^0, x^1, \cdots, x^n$ be $n+1$ distinct points of $\mathcal{B}$. They are called *monotone*, if $0 = x^0 < x^1 < \cdots < x^n = \mathbf{1}$. We call the corresponding ordered set $(x^0, \cdots, x^n) \subseteq \mathcal{B}$ a *monotone chain* in $\mathcal{B}$. Therefore, we use a monotone chain to represent a set of monotone points. Then, we have the following observation.

Proposition 4. *The set of monotone chains is in one-to-one correspondence to the set Π of permutations via the map V defined as follows: for all $\pi \in \Pi$, $V(\pi) := (v^i(\pi) \mid i \in \mathcal{N} \cup \{0\})$.*

We find that permutations and monotone chains are indeed equivalent. We note that any $n+1$ distinct points from $\mathrm{gra}_{\mathcal{B}}(f)$ are affinely independent in $\mathbb{R}^{n+1}$ and hence support a hyperplane in $\mathbb{R}^{n+1}$. Thereby, we can infer from Prop. 1 and Prop. 4 that any $n+1$ distinct points $(x^0, f(x^0)), \cdots, (x^n, f(x^n))$ of $\mathrm{gra}_{\mathcal{B}}(f)$ define (or support) a facet of the extended envelope epigraph, if $(x^0, \cdots, x^n)$ is a monotone chain in $\mathcal{B}$. We say that this monotone chain *induces the facet.*

We give an example to show that we can even remove more nontrivial facets that are induced by monotone chains from EE_f, and the new enlarged polyhedron remains $\text{hypo}_{\mathcal{B}}(f)$-free. This example also disproves the maximality of EE_f. Consider $n = 3$, $\mathcal{B} = \{0,1\}^3$, there are 6 permutations, and 6 monotone chains. We assume that, in a non-degenerate case, the associated extended envelope epigraph EE_f has 6 facets induced by 6 chains respectively. The vertices $(0,0,0)$ and $(1,1,1)$ are visited by all the chains, while the other vertices are visited twice each. Therefore, a chain cannot 'exclusively' visit a vertex, so the corresponding facet cannot contain one point of $\text{gra}_{\mathcal{B}}(f)$ in its relative interior. In fact, we can remove some facets from the extended envelope epigraph. We keep three chains:

$$\begin{aligned}
&(\,(0,0,0),(0,0,1),(0,1,1),(1,1,1)\,)\,,\\
&(\,(0,0,0),(0,1,0),(1,1,0),(1,1,1)\,)\,,\\
&(\,(0,0,0),(1,0,0),(1,0,1),(1,1,1)\,)\,.
\end{aligned}$$

These chains induce 3 facets such that at least one point of $\text{gra}_{\mathcal{B}}(f)$ is in the relative interior of each facet and each point of $\mathcal{B}$ is in these 3 facets, so the polyhedron defined by these 3 facets is a $\text{hypo}_{\mathcal{B}}(f)$-free set larger than EE_f.

Therefore, in order to enlarge the extended envelope epigraph, we should remove some of its facet-defining inequalities. This results in a subset of chains (permutations) that suffices to induce a maximal relaxation of the extended envelope epigraph. A subset Π' of permutations of Π is called a cover, if $\bigcup_{\pi\in\Pi'} V(\pi) = \mathcal{B}$; moreover, Π' is called a minimal cover, if additionally, for all $\pi \in \Pi'$, $V(\pi) \setminus \bigcup_{\pi'\in\Pi':\pi'\neq\pi} V(\pi')$ is not empty. It is obvious that Π is always a cover but not a minimal cover. Let Π' be a subset of permutations of Π, $\mathcal{C}(\Pi') := \{(x,t) : \forall\pi \in \Pi'\, \sigma(\pi)x \leq t\}$ denotes the relaxation of the extended envelope epigraph induced by Π'. The following corollary characterizes $\mathcal{C}(\Pi)$.

