Math. Program., Ser. A
DOI 10.1007/s10107-017-1171-2

FULL LENGTH PAPER

Polynomial combinatorial algorithms for skew-bisubmodular function minimization

Satoru Fujishige[1] · **Shin-ichi Tanigawa**[1,2]

Received: 23 October 2015 / Accepted: 15 June 2017

Abstract Huber et al. (SIAM J Comput 43:1064–1084, 2014) 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, and Huber and Krokhin (SIAM J Discrete Math 28:1828–1837, 2014) showed the oracle tractability of minimization of skew-bisubmodular functions. Fujishige et al. (Discrete Optim 12:1–9, 2014) also showed a min–max theorem that characterizes the skew-bisubmodular function minimization, but devising a combinatorial polynomial algorithm for skew-bisubmodular function minimization was left open. In the present paper we give first combinatorial (weakly and strongly) polynomial algorithms for skew-bisubmodular function minimization.

Keywords Skew-bisubmodular functions · Submodular functions · Discrete convexity · Combinatorial algorithms · Strongly polynomial algorithms

Mathematics Subject Classification 90C27 · 52B40 · 68W40

1 Introduction

The concept of bisubmodularity was independently introduced by Bouchet [3] and Chandrasekaran–Kabadi [5] (also see [1,2,6,7,22]), and has been extensively studied

✉ Shin-ichi Tanigawa
tanigawa@kurims.kyoto-u.ac.jp

Satoru Fujishige
fujishig@kurims.kyoto-u.ac.jp

1 Research Institute for Mathematical Sciences, Kyoto University, Kyoto, Japan

2 Centrum Wiskunde & Informatica, Amsterdam, Netherlands

Published online: 22 June 2017

in combinatorial optimization as a generalization of submodular functions (see, e.g., [4]). As a further generalization of bisubmodularity, the concept of skew-bisubmodular function was recently introduced by Huber et al. [16] in their complexity dichotomy theorem for the valued constraint satisfaction problems (VCSPs) over the three-value domain (cf. [24]).

Let V be a finite nonempty set of n elements and $3^V = \{(X, Y) \mid X, Y \subseteq V, X \cap Y = \emptyset\}$. Let $\alpha \in (0, 1]$. A function $f\colon 3^V \to \mathbb{R}$ is called *α-bisubmodular* [16] if, for every $\mathbf{Z}_1 = (X_1, Y_1)$ and $\mathbf{Z}_2 = (X_2, Y_2) \in 3^V$, f satisfies

$$f(\mathbf{Z}_1) + f(\mathbf{Z}_2) \geq f(\mathbf{Z}_1 \cap \mathbf{Z}_2) + \alpha f(\mathbf{Z}_1 \cup_0 \mathbf{Z}_2) + (1 - \alpha) f(\mathbf{Z}_1 \cup_1 \mathbf{Z}_2),$$

where $\mathbf{Z}_1 \cap \mathbf{Z}_2 = (X_1 \cap X_2, Y_1 \cap Y_2)$, $\mathbf{Z}_1 \cup_0 \mathbf{Z}_2 = ((X_1 \cup X_2)\backslash(Y_1 \cup Y_2), (Y_1 \cup Y_2)\backslash(X_1 \cup X_2))$, and $\mathbf{Z}_1 \cup_1 \mathbf{Z}_2 = (X_1 \cup X_2, (Y_1 \cup Y_2)\backslash(X_1 \cup X_2))$. When $\alpha = 1$, 1-bisubmodularity is exactly bisubmodularity. A function $f\colon 3^V \to \mathbb{R}$ is called *skew-bisubmodular* [16] if it is α-bisubmodular for some $\alpha \in (0, 1]$. Huber and Krokhin [15] pointed out that the minimization of skew-bisubmodular functions is tractable via the ellipsoid method as in the work by Qi [22] for bisubmodular functions. However, developing a combinatorial algorithm remains unsolved.

In this paper we give first combinatorial weakly and strongly polynomial algorithms for skew bisubmodular function minimization. In [12] the concept of skew-bisubmodularity was slightly generalized, and a min–max relation characterizing the minimum of a (generalized) skew-bisubmodular function was shown by introducing skew-scaled bisubmodular polyhedra. Building on those polyhedral backgrounds, our algorithms are adaptations of the combinatorial algorithms for bisubmodular function minimization by Fujishige and Iwata [9] and McCormick and Fujishige [21], which are built on the Iwata–Fleischer–Fujishige algorithm [17] for submodular function minimization. However, a simple adaptation causes several technical problems. Two major obstacles, which seem worth emphasizing here, are listed below.

1. The Fujishige–Iwata weakly polynomial algorithm [9] makes use of the boundary operator of skew-symmetric digraphs to describe edge vectors of the associated bisubmodular polyhedron, and their analysis implicitly relies on the symmetry of the operator. In the skew-bisubmodular case, the associated polyhedra are scaled ("skewed") and the edge vectors are best described by scaled boundary of skew-symmetric digraphs. This, however, makes the boundary operator asymmetric, and we cannot directly apply the arguments of [9] and [21]. We will overcome the difficulty by introducing a new augmentation concept, called *augmenting path-sequence*.
2. Given a partition $\Pi = \{X_1, \ldots, X_k\}$ of V, one can define the *aggregation* of a submodular function $f\colon 2^V \to \mathbb{R}$ as a function $\hat{f}$ on 2^Π defined by $\hat{f}(S) = f(\bigcup_{X \in S} X)$ for $S \subseteq \Pi$. This operation can naturally be extended to bisubmodular functions, and as in the Iwata–Fleischer–Fujishige algorithm [17] for submodular functions, the McCormick–Fujishige strongly polynomial algorithm [21] makes use of aggregation as a crucial tool to control the size of entry values of bases in the intermediate steps. This operation, however, cannot be extended to skew-bisubmodular functions (at least in an obvious manner). This difficulty will be

overcome by introducing a new technique to find a base of the associated polyhedron with small duality gap with the aid of the ordinary submodular function minimization as a subroutine.

Our quest of extending combinatorial algorithms for submodular functions is motivated by questions about the tractability of submodular function minimization defined on general discrete structures such as semilattices and sets of transversals (see, e.g., [10,11,13,14,18–20]), where bisubmodular functions are special cases of submodular functions on a semilattice [13]. New techniques presented here might also be useful for other classes of submodular functions.

The rest of the paper is organized as follows. In Sect. 2 we list preliminary facts on skew-bisubmodular functions given in [12] and introduce necessary notation. In Sect. 3 we first give a combinatorial weakly polynomial algorithm. In Sect. 4 we give a combinatorial strongly polynomial algorithm by using the main body of the weakly polynomial algorithm as a subroutine.

2 Definitions and preliminaries

For each $v \in V$ let $\chi_v \in \mathbb{R}^V$ be the characteristic vector of the singleton set $\{v\}$, i.e., $\chi_v(v) = 1$ and $\chi_v(w) = 0$ for $w \in V \setminus \{v\}$. Each $(X, Y) \in 3^V$ is called a *signed set* and is identified with a $\{0, \pm 1\}$-vector $\sum_{v \in X} \chi_v - \sum_{v \in Y} \chi_v$.

The original definition of skew-bisubmodular function of Huber et al. [16] was slightly generalized in [12] as follows.

Let $\alpha = (\alpha^+, \alpha^-)$ with $\alpha^+ \colon V \to \mathbb{R}_{>0}$ and $\alpha^- \colon V \to \mathbb{R}_{>0}$. For simplicity we assume

$$\alpha^+(v) \geq \alpha^-(v) \qquad (\forall v \in V)$$

without loss of generality.[1] Let $\beta = \max\{\frac{\alpha^+(v)}{\alpha^-(v)} \mid v \in V\}$. Note that by the assumption we have $0 < \frac{\alpha^-(v)}{\alpha^+(v)} \leq 1$ for all $v \in V$ and hence $\beta \geq 1$. For each $t \in [0, 1)$, let $V_t = \left\{ v \in V \,\middle|\, \frac{\alpha^-(v)}{\alpha^+(v)} \leq t \right\}$ and define a binary operation $\sqcup_t$ on 3^V by

$$(X_1, Y_1) \sqcup_t (X_2, Y_2) = (((X_1 \cup X_2) \setminus \Delta) \cup (V_t \cap \Delta), (Y_1 \cup Y_2) \setminus \Delta)$$

where $\Delta = (X_1 \cup X_2) \cap (Y_1 \cup Y_2)$ (see Fig. 1). Note that V_t is monotone nondecreasing in $t \in [0, 1)$.

The (*generalized*) *skew-bisubmodular function* is defined based on binary operations $\sqcap$ and $\sqcup_t$ $(t \in [0, 1))$ on 3^V as follows, by generalizing $\sqcup_0$ and $\sqcup_1$ given in the introduction.

[1] If $\alpha^+(v) < \alpha^-(v)$ for some $v \in V$, consider a reflection of f by element v given by $f^v(X, Y) = f(X \setminus \{v\}, Y \cup \{v\})$ if $v \in X$, $f^v(X, Y) = f(X \cup \{v\}, Y \setminus \{v\})$ if $v \in Y$, and $f^v(X, Y) = f(X, Y)$ otherwise. Also consider old $\alpha^+(v)$ and $\alpha^-(v)$ as new $\alpha^-(v)$ and $\alpha^+(v)$, respectively.

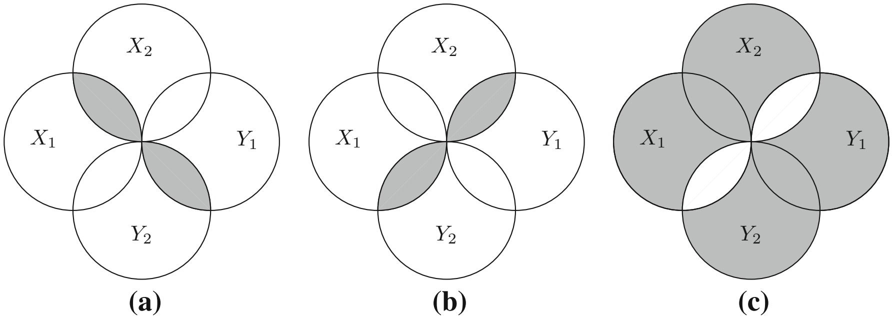

Fig. 1 The *shaded regions* correspond to **a** $(X_1, Y_1) \sqcap (X_2, Y_2)$, **b** $\Delta = (X_1 \cup X_2) \cap (Y_1 \cup Y_2)$, and **c** $(X_1, Y_1) \sqcup_0 (X_2, Y_2)$

Definition 1 For given V and α, define a set $T = \{\frac{\alpha^-(v)}{\alpha^+(v)} \mid v \in V\} \cup \{0, 1\}$ and arrange the distinct elements of T in the increasing order of magnitude as $0 = t_0 < t_1 < t_2 < \cdots < t_{p+1} = 1$. Then a function $f : 3^V \to \mathbb{R}$ is called α-*bisubmodular* if

$$f(X_1, Y_1) + f(X_2, Y_2) \geq f\left((X_1, Y_1) \sqcap (X_2, Y_2)\right) + \sum_{i=0}^{p} (t_{i+1} - t_i) f\left((X_1, Y_1) \sqcup_{t_i} (X_2, Y_2)\right)$$

for all $(X_1, Y_1), (X_2, Y_2) \in 3^V$. We assume $f(\emptyset, \emptyset) = 0$.

We consider the problem of minimizing an α-bisubmodular function f.

