RIMS-1781

Generalized Skew Bisubmodularity: A Characterization and a Min-Max Theorem

By

Satoru FUJISHIGE, Shin-ichi TANIGAWA and Yuichi YOSHIDA

May 2013

京都大学　数理解析研究所

RESEARCH INSTITUTE FOR MATHEMATICAL SCIENCES

KYOTO UNIVERSITY, Kyoto, Japan

Generalized Skew Bisubmodularity: A Characterization and a Min-Max Theorem

Satoru Fujishige[*] Shin-ichi Tanigawa[†] Yuichi Yoshida[‡]

May 15, 2013

Abstract

Huber, Krokhin, and Powell (Proc. SODA2013) introduced a concept of skew bisubmodularity, as a generalization of bisubmodularity, in their complexity dichotomy theorem for valued constraint satisfaction problems over the three-value domain. In this paper we consider a natural generalization of the concept of skew bisubmodularity and show a connection between the generalized skew bisubmodularity and a convex extension over rectangles. We also analyze the dual polyhedra, called skew bisubmodular polyhedra, associated with generalized skew bisubmodular functions and derive a min-max theorem that characterizes the minimum value of a generalized skew bisubmodular function in terms of a minimum-norm point in the associated skew bisubmodular polyhedron.

1 Introduction

For a finite set V let 2^V be the set of all subsets of V and 3^V be the set of all the ordered pairs of disjoint subsets of V. A function $f : 3^V \to \mathbb{R}$ is called *bisubmodular* if

$$f(X_+, X_-) + f(Y_+, Y_-) \geq f(X_+ \cap Y_+, X_- \cap Y_-) + f((X_+ \cup Y_+) \setminus (X_- \cup Y_-), (X_- \cup Y_-) \setminus (X_+ \cup Y_+))$$

for all $(X_+, X_-), (Y_+, Y_-) \in 3^V$. The concept of bisubmodularity was introduced in the study of Δ-matroids by Bouchet [3] and independently by Chandrasekaran–Kabadi [5] (also see [6, 1]). Examples of Δ-matroids include the base family of a matroid as well as the family of matchable vertex sets in a graph, and bisubmodularity plays an important rôle in combinatorial optimization for establishing the common generalization of matroid theory and matching theory from the optimization view point (see, e.g., [4]).

Bisubmodularity generalizes the well-known concept of submodularity. A function $f : 2^V \to \mathbb{R}$ is called *submodular* if

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$$

[*]Research Institute for Mathematical Sciences, Kyoto University, `fujishig@kurims.kyoto-u.ac.jp`, Supported by JSPS Grant-in-Aid for Scientific Research (B) 25280004.

[†]Research Institute for Mathematical Sciences, Kyoto University, `tanigawa@kurims.kyoto-u.ac.jp`, Supported by JSPS Grant-in-Aid for Scientific Research (B) 25280004.

[‡]National Institute of Informatics, and Preferred Infrastructure, Inc., `yyoshida@nii.ac.jp`, Supported by JSPS Grant-in-Aid for Research Activity Start-up (24800082), MEXT Grant-in-Aid for Scientific Research on Innovative Areas (24106001), and JST, ERATO, Kawarabayashi Large Graph Project.

for all $X, Y \in 2^V$. The *Lovász extension* $\widehat{f}$ (or the *Choquet integral*) of a submodular function $f : 2^V \to \mathbb{R}$ is a convex extension over $[0,1]^V$, which plays a fundamental rôle in minimizing submodular functions as well as generalizing the submodular analysis to discrete convex analysis. In fact Grötschel, Lovász, and Schrijver [13, Chapter 10] pointed out that one can minimize f by applying the ellipsoid method to $\widehat{f}$, which led to the first weakly and strongly polynomial-time algorithms for minimizing submodular functions [12, 13]. Later, Iwata, Fleischer, and Fujishige [18] and Schrijver [22] independently gave combinatorial, strongly polynomial-time algorithms for minimizing submodular functions.

Algorithms for bisubmodular function minimization showed a similar historical development following submodular function minimization. Qi [21] proposed a convex extension of a bisubmodular function over $[-1,1]^V$ and adapted the argument of Grötschel, Lovász, and Schrijver [13] to bisubmodular functions. Fujishige and Iwata [10] extended their submodular function minimization algorithm to bisubmodular function minimization. The time complexity of their algorithm is not strongly polynomial, but later a combinatorial, strongly polynomial-time algorithm was developed by McCormick and Fujishige [20].

Huber, Krokhin, and Powell [17] introduced a generalization of bisubmodularity, called skew bisubmodularity, in their complexity dichotomy theorem for the valued constraint satisfaction problems (VCSPs) over the three-value domain. Let α be a number with $0 < \alpha \leq 1$. A function $f : 3^V \to \mathbb{R}$ is called *α-bisubmodular* if, for every $\mathbf{X} = (X_+, X_-)$ and $\mathbf{Y} = (Y_+, Y_-) \in 3^V$,

$$f(\mathbf{X}) + f(\mathbf{Y}) \geq f(\mathbf{X} \cap \mathbf{Y}) + \alpha f(\mathbf{X} \cup_0 \mathbf{Y}) + (1-\alpha) f(\mathbf{X} \cup_1 \mathbf{Y}), \tag{1}$$

where $\mathbf{X} \cap \mathbf{Y} = (X_+ \cap Y_+, X_- \cap Y_-)$, $\mathbf{X} \cup_0 \mathbf{Y} = ((X_+ \cup Y_+) \setminus (X_- \cup Y_-), (X_- \cup Y_-) \setminus (X_+ \cup Y_+))$, and $\mathbf{X} \cup_1 \mathbf{Y} = (X_+ \cup Y_+, (X_- \cup Y_-) \setminus (X_+ \cup Y_+))$. 1-bisubmodularity is nothing but bisubmodularity. A function $f : 3^V \to \mathbb{R}$ is called *skew bisubmodular* if it is α-bisubmodular for some $\alpha \in (0,1]$. It was left open in the proceedings paper [17] to decide whether α-bisubmodular functions could be minimized in polynomial time for any $\alpha \in (0,1)$ in the value oracle model, but very recently we have been informed that Huber and Krokhin [16] showed that the minimization problem is indeed tractable via a convex extension.[1]

In this paper we introduce a further natural generalization of the concept of skew bisubmodularity, and reveal the importance of (generalized) skew bisubmodularity from the point of view of discrete convex analysis. We examine an analog of the Lovász extension over general n-dimensional rectangles and show that a necessary and sufficient condition for such an extension to be convex is the generalized skew bisubmodularity, where α-bisubmodularity introduced in [17] shows up as a special case when the rectangle is of form $[-\alpha, 1]^V$. This implies that the generalized skew bisubmodular functions can also be minimized in strongly polynomial time by the ellipsoid method. We also analyze the dual polyhedra, called *skew bisubmodular polyhedra*, associated with skew bisubmodular functions. It turns out that each orthant of a skew bisubmodular polyhedron forms a submodular polyhedron scaled by parameters, and skew bisubmodular polyhedra are special cases of *polybasic polyhedra* examined by Fujishige, Makino, Takabatake, and Kashiwabara [11]. Also skew bisubmodularity can be viewed as a special case of the discrete convexity defined within the general framework recently developed by Hirai [14, 15], while his general framework does not directly imply the oracle tractability of skew bisubmodular function minimization.