Corollary 3. *Let Π' be a subset of permutations of Π. $\mathcal{C}(\Pi')$ is $\text{hypo}_{\mathcal{B}}(f)$-free if Π' is a cover. $\mathcal{C}(\Pi')$ is maximally $\text{hypo}_{\mathcal{B}}(f)$-free if Π' is a minimal cover.*

Let us solve the submodular maximization problem (1) via a polyhedral outer approximation $\mathcal{P}$ of $\text{hypo}_{\mathcal{B}}(f)$, and let X be the orthogonal projection of $\mathcal{P}$ on x-space. We remark that, within a branch-and-cut algorithm, X might happen to be within a low-dimensional face of $\bar{\mathcal{B}}$. Let $\tilde{z} := (\tilde{x},\tilde{t})$ be a solution to the LP relaxation $\max_{(x,t)\in\mathcal{P}} t$. We assume that $\tilde{x} \notin \mathcal{B}$, otherwise, $\tilde{x}$ is already an optimal solution to (1). The polyhedral outer approximation $\mathcal{P}$ gives rise to a piece-wise linear concave overestimating function of f over X: $\bar{f}(x) := \max_{(x,t)\in\mathcal{P}} t$, such that $\max_{(x,t)\in\mathcal{P}} t = \max_{x\in X} \bar{f}(x)$. We then have the following observation.

Proposition 5. *Assume that f is not affine over X, and let $x^* \in \text{relint}(X)$. Then $\bar{f}(x^*) > \mathsf{F}(x^*)$, i.e., $(x^*, \bar{f}(x^*)) \in \text{int}(EE_f)$.*

The measure of the relative boundary $\text{relbd}(X)$ is zero, so we can assume that a mild *relative interior condition* that $\tilde{x} \in \text{relint}(X)$ holds with probability one. Then, the relaxation point $\tilde{z}$ is in the relative interior of the extended envelope epigraph with probability one.

4 Hypograph-free and Superlevel-free Sets for SS Functions

We consider hypograph and superlevel sets for an SS function $f := f_1 - f_2$, where f_1 and f_2 are two submodular functions, and let $\mathcal{S} := \{(x,t) \in \mathcal{B} \times \mathbb{R} : f(x) \geq \ell t\}$, with $\ell \in \{0,1\}$. We want to find cutting planes separating $(\tilde{x}, \tilde{t})$ from $\mathcal{S}$.

Let $\mathsf{F}_1 := \max_{s \in EPM_{f_1}} sx$ and $\mathsf{F}_2 := \max_{s \in EPM_{f_2}} sx$ be extended envelopes of f_1, f_2, respectively. As F_1 (resp. F_2) is a convex extension of f_1 (resp. f_2), we have that $\mathcal{S} = \{(x,t) \in \mathcal{B} \times \mathbb{R} : \mathsf{F}_1(x) - \mathsf{F}_2(x) \geq \ell t\}$. By relaxing $\mathcal{B}$ to $\mathbb{R}^n$, a (non-convex) continuous outer approximation of $\mathcal{S}$ is $\overline{\mathcal{S}} := \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \mathsf{F}_1(x) - \mathsf{F}_2(x) \geq \ell t\}$. Moreover, for all $x \in \mathcal{B}$, $(x,t) \in \overline{\mathcal{S}}$ if and only if $(x,t) \in \mathcal{S}$.

Special cases. When $\ell = 1$, $\mathcal{S}$ is the hypograph of the SS function f; when $\ell = 0$, $\mathcal{S}$ is the 0-superlevel set of the SS function f. Setting $f_2 = 0$ and $\ell = 1$, the set $\mathcal{S}$ becomes $\{(x,t) \in \mathcal{B} \times \mathbb{R} : f_1(x) \geq t\}$, which is studied in the previous section. Setting $f_1 = 0$, the relaxed set $\overline{\mathcal{S}}$ becomes $\{(x,t) \in \mathcal{B} \times \mathbb{R} : \mathsf{F}_2(x) \leq -\ell t\}$. If $(\tilde{x}, \tilde{t}) \notin \overline{\mathcal{S}}$, since $\mathsf{F}_2(x) \geq \gamma^* x$ and $\mathsf{F}_2(\tilde{x}) = \gamma^* \tilde{x}$ for any $\gamma^* \in \partial \mathsf{F}_2(\tilde{x})$, the simple outer approximation cut $\gamma^* x \leq t$ is a valid inequality for $\overline{\mathcal{S}}$ (hence for $\mathcal{S}$).

In general, we should separate intersection cuts specifically for SS functions. Let $\gamma^* \in \partial \mathsf{F}_2(\tilde{x})$ be a solution to (3) associated with f_2, and we define the set

$$\mathcal{C}_{\tilde{x}} := \{(x,t) \in \mathcal{B} \times \mathbb{R} : \mathsf{F}_1(x) - \gamma^* x \leq \ell t\}. \tag{4}$$

The following proposition gives $\mathcal{S}$-free sets.