We give some additional definitions and notation, and then review basic facts about α-bisubmodular functions shown in [12]. Throughout the paper, we prepare the signed copies v^+ and v^- for each $v \in V$. For any $X \subseteq V$ define $X^+ = \{v^+ \mid v \in X\}$ and $X^- = \{v^- \mid v \in X\}$. Every signed set $(X, Y) \in 3^V$ is identified with $X^+ \cup Y^-$ if it is clear from the context. A subset Z of $V^+ \cup V^-$ is called *consistent* if there exists no $v \in V$ such that $\{v^+, v^-\} \subseteq Z$. Note that there is a natural bijection between 3^V and the set of all consistent subsets of $V^+ \cup V^-$. For any $(X_1, Y_1), (X_2, Y_2) \in 3^V$ we say that (X_1, Y_1) and (X_2, Y_2) are *compatible* if $X_1 \cap Y_2 = \emptyset$ and $X_2 \cap Y_1 = \emptyset$. For any compatible (X_1, Y_1) and (X_2, Y_2) we write $(X_1, Y_1) \cup (X_2, Y_2) = (X_1 \cup X_2, Y_1 \cup Y_2)$. Also we write $(X_1, Y_1) \subseteq (X_2, Y_2)$ if $X_1 \subseteq X_2$ and $Y_1 \subseteq Y_2$. When $(X_1, Y_1) \subseteq (X_2, Y_2)$, define $(X_2, Y_2) \backslash (X_1, Y_1) = (X_2 \backslash X_1, Y_2 \backslash Y_1)$.

For any $(A, B) \in 3^V$ define the *contraction* $f_{(A,B)}$ of f by (A, B) as follows: the domain of $f_{(A,B)}$ is given by $3^{V \backslash (A \cup B)}$ and for each $(X, Y) \in 3^{V \backslash (A \cup B)}$ $f_{(A,B)}(X, Y) = f((X, Y) \cup (A, B)) - f(A, B)$. The contraction $f_{(A,B)}$ is α-bisubmodular.

For any $(X, Y) \in 3^V$ define a vector $\chi^\alpha_{(X,Y)}$ in $\mathbb{R}^V$ by

$$\chi^\alpha_{(X,Y)} = \sum_{v \in X} \alpha^+(v) \chi_v - \sum_{v \in Y} \alpha^-(v) \chi_v$$

which can be regarded as a signed α-scaled characteristic vector of signed set (X, Y). Note that the canonical inner product of $x \in \mathbb{R}^V$ and $\chi^\alpha_{(X,Y)}$ is given by

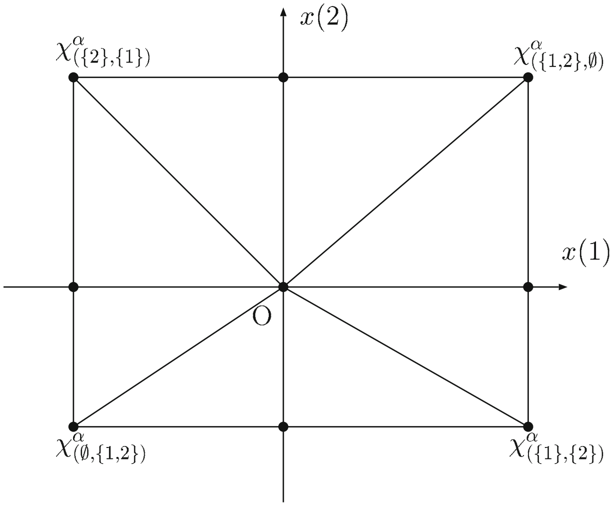

Fig. 2 A simplicial division that determines the convex extension of f

$$\begin{aligned}\left\langle x, \chi^{\alpha}_{(X,Y)}\right\rangle &= \sum_{v\in X} \alpha^{+}(v)x(v) - \sum_{v\in Y} \alpha^{-}(v)x(v)\\ &= \sum_{v^{\sigma}\in(X,Y)} \sigma\alpha^{\sigma}(v)x(v).\end{aligned}$$

The α-*bisubmodular polyhedron* associated with an α-bisubmodular function f is defined by

$$\mathrm{P}^{\alpha}(f) = \left\{x \in \mathbb{R}^{V} \mid \forall (X,Y) \in 3^{V} \colon \left\langle x, \chi^{\alpha}_{(X,Y)}\right\rangle \le f(X,Y)\right\}.$$

A signed set $(S,T) \in 3^{V}$ with $S \cup T = V$ is called an *orthant*. For each $(S,T) \in 3^{V}$, f restricted on $2^{(S,T)} := \{(X,Y) \mid (X,Y) \subseteq (S,T)\}$ is an ordinary submodular function. Hence, in each orthant (S,T) we have the α-scaled submodular polyhedron given by

$$\mathrm{P}^{\alpha}_{(S,T)}(f) = \left\{x \in \mathbb{R}^{V} \mid \forall (X,Y) \subseteq (S,T) \colon \left\langle x, \chi^{\alpha}_{(X,Y)}\right\rangle \le f(X,Y)\right\},$$

and the α-scaled base polyhedron by

$$\mathrm{B}^{\alpha}_{(S,T)}(f) = \left\{x \in \mathrm{P}^{\alpha}_{(S,T)}(f) \mid \langle x, \chi^{\alpha}_{(S,T)}\rangle = f(S,T)\right\}.$$

[Compare them with ordinary submodular polyhedra and base polyhedra (see [8])].

Figures 2, 3 show two-dimensional examples with $V = \{1,2\}$. Figure 2 gives a simplicial division of the rectangle (the convex hull of points $\chi^{\alpha}_{(X,Y)}$ $((X,Y) \in 3^{V})$) that determines the convex extension of f. Note that the extension of f is convex if and only if f is α-bisubmodular [12,15]. Figure 3 shows an example of the α-bisubmodular polyhedron $\mathrm{P}^{\alpha}(f)$, which is the subdifferential of the convex extension

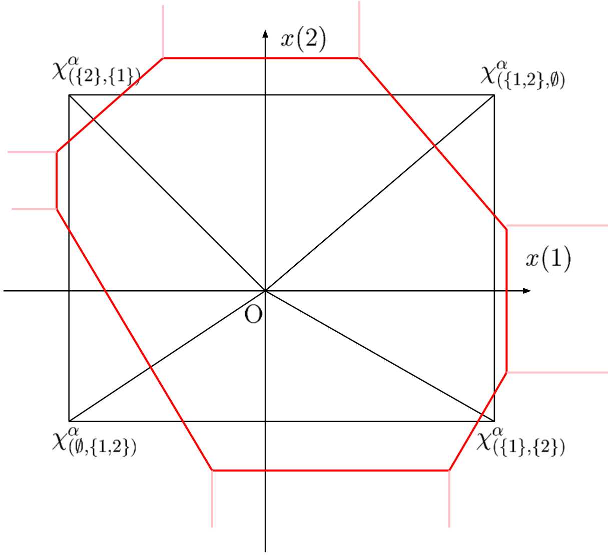

Fig. 3 An example of the α-bisubmodular polyhedron $\mathrm{P}^{\alpha}(f)$

of f at the origin. This can be seen by the defining inequalities for $x \in \mathrm{P}^{\alpha}(f)$: $\forall (X, Y) \in 3^V : \langle x, \chi^{\alpha}_{(X,Y)} - \chi^{\alpha}_{(\emptyset,\emptyset)} \rangle \leq f(X, Y) - f(\emptyset, \emptyset)$.

Let $\sigma : V \to \{+, -\}$ be a sign function. For any $X \subseteq V$, $X|\sigma$ denotes $(X_{\sigma^+}, X_{\sigma^-}) \in 3^V$ with $X_{\sigma^+} = \{v \in X \mid \sigma(v) = +\}$ and $X_{\sigma^-} = \{v \in X \mid \sigma(v) = -\}$. Let $L = (v_1, \ldots, v_n)$ be a linear ordering of V with $|V| = n$. For each $i = 1, \ldots, n$ define $L(v_i) = \{v_1, \ldots, v_i\}$.

For a linear ordering $L = (v_1, \ldots, v_n)$ of V and a sign function $\sigma : V \to \{+, -\}$, let $y \in \mathbb{R}^V$ be given by

$$y(v_i) = \sigma(v_i) \frac{f(L(v_i)|\sigma) - f(L(v_{i-1})|\sigma)}{\alpha^{\sigma(v_i)}(v_i)} \tag{1}$$

for $i = 1, \ldots, n$, where we define $L(v_0) = \emptyset$. Then y is an extreme point of $\mathrm{P}^{\alpha}(f)$, which is called *the extreme point generated by L and σ*. Conversely, every extreme point of $\mathrm{P}^{\alpha}(f)$ can be generated by some L and σ through (1). Note that y is determined by a signed, α-scaled version of the *greedy algorithm* by (1).

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

$$\begin{aligned} \|x\|_{\alpha} &:= - \sum_{v \in V : x(v) < 0} \alpha^{+}(v) x(v) + \sum_{v \in V : x(v) > 0} \alpha^{-}(v) x(v) \\ &= - \sum_{v \in V : \sigma x(v) < 0} \sigma \alpha^{\sigma}(v) x(v) \\ &= \sum_{v \in V : \sigma x(v) < 0} \alpha^{\sigma}(v) |x(v)|, \end{aligned}$$

which is an asymmetric norm (positively homogeneous convex function) of x.

The following min–max theorem characterizing the minimum value of α-bisubmodular function f was shown in [12].

Theorem 1 *For any α-bisubmodular function $f: 3^V \to \mathbb{R}$ (with $f(\emptyset,\emptyset)=0$),*

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

3 Weakly polynomial algorithm

In this section we describe a weakly polynomial algorithm for minimizing an α-bisubmodular function f. The algorithm is designed for any real-valued α-bisubmodular functions, and its main subroutine will be also used for the strongly polynomial algorithm in the next section.

3.1 Algorithm description

Let $\mathrm{K}_{V^+\cup V^-}$ be the complete digraph with vertex set $V^+ \cup V^-$, where recall that V^+ is the positive copy of V and V^- is the negative copy of V.

During the execution of our algorithm we keep the following:

- a positive number δ, which will be used as a parameter of the scaling.
- a vector $x \in \mathrm{P}^\alpha(f)$ along with its expression as a convex combination of extreme points y_i of $\mathrm{P}^\alpha(f)$ indexed by a finite set J, i.e.,

$$x = \sum_{i\in J} \lambda_i y_i \quad \text{with } \lambda_i > 0\,(i \in J) \text{ and } \sum_{i\in J} \lambda_i = 1. \tag{2}$$

  Here each y_i is represented by a pair of a linear ordering L_i of V and a sign function σ_i on V [with which y_i is computed by (1)]. It should be noted that each y_i computed as such is an extreme point of $\mathrm{P}^\alpha(f)$.
- a nonnegative function $\psi: (V^+ \cup V^-) \times (V^+ \cup V^-) \to \mathbb{R}_{\geq 0}$.

Such a function $\psi: (V^+ \cup V^-) \times (V^+ \cup V^-) \to \mathbb{R}_{\geq 0}$ is called a *flow* in $\mathrm{K}_{V^+\cup V^-}$. The algorithm starts with

- some positive δ and $x \in \mathrm{P}^\alpha(f)$, which will be specified later, and
- $\psi = \mathbf{0}$.

The algorithm is controlled by the scaling parameter δ. At each scaling phase with parameter δ we keep ψ being δ-*feasible*, which by definition satisfies the following for all $u, v \in V^+ \cup V^-$:

- $0 \leq \psi(u,v) \leq \delta$,
- $\psi(u,v) = 0$ or $\psi(v,u) = 0$.

