Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Characterization of Simultaneous Optimization, Majorization, and (Bi-)Submodular Polyhedra

Martijn H. H. Schoot Uiterkamp

informs
https://pubsonline.informs.org/journal/moor

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 50, No. 1, February 2025, pp. 252–276
ISSN 0364-765X (print), ISSN 1526-5471 (online)

A Characterization of Simultaneous Optimization, Majorization, and (Bi-)Submodular Polyhedra

Martijn H. H. Schoot Uiterkamp[a]

[a] Department of Econometrics and Operations Research, Tilburg University, 5000 LE Tilburg, Netherlands

Contact: m.h.h.schootuiterkamp@tilburguniversity.edu, https://orcid.org/0000-0002-8125-0479 (MHHSU)

Received: February 27, 2023
Revised: October 31, 2023
Accepted: January 1, 2024
Published Online in Articles in Advance: February 20, 2024

MSC2020 Subject Classifications: Primary: 90C27; secondary: 62J07, 91A12, 91B32

https://doi.org/10.1287/moor.2023.0054

Abstract. Motivated by resource allocation problems (RAPs) in power management applications, we investigate the existence of solutions to optimization problems that simultaneously minimize the class of Schur-convex functions, also called least-majorized elements. For this, we introduce a generalization of majorization and least-majorized elements, called (a, b)-majorization and least (a, b)-majorized elements, and characterize the feasible sets of problems that have such elements in terms of base and (bi-)submodular polyhedra. Hereby, we also obtain new characterizations of these polyhedra that extend classical characterizations in terms of optimal greedy algorithms from the 1970s. We discuss the implications of our results for RAPs in power management applications and derive a new characterization of convex cooperative games and new properties of optimal estimators of specific regularized regression problems. In general, our results highlight the combinatorial nature of simultaneously optimizing solutions and provide a theoretical explanation for why such solutions generally do not exist.

Keywords: simultaneous optimization • majorization • Schur convexity • least-majorized element • submodular polyhedra • bisubmodular polyhedra • resource allocation problem • convex cooperative game • regularized regression

1. Introduction

1.1. Resource Allocation and Simultaneous Optimization

Efficient resource allocation is an important challenge for many modern computing systems. The general goal in such resource allocation problems (RAPs) is to divide the available resource over individual users or system components as to maximize the overall system utility of this division (or equivalently, to minimize the overall system cost). These problems occur in many different engineering applications, such as telecommunications (Palomar and Fonollosa [39], Shams et al. [44]), processor scheduling (Gerards et al. [26]), and regularized learning (Dai and Fletcher [11], Mairal et al. [35]; see also the overview in Patriksson [40]).

For many allocation problems, there has been interest in the development of algorithms that find allocations with good optimality guarantees that hold for multiple utility or cost functions simultaneously. One motivation for this is that different definitions of the cost of an allocation may exist that are, ideally, minimized simultaneously. For instance, in the context of smart storage systems in distribution grids, it is crucial that the system is operated so that the stress put on the grid is reduced as much as possible. This stress can be reduced in several ways: for instance, by minimizing the peak energy consumption or flattening the overall load profile as much as possible. This leads to different objectives and corresponding cost functions that are preferably optimized at the same time (see also Schoot Uiterkamp et al. [42]).

Another motivation is that it is often hard in practice to specify the utility or cost function exactly because of, for example, missing data or the absence of objective measures for such a function. Instead, usually only some structural properties, such as concavity or symmetry, can be observed or assumed with reasonable certainty. Therefore, it is desirable to have a solution procedure that provides good solutions regardless of the exact description of the objective function. One example of such an application is speed scaling, where the goal is to schedule tasks on a processor and determine execution speeds for these tasks while minimizing the total energy usage of the processor and respecting any restrictions on the execution of the tasks, such as deadlines. A common observation in this area is that the total energy usage depends primarily on the chosen speeds and that this dependence is convex. However, the exact relation between these two quantities depends on the specific properties of the used processor and is, therefore, generally unknown (see also Gerards et al. [26]).

Most research on finding allocations with good performance guarantees for whole classes of objective functions focus on approximate solutions rather than optimal ones (see, e.g., Goel and Meyerson [27]) The main reason for this is that for many problems of interest, there is no solution that simultaneously optimizes whole classes of utility or cost functions. However, little is known about why this is the case and whether there exists a characterization of problems that do have such a solution. These questions are the main motivation for the research in this article.

1.2. Simultaneous Optimization, Majorization, and (Bi-)Submodularity

The existence of solutions that minimize whole classes of objective functions is closely related to the concept of majorization (Marshall et al. [36]). Majorization is a formalization of the vague notion that the components of one vector are less spread out (or closer together) than those of another vector. We say that, in this case, the former vector is *majorized* by the latter. Majorization can be characterized in several ways, one of which is via Schur-convex functions. More precisely, given two vectors $x, y \in \mathbb{R}^n$, x is majorized by y if and only if $\Phi(x) \leq \Phi(y)$ for all Schur-convex functions Φ. Several variants of majorization exist that can also be characterized in terms of better objective values for particular classes of Schur-convex functions. Examples of these are weak submajorization, weak supermajorization, and weak absolute majorization, which are characterized by better objective values for nondecreasing, nonincreasing, and element-wise even Schur-convex functions, respectively (see also Marshall et al. [36] and Section 3.1).

The class of Schur-convex functions is broad and includes, for example, symmetric quasiconvex functions. In particular, many common utility and cost functions in RAPs are special cases of Schur-convex functions. As a consequence, majorization and its characterization in terms of Schur-convex functions have many applications in, for example, telecommunications (Jorswieck and Boche [34]) and economics (Arnold and Sarabia [4]; see also Marshall et al. [36]).

Solutions that are majorized by all other vectors in a given feasible set and thereby, also simultaneously minimize all Schur-convex functions over this set are known as least-majorized elements. The existence of such elements has hardly been investigated in the literature. In fact, the only nontrivial sets for which the existence of such elements has been established are (submodular) base polyhedra (Dutta and Ray [14]; see also the historical overview in Frank and Murota [19]).

Several other connections exist between majorization and (bi-)submodularity. For instance, majorization can be characterized in terms of submodular functions (see, e.g., Fujishige [22, p. 44]), and a similar characterization of weak absolute majorization exists in terms of bisubmodular functions (Zhan [48]). Furthermore, several combinatorial sets that are in some way related to base polyhedra have least weakly sub- or supermajorized elements (i.e., vectors that are weakly sub- or supermajorized by all other vectors in the given set). Examples of such sets are bounded generalized polymatroids (Tamir [45]) and jump systems (Ando [1]).

Majorization and least-majorized elements are also important concepts in economic theories of fair allocation of resources (see, e.g., Arnold and Sarabia [4] for an overview). In particular, they play an important role in convex cooperative games, which have the unique property that their core (i.e., the set of all payoffs where no group of players has an incentive to form their own competitive coalition) is a base polyhedron (see also Ichiishi [31]). Therefore, several key properties of submodular functions and base polyhedra can be proven using tools from cooperative game theory. In particular, existence results for particular solution vectors in the core of convex games translate directly into existence results of such vectors for base polyhedra. Examples of these are the egalitarian solution (Dutta and Ray [14]) and min-max fair and utilitarian solutions for uniform utility Nash and nonsymmetric bargaining games (Chakrabarty et al. [10]) (we come back to this in Section 6.2).

1.3. Contributions

A natural question is whether the existence of least (weakly absolutely) majorized and least weakly sub- or supermajorized elements is unique for base polyhedra and related sets. As discussed in the previous subsection, this directly relates to the question to what extent a characterization exists of optimization problems and solutions to these problems that simultaneously optimize multiple (Schur-convex) objectives. In this article, we provide an answer to these questions and in the process, reveal a strong relation between majorization and base, submodular, supermodular, and bisubmodular polyhedra.

If we consider the "classical" concept of majorization as discussed in the previous subsection, the answer to the first question is negative. For instance, given $R \in \mathbb{R}$, the vector $R\left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$ is a least-majorized element of any subset of $\mathbb{R}^n$ that contains it and whose vectors have the same element sum R. This observation follows directly from Jensen's inequality. Similarly, the zero vector is a least weakly absolutely majorized element of any subset of $\mathbb{R}^n$ that contains it. These observations show that sets admitting least-majorized elements are not necessarily "nice" (i.e., polyhedral or convex) and thus, also not necessarily base polyhedra. Therefore, it seems unlikely that a characterization can be given of sets that have least-majorized elements.

These observations motivate us to introduce a more general notion of majorization that we call *(a, b)-majorization*. This concept can be interpreted as majorization relative to a positive scaling vector a and a shifting vector b and is a slight generalization of the concept of d-majorization in Veinott [47]. In terms of optimization, (a, b)-majorization is concerned with order-preserving objective functions of the form $\sum_{i\in N} a_i \phi\left(\frac{x_i+b_i}{a_i}\right)$, where ϕ is a continuous convex function. Such objective functions naturally occur in energy scheduling applications, such as power allocation for multichannel communication systems and speed scaling for computer processors (Schoot Uiterkamp et al. [42]) and price mechanism design (Reijnders et al. [41]). Analogously to "classical" majorization, we say that x is (a, b) majorized by y if $\sum_{i\in N} a_i \phi\left(\frac{x_i+b_i}{a_i}\right) \leq \sum_{i\in N} a_i \phi\left(\frac{y_i+b_i}{a_i}\right)$ for all continuous convex functions ϕ. Moreover, given a set $C \subseteq \mathbb{R}^n$, we say that $x^* \in C$ is a least (a, b)-majorized element of C if it is an optimal solution to the problem $\min_{x\in C} \sum_{i\in N} a_i \phi\left(\frac{x_i+b_i}{a_i}\right)$ for any choice of continuous convex function ϕ. Analogously to "classical" majorization, we also define similar generalizations of weak submajorization, weak supermajorization, and weak absolute majorization and of the corresponding least-majorized elements (see Section 3.2).

This generalization of majorization allows us to obtain a positive answer to the questions posed at the start of this subsection. For this, we build upon a result in earlier work (Schoot Uiterkamp et al. [42]) that shows the existence of least (a, b)-majorized elements for base polyhedra. As a first contribution, in the present article, we extend this result by establishing the existence of least weakly (a, b)-submajorized and (a, b)-supermajorized and least weakly absolutely (a, b)-majorized elements for several related sets, including submodular, supermodular, and bisubmodular polyhedra. Thereby, we generalize some results in Dutta and Ray [14] and Tamir [45] on the existence of least weakly sub- and supermajorized elements for special cases of these polyhedra. In fact, for the case of weak absolute majorization, no such existence results were known for these sets, even for the case of "classical" majorization (i.e., when $a = 1$ and $b = 0$), and we are the first to establish them.

As our second and main contribution, we show that the existence of least (a, b)-majorized elements for all choices of scaling and shifting vectors a and b is a unique property of base polyhedra within the class of compact convex sets. Hereby, we prove a reverse of the earlier-mentioned result in Schoot Uiterkamp et al. [42]. Moreover, still within the class of compact convex sets, we show that the existence of least weakly absolutely (a, b)-majorized elements for all choices of a and b is a unique property of bisubmodular polyhedra. Thus, we obtain new characterizations of base and bisubmodular polyhedra in terms of (weak absolute) (a, b)-majorization. Additionally, we completely characterize sets admitting least weakly (a, b)-submajorized or (a, b)-supermajorized elements for all choices of a and b as those that are contained in a super- or submodular polyhedron, respectively, and contain the corresponding base polyhedron.

For $a = 1$, all these results carry over to the integral versions of these sets (i.e., the collection of integral points in the set). Here, the class of compact convex sets is replaced by that of bounded integral hole-free sets (i.e., bounded sets that contain exactly all integral points in their convex hull). All characterization results for both set versions are summarized in Theorems 1–8.

The existence results are proven based on the earlier-mentioned result for base polyhedra in Schoot Uiterkamp et al. [42] and on the relation of the aforementioned sets with these polyhedra. For instance, the existence of least weakly (a, b)-supermajorized elements for submodular polyhedra can be proven using the fact that such polyhedra contain vectors that are in some sense maximal and thus, necessarily minimize any nonincreasing function over such polyhedra. These maximal vectors form a base polyhedron within the given submodular polyhedron. Furthermore, the existence of least weakly absolutely (a, b)-majorized elements for bisubmodular polyhedra uses the fact that such polyhedra are the intersection of (reflections of) particular submodular polyhedra and contain all corresponding (reflected) base polyhedra (see also Ando and Fujishige [2], Fujishige [22]).

The proof of the characterization results for base and bisubmodular polyhedra is inspired by classical characterizations for the optimality of greedy algorithms for linear optimization over these polyhedra (see Dunstan and Welsh [13], Edmonds [16], Nakamura [37]). The latter characterizations state that for each permutation of the index set, there exists a vector that simultaneously minimizes any linear cost function over these polyhedra whose coefficients are monotonically increasing or decreasing under this permutation. In this light, our characterization results can be seen as an extension of these classical results to (Schur-)convex objective functions.

We demonstrate the impact of our results in three different fields. First, we focus on RAPs in power management and in particular, on energy storage scheduling problems. In these problems, the utility or cost of an allocation depends on the actual allocation plus a fixed (uncontrollable) load, such as static energy consumption (see also van der Klauw et al. [46]). Our results show that RAPs over base polyhedra are the only convex (concave) RAPs where for any given fixed load, there exists an allocation that simultaneously optimizes any symmetric (quasi-)convex ((quasi-)concave) objective function of the combined load and allocation. This work, therefore, provides a theoretical explanation for the necessity of objective trade-offs for more general RAPs, even when those objectives do not conflict at first glance (e.g., when the considered objective functions are all symmetric utility functions). This motivates the necessity and development of multiobjective optimization algorithms for the operation of the corresponding systems.

The second application is in cooperative game theory. We obtain a new characterization of convex cooperative games with transferable utility (TU) in terms of the existence of egalitarian solutions. Although this characterization is simply a direct reformulation of one of our main results (Theorem 5) in game-theoretical terms, no similar characterizations have been proposed before to the best of our knowledge.

Finally, we focus on regularized regression problems and their corresponding optimal estimators. We show that for certain regression problems arising in orthogonal design experiments, the optimal regression estimators for least absolute shrinkage and selection operator (LASSO) and max-norm regularization are least weakly absolutely (a, b)-majorized elements, whereas the optimal estimator for ridge regression is not. These observations could be a stepping stone to demonstrate the optimality of these or related estimators for regression problems with more general loss functions.

Summarizing, our concrete contributions are as follows.

1. We introduce a natural generalization of majorization, called (a, b)-majorization, and investigate its properties and relation to "classical" majorization.

2. We completely characterize all compact convex sets and bounded integral hole-free sets that have least weakly (a, b)-submajorized or (a, b)-supermajorized or least (weakly absolutely) (a, b)-majorized elements for all pairs of scaling vectors a and shifting vectors b in terms of submodular, supermodular, base, and bisubmodular polyhedra.

3. We characterize base and bisubmodular polyhedra in terms of the existence of least (a, b) majorized and least weakly absolutely (a, b)-majorized elements for all pairs of scaling vectors a and shifting vectors b, respectively.

4. On the application side, we provide a theoretical explanation for the presence of conflicting utility or cost objectives in more general RAPs, give a new characterization of convex cooperative games, and establish new properties of a specific class of regression estimators.

The outline of the remainder of this article is as follows. In Section 2, we introduce the necessary concepts and results regarding (bi-)submodular functions and the related polyhedral, and in Section 3, we introduce (a, b)-majorization. In Section 4, we prove the existence of least (a, b)-majorized and related elements for submodular, supermodular, and bisubmodular polyhedra and related sets, and in Section 5, we derive a complete characterization of these elements in terms of these polyhedra. In Section 6, we demonstrate the impact of these characterization results in several applications, and in Section 7, we state our conclusions and directions for future research.

2. Preliminaries

In this section, we introduce some general notation and all concepts and known results on (bi-)submodularity that we require for this article.

2.1. General Notation

Throughout, we denote the index set by $N := \{1, \dots, n\}$ and the power set of N by $2^N := \{S | S \subseteq N\}$. Given a vector $x \in \mathbb{R}^n$ and a subset $S \in 2^N$, we denote the sum of all elements of x whose indices are in S by $x(S) := \sum_{i \in S} x_i$. We denote the convex hull of C by $\text{co}(C)$. For convenience, we say that a given vector, set, or function is *integral* if it is integer valued. We call an integral set $C \subseteq \mathbb{Z}^n$ *hole free* if it consists of all integral points in its convex hull (i.e., if $C = \text{co}(C) \cap \mathbb{Z}^n$). Moreover, we say that a function Φ is continuous and convex on C if there exists a continuous and convex function $\tilde{\Phi}$ on $\mathbb{R}^n$ that coincides with Φ on C. Given two vectors $x, y \in \mathbb{R}^n$, we write $x \leq y$ if $x_i \leq y_i$ for all $i \in N$ and $x < y$ if, in addition, we have $x_i < y_i$ for at least one index $i \in N$. For each $i \in N$, we denote by e^i the unit base vector of dimension n corresponding to i (i.e., $e^i_k = 1$ if $k = i$ and $e^i_k = 0$ if $k \neq i$).

2.2. Submodularity

A set function $f : 2^N \to \mathbb{R}^n$ is *submodular* if $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$ for all $A, B \in 2^N$, where we assume that $f(\emptyset) = 0$. Moreover, we say that a set function g is *supermodular* if $-g$ is submodular. Note that if f is submodular, the function $f^{\#}(A) := f(N) - f(N \backslash A)$ is supermodular, and we call $f^{\#}$ the *dual supermodular function* of f. Analogously, we call the submodular function $g^{\#}(A) := g(N) - g(N \backslash A)$ the *dual submodular function* of g.