Proposition 6. *The set $\mathcal{C}_{\tilde{x}}$ is an $\overline{\mathcal{S}}$-free set and hence an $\mathcal{S}$-free set. Moreover, if $(\tilde{x}, \tilde{t}) \notin \overline{\mathcal{S}}$, then $\mathcal{C}_{\tilde{x}}$ does not contain $\tilde{x}$ in its interior.*

Proof. We first prove that $\mathcal{C}_{\tilde{x}}$ is $\overline{\mathcal{S}}$-free. By definition, $\gamma^* x \leq \mathsf{F}_2(x)$, which implies that $\mathsf{F}_1(x) - \gamma^* x \geq \mathsf{F}_1(x) - \mathsf{F}_2(x)$. Therefore, for $(x,t) \in \operatorname{int}(\mathcal{C}_{\tilde{x}})$, we have that $\ell t > \mathsf{F}_1(x) - \gamma^* x \geq \mathsf{F}_1(x) - \mathsf{F}_2(x)$, which implies that $(x,t) \notin \overline{\mathcal{S}}$. Hence, $\operatorname{int}(\mathcal{C}_{\tilde{x}}) \cap \overline{\mathcal{S}} = \varnothing$. Additionally, $\mathcal{C}_{\tilde{x}}$ is convex. Two facts imply that $\mathcal{C}_{\tilde{x}}$ is $\overline{\mathcal{S}}$-free. Next, assume that $(\tilde{x}, \tilde{t}) \notin \overline{\mathcal{S}}$, then $\ell \tilde{t} > \mathsf{F}_1(\tilde{x}) - \mathsf{F}_2(\tilde{x}) \leq \mathsf{F}_1(\tilde{x}) - \gamma^* \tilde{x}$, so $(\tilde{x}, \tilde{t}) \in \operatorname{int}(\mathcal{C}_{\tilde{x}})$. $\square$

We discuss applications of previous results to Boolean Multilinear Programming (BMP) problems. We have the following observation.

Proposition 7. *Given a Boolean multilinear function $f : \mathcal{B} \to \mathbb{R} : x \to f(x) := \sum_{k \in [K]} a_k \prod_{j \in A_k} x_j$ ($A_k \subseteq \mathcal{N}$) with K multilinear terms, let $f = f_1 - f_2$ where $f_1(x) := \sum_{k \in [K]: a_k < 0} a_k \prod_{j \in A_k} x_j$ and $f_2(x) := \sum_{k \in [K]: a_k > 0} -a_k \prod_{j \in A_k} x_j$. Then, f_1, f_2 are submodular over $\mathcal{B}$.*

Proof. Given a Cartesian product set $D := \prod_{j \in \mathcal{N}} D_j$ ($D_j \subseteq \mathbb{R}$), a function $g : D \to \mathbb{R}$ is a generalized supermodular over D, if for every $x, y \in D$, $g(x) + g(y) \leq g(x \vee y) + g(x \wedge y)$. Each multilinear term function $\prod_{j \in A_k} x_j$ is a Cobb-Douglas function [21], which is a generalized supermodular function over $\mathbb{R}^n$. It is known [21] that, if restricting the domain (e.g. $\mathbb{R}^n$) to its subdomain (e.g. $\mathcal{B}$) still yields a Cartesian product set, then the supermodularity is preserved. Moreover, a negative combination of supermodular functions is a submodular function. Therefore, f_1, f_2 are submodular functions over $\mathcal{B}$. $\square$

Since every Boolean multilinear function is an SS function, We can construct intersection cuts for a BMP problem in the following form:

$$\max_{t\in\mathbb{R},x\in\mathcal{B}} t \text{ s.t. } \sum_{k\in\mathcal{K}_0} a_{ik} \prod_{j\in A_k} x_j \geq t,\ \forall i \in [m] \sum_{k\in\mathcal{K}_i} a_{ik} \prod_{j\in A_k} x_j \geq 0, \tag{5}$$

where m is the number of constraints, $\mathcal{K} := [K]$, $\mathcal{K}_i \subseteq \mathcal{K}$ is the index set of multilinear terms in the i-th constraint (0 for objective). For all $i \in [m]$ or $i = 0$, denote $f_i(x) := \sum_{k\in\mathcal{K}_i} a_{ik} \prod_{j\in A_k} x_j$, and let $f_i(x) = f_{i1}(x) - f_{i2}(x)$, where $f_{i1} = \sum_{k\in\mathcal{K}_i:a_{ik}<0} a_{ik} \prod_{j\in A_k} x_j$ and $f_{i2} = \sum_{k\in\mathcal{K}_i:a_{ik}>0} -a_{ik} \prod_{j\in A_k} x_j$ are two corresponding submodular functions. The objective and constraints of (5) can be represented as $f_{i1}(x) - f_{i2}(x) \geq \ell_i t$ ($\ell_i = 0$ for $i \in [m]$, and $\ell_0 = 1$). The method to construct intersection cuts for the BMP is straightforward: we directly separate intersection cuts constructed from the $\mathcal{S}$-free sets given by Prop. 6.