[1] The oracle tractability was announced at the Dagstuhl Seminar in November 2012 (see the slides of Anna Huber: *VCSPs on Three Elements.* Seminar 12451 on "The Constraint Satisfaction Problem: Complexity and Approximability").

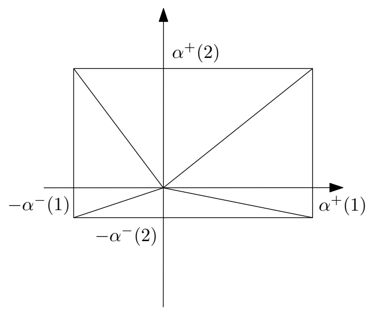

Figure 1: The simplicial division of $[-\boldsymbol{\alpha}^-, \boldsymbol{\alpha}^+]$ for $n = 2$.

Throughout the present paper we sometimes use bold-faced capital letters to denote elements in 3^V. For $(X_+, X_-) \in 3^V$, for example, we use the bold-faced $\mathbf{X}$ to designate the pair (X_+, X_-) and we define $(\mathbf{X})_+ = X_+$ and $(\mathbf{X})_- = X_-$. We adopt this convention for other letters as well. By $\mathbf{X} \subseteq \mathbf{Y}$ we mean $X_+ \subseteq Y_+$ and $X_- \subseteq Y_-$, and by $\mathbf{X} \subset \mathbf{Y}$ we mean $\mathbf{X} \subseteq \mathbf{Y}$ and $\mathbf{X} \neq \mathbf{Y}$.

For any $X \subseteq V$, χ_X denotes the characteristic vector of X in $\mathbb{R}^V$.

If $f(\emptyset, \emptyset) \neq 0$, one can apply arguments to $f - f(\emptyset, \emptyset)$ instead of f and derive the corresponding statements, so that we assume in the sequel that any function $f : 3^V \to \mathbb{R}$ satisfies $f(\emptyset, \emptyset) = 0$.

2 A Generalization of Skew Bisubmodularity

In this section we shall introduce an extension $\widehat{f}$ of a function $f : 3^V \to \mathbb{R}$ over rectangles in Section 2.1 and then introduce generalized skew bisubmodular functions in Section 2.2. A relation between these two concepts is clarified in Section 3.

2.1 A simplicial division and an extension

For a finite set V of n elements let $\boldsymbol{\alpha} = (\boldsymbol{\alpha}^+, \boldsymbol{\alpha}^-)$ be a pair of positive vectors $\boldsymbol{\alpha}^+, \boldsymbol{\alpha}^- : V \to \mathbb{R}_{>0}$ and consider the n-dimensional rectangle $[-\boldsymbol{\alpha}^-, \boldsymbol{\alpha}^+] = \{x \in \mathbb{R}^V \mid -\boldsymbol{\alpha}^- \leq x \leq \boldsymbol{\alpha}^+\}$.

For any $\mathbf{X} \in 3^V$ define

$$\chi^{\boldsymbol{\alpha}}_{\mathbf{X}} = \sum_{v \in X_+} \boldsymbol{\alpha}^+(v) \chi_{\{v\}} - \sum_{v \in X_-} \boldsymbol{\alpha}^-(v) \chi_{\{v\}}. \tag{2}$$

Then, for each chain $\mathbf{A}_1 \subset \cdots \subset \mathbf{A}_k$ in 3^V the convex hull of $\{\chi^{\boldsymbol{\alpha}}_{\mathbf{A}_i} \mid 1 \leq i \leq k\}$ is a simplex and such simplices for all the maximal chains induce a simplicial division of rectangle $[-\boldsymbol{\alpha}^-, \boldsymbol{\alpha}^+]$. See Figure 1 for a two-dimensional example. This leads us to the the following essential fact.

Proposition 1. *For any $c \in \mathbb{R}^V \setminus \{\mathbf{0}\}$, there uniquely exist a chain $(\emptyset, \emptyset) \neq \mathbf{A}_1 \subset \cdots \subset \mathbf{A}_k$ and coefficients $\lambda_1, \ldots, \lambda_k \in \mathbb{R}_{>0}$ such that*

$$c = \sum_{i=1}^{k} \lambda_i \chi^{\boldsymbol{\alpha}}_{\mathbf{A}_i}. \tag{3}$$

By using the unique chain $\mathbf{A}_1 \subset \mathbf{A}_2 \subset \cdots \subset \mathbf{A}_k$ and coefficients $\lambda_1, \ldots, \lambda_k$ appearing in (3) for $c \in \mathbb{R}^V \setminus \{\mathbf{0}\}$, we define an extension $\widehat{f} : \mathbb{R}^V \to \mathbb{R}$ of a function $f : 3^V \to \mathbb{R}$ by

$$\widehat{f}(c) = \sum_{i=1}^{k} \lambda_i f(\mathbf{A}_i) \qquad (c \in \mathbb{R}^V \setminus \{\mathbf{0}\}) \tag{4}$$

and $\widehat{f}(\mathbf{0}) = f(\emptyset, \emptyset) = 0$.

2.2 Generalized skew bisubmodular functions

The key observation to analyze $\widehat{f}$ is a modular equation among the scaled characteristic vectors $\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}$. This relation can be derived by checking how $c \equiv \chi^{\boldsymbol{\alpha}}_{\mathbf{X}} + \chi^{\boldsymbol{\alpha}}_{\mathbf{Y}}$ can be expressed in the form of (3) for $\mathbf{X}, \mathbf{Y} \in 3^V$, i.e., we shall compute $\lambda_1, \ldots, \lambda_k$ and $\mathbf{A}_1 \subset \cdots \subset \mathbf{A}_k$ for $\chi^{\boldsymbol{\alpha}}_{\mathbf{X}} + \chi^{\boldsymbol{\alpha}}_{\mathbf{Y}}$. The chain and coefficients can be written by an explicit formula by using binary operations $\cup_t$ on 3^V for $t \in (0, 1)$ defined as follows: For each $t \in (0, 1)$ define

- $\mathbf{V}_t = (V_{t,+}, V_{t,-}) \in 3^V$ by

$$V_{t,+} = \left\{ v \in V \;\middle|\; \frac{\boldsymbol{\alpha}^-(v)}{\boldsymbol{\alpha}^+(v)} \leq t \right\}, \qquad V_{t,-} = \left\{ v \in V \;\middle|\; \frac{\boldsymbol{\alpha}^+(v)}{\boldsymbol{\alpha}^-(v)} \leq t \right\}$$