Given a submodular function f, we denote by $P(f)$ and $B(f)$ the submodular and base polyhedron defined by f, which are, respectively, given by

$$P(f) := \{x \in \mathbb{R}^n | x(A) \leq f(A) \ \ \forall A \in 2^N\};$$
$$B(f) := \{x \in P(f) | x(N) = f(N)\}.$$

Analogously, given a supermodular function g, we denote by $P_{\text{sup}}(g)$ and $B_{\text{sup}}(g)$ the supermodular and base polyhedron defined by g, which are, respectively, given by

$$P_{\text{sup}}(g) := \{x \in \mathbb{R}^n | x(A) \geq g(A) \ \ \forall A \in 2^N\};$$
$$B_{\text{sup}}(g) := \{x \in P_{\text{sup}}(g) | x(N) = g(N)\}.$$

Note that $B(f) = B_{\text{sup}}(f^{\#})$ for any submodular function f because $f(N) = f^{\#}(N)$ and because for any $A \in 2^N$, $x \in B(f)$ implies $x(A) = f(N) - x(N \backslash A) \geq f(N) - f(N \backslash A) = f^{\#}(A)$ and $x \in B_{\text{sup}}(f^{\#})$ implies $x(A) = f^{\#}(N) - x(N \backslash A) \leq f^{\#}(N) - f^{\#}(N \backslash A) = f(A)$. Analogously, $B_{\text{sup}}(g) = B(g^{\#})$ for any supermodular function g.

We also define the *integral* submodular (supermodular) and base polyhedron defined by an integral submodular function f (integral supermodular function g) as the set of all integer-valued points in $P(f)$, $B(f)$, $P_{\text{sup}}(g)$, and $B_{\text{sup}}(g)$, respectively:

$$P^{\mathbb{Z}}(f) := P(f) \cap \mathbb{Z}^n; \qquad B^{\mathbb{Z}}(f) := B(f) \cap \mathbb{Z}^n;$$
$$P^{\mathbb{Z}}_{\text{sup}}(g) := P_{\text{sup}}(g) \cap \mathbb{Z}^n; \qquad B^{\mathbb{Z}}_{\text{sup}}(g) := B_{\text{sup}}(g) \cap \mathbb{Z}^n.$$

Note that all these sets are hole free.

Given a submodular function f and a vector $v \in \mathbb{R}^n$, we define the *reduction* of f by v as the set function

$$f^v(A) := \min(f(Z) + v(A \backslash Z) | Z \subseteq A), \quad A \in 2^N.$$

It follows from, for example, Fujishige [22, theorem 3.3] that also f^v is submodular and that

$$P(f^v) = \{x \in P(f) | x \leq v\}; \qquad P^{\mathbb{Z}}(f^v) = \{x \in P^{\mathbb{Z}}(f) | x \leq v\}.$$

Analogously, given a supermodular function g, the reduction of g by v is defined as

$$g^v(A) := \max(g(Z) + v(A \backslash Z) | Z \subseteq A), \quad A \in 2^N.$$

Moreover, g^v is supermodular and

$$P_{\text{sup}}(g^v) = \{x \in P_{\text{sup}}(g) | x \geq v\}; \qquad P^{\mathbb{Z}}_{\text{sup}}(g^v) = \{x \in P^{\mathbb{Z}}_{\text{sup}}(g) | x \geq v\}.$$

Lemma 1 shows that we can always increase (decrease) the entries of a vector in a submodular (supermodular) polyhedron in such a way that it becomes an element of the corresponding base polyhedron.

Lemma 1 (e.g., Fujishige [22, theorem 2.3]). *Let a (integral) submodular function f and a (integral) supermodular function g be given. Then, the following hold.*
- *For each $x \in P(f)$ ($x \in P^{\mathbb{Z}}(f)$), there exists $y \in B(f)$ ($y \in B^{\mathbb{Z}}(f)$) such that $x \leq y$.*
- *For each $x \in P_{\text{sup}}(g)$ ($x \in P^{\mathbb{Z}}_{\text{sup}}(g)$), there exists $y \in B_{\text{sup}}(g)$ ($y \in B^{\mathbb{Z}}_{\text{sup}}(g)$) such that $x \geq y$.*

Finally, in Lemma 2, we state the known result that the extreme points of a base polyhedron can be expressed analytically in terms of the submodular function defining the base polyhedron.

Lemma 2 (e.g., Fujishige and Tomizawa [23, theorem 3.4]). *For each extreme point z of a (integral) base polyhedron defined by a (integral) submodular function f, there exists a permutation π of N such that $z_{\pi(1)} = f(\{\pi(1)\})$ and $z_{\pi(k)} = f(\{\pi(1), \ldots, \pi(k)\}) - f(\{\pi(1), \ldots, \pi(k-1)\}$ for each $k \in \{2, \ldots, n\}$.*

For more background on submodular functions in general and the lemmas in particular, we refer to Fujishige [22].

2.3. Bisubmodularity

We denote the set of all ordered pairs (S, T) of disjoint subsets of N by $3^N := \{(S, T) | S, T \subseteq N, S \cap T = \emptyset\}$. Given an element $U = (S, T) \in 3^N$, we use the notation U^+ to denote S and U^- to denote T. Moreover, for any $X \subseteq N$ and vector $s \in \{-1, 1\}^n$, we use the notation $X|s$ to denote the biset $(\{i \in X | s_i = 1\}, \{i \in X | s_i = -1\})$. We use the symbols $\sqsubseteq$ and $\sqsupseteq$ to define the subset and superset relations on two elements of 3^N (i.e., for any two pairs $(S_1, T_1), (S_2, T_2) \in 3^N$, we write $(S_1, T_1) \sqsubseteq (S_2, T_2)$ if $S_1 \subseteq S_2$ and $T_1 \subseteq T_2$, and we have $(S_1, T_1) \sqsupseteq (S_2, T_2)$ if $S_1 \supseteq S_2$ and

$T_1 \supseteq T_2$). Moreover, we define the reduced union and intersection of (S_1, T_1) and (S_2, T_2) as

$$(S_1, T_1) \sqcup (S_2, T_2) := ((S_1 \cup S_2) \backslash (T_1 \cup T_2), (T_1 \cup T_2) \backslash (S_1 \cup S_2));$$
$$(S_1, T_1) \sqcap (S_2, T_2) := (S_1 \cap S_2, T_1 \cap T_2),$$

which can be seen as a generalization of the usual union and intersection operations, respectively.

A biset function $h : 3^N \to \mathbb{R}$ with $h(\emptyset, \emptyset) := 0$ is *bisubmodular* if for each two pairs $(S_1, T_1), (S_2, T_2) \in 3^N$, we have

$$h(S_1, T_1) + h(S_2, T_2) \geq h((S_1, T_1) \sqcup (S_2, T_2)) + h((S_1, T_1) \sqcap (S_2, T_2)).$$

Given a bisubmodular function h, the *bisubmodular polyhedron* defined by h is the set

$$\tilde{B}(h) := \{x \in \mathbb{R}^n | x(S) - x(T) \leq h(S, T) \ \forall (S, T) \in 3^N\}.$$

Analogously to submodular functions and polyhedra, we define the *integral* bisubmodular polyhedron defined by an integral bisubmodular function h as the set of all integral points in $\tilde{B}(h)$ (i.e., $\tilde{B}^{\mathbb{Z}}(h) := \tilde{B}(h) \cap \mathbb{Z}^n$). Note that $\tilde{B}^{\mathbb{Z}}(h)$ is hole free.

Given a pair $(S, T) \in 3^N$ and a bisubmodular function h, we denote by $2^{(S,T)}$ the set of all pairs (X, Y) with $(X, Y) \sqsubseteq (S, T)$, and we define the set function $h_{(S,T)}(X) := h(S \cap X, T \cap X)$. Note that $h_{(S,T)}$ is submodular because we have the following for any $X, Y \in 2^N$:

$$\begin{aligned} & h_{(S,T)}(X) + h_{(S,T)}(Y) \\ &= h(S \cap X, T \cap X) + h(S \cap Y, T \cap Y) \\ &\geq h((S \cap X, T \cap X) \sqcup (S \cap Y, T \cap Y)) + h((S \cap X, T \cap X) \sqcap (S \cap Y, T \cap Y)) \\ &= h(S \cap (X \cup Y), T \cap (X \cup Y)) + h(S \cap X \cap Y, T \cap X \cap Y) \\ &= h_{(S,T)}(X \cup Y) + h_{(S,T)}(X \cap Y). \end{aligned}$$

If $S \cup T = N$, we call (S, T) an *orthant* of $\mathbb{R}^n$. We define the submodular and base polyhedron in the orthant (S, T) as

$$P_{(S,T)}(h) := \{x \in \mathbb{R}^n | x(X) - x(Y) \leq h(X, Y) \ \forall (X, Y) \in 2^{(S,T)}\};$$
$$B_{(S,T)}(h) := \{x \in P_{(S,T)}(h) | x(S) - x(T) = h(S, T)\}.$$

Note that, by definition of $h_{(S,T)}$, we have for any orthant (S, T) of $\mathbb{R}^n$ that

$$P_{(S,T)}(h) := \left\{ x \in \mathbb{R}^n | \exists y \in P(h_{(S,T)}) \text{ such that } x_i = \begin{cases} y_i & \text{if } i \in S; \\ -y_i & \text{if } i \in T. \end{cases} \right\}$$

This implies for any $x \in \mathbb{R}^n$ that $x \in P_{(S,T)}(h)$ ($x \in B_{(S,T)}(h)$) if and only if $\tilde{x} \in P(h_{(S,T)})$ ($\tilde{x} \in B(h_{(S,T)})$), where $\tilde{x}_i = x_i$ if $i \in S$ and $\tilde{x}_i = -x_i$ if $i \in T$. Note that $\tilde{B}(h)$ is the intersection of all sets $P_{(S,T)}(h)$ for all orthants (S, T) of $\mathbb{R}^n$. Moreover, for any such orthant, the corresponding base polyhedron $B_{(S,T)}(h)$ is contained in $\tilde{B}(h)$ (see, e.g., Fujishige [22, lemma 3.60]).

Finally, analogously to base polyhedra and Lemma 2, the extreme points of a bisubmodular polyhedron can be expressed analytically in terms of the underlying bisubmodular function.

Lemma 3 (e.g., Ando and Fujishige [2, theorem 3.4]). *For each extreme point z of a bisubmodular polyhedron defined by a bisubmodular function h, there exists a permutation π of N and a sign vector $s \in \{-1, 1\}^n$ such that $z_{\pi(1)} = s_{\pi(1)} h(\{\pi(1)\}|s)$ and $z_{\pi(k)} = s_{\pi(k)}(h(\{\pi(1), \ldots, \pi(k)\}|s) - h(\{\pi(1), \ldots, \pi(k-1)\}|s))$ for each $k \in \{2, \ldots, n\}$.*

For more background on bisubmodular functions and polyhedra and on their discussed properties, we refer to Ando and Fujishige [2], Bouchet and Cunningham [8], and Fujishige [21].

3. Majorization and (*a*, *b*)-Majorization

In this section, we discuss the classical concept of majorization (Section 3.1) and introduce a more general version of this concept that we call *(a, b)-majorization* (Section 3.2).

3.1. Classical Majorization

Given a vector $x \in \mathbb{R}^n$, we denote the nonincreasing ordering of the elements in x by $x^{\downarrow}$ (so that $x_1^{\downarrow} \geq \cdots \geq x_n^{\downarrow}$) and the nondecreasing order by $x^{\uparrow}$ (so that $x_1^{\uparrow} \leq \cdots \leq x_n^{\uparrow}$). Given also another vector $y \in \mathbb{R}^n$, we say that x is

- weakly submajorized by y, denoted by $x \prec^{\downarrow} y$, if $\sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow}$ for all $k \in N$;
- weakly supermajorized by y, denoted by $x \prec^{\uparrow} y$, if $\sum_{i=1}^{k} x_i^{\uparrow} \geq \sum_{i=1}^{k} y_i^{\uparrow}$ for all $k \in N$; and

- majorized by y, denoted by $x \prec y$, if x is both weakly sub- and supermajorized by y or equivalently, $x \prec^{\downarrow} y$ and $x(N) = y(N)$.

Several characterizations of these majorization concepts exist. One characterization that is particularly useful in the context of optimization problems is in terms of objective values for Schur-convex functions. A function $\Phi : \mathbb{R}^n \to \mathbb{R}$ is said to be Schur convex if it preserves the order of majorization (i.e., if $\Phi(x) \leq \Phi(y)$ whenever $x \prec y$). The following well-known characterizations of majorization and Schur-convex functions will be useful in the remainder of this article.

Lemma 4. *Given vectors $x, y \in \mathbb{R}^n$, the following statements are equivalent*:
1. $x \prec^{\downarrow} y$;
2. $\Phi(x) \leq \Phi(y)$ *for all nondecreasing continuous Schur-convex functions* Φ;
3. $\sum_{i\in N}\phi(x_i) \leq \sum_{i\in N}\phi(y_i)$ *for all nondecreasing continuous convex functions* ϕ; *and*
4. $\sum_{i\in N}\max(0, \alpha + x_i) \leq \sum_{i\in N}\max(0, \alpha + y_i)$ *for all* $\alpha \in \mathbb{R}$.

Lemma 5. *Given vectors $x, y \in \mathbb{R}^n$, the following statements are equivalent*:
1. $x \prec^{\uparrow} y$;
2. $\Phi(x) \leq \Phi(y)$ *for all nonincreasing continuous Schur-convex functions* Φ;
3. $\sum_{i\in N}\phi(x_i) \leq \sum_{i\in N}\phi(y_i)$ *for all nonincreasing continuous convex functions* ϕ; *and*
4. $\sum_{i\in N}\max(0, \alpha - x_i) \leq \sum_{i\in N}\max(0, \alpha - y_i)$ *for all* $\alpha \in \mathbb{R}$.

Lemma 6. *Given vectors $x, y \in \mathbb{R}^n$, the following statements are equivalent*:
1. $x \prec y$;
2. $\Phi(x) \leq \Phi(y)$ *for all continuous Schur-convex functions* Φ;
3. $\sum_{i\in N}\phi(x_i) \leq \sum_{i\in N}\phi(y_i)$ *for all continuous convex functions* ϕ; *and*
4. $x(N) = y(N)$ *and* $\sum_{i\in N}\max(0, \alpha - x_i) \leq \sum_{i\in N}\max(0, \alpha - y_i)$ *for all* $\alpha \in \mathbb{R}$.

We refer to Marshall et al. [36, propositions 1.A.2, 1.A.8, and 4.B.1–4.B.4] for the proofs of these results and to Marshall et al. [36] in general for more background on majorization.

Alongside these well-known majorization concepts, we also consider the concept of weak *absolute* majorization. Given a vector $x \in \mathbb{R}^n$, we denote the nonincreasing ordering of *the absolute values of* the elements of x by x^{abs}. We say that x is *weakly absolutely* majorized by y, denoted by $x \prec^{\text{abs}} y$, if $\sum_{i=1}^k |x_i^{\text{abs}}| \leq \sum_{i=1}^k |y_i^{\text{abs}}|$ for all $k \in N$ (see also Eaton [15], Niezgoda [38], Zhan [48]). Although this majorization concept is less known than the three other concepts discussed, it occurs regularly as an illustrative example when studying majorization from the perspective of ordered groups (see, e.g., Francis and Wynn [18]).

Lemma 7 provides a characterization of weak absolute majorization in terms of a particular class of Schur-convex functions that we call *monotonically even*. This class consists of all Schur-convex functions Φ for which
- $\Phi(x_1, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_n) = \Phi(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$ for all $x \in \mathbb{R}^n$ and $i \in N$ (Φ is even);
- $\Phi(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \leq \Phi(x_1, \ldots, x_{i-1}, x_i - \epsilon, x_{i+1}, \ldots, x_n)$ for all $\epsilon > 0$ and $x \in \mathbb{R}^n$ with $x_i \leq 0$ (Φ is elementwise nonincreasing for negative inputs); and
- $\Phi(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \leq \Phi(x_1, \ldots, x_{i-1}, x_i + \epsilon, x_{i+1}, \ldots, x_n)$ for all $\epsilon > 0$ and $x \in \mathbb{R}^n$ with $x_i \geq 0$ (Φ is elementwise nondecreasing for positive inputs).

This definition directly leads to a characterization in terms of separable *even* convex objective functions (i.e., functions of the form $\sum_{i\in N}\phi(x_i)$, where each ϕ is convex and $\phi(-\zeta) = \phi(\zeta)$ for all $\zeta \in \mathbb{R}$ (note that evenness and convexity of ϕ directly ensure that ϕ is also monotonically even)). This characterization and its proof are analogous to those corresponding to Lemmas 4–6.

Lemma 7. *Given vectors $x, y \in \mathbb{R}^n$, the following statements are equivalent*:
1. $x \prec^{\text{abs}} y$;
2. $\Phi(x) \leq \Phi(y)$ *for all monotonically even continuous Schur-convex functions* Φ;
3. $\sum_{i\in N}\phi(x_i) \leq \sum_{i\in N}\phi(y_i)$ *for all even continuous convex functions* ϕ; *and*
4. $\sum_{i\in N}\max(\alpha - x_i, 0, \alpha + x_i) \leq \sum_{i\in N}\max(\alpha - y_i, 0, \alpha + y_i)$ *for all* $\alpha \in \mathbb{R}$.