5 Separation Problem

In this section, we consider the separation problem to generate an intersection cut using an $\mathcal{S}$-free set. From the previous sections, the $\mathcal{S}$-free set is assumed to have the form of $\mathcal{C} := \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \mathsf{G}(x) \leq \ell t\}$, where $\mathsf{G}(x) = \max_{s\in \mathrm{ext}(EPM_g)} sx$ is the extended envelope of some submodular function g over $\mathcal{B}$ and $\ell \in \{0, 1\}$. We remark that the extended envelope epigraph EE_f (2) is a special case with $\ell = 1$ and $g = f$; the set $\mathcal{C}_{\tilde{x}}$ (4) is also a special case that $g(x) = f_1(x) - \gamma^* x$.

Assume that $z^* := (\tilde{x}, \tilde{t})$ is a vertex of the corner polyhedron $\mathcal{R}$, and $z^* \in \mathrm{int}(\mathcal{C})$. Recall Sect. 2, and the separation problem is reduced to calculate the step length along each ray r^j:

$$\eta_j^* = \sup_{\eta_j\geq 0} \{\eta_j : z^* + \eta_j r^j \in \mathcal{C}\}. \tag{6}$$

This is a line search problem that computes the step length to the border of $\mathcal{C}$ along the ray r^j from the interior point z^*. We denote by r_x^j, r_t^j the projection of r^j on x- and t- spaces. We define the following univariate function $\zeta^j : \mathbb{R}_+ \to \mathbb{R} : \eta_j \to \zeta^j(\eta_j) := \ell(\tilde{t} + r_t^j \eta_j) - \mathsf{G}(\tilde{x} + r_x^j \eta_j)$.

Proposition 8. *ζ^j is a concave piece-wise linear function over $[0, +\infty]$ with $\zeta^j(0) > 0$. If $\eta_j^* < \infty$ and there exists an $\eta_j' > 0$ with $\zeta^j(\eta_j') = 0$, then $\eta_j' = \eta_j^*$, i.e., the root η_j^* must be unique. For all $s^* \in \mathrm{argmax}_{s\in\mathrm{ext}(EPM_g)} s(\tilde{x} + \eta_j r_x^j)$, $\ell r_t^j - s^* r_x^j$ is a subgradient in $\partial\zeta^j(\eta_j)$. For $\eta_j > \eta_j^*$, $\partial\zeta^j(\eta_j) \leq \partial\zeta^j(\eta_j^*)$.*

The evaluation and subgradient computation of ζ^j are reduced to the evaluation and subgradient computation of G, which can be solved by a sorting algorithm (Prop. 3). By Prop. 8, we can reformulate the line search problem (6) into solving univariate nonlinear equations: for each ray r^j, compute the unique root (or zero) of the univariate function ζ^j, or certificate no such root exists. Thereby, to solve the solving univariate nonlinear equations, it is natural to deploy the Newton like algorithm, which requires a step-length decision. Because the function ζ^j is piece-wise linear and uses a sorting algorithm to compute subgradients, our implementation is similar to the *discrete Newton algorithm* in [13].

6 Computational Results

In this section, we conduct computational experiments to test the efficiency of the proposed cuts.

Setup. The experiments are conducted on a server with Intel Xeon W-2245 CPU @ 3.90GHz and 126GB main memory. We use `SCIP` 8.0.3 [7] as a framework to construct LP relaxations of MINLPs and perform cut separation. `SCIP` is equipped with `CPLEX` 22.1 as an LP solver, `IPOPT` 3.14.7 as an NLP solver. The time limit of each test is set to 3600 CPU seconds. We use `SCIP` to solve the natural formulation of a test problem, and intersection cuts are added via a *cut separator*. By Thm. 1, the simple lifted-split $H_j := \{x \in \mathbb{R}^n : 0 \leq x_j \leq 1\} \times \mathbb{R}$ is a maximal $\text{hypo}_{\mathcal{B}}(f)$-free set for any f. We test intersection cuts derived from both EE_f and H_j, where the splitting variable x_j is chosen to be the most fractional entry of the relaxation solution.