For a flow $\psi: (V^+ \cup V^-) \times (V^+ \cup V^-) \to \mathbb{R}_{\geq 0}$, define $\partial_\alpha \psi \in \mathbb{R}^V$ by

$$\partial_\alpha \psi = \sum_{(u^{\tau_1}, v^{\tau_2})} \left(\tau_1 \frac{1}{\alpha^{\tau_1}(u)} \chi_u - \tau_2 \frac{1}{\alpha^{\tau_2}(v)} \chi_v\right) \psi(u^{\tau_1}, v^{\tau_2}), \tag{3}$$

where the sum is taken over all arcs (u^{τ_1}, v^{τ_2}) of $\mathrm{K}_{V^+\cup V^-}$ with $\tau_1, \tau_2 \in \{+, -\}$. Note that (3) can also be written as follows.

$$\partial_\alpha \psi = \sum_{v\in V} \frac{\partial\psi(v^+)}{\alpha^+(v)}\chi_v - \sum_{v\in V} \frac{\partial\psi(v^-)}{\alpha^-(v)}\chi_v,$$

where $\partial\psi(v^\pm)$ is the ordinary flow boundary (the net out-flow value) of ψ at vertex $v^\pm$ in $\mathrm{K}_{V^+\cup V^-}$ defined by

$$\partial\psi(v) = \sum_{w\in V^+\cup V^-} \psi(v, w) - \sum_{w\in V^+\cup V^-} \psi(w, v) \qquad (\forall v \in V^+ \cup V^-).$$

(The α-boundary $\partial_\alpha\psi$ of ψ is sort of signed, inversely α-scaled flow boundary of ψ.) Then put

$$z := x + \partial_\alpha \psi.$$

At each phase we try to minimize $\|z\|_\alpha$. The gap between $\|x\|_\alpha$ and $\|z\|_\alpha$ can be estimated by the δ-feasible flow ψ, and it becomes close to zero for small $\delta > 0$ since ψ is δ-feasible. We will show that when δ becomes small enough, then obtained x gives a minimizer of f if f is integer-valued.

We now describe the algorithm. Each phase starts by cutting the value of δ in half, and then it modifies ψ to make it δ-feasible. This can be done by setting each $\psi(u, v)$ to δ if the value is more than δ.

In order to decrease $\|z\|_\alpha$ we introduce an auxiliary graph and define augmenting paths. The *auxiliary graph* with respect to ψ, denoted by $G(\psi)$, is the subgraph of $\mathrm{K}_{V^+\cup V^-}$ consisting of arcs (u^{τ_1}, v^{τ_2}) with $\psi(u^{\tau_1}, v^{\tau_2}) = 0$. Define the following four disjoint subsets of $V^+ \cup V^-$.

$$S^+ = \left\{v^+ \in V^+ \mid z(v) \le -\frac{\delta}{\alpha^+(v)}\right\}, \qquad \tilde{S}^- = \left\{v^- \in V^- \mid z(v) \le -\frac{\delta}{\alpha^-(v)}\right\},$$
$$\tilde{T}^+ = \left\{v^+ \in V^+ \mid z(v) \ge \frac{\delta}{\alpha^+(v)}\right\}, \qquad T^- = \left\{v^- \in V^- \mid z(v) \ge \frac{\delta}{\alpha^-(v)}\right\}. \tag{4}$$

A simple directed path (dipath) P in $G(\psi)$ from $S^+ \cup T^-$ to $\tilde{S}^- \cup \tilde{T}^+$ is called an *augmenting path*. The following procedure Single_Augment(δ', P, ψ) updates the flow ψ through a dipath P so that $\|z\|_\alpha$ gets smaller. For later use we prepare Procedure Single_Augment for any dipath P in $K_{V^+\cup V^-}$ (which may not be in $G(\psi)$).

Lemma 1 *Let ψ' be the flow obtained from ψ by* Single_Augment(δ', P, ψ). *Then we have $\partial_\alpha(\psi' - \psi) = \tau_1 \frac{\delta'}{\alpha^{\tau_1}(u)}\chi_u - \tau_2 \frac{\delta'}{\alpha^{\tau_2}(w)}\chi_w$, where u^{τ_1} and w^{τ_2} denote the initial vertex and the terminal vertex of P, respectively.*

Proof For any intermediate vertex v^{τ_3} in P, we have $\partial_\alpha(\psi' - \psi)(v) = -\tau_3 \frac{\delta'}{\alpha^{\tau_3}(v)} + \tau_3 \frac{\delta'}{\alpha^{\tau_3}(v)} = 0$. Also, $\partial_\alpha(\psi' - \psi)(u) = \tau_1 \frac{\delta'}{\alpha^{\tau_1}(u)}$ and $\partial_\alpha(\psi' - \psi)(w) = -\tau_2 \frac{\delta'}{\alpha^{\tau_2}(w)}$. □

Algorithm 1 Single_Augment(δ', P, ψ)

Input: A simple dipath P in $\mathrm{K}_{V^+ \cup V^-}$, a flow ψ in $\mathrm{K}_{V^+ \cup V^-}$, and $\delta' \in \mathbb{R}_{>0}$.
for each $(u^{\tau_1}, v^{\tau_2}) \in P$ **do**
  if $\psi(v^{\tau_2}, u^{\tau_1}) > 0$ **then**
    if $\delta' \leq \psi(v^{\tau_2}, u^{\tau_1})$ **then**
      $\psi(v^{\tau_2}, u^{\tau_1}) := \psi(v^{\tau_2}, u^{\tau_1}) - \delta'$
    else
      $\psi(v^{\tau_2}, u^{\tau_1}) := 0$ and $\psi(u^{\tau_1}, v^{\tau_2}) := \delta' - \psi(v^{\tau_2}, u^{\tau_1})$
  else
    $\psi(u^{\tau_1}, v^{\tau_2}) := \psi(u^{\tau_1}, v^{\tau_2}) + \delta'$
return ψ

Using the concept of augmenting path, a presumable algorithm would be described as follows. First, check whether $G(\psi)$ has an augmenting path P and call Single_Augment if such P exists. If there is no augmenting path, then we take the set W of vertices in $G(\psi)$ reachable from $S^+ \cup T^-$, and we would expect that W relates $\|x\|_\alpha$ and the minimum value of f within a tolerance measured by the scaling parameter δ. The following lemma more explicitly shows how W can be used.

Lemma 2 *Suppose that we have a δ-feasible flow ψ in $\mathrm{K}_{V^+ \cup V^-}$ and a vector $x \in \mathrm{P}^\alpha(f)$ expressed by (2), and that there is no augmenting path in $G(\psi)$. Let W be the set of vertices in $G(\psi)$ reachable from $S^+ \cup T^-$, and let $A = \{v \in V \mid v^+ \in W\}$ and $B = \{v \in V \mid v^- \in W\}$. Suppose that the following three conditions are satisfied:*

(W1) $(A, B) \in 3^V$,
(W2) *for each $i \in J$, $A \cup B$ precedes $V \backslash (A \cup B)$ in L_i, and*
(W3) *for each $i \in J$, $\sigma_i(v) = +$ for all $v \in A$ and $\sigma_i(v) = -$ for all $v \in B$.*

Then $\|z\|_\alpha \leq 4\beta n^2 \delta - f(A, B)$ and $\|x\|_\alpha \leq 6\beta n^2 \delta - f(A, B)$. Moreover, if $\delta < 1/(6\beta n^2)$ and f is integer-valued, then (A, B) is a minimizer of f.

Proof Due to the three conditions, we have $\langle y_i, \chi^\alpha_{(A,B)} \rangle = f(A, B)$ for all $i \in J$, and hence $\langle x, \chi^\alpha_{(A,B)} \rangle = f(A, B)$ by (2). Also note that $S^+ \cup T^- \subseteq W$, and hence from (4) $z(v) > -\delta/\alpha^+(v)$ for $v \notin A$ and $z(v) < \delta/\alpha^-(v)$ for $v \notin B$. Therefore we have

$$\begin{aligned}
\|z\|_\alpha &= - \sum_{v \in V : z(v) < 0} \alpha^+(v) z(v) + \sum_{v \in V : z(v) > 0} \alpha^-(v) z(v) \\
&\leq - \sum_{v \in A} \alpha^+(v) z(v) + \sum_{v \in B} \alpha^-(v) z(v) + 2\beta n \delta \\
&= -\langle x, \chi^\alpha_{(A,B)} \rangle - \langle \partial_\alpha \psi, \chi^\alpha_{(A,B)} \rangle + 2\beta n \delta \\
&\leq -f(A, B) + 2\beta n^2 \delta + 2\beta n \delta \\
&\leq -f(A, B) + 4\beta n^2 \delta.
\end{aligned} \tag{5}$$

Moreover, since $z = x + \partial_\alpha \psi$ and $\| - \partial_\alpha \psi \|_\alpha \leq 2\beta n^2 \delta$, it follows from (5) that

$$\|x\|_\alpha \leq \|x + \partial_\alpha \psi\|_\alpha + \| - \partial_\alpha \psi \|_\alpha \leq 6\beta n^2 \delta - f(A, B). \tag{6}$$

If $\delta < 1/(6\beta n^2)$, inequality (6) implies $f(A, B) - (-\|x\|_\alpha) < 1$. It follows from Theorem 1 that (A, B) is a minimizer of f if f is integer-valued. $\square$

Hence we now focus on how to achieve the conditions of Lemma 2 for W. It will turn out that in order to achieve the three conditions for W in Lemma 2 we need to introduce a stronger augmentation procedure beyond those used in bisubmodular function minimization [9,21]. This is because of the lack of skew-symmetry of $G(\psi)$.

Remark If there exist two dipaths such as

$$\begin{aligned} &v_1^{\tau_1} \to v_2^{\tau_2} \to \cdots \to v_\ell^{\tau_\ell} \\ &v_\ell^{-\tau_\ell} = u_{\ell'}^{\sigma_{\ell'}} \leftarrow \cdots \leftarrow u_2^{\sigma_2} \leftarrow u_1^{\sigma_1} \end{aligned}$$

then we can compose them so that the α-boundary at v_ℓ is equal to zero, and we may achieve an augmentation. Here, note that if $\tau_\ell = +$, to guarantee the δ-feasibility of updated flow ψ the value of augmentation for the second path should be $\delta \times \frac{\alpha^-(v_\ell)}{\alpha^+(v_\ell)}$, which may be $\delta \times \frac{1}{\beta}$.

For a simple dipath $P = (v_1^{\tau_1}, v_2^{\tau_2}, \ldots, v_\ell^{\tau_\ell})$ in $G(\psi)$, define $P^{-1} = (v_\ell^{-\tau_\ell}, v_{\ell-1}^{-\tau_{\ell-1}}, \ldots, v_1^{-\tau_1})$, a simple dipath in $\mathrm{K}_{V^+\cup V^-}$, where we do not care about whether P^{-1} exists in $G(\psi)$ as a dipath. For two dipaths P_1 and P_2 in $\mathrm{K}_{V^+\cup V^-}$ such that the terminal vertex of P_1 is the initial vertex of P_2, let $P_1 \circ P_2$ denote the concatenation of P_1 and P_2. We can define the concatenation of more than two dipaths in a natural way since the binary operation $\circ$ of concatenation is associative.

Let $\mathbf{P} = (P_1, \ldots, P_k)$ be a sequence of dipaths in $G(\psi)$. Suppose that $P_1 \circ P_2^{-1} \circ P_3 \circ \cdots \circ P_k^{(-1)^{k-1}}$ forms

- a walk in $\mathrm{K}_{V^+\cup V^-}$ from a vertex in $S^+ \cup T^-$ to a vertex in $\tilde{S}^- \cup \tilde{T}^+$ if k is odd, or
- a walk in $\mathrm{K}_{V^+\cup V^-}$ from a vertex in $S^+ \cup T^-$ to a vertex in $S^- \cup T^+$ if k is even.