Proof of (1) ⇒ (2). Given a monotonically even continuous Schur-convex function Φ, consider the nondecreasing Schur-convex function

$$\tilde{\Phi}(x) = \begin{cases} \Phi(0) & \text{if } x \not\geq 0, \\ \Phi(x) & \text{if } x \geq 0. \end{cases}$$

Let $\tilde{x} := (|x_i|)_{i\in N}$ and $\tilde{y} := (|y_i|)_{i\in N}$. If $x \prec^{\text{abs}} y$, then $\tilde{x} \prec^{\downarrow} \tilde{y}$, and thus, $\tilde{\Phi}(\tilde{x}) \leq \tilde{\Phi}(\tilde{y})$. Because $\tilde{x}$ and $\tilde{y}$ are nonnegative, it follows that $\Phi(\tilde{x}) \leq \Phi(\tilde{y})$ and because Φ is even, that $\Phi(x) \leq \Phi(y)$. □

Proof of (2) ⇒ (3). Given a continuous convex function ϕ, the function $\sum_{i\in N}\phi(x_i)$ is Schur convex. If ϕ is also even, then $\sum_{i\in N}\phi(x_i)$ is Schur convex and monotonically even, and thus, (3) is a special case of (2). □

Proof of (3) ⇒ (4). Note that for any $\alpha \in \mathbb{R}$, the function $\max(\alpha - z, 0, \alpha + z)$ is even and convex. Thus, (4) is a special case of (3). □

Proof of (4) ⇒ (1). We define for a given $k \in N$ the even convex function $\tilde{\phi}^k(z) := \max(-z - |y_k^{\text{abs}}|, 0, z - |y_k^{\text{abs}}|)$. Then, we have $\sum_{i=1}^k \tilde{\phi}^k(y_i^{\text{abs}}) = \sum_{i=1}^k |y_i^{\text{abs}}| - k|y_k^{\text{abs}}|$ and $\sum_{i=k+1}^n \tilde{\phi}^k(y_i^{\text{abs}}) = 0$. Note that, by assumption, we have $\sum_{i\in N}\tilde{\phi}^k(x_i) \leq \sum_{i\in N}\tilde{\phi}^k(y_i)$. Moreover, for any $z \in \mathbb{R}$, we have $\tilde{\phi}^k(z) \geq 0$ and $\tilde{\phi}^k(z) \geq |z| - |y_k^{\text{abs}}|$. It follows that

$$\sum_{i=1}^k |y_i^{\text{abs}}| - k|y_k^{\text{abs}}| = \sum_{i\in N}\tilde{\phi}^k(y_i) \geq \sum_{i\in N}\tilde{\phi}^k(x_i)$$
$$\geq \sum_{i=1}^k \tilde{\phi}^k(x_i^{\text{abs}}) \geq \sum_{i=1}^k |x_i^{\text{abs}}| - k|y_k^{\text{abs}}|$$

which implies that $\sum_{i=1}^k |x_i^{\text{abs}}| \leq \sum_{i=1}^k |y_i^{\text{abs}}|$. Because this holds for all $k \in N$, it follows that $x \prec^{\text{abs}} y$. □

3.2. *(a, b)*-Majorization

We introduce a more general version of all four majorization concepts discussed in Section 3.1. This generalization, which we call *(a, b)-majorization*, can be interpreted as majorization relative to scaling by a (positive) vector a and shifting by a vector b. More precisely, for fixed scaling and shifting vectors $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$ and given two vectors $x, y \in \mathbb{R}^n$, we say that x is

- (a, b)-majorized by y, denoted by $x \prec_{(a,b)} y$, if $x(N) = y(N)$ and for all continuous convex functions ϕ, we have $\sum_{i\in N} a_i\phi\left(\frac{x_i+b_i}{a_i}\right) \leq \sum_{i\in N} a_i\phi\left(\frac{y_i+b_i}{a_i}\right)$;
- weakly (a, b)-submajorized by y, denoted by $x \prec^{\downarrow}_{(a,b)} y$, if for all nondecreasing continuous convex functions ϕ, we have $\sum_{i\in N} a_i\phi\left(\frac{x_i+b_i}{a_i}\right) \leq \sum_{i\in N} a_i\phi\left(\frac{y_i+b_i}{a_i}\right)$;
- weakly (a, b)-supermajorized by y, denoted by $x \prec^{\uparrow}_{(a,b)} y$, if for all nonincreasing continuous convex functions ϕ, we have $\sum_{i\in N} a_i\phi\left(\frac{x_i+b_i}{a_i}\right) \leq \sum_{i\in N} a_i\phi\left(\frac{y_i+b_i}{a_i}\right)$; and
- weakly absolutely (a, b)-majorized by y, denoted by $x \prec^{\text{abs}}_{(a,b)} y$, if for all even continuous convex functions ϕ, we have $\sum_{i\in N} a_i\phi\left(\frac{x_i+b_i}{a_i}\right) \leq \sum_{i\in N} a_i\phi\left(\frac{y_i+b_i}{a_i}\right)$.

Moreover, given a set $C \subseteq \mathbb{R}^n$, a vector $x \in C$ is a

- *least (a, b)-majorized* element of C if $x \prec_{(a,b)} y$ for all $y \in C$;
- *least weakly (a, b)-submajorized* element of C if $x \prec^{\downarrow}_{(a,b)} y$ for all $y \in C$;
- *least weakly (a, b)-supermajorized* element of C if $x \prec^{\uparrow}_{(a,b)} y$ for all $y \in C$; and
- *least weakly absolutely (a, b)-majorized* element of C if $x \prec^{\text{abs}}_{(a,b)} y$ for all $y \in C$.

For convenience, if the specific type of majorized element is clear from the context or not relevant, we call all these four types of elements least (a, b)-majorized elements. If the vectors a and b are also not specified, we call them least-majorized elements.

The concept of (a, b)-majorization is a slight generalization of the earlier-mentioned concept of d-majorization (Veinott [47]) (i.e., d-majorization is equivalent to $(d, 0)$-majorization). Although the inclusion of a shift vector b may not seem as a significant generalization, it is a crucial factor when proving our main characterization results in Section 5 with regard to submodular, supermodular, base, and bisubmodular polyhedra.

Note that, as opposed to the "classical" majorization concepts of Section 3.1, a definition of (a, b)-majorization cannot be given in terms of partial sums of nonincreasing, nondecreasing, or absolute orders. Similar generalizations of majorization, such as d-majorization (Veinott [47]), are introduced via an alternative definition of majorization involving doubly stochastic matrices (see also Marshall et al. [36, chapter 2 and section 14.B]) or in terms of optimality for particular classes of separable convex objective functions (see, e.g., Joe [33]). Because we focus in this article on optimization problems, we follow the latter approach to define and interpret (a, b)-majorization.

Despite this discrepancy between "classical" majorization and (a, b)-majorization, Lemmas 4–7 can be partially generalized to the case of (a, b)-majorization. More precisely, to prove that a vector $x \in \mathbb{R}^n$ is (a, b)-majorized by another vector $y \in \mathbb{R}^n$ in one of the four senses, it suffices to check the majorization property for only a specific class of max-functions.

Lemma 8. *Given $x, y \in \mathbb{R}^n$, the following hold:*
1. $x \prec^{\downarrow}_{(a,b)} y$ *if and only if* $\sum_{i \in N} a_i \max\left(0, \alpha + \frac{x_i + b_i}{a_i}\right) \leq \sum_{i \in N} a_i \max\left(0, \alpha + \frac{y_i + b_i}{a_i}\right)$ *for all* $\alpha \in \mathbb{R}$;
2. $x \prec^{\uparrow}_{(a,b)} y$ *if and only if* $\sum_{i \in N} a_i \max\left(0, \alpha - \frac{x_i + b_i}{a_i}\right) \leq \sum_{i \in N} a_i \max\left(0, \alpha - \frac{y_i + b_i}{a_i}\right)$ *for all* $\alpha \in \mathbb{R}$;
3. $x \prec_{(a,b)} y$ *if and only if* $x \prec^{\downarrow}_{(a,b)} y$ *and* $x \prec^{\uparrow}_{(a,b)} y$; *and*
4. $x \prec^{\text{abs}}_{(a,b)} y$ *if and only if* $\sum_{i \in N} a_i \max\left(\alpha - \frac{x_i + b_i}{a_i}, 0, \alpha + \frac{x_i + b_i}{a_i}\right) \leq \sum_{i \in N} a_i \max\left(\alpha - \frac{y_i + b_i}{a_i}, 0, \alpha + \frac{y_i + b_i}{a_i}\right)$ *for all* $\alpha \in \mathbb{R}$.

Proof of (1). The "only if" part follows directly from the definition of weak (a, b)-submajorization because $\max(0, \alpha + \zeta)$ is nondecreasing and convex in ζ. Thus, we proceed to prove the "if" part. Given a nondecreasing continuous convex function ϕ, consider the piecewise linear approximation $\hat{\phi}$ of ϕ whose breakpoints are in the set $A(x, y) := \left\{\frac{x_1 + b_1}{a_1}, \ldots, \frac{x_n + b_n}{a_n}, \frac{y_1 + b_1}{a_1}, \ldots, \frac{y_n + b_n}{a_n}\right\}$ and that agrees with ϕ on these breakpoints (i.e., $\hat{\phi}(z) = \phi(z)$ for all $z \in A(x, y)$). Moreover, let $\delta' := \min(A(x, y))$. Observe that for any $\zeta \in \mathbb{R}$, $\hat{\phi}(\zeta)$ can be written as $\phi(\delta') + \sum_{\delta \in A(x,y)} \beta_\delta \max(0, \alpha_\delta + \zeta)$ for some values $\alpha_\delta \in \mathbb{R}$ and $\beta_\delta \in \mathbb{R}_{>0}$. Thus, we have

$$\begin{aligned}
\sum_{i \in N} a_i \phi\left(\frac{x_i + b_i}{a_i}\right) &= \sum_{i \in N} a_i \left(\phi(\delta') + \sum_{\delta \in A(x,y)} \beta_\delta \max\left(0, \alpha_\delta + \frac{x_i + b_i}{a_i}\right)\right) \\
&\leq \sum_{i \in N} a_i \left(\phi(\delta') + \sum_{\delta \in A(x,y)} \beta_\delta \max\left(0, \alpha_\delta + \frac{y_i + b_i}{a_i}\right)\right) \\
&= \sum_{i \in N} a_i \phi\left(\frac{y_i + b_i}{a_i}\right).
\end{aligned}$$

It follows that $x \prec^{\downarrow}_{(a,b)} y$. □

Proof of (2). This proof is analogous to the proof of (1). □

Proof of (3). This proof is analogous to the proof of (1), where we instead use the observation that the piecewise linear approximation of any continuous convex function ϕ with breakpoints in set $A(x, y)$ can be written as $\phi(\delta') + \sum_{\delta \in A(x,y); \delta \geq \delta'} \beta_\delta \max(0, \alpha_\delta - z) + \sum_{\delta \in A(x,y); \delta \leq \delta'} \beta_\delta \max(0, \alpha_\delta + z)$ for some values $\alpha_\delta \in \mathbb{R}$ and $\beta_\delta \in \mathbb{R}_{>0}$, where $\delta' := \arg\min_{\delta \in A(x,y)} \hat{\phi}(\delta)$. Furthermore, letting $\alpha' := \max_{\delta \in A(x,y)} |\delta|$, the fact that $x \prec^{\downarrow}_{(a,b)} y$ implies that

$$\begin{aligned}
a(N)\alpha' + x(N) + b(N) = \sum_{i \in N} a_i \max\left(0, \alpha' + \frac{x_i + b_i}{a_i}\right) &\leq \sum_{i \in N} a_i \max\left(0, \alpha' + \frac{y_i + b_i}{a_i}\right) \\
&= a(N)\alpha' + y(N) + b(N).
\end{aligned}$$

It follows that $x(N) \leq y(N)$. Analogously, $x \prec^{\uparrow}_{(a,b)} y$ implies that $x(N) \geq y(N)$, and thus, we have $x(N) = y(N)$. □

Proof of (4). This proof is analogous to the proof of (1), where we instead use the observation that the piecewise linear approximation of any even convex function ϕ with breakpoints in the set $A(x, y)$ can be written as $\phi(\delta') + \sum_{\delta \in A(x,y)} \beta_\delta \max(\alpha_\delta - z, 0, \alpha_\delta + z)$ for some values $\alpha_\delta \in \mathbb{R}$ and $\beta_\delta \in \mathbb{R}_{>0}$, where $\delta' := \arg\min_{\delta \in A(x,y)} \hat{\phi}(\delta)$. □

Finally, we note that for any $b \in \mathbb{R}^n$, a least $(1, b)$-majorized element of some set $C \subseteq \mathbb{R}^n$ also minimizes $\Phi(x + b)$ over C for any choice of continuous Schur-convex function Φ. Analogously, a least weakly $(1, b)$-submajorized or $(1, b)$-supermajorized element or a least weakly absolutely $(1, b)$-majorized element of C minimizes $\Phi(x + b)$ for any choice of nondecreasing, nonincreasing, or monotonically even Schur-convex function Φ, respectively.

4. Identifying Sets with Least (a, b)-Majorized Elements

In this section, we establish the existence of least (a, b)-majorized, least weakly (a, b)-submajorized and (a, b)-supermajorized, and least weakly absolutely (a, b)-majorized elements for several sets that are in some way related to (bi-)submodular functions and the related polyhedra. The starting point for our investigations is a recently proved result in Schoot Uiterkamp et al. [42] that establishes the existence of least (a, b)-majorized elements for base polyhedra and of least $(1, b)$-majorized elements for integral base polyhedral.

Lemma 9 (Schoot Uiterkamp et al. [42, condition 1 and theorem 1]). *Let f be a submodular function. For each pair of vectors $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$, the base polyhedron $B(f)$ has a unique least (a, b)-majorized element, namely the optimal solution to the problem $\min_{x \in B(f)} \sum_{i \in N} \frac{1}{2} \frac{(x_i + b_i)^2}{a_i}$. Moreover, if f and b are integral, then any optimal solution to the problem $\min_{x \in B^{\mathbb{Z}}(f)} \sum_{i \in N} \frac{1}{2}(x_i + b_i)^2$ is a least $(1, b)$-majorized element of the integral base polyhedron $B^{\mathbb{Z}}(f)$.*

A natural question is whether base polyhedra are the only sets containing least (a, b)-majorized elements. The answer to this question is no; any least (a, b)-majorized element of a given base polyhedron is also a least (a, b)-majorized element of any subset of the base polyhedron that contains this element. Thus, the existence of least (a, b)-majorized elements is not limited to "nice" sets, such as polyhedra and convex sets. However, we do show in Section 5.3 that base polyhedra are the only compact convex subsets of $\mathbb{R}^n$ that have least (a, b)-majorized elements for *all* pairs (a, b).

Lemma 9 forms the basis for all existence results in the remainder of this section. We present our existence results for least weakly (a, b)-submajorized and (a, b)-supermajorized elements in Section 4.1 and those for least weakly absolutely (a, b)-majorized elements in Section 4.2.

4.1. Existence of Least Weakly (*a*, *b*)-Submajorized and (*a*, *b*)-Supermajorized Elements

In this subsection, we focus on the existence of least weakly (a, b)-submajorized and (a, b)-supermajorized elements. First, we use the initial existence result in Lemma 9 to prove that submodular and supermodular polyhedra have least weakly (a, b)-supermajorized and (a, b)-submajorized elements, respectively.

Lemma 10. *Let a submodular function f and vectors $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$ be given.*

- *Any least (a, b)-majorized element of $B(f)$ is also a least weakly (a, b)-supermajorized element of $P(f)$ and a least weakly (a, b)-submajorized element of $P_{\sup}(f^{\#})$.*
- *If f and b are integral, then any least $(1, b)$-majorized element of $B^{\mathbb{Z}}(f)$ is also a least weakly $(1, b)$-supermajorized element of $P^{\mathbb{Z}}(f)$ and a least weakly $(1, b)$-submajorized element of $P^{\mathbb{Z}}_{\sup}(f^{\#})$.*

Proof. We prove the first statement for the case of $P(f)$ (the proofs for the cases of $P_{\sup}(f^{\#})$ and the second statement are analogous). Let x^* be a least (a, b)-majorized element of $B(f)$, which exists because of Lemma 9. Given a vector $z \in P(f)$, it follows from Lemma 1 that there exists a vector $y \in B(f)$ with $z \leq y$. Then, we have for any nonincreasing continuous convex function ϕ that $\sum_{i \in N} a_i \phi\left(\frac{x^*_i + b_i}{a_i}\right) \leq \sum_{i \in N} a_i \phi\left(\frac{y_i + b_i}{a_i}\right) \leq \sum_{i \in N} a_i \phi\left(\frac{z_i + b_i}{a_i}\right)$, where the first inequality follows because x^* is a least (a, b)-majorized element of $B(f)$ and the second inequality follows because $z \leq y$ and ϕ is nonincreasing. Because $z \in P(f)$ was chosen arbitrarily, it follows that x^* is a least weakly (a, b)-supermajorized element of $P(f)$. □

It follows directly from Lemma 10 that any set that contains a least (a, b)-majorized element of a base polyhedron and that is contained in the corresponding submodular or supermodular polyhedron has a least weakly (a, b)-supermajorized or (a, b)-submajorized element, respectively.