Performance metrics. We focus on the root node performance of intersection cuts: we measure the performance by the *closed root gap*. Let d_1 be the first solution value of the initial LP relaxation (without cuts added), and let d_2 be the dual bound obtained when the algorithm finishes at the root node (after cuts are added), and let p be a reference primal bound, the closed root gap $(d_2 - d_1)/(p - d_1)$ is the gap closed improvement of d_2 with respect to d_1. In addition, we record the number of cuts separated, the relative improvement to the `SCIP`'s default setting and the total running time. With each benchmark, we compute the shifted geometric mean (shift value: 1) of these statistics.

Experiment 1: Max cut. Given an undirected graph $G = (\mathcal{N}, E, w)$, where $\mathcal{N}$ is the set of vertices, E is the set of edges, and w is a weight function over E. For a subset S of $\mathcal{N}$, its associated cut capacity is the sum of the weights of edges with end nodes both in S and $\mathcal{N} \setminus S$. The MAX CUT problem aims to find a subset with maximum cut capacity. Let binary variable $x \in \mathcal{B}$ indicate whether vertices belong to S, then the problem can be formulated as the following Quadratic Unconstrained Binary Optimization (QUBO) problem: $\max_{x\in\mathcal{B}} \sum_{\{i,j\}\in E} w(\{i,j\})((1-x_i)x_j + x_i(1-x_j))$, where the objective function is the cut capacity of S. When w is positive (i.e., $\forall\{i,j\} \in E \; w(\{i,j\}) \geq 0$), the cut capacity function is non-monotone and submodular. Here, we consider this specific case. Our benchmark contains 30 "g05" and 30 "pw" instances from Biq Mac library. We use the best-known primal bound from Biq Mac library as the reference primal bound. We encode the hypograph reformulation (1) of the QUBO in `SCIP`. `SCIP` will automatically reformulate the problem into an MILP via Reformulation-and-Linearization Technique [1]. In Table 1, we report the computational result.

Experiment 2: Pseudo Boolean maximization. Every pseudo Boolean function from $\mathcal{B}$ to $\mathbb{R}$ can be written uniquely as a Boolean multilinear function

Default		Submodular cut				Split cut			
closed	time	closed	relative	time	cuts	closed	relative	time	cuts
0.22	8.31	0.27	1.22	152.20	76.24	0.27	1.21	31.76	45.15

Table 1. Summary of MAX CUT results

over $\mathcal{B}$. Therefore, PSEUDO BOOLEAN MAXIMIZATION is usually formulated as a Multilinear Unconstrained Binary Optimization (MUBO) problem, a generalization of QUBO. Since then MUBO is an unconstrained BMP, we can use techniques from Sect. 4. `SCIP` constructs an extended formulation of the MUBO by lifting each multilinear term and applies the standard linearization [9] to obtain an equivalent MILP reformulation. Some hypergraph-theoretical based linear inequalities [10,11] are valid for the extended formulation. Our benchmark contains 44 highly dense "autocorr_bern" MUBO instances from MINLPLib. We use the best-known primal bound from MINLPLib as the reference primal bound. In Table 2, we report the computational result.

Default		Submodular cut				Split cut			
closed	time	closed	relative	time	cuts	closed	relative	time	cuts
0.10	15.33	0.11	1.13	35.33	14.52	0.10	1.01	18.09	4.36

Table 2. Summary of PSEUDO BOOLEAN MAXIMIZATION results

7 Conclusion

We construct hypograph-free sets for submodular functions, and derive intersection cuts for submodular maximization. Although these sets are not inclusion-wise maximal, we give a characterization of maximal hypograph-free sets including these hypograph-free sets. We generalize our results for submodular-supermodular functions, and apply them to Boolean Multilinear Programming. Future studies will extend our methods for discrete submodular functions on general lattices [4], settle the complexity status of the discrete Newton algorithm and test more problems.