Then we call $\mathbf{P} = (P_1, \ldots, P_k)$ an *augmenting path-sequence*. The number k is called the *length* of the augmenting path-sequence $\mathbf{P}$. (An augmenting path-sequence of length one is an augmenting path.) Let $v_0^{-\tau_0}$ be the initial vertex of P_1 and $v_i^{\tau_i}$ be the terminal vertex of $P_i^{(-1)^{i-1}}$ for each $i = 1, \ldots, k$. Then, if i is odd, P_i is a path from $v_{i-1}^{-\tau_{i-1}}$ to $v_i^{\tau_i}$, and otherwise P_i is a path from $v_i^{\tau_i}$ to $v_{i-1}^{-\tau_{i-1}}$. Define $p(i)$ $(i = 1, 2, \ldots, k)$ by

$$p(1) = 1, \qquad p(i) = \prod_{j=1}^{i-1} \frac{\alpha^{-\tau_j}(v_j)}{\alpha^{\tau_j}(v_j)} \quad (i = 2, \ldots, k). \tag{7}$$

We augment an appropriate flow value through each path P_i of the augmenting path-sequence $\mathbf{P} = (P_1, \ldots, P_k)$ so that non-zero α-boundary of the flow changing can appear only at the initial vertex of P_1 and the terminal (or initial) vertex of P_k when k is odd (or even). The details of the procedure are given in Augment as follows.

The following lemma shows how we can decrease $\|z\|_\alpha$ by calling Augment when we are given an augmenting path-sequence $(P_1, \ldots, P_k)$.

Algorithm 2 Augment$(\delta, (P_1, \ldots, P_k), \psi)$

Input: An augmenting path-sequence $(P_1, \ldots, P_k)$, where P_i is a path in $G(\psi)$ from $v_{i-1}^{-\tau_{i-1}}$ to $v_i^{\tau_i}$ for odd $i \in \{1, \ldots, k\}$ and P_i is from $v_i^{\tau_i}$ to $v_{i-1}^{-\tau_{i-1}}$ for even $i \in \{1, \ldots, k\}$.
1: $p(1) := 1$
2: **for** $i = 2, \ldots, k$ **do**
3: $\quad p(i) := \frac{\alpha^{-\tau_{i-1}}(v_{i-1})}{\alpha^{\tau_{i-1}}(v_{i-1})} p(i-1)$
4: $\pi := \max\{p(i) \mid 1 \leq i \leq k\}$
5: Single_Augment$\left(\frac{p(i)}{k\pi}\delta, P_i, \psi\right)$ for $i = 1, \ldots, k$.
6: **return** ψ.

Lemma 3 *Let ψ' be the new flow in $\mathrm{K}_{V^+ \cup V^-}$ obtained by* Augment$(\delta, (P_1, \ldots, P_k), \psi)$ *from ψ through an augmenting path-sequence $(P_1, \ldots, P_k)$. Then ψ' is δ-feasible and $\|z\|_\alpha$ decreases by at least $\delta/(k\beta^{\lceil \frac{k}{2} \rceil - 1})$, where $\beta = \max\{\frac{\alpha^+(v)}{\alpha^-(v)} \mid v \in V\} (\geq 1)$.*

Proof Define $\delta_i = p(i)\delta/(k\pi)$ for each $i = 1, \ldots, k$. Then $\delta_i \leq \delta/k$ holds for all i since $p(i) \leq \pi$. Since each arc appears at most k times in total in the augmenting path-sequence of length k, ψ' is δ-feasible by the way of computing ψ' through Augment$(\delta, (P_1, \ldots, P_k), \psi)$. Also observe that from (7) and the definition of π we have

$$\delta_{i+1} = \frac{\alpha^{-\tau_i}(v_i)}{\alpha^{\tau_i}(v_i)} \delta_i \qquad (\forall i = 1, \ldots, k-1), \tag{8}$$

$$\delta_1 \geq \frac{\delta}{k\beta^{j-1}}, \quad \delta_k \geq \frac{\delta}{k\beta^{k-j}} \qquad (\forall j \in \mathrm{Argmax}\{p(i) \mid i \in \{1, \ldots, k\}\}). \tag{9}$$

Let us evaluate $\partial_\alpha \psi' - \partial_\alpha \psi$. By Lemma 1 and (8),

$$\begin{aligned}
&\partial_\alpha \psi' - \partial_\alpha \psi \ (= \partial_\alpha(\psi' - \psi)) \\
&= \sum_{i=1}^{k} (-1)^{i-1} \left(\frac{\tau_{i-1}\delta_i}{\alpha^{-\tau_{i-1}}(v_{i-1})} \chi_{v_{i-1}} - \frac{\tau_i \delta_i}{\alpha^{\tau_i}(v_i)} \chi_{v_i} \right) \\
&= \frac{\tau_0 \delta_1}{\alpha^{-\tau_0}(v_0)} \chi_{v_0} - (-1)^{k-1} \frac{\tau_k \delta_k}{\alpha^{\tau_k}(v_k)} \chi_{v_k} + \sum_{i=1}^{k-1} (-1)^{i-1} \left(-\frac{\tau_i \delta_i}{\alpha^{\tau_i}(v_i)} + \frac{\tau_i \delta_{i+1}}{\alpha^{-\tau_i}(v_i)} \right) \chi_{v_i} \\
&= \frac{\tau_0 \delta_1}{\alpha^{-\tau_0}(v_0)} \chi_{v_0} - (-1)^{k-1} \frac{\tau_k \delta_k}{\alpha^{\tau_k}(v_k)} \chi_{v_k}.
\end{aligned} \tag{10}$$

Putting $z' := x + \partial_\alpha \psi'$, we have $z(v) = z'(v)$ for all $v \in V \backslash \{v_0, v_k\}$ by (10). Observe also that the sign of $z'(v_0)$ is equal to that of $z(v_0)$, which is equal to τ_0 since $v_0^{-\tau_0} \in S^+ \cup T^-$. Similarly, the sign of $z'(v_k)$ is equal to that of $z(v_k)$, which is equal to $(-1)^{k-1}\tau_k$ since $v_k^{\tau_k} \in S^+ \cup T^-$ if k is even and otherwise $v_k^{\tau_k} \in \tilde{S}^- \cup \tilde{T}^+$. Hence, we get

$$\|z'\|_\alpha - \|z\|_\alpha = -|\alpha^{-\tau_0}(v_0)\partial_\alpha(\psi' - \psi)(v_0)| - |\alpha^{(-1)^k \tau_k}(v_k)\partial_\alpha(\psi' - \psi)(v_k)|.$$

Combining this with (10) and (9), we have

$$\|z'\|_\alpha - \|z\|_\alpha = \begin{cases} -\delta_1 - \delta_k \le -\left(\frac{1}{\beta^{j-1}} + \frac{1}{\beta^{k-j}}\right)\frac{\delta}{k} & \text{(if } k \text{ is even)} \\ -\delta_1 - \frac{\alpha^{-\tau_k}(v_k)}{\alpha^{\tau_k}(v_k)}\delta_k \le -\left(\frac{1}{\beta^{j-1}} + \frac{1}{\beta^{k-j+1}}\right)\frac{\delta}{k} & \text{(if } k \text{ is odd)} \end{cases}$$

for some j with $1 \le j \le k$. Note that $\min\{j-1, k-j\} \le k/2 - 1$ if k is even and $\min\{j-1, k-j+1\} \le k/2 - 1/2 = \lceil k/2 \rceil - 1$ if k is odd. Hence, we get

$$\|z\|_\alpha - \|z'\|_\alpha \ge \frac{\delta}{k\beta^{\lceil \frac{k}{2} \rceil - 1}}.$$

□

Lemma 3 implies that the value of augmentation may be exponentially small, which causes a trouble in constructing a polynomial algorithm. However, fortunately we can show a crucial fact that it suffices to consider augmenting path-sequences of length at most four. Our algorithm checks whether $G(\psi)$ has an augmenting path-sequence of length $k \le 4$. If there exists such an augmenting path-sequence, the algorithm calls Augment to update ψ. On the other hand, if there exists no augmenting path-sequence of length $k \le 4$, then we compute the set W of vertices in $G(\psi)$ reachable from $S^+ \cup T^-$ [by dipaths in $G(\psi)$].

Let R be the set of vertices in $G(\psi)$ from which we can reach some vertex in $\{v^{-\tau} \in V^+ \cup V^- \mid v^\tau \in W\}$. We then have the following.

Lemma 4 *$G(\psi)$ has an augmenting path-sequence of length $k \le 4$ if one of the following three holds*:

(i) $W \cap R \neq \emptyset$;
(ii) *W is not consistent*;
(iii) *R is not consistent.*

Proof (ii): Suppose that there is a vertex $v \in V$ with $\{v^+, v^-\} \subseteq W$. Then there are a path P_1 from $S^+ \cup T^-$ to v^+ and a path P_2 from $S^+ \cup T^-$ to v^-, so that (P_1, P_2) is an augmenting path-sequence of length 2.

(i): This follows from (ii) since a vertex that is reachable from $W \cap R$ is still contained in W.

(iii): Suppose that there is a vertex $v \in V$ with $\{v^+, v^-\} \subseteq R$. Then there are a path P_2 from v^+ to a vertex u^{τ_1} with $u^{-\tau_1} \in W$ and a path P_3 from v^- to a vertex w^{τ_2} with $w^{-\tau_2} \in W$. Hence there are a path P_1 from $S^+ \cup T^-$ to $u^{-\tau_1}$ and a path P_4 from $S^+ \cup T^-$ to $w^{-\tau_2}$. Observe that (P_1, P_2, P_3, P_4) is an augmenting path-sequence of length 4. □

It follows from Lemma 4 (ii) that if there is no augmenting path-sequence of length $k \le 4$, then W satisfies condition (W1) of Lemma 2. If W violates (W2) or (W3), we call procedure Double_Exchange or Tail_Exchange defined below to improve the situation. These two procedures are direct adaptations of those devised for bisubmodular function minimization in [9], where Double_Exchange originally appeared in [17].

Suppose that we are given an expression (2) for current $x \in \mathrm{P}_\alpha(f)$, where recall that each extreme point y_i $(i \in J)$ of $\mathrm{P}_\alpha(f)$ is generated by a linear ordering L_i and a sign function σ_i on V. We say that a triple (i, u, v) with $i \in J$ and $u, v \in V$ is *active* if

(a) u immediately succeeds v in L_i and
(b) $u^{\sigma_i(u)} \in W$ and $v^{\sigma_i(v)} \notin W$, or $u^{\sigma_i(u)} \notin R$ and $v^{\sigma_i(v)} \in R$.

If such an active triple exists, we perform procedure $\mathsf{Double\_Exchange}(i, u, v)$.

Algorithm 3 $\mathsf{Double\_Exchange}(i, u, v)$

Input: An active triple (i, u, v)
1: $t := f(L_i(u)\backslash\{v\}|\sigma_i) - f(L_i(u)|\sigma_i) + \sigma_i(v)\alpha^{\sigma_i(v)}y_i(v)$
2: $s := \min\{\delta, \lambda_i t\}$ (where λ_i is as given in (2))
3: **if** $s < \lambda_i t$ **then**
4: $\quad k$: a new index
5: $\quad J := J \cup \{k\}$
6: $\quad \lambda_k := \lambda_i - s/t$
7: $\quad \lambda_i := s/t$
8: $\quad y_k := y_i$
9: $\quad L_k := L_i$
10: $\quad \sigma_k := \sigma_i$
11: Update L_i to be the linear ordering obtained from L_i by interchanging u and v.
12: $y_i := y_i + t\left(\frac{\sigma_i(u)\chi_u}{\alpha^{\sigma_i(u)}(u)} - \frac{\sigma_i(v)\chi_v}{\alpha^{\sigma_i(v)}(v)}\right)$
13: $x := x + s\left(\frac{\sigma_i(u)\chi_u}{\alpha^{\sigma_i(u)}(u)} - \frac{\sigma_i(v)\chi_v}{\alpha^{\sigma_i(v)}(v)}\right)$
14: **if** $s \geq \psi(u^{\sigma_i(u)}, v^{\sigma_i}(v))$ **then**
15: $\quad \psi(v^{\sigma_i(v)}, u^{\sigma_i(u)}) := s - \psi(u^{\sigma_i(u)}, v^{\sigma_i(v)})$
16: $\quad \psi(u^{\sigma_i(u)}, v^{\sigma_i(v)}) := 0$
17: **else**
18: $\quad \psi(u^{\sigma_i(u)}, v^{\sigma_i(v)}) := \psi(u^{\sigma_i(u)}, v^{\sigma_i(v)}) - s$

We now give the detail of procedure $\mathsf{Double\_Exchange}(i, u, v)$. Given an active triple (i, u, v), let L'_i be the linear ordering obtained from L_i by interchanging u and v, and let y'_i be the extreme point associated with L'_i and σ_i. Then the vector

$$\sigma_i(u)\frac{1}{\alpha^{\sigma_i(u)}(u)}\chi_u - \sigma_i(v)\frac{1}{\alpha^{\sigma_i(v)}(v)}\chi_v$$

is an edge vector of the edge of $\mathrm{P}^\alpha(f)$ connecting adjacent y_i and y'_i unless $y_i = y'_i$. The number t defined in Line 1 of $\mathsf{Double\_Exchange}(i, u, v)$ is nothing but the one satisfying

$$y'_i = y_i + t\left(\sigma_i(u)\frac{1}{\alpha^{\sigma_i(u)}(u)}\chi_u - \sigma_i(v)\frac{1}{\alpha^{\sigma_i(v)}(v)}\chi_v\right).$$

If $t \neq 0$, $\lambda_i y_i$ is updated to $(\lambda_i - \frac{s}{t})y_i + \frac{s}{t}y'_i$ with s defined in Line 2, and ψ is updated so that z does not change, as will be shown in the following lemma.