Corollary 1. *Let a submodular function f and vectors $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$ be given. Any least (a, b)-majorized element x^* of $B(f)$ is also*

- *a least weakly (a, b)-supermajorized element of any subset of $P(f)$ that contains x^* and*
- *a least weakly (a, b)-submajorized element of any subset of $P_{\sup}(f^{\#})$ that contains x^*.*

Moreover, if f and b are integral, then any least $(1, b)$-majorized element x^* of $B^{\mathbb{Z}}(f)$ is also

- *a least weakly $(1, b)$-supermajorized element of any subset of $P^{\mathbb{Z}}(f)$ that contains x^* and*
- *a least weakly $(1, b)$-submajorized element of any subset of $P^{\mathbb{Z}}_{\sup}(f^{\#})$ that contains x^*.*

Analogously to Lemma 9, Corollary 1 implies that sets containing least weakly (a, b)-submajorized or (a, b)-supermajorized elements need not be "nice."

Corollary 1 allows us to prove the existence of least weakly (a, b)-submajorized and (a, b)-supermajorized elements for sets that are extensions or generalizations of submodular and supermodular polyhedra. In particular, this applies to bisubmodular polyhedral.

Lemma 11. *Let a bisubmodular function h and vectors $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$ be given. Then, the following hold.*

- *Any least (a, b)-majorized element of the base polyhedron $B(h_{(N,\emptyset)})$ is also a least weakly (a, b)-supermajorized element of $\tilde{B}(h)$.*

- *Any least* (a, b)*-majorized element of the base polyhedron* $B_{\sup}(-h_{(\emptyset,N)})$ *is also a least weakly* (a, b)*-submajorized element of* $\tilde{B}(h)$.

Moreover, if h and b are integral, then also the following hold.

- *Any least* $(1, b)$*-majorized element of the integral base polyhedron* $B^{\mathbb{Z}}(h_{(N,\emptyset)})$ *is also a least weakly* $(1, b)$*-supermajorized element of* $\tilde{B}^{\mathbb{Z}}(h)$.
- *Any least* $(1, b)$*-majorized element of the integral base polyhedron* $B^{\mathbb{Z}}_{\sup}(-h_{(\emptyset,N)})$ *is also a least weakly* $(1, b)$*-submajorized element of* $\tilde{B}^{\mathbb{Z}}(h)$.

Proof. We prove that any least (a, b)-majorized element of $B(h_{(N,\emptyset)})$ is also a least weakly (a, b)-supermajorized element of $\tilde{B}(h)$ (the proofs of the other three statements are analogous). Recall from Section 2.3 that because $(N,\emptyset)$ is an orthant, we have $B_{(N,\emptyset)}(h) \subseteq \tilde{B}(h) \subseteq P_{(N,\emptyset)}(h)$. Also, note that for the specific orthant $(N,\emptyset)$, we have $P_{(N,\emptyset)}(h) = P(h_{(N,\emptyset)})$ and $B_{(N,\emptyset)}(h) = B(h_{(N,\emptyset)})$. Thus, $\tilde{B}(h)$ is a subset of the submodular polyhedron $P(h_{(N,\emptyset)})$ and contains the corresponding base polyhedron $B(h_{(N,\emptyset)})$. Now, the result of the lemma follows from Corollary 1. □

Alternatively, for integral bisubmodular polyhedra, Lemma 11 can be proven by recognizing such polyhedra as special cases of jump systems (Bouchet and Cunningham [8]) and slightly adjusting Ando [1, proof of theorem 2.1], which shows that jump systems have both a least weakly (1, 0)-submajorized element and a least weakly (1, 0)-supermajorized element.

Lemma 11 also implies that all special cases of bisubmodular polyhedra have both a least weakly (a, b)-submajorized and (a, b)-supermajorized element. One example of such a special case is (bounded) generalized polymatroids (see, e.g., Bouchet and Cunningham [8]). In particular, Lemma 11 generalizes Tamir [45, corollary 3.3], where it is proven that bounded generalized polymatroids have both a least weakly (1, 0)-submajorized element and a least weakly (1, 0)-supermajorized element.

4.2. Existence of Least Weakly Absolutely (*a*, *b*)-Majorized Elements

In this subsection, we prove the existence of least weakly absolutely (a, b)-majorized elements for submodular and supermodular polyhedra (Lemma 12) and bisubmodular polyhedra (Lemma 13).

Lemma 12. *Let a submodular function f, a supermodular function g, and vectors* $a \in \mathbb{R}^n_{>0}$ *and* $b \in \mathbb{R}^n$ *be given. Then, the following hold.*

- *Any least weakly* (a, b)*-supermajorized element of* $P(f^{-b})$ *is also a least weakly absolutely* (a, b)*-majorized element of* $P(f)$.
- *Any least weakly* (a, b)*-submajorized element of* $P_{\sup}(g^{-b})$ *is also a least weakly absolutely* (a, b)*-majorized element of* $P_{\sup}(g)$.

Moreover, if f, g, and b are integral, then the following also hold.

- *Any least weakly* $(1, b)$*-supermajorized element of* $P^{\mathbb{Z}}(f^{-b})$ *is also a least weakly absolutely* $(1, b)$*-majorized element of* $P^{\mathbb{Z}}(f)$.
- *Any least weakly* $(1, b)$*-submajorized element of* $P^{\mathbb{Z}}_{\sup}(g^{-b})$ *is also a least weakly absolutely* $(1, b)$*-majorized element of* $P^{\mathbb{Z}}_{\sup}(g)$.

Proof. We prove that any least weakly (a, b)-supermajorized element of $P(f^{-b})$ is also a least weakly absolutely (a, b)-majorized element of $P(f)$ (the proofs of the other three statements are analogous). Consider the reduction f^{-b}, and let x^* be a least weakly (a, b)-supermajorized element of $P(f^{-b})$ (recall that f^{-b} is submodular and that x^* exists by Lemma 10). Note that, by definition of $P(f^{-b})$, we have $x^* \leq -b$ and thus, $\frac{x^*_i+b_i}{a_i} \leq 0$ for all $i \in N$.

Let a vector $x \in P(f)$ be given, and define the vector $\tilde{x} \in \mathbb{R}^n$ as $\tilde{x}_i := \min(x_i, -b_i)$ for $i \in N$. Note that $\tilde{x} \in P(f)$ because $\tilde{x} \leq x$ and $x \in P(f)$. It follows that $\tilde{x} \in P(f^{-b})$ because $\tilde{x} \leq -b$. Moreover, we have $\frac{\tilde{x}_i+b_i}{a_i} \leq 0$. Let an even continuous convex function ϕ be given, and let

$$\tilde{\phi}(\zeta) := \begin{cases} \phi(\zeta) & \text{if } \zeta \leq 0; \\ \phi(0) & \text{if } \zeta \geq 0. \end{cases}$$

We can now derive the following:

$$\sum_{i \in N} a_i \phi\left(\frac{x^*_i + b_i}{a_i}\right) = \sum_{i \in N} a_i \tilde{\phi}\left(\frac{x^*_i + b_i}{a_i}\right) \tag{1a}$$

$$\leq \sum_{i \in N} a_i \tilde{\phi}\left(\frac{\tilde{x}_i + b_i}{a_i}\right) \tag{1b}$$

$$= \sum_{i \in N} a_i \phi\left(\frac{\tilde{x}_i + b_i}{a_i}\right)$$
$$= \sum_{i:\tilde{x}_i = x_i} a_i \phi\left(\frac{x_i + b_i}{a_i}\right) + \sum_{i:\tilde{x}_i = -b_i < x_i} a_i \phi(0) \leq \sum_{i \in N} a_i \phi\left(\frac{x_i + b_i}{a_i}\right), \tag{1c}$$

where (1a) follows because $\frac{x_i^* + b_i}{a_i} \leq 0$ for all $i \in N$; (1b) follows because x^* is a least weakly (a, b)-supermajorized element of $P(f^{-b})$, $\tilde{\phi}$ is nonincreasing, and $\tilde{x} \in P(f^{-b})$; and (1c) follows because $\frac{\tilde{x}_i + b_i}{a_i} \leq 0$ for all $i \in N$. Because both x and ϕ were chosen arbitrarily, it follows that x^* is a least weakly absolutely (a, b)-majorized element of $P(f)$. □

Note that contrary to least weakly (a, b)-submajorized and (a, b)-supermajorized elements, a least weakly absolutely (a, b)-majorized element of a sub- or supermodular polyhedron does not necessarily lie in the corresponding base polyhedron. This happens, for instance, when the base polyhedron is disjoint from the reduced sub- or supermodular polyhedron that contains the least weakly absolutely (a, b)-majorized element as specified in Lemma 12. In particular, this occurs when $-b_i < \min_{x \in B(f)} x_i$ for some $i \in N$ because in that case, we have for all $x \in B(f)$ that $x_i > -b_i$, and thus, $x \notin P(f^{-b})$.

We are now ready to prove that bisubmodular polyhedra also have least weakly absolutely (a, b)-majorized elements.

Lemma 13. *Let a bisubmodular function h and vectors $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$ be given. Then, the following hold.*

- *Any solution to the problem $\min_{x \in \tilde{B}(h)} \sum_{i \in N} \frac{1}{2} \frac{(x_i + b_i)^2}{a_i}$ is a least weakly absolutely (a, b)-majorized element of $\tilde{B}(h)$.*
- *If h and b are integral, then any solution to the problem $\min_{x \in \tilde{B}^{\mathbb{Z}}(h)} \sum_{i \in N} \frac{1}{2}(x_i + b_i)^2$ is a least weakly absolutely $(1, b)$-majorized element of $\tilde{B}^{\mathbb{Z}}(h)$.*

Proof. We prove the lemma for the case of bisubmodular polyhedra (the proof for the case of integral bisubmodular polyhedra is analogous). Let h be a bisubmodular function. If $-b \in \tilde{B}(h)$, then $-b$ is the unique least weakly absolutely (a, b)-majorized element of $\tilde{B}(h)$ because for any $x \in \mathbb{R}^n$ and even continuous convex function ϕ, we have $\sum_{i \in N} a_i \phi\left(\frac{x_i + b_i}{a_i}\right) \geq \sum_{i \in N} a_i \phi(0) = \sum_{i \in N} a_i \phi\left(\frac{-b_i + b_i}{a_i}\right)$. Thus, suppose that $-b \notin \tilde{B}(h)$, and let x^* be an optimal solution to the problem $\min_{x \in \tilde{B}(h)} \sum_{i \in N} \frac{1}{2} \frac{(x_i + b_i)^2}{a_i}$. We define the following partition of the index set N based on the difference between x^* and $-b$:

$$A_+ := \{i \in N | x_i^* < -b_i\}; \quad A_- := \{i \in N | x_i^* > -b_i\}; \quad A_0 := \{i \in N | x_i^* = -b_i\}.$$

It can be shown by means of an exchange argument that there exists a pair $(\tilde{X}, \tilde{Y}) \in 3^N$ such that $(\tilde{X}, \tilde{Y}) \sqsupseteq (A_+, A_-)$ and $x^*(\tilde{X}) - x^*(\tilde{Y}) = h(\tilde{X}, \tilde{Y})$ (see, e.g., Fujishige [21, theorem 3.1]). Let (S, T) be an orthant that contains $(\tilde{X}, \tilde{Y})$. Finally, we introduce the following notation; for a given vector $v \in \mathbb{R}^n$, let $\tilde{v}$ be a corresponding vector with $\tilde{v}_i = -v_i$ if $i \in \tilde{Y}$ and $\tilde{v}_i = v_i$ otherwise.

The proof consists of two parts that we separately prove, namely showing that (i) $\tilde{x}^* \in B((h_{(S,T)})^{-\tilde{b}})$ and that (ii) for any vector $x \in \mathbb{R}^n$, $\tilde{x} \in B((h_{(S,T)})^{-\tilde{b}})$ implies $x \in \tilde{B}(h)$. By definition of x^*, we have for any $x \in \tilde{B}(h)$, that $\sum_{i \in N} \frac{1}{2} \frac{(x_i^* + b_i)^2}{a_i} \leq \sum_{i \in N} \frac{1}{2} \frac{(x_i + b_i)^2}{a_i}$ or equivalently, $\sum_{i \in N} \frac{1}{2} \frac{(\tilde{x}_i^* + \tilde{b}_i)^2}{a_i} \leq \sum_{i \in N} \frac{1}{2} \frac{(\tilde{x}_i + \tilde{b}_i)^2}{a_i}$. In particular, the latter holds for all $\tilde{x} \in B((h_{(S,T)})^{-\tilde{b}})$ by (ii). Because $B((h_{(S,T)})^{-\tilde{b}})$ is a base polyhedron containing $\tilde{x}^*$ (by (i)), this implies via Lemma 9 that $\tilde{x}^*$ is a least $(a, \tilde{b})$-majorized element of $B((h_{(S,T)})^{-\tilde{b}})$. It follows from Lemma 10 that $\tilde{x}^*$ is a least weakly $(a, \tilde{b})$-supermajorized element of $P((h_{(S,T)})^{-\tilde{b}})$ and subsequently from Lemma 12 that $\tilde{x}^*$ is a least weakly absolutely $(a, \tilde{b})$-majorized element of $P(h_{(S,T)})$. Thus, for any vector $\tilde{x} \in P(h_{(S,T)})$, we have $\sum_{i \in N} a_i \phi\left(\frac{\tilde{x}_i^* + \tilde{b}_i}{a_i}\right) \leq \sum_{i \in N} a_i \phi\left(\frac{\tilde{x}_i + \tilde{b}_i}{a_i}\right)$ for any even continuous convex function ϕ. Recall from Section 2.3 that $\tilde{x} \in P(h_{(S,T)})$ if and only if $x \in P_{(S,T)}(h)$. This implies that $\sum_{i \in N} a_i \phi\left(\frac{x_i^* + b_i}{a_i}\right) \leq \sum_{i \in N} a_i \phi\left(\frac{x_i + b_i}{a_i}\right)$ for all $x \in P_{(S,T)}(h)$ and any even continuous convex function ϕ, where we explicitly use the fact that ϕ is even. Thus, x^* is a least weakly absolutely (a, b)-majorized element of $P_{(S,T)}(h)$ and also of $\tilde{B}(h)$ because $\tilde{B}(h) \subseteq P_{(S,T)}(h)$.

For the first part, note that $\tilde{x}^* \leq -\tilde{b}$. Also, we have $\tilde{x}^* \in P(h_{(S,T)})$. Together, this means that $\tilde{x}^* \in P((h_{(S,T)})^{-\tilde{b}})$. Furthermore, we have that

$$\tilde{x}^*(N) \leq (h_{(S,T)})^{-\tilde{b}}(N) = \min(h_{(S,T)}(Z) - \tilde{b}(N \backslash Z) | Z \subseteq N) \leq h_{(S,T)}(\tilde{X} \cup \tilde{Y}) - \tilde{b}(N \backslash (\tilde{X} \cup \tilde{Y}))$$
$$= h(\tilde{X}, \tilde{Y}) - b(N \backslash (\tilde{X} \cup \tilde{Y})) = x^*(\tilde{X}) - x^*(\tilde{Y}) + x^*(N \backslash (\tilde{X} \cup \tilde{Y})) = \tilde{x}^*(N).$$

Thus, $\tilde{x}^*(N) = (h_{(S,T)})^{-\tilde{b}}(N)$, and we have $\tilde{x}^* \in B((h_{(S,T)})^{-\tilde{b}})$.