References

1. Adams, W.P., Sherali, H.D.: A tight linearization and an algorithm for zero-one quadratic programming problems. Management Science **32**(10), 1274–1290 (1986)
2. Ahmed, S., Atamtürk, A.: Maximizing a class of submodular utility functions. Mathematical programming **128**(1), 149–169 (2011)
3. Atamtürk, A., Narayanan, V.: Submodular function minimization and polarity. Mathematical Programming (2021)

4. Bach, F.: Submodular functions: from discrete to continuous domains. Mathematical Programming **175**(1), 419–459 (2019)
5. Balas, E.: Intersection cuts—a new type of cutting planes for integer programming. Operations Research **19**(1), 19–39 (1971)
6. Basu, A., Conforti, M., Cornuéjols, G., Zambelli, G.: Maximal lattice-free convex sets in linear subspaces. Mathematics of Operations Research **35**(3), 704–720 (2010)
7. Bestuzheva, K., Besançon, M., Chen, W.K., Chmiela, A., Donkiewicz, T., van Doornmalen, J., Eifler, L., Gaul, O., Gamrath, G., Gleixner, A., et al.: The scip optimization suite 8.0. arXiv preprint arXiv:2112.08872 (2021)
8. Conforti, M., Cornuéjols, G., Zambelli, G.: Integer programming. Springer International Publishing, Cham (2014)
9. Crama, Y.: Concave extensions for nonlinear 0–1 maximization problems. Mathematical Programming **61**(1), 53–60 (1993)
10. Del Pia, A., Khajavirad, A.: The multilinear polytope for acyclic hypergraphs. SIAM Journal on Optimization **28**(2), 1049–1076 (2018)
11. Del Pia, A., Walter, M.: Simple odd-cycle inequalities for binary polynomial optimization. In: International Conference on Integer Programming and Combinatorial Optimization. pp. 181–194. Springer (2022)
12. Edmonds, J.: Submodular functions, matroids, and certain polyhedra. In: Combinatorial Optimization—Eureka, You Shrink!, pp. 11–26. Springer (2003)
13. Goemans, M.X., Gupta, S., Jaillet, P.: Discrete newton's algorithm for parametric submodular function minimization. In: Eisenbrand, F., Koenemann, J. (eds.) Integer Programming and Combinatorial Optimization. pp. 212–227. Springer International Publishing, Cham (2017)
14. Hiriart-Urruty, J.B., Lemaréchal, C.: Fundamentals of convex analysis. Springer Science & Business Media (2004)
15. Liberti, L.: Spherical cuts for integer programming problems. International Transactions in Operational Research **15**(3), 283–294 (2008)
16. Lovász, L.: Submodular functions and convexity. In: Mathematical programming the state of the art, pp. 235–257. Springer (1983)
17. Murota, K.: Discrete convex analysis. Mathematical Programming **83**(1), 313–371 (1998)
18. Nemhauser, G.L., Wolsey, L.A., Fisher, M.L.: An analysis of approximations for maximizing submodular set functions—i. Mathematical programming **14**(1), 265–294 (1978)
19. Schrijver, A., et al.: Combinatorial optimization: polyhedra and efficiency, vol. 24. Springer (2003)
20. Shi, X., Prokopyev, O.A., Zeng, B.: Sequence independent lifting for a set of submodular maximization problems. Mathematical Programming pp. 1–46 (2022)
21. Topkis, D.M.: Supermodularity and complementarity. Princeton university press, Princeton (2011)
22. Tuy, H.: Concave programming with linear constraints. In: Doklady Akademii Nauk. vol. 159, pp. 32–35. Russian Academy of Sciences (1964)

A Appendix

A.1 Proof for Thm. 1

Proof. We note that $\text{int}(\mathcal{C}) = \text{int}(\mathcal{K}) \times \mathbb{R}$. It is easy to show that $\mathcal{C}$ is $\text{hypo}_{\mathcal{B}}(f)$-free, since $\text{int}(\mathcal{C}) \cap \text{hypo}_{\mathcal{B}}(f) = \varnothing$. Assume that there exists a $\text{hypo}_{\mathcal{B}}(f)$-free set $\mathcal{C}'$ including $\mathcal{C}$. Then, the recession cone of $\mathcal{C}'$ must include that of $\mathcal{C}$, so $\mathcal{C}' = \mathcal{K}' \times \mathbb{R}$ for some closed convex set $\mathcal{K}'$ including $\mathcal{K}$. Moreover, $\mathcal{K}'$ must be a $\mathcal{B}$-free set, otherwise, there exists a point $x \in \mathcal{B} \cap \mathrm{int}(\mathcal{K}')$ such that $(x, f(x)) \in \mathrm{int}(\mathcal{K}') \times \mathbb{R} = \mathrm{int}(\mathcal{C}')$. However, since $\mathcal{K}$ is maximally $\mathcal{B}$-free, this implies that $\mathcal{K} = \mathcal{K}'$. As a result, $\mathcal{C} = \mathcal{C}'$, so $\mathcal{C}$ is maximal. □