We say that $\mathsf{Double\_Exchange}(i, u, v)$ is *saturating* if $s = \lambda_i t$ holds at Line 2, and otherwise *non-saturating*.

Lemma 5 *Vector z remains the same by* $\mathsf{Double\_Exchange}(i, u, v)$. *Moreover, a new vertex joins W or R after non-saturating* $\mathsf{Double\_Exchange}(i, u, v)$*,*

Proof Let z, x and ψ be those obtained before performing $\mathsf{Double\_Exchange}(i, u, v)$ and let z', x' and ψ' be the new ones obtained after $\mathsf{Double\_Exchange}(i, u, v)$. Then,

$$x' = x + s\left(\sigma_i(u)\frac{1}{\alpha^{\sigma_i(u)}(u)}\chi_u - \sigma_i(v)\frac{1}{\alpha^{\sigma_i(v)}(v)}\chi_v\right).$$

Also, $\psi(u^{\sigma_i(u)}, v^{\sigma_i(v)})$ is decreased by s, where in effect, if $s \geq \psi(u^{\sigma_i(u)}, v^{\sigma_i(v)})$, the flow value $\psi(u^{\sigma_i(u)}, v^{\sigma_i(v)})$ is put to be zero and $\psi(v^{\sigma_i(v)}, u^{\sigma_i(u)})$ of the reversed arc $(v^{\sigma_i(v)}, u^{\sigma_i(u)})$ is increased from zero to $s - \psi(u^{\sigma_i(u)}, v^{\sigma_i(v)})$, to keep $\psi \geq \mathbf{0}$. Therefore,

$$\partial_\alpha \psi' - \partial_\alpha \psi = -s\left(\sigma_i(u)\frac{1}{\alpha^{\sigma_i(u)}(u)}\chi_u - \sigma_i(v)\frac{1}{\alpha^{\sigma_i(v)}(v)}\chi_v\right)$$

due to definition (3) of ∂_α. This implies $z' = x' + \partial_\alpha \psi' = x + \partial_\alpha \psi = z$.

To see the second statement, suppose $\mathsf{Double\_Exchange}(i, u, v)$ is non-saturating. Then $s = \delta$ holds at Line 2. Hence $\psi(u^{\sigma_i(u)}, v^{\sigma_i(v)}) = 0$ holds at Line 16, and a new arc $(u^{\sigma_i}(u), v^{\sigma_i(v)})$ emerges in updated $G(\psi)$. If $u^{\sigma_i(u)} \in W$ and $v^{\sigma_i(v)} \notin W$, $v^{\sigma_i(v)}$ is newly included in W, while if $u^{\sigma_i(u)} \notin R$ and $v^{\sigma_i(v)} \in R$, then $u^{\sigma_i(u)}$ is newly included in R. $\square$

A pair (i, v) of $i \in J$ and $v \in V$ is called *active* if v is the last element in L_i and $v^{\sigma_i(v)} \in R$. If such an active pair exists, we perform $\mathsf{Tail\_Exchange}(i, v)$.

Given an active pair (i, v), let σ_i' be the sign function obtained from σ_i by changing the sign of $\sigma_i(v)$, and let y_i' be the extreme point associated with L_i and σ_i'. Then t computed in Line 2 of $\mathsf{Tail\_Exchange}$ is determined so that the following relation holds:

$$y_i'(v) = y_i(v) + t\sigma_i'(v)\left(\frac{1}{\alpha^{\sigma_i'(v)}} + \frac{1}{\alpha^{-\sigma_i'(v)}}\right)\chi_v.$$

We say that $\mathsf{Tail\_Exchange}(i, v)$ is *saturating* if $s = \lambda_i t$ holds at Line 3, and otherwise *non-saturating*.

Lemma 6 *Vector z remains the same by* $\mathsf{Tail\_Exchange}(i, v)$. *Moreover, a new augmenting path-sequence of length four appears as a result of non-saturating* $\mathsf{Tail\_Exchange}(i, v)$.

Proof The first claim can be checked in the same manner as in the proof of Lemma 5.

To see the second claim, let (i, v) be the active pair on which $\mathsf{Tail\_Exchange}$ is performed with $\tau := \sigma_i(v)$. If the present $\mathsf{Tail\_Exchange}(i, v)$ is non-saturating, then we have $s = \delta$ at Line 3. Also, in the case of non-saturating $\mathsf{Tail\_Exchange}(i, v)$,

Algorithm 4 Tail_Exchange(i, v)

Input: An active pair (i, v)

1: $\sigma_i(v) := -\sigma_i(v)$

2: $t := \frac{\alpha^{-\sigma_i(v)}}{\alpha^{\sigma_i(v)}+\alpha^{-\sigma_i(v)}}[f(V \mid \sigma_i) - f(V\backslash\{v\} \mid \sigma_i)] - \frac{\sigma_i(v)\alpha^{\sigma_i(v)}\alpha^{-\sigma_i(v)}}{\alpha^{\sigma_i(v)}+\alpha^{-\sigma_i(v)}} y_i(v)$

3: $s := \min\{\delta, \lambda_i t\}$

4: **if** $s < \lambda_i t$ **then**

5: $\quad k$: a new index

6: $\quad J := J \cup \{k\}$

7: $\quad \lambda_k := \lambda_i - s/t$

8: $\quad \lambda_i := s/t$

9: $\quad y_k := y_i$

10: $\quad L_k := L_i$

11: $\quad \sigma_k := \sigma_i$ and $\sigma_k(v) := -\sigma_k(v)$

12: $y_i := y_i + t\sigma_i(v)\left(\frac{1}{\alpha^{\sigma_i(v)}} + \frac{1}{\alpha^{-\sigma_i(v)}}\right)\chi_v$

13: $x := x + s\sigma_i(v)\left(\frac{1}{\alpha^{\sigma_i(v)}} + \frac{1}{\alpha^{-\sigma_i(v)}}\right)\chi_v$

14: **if** $s \geq \psi(v^{\sigma_i(v)}, v^{-\sigma_i(v)})$ **then**

15: $\quad \psi(v^{-\sigma_i(v)}, v^{\sigma_i(v)}) := s - \psi(v^{-\sigma_i(v)}, v^{\sigma_i(v)}))$

16: $\quad \psi(v^{\sigma_i(v)}, v^{-\sigma_i(v)}) := 0$

17: **else**

18: $\quad \psi(v^{\sigma_i(v)}, v^{-\sigma_i(v)}) := \psi(v^{\sigma_i(v)}, v^{-\sigma_i(v)}) - s$

$\psi(v^{-\tau}, v^{\tau}) = 0$ holds at Line 16, which means that a new arc $(v^{-\tau}, v^{\tau})$ emerges in updated $G(\psi)$. Hence, in the resulting $G(\psi)$, we have $\{v^-, v^+\} \subseteq R$. This implies that $G(\psi)$ has an augmenting path-sequence of length at most four by Lemma 4 (iii). □

Moreover, we have the following.

Lemma 7 *Let W be the set of vertices in $G(\psi)$ reachable from $S^+ \cup T^-$. Suppose that there is no augmenting path-sequence of length $k \leq 4$ and there is neither an active triple nor an active pair. Then, letting $A = \{v \in V \mid v^+ \in W\}$ and $B = \{v \in V \mid v^- \in W\}$, (A, B) together with L_i and σ_i for all $i \in J$ satisfies the three conditions* (W1)*,* (W2) *and* (W3) *in* Lemma 2.

Proof It follows from the present assumption and (ii) in Lemma 4 that there is no $v \in V$ with $\{v^+, v^-\} \subseteq W$, which means that condition (W1) holds, i.e., $(A, B) \in 3^V$.

Condition (W2) of Lemma 2 easily follows as there is no active triple.

To see that condition (W3) of Lemma 2 is satisfied, suppose to the contrary that there are $i \in J$ and $v \in V$ such that $v^{-\sigma_i(v)} \in W$. Then $v^{\sigma_i(v)} \in R$. Since there is no active triple, there should hold $u^{\sigma_i(u)} \in R$ for the element u next to v in L_i. Hence, continuing this argument, we conclude that $w^{\sigma_i(w)} \in R$ for the last element w in L_i. However, this implies that (i, w) is an active pair, which contradicts the assumption, so that condition (W3) of Lemma 2 holds. □

Summarizing the discussion so far, we are now ready to describe the whole algorithm, weakly-ABSFM(f). The main body of the algorithm will also be used in the strongly polynomial time algorithm given in the next section, and hence we shall refer to it as REFINE. An iteration of the while-loop in REFINE (i.e., lines 3–21) corresponds to a *scaling phase* with a scaling parameter δ discussed above.

Algorithm 5 weakly-ABSFM (f)

L_0 : a linear ordering on V
σ_0 : a sign function on V
x : an extreme point of $\mathrm{P}^\alpha(f)$ generated by L_0 and σ_0
$J := \{1\}$, $y_1 := x$, $\lambda_1 := 1$, $\psi = \mathbf{0}$
$\delta := \frac{\|x\|_\alpha}{\beta n^2}$
$\zeta := \frac{1}{6\beta n^2}$
return REFINE (f, x, δ, ζ)

Algorithm 6 REFINE (f, x, δ, ζ)

Input: an α-bisubmodular function f, a point $x \in \mathrm{P}^\alpha(f)$ along with its expression as a convex combination of extreme points of $\mathrm{P}^\alpha(f)$ as in (2), and $\delta > \zeta > 0$.
$\psi := 0$
while $\delta \geq \zeta$ **do**
  $\delta := \delta/2$
  for all $(u^{\tau_1}, v^{\tau_2}) \in V^+ \cup V^- \times V^+ \cup V^-$ **do**
    $\psi(u^{\tau_1}, v^{\tau_2}) := \delta$ if $\psi(u^{\tau_1}, v^{\tau_2}) > \delta$
  repeat
    $S^+ := \{v^+ \in V^+ \mid x(v) + \partial_\alpha\psi(v) \leq -\frac{\delta}{\alpha^+(v)}\}$
    $T^- := \{v^- \in V^- \mid x(v) + \partial_\alpha\psi(v) \geq \frac{\delta}{\alpha^-(v)}\}$
    W : the set of vertices reachable from $S^+ \cup T^-$ in $G(\psi)$
    R : the set of vertices from which we can reach $\{v^{-\tau} \in V^+ \cup V^- \mid v^\tau \in W\}$
    $A := \{v \in V \mid v^+ \in W\}$
    $B := \{v \in V \mid v^- \in W\}$
    if $\exists(P_1, \ldots, P_k)$: an augmenting path-sequence of length $k \leq 4$ **then**
      Augment $(\delta, (P_1, \ldots, P_k), \psi)$
      Reduce x (i.e., express x as a convex combination of at most $|V| + 1$ extreme points)
    else
      Compute the set Q of active pairs and active triples in $G(\psi)$.
      if $Q \neq \emptyset$ **then**
        Take $(i, u, v) \in Q$ or $(i, v) \in Q$.
        Double_Exchange (i, u, v) or Tail_Exchange (i, v).
  until $\nexists$augmenting path-sequence of length at most four and $Q = \emptyset$
return (A, B) and x

Although the above algorithm checks the existence of augmenting path-sequences of length at most four, according to the correctness proof it actually works even if allowable sequences are restricted to those of length two or four.