- and a binary operation $\cup_t$ on 3^V by

$$\begin{aligned}(\mathbf{X} \cup_t \mathbf{Y})_+ &= (\mathbf{X} \cup_0 \mathbf{Y})_+ \cup (V_{t,+} \cap (X_+ \cup Y_+) \cap (X_- \cup Y_-)),\\ (\mathbf{X} \cup_t \mathbf{Y})_- &= (\mathbf{X} \cup_0 \mathbf{Y})_- \cup (V_{t,-} \cap (X_+ \cup Y_+) \cap (X_- \cup Y_-)).\end{aligned}$$

Example. If $V = \{1, 2, 3, 4\}$, $\frac{\boldsymbol{\alpha}^+(1)}{\boldsymbol{\alpha}^-(1)} = \frac{2}{3}$, $\frac{\boldsymbol{\alpha}^+(2)}{\boldsymbol{\alpha}^-(2)} = \frac{1}{3}$, $\frac{\boldsymbol{\alpha}^-(3)}{\boldsymbol{\alpha}^+(3)} = \frac{2}{3}$, and $\frac{\boldsymbol{\alpha}^-(4)}{\boldsymbol{\alpha}^+(4)} = \frac{1}{2}$, then $\mathbf{V}_{\frac{1}{3}} = (\emptyset, \{2\})$, $\mathbf{V}_{\frac{1}{2}} = (\{4\}, \{2\})$, $\mathbf{V}_{\frac{2}{3}} = (\{3, 4\}, \{1, 2\})$, and $(\{1, 3\}, \{2, 4\}) \cup_0 (\{2, 4\}, \{3\}) = (\{1\}, \emptyset)$. We have $(\{1, 3\}, \{2, 4\}) \cup_{\frac{1}{3}} (\{2, 4\}, \{3\}) = (\{1\}, \{2\})$, $(\{1, 3\}, \{2, 4\}) \cup_{\frac{1}{2}} (\{2, 4\}, \{3\}) = (\{1, 4\}, \{2\})$, and $(\{1, 3\}, \{2, 4\}) \cup_{\frac{2}{3}} (\{2, 4\}, \{3\}) = (\{1, 3, 4\}, \{2\})$. $\square$

Using $\cup_0$ and $\cup_1$ defined in Section 1, we have defined binary operations $\cup_t$ for all $t \in [0, 1]$. Note that $\mathbf{V}_t \subseteq \mathbf{V}_{t'}$ if $t \leq t'$ and that these binary operations $\cup_t$ are determined once we fix $\boldsymbol{\alpha}$.

We now have the following.

Lemma 2. *For given V and $\boldsymbol{\alpha}$, define a set $T = \left\{ \min\left\{ \frac{\boldsymbol{\alpha}^-(v)}{\boldsymbol{\alpha}^+(v)}, \frac{\boldsymbol{\alpha}^+(v)}{\boldsymbol{\alpha}^-(v)} \right\} \middle| v \in V \right\} \cup \{0, 1\}$ and arrange the distinct elements of T in the increasing order of magnitude as $0 = t_0 < t_1 < t_2 < \cdots < t_{k+1} = 1$. Then we have*

$$\chi^{\boldsymbol{\alpha}}_{\mathbf{X}} + \chi^{\boldsymbol{\alpha}}_{\mathbf{Y}} = \chi^{\boldsymbol{\alpha}}_{\mathbf{X} \cap \mathbf{Y}} + \sum_{i=0}^{k} (t_{i+1} - t_i) \chi^{\boldsymbol{\alpha}}_{\mathbf{X} \cup_{t_i} \mathbf{Y}}. \tag{5}$$

Proof. Denote the vector on the left-hand side of (5) by LH and that on the right-hand side by RH. We show $LH(v) = RH(v)$ for all $v \in V$.

Choose any $v \in V$.

(I) If $v \notin X_+ \cup X_- \cup Y_+ \cup Y_-$, then we have $LH(v) = 0 = RH(v)$.

(II) If $v \in X_+ \cap Y_+$, then $LH(v) = 2\boldsymbol{\alpha}^+(v)$. Since $v \in X_+ \cap Y_+$ and $v \in (\mathbf{X} \cup_0 \mathbf{Y})_+ \subseteq (\mathbf{X} \cup_{t_i} \mathbf{Y})_+$ for all i, we also have $RH(v) = 2\boldsymbol{\alpha}^+(v)$.

(III) If $v \in X_+ \setminus (Y_+ \cup Y_-)$, then $LH(v) = \boldsymbol{\alpha}^+(v)$. Since $v \notin (\mathbf{X} \cap \mathbf{Y})_+ \cup (\mathbf{X} \cap \mathbf{Y})_-$ and $v \in (\mathbf{X} \cup_0 \mathbf{Y})_+ \subseteq (\mathbf{X} \cup_{t_i} \mathbf{Y})_+$ for all i, we also have $RH(v) = \boldsymbol{\alpha}^+(v)$.

(IV) Because of the symmetry we assume that the remaining case is when $v \in X_+ \cap Y_-$. Then, $LH(v) = \boldsymbol{\alpha}^+(v) - \boldsymbol{\alpha}^-(v)$. Suppose that $\boldsymbol{\alpha}^+(v) \geq \boldsymbol{\alpha}^-(v)$. Then, $v \notin (\mathbf{X} \cup_t \mathbf{Y})_-$ for any $t \in [0, 1)$, and $v \in (\mathbf{X} \cup_t \mathbf{Y})_+$ if and only if $\frac{\boldsymbol{\alpha}^-(v)}{\boldsymbol{\alpha}^+(v)} \leq t$. By definition, there is an index j such that $t_j = \frac{\boldsymbol{\alpha}^-(v)}{\boldsymbol{\alpha}^+(v)}$. Since $v \notin (\mathbf{X} \cap \mathbf{Y})_+ \cup (\mathbf{X} \cap \mathbf{Y})_-$, we thus have $RH(v) = \sum_{i=0}^{k} (t_{i+1} - t_i) \chi^{\boldsymbol{\alpha}}_{\mathbf{X} \cup_{t_i} \mathbf{Y}}(v) = \sum_{i=j}^{k} (t_{i+1} - t_i) \chi^{\boldsymbol{\alpha}}_{\mathbf{X} \cup_{t_i} \mathbf{Y}}(v) = \sum_{i=j}^{k} (t_{i+1} - t_i) \boldsymbol{\alpha}^+(v) = (t_{k+1} - t_j)\boldsymbol{\alpha}^+(v) = \boldsymbol{\alpha}^+(v) - \boldsymbol{\alpha}^-(v) = LH(v)$. The same argument can also be applied to the case when $\boldsymbol{\alpha}^+(v) < \boldsymbol{\alpha}^-(v)$.