Proof for Prop. 1

Proof.

$$\sigma(\pi)v^i(\pi) = \sum_{j\in\mathcal{N}_i} \sigma(\pi)_{\pi(j)} = \sum_{j\in\mathcal{N}_i} \Big(f(v^j(\pi)) - f(v^j(\pi))\Big) = f^i(v^i(\pi)) - f(0) = f(v^i(\pi)),$$

where the first equation follows from Defn. 2, the second equation follows from Lemma 2, and the last two equations follow from the expansion of the sum. □

A.2 Proof for Prop. 2

Proof. As $\mathsf{F}(x) = \max_{s\in\mathrm{ext}(EPM_f)} sx$, F is the maximum of a set of linear functions. This implies that it is positive-homogeneous of degree-1 and convex, and it is easy to show the other results. □

A.3 Proof for Prop. 3

Proof. Let π^* be the permutation found by the sorting algorithm. By Lemma 2, $\sigma(\pi^*)$ is in $\mathrm{ext}(EPM_f)$ and hence a feasible solution to (3). Next, we prove the optimality of $\sigma(\pi^*)$. Let $d = \min_{i\in\mathcal{N}} \tilde{x}_i$, then $\tilde{x} + d\mathbf{1} \geq 0$. The following inequalities hold:

$$\sigma(\pi^*)\tilde{x} \leq \max_{s\in\mathrm{ext}(EPM_f)} s(\tilde{x}+d\mathbf{1}-d\mathbf{1}) \leq \max_{s\in\mathrm{ext}(EPM_f)} s(\tilde{x}+d\mathbf{1}) + \max_{s\in\mathrm{ext}(EPM_f)} s(-d\mathbf{1}). \tag{7}$$

It is easy to show that $(\tilde{x} + d\mathbf{1})_{\pi^*(1)} \geq \cdots \geq (\tilde{x} + d\mathbf{1})_{\pi^*(n)}$, as $\tilde{x} + d\mathbf{1} \geq 0$, by the sorting algorithm, $\sigma(\pi^*)(\tilde{x}+d\mathbf{1}) = \max_{s\in\mathrm{ext}(EPM_f)} s(\tilde{x}+d\mathbf{1})$. As the entries of $-d\mathbf{1}$ are identical, it follows from Prop. 1 that for any $\pi \in \Pi$, $\sigma(\pi)(-d\mathbf{1}) = -df(\mathbf{1})$. Therefore, for any $s \in \mathrm{ext}(EPM_f)$, $s(-d\mathbf{1}) = -df(\mathbf{1})$, so $\sigma(\pi^*)(-d\mathbf{1}) = \max_{s\in\mathrm{ext}(EPM_f)} s(-d\mathbf{1})$. Looking at the inequalities (7), they become equations, because

$$\sigma(\pi^*)\tilde{x} \leq \max_{s\in\mathrm{ext}(EPM_f)} s\tilde{x} \leq \sigma(\pi^*)(\tilde{x} + d\mathbf{1}) + \sigma(\pi^*)(-d\mathbf{1}) = \sigma(\pi^*)\tilde{x}.$$

Therefore, $\sigma(\pi^*)$ is an optimal solution to (3). □

A.4 Proof for Thm. 2

Proof. We consider a point $v \in \mathcal{B}$ and look at the line $\ell = \{(v,t) : t \in \mathbb{R}\}$. It can be separated into the restricted epigraph $\ell_+ := \{(v,t) : f(v) \leq t\}$ and the restricted hypograph $\ell_- := \{(v,t) : f(v) \geq t\}$, as $\ell_+ \cap \ell_- = (v, f(v))$ and $\ell = \ell_+ \cup \ell_-$. First, we know that, by definition of Q_f and Lemma 1, $\ell_+ \subseteq Q_f \subseteq EE_f$. Second, by Prop. 1, the point $(v, f(v))$ is supported by some facets of EE_f, so the point (v,t) with $t < f(v)$ is separated by these facets from EE_f. Thereby, we know that $\ell_- \cap EE_f = \{(v, f(v))\}$. To summarize, we know that $EE_f \cap \ell = \ell_+$ and $(v, f(v)) \in \mathrm{bd}(EE_f)$. As $\mathrm{gra}_{\mathcal{B}}(f) = \cup_{v \in B}\{(v, f(v))\}$, we have that $\mathrm{gra}_{\mathcal{B}}(f) \subseteq \mathrm{bd}(EE_f)$. As the hypograph $\mathrm{hypo}_{\mathcal{B}}(f) = \cup_{v \in \mathcal{B}}\{(v,t) : f(v) \geq t\}$ (union of restricted hypographs), we have that $EE_f \cap \mathrm{hypo}_{\mathcal{B}}(f) = \mathrm{gra}_{\mathcal{B}}(f)$. $\square$