3.2 Analysis

We still assume that f is real-valued. Lemmas 8–10 and Theorem 2 hold for real-valued f.

Lemma 8 *At the end of each scaling phase of* REFINE, *we have* $(A, B) \in 3^V$, *and* $z := x + \partial_\alpha\psi$ *satisfies* $\|z\|_\alpha \leq 4\beta n^2\delta - f(A, B)$ *and* $\|x\|_\alpha \leq 6\beta n^2\delta - f(A, B)$.

Proof The present lemma follows from Lemmas 2 and 7. □

Lemma 9 *Suppose that* $x \in \mathrm{P}^\alpha(f)$ *and* $\delta > 0$ *satisfy* $\|x\|_\alpha + f(A, B) \leq 6\beta n^2\delta$ *for some* $(A, B) \in 3^V$. *Then each scaling phase of* REFINE(f, x, δ, ζ) *carries out* $\mathrm{O}(\beta^2 n^2)$ *augmentations.*

Proof Observe first that, at the beginning of a scaling phase (before reducing δ by half),

$$\|z\|_\alpha + f(A, B) \leq 6n^2\beta\delta \tag{11}$$

for some $(A, B) \in 3^V$. Indeed, if the scaling phase is the initial phase of the algorithm, then (11) follows from the lemma assumption and $x = z$. Otherwise by Lemma 8 the pair (A, B) obtained at the end of the previous scaling phase satisfies $\|z\|_\alpha \leq 6n^2\beta\delta - f(A, B)$.

Now the first step of the scaling phase reduces δ by half, and hence $\|z\|_\alpha \leq 12n^2\beta\delta - f(A, B)$ for the new δ.

At the end of the scaling phase we have $\|z\|_\alpha \geq -\langle z, \chi^\alpha_{(A,B)}\rangle \geq -\langle x, \chi^\alpha_{(A,B)}\rangle - 2\beta n^2\delta \geq -f(A, B) - 2\beta n^2\delta$. Therefore, $\|z\|_\alpha$ decreases by at most $14\beta n^2\delta$. Since $\|z\|_\alpha$ decreases by at least $\delta/(4\beta)$ by each Augment through an augmenting path-sequence of length $k \leq 4$, the number of augmentations is bounded by $\mathrm{O}(\beta^2 n^2)$. □

Lemma 10 REFINE *carries out saturating* Double_Exchange $\mathrm{O}(n^3)$ *times, non-saturating* Double_Exchange $\mathrm{O}(n)$ *times, saturating* Tail_Exchange $\mathrm{O}(n^2)$ *times, and non-saturating* Tail_Exchange *at most once, between consecutive augmentations.*

Proof We should remark that, due to Reduce, $|J| = \mathrm{O}(n)$ holds after every augmentation.

By Lemma 6, the algorithm carries out non-saturating Tail_Exchange at most once between augmentations. By Lemma 5, $W \cup R$ becomes larger after a non-saturating Double_Exchange. Hence non-saturating Double_Exchange is performed at most $2n$ times. Since new L_k and σ_k arise only as a result of non-saturating Double_Exchange, $|J| = \mathrm{O}(n)$ holds between augmentations.

Notice that, if Double_Exchange(i, u, v) for an active triple (i, u, v) is performed and is saturating, then triple (i, u, v) never becomes active again till the next augmentation. This means that saturating Double_Exchange is performed $\mathrm{O}(n^3)$ times since $|J| = \mathrm{O}(n)$. Similarly, saturating Tail_Exchange is performed $\mathrm{O}(n^2)$ times between augmentations. □

Theorem 2 *Let* $f : 3^V \to \mathbb{R}$ *be an* α*-bisubmodular function with* $f(\emptyset, \emptyset) = 0$, $y \in \mathrm{P}^\alpha(f)$, *and* $\delta > \zeta > 0$. *If*

$$\|y\|_\alpha + f(S, T) \leq 6\beta n^2\delta$$

for some $(S, T) \in 3^V$, *then* REFINE (f, y, δ, ζ) *outputs* $(A, B) \in 3^V$ *and* $x \in \mathrm{P}^\alpha(f)$ *such that*

$$\|x\|_\alpha + f(A, B) \leq 6\beta n^2\zeta$$

with $\mathrm{O}(\beta^2 n^5 \log \frac{\delta}{\zeta})$ *function evaluations and arithmetic operations.*

Proof The algorithm has $\mathrm{O}(\log \frac{\delta}{\zeta})$ scaling phases. In each scaling phase, by Lemma 9, the algorithm carries out Augment and Reduce $\mathrm{O}(\beta^2 n^2)$ times. Each Reduce takes $\mathrm{O}(n^3)$ running time, while each Augment requires $\mathrm{O}(n)$ running time. By Lemma 10, between consecutive augmentations the algorithm carries out Double_Exchange and Tail_Exchange $\mathrm{O}(n^3)$ times. Since $|J| = \mathrm{O}(n)$, the total running time for updating S^+, T^-, A, B, and Q between consecutive augmentations is $\mathrm{O}(n^3)$. Therefore, the number of function evaluations and arithmetic operations is bounded as stated in the present theorem.

Moreover, by Lemma 8 we have $\|x\|_\alpha < 6\beta n^2 \zeta - f(A, B)$ at the end. □

Now we assume that f is integer-valued.

Theorem 3 *Let* $f: 3^V \to \mathbb{Z}$ *be an* α*-bisubmodular function with* $f(\emptyset, \emptyset) = 0$. *Then* weakly-ABSFM (f) *finds a minimizer of* f *in* $\mathrm{O}(\beta^2 n^5 \log \beta n M)$ *function evaluations and arithmetic operations, where* $M = \max\{f(X, Y) \mid (X, Y) \in 3^V\}$.

Proof At the end of the algorithm, we have $\|x\|_\alpha \le 6\beta n^2 \zeta - f(A, B) < 1 - f(A, B)$ by Theorem 2. The present theorem follows from Theorem 1 since f is integer-valued. □

4 Strongly polynomial algorithm

In this section we show how to make the weakly polynomial algorithm given in the previous section strongly polynomial for real-valued α-bisubmodular functions.

Let us consider an α-bisubmodular function $f: 3^V \to \mathbb{R}$ as before. As in the bisubmodular function minimization, the algorithm tries to collect two types of information: elements which are not included in any minimizer of f and pairs of elements for which every minimizer containing one always contains the other. This information will be stored in a set U_{e} of excluded elements and a *conditioning graph* $H = (W, C)$, which will be explained in the next subsection. A key parameter that controls the next procedure in the algorithm is δ_1, which is defined based on the marginal gain of f on the strongly connected components of H. By definition δ_1 is nonnegative, and we show that, if $\delta_1 = 0$, then a signed set that corresponds to a maximal consistent ideal in H is a minimizer of f. On the other hand, if $\delta_1 > 0$, H can be updated (by adding a new arc or deleting at least one node) by using REFINE given in the last section. A detailed description will be given in Sect. 4.3.

4.1 Conditioning graph

The algorithm keeps $U_{\mathrm{e}} \subseteq V$ and a digraph $H = (W, C)$ on $W := (V \backslash U_{\mathrm{e}})^+ \cup (V \backslash U_{\mathrm{e}})^-$. The set U_{e} denotes a set of elements which are currently known to be included in none of the minimizers of f, while H denotes the diagram of logical implications such that

$$(u^\sigma, v^\tau) \in C \text{ implies that every minimizer of } f \text{ containing } u^\sigma \text{ contains } v^\tau. \quad (12)$$

Since elements of U_e do not affect the set of minimizers, we may always update $V \leftarrow V \setminus U_e$, and omit to mention U_e if it is clear from the context.

Initially we have a conditioning graph $H = (W, C)$ with $C = \emptyset$. Assuming that H keeps property (12) we can impose extra properties of H. The following lemma, which is a generalization of [21, Lemma 2.1], is used to ensure those properties.

Lemma 11 *For any distinct $u, v \in V$, if every minimizer of f_{u^σ} contains v^τ, then every minimizer of f containing $v^{-\tau}$ contains $u^{-\sigma}$.*

Proof Suppose to the contrary that there exists a minimizer (X, Y) of f that contains $v^{-\tau}$ but not $u^{-\sigma}$. Let (S, T) be a minimizer of f_{u^σ}. Then we have $u^\sigma, v^\tau \in (S, T)$, due to the assumption.

Note that u^σ is contained in $(S, T) \cup_{t_i} (X, Y)$ for all $i = 0, \ldots, p$, and hence

$$f(S, T) \le f((S, T) \cup_{t_i} (X, Y)) \qquad (\forall i = 0, \ldots, p). \tag{13}$$

On the other hand, v^τ is not contained in $(S, T) \cup_{t_i} (X, Y)$ for any i such that $0 \le t_i < \frac{\alpha^-(v)}{\alpha^+(v)}$. (Note that $i = 0$ is always among those is.) For such i the inequality (13) holds with strict inequality by the assumption. Hence we have

$$f(S, T) < \sum_{i=0}^{p} (t_{i+1} - t_i) f((S, T) \cup_{t_i} (X, Y)). \tag{14}$$

By the α-bisubmodularity of f we have

$$f(S, T) + f(X, Y) \ge f((S, T) \cap (X, Y)) + \sum_{i=0}^{p} (t_{i+1} - t_i) f((S, T) \cup_{t_i} (X, Y)). \tag{15}$$

It follows from (14) and (15) that $f(X, Y) > f((S, T) \cap (X, Y))$, which contradicts that (X, Y) is a minimizer of f. $\square$

For a vertex v^σ in H, let $R(v^\sigma)$ be the set of vertices reachable from v^σ in H. We say that H is *skew-symmetric* if $(u^\tau, v^\sigma) \in C$ implies $(v^{-\sigma}, u^{-\tau}) \in C$ for $u \ne v$. Starting from $C = \emptyset$, the algorithm will insert new arcs in H keeping the skew-symmetry. More specifically, Algorithm 17 given in Sect. 4.3 has two possible cases for the update of H:

Case 1: (Line 12–13) It finds an element u^τ that is contained in every minimizer of $f_{R(D)}$, where D is a strongly connected component in H, $R(D)$ denotes the set of vertices reachable from some vertex in D, and $f_{R(D)}$ denotes the contraction of f by $R(D)$. This implies that every minimizer of f_{v^σ} also contains u^τ for every $v^\sigma \in D$ by (12). Thus every minimizer of f containing v^σ contains u^τ, and we can add (v^σ, u^τ) to C. By Lemma 11, we can also add $(u^{-\tau}, v^{-\sigma})$ to C to keep the skew-symmetry of H.

Case 2: (Line 14–16) It finds a set F of elements that are not contained in any minimizer of f. We can add $(u^\tau, u^{-\tau})$ to C for each $u^\tau \in F$. (Clearly H is still skew-symmetric).