This completes the proof. $\square$

Motivated by Lemma 2, we say that a function $f : 3^V \to \mathbb{R}$ is $\boldsymbol{\alpha}$*-bisubmodular* if

$$f(\mathbf{X}) + f(\mathbf{Y}) \geq f(\mathbf{X} \cap \mathbf{Y}) + \sum_{i=0}^{k} (t_{i+1} - t_i) f(\mathbf{X} \cup_{t_i} \mathbf{Y}) \tag{6}$$

for all $\mathbf{X}, \mathbf{Y} \in 3^V$, where t_i $(i = 0, \ldots, k+1)$ are those defined in Lemma 2. When $\boldsymbol{\alpha}^+(v) = 1$ and $\boldsymbol{\alpha}^-(v) = \alpha$ for all $v \in V$ for some $\alpha \in (0, 1]$, $\boldsymbol{\alpha}$-bisubmodularity becomes α-bisubmodularity in [17] defined by (1).

3 Skew Bisubmodular Polyhedron and Convexity of $\widehat{f}$

Let $\boldsymbol{\alpha} = (\boldsymbol{\alpha}^+, \boldsymbol{\alpha}^-)$ with $\boldsymbol{\alpha}^+ : V \to \mathbb{R}_{>0}$ and $\boldsymbol{\alpha}^- : V \to \mathbb{R}_{>0}$. For any $x \in \mathbb{R}^V$ and $\mathbf{X} \in 3^V$ define $x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) = \sum_{v \in V} x(v) \chi^{\boldsymbol{\alpha}}_{\mathbf{X}}(v)$, which is the canonical inner product $\langle x, \chi^{\boldsymbol{\alpha}}_{\mathbf{X}} \rangle$ of x and $\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}$ in (2). Hence,

$$x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) = \sum_{v \in X_+} \boldsymbol{\alpha}^+(v) x(v) - \sum_{v \in X_-} \boldsymbol{\alpha}^-(v) x(v). \tag{7}$$

Also define the $\boldsymbol{\alpha}$*-bisubmodular polyhedron* $\mathrm{P}(f)$ associated with an $\boldsymbol{\alpha}$-bisubmodular function f by

$$\mathrm{P}(f) = \{x \in \mathbb{R}^V \mid \forall \mathbf{X} \in 3^V : x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) \leq f(\mathbf{X})\}. \tag{8}$$

We show that $\widehat{f}$ defined by (4) is the support function of $\mathrm{P}(f)$, i.e., for any $c \in \mathbb{R}^V$, $\widehat{f}(c) = \max\{\langle c, x \rangle \mid x \in \mathrm{P}(f)\}$. This implies that $\boldsymbol{\alpha}$-bisubmodularity is a necessary and sufficient condition for the convexity of $\widehat{f}$ (Theorem 7 shown below). The argument given here is essentially an adaptation of bisubmodular analysis given in [9].

Let us proceed to the detailed description. For any given $c \in \mathbb{R}^V$ consider the following linear programming problem.

$$\begin{array}{lll} \text{(P)} & \text{Maximize} & \langle c, x \rangle \\ & \text{subject to} & x \in \mathrm{P}(f). \end{array}$$

To show that a dual optimal solution of this problem forms a chain, we first consider a relaxation of the system of linear inequalities defining $\mathrm{P}(f)$ in (8).

A pair $\mathbf{S} = (S_+, S_-) \in 3^V$ is called an *orthant* if $S_+ \cup S_- = V$. The set of all the pairs $\mathbf{X} = (X_+, X_-)$ such that $X_+ \subseteq S_+$ and $X_- \subseteq S_-$ is denoted by $2^{\mathbf{S}}$. We define a superset $\mathrm{P}_{\mathbf{S}}(f)$ of $\mathrm{P}(f)$ by

$$\mathrm{P}_{\mathbf{S}}(f) = \{x \in \mathbb{R}^V \mid \forall \mathbf{X} \in 2^{\mathbf{S}} \colon x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) \le f(\mathbf{X})\},$$

which is obtained from $\mathrm{P}(f)$ by discarding constraints not related to $2^{\mathbf{S}}$.

The advantage of introducing orthants is that the maximization over $\mathrm{P}_{\mathbf{S}}(f)$ is equivalent to the maximization over a submodular polyhedron. Let us explain this fact now. Notice that, once we fix an orthant $\mathbf{S}$, f becomes submodular on $2^{\mathbf{S}}$. In other words, by defining $f_{\mathbf{S}} : 2^V \to \mathbb{R}$ by

$$f_{\mathbf{S}}(X) = f(S_+ \cap X, S_- \cap X) \qquad (X \subseteq V),$$

$f_{\mathbf{S}}$ is submodular on 2^V. Consider the submodular polyhedron $\mathrm{P}(f_{\mathbf{S}})$, which is given by

$$\mathrm{P}(f_{\mathbf{S}}) = \{x \in \mathbb{R}^V \mid \forall X \subseteq V \colon x(X) \le f_{\mathbf{S}}(X)\}.$$

Then, observe

$$\mathrm{P}_{\mathbf{S}}(f) = \{x \in \mathbb{R}^V \mid \exists y \in \mathrm{P}(f_{\mathbf{S}}), \forall v \in S_+ \colon \boldsymbol{\alpha}^+(v)x(v) = y(v), \forall v \in S_- \colon -\boldsymbol{\alpha}^-(v)x(v) = y(v)\}.$$

This implies that $\mathrm{P}_{\mathbf{S}}(f)$ can be obtained from $\mathrm{P}(f_{\mathbf{S}})$ by reflections and scaling along axes, and $\mathrm{P}_{\mathbf{S}}(f)$ is combinatorially equivalent to $\mathrm{P}(f_{\mathbf{S}})$. Recall that a greedy algorithm solves the maximization problem over any submodular polyhedron (see [7, 9]). In terms of $\mathrm{P}_{\mathbf{S}}(f)$ we obtain a variant of the greedy algorithm, Greedy_Algorithm, which actually computes an optimal solution of (P) together with the relevant orthant $\mathbf{S}$ (see Theorem 5 shown below).

Greedy_Algorithm

Input: An $\boldsymbol{\alpha}$-bisubmodular function $f : 3^V \to \mathbb{R}_+$ on a finite set V, and a vector $c \in \mathbb{R}^V$.
Output: An optimal solution x^* of (P).