For the second part, note that for all $\tilde{x} \in B((h_{(S,T)})^{-\tilde{b}})$, we have $\tilde{x} \in P(h_{(S,T)})$. Moreover, the proof of the first part implies that $\tilde{x}(N) = (h_{(S,T)})^{-\tilde{b}}(N) = h_{(S,T)}(\tilde{X} \cup \tilde{Y}) - \tilde{b}(N \backslash (\tilde{X} \cup \tilde{Y}))$. Because $\tilde{x}(\tilde{X} \cup \tilde{Y}) \leq h_{(S,T)}(\tilde{X} \cup \tilde{Y})$ and $\tilde{x} \leq -\tilde{b}$, it follows that $\tilde{x}(\tilde{X} \cup \tilde{Y}) = h_{(S,T)}(\tilde{X} \cup \tilde{Y})$ and $\tilde{x}_i = -\tilde{b}_i$ for all $i \notin \tilde{X} \cup \tilde{Y}$. Thus, $x \in P_{(S,T)}(h)$, $x(\tilde{X}) - x(\tilde{Y}) = h(\tilde{X}, \tilde{Y})$, and $x_i = -b_i$ for all $i \notin \tilde{X} \cup \tilde{Y}$. We show that these three properties of x imply $x \in \tilde{B}(h)$.[1] For this, we choose an arbitrary $(X, Y) \in 3^N$ and define $X_1 := X \cap (\tilde{X} \cup \tilde{Y})$, $X_2 := X \backslash (\tilde{X} \cup \tilde{Y})$, $Y_1 := Y \cap (\tilde{X} \cup \tilde{Y})$, and $Y_2 := Y \backslash (\tilde{X} \cup \tilde{Y})$ (note that all sets X_1, X_2, Y_1, Y_2 are disjoint). We prove that $x(X) - x(Y) \leq h(X, Y)$. First, if $(X_1, Y_1) \sqsubseteq (\tilde{X}, \tilde{Y})$, then

$$x(X) - x(Y) = x(X_1) - x(Y_1) - b(X_2) + b(Y_2) \tag{2a}$$

$$= x(X_1) - x(Y_1) + x^*(\tilde{X} \cup X_2) - x^*(\tilde{X}) - x^*(\tilde{Y} \cup Y_2) + x^*(\tilde{Y}) \tag{2b}$$

$$\leq h(X_1, Y_1) + h(\tilde{X} \cup X_2, \tilde{Y} \cup Y_2) - h(\tilde{X}, \tilde{Y}) \tag{2c}$$

$$\leq h(X, Y), \tag{2d}$$

where (2a) follows because X_2, Y_2, and $\tilde{X} \cup \tilde{Y}$ are disjoint and $x_i = -b_i$ for all $i \notin \tilde{X} \cup \tilde{Y}$; (2b) follows because additionally $(A_+, A_-) \sqsubseteq (\tilde{X}, \tilde{Y})$ and thus, $X_2, Y_2 \subseteq A_0$; (2c) follows because $x \in P_{(S,T)}(h)$, $(X_1, Y_1) \sqsubseteq (\tilde{X}, \tilde{Y}) \sqsubseteq (S, T)$, $x^* \in \tilde{B}(h)$, and $x(\tilde{X}) - x(\tilde{Y}) = h(\tilde{X}, \tilde{Y})$; and (2d) follows by bisubmodularity of h. Second, if $(X_1, Y_1) \not\sqsubseteq (\tilde{X}, \tilde{Y})$, we have that

$$(X \cup \tilde{X}) \backslash (Y \cup \tilde{Y}) = (\tilde{X} \backslash Y_1) \cup X_2 \tag{3}$$

and analogously, $(Y \cup \tilde{Y}) \backslash (X \cup \tilde{X}) = (\tilde{Y} \backslash X_1) \cup Y_2$ (see Figure 1). Moreover, because $((\tilde{X} \backslash Y_1) \cup X_2) \cap (\tilde{X} \cup \tilde{Y}) = \tilde{X} \backslash Y \subseteq \tilde{X}$ and $((\tilde{Y} \backslash X_1) \cup Y_2) \cap (\tilde{X} \cup \tilde{Y}) = \tilde{Y} \backslash X \subseteq \tilde{Y}$, it follows from (2) that

$$x((X \cup \tilde{X}) \backslash (Y \cup \tilde{Y})) - x((Y \cup \tilde{Y}) \backslash (X \cup \tilde{X})) \leq h((X \cup \tilde{X}) \backslash (Y \cup \tilde{Y}), (Y \cup \tilde{Y}) \backslash (X \cup \tilde{X})) \tag{4}$$

and also, $x(X \cap \tilde{X}) - x(Y \cap \tilde{Y}) \leq h(X \cap \tilde{X}, Y \cap \tilde{Y})$. We are now ready to derive the following:

$$\begin{aligned} & x(X) - x(Y) \\ &= x(X_1 \cap \tilde{X}) + x(X_1 \cap \tilde{Y}) - x(Y_1 \cap \tilde{X}) - x(Y_1 \cap \tilde{Y}) + x(X_2) - x(Y_2) \end{aligned} \tag{5a}$$

$$= x(X \cap \tilde{X}) + x(X_1 \cap \tilde{Y}) - x(Y_1 \cap \tilde{X}) - x(Y \cap \tilde{Y}) + x(X_2) - x(Y_2) \tag{5b}$$

$$= x(X \cap \tilde{X}) - x(Y \cap \tilde{Y}) - h(\tilde{X}, \tilde{Y}) + x(\tilde{X} \backslash Y_1) - x(\tilde{Y} \backslash X_1) + x(X_2) - x(Y_2) \tag{5c}$$

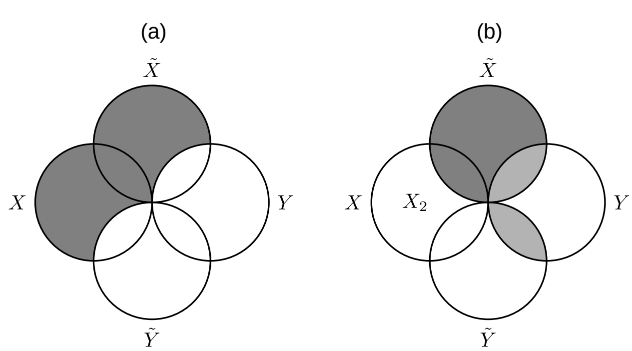
$$= x(X \cap \tilde{X}) - x(Y \cap \tilde{Y}) - h(\tilde{X}, \tilde{Y}) + x((X \cup \tilde{X}) \backslash (Y \cup \tilde{Y})) - x((Y \cup \tilde{Y}) \backslash (X \cup \tilde{X})) \tag{5d}$$

$$\leq h(X \cap \tilde{X}, Y \cap \tilde{Y}) - h(\tilde{X}, \tilde{Y}) + h((X \cup \tilde{X}) \backslash (Y \cup \tilde{Y}), (Y \cup \tilde{Y}) \backslash (X \cup \tilde{X})) \tag{5e}$$

$$\leq h(X, Y), \tag{5f}$$

where (5a) follows because $X_1, X_2 \subseteq \tilde{X} \cup \tilde{Y}$, (5b) follows because $X_2 \cap \tilde{X} = Y_2 \cap \tilde{Y} = \emptyset$, (5c) follows because $x(\tilde{X}) - x(\tilde{Y}) = h(\tilde{X}, \tilde{Y})$, (5d) follows from (3), (5e) follows from (4), and (5f) follows by bisubmodularity of h. Thus, we have $x(X) - x(Y) \leq h(X, Y)$, and because (X, Y) was chosen arbitrarily from 3^N, it follows that $x \in \tilde{B}(h)$. □

Figure 1. Venn diagrams of X, Y, $\tilde{X}$, and $\tilde{Y}$ in the first part of the proof of Lemma 13. (a) $(X \cup \tilde{X}) \backslash (Y \cup \tilde{Y})$. (b) Y_1 (light gray), $\tilde{X} \backslash Y_1$ (dark gray), and X_2.



5. From (*a*, *b*)-Majorization to (Bi-)Submodularity

In the previous section, we showed the existence of least-majorized elements for several sets, including submodular, supermodular, base, and bisubmodular polyhedra. In this section, we show that the existence of such elements is limited to these sets within the class of compact convex or bounded hole-free sets. Together with the existence results of the previous section, this yields characterizations of the existence of least-majorized elements in terms of submodular, supermodular, base, and bisubmodular polyhedra. We characterize the existence of least weakly (a, b)-supermajorized and (a, b)-submajorized, least (a, b)-majorized, and least weakly absolutely (a, b)-majorized elements in Sections 5.1–5.4, respectively.

5.1. Least Weakly (*a*, *b*)-Supermajorized Elements and Submodular Polyhedra

To characterize the existence of least weakly (a, b)-supermajorized elements, we first prove two intermediate results. First, in Lemma 14, we prove that for any closed and bounded subset $C \subset \mathbb{R}^n$, the existence of a least $(1, b)$-majorized element for all $b \in \mathbb{R}^n$ implies the existence of vectors whose nested sums are maximal among vectors in C.

Lemma 14. *Let $C \subset \mathbb{R}^n$ be a closed and bounded set. Suppose that C has a least weakly $(1, b)$-supermajorized element for each $b \in \mathbb{R}^n$. Then, for each permutation π of N, there exists a vector $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \geq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$.*

Proof. For each $i \in N$, we define l_i and u_i as the minimum and maximum value of a variable x_i in C (i.e., $l_i := \min_{x \in C} x_i$ and $u_i := \max_{x \in C} x_i$). Given a permutation π of N, we recursively define the following vector:

$$\hat{b}_{\pi(1)} := l_{\pi(1)}; \quad \hat{b}_{\pi(i+1)} := \hat{b}_{\pi(i)} + u_{\pi(i)} - l_{\pi(i+1)} + 1, \quad i \in \{1, \ldots, n-1\}.$$

Moreover, we define for each $k \in N$ the convex function $\hat{\phi}^k(\zeta) := \max(0, \hat{b}_{\pi(k)} + u_{\pi(k)} - \zeta)$. Note that $\hat{b}_{\pi(i+1)} + u_{\pi(i+1)} \geq \hat{b}_{\pi(i)} + u_{\pi(i)} + 1$ for all $i \in \{1, \ldots, n-1\}$ and thus, that for each $k \in N$,

$$\sum_{i \in N} \hat{\phi}^k(x_i + \hat{b}_i) = -\sum_{i=1}^k (\hat{b}_{\pi(i)} + x_{\pi(i)}) + k(\hat{b}_{\pi(k)} + u_{\pi(k)}). \tag{6}$$

It follows from Lemma 8 that any least weakly $(1, \hat{b})$-supermajorized element $x^* \in C$ minimizes $\sum_{i \in N} \hat{\phi}^k(x_i + \hat{b}_i)$ over C for all $k \in N$. Thus, for all $x \in C$ and $k \in N$, we have

$$\begin{aligned} -\sum_{i=1}^k (x^*_{\pi(i)} + \hat{b}_{\pi(i)}) + k(\hat{b}_{\pi(k)} + u_{\pi(k)}) = \sum_{i \in N} \hat{\phi}^k(x^*_i + \hat{b}_i) &\leq \sum_{i \in N} \hat{\phi}^k(x_i + \hat{b}_i) \\ &= -\sum_{i=1}^k (x_{\pi(i)} + \hat{b}_{\pi(i)}) + k(\hat{b}_{\pi(k)} + u_{\pi(k)}), \end{aligned} \tag{7}$$

which implies that $\sum_{i=1}^k x^*_{\pi(i)} \geq \sum_{i=1}^k x_{\pi(i)}$. □

Remark 1. Note that the functions $\hat{\phi}^k$ as defined in the proof of this lemma are not (continuously) differentiable. Alternatively, we can prove the result of the lemma starting from the assumption that for any $b \in \mathbb{R}^n$, there exists a vector in C that is an optimal solution to $\min_{x \in C} \sum_{i \in N} \phi(x_i + b_i)$ for any *continuously differentiable* convex function ϕ. The corresponding proof is equal to that of Lemma 14, except that for each $k \in N$, we choose $\hat{\phi}^k$ as the following continuously differentiable and convex function:

$$\hat{\phi}^k(\zeta) = \begin{cases} \hat{b}_{\pi(k)} - u_{\pi(k)} - \zeta + \frac{1}{2} & \text{if } \zeta \leq \hat{b}_{\pi(k)} - u_{\pi(k)}; \\ -\frac{1}{2}(\hat{b}_{\pi(k)} - u_{\pi(k)} + 1 - \zeta)^4 & \text{if } \hat{b}_{\pi(k)} - u_{\pi(k)} \leq \zeta \\ \quad +(\hat{b}_{\pi(k)} - u_{\pi(k)} + 1 - \zeta)^3 & \quad \leq \tilde{b}_{\pi(k)} - u_{\pi(k)} + 1; \\ 0 & \text{if } \zeta \geq \tilde{b}_{\pi(k)} - u_{\pi(k)} + 1. \end{cases}$$

To see this, note that for all $x \in C$ and $i \in N$, we have by definition of $\hat{b}$ that $x_i + \hat{b}_i \notin (\hat{b}_{\pi(k)} - u_{\pi(k)}, \hat{b}_{\pi(k)} - u_{\pi(k)} + 1)$ for any $k \in N$. It follows that $\hat{\phi}^k(x_i + \hat{b}_i) = \max(0, \hat{b}_{\pi(k)} + u_{\pi(k)} - x_i - \hat{b}_i + \frac{1}{2})$. As a consequence, all steps in the proof of Lemma 14 also hold for this choice of $\hat{\phi}^k$, where we only need to add the term $\frac{1}{2}k$ to the right-hand side of (6) and to the left- and right-hand sides of (7).

In a second step, we prove in Lemma 15 that the convex hull of any closed and bounded set satisfying the result of Lemma 14 is contained in a submodular polyhedron and contains the corresponding base polyhedron. The proof of this lemma is inspired by Nakamura [37, proof of theorem 1], where the optimality of Edmonds' classical greedy algorithm for linear optimization (Edmonds [16]) is characterized in terms of submodular polyhedra.

Lemma 15. *Let $C \subset \mathbb{R}^n$ be a closed and bounded set. If for each permutation π of N, there exists a vector $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \geq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$, then the set function $f(A) := \max_{x \in C} x(A)$ is submodular, and we have $B(f) \subseteq \mathrm{co}(C) \subseteq P(f)$.*

Proof. We first prove that the set function f as defined in the lemma is submodular. Given two sets $A, B \in 2^N$, we define $X := A \cap B$, $Y := A \backslash B$, and $Z := B \backslash A$. Let π be a permutation whose first $|X|$ elements are indices in X and whose next $|Y \cup Z|$ elements are indices in $Y \cup Z$. By assumption, there exists a solution x^* such that $x^*(X) \geq x(X)$ and $x^*(X \cup Y \cup Z) \geq x(X \cup Y \cup Z)$ for all $x \in C$. It follows that

$$\begin{aligned} f(A) + f(B) &= f(X \cup Y) + f(X \cup Z) = \max_{x \in C} x(X \cup Y) + \max_{x \in C} x(X \cup Z) \\ &\geq x^*(X \cup Y) + x^*(X \cup Z) = x^*(X) + x^*(X \cup Y \cup Z) \\ &= \max_{x \in C} x(X) + \max_{x \in C} x(X \cup Y \cup Z) = f(X) + f(X \cup Y \cup Z) \\ &= f(A \cap B) + f(A \cup B), \end{aligned}$$

which implies that f is submodular.

Finally, we prove that $B(f) \subseteq \mathrm{co}(C) \subseteq P(f)$. To prove that $\mathrm{co}(C) \subseteq P(f)$, note that $C \subseteq P(f)$ by definition of f and $P(f)$. Because $P(f)$ is a convex set, it follows that $\mathrm{co}(C) \subseteq P(f)$. To prove that $\mathrm{co}(C) \supseteq B(f)$, suppose that there exists a vector $z \in B(f)$ that is not in $\mathrm{co}(C)$. Because $\mathrm{co}(C)$ is compact and convex, we may assume without loss of generality that z is an extreme point of $B(f)$. It follows from Lemma 2 that there exists a permutation π of N such that $z_{\pi(1)} = f(\{\pi(1)\})$, and for each $k \geq 1$, we have $z_{\pi(k+1)} = f(\{\pi(1), \ldots, \pi(k+1)\}) - f(\{\pi(1), \ldots, \pi(k)\})$. By assumption, there exists a vector $y^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k y^*_{\pi(i)} \geq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$. It follows that

$$z_{\pi(1)} = f(\{\pi(1)\}) = \max_{x \in C} x_{\pi(1)} = y^*_{\pi(1)},$$

and for each $k \geq 1$, we have

$$\begin{aligned} z_{\pi(k+1)} &= f(\{\pi(1), \ldots, \pi(k+1)\}) - f(\{\pi(1), \ldots, \pi(k)\}) \\ &= \max_{x \in C} \sum_{i=1}^{k+1} x_{\pi(i)} - \max_{x \in C} \sum_{i=1}^{k} x_{\pi(i)} = \sum_{i=1}^{k+1} y^*_{\pi(i)} - \sum_{i=1}^{k} y^*_{\pi(i)} = y^*_{\pi(k+1)}. \end{aligned}$$

This implies that $z = y^*$ and thus, that $z \in \mathrm{co}(C)$. This is a contradiction, which means that we have $\mathrm{co}(C) \supseteq B(f)$. □

When the set C in Lemma 15 is convex, it follows that C itself is contained in a submodular polyhedron and contains the corresponding base polyhedron because $\mathrm{co}(C) = C$. Furthermore, we show in Corollary 2 that when C is an integral hole-free set, we can adjust the proof of the lemma slightly so that we may conclude that C is contained in an integral submodular polyhedron and contains the corresponding integral base polyhedron.

Corollary 2. *Let $C \subseteq \mathbb{Z}^n$ be a bounded integral hole-free set. If for each permutation π of N, there exists a vector $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \geq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$, then the set function $f(A) := \max_{x \in C} x(A)$ is submodular, and we have $B^{\mathbb{Z}}(f) \subseteq C \subseteq P^{\mathbb{Z}}(f)$.*

Proof. The result follows from two slight adjustments of the proof of Lemma 15, where we now aim to prove that $B^{\mathbb{Z}}(f) \subseteq C \subseteq P^{\mathbb{Z}}(f)$. First, we have $C \subseteq P^{\mathbb{Z}}(f)$ because for each $y \in C$ and $A \in 2^N$, it holds that $y(A) \leq \max_{x \in C} x(A) = f(A)$. Second, to prove that $C \supseteq B^{\mathbb{Z}}(f)$, we now suppose that there exists a vector $z \in B^{\mathbb{Z}}(f)$ that is not in C. Because C is bounded and hole free, we may assume without loss of generality that z is an extreme point of $B(f)$. Following the remainder of the proof, we may conclude that $z \in C$. □

We are now ready to prove our two characterizations of the existence of least weakly (a, b)-supermajorized elements.