Since H is skew-symmetric, we have the following implication:

$$\text{If } (u^\sigma, v^\tau) \in C \text{ and } (u^{-\sigma}, u^\sigma) \in C, \text{ then no minimizer of } f \text{ contains } v^{-\tau}.$$

Indeed, by the skew-symmetry of H, $(u^\sigma, v^\tau) \in C$ implies $(v^{-\tau}, u^{-\sigma}) \in C$ and hence any minimizer containing $v^{-\tau}$ would contain $u^{-\sigma}$, contradicting $(u^{-\sigma}, u^\sigma) \in C$. Thus we may further perform the following update keeping (12):

- If $(u^\tau, v^\sigma) \in C$ and $(u^{-\tau}, u^\tau) \in C$, then add $(v^{-\sigma}, v^\sigma)$ to C (if it does not exist).
- If $R(u^\tau)$ is not consistent (i.e., $\exists v \in V$ with $\{v^+, v^-\} \subseteq R(u^\tau)$), then add $(u^\tau, u^{-\tau})$ to C (if it does not exist).
- If $(u^\tau, u^{-\tau}) \in C$ and $(u^{-\tau}, u^\tau) \in C$, then delete u from the ground set (i.e., add u to U_e and update V).

In total, every time H gets new arcs, the algorithm performs the above update of H so that it satisfies the following four extra properties:

$$\begin{aligned} &H \text{ is skew-symmetric.} \\ &\text{If } R(u^\tau) \text{ is not consistent, } (u^\tau, u^{-\tau}) \in C. \\ &\text{There is no } u \in V \text{ with } (u^\tau, u^{-\tau}), (u^{-\tau}, u^\tau) \in C. \\ &\text{If } (u^{-\tau}, u^\tau) \in C, \text{ then } (v^{-\sigma}, v^\sigma) \in C \text{ for every } v^\sigma \in R(u^\tau). \end{aligned} \tag{16}$$

4.2 Parameter δ_1

In the subsequent discussion, we shall assign a label i for each strongly connected component H_i in H. For each component H_i, the vertex set and the edge set of H_i are denoted by W_i and C_i, respectively, and the set of vertices that are reachable from W_i in H is denoted by D_i. We set

$$I := \{i : D_i \text{ is consistent}\}.$$

We say that $Z \subseteq W$ is an *ideal* of $H = (W, C)$ if there is no arc $(u^\sigma, v^\tau) \in C$ leaving Z. It is known that the collection $\mathscr{R}(H)$ of all consistent ideals of H (regarded as signed subsets of V) is closed with respect to binary operations $\cap$ and $\cup_0$, i.e., $\mathscr{R}(H)$ is a *signed ring family*. However, $\mathscr{R}(H)$ may not be closed with respect to $\cup_t$ in general.

We say $(X, Y) \in 3^V$ (or its corresponding $X^+ \cup Y^-$) *spans* V if $X \cup Y = V$. We remark the following.

Lemma 12 *Any maximal consistent ideal of H spans V.*

Proof Let $U_0 = \{v^\sigma \in W : (v^{-\sigma}, v^\sigma) \in C\}$. By (16), U_0 is consistent and there is no arc from U_0 to $W \setminus U_0$.

We show that, if a consistent ideal U does not span V, then it is not maximal. Suppose that U does not span V. Then we can take $u^\tau \in W\backslash(U \cup U^-)$. We claim that $U \cup R(u^\tau)$ is a larger consistent ideal. If $u^\tau \in U_0$, then the claim holds since $U \cap U_0^- = \emptyset$ and $R(u^\tau) \subseteq U_0$ (as there is no arc from U_0 to $W\backslash U_0$). Otherwise, $U \cup R(u^\tau)$ becomes consistent because there is no arc from $W\backslash(U \cup U^-)$ to U^- since otherwise U cannot be ideal due to the skew-symmetry of H. □

For each $i \in I$, let $f_i: 2^{W_i} \to \mathbb{R}$ be the minor $f^{D_i}_{D_i\backslash W_i}$ obtained from f by the restriction to D_i and the contraction by $D_i\backslash W_i$. We define δ_1 by

$$\delta_1 = \max_{i\in I} \left\{ f_i(W_i) - \min_{X\subseteq W_i} \{f_i(X)\} \right\}. \tag{17}$$

It should be noted that we always have $\delta_1 \geq 0$ and that if $\delta_1 = 0$, then W_i is a minimizer of f_i for all $i \in I$.

It should also be noted here that f_i is a submodular (set) function on 2^{W_i} with $f_i(\emptyset, \emptyset) = 0$. Thus we can employ a submodular function minimization algorithm to compute a minimizer of each f_i and hence δ_1 can be computed in time proportional to that required for a single submodular function minimization with an underlying set of size $|V| = n$.

Let $\mathrm{B}(f_i)$ be the base polyhedron associated with f_i. That is,

$$\mathrm{B}(f_i) := \left\{ x \in \mathbb{R}^{W_i} \mid \forall X \subseteq W_i\colon \sum_{v^\tau\in X} \tau x(v) \leq f_i(X), \sum_{v^\tau\in W_i} \tau x(v) = f_i(W_i) \right\}.$$

Applying an existing algorithm for the ordinary submodular function minimization (e.g., [23]), we have the following.

Lemma 13 *For each $i \in I$, there exists $x_i \in \mathrm{B}(f_i)$ such that x_i is a maximizer of*

$$\max \left\{ \sum_{v^\tau\in W_i\colon \tau x(v)<0} \tau x(v) \;\middle|\; x \in \mathrm{B}(f_i) \right\}$$

and

$$\tau x_i(v) \begin{cases} \leq 0 & \text{for all } v^\tau \in M_i \\ \geq 0 & \text{for all } v^\tau \in W_i\backslash M_i \end{cases}$$

where M_i is any minimizer of f_i. Moreover, a submodular function minimization algorithm can compute such $x_i \in \mathrm{B}(f_i)$, together with an expression $x_i = \sum_{i\in J_i} \lambda_j y_j$, a convex combination of extreme bases $y_j \in \mathrm{B}_{W_i}(f_i)$ ($j \in J_i$), each corresponding to a linear ordering $L_j|\sigma_j$ of W_i, where $|J_i| \leq |W_i|$.

4.3 Algorithm description

We now give an algorithm description. In order to understand the whole picture of the algorithm, we also state key lemmas, whose proofs will be given in the next subsections.

The algorithm first computes δ_1 defined in the last subsection, and decides the next procedure depending on whether $\delta_1 = 0$ or $\delta_1 > 0$. If $\delta_1 = 0$, we have the following.

Lemma 14 *Suppose $\delta_1 = 0$. Then, any consistent ideal of H that spans V is a minimizer of f.*

Hence, in this case, we can output a minimizer of f by computing a maximal consistent ideal of H by Lemma 12. On the other hand, if $\delta_1 > 0$, then we further split the case into two subcases as follows.

Let $i^* \in I$ be a maximizer of (17), let $f^* = f_{D_{i^*}}$ be the contraction of f by D_{i^*}, and let $V^* \subseteq V$ be the ground set of f^*. For $\delta > 0$ we call $(X, Y) \in 3^{V^*}$ *δ-highly negative* for f^* if $f^*(X, Y) \leq -\delta$. The following lemma is adapted from [21, Lemma 3.8].

Lemma 15 *Let $i^* \in I$ be a maximizer of (17). Suppose that $\delta_1 > 0$ and that there is no δ_1-highly negative element for f^*. Then there exists no minimizer (X, Y) of f such that $(X, Y) \supseteq W_{i^*}$.*

On the other hand, if there is a δ_1-highly negative element, we have the following.

Lemma 16 *Suppose that $\delta_1 > 0$ and that there exists a δ_1-highly negative element for f^*. Let x be the output of* REFINE *for $f^* = f_{D_{i^*}}$ with $\delta = \delta_1$ and $\zeta = \delta_1/(12\beta n^3)$. Then there exist $u \in V^*$ and $\tau \in \{-, +\}$ such that*

$$\tau\alpha^\tau(u)x(u) \leq -\frac{\delta_1}{n}. \tag{18}$$

Moreover, if u^τ satisfies (18), then u^τ is contained in every minimizer of f^*.

Hence, from Lemmas 15 and 16, after applying REFINE for $f^* = f_{D_{i^*}}$ with $\delta = \delta_1$ and $\zeta = \delta_1/(6\beta n^3)$ we can determine one of the following two:

(I) There exists no minimizer of f that contains elements of W_{i^*}.
(II) There exists some $j \in I \backslash \{i^*\}$ such that every minimizer of f containing W_{i^*} contains W_j.

Now we are ready to describe our algorithm strongly-ABSFM (f).

For Line 9 of strongly-ABSFM (f) we have the following.

Lemma 17 *There is an algorithm that computes $y^* \in \mathrm{P}^\alpha(f^*)$ with $\|y^*\|_\alpha + f^*(S, T) \leq 2n\delta_1$ for some $(S, T) \in 3^{V^*}$, along with the expression of y^* as a convex combination of extreme points of $\mathrm{P}^\alpha(f^*)$, in $\mathrm{O}(n^2 + \mathbf{SFM}(n))$ time, where $\mathbf{SFM}(n)$ denotes the complexity of ordinary submodular function minimization with the underlying set of size n.*

Assuming the correctness of above lemmas, we now have the following theorem.

Algorithm 7 strongly-ABSFM (f)

Initialize $H = (W, C)$ to be the graph on $W = V^+ \cup V^-$ with no arc, and $U_\mathrm{e} = \emptyset$.
while $U_\mathrm{e} \neq V$ **do**
  Compute δ_1
  if $\delta_1 = 0$ **then**
    Compute any maximal consistent ideal of H and return the corresponding signed set of 3^V.
  else
    $i^* :=$ a maximizer of (17) for δ_1.
    Let $f^* = f_{D_{i^*}}$ and let $V^* \subseteq V$ be the ground set of f^*.
    Compute $y^* \in \mathrm{P}^\alpha(f^*)$ with $\|y^*\|_\alpha + f^*(S, T) \leq 2n\delta_1$ for some $(S, T) \in 3^{V^*}$.
    REFINE$(f^*, y^*, \delta = \delta_1, \zeta = \delta_1/(12\beta n^3))$.
    Let (S, T) and x be the output of REFINE.
    if there are $u \in V^*$ and $\tau \in \{-, +\}$ with $\tau\alpha^\tau(u)x(u) \leq -\frac{\delta_1}{n}$ **then**
      Add arcs from W_{i^*} to u^τ in H.
    else
      Compute the set F of vertices from which we can reach W_{i^*} in H.
      For each $u^\tau \in F$, add $(u^\tau, u^{-\tau})$ to H.
    Update U_e and H so that it satisfies (16).

Theorem 4 *Let* $f : 3^V \to \mathbb{R}$ *with* $f(\emptyset, \emptyset) = 0$. strongly-ABSFM (f) *returns a minimizer of* f *in* $\mathrm{O}(n^2(n^5\mathbf{EO}\beta^2 \log \beta n + \mathbf{SFM}(n)))$ *time, where* $\mathbf{EO}$ *denotes the oracle time for the function evaluation of* f.

Proof The correctness follows from the above arguments.

Let us check the time complexity. The number of while-loop iterations is $\mathrm{O}(n^2)$ since in each iteration the algorithm adds a new arc or delete at least one node.

In each iteration of the while-loop the running time of $\mathbf{SFM}(n)$ is required for computing δ_1 and y^* with additional $\mathrm{O}(n^2)$ time, while each REFINE $(f^*, y^*, \delta_1, \delta_1/(12\beta n^3))$ requires $\mathrm{O}(n^5\mathbf{EO}\beta^2 \log \beta n)$ time. □

Thus the remaining two subsections are devoted to giving the missing proofs.

4.4 Concatenating linear orderings and proof of Lemma 17

Before going to the proofs of the lemmas, we give a technique for concatenating the linear orderings on strongly connected components given in Lemma 13 to be linear orderings of the whole set.