A.5 Proof for Cor. 2

Proof. Since $\mathrm{gra}_{\mathcal{B}}(f) \subseteq \mathrm{bd}(EE_f)$, we conclude that $EE_f \cap \mathrm{hypo}_{\mathcal{B}}(f) \subseteq \mathrm{bd}(EE_f)$ and hence $\mathrm{int}(EE_f) \cap \mathrm{hypo}_{\mathcal{B}}(f) = \varnothing$. Additionally, EE_f is convex and hence $\mathrm{hypo}_{\mathcal{B}}(f)$-free. As $Q_f \subseteq EE_f$, Q_f is $\mathrm{hypo}_{\mathcal{B}}(f)$-free set. $\square$

A.6 Proof for Prop. 4

Proof. By Prop. 1, since $\varnothing = \pi(\mathcal{N}_0) \subsetneq \cdots \subsetneq \pi(\mathcal{N}_n) = \mathcal{N}$, by Defn. 2, $0 = v^0(\pi) < \cdots < v^n(\pi) = 1$, so $V(\pi)$ is a monotone chain. Conversely, given a monotone chain $(x^0, \cdots, x^n)$, its inverse map π exists and satisfies that $\pi(0) = 0$; and for all $i \in \mathcal{N}$, $\pi(i)$ is the index of the non-zero entry of $x^i - x^{i-1}$. $\square$

A.7 Proof for Cor. 3

Proof. First, we note that $C(\Pi')$, as a relaxation of EE_f includes $\mathrm{gra}_{\mathcal{B}}(f)$. Next, we assume that Π' is a cover. Then, points of $\mathrm{gra}_{\mathcal{B}}(f)$ are supported by facets of $C(\Pi')$. By Thm. 3, $C(\Pi')$ is $\mathrm{hypo}_{\mathcal{B}}(f)$-free. Finally, if Π' is a minimal cover, then each facet of $C(\Pi')$ has a point of $\mathrm{gra}_{\mathcal{B}}(f)$ in its interior. By Thm. 3, $C(\Pi')$ is maximally $\mathrm{hypo}_{\mathcal{B}}(f)$-free. $\square$

A.8 Proof for Prop. 5

Proof. As $\bar{f}$ is concave overestimator of f over X and F is convex underestimator of f over X, $\bar{f} \geq \mathsf{F}$ over X. Suppose, to aim at a contradiction, that $\bar{f}(x^*) = \mathsf{F}(x^*)$. Define a concave function $g := \bar{f} - \mathsf{F}$, then for all $x \in X$, $g(x) \geq 0$, and $g(x^*) = 0$. By its concavity, there exists an affine overestimating function a of g, such that $g(x^*) = a(x^*) = 0$, and for all $x \in X$, $0 \leq g(x) \leq a(x)$. As $x^* \in \mathrm{relint}(X)$, the affinity of a implies that $a = g = 0$ over X, i.e., $\bar{f} = \mathsf{F}$ over X. So f is both concave and convex over X and thus affine over X, which is a contradiction. $\square$

A.9 Proof for Prop. 8

Proof. Since the extended envelope G is the maximum of linear functions, it is convex and piece-wise linear, and so ζ^j is concave and piece-wise linear. Since $\zeta^j(0) = \ell\tilde{t} - \mathsf{G}(\tilde{x})$, it follows from the assumption $z^* \in \operatorname{int}(\mathcal{C})$ that $\tilde{t} > \mathsf{G}(\tilde{x})$ and thus $\zeta^j(0) > 0$. Since $\mathcal{C}$ is closed and convex, $\eta'_j = \eta^*_j$ if and only if $z^* + \eta'_j r^j \in \operatorname{bd}(\mathcal{C})$. That is $\mathsf{G}(r^j_x\eta_j + \tilde{x}) = \mathsf{G}(\tilde{x}) + r^j_t\eta'_j$, i.e., $\zeta^j(\eta'_j) = 0$. Since $s^* \in \partial\mathsf{G}(\tilde{x} + r^j_x\eta_j)$, by the chain rule, $\ell r^j_t - s^* r^j_x$ is a subgradient of ζ^j. By the concavity of ζ^j, its subgradients are non-increasing. □