Theorem 1. *Let $C \subset \mathbb{R}^n$ be a compact convex set. The following statements are equivalent.*

1. *C has a least weakly (a, b)-supermajorized element for each $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$.*

2. *C has a least weakly $(1, b)$-supermajorized element for each $b \in \mathbb{R}^n$.*
3. *For each $b \in \mathbb{R}^n$, there exists $x^* \in C$ that is an optimal solution to $\min_{x \in C} \Phi(x + b)$ for any choice of nonincreasing continuous Schur-convex function Φ.*
4. *For each permutation π of N, there exists $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \geq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$.*
5. *The function $f(A) := \max_{x \in C} x(A)$ is submodular and $B(f) \subseteq C \subseteq P(f)$.*

Proof. (2) is a special case of (1), (2) and (3) are equivalent because of Lemma 5, (2) implies (4) via Lemma 14, (4) implies (5) via Lemma 15 because $\text{co}(C) = C$, and (5) implies (1) via Lemma 9 and Corollary 1. □

Theorem 2. *Let $C \subset \mathbb{Z}^n$ be a bounded integral hole-free set. The following statements are equivalent.*
1. *C has a least weakly $(1, b)$-supermajorized element for each $b \in \mathbb{Z}^n$.*
2. *For each $b \in \mathbb{Z}^n$, there exists $x^* \in C$ that is an optimal solution to $\min_{x \in C} \Phi(x + b)$ for any choice of nonincreasing continuous Schur-convex function Φ.*
3. *For each permutation π of N, there exists $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \geq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$.*
4. *The function $f(A) := \max_{x \in C} x(A)$ is integral and submodular and $B^{\mathbb{Z}}(f) \subseteq C \subseteq P^{\mathbb{Z}}(f)$.*

Proof. (1) and (2) are equivalent because of Lemma 5, (1) implies (3) via Lemma 14, (3) implies (4) via Corollary 2, and (4) implies (1) via Corollary 1. □

5.2. Least Weakly (*a*, *b*)-Submajorized Elements and Supermodular Polyhedra

A characterization analogous to Theorems 1 and 2 can be proven for the existence of least weakly (a, b)-submajorized elements. This proof relies on the following intermediate results, whose proofs are analogous to those of Lemmas 14 and 15 and Corollary 2, respectively.

Lemma 16. *Let $C \subset \mathbb{R}^n$ be a closed and bounded set. Suppose that C has a least weakly $(1, b)$-submajorized element for each $b \in \mathbb{R}^n$. Then, for each permutation π of N, there exists a vector $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \leq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$.*

Lemma 17. *Let $C \subset \mathbb{R}^n$ be a closed and bounded set. If for each permutation π of N, there exists a vector $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \leq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$, then the set function $g(A) := \min_{x \in C} x(A)$ is supermodular, and we have $B_{\text{sup}}(g) \subseteq \text{co}(C) \subseteq P_{\text{sup}}(g)$.*

Corollary 3. *Let $C \subseteq \mathbb{Z}^n$ be a bounded integral hole-free set. If for each permutation π of N, there exists a vector $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \leq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$, then the set function $g(A) := \min_{x \in C} x(A)$ is supermodular, and we have $B^{\mathbb{Z}}_{\text{sup}}(g) \subseteq C \subseteq P^{\mathbb{Z}}_{\text{sup}}(g)$.*

These results lead to the following characterization of least weakly (a, b)-submajorized elements, whose proofs are analogous to those of Theorems 1 and 2, respectively.

Theorem 3. *Let $C \subset \mathbb{R}^n$ be a compact convex set. The following statements are equivalent.*
1. *C has a least weakly (a, b)-submajorized element for each $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$.*
2. *C has a least weakly $(1, b)$-submajorized element for each $b \in \mathbb{R}^n$.*
3. *For each $b \in \mathbb{R}^n$, there exists $x^* \in C$ that is an optimal solution to $\min_{x \in C} \Phi(x + b)$ for any choice of nondecreasing continuous Schur-convex function Φ.*
4. *For each permutation π of N, there exists $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \leq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$.*
5. *The function $g(A) := \min_{x \in C} x(A)$ is supermodular, and $B_{\text{sup}}(g) \subseteq C \subseteq P_{\text{sup}}(g)$.*

Theorem 4. *Let $C \subset \mathbb{Z}^n$ be a bounded integral hole-free set. The following statements are equivalent.*
1. *C has a least weakly $(1, b)$-submajorized element for each $b \in \mathbb{Z}^n$.*
2. *For each $b \in \mathbb{Z}^n$, there exists $x^* \in C$ that is an optimal solution to $\min_{x \in C} \Phi(x + b)$ for any choice of nondecreasing continuous Schur-convex function Φ.*
3. *For each permutation π of N, there exists $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k x^*_{\pi(i)} \leq \sum_{i=1}^k x_{\pi(i)}$ for all $k \in N$.*
4. *The function $g(A) := \min_{x \in C} x(A)$ is integral and supermodular, and $B^{\mathbb{Z}}_{\text{sup}}(g) \subseteq C \subseteq P^{\mathbb{Z}}_{\text{sup}}(g)$.*

5.3. Least (*a*, *b*)-Majorized Elements and Base Polyhedra

Using the characterization of the existence of least weakly (a, b)-submajorized elements in Theorems 1 and 2, we obtain the following characterization of least (a, b)-majorized elements in terms of (integral) base polyhedral.

Theorem 5. *Let $C \subset \mathbb{R}^n$ be a compact convex set. The following statements are equivalent.*
1. *C has a least (a, b)-majorized element for each $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$.*

2. *C has a least $(1, b)$-majorized element for each $b \in \mathbb{R}^n$.*
3. *For each $b \in \mathbb{R}^n$, there exists $x^* \in C$ that is an optimal solution to $\min_{x \in C} \Phi(x + b)$ for any choice of continuous Schur-convex function Φ.*
4. *The function $f(A) := \max_{x \in C} x(A)$ is submodular, and $C = B(f)$.*

Proof. (2) is a special case of (1), (2) and (3) are equivalent because of Lemma 6, and (4) implies (1) by Lemma 9. To prove that (2) implies (4), note that (2) implies that all elements in C have the same element sum. Moreover, it follows from Theorem 1 that $B(f) \subseteq C \subseteq P(f)$ for the submodular function $f(A) := \max_{x \in C} x(A)$. Because $x(N) = f(N)$ for all $x \in B(f)$, it follows that $x(N) = f(N)$ for all $x \in C$, and thus, $C \subseteq B(f)$. We may, therefore, conclude that $C = B(f)$. □

Theorem 6. *Let $C \subset \mathbb{Z}^n$ be a bounded integral hole-free set. The following statements are equivalent.*
1. *C has a least $(1, b)$-majorized element for each $b \in \mathbb{Z}^n$.*
2. *For each $b \in \mathbb{Z}^n$, there exists $x^* \in C$ that is an optimal solution to $\min_{x \in C} \Phi(x + b)$ for any choice of continuous Schur-convex function Φ.*
3. *The function $f(A) := \max_{x \in C} x(A)$ is integral and submodular, and $C = B^{\mathbb{Z}}(f)$.*

Proof. The proof is analogous to that of Theorem 5. □

We conclude this subsection with a result that may be of independent interest. First, for a given feasible set C, a vector $x \in C$, and continuous convex functions ϕ_i for each $i \in N$, we state the following condition on x.

Condition 1. *For each pair $(i, k) \in N^2$ such that $x_i + \alpha(e^k - e^i) \in C$ for some $\alpha > 0$, we have $\phi_k^+(x_k) \geq \phi_i^-(x_i)$, where ϕ_k^+ is the right derivative of ϕ_k and ϕ_i^- is the left derivative of ϕ_i.*

The proof of the existence of least (a, b)-majorized elements in (integral) base polyhedra (see also Lemma 9) is directly based on Condition 1. More generally, Condition 1 is a valid optimality condition for minimizing separable convex functions over (integral) base polyhedra (see, e.g., Fujishige [22, theorem 8.1]). Here, we prove that this condition is unique in the sense that (integral) base polyhedra are the only compact convex (bounded integral hole-free) sets for which Condition 1 is a valid optimality condition for any choice of continuous convex functions ϕ_i.

Corollary 4. *Let $C_1 \subset \mathbb{R}^n$ be a compact convex set and $C_2 \subset \mathbb{Z}^n$ be a bounded integral hole-free set. Then, the following hold.*
- *Condition 1 is a valid optimality condition for the problem $\min_{x \in C_1} \sum_{i \in N} \phi_i(x_i)$ for any choice of continuous convex functions ϕ_i, $i \in N$, if and only if C_1 is a base polyhedron.*
- *Condition 1 is a valid optimality condition for the problem $\min_{x \in C_2} \sum_{i \in N} \phi_i(x_i)$ for any choice of continuous convex functions ϕ_i, $i \in N$, if and only if C_2 is an integral base polyhedron.*

Proof. The "if" parts follow from, for example, Fujishige [22, theorem 8.1]. Regarding the "only if" parts, Condition 1 implies via Schoot Uiterkamp et al. [42, theorem 1] the existence of least (a, b)-majorized elements for each pair of vectors $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$ (see also Lemma 9). It follows from parts (1) and (4) of Theorem 5 that C_1 is a base polyhedron and from parts (1) and (3) of Theorem 6 that C_2 is an integral base polyhedron. □

5.4. Least Weakly Absolutely (*a*, *b*)-Majorized Elements and Bisubmodular Polyhedra

In this section, we prove that bisubmodular polyhedra are the only compact convex sets with least weakly absolutely (a, b)-majorized elements for all $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$ and that integral bisubmodular polyhedra are the only bounded integral hole-free sets with least weakly absolutely $(1, b)$-majorized elements for all $b \in \mathbb{Z}^n$. For this, we follow the same proof idea as in Section 5.1 for the case of least weakly (a, b)-supermajorized elements and submodular polyhedra. First, we prove in Lemma 18 that, given a closed and bounded set $C \subset \mathbb{R}^n$, the existence of least weakly absolutely $(1, b)$-elements for all $b \in \mathbb{R}^n$ implies for each permutation π and orthant (S, T) the existence of a vector in C that maximizes functions of the form $x(X) - x(Y)$, where $(X, Y) \sqsubseteq (S, T)$.

Lemma 18. *Let $C \subset \mathbb{R}^n$ be a closed and bounded set. Suppose that C has a least weakly absolutely $(1, b)$-majorized element for each $b \in \mathbb{R}^n$. Then, for each permutation π of N and sign vector $s \in \{-1, 1\}^n$, there exists $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^k s_{\pi(i)} x^*_{\pi(i)} \geq \sum_{i=1}^k s_{\pi(i)} x_{\pi(i)}$ for all $k \in N$.*

Proof. Let a permutation π and a sign vector s be given, and let $(S, T) \in 3^N$ be the unique orthant corresponding to s: that is, $(S, T) := N|s$ (i.e., $S := \{i \in N | s_i = 1\}$ and $T := \{i \in N | s_i = -1\}$ (see also Section 2.3)). Moreover, let

$l_i := \min_{x \in C} x_i$ and $u_i := \max_{x \in C} x_i$ for $i \in N$. Based on (S, T), we define alternative bound vectors $\hat{l}, \hat{u} \in \mathbb{R}^n$ as follows:

$$\hat{l}_i := \begin{cases} u_i & \text{if } i \in S; \\ l_i & \text{if } i \in T; \end{cases} \qquad \hat{u}_i := \begin{cases} l_i & \text{if } i \in S; \\ u_i & \text{if } i \in T. \end{cases}$$

Note that $s_i \hat{l}_i \geq s_i \hat{u}_i$ for all $i \in N$. We recursively define the vector $\hat{b}$ as follows:

$$\hat{b}_{\pi(n)} := -\hat{l}_{\pi(n)} - s_{\pi(n)};$$
$$\hat{b}_{\pi(i)} := -s_{\pi(i)}(|\hat{b}_{\pi(i+1)} + \hat{u}_{\pi(i+1)}| + 1) - \hat{l}_{\pi(i)}, \quad i \in \{1, \ldots, n-1\}.$$

Note that for all i with $\pi(i) \in S$, we have $0 \geq \hat{b}_{\pi(i)} + \hat{l}_{\pi(i)} \geq \hat{b}_{\pi(i)} + \hat{u}_{\pi(i)}$ and that for all i with $\pi(i) \in T$, we have $0 \leq \hat{b}_{\pi(i)} + \hat{l}_{\pi(i)} \leq \hat{b}_{\pi(i)} + \hat{u}_{\pi(i)}$. It follows that for any $x \in C$, we have for i with $\pi(i) \in S$ that

$$-|\hat{b}_{\pi(i)} + \hat{l}_{\pi(i)}| = \hat{b}_{\pi(i)} + \hat{l}_{\pi(i)} \geq \hat{b}_{\pi(i)} + x_{\pi(i)} \geq \hat{b}_{\pi(i)} + \hat{u}_{\pi(i)} = -|\hat{b}_{\pi(i)} + \hat{u}_{\pi(i)}| \tag{8}$$

and for i with $\pi(i) \in T$ that

$$|\hat{b}_{\pi(i)} + \hat{l}_{\pi(i)}| = \hat{b}_{\pi(i)} + \hat{l}_{\pi(i)} \leq \hat{b}_{\pi(i)} + x_{\pi(i)} \leq \hat{b}_{\pi(i)} + \hat{u}_{\pi(i)} = |\hat{b}_{\pi(i)} + \hat{u}_{\pi(i)}|. \tag{9}$$

Also, note that for all $i < n$, we have

$$|\hat{b}_{\pi(i)} + \hat{l}_{\pi(i)}| = |\hat{b}_{\pi(i+1)} + \hat{u}_{\pi(i+1)}| + 1 > |\hat{b}_{\pi(i+1)} + \hat{l}_{\pi(i+1)}| \tag{10}$$

and that for all $i > 1$, we have

$$|\hat{b}_{\pi(i)} + \hat{u}_{\pi(i)}| = -s_{\pi(i-1)}(\hat{b}_{\pi(i-1)} + \hat{l}_{\pi(i-1)}) - 1 < |\hat{b}_{\pi(i-1)} + \hat{l}_{\pi(i-1)}|. \tag{11}$$

We define for each $k \in N$ the function $\hat{\phi}^k(y) := \max(-|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| - y, 0, -|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| + y)$. For each $k \in N$, we have for i with $\pi(i) \in S$ that

$$\begin{aligned} \hat{\phi}^k(x_{\pi(i)} + \hat{b}_{\pi(i)}) &= \max(-|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| - x_{\pi(i)} - \hat{b}_{\pi(i)}, 0, -|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| + x_{\pi(i)} + \hat{b}_{\pi(i)}) \\ &= \max(-|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| - x_{\pi(i)} - \hat{b}_{\pi(i)}, 0) \\ &= \begin{cases} -|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| - x_{\pi(i)} - \hat{b}_{\pi(i)} & \text{if } i \leq k; \\ 0 & \text{if } i > k, \end{cases} \end{aligned}$$

where the second equality follows because $x_{\pi(i)} + \hat{b}_{\pi(i)} \leq 0$ by (8) and the third equality follows from (8) and (10) (for $i \leq k$) and (8), (11), and (10) (for $i > k$). Analogously, we have for i with $\pi(i) \in T$ that

$$\begin{aligned} \hat{\phi}^k(x_{\pi(i)} + \hat{b}_{\pi(i)}) &= \max(-|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| - x_{\pi(i)} - \hat{b}_{\pi(i)}, 0, -|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| + x_{\pi(i)} + \hat{b}_{\pi(i)}) \\ &= \max(0, -|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| + x_{\pi(i)} + \hat{b}_{\pi(i)}) \\ &= \begin{cases} -|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| + x_{\pi(i)} + \hat{b}_{\pi(i)} & \text{if } i \leq k; \\ 0 & \text{if } i > k, \end{cases} \end{aligned}$$

where the second equality follows because $x_{\pi(i)} + \hat{b}_{\pi(i)} \geq 0$ by Equation (9) and the third equality follows from (9) and (10) (for $i \leq k$) and (9), (11), and (10) (for $i > k$). It follows that

$$\sum_{i \in N} \hat{\phi}^k(x_i + \hat{b}_i) = -\sum_{i=1}^{k} s_{\pi(i)}(x_{\pi(i)} + \hat{b}_{\pi(i)}) - k|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}|.$$

By assumption, there exists a vector $x^* \in C$ that simultaneously minimizes these functions over C for all $k \in N$. This means that for any $x \in C$, we have

$$-\sum_{i=1}^{k} s_{\pi(i)}(x^*_{\pi(i)} + \hat{b}_{\pi(i)}) - k|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}| = \sum_{i\in N} \hat{\phi}^k(x^*_i + \hat{b}_i) \le \sum_{i\in N} \hat{\phi}^k(x_i + \hat{b}_i)$$
$$= -\sum_{i=1}^{k} s_{\pi(i)}(x_{\pi(i)} + \hat{b}_{\pi(i)}) - k|\hat{b}_{\pi(k)} + \hat{l}_{\pi(k)}|,$$

which implies that $\sum_{i=1}^{k} s_{\pi(i)} x^*_{\pi(i)} \ge \sum_{i=1}^{k} s_{\pi(i)} x_{\pi(i)}$ for all $k \in N$. □

Remark 2. Note that analogously to Lemma 14 for least $(1,b)$-majorized elements, we can prove the result of Lemma 18 starting from the assumption that for each $b \in \mathbb{R}^n$, there exists a vector in C that is an optimal solution to $\min_{x\in C} \sum_{i\in N} \phi(x_i + b_i)$ for any *continuously differentiable* convex function ϕ (see also Remark 1).

In a second step, we prove in Lemma 19 that the convex hull of any closed and bounded set satisfying the result of Lemma 18 is a bisubmodular polyhedron. Analogously to the proof of Lemma 15, the proof of this lemma is inspired by Nakamura [37, proof of theorem 2], where the optimality of the greedy algorithm in Dunstan and Welsh [13] for linear optimization is characterized in terms of bisubmodular polyhedra.