1: Compute an orthant $\mathbf{S} = (\{v \in V \mid c(v) \ge 0\}, \{v \in V \mid c(v) < 0\})$ and a vector $c_{\boldsymbol{\alpha}} \in \mathbb{R}^V$ by

$$c_{\boldsymbol{\alpha}}(v) = \begin{cases} \frac{c(v)}{\boldsymbol{\alpha}^+(v)} & \text{if } v \in S_+ \\ -\frac{c(v)}{\boldsymbol{\alpha}^-(v)} & \text{if } v \in S_- \end{cases} \qquad (v \in V). \tag{9}$$

2: Find a total ordering $L = (v_1, v_2, \ldots, v_n)$ of V such that $c_{\boldsymbol{\alpha}}(v_1) \ge c_{\boldsymbol{\alpha}}(v_2) \ge \cdots \ge c_{\boldsymbol{\alpha}}(v_n)$.
3: Compute a vector $x^* \in \mathbb{R}^V$ by

$$x^*(v_i) = \begin{cases} \frac{1}{\boldsymbol{\alpha}^+(v_i)}(f(\mathbf{X}_i) - f(\mathbf{X}_{i-1})) & \text{if } v_i \in S_+ \\ -\frac{1}{\boldsymbol{\alpha}^-(v_i)}(f(\mathbf{X}_i) - f(\mathbf{X}_{i-1})) & \text{if } v_i \in S_- \end{cases} \qquad (1 \le i \le n), \tag{10}$$

where $\mathbf{X}_i$ is the restriction of $\mathbf{S}$ to $\{v_1, \ldots, v_i\}$ and $\mathbf{X}_0 = (\emptyset, \emptyset)$.
4: Return x^*.

Proposition 3. *Let $f : 3^V \to \mathbb{R}$ be an $\boldsymbol{\alpha}$-bisubmodular function. For $c \in \mathbb{R}^V$, let x^* be the vector and $\mathbf{S}$ be the orthant computed by* Greedy_Algorithm*. Then x^* is an extreme point of*

$$\mathrm{B}_{\mathbf{S}}(f) := \{x \in \mathbb{R}^V \mid x \in \mathrm{P}_{\mathbf{S}}(f),\ x(\chi^{\boldsymbol{\alpha}}_{\mathbf{S}}) = f(\mathbf{S})\},$$

and $\langle c, x^* \rangle \ge \langle c, x \rangle$ for all $x \in \mathrm{P}_{\mathbf{S}}(f)$.

Following the argument in [9, Section 3.5(b)], we now show that x^* is indeed an optimal solution not only over $\mathrm{P}_{\mathbf{S}}(f)$ but also over $\mathrm{P}(f)$. To see this we need one more technical lemma, which is an analogue of [9, Lemma 3.60] for bisubmodular analysis.

Lemma 4. *Let $f : 3^V \to \mathbb{R}$ be an $\boldsymbol{\alpha}$-bisubmodular function. For each orthant $\mathbf{S} \in 3^V$ we have $\mathrm{B}_{\mathbf{S}}(f) \subseteq \mathrm{P}(f)$.*

Proof. Let $x \in \mathrm{B}_{\mathbf{S}}(f)$. It then follows from $\boldsymbol{\alpha}$-bisubmodularity of f and Lemma 2 that for any $\mathbf{X} \in 3^V$ we have

$$
\begin{aligned}
& x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) - f(\mathbf{X}) \\
& = x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) - f(\mathbf{X}) + x(\chi^{\boldsymbol{\alpha}}_{\mathbf{S}}) - f(\mathbf{S}) \qquad (\text{by } x \in \mathrm{B}(f)) \\
& \leq x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cap\mathbf{S}}) + \sum_{i=0}^{k}(t_{i+1} - t_i)x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cup_{t_i}\mathbf{S}}) - [f(\mathbf{X}\cap\mathbf{S}) + \sum_{i=0}^{k}(t_{i+1} - t_i)f(\mathbf{X}\cup_{t_i}\mathbf{S})] \\
& \leq 0 \qquad (\text{by } x \in \mathrm{P}(f)),
\end{aligned}
$$

which implies $x \in \mathrm{P}(f)$. □

Theorem 5. *Let $f : 3^V \to \mathbb{R}$ be an $\boldsymbol{\alpha}$-bisubmodular function. For $c \in \mathbb{R}^V$, let x^* be the vector obtained by* Greedy_Algorithm*. Then we have $\langle c, x^* \rangle \geq \langle c, x \rangle$ for all $x \in \mathrm{P}(f)$.*

Proof. Let $\mathbf{S} = (\{v \in V \mid c(v) \geq 0\}, \{v \in V \mid c(v) < 0\})$ be the orthant computed by Greedy_Algorithm. Note that $\mathrm{P}(f) \subseteq \mathrm{P}_{\mathbf{S}}(f)$. Combining this relation with Lemma 4, we have

$$\max\{\langle c, x \rangle \mid x \in \mathrm{P}_{\mathbf{S}}(f)\} \geq \max\{\langle c, x \rangle \mid x \in \mathrm{P}(f)\} \geq \max\{\langle c, x \rangle \mid x \in \mathrm{B}_{\mathbf{S}}(f)\}.$$

However, Proposition 3 implies $\max\{\langle c, x \rangle \mid x \in \mathrm{P}_{\mathbf{S}}(f)\} = \max\{\langle c, x \rangle \mid x \in \mathrm{B}_{\mathbf{S}}(f)\} = \langle c, x^* \rangle$. We thus have $\langle c, x^* \rangle \geq \langle c, x \rangle$ for any $x \in \mathrm{P}(f)$. □

Corollary 6. *Let $f : 3^V \to \mathbb{R}$ be an $\boldsymbol{\alpha}$-bisubmodular function. Then, for any $c \in \mathbb{R}^V$ we have*

$$\widehat{f}(c) = \max\{\langle c, x \rangle \mid x \in \mathrm{P}(f)\} \qquad (c \in \mathbb{R}^V). \tag{11}$$

Proof. Let vectors $c_{\boldsymbol{\alpha}}$ and x^* and chain $\mathbf{X}_1 \subset \mathbf{X}_2 \subset \cdots \subset \mathbf{X}_n$ be those computed by Greedy_Algorithm. Define λ_i by $\lambda_i = c_{\boldsymbol{\alpha}}(v_i) - c_{\boldsymbol{\alpha}}(v_{i+1})$ for $1 \leq i \leq k-1$ and by $\lambda_k = c_{\boldsymbol{\alpha}}(v_k)$. Then it can easily be checked that $\langle c, x^* \rangle = \sum_{i=1}^{n} \lambda_i f(\mathbf{X}_i)$ and $c = \sum_{i=1}^{n} \lambda_i \chi^{\boldsymbol{\alpha}}_{\mathbf{X}_i}$. Therefore, we obtain (11) because of Proposition 1, the definition of $\widehat{f}$, and Theorem 5. □

We now show a main theorem of this section. We remark that, from the definition of $\widehat{f}$ in (4), $\widehat{f}$ is *positively homogeneous* (i.e., $\widehat{f}(\lambda c) = \lambda \widehat{f}(c)$ for any $\lambda > 0$ and $c \in \mathbb{R}^V$).