Lemma 19. *Let $C \subset \mathbb{R}^n$ be a closed and bounded set. If for each permutation π of N and sign vector $s \in \{-1,1\}^n$, there exists $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^{k} s_{\pi(i)} x^*_{\pi(i)} \ge \sum_{i=1}^{k} s_{\pi(i)} x_{\pi(i)}$ for all $k \in N$, then the convex hull of C is a bisubmodular polyhedron defined by the biset function $h(X,Y) := \max_{x\in C}(x(X) - x(Y))$.*

Proof. First, we prove that h is bisubmodular. For this, choose any two pairs $(S_1,T_1),(S_2,T_2) \in 3^N$. Because $S_1 \cap T_1 = S_2 \cap T_2 = \emptyset$, we have that $i \in S_1 \cap S_2$ implies $i \notin T_1 \cup T_2$ and thus, that $S_1 \cap S_2 \subseteq (S_1 \cup S_2)\backslash(T_1 \cup T_2)$. Analogously, we have $T_1 \cap T_2 \subseteq (T_1 \cup T_2)\backslash(S_1 \cup S_2)$. This means that there exists a permutation π of N that satisfies the following properties:

- $\pi(i) \in (S_1 \cap S_2) \cup (T_1 \cap T_2)$ if $1 \le i \le |S_1 \cap S_2| + |T_1 \cap T_2|$ and
- $\pi(i) \in (S_1 \cup S_2)\backslash(T_1 \cup T_2) \cup (T_1 \cup T_2)\backslash(S_1 \cup S_2)$ if $1 \le i \le |(S_1 \cup S_2)\backslash(T_1 \cup T_2) \cup (T_1 \cup T_2)\backslash(S_1 \cup S_2)|$.

By assumption, there exists a vector $x^* \in C$ such that for all $k \in N$, we have $\sum_{i=1}^{k} s_{\pi(i)} x^*_{\pi(i)} \ge \sum_{i=1}^{k} s_{\pi(i)} x_{\pi(i)}$ for all $x \in C$. In particular, this holds for $k = |S_1 \cap S_2| + |T_1 \cap T_2|$ and $k = |(S_1 \cup S_2)\backslash(T_1 \cup T_2)| + |(T_1 \cup T_2)\backslash(S_1 \cup S_2)|$, which implies that $x^*(S_1 \cap S_2) - x^*(T_1 \cap T_2) \ge x(S_1 \cap S_2) - x(T_1 \cap T_2)$ and $x^*((S_1 \cup S_2)\backslash(T_1 \cup T_2)) - x^*((T_1 \cup T_2)\backslash(S_1 \cup S_2)) \ge x((S_1 \cup S_2)\backslash(T_1 \cup T_2)) - x((T_1 \cup T_2)\backslash(S_1 \cup S_2))$ for all $x \in C$. It follows that

$$\begin{aligned}
&h(S_1,T_1) + h(S_2,T_2)\\
&= \max_{x\in C}(x(S_1) - x(T_1)) + \max_{x\in C}(x(S_2) - x(T_2))\\
&\ge x^*(S_1) - x^*(T_1) + x^*(S_2) - x^*(T_2)\\
&= x^*(S_1 \cup S_2) + x^*(S_1 \cap S_2) - x^*(T_1 \cup T_2) - x^*(T_1 \cap T_2)\\
&= x^*(S_1 \cup S_2) - x^*(S_1 \cup S_2 \cup T_1 \cup T_2) + x^*(S_1 \cup S_2 \cup T_1 \cup T_2) - x^*(T_1 \cup T_2)\\
&\quad + \max_{x\in C}(x(S_1 \cap S_2) - x(T_1 \cap T_2))\\
&= -x^*((T_1 \cup T_2)\backslash(S_1 \cup S_2)) + x^*((S_1 \cup S_2)\backslash(T_1 \cup T_2)) + h(S_1 \cap S_2, T_1 \cap T_2)\\
&= \max_{x\in C}(x((S_1 \cup S_2)\backslash(T_1 \cup T_2)) - x((T_1 \cup T_2)\backslash(S_1 \cup S_2))) + h(S_1 \cap S_2, T_1 \cap T_2)\\
&= h((S_1 \cup S_2)\backslash(T_1 \cup T_2), (T_1 \cup T_2)\backslash(S_1 \cup S_2)) + h(S_1 \cap S_2, T_1 \cap T_2).
\end{aligned}$$

Thus, h is bisubmodular.

Finally, we prove that $\text{co}(C) = \tilde{B}(h)$. For this, first note that $C \subseteq \tilde{B}(h)$ by definition of h and $\tilde{B}(h)$. Because $\tilde{B}(h)$ is a convex set, it follows that $\text{co}(C) \subseteq \tilde{B}(h)$. To prove that $\text{co}(C) \supseteq \tilde{B}(g)$, suppose that there exists a vector $z \in \tilde{B}(h)$ that is not in $\text{co}(C)$. Because $\text{co}(C)$ is convex, we may assume without loss of generality that z is an extreme point of $\tilde{B}(h)$. It follows from Lemma 3 that there exists a permutation π of N and a sign vector $s \in \{-1,1\}^n$ such that $z_{\pi(1)} = s_{\pi(1)} h(\{\pi(1)\}|s)$, and for each $k > 1$, we have

$$z_{\pi(k)} = s_{\pi(k)}(h(\{\pi(1),\dots,\pi(k)\}|s) - h(\{\pi(1),\dots,\pi(k-1)\}|s)).$$

By assumption, there exists a vector $y^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^{k} s_{\pi(i)} y^*_{\pi(i)} \geq \sum_{i=1}^{k} s_{\pi(i)} x_{\pi(i)}$ for all $k \in N$. It follows that

$$z_{\pi(1)} = s_{\pi(1)} h(\{\pi(1)\}|s) = s_{\pi(1)} \max_{x \in C} x(\{\pi(1)\}|s) = y^*_{\pi(1)},$$

and

$$\begin{aligned} z_{\pi(k)} &= s_{\pi(k)}(h(\{\pi(1),\ldots,\pi(k)\}|s) - h(\{\pi(1),\ldots,\pi(k-1)\}|s) \\ &= s_{\pi(k)} \max_{x \in C}(x(\{\pi(1),\ldots,\pi(k)\}^+) - x(\{\pi(1),\ldots,\pi(k)\}^-)) \\ &\quad - s_{\pi(k)} \max_{x \in C}(x(\{\pi(1),\ldots,\pi(k-1)\}^+) - x(\{\pi(1),\ldots,\pi(k-1)\}^-)) \\ &= s_{\pi(k)} \max_{x \in C} \sum_{i=1}^{k} s_{\pi(i)} x_{\pi(i)} - s_{\pi(k)} \max_{x \in C} \sum_{i=1}^{k-1} s_{\pi(i)} x_{\pi(i)} \\ &= s_{\pi(k)} \sum_{i=1}^{k} s_{\pi(i)} y^*_{\pi(i)} - s_{\pi(k)} \sum_{i=1}^{k-1} s_{\pi(i)} y^*_{\pi(i)} \\ &= s^2_{\pi(k)} y^*_{\pi(k)} = y^*_{\pi(k)}. \end{aligned}$$

This implies that $z = y^*$ and thus, that $z \in \mathrm{co}(C)$. This is a contradiction, which means that we have $\mathrm{co}(C) \supseteq \tilde{B}(h)$, and thus, $\mathrm{co}(C) = \tilde{B}(h)$. □

Analogously to the case of submodular and base polyhedra in Lemma 15 and Corollary 2, it follows that any *convex* set C satisfying the requirements of Lemma 19 is necessarily a bisubmodular polyhedron because it equals its convex hull and that any *integral hole-free* set C satisfying the requirements of the lemma is an integral bisubmodular polyhedron because it equals the integral points of its convex hull. As a consequence, we can now prove our two main characterization results for the existence of least weakly absolutely (a, b)-majorized elements and (integral) bisubmodular polyhedral.

Theorem 7. *Let $C \subset \mathbb{R}^n$ be a compact convex set. The following statements are equivalent.*
1. *C has a least weakly absolutely (a, b)-majorized element for each $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$.*
2. *C has a least weakly absolutely $(1, b)$-majorized element for each $b \in \mathbb{R}^n$.*
3. *For each $b \in \mathbb{R}^n$, there exists $x^* \in C$ that is an optimal solution to $\min_{x \in C} \Phi(x+b)$ for any choice of monotonically even continuous Schur-convex function Φ.*
4. *For each permutation π of N and sign vector $s \in \{-1,1\}^n$, there exists $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^{k} s_{\pi(i)} x^*_{\pi(i)} \geq \sum_{i=1}^{k} s_{\pi(i)} x_{\pi(i)}$ for all $k \in N$.*
5. *The function $h(X,Y) := \max_{x \in C}(x(X) - x(Y))$ is bisubmodular, and $C = \tilde{B}(h)$.*

Proof. (2) is a special case of (1), (2) and (3) are equivalent because of Lemma 7, (2) implies (4) via Lemma 18, (4) implies (5) via Lemma 19, and (5) implies (1) via Lemma 13. □

Theorem 8. *Let $C \subset \mathbb{Z}^n$ be a bounded integral hole-free set. The following statements are equivalent.*
1. *C has a least weakly absolutely $(1, b)$-majorized element for each $b \in \mathbb{Z}^n$.*
2. *For each $b \in \mathbb{Z}^n$, there exists $x^* \in C$ that is an optimal solution to $\min_{x \in C} \Phi(x+b)$ for any choice of monotonically even continuous Schur-convex function Φ.*
3. *For each permutation π of N and sign vector $s \in \{-1,1\}^n$, there exists $x^* \in C$ such that for all $x \in C$, we have $\sum_{i=1}^{k} s_{\pi(i)} x^*_{\pi(i)} \geq \sum_{i=1}^{k} s_{\pi(i)} x_{\pi(i)}$ for all $k \in N$.*
4. *The function $h(X,Y) := \max_{x \in C}(x(X) - x(Y))$ is integral and bisubmodular, and $C = \tilde{B}^{\mathbb{Z}}(h)$.*

Proof. The proof is analogous to that of Theorem 7. □

We conclude this section by noting that in statement (3) of Theorems 1, 3, 5, and 7 and statement (2) of Theorems 2, 4, 6, and 8, we may restrict the stated choices of Φ that need to be checked to those that are, additionally, continuously differentiable. This is because of Remarks 1 and 2 that state that in the proofs of Lemmas 14 and 18 (and thus also in a proof for Lemma 16), it suffices that the least $(1, b)$-majorized element x^* is optimal for the problem $\min_{x \in C} \sum_{i \in N} \phi(x_i + b_i)$ only when ϕ is a *continuously differentiable* convex function.

6. Applications

In this section, we demonstrate the impact of our characterization results in several fields other than combinatorial optimization. In particular, we highlight the insights that the results provide in three application areas, namely power management, cooperative game theory, and regularized regression.

6.1. Energy Storage Scheduling and Power Management

To illustrate the impact of our characterization results on RAPs in power management applications, we take as an example the energy storage scheduling problem described in Schoot Uiterkamp et al. [42, section 5.2]. Given a time horizon consisting of n equidistant time intervals of length Δt indexed by the set $N := \{1, \ldots, n\}$, we determine for each interval $i \in N$ the (dis-)charging power x_i of the storage system during this interval. The goal is to optimize a given system objective function of the form $\Phi(x+p)$, where $p \in \mathbb{R}^n$ denotes the uncontrollable energy usage of the system and its environment (e.g., a household or neighborhood). Given the initial and target amounts of energy S_{start} and S_{end} at the start and end of the time horizon, respectively, the scheduling problem is formulated as follows:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \Phi(x+p) \\ \text{s.t.} \quad & 0 \leq S_{\text{start}} + \Delta t \cdot x(\{1, \ldots, j\}) \leq D, \quad j \in N \backslash \{n\}, \\ & S_{\text{start}} + \Delta t \cdot x(N) = S_{\text{end}}, \\ & X_{\min} \leq x_i \leq X_{\max}, \quad i \in N. \end{aligned}$$

The feasible region of this problem is a base polyhedron, and thus, Lemma 9 (Schoot Uiterkamp et al. [42, condition 1 and theorem 1]) applies. Using this observation, it was established that the following objectives can be optimized simultaneously:

- minimizing exchange with the main grid: $\Phi(x+p) = \sum_{i \in N} |x_i + p_i|$;
- load profile flattening: $\Phi(x+p) = \sum_{i \in N} (x_i + p_i)^2$; and
- peak shaving under a threshold M: $\Phi(x+p) = \sum_{i \in N} \max(0, \underline{f}(x_i + p_i))$, where $\underline{f}$ is a convex nondecreasing function with $\underline{f}(M) = 0$.

Now, utilizing our characterization result in Theorem 5, we expand upon this result in two directions. First, we may extend the collection of equivalent objectives with general norms. In particular, to model the widely used objective of (general) peak shaving (i.e., without a threshold M), we may consider the max-norm $\Phi(x+p) = \max_{i \in N} |x_i + p_i|$. We may thus conclude that a solution that optimizes the (quadratic) objective of load profile flattening also optimizes the general peak-shaving objective.

Second, however, the characterization also suggests that these properties might not hold anymore for extensions and variants of the basic energy storage model that destroy the submodular structure. Examples of these are the inclusion of conversion losses (Guo and Fang [28]) or restricting the charging rate to a finite set of base rates (van der Klauw et al. [46]). For these cases, there exist Schur-convex choices of Φ for which an optimizer of the load profile-flattening objective is not optimal. An interesting question for future work is to investigate whether this includes the aforementioned choices for Φ that are relevant in energy storage scheduling.

Finally, we note that the observations in this subsection also apply to other device scheduling problems, such as electric vehicles and heat pumps (van der Klauw et al. [46]), as well as to other power management problems with submodular structure, including power allocation in multichannel communication systems, vessel speed optimization, and speed scaling (see also Schoot Uiterkamp et al. [42, section 5]).

6.2. Convex Cooperative Games with Transferable Utility

In cooperative game theory, one important subclass of games consists of those with TU. These games are defined by a player set N (the grand coalition) and a set function v that assigns a value to each coalition. The core of the game consists of all payoff allocations where no coalition has an incentive to split from the grand coalition N, meaning that each coalition $A \in 2^N$ receives at least their value $v(A)$ and the total payoff equals the value $v(N)$ of the grand coalition. Two n-player TU games (N, v) and (N, w) are said to be strategic equivalent (S-equivalent) if there exist a value $k > 0$ and a vector $b \in \mathbb{R}^n$ such that $w(A) = kv(A) + b(A)$. Moreover, a TU game is called convex if its value function is supermodular, in which case its core is a base polyhedron. Several characterizations of convexity of a game exist: for example, in terms of extreme points of the core (Ichiishi [31]).

Here, we obtain a new characterization of convex games in terms of the existence of so-called egalitarian solutions (Dutta and Ray [14]). Egalitarian solutions aim to distribute the payoff of the grand coalition as equally as possible over the players. By definition, such solutions are exactly the least-majorized elements of the considered allocation space (see, e.g., Arin et al. [3]). For convex games, an egalitarian solution always exists and is a member of the core. However, the existence of such an egalitarian core solution (i.e., a core solution that is egalitarian with respect to other core solutions) is not sufficient for convexity of the game (Dutta and Ray [14]). Therefore, this property on its own does not yield a characterization of convex games. However, using Theorem 5, we do obtain a characterization when also requiring that each corresponding S-equivalent game has an egalitarian core solution.

Theorem 9. *Let (N, v) be a cooperative TU game with n players. Then, (N, v) is a convex game if and only if every TU game that is S-equivalent to (N, v) has an egalitarian core solution.*

Proof. Note that given v, the fact that every TU game S-equivalent to (N, v) has an egalitarian core solution is equivalent to the fact that for each $k > 0$ and $b \in \mathbb{R}^n$, the base polyhedron $B_{\sup}(g)(= B(g^{\#}))$ with $g(A) := kv(A) + b(A)$ has a least-majorized element.[2] Also note that given k and b, $B(g^{\#})$ has a least-majorized element if and only if $B_{\sup}(v)(= B(v^{\#}))$ has a least (a, b)-majorized element, where $a_i = \frac{1}{k}$ for all $i \in N$. It follows that the theorem is equivalent to stating that v is supermodular (or equivalently, $v^{\#}$ is submodular) if and only if $B(v^{\#})$ has a least (a^k, b)-majorized element for each $k > 0$ and $b \in \mathbb{R}^n$, where $a^k \in \mathbb{R}^n$ with $a_i^k := \frac{1}{k}$. Thus, the "if" part follows from the equivalence between parts (2) and (4) of Theorem 5 by considering the special case $k = 1$, and the "only if" part follows directly from the equivalence of parts (1) and (4) of the same theorem. □

Note that the proof of Theorem 9 suggests that it suffices to only check the existence of egalitarian core solutions for S-equivalent games with $k = 1$. Moreover, note that in contrast with other solution concepts such as the Shapley value, egalitarian core solutions are not invariant with regard to S-equivalence. This means that such solutions are generally not simply linear combinations of egalitarian core solutions of the original game.

One interesting question is whether the characterization result in Theorem 7 for bisubmodular polyhedra can be used to characterize bicooperative games (Bilbao et al. [7]), analogously to Theorem 9. A first step for this, which we leave for future work, would be to find a suitable definition of egalitarian solutions for this type of game and investigate how this definition corresponds with least weakly absolute majorization.

6.3. Regularized Regression Estimators

Regression is an important and widely used method for establishing relationships between dependent and independent variables. In regression problems, regularization is often applied to enforce a specific desirable structure among the regression coefficients (see, e.g., Hastie et al. [29, chapter 3]). Examples of this are reducing the number of relevant independent variables (structured sparsity) or adhering to a known structure within the input data (e.g., spatiotemporal relationships between independent variables).

The goal of this section is to demonstrate the potential of our characterization results in establishing new properties of common regression estimators, such as the LASSO and ridge regression estimators. To this end, we consider the regularized least squares problem with input matrix $X \in \mathbb{R}^{m\times n}$, vector $y \in \mathbb{R}^m$ of outcomes, a vector $\beta \in \mathbb{R}^n$ of coefficients, a regularizer Ω, and a regularization parameter t:

$$\min_{\beta\in\mathbb{R}^n} \|y - X\beta\|_2^2 \text{ such that } \Omega(\beta) \le t. \tag{12}$$

Several special cases of Problem (12) and the corresponding regression methods are obtained when Ω is a p-norm with $p = 1$ (LASSO), $p = 2$ (ridge regression), and $p = \infty$ (max-norm regression). We also consider the following problem, which is closely related to the regularized regression problem for general loss functions Φ:

$$\min_{\beta\in\mathbb{R}^n} \Phi(X^\top y - \beta) \text{ such that } \Omega(\beta) \le t. \tag{13}$$

One important application of regression problems is when performing design experiments. In such an experiment, the behavior of a system and the impact of independent variables are learned by choosing as input specific combinations of settings of the independent variables rather than observed measurement data. When testing for all combinations of settings for a given subset of the independent variables, the resulting design is called orthogonal, and the input matrix X is orthonormal (i.e., $X^\top X = I$ (see, e.g., Bailey [6], Zurovac and Brown [49])). We show in Theorem 10 that in this case, the optimal estimator β^* of the regularized regression Problem (12) is also optimal for Problem (13) for particular combinations of regularizers Ω and loss functions Φ.

Theorem 10. *Let β^* denote the optimal estimator in Problem (12). If the input matrix is orthonormal, the following hold.*

1. *For $\Omega(\beta) = \|\beta\|_1$ or $\Omega(\beta) = \|\beta\|_\infty = \max_{i\in N}|\beta_i|$, β^* is also an optimal estimator for Problem (13) for any choice of monotonically even continuous Schur-convex function Φ.*

2. *If $\Omega(\beta) = \|\beta\|_2$, there exists an outcome vector $y \in \mathbb{R}^n$ and a continuous Schur-convex function Φ that is nonincreasing, nondecreasing, or monotonically even such that β^* is not an optimal estimator for Problem* (13).

Proof. If $X^\top X = I$, the objective function of Problem (12) reduces to

$$\|y - X\beta\|_2^2 = y^\top y - 2y^\top X\beta + \beta^\top X^\top X\beta = y^\top X^\top X y - 2y^\top X\beta + \beta^\top \beta$$
$$= \|X^\top y - \beta\|_2^2 = \sum_{i \in N} (x_i^\top y - \beta_i)^2,$$

where x_i is the ith column of X. Thus, β^* optimizes a separable quadratic function over the region defined by $\Omega(\beta)$. This allows us to prove each part of the lemma as follows.

1. Each of the regions defined by the constraints $\|\beta\|_1 \leq t$ and $\|\beta\|_\infty \leq t$ can be reformulated as a bisubmodular polyhedron $\tilde{B}(h)$, where $h(X, Y) = t$ and $h(X, Y) = t|X \cup Y|$ for all $(X, Y) \in 3^N$, respectively. In both cases, it follows from Lemma 13 that β^* is the (unique) least weakly absolutely $(1, X^\top y)$-majorized element of $\tilde{B}(h)$, and the result follows from the characterization in Theorem 7.

2. The region defined by the constraint $\|\beta\|_2 \leq t$ is not polyhedral and thus, cannot be a base or bisubmodular polyhedron. Moreover, it cannot simultaneously be contained in a sub- or supermodular polyhedron and contain the corresponding base polyhedron. The result follows from the characterizations in Theorems 1, 3, and 7. □

The objective function of Problem (13) closely resembles that of a regression problem with Schur-convex loss function. An interesting question, which we leave for future work, is whether Theorem 10 can be extended to such problems.

7. Conclusions and Directions for Future Research

In this article, we studied the existence of solutions to optimization problems that optimize whole classes of utility or cost functions simultaneously. In particular, motivated by applications in power management, we aimed to characterize the set of problems for which such solutions, called least-majorized elements, exist. To answer this question, we introduced a new natural generalization of majorization that is determined by two input vectors a and b, called (a, b)-majorization, and the corresponding least (a, b)-majorized elements. We showed that such elements exist for any valid choice of a and b if and only if the feasible set of the optimization problem is a base polyhedron. Similar characterizations were obtained for weaker concepts of least (a, b)-majorized elements and sets related to base polyhedral, such as submodular, supermodular, and bisubmodular polyhedra. On the one hand, these characterizations reveal new and unique properties of these polyhedra. On the other hand, our results suggest that for most optimization problems arising in applications, no solutions exist that simultaneously optimize classes of objective functions. Although this observation is usually easy to confirm empirically for a given optimization problem, we now provide theoretical insight into why this is the case.

Given an optimization problem whose feasible set does not fall in one of the classes described, the perhaps next best thing that one might hope for is the existence of solutions that simultaneously *approximately minimize* entire classes of objective functions, instead of *minimize*. Therefore, our main direction of future research will be to investigate the former topic further. More specifically, we will try to find a complete characterization of problems that do have such simultaneously approximately minimizing solutions, parametrized by the desired error factor or term of approximation. Based on existing work in this direction (e.g., Goel and Meyerson [27]), we expect that such a characterization will include very general classes of resource allocation problems and will thus be useful in many applications in telecommunications and energy-efficient scheduling.

We conclude this article by listing three other directions for future research.

1. One limitation of this work is that we focused on the case where least (a, b)-majorized elements exist for *each* choice of a and b. Moreover, the definition of these elements requires them to be optimal for *any* choice of separable convex objective function. Thus, it would be interesting to investigate whether the existence of least (a, b)-majorized elements for specific (sets of) values of a and b can be characterized. In particular, an interesting question is whether such a characterization exists for $b = 0$ (i.e., when we only consider scaled objective functions).

2. Another limitation is that we focused on the case where the feasible sets are either compact and convex or bounded and hole free (i.e., contain exactly all integral points in its convex hull). An interesting direction for future research is to investigate whether our results can be extended to feasible sets that do not have such a convexity property. Promising candidates for such sets and corresponding optimization problems include extensions of base polyhedra and RAPs (e.g., jump systems (Bouchet and Cunningham [8]) and semicontinuous knapsack polytopes (de Farias and Zhao [12], Schoot Uiterkamp et al. [43])). Another promising class of problems consists of those

whose structure and optimal solutions depend primarily on the convexity of the objective function rather than explicitly on the function itself (e.g., discrete speed scaling problems (Gerards et al. [26])).

3. Our results indicate a natural connection between (a, b)-majorization and base polyhedra and between special cases of (a, b)-majorization and generalizations of base polyhedra. A natural follow-up question is whether more of such pairs exist. A first class of suitable candidates for sets in such a pair is polyhedra that are in some sense obtained from submodular or base polyhedra, examples of which not already discussed in this article are polybasic polyhedra (Fujishige et al. [25]), skew-bisubmodular polyhedra (Fujishige et al. [24]), and k-submodular polyhedra (Huber and Kolmogorov [30]). This is because all our characterization results depend on the existence of least (a, b)-majorized elements for base polyhedra.

Another way to identify promising candidate sets are those for which a greedy algorithm in the style of Edmonds [16] is optimal for linear optimization. This can be seen by considering the role of linear optimization in the conditional gradient method for solving convex optimization problems with continuously differentiable objective functions (see also Bach [5], Jaggi [32]). It can be shown that, when optimizing the function $\sum_{i\in N} a_i \phi\left(\frac{x_i+b_i}{a_i}\right)$ over a base (bisubmodular) polyhedron for given vectors a and b, there exists a sequence of iterate solutions that is a valid possible outcome of the classical Frank–Wolfe algorithm (Frank and Wolfe [20]) for any choice of continuously differentiable (even) convex function ϕ. We expect that similar analyses can be done for polyhedra for which similar greedy algorithms are optimal for linear optimization, such as (generalized) skew-bisubmodular polyhedra (Fujishige et al. [24]) and polyhedra defined by Monge (Burkard et al. [9]) and greedy (Faigle et al. [17]) matrices.

Acknowledgments

The author thank the editors and reviewers for their thorough assessment of the manuscript and for their helpful suggestions and feedback. Parts of this research were conducted while the author was affiliated with Eindhoven University of Technology.

Endnotes

[1] Note that if $(\tilde{X}, \tilde{Y})$ is an orthant (meaning $(S, T) = (\tilde{X}, \tilde{Y})$), then $x \in B_{(S,T)}(h)$. Thus, we already know that $x \in \tilde{B}(h)$ because $B_{(S,T)}(h) \subseteq \tilde{B}(h)$. However, it is not guaranteed that $(\tilde{X}, \tilde{Y})$ is an orthant. Hence, an extended proof is necessary.

[2] Here, we slightly abuse the exact definition of base polyhedra as introduced in Section 2.2 because g and v are not assumed to be supermodular at this point.

References

[1] Ando K (1996) Weak majorization on finite jump systems. Technical report, Institute of Socio-Economic Planning, University of Tsukuba, Tsukuba, Japan.
[2] Ando K, Fujishige S (1996) On structures of bisubmodular polyhedra. *Math. Programming* 74(3):293–317.
[3] Arin J, Kuipers J, Vermeulen D (2003) Some characterizations of egalitarian solutions on classes of TU-games. *Math. Soc. Sci.* 46(3):327–345.
[4] Arnold BC, Sarabia JM (2018) *Majorization and the Lorenz Order with Applications in Applied Mathematics and Economics*, 1st ed. (Springer, Cham, Switzerland).
[5] Bach F (2013) Learning with submodular functions: A convex optimization perspective. *Foundations Trends Machine Learn.* 6(2–3):145–373.
[6] Bailey RA (2008) *Design of Comparative Experiments*, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, Cambridge, UK).
[7] Bilbao JM, Fernández JR, Jiménez N, López JJ (2008) A survey of bicooperative games. Chinchuluun A, Pardalos PM, Migdalas A, Pitsoulis L, eds. *Pareto Optimality, Game Theory and Equilibria*, Springer Optimization and Its Applications, vol. 17 (Springer, New York), 187–216.
[8] Bouchet A, Cunningham WH (1995) Delta-matroids, jump systems, and bisubmodular polyhedra. *SIAM J. Discrete Math.* 8(1):17–32.
[9] Burkard RE, Klinz B, Rudolf R (1996) Perspectives of Monge properties in optimization. *Discrete Appl. Math.* 70(2):95–161.
[10] Chakrabarty D, Goel G, Vazirani VV, Wang L, Yu C (2014) Submodularity helps in Nash and nonsymmetric bargaining games. *SIAM J. Discrete Math.* 28(1):99–115.
[11] Dai YH, Fletcher R (2006) New algorithms for singly linearly constrained quadratic programs subject to lower and upper bounds. *Math. Programming* 106(3):403–421.
[12] de Farias IR, Zhao M (2013) A polyhedral study of the semi-continuous knapsack problem. *Math. Programming* 142(1–2):169–203.
[13] Dunstan FDJ, Welsh DJA (1973) A greedy algorithm for solving a certain class of linear programmes. *Math. Programming* 5(1):338–353.
[14] Dutta B, Ray D (1989) A concept of egalitarianism under participation constraints. *Econometrica* 57(3):615–635.
[15] Eaton ML (1984) On group induced orderings, monotone functions, and convolution theorems. Tong YL, ed. *Inequalities in Statistics and Probability*, IMS Lecture Notes—Monographic Series, vol. 5 (Institute of Mathematical Statistics, Hayward, CA), 13–25.
[16] Edmonds J (2003) Submodular functions, matroids, and certain polyhedra. Jünger M, Reinelt G, Rinaldi G, eds. *Combinatorial Optimization—Eureka, You Shrink!* (Springer, Berlin), 11–26.

[17] Faigle U, Kern W, Peis B (2011) On greedy and submodular matrices. Marchetti-Spaccamela A, Segal M, eds. *Theory and Practice of Algorithms in (Computer) Systems* (Springer, Berlin), 116–126.
[18] Francis AR, Wynn HP (2014) Subgroup majorization. *Linear Algebra Appl.* 444:53–66.
[19] Frank A, Murota K (2020) Discrete decreasing minimization, part II: Views from discrete convex analysis. Preprint, submitted June 30, https://arxiv.org/abs/1808.08477.
[20] Frank M, Wolfe P (1956) An algorithm for quadratic programming. *Naval Res. Logist. Quart.* 3(1–2):95–110.
[21] Fujishige S (1997) A min–max theorem for bisubmodular polyhedra. *SIAM J. Discrete Math.* 10(2):294–308.
[22] Fujishige S (2005) *Submodular Functions and Optimization*, 2nd ed. (Elsevier, Amsterdam).
[23] Fujishige S, Tomizawa N (1983) A note on submodular functions on distributive lattices. *J. Oper. Res. Soc. Japan* 26(4):309–318.
[24] Fujishige S, Tanigawa SI, Yoshida Y (2014) Generalized skew bisubmodularity: A characterization and a min–max theorem. *Discrete Optim.* 12:1–9.
[25] Fujishige S, Makino K, Takabatake T, Kashiwabara K (2004) Polybasic polyhedra: Structure of polyhedra with edge vectors of support size at most 2. *Discrete Math* 280(1):13–27.
[26] Gerards MET, Hurink JL, Hölzenspies PKF (2016) A survey of offline algorithms for energy minimization under deadline constraints. *J. Scheduling* 19(1):3–19.
[27] Goel A, Meyerson A (2006) Simultaneous optimization via approximate majorization for concave profits or convex costs. *Algorithmica* 44(4):301–323.
[28] Guo Y, Fang Y (2013) Electricity cost saving strategy in data centers by using energy storage. *IEEE Trans. Parallel Distributed Systems* 24(6):1149–1160.
[29] Hastie T, Tibshirani R, Friedman J (2009) *The Elements of Statistical Learning*, 2nd ed. (Springer, New York).
[30] Huber A, Kolmogorov V (2012) Toward minimizing k-submodular functions. Mahjoub AR, Markakis V, Milis I, Paschos VT, eds. *Combinatorial Optimization (ISCO 2012)*, Lecture Notes in Computer Science, vol. 7422 (Springer, Berlin), 451–462.
[31] Ichiishi T (1981) Super-modularity: Applications to convex games and to the greedy algorithm for LP. *J. Econom. Theory* 25(2):283–286.
[32] Jaggi M (2013) Revisiting Frank-Wolfe: Projection-free sparse convex optimization. Dasgupta S, McAllester D, eds. *Proc. 30th Internat. Conf. Machine Learn.*, vol. 28 (PMLR, New York), 427–435.
[33] Joe H (1990) Majorization and divergence. *J. Math Anal. Appl.* 148(2):287–305.
[34] Jorswieck E, Boche H (2006) Majorization and matrix-monotone functions in wireless communications. *Foundations Trends Comm. Inform. Theory* 3(6):553–701.
[35] Mairal J, Jenatton R, Obozinski G, Bach F (2011) Convex and network flow optimization for structured sparsity. *J. Machine Learn. Res.* 12(81):2681–2720.
[36] Marshall AW, Olkin I, Arnold BC (2011) *Inequalities: Theory of Majorization and Its Applications*, 2nd ed. (Springer, New York).
[37] Nakamura M (1988) A characterization of greedy sets: Universal polymatroids. *Sci. Papers College Arts Sci. Univ. Tokyo* 38(2):155–167.
[38] Niezgoda M (2012) Cone orderings, group majorizations and similarly separable vectors. *Linear Algebra Appl.* 436(3):579–594.
[39] Palomar DP, Fonollosa JR (2005) Practical algorithms for a family of waterfilling solutions. *IEEE Trans. Signal Processing* 53(2):686–695.
[40] Patriksson M (2008) A survey on the continuous nonlinear resource allocation problem. *Eur. J. Oper. Res.* 185(1):1–46.
[41] Reijnders VMJJ, Gerards MET, Hurink JL (2022) A hybrid electricity pricing mechanism for joint system optimization and social acceptance within energy communities. *Energy Rep.* 8:13281–13292.
[42] Schoot Uiterkamp MHH, Gerards MET, Hurink JL (2022) On a reduction for a class of resource allocation problems. *INFORMS J. Comput.* 34(3):1387–1402.
[43] Schoot Uiterkamp MHH, van der Klauw T, Gerards MET, Hurink JL (2018) Offline and online scheduling of electric vehicle charging with a minimum charging threshold. *2018 IEEE Internat. Conf. Comm. Control Comput. Tech. Smart Grids (SmartGridComm)* (IEEE, Piscataway, NJ).
[44] Shams F, Bacci G, Luise M (2014) A survey on resource allocation techniques in OFDM(A) networks. *Comput. Networks* 65:129–150.
[45] Tamir A (1995) Least majorized elements and generalized polymatroids. *Math Oper. Res.* 20(3):583–589.
[46] van der Klauw T, Gerards MET, Hurink JL (2017) Resource allocation problems in decentralized energy management. *OR Spectrum* 39(3):749–773.
[47] Veinott AF (1971) Least d-majorized network flows with inventory and statistical applications. *Management Sci.* 17(9):547–567.
[48] Zhan P (2005) Polyhedra and optimization related to a weak absolute majorization ordering. *J. Oper. Res. Soc. Japan* 48(2):90–96.
[49] Zurovac J, Brown R (2012) Orthogonal design: A powerful method for comparative effectiveness research with multiple interventions. Technical report, Mathematica Policy Research, Princeton, NJ.