Theorem 7. *Let V be a finite set and $\boldsymbol{\alpha} = (\boldsymbol{\alpha}^+, \boldsymbol{\alpha}^-)$ be a pair of vectors $\boldsymbol{\alpha}^+ : V \to \mathbb{R}_{>0}$ and $\boldsymbol{\alpha}^- : V \to \mathbb{R}_{>0}$. Then, for any $f : 3^V \to \mathbb{R}$, $\widehat{f}$ is convex if and only if f is $\boldsymbol{\alpha}$-bisubmodular.*

Proof. The proof is essentially the same as that of the corresponding theorem for submodular functions given in [9, Theorem 6.13]. Let us give it for the sake of completeness.

For each $c \in \mathbb{R}^V$, let x_c be a maximizer of the right-hand side of (11). Then, for any $c, c' \in \mathbb{R}^V$, we have $2\widehat{f}(\frac{c+c'}{2}) = \widehat{f}(c + c') = \langle c + c', x_{c+c'} \rangle \leq \langle c, x_c \rangle + \langle c', x_{c'} \rangle = \widehat{f}(c) + \widehat{f}(c')$, which implies the convexity of $\widehat{f}$.

Conversely, suppose that $\widehat{f}$ is convex. To show $\boldsymbol{\alpha}$-bisubmodularity of f, take any $\mathbf{X}$ and $\mathbf{Y}$ in 3^V. Since $\widehat{f}$ is positively homogeneous and convex, we have

$$\frac{\widehat{f}(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}} + \chi^{\boldsymbol{\alpha}}_{\mathbf{Y}})}{2} = \widehat{f}\left(\frac{\chi^{\boldsymbol{\alpha}}_{\mathbf{X}} + \chi^{\boldsymbol{\alpha}}_{\mathbf{Y}}}{2}\right) \leq \frac{\widehat{f}(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) + \widehat{f}(\chi^{\boldsymbol{\alpha}}_{\mathbf{Y}})}{2}. \tag{12}$$

On the other hand, since $\mathbf{X} \cap \mathbf{Y} \subseteq \mathbf{X} \cup_{t_0} \mathbf{Y} \subseteq \mathbf{X} \cup_{t_1} \mathbf{Y} \subseteq \cdots \subseteq \mathbf{X} \cup_{t_k} \mathbf{Y}$, it follows from the definition of $\widehat{f}$ in (4) that

$$\widehat{f}\left(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cap\mathbf{Y}} + \sum_{i=0}^{k}(t_{i+1} - t_i)\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cup_{t_i}\mathbf{Y}}\right) = \widehat{f}(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cap\mathbf{Y}}) + \sum_{i=0}^{k}\widehat{f}((t_{i+1} - t_i)\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cup_{t_i}\mathbf{Y}}). \tag{13}$$

Therefore,

$$\begin{aligned}
f(\mathbf{X}) + f(\mathbf{Y}) &= \widehat{f}(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) + \widehat{f}(\chi^{\boldsymbol{\alpha}}_{\mathbf{Y}}) \qquad \text{(since } \widehat{f} \text{ is an extension of } f\text{)} \\
&\geq \widehat{f}\,(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}} + \chi^{\boldsymbol{\alpha}}_{\mathbf{Y}}) \qquad \text{(by (12))} \\
&= \widehat{f}\left(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cap\mathbf{Y}} + \sum_{i=0}^{k}(t_{i+1} - t_i)\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cup_{t_i}\mathbf{Y}}\right) \qquad \text{(by (5))} \\
&= \widehat{f}(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cap\mathbf{Y}}) + \sum_{i=0}^{k}\widehat{f}((t_{i+1} - t_i)\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cup_{t_i}\mathbf{Y}}) \qquad \text{(by (13))} \\
&= \widehat{f}(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cap\mathbf{Y}}) + \sum_{i=0}^{k}(t_{i+1} - t_i)\widehat{f}(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cup_{t_i}\mathbf{Y}}) \qquad \text{(since } \widehat{f} \text{ is positively homogeneous)} \\
&= f(\mathbf{X} \cap \mathbf{Y}) + \sum_{i=0}^{k}(t_{i+1} - t_i)f(\mathbf{X} \cup_{t_i} \mathbf{Y}) \qquad \text{(since } \widehat{f} \text{ is an extension of } f\text{)}
\end{aligned}$$

Hence f is $\boldsymbol{\alpha}$-bisubmodular. □

We also have the following theorem (see [2, 17] for special cases of bisubmodular and α-bisubmodular functions; also see [14, Proposition 4.11] for more general functions).

Theorem 8. *Under the same assumption as in Theorem 7, $f : 3^V \to \mathbb{R}$ is $\boldsymbol{\alpha}$-bisubmodular if and only if*

(a) *for every orthant $\mathbf{S}$, f restricted on $2^{\mathbf{S}}$ is submodular, and*

(b) *for every $v \in V$ and $U \subseteq V \setminus \{v\}$, putting $W = V \setminus (\{v\} \cup U)$, we have*

$$\alpha^-(v)f(U \cup \{v\}, W) + \alpha^+(v)f(U, W \cup \{v\}) \geq (\alpha^+(v) + \alpha^-(v))f(U, W).$$

Proof. We can easily see that the $\boldsymbol{\alpha}$-bisubmodularity of f implies (a) and (b). Hence it suffices to show the if part.

Suppose that (a) and (b) hold. It follows from (a) that the extension $\widehat{f}$ defined by (4) is convex on the cone $\mathbb{R}^{S_+}_{\geq 0} \times \mathbb{R}^{S_-}_{\leq 0}$ of every orthant $\mathbf{S}$ (see [19]). Moreover, (b) implies the convexity of $\widehat{f}$ on

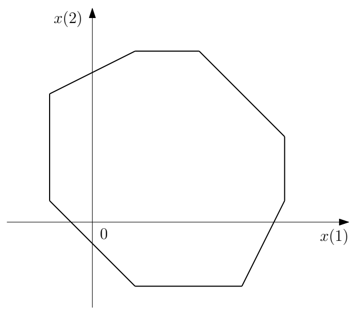

Figure 2: A $\frac{1}{2}$-bisubmodular polyhedron.

the union of adjacent simplices (having common facet $x(v) = 0$) that correspond to maximal chains of 3^V:

$$\mathbf{X}_0(= (\emptyset, \emptyset)) \subset \cdots \subset \mathbf{X}_{n-1}(= (U, W)) \subset \mathbf{X}_n = (U \cup \{v\}, W),$$
$$\mathbf{X}_0(= (\emptyset, \emptyset)) \subset \cdots \subset \mathbf{X}_{n-1}(= (U, W)) \subset \mathbf{X}'_n = (U, W \cup \{v\}),$$

where note that only the last elements (adjacent orthants) are different. Hence $\widehat{f}$ is convex, so that f is $\boldsymbol{\alpha}$-bisubmodular due to Theorem 7. □

For a submodular function $f : 2^V \to \mathbb{R}$, let $\widehat{f}$ be the Lovász extension of f ([19]). As shown by Grötschel, Lovász, and Schrijver [13], one can develop a polynomial-time (weak) separation oracle that separates a point $p \in \mathbb{R}^V \setminus F$ from the set F of minimizers of $\widehat{f}$, which implies that one can find a minimizer of $\widehat{f}$ in polynomial time. Since $\widehat{f}$ is linear on each cell of the simplicial division, one can also find a minimizer of f. Qi [21] extended this argument to bisubmodular functions, and here we can adopt the same argument for $\boldsymbol{\alpha}$-bisubmodular function f due to the convexity of $\widehat{f}$.

Corollary 9. *Any $\boldsymbol{\alpha}$-bisubmodular function $f : 3^V \to \mathbb{R}$ can be minimized in strongly polynomial time.*

It follows from Proposition 3 and Theorem 5 that, for a fixed orthant $\mathbf{S}$, $\mathrm{P}_{\mathbf{S}}(f)$ is the one obtained from a submodular polyhedron by a reflection and scaling. It is known that each edge vector (i.e., a direction vector of each edge) of a bisubmodular polyhedron is one of the following forms

$$(0, \ldots, 0, 1, 0 \ldots, 0),\ (0, \ldots, 0, \pm 1, 0, \ldots, 0, \mp 1, 0, \ldots, 0),\ (0, \ldots, 0, \pm 1, 0, \ldots, 0, \pm 1, 0, \ldots, 0)$$

and hence each edge vector of an $\boldsymbol{\alpha}$-bisubmodular polyhedron has the support of size at most two. See Figure 2 for an example.

The concept of a *polybasic polyhedron* is introduced in [11], where a convex polyhedron is polybasic if every edge vector has a support of size at most two. Hence, skew bisubmodular polyhedra are special cases of polybasic polyhedra.

4 A Min-Max Theorem

For any $x \in \mathbb{R}^V$ let us define

$$\|x\|_{\boldsymbol{\alpha}} = \sum_{v \in V: x(v)>0} \boldsymbol{\alpha}^-(v)x(v) - \sum_{v \in V: x(v)<0} \boldsymbol{\alpha}^+(v)x(v).$$

It is not difficult to see that $\|\cdot\|_{\boldsymbol{\alpha}}$ is a norm on $\mathbb{R}^V$. The following extension of a theorem given in [8] implies that the $\boldsymbol{\alpha}$-bisubmodular function minimization can be reduced to finding a minimum-norm point with respect to $\|\cdot\|_{\boldsymbol{\alpha}}$ in the $\boldsymbol{\alpha}$-bisubmodular polyhedron $\mathrm{P}(f)$.

To show this we need one technical lemma. For $x \in \mathrm{P}(f)$, $\mathbf{X}$ is called *x-tight* if $x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) = f(\mathbf{X})$.

Lemma 10. *Let $x \in \mathrm{P}(f)$. If $\mathbf{X}$ and $\mathbf{Y}$ are x-tight, then $\mathbf{X} \cap \mathbf{Y}$ and $\mathbf{X} \cup_{t_i} \mathbf{Y}$ $(i = 0, \ldots, k+1)$ are also x-tight.*

Proof. By using Lemma 2, $x \in \mathrm{P}(f)$, $\boldsymbol{\alpha}$-bisubmodularity of f, and the x-tightness of $\mathbf{X}$ and $\mathbf{Y}$, we have $x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) + x(\chi^{\boldsymbol{\alpha}}_{\mathbf{Y}}) = x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cap\mathbf{Y}}) + \sum_{i=0}^{k}(t_{i+1}-t_i)x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}\cup_{t_i}\mathbf{Y}}) \leq f(\mathbf{X}\cap\mathbf{Y}) + \sum_{i=0}^{k}(t_{i+1}-t_i)f(\mathbf{X}\cup_{t_i}\mathbf{Y}) \leq f(\mathbf{X}) + f(\mathbf{Y}) = x(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}}) + x(\chi^{\boldsymbol{\alpha}}_{\mathbf{Y}})$. Hence, the inequalities must hold with equality, from which follows the present lemma. □

Theorem 11. *For any $\boldsymbol{\alpha}$-bisubmodular function $f : 3^V \to \mathbb{R}$,*

$$\min\{\|x\|_{\boldsymbol{\alpha}} \mid x \in \mathrm{P}(f)\} = \max\{-f(\mathbf{X}) \mid \mathbf{X} \in 3^V\}. \tag{14}$$

Proof. For any $x \in \mathrm{P}(f)$ and $\mathbf{X} = (X_+, X_-) \in 3^V$, we always have $\|x\|_{\boldsymbol{\alpha}} \geq -f(\mathbf{X})$ since

$$\|x\|_{\boldsymbol{\alpha}} = \sum_{v \in V: x(v)>0} \boldsymbol{\alpha}^-(v)x(v) - \sum_{v \in V: x(v)<0} \boldsymbol{\alpha}^+(v)x(v) \geq \sum_{v \in X_-} \boldsymbol{\alpha}^-(v)x(v) - \sum_{v \in X_+} \boldsymbol{\alpha}^+(v)x(v) \geq -f(\mathbf{X}).$$

Hence it suffices to show that $\|x\|_{\boldsymbol{\alpha}} = -f(\mathbf{X})$ for some $x \in \mathrm{P}(f)$ and $\mathbf{X} \in 3^V$.

Let $\hat{x}$ be a minimizer of the left-hand side of (14), and let $A_+ = \{v \in V \mid \hat{x}(v) < 0\}$, $A_- = \{v \in V \mid \hat{x}(v) > 0\}$, and $\mathbf{A} = (A_+, A_-)$. Note that for any $u \in A_+$ and $v \in A_-$ there exist $\hat{x}$-tight $\mathbf{X}$ and $\mathbf{Y}$ such that $u \in X_+$ and $v \in Y_-$.

Take any $u \in A_+$. For each $v \in A_-$, if every $\hat{x}$-tight $\mathbf{X}$ with $u \in X_+$ satisfies $v \in X_+$, then for a sufficiently small positive number ϵ, we can obtain a better solution than $\hat{x}$ in the minimization problem by increasing $\hat{x}(u)$ by $\epsilon/\boldsymbol{\alpha}^+(u)$ and decreasing $\hat{x}(v)$ by $\epsilon/\boldsymbol{\alpha}^+(v)$. Therefore, for each $v \in A_-$, there exists an $\hat{x}$-tight $\mathbf{X}^{uv}$ such that $u \in X^{uv}_+$ and $v \notin X^{uv}_+$. Similarly, for each $v \in A_+ \setminus \{u\}$, there exists an $\hat{x}$-tight set $\mathbf{X}^{uv}$ such that $u \in X^{uv}_+$ and $v \notin X^{uv}_-$, since otherwise (i.e., no such $\hat{x}$-tight set exists) for a sufficiently small positive number ϵ, increasing $\hat{x}(u)$ by $\epsilon/\boldsymbol{\alpha}^+(u)$ and $\hat{x}(v)$ by $\epsilon/\boldsymbol{\alpha}^-(v)$ gives a better solution again. Put $\mathbf{X}^u = \bigcap_{v \in (A_+\setminus\{u\})\cup A_-} \mathbf{X}^{uv}$. It follows from Lemma 10 that $\mathbf{X}^u$ is $\hat{x}$-tight with $u \in X^u_+$, $X^u_+ \cap A_- = \emptyset$, and $X^u_- \cap A_+ = \emptyset$.

By a symmetric argument we see that for any $u \in A_-$ there is an $\hat{x}$-tight $\mathbf{X}^u$ such that $u \in X^u_-$, $X^u_+ \cap A_- = \emptyset$, and $X^u_- \cap A_+ = \emptyset$.

Put $\mathbf{X}^* = \bigcup_0 \mathbf{X}^u$, where $\cup_0$ is taken over all $u \in A_+ \cup A_-$. Then, it follows from Lemma 10 that $\mathbf{X}^*$ is $\hat{x}$-tight with $A_+ \subseteq X^*_+$ and $A_- \subseteq X^*_-$. Moreover, since $\hat{x}(v) = 0$ for all $v \in V \setminus (A_+ \cup A_-)$ by the definition of A_+ and A_-, we have

$$\begin{aligned}\|\hat{x}\|_{\boldsymbol{\alpha}} &= \sum_{v \in V : \hat{x}(v) > 0} \alpha^-(v)\hat{x}(v) - \sum_{v \in V : \hat{x}(v) < 0} \alpha^+(v)\hat{x}(v) = \sum_{v \in A_-} \alpha^-(v)\hat{x}(v) - \sum_{v \in A_+} \alpha^+(v)\hat{x}(v) \\ &= \sum_{v \in X^*_-} \alpha^-(v)\hat{x}(v) - \sum_{v \in X^*_+} \alpha^+(v)\hat{x}(v) = -\hat{x}(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}^*})\end{aligned}$$

Consequently, by the $\hat{x}$-tightness of $\mathbf{X}^*$, we obtain $\|\hat{x}\|_{\boldsymbol{\alpha}} = -\hat{x}(\chi^{\boldsymbol{\alpha}}_{\mathbf{X}^*}) = -f(\mathbf{X}^*)$. This completes the proof. □

5 Concluding Remarks

We have considered a natural generalization of the concept of skew bisubmodularity. We have shown a characterization of the generalized skew bisubmodularity in terms of its convex extension over rectangles, where an important rôle is played by skew bisubmodular polyhedra associated with skew bisubmodular functions. We have also derived a min-max theorem (Theorem 11) that relates the minimum value of a skew bisubmodular function to a minimum-norm point in the associated skew bisubmodular polyhedron. All the existing combinatorial algorithms for minimizing submodular functions or bisubmodular functions are based on min-max theorems corresponding to Theorem 11. Devising a combinatorial polynomial-time algorithm for skew bisubmodular function minimization will be discussed elsewhere.

References

[1] K. Ando and S. Fujishige: On structures of bisubmodular polyhedra. *Mathematical Programming*, 74:293–317 (1996).

[2] K. Ando, S. Fujishige and T. Naitoh: A characterization of bisubmodular functions. *Discrete Mathematics*, 148:299–303 (1996).

[3] A. Bouchet. Greedy algorithm and symmetric matroids. *Mathematical Programming*, 38(2):147–159, 1987.

[4] A. Bouchet and W. H. Cunningham. Delta-matroids, jump systems, and bisubmodular polyhedra. *SIAM Journal on Discrete Mathematics*, 8(1):17–32, 1995.

[5] R. Chandrasekaran and S. N. Kabadi. Pseudomatroids. *Discrete Mathematics*, 71(3):205–217, 1988.

[6] F. D. J. Dunstan and D. J. A. Welsh. A greedy algorithm for solving a certain class of linear programmes. *Mathematical Programming*, 62: 338–353, 1973.

[7] J. Edmonds. Submodular functions, matroids, and certain polyhedra. *Proceedings of the Calgary International Conference on Combinatorial Structures and Their Applications* (R. Guy, H. Hanani, N. Sauer and J. Schönheim, eds., Gordon and Breach, New York, 1970), pp. 69–87; also in: *Combinatorial Optimization—Eureka, You Shrink!* (M. Jünger, G. Reinelt and G. Rinaldi, eds., Lecture Notes in Computer Science **2570**, Springer, Berlin, 2003), pp. 11–26.

[8] S. Fujishige. A min–max theorem for bisubmodular polyhedra. *SIAM Journal on Discrete Mathematics*, 10(2):294–308, 1997.

[9] S. Fujishige. *Submodular Functions and Optimization*, volume 58 of *Annals of Discrete Mathematics*. Elsevier, 2nd edition, 2005.

[10] S. Fujishige and S. Iwata. Bisubmodular function minimization. *SIAM Journal on Discrete Mathematics*, 19(4):1065–1073, 2005.

[11] S. Fujishige, K. Makino, T. Takabatake, and K. Kashiwabara. Polybasic polyhedra: structure of polyhedra with edge vectors of support size at most 2. *Discrete Mathematics*, 280(1):13–27, 2004.

[12] M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981.

[13] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics*. Springer-Verlag, 1st edition, 1988.

[14] H. Hirai. *Discrete convexity and polynomial solvability in minimum 0-extension problems.* Mathematical Engineering Technical Reports, METR 2012-18, Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, October 2012.

[15] H. Hirai. *Discrete convexity for multiflows and 0-extensions.* Proceedings of the 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, June 4-7, 2013 in Veszprém, Hungary.

[16] A. Huber and A. Krokhin. A Lovász extension for skew bisubmodular functions. Unpublished manuscript, October 30, 2012.

[17] A. Huber, A. Krokhin, and R. Powell. Skew bisubmodularity and valued CSPs. In *SODA'13: Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1296–1305, 2013.

[18] S. Iwata, L. Fleischer, and S. Fujishige. A combinatorial strongly polynomial algorithm for minimizing submodular functions. *Journal of the ACM*, 48(4):761–777, 2001.

[19] L. Lovász.: Submodular functions and convexity. In: *Mathematical Programming—The State of the Art* (A. Bachem, M. Grötschel and B. Korte, eds., Springer, Berlin, 1983), pp. 235–257.

[20] S. T. McCormick and S. Fujishige. Strongly polynomial and fully combinatorial algorithms for bisubmodular function minimization. *Mathematical Programming*, 122(1):87–120, 2008.

[21] L. Qi. Directed submodularity, ditroids and directed submodular flows. *Mathematical Programming*, 42(1-3):579–599, 1988.

[22] A. Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *Journal of Combinatorial Theory, Series B*, 80(2):346–355, 2000.