Choose any maximal chain

$$\mathscr{C} : (\emptyset, \emptyset) = (S_0, T_0) \subsetneq \cdots \subsetneq (S_k, T_k). \tag{19}$$

of consistent ideals of H. By Lemma 12, (S_k, T_k) spans V. Here, note that for each $\ell = 1, \ldots, k$ there uniquely exists $i_\ell \in I$ such that $(S_\ell, T_\ell) \setminus (S_{\ell-1}, T_{\ell-1}) = W_{i_\ell}$.

By Lemma 13 we have extreme bases $y_j \in \mathrm{B}(f_i)$ corresponding to linear orderings $L_j|\sigma_j$ $(j \in J_i)$ of W_i and positive numbers λ_j $(j \in J_i)$ with $\sum_{j \in J_i} \lambda_j = 1$. Those linear orderings can be concatenated to be linear orderings $L'_q|\sigma'_q$ $(q \in Q)$ of (S_k, T_k) with the index set Q such that

$$\sum_{q \in Q} \mu_q \hat{y}_q \in \mathrm{P}^{\alpha}(f),$$

where $\hat{y}_q$ is the extreme base of $\mathrm{P}^{\alpha}(f)$ generated by $L'_q|\sigma'_q$ and μ_q is a positive scaler for each $q \in Q$ satisfying

$$\lambda_j = \sum_{q: L'_q|\sigma'_q \text{ coincides with } L_j|\sigma_j \text{ on } W_{i_\ell}} \mu_q \qquad (1 \leq \forall \ell \leq k, \forall j \in J_{i_\ell}) \tag{20}$$

The following procedure gives an explicit construction of such $L'_q|\sigma'_q$ and μ_q.

(P) Let $J_* = \cup_{\ell=1}^{k} J_{i_\ell}$, where we assume J_{i_ℓ}s are disjoint. Let $Q = \emptyset$.
Repeat the following until $J_* = \emptyset$.
1. Find $j_* \in J_*$ such that $\lambda_{j_*} = \min\{\lambda_j \mid j \in J_*\}$. Suppose $j_* \in J_{i_*}$.
2. Put $\mu_{j_*} = \lambda_{j_*}$ and $Q \leftarrow Q \cup \{j_*\}$.
3. For each $\ell \in \{1, \ldots, k\} \backslash \{i_*\}$ choose one $j_\ell \in J_{i_\ell}$. Also, put $j_{i_*} = j_*$ for $\ell = i^*$.
4. For each $\ell \in \{1, \ldots, k\}$ do:
   $\lambda_{j_\ell} \leftarrow \lambda_{j_\ell} - \lambda_{j_*}$
   if $\lambda_{j_\ell} = 0$ then $J_\ell \leftarrow J_\ell \backslash \{j_\ell\}$ and $J_* \leftarrow J_* \backslash \{j_\ell\}$.
5. Let $L'_{j_*}|\sigma'_{j_*}$ be a signed linear ordering of V such that $L_{j_\ell}|\sigma_{j_\ell}$ $(\ell = 1, \ldots, k)$ appear in $L'_{j_*}|\sigma'_{j_*}$, each as an interval, in the order of ℓ.

Note that, since $\sum_{j \in J_{i_\ell}} \lambda_j = 1$ for each ℓ and the procedure decreases $\sum_{j \in J_{i_\ell}} \lambda_j$ by the same amount for all ℓ at Line 4, J_{i_ℓ} becomes empty for some ℓ if and only if J_{i_ℓ} becomes empty for all ℓ. In other words, $J_{i_\ell} \neq \emptyset$ for all ℓ at Line 3, and the procedure works in $\mathrm{O}(n^2)$ time.

Suppose that we are given (L'_q, σ'_q) $(q \in Q)$ and μ_q $(q \in Q)$ by procedure **(P)**. For each $q \in Q$ let y_q be the base of $\mathrm{B}(f^{(S_k, T_k)})$ determined by (L'_q, σ'_q), and define $y \in \mathbb{R}^V$ by

$$y = \sum_{q \in Q} \mu_q y_q. \tag{21}$$

Lemma 18 *Let x_i $(i \in I)$ be given as in* Lemma 13 *and let y be defined by* (21). *Then for all $\ell = 1, \ldots, k$ and $v^\tau \in W_{i_\ell}$ we have*

$$\tau y(v) \leq \tau x_{i_\ell}(v).$$

Proof Consider $f^{(S_\ell, T_\ell)}_{(S_{\ell-1}, T_{\ell-1})}$. This is submodular on W_{i_ℓ}. Moreover, since $(D_{i_\ell} \backslash W_{i_\ell}) \subseteq (S_{\ell-1}, T_{\ell-1})$, we have $f^{(S_\ell, T_\ell)}_{(S_{\ell-1}, T_{\ell-1})} \leq f_{i_\ell}$ by the submodularity of $f^{(S_\ell, T_\ell)}$. Therefore, for each linear ordering $L_j|\sigma_j$ of W_{i_ℓ}, we have $\tau y_j(v) \leq \tau x_j(v)$, where y_j and x_j are bases of $\mathrm{B}(f^{(S_\ell, T_\ell)}_{(S_{\ell-1}, T_{\ell-1})})$ and $\mathrm{B}(f_{i_\ell})$ generated by $L_j|\sigma_j$, respectively. Therefore, by (20), we have

$$\begin{aligned}\tau x_{i_\ell}(v) &= \sum_{j \in J_{i_\ell}} \lambda_j \tau x_j(v) \geq \sum_{j \in J_{i_\ell}} \sum_{q: L'_q|\sigma'_q \text{ coincides with } L_j|\sigma_j \text{ on } W_{i_\ell}} \mu_q \tau y_j(v) \\ &= \sum_{q \in Q} \mu_q \tau y_q(v) = \tau y(v)\end{aligned}$$

for all $\ell = 1, \ldots, k$ and $v^\tau \in W_{i_\ell}$. □

From y, let us further define $\hat{y} \in \mathbb{R}^V$ by

$$\hat{y}(v) = \frac{1}{\alpha^\tau(v)} y(v) \qquad (\forall v^\tau \in (S_k, T_k)). \tag{22}$$

Lemma 19 *Let $\hat{y}$ be defined by (22). Then $\hat{y} \in \mathrm{B}^\alpha_{(S_k,T_k)}(f)$.*

Proof Since $\mathrm{B}^\alpha_{(S_k,T_k)}(f)$ is obtained from $\mathrm{B}(f^{(S_k,T_k)})$ by appropriate scaling, the statement follows from Lemma 18. □

For proving Lemma 17, we need one more technical lemma.

Lemma 20 *Let $\hat{y}$ be defined by (22). Then for each $\ell = 1, \ldots, k$ we have*

$$\sum_{v^\tau \in W_{i_\ell}: \tau \hat{y}(v) > 0} \tau \alpha^\tau(v) \hat{y}(v) \leq \delta_1.$$

Proof Denote x_{i_ℓ} by x for simplicity in the present proof. Since $x \in \mathrm{B}(f_{i_\ell})$, we have

$$\sum_{v^\tau \in W_{i_\ell}} \tau x(v) = f_{i_\ell}(W_{i_\ell}). \tag{23}$$

On the other hand, due to the min–max relation for the submodular function minimization,

$$\sum_{v^\tau \in W_{i_\ell}: \tau x(v) < 0} \tau x(v) = f_{i_\ell}(M_{i_\ell}), \tag{24}$$

where M_{i_ℓ} is a minimizer of f_{i_ℓ}. It follows from (23) and (24) that

$$\sum_{v^\tau \in W_{i_\ell}: \tau x(v) > 0} \tau x(v) = f_{i_\ell}(W_{i_\ell}) - f_{i_\ell}(M_{i_\ell}) \leq \delta_1.$$

Also, by Lemma 18, $\tau \hat{y}(v) > 0$ holds only if $\tau x(v) > 0$ for each $v^\tau \in W_{i_\ell}$. Therefore we get

$$\sum_{v^\tau \in W_{i_\ell}: \tau \hat{y}(v) > 0} \tau \alpha^\tau \hat{y}(v) \leq \sum_{v^\tau \in W_{i_\ell}: \tau x(v) > 0} \tau x(v) \leq \delta_1.$$

□

Now we are ready to prove Lemma 17.

Proof of Lemma 17 We can assume that the maximal chain $\mathscr{C}$ in (19) contains $(S_\ell, T_\ell) = D_{i^*}$ for some $\ell \in \{1, \ldots, k\}$. Let $\hat{y}$ be defined by (22). Then we have $f(D_{i^*}) = \langle \hat{y}, \chi^\alpha_{D_{i^*}} \rangle$. Hence, putting $(S'_k, T'_k) := (S_k, T_k) \backslash D_{i^*}$ and letting y^* be the restriction of $\hat{y}$ on (S'_k, T'_k), by Lemma 19 we have $y^* \in \mathrm{B}^\alpha_{(S'_k, T'_k)}(f^*)$ and

$$\sum_{v^\tau \in (S'_k, T'_k)} \tau\alpha^\tau \hat{y}(v) = f^*(S'_k, T'_k).$$

Therefore,

$$\begin{aligned}
f^*(S'_k, T'_k) + \|y^*\|_\alpha &= f^*(S'_k, T'_k) - \sum_{v \in V^* : \hat{y}(v) < 0} \alpha^+(v)\hat{y}(v) + \sum_{v \in V^* : \hat{y}(v) > 0} \alpha^-(v)\hat{y}(v) \\
&= f^*(S'_k, T'_k) - \sum_{v^\tau \in (S'_k, T'_k) : \tau\hat{y}(v) < 0} \tau\alpha^\tau \hat{y}(v) \\
&\quad + \sum_{v^\tau \in (S'_k, T'_k) : \tau\hat{y}(v) > 0} \tau\alpha^\tau \hat{y}(v) \\
&= 2\left(\sum_{v^\tau \in (S'_k, T'_k) : \tau\hat{y}(v) > 0} \tau\alpha^\tau \hat{y}(v) \right) \\
&\le 2n\delta_1,
\end{aligned}$$

where the last inequality follows from Lemma 20 and recall that $V^* = S'_k \cup T'_k$. □

4.5 Proofs of Lemmas 14, 15, and 16

Proof of Lemma 14 Since $\delta_1 = 0$, we see that for all $i \in I$ W_i is a minimizer of f_i. Hence we have a base $x_i \in \mathrm{B}(f_i)$ such that $\tau x_i(u) \le 0$ $(\forall u^\tau \in W_i)$ by Lemma 13.

Now, let (A, B) be an arbitrary consistent ideal of H that spans V. Then there is a maximal chain $\mathscr{C}$ of consistent ideals of H whose last element is (A, B), and let $y \in \mathrm{B}(f^{(A,B)})$ be the vector constructed in (21) in Sect. 4.4 with respect to chain $\mathscr{C}$. Then by Lemma 18 we have

$$\tau y(v) \le \tau x_i(v) \le 0 \qquad \forall v^\tau \in A^+ \cup B^-.$$

This means that $\{v \in V : y(v) < 0\} \subseteq A$ and $\{v \in V : y(v) > 0\} \subseteq B$. Therefore, setting $\hat{y}$ as in (22), we get

$$\begin{aligned}
\|\hat{y}\|_\alpha &= \sum_{u \in V : y(u) < 0} \alpha^+(u)\hat{y}(u) - \sum_{u \in V : y(u) > 0} \alpha^-(u)\hat{y}(u) \\
&= \sum_{u \in A} \alpha^+(u)\hat{y}(u) - \sum_{u \in B} \alpha^-(u)\hat{y}(u) = \langle \hat{y}, \chi^\alpha_{(A,B)} \rangle = f(A, B).
\end{aligned}$$

where the last equation follows from Lemma 19. It follows from Theorem 1 that (A, B) is a minimizer of f. ◻

Proof of Lemma 15 Let M_{i^*} be a minimizer of f_{i^*}, and let $E_{i^*} = M_{i^*} \cup (D_{i^*} \setminus W_{i^*})$. Since $\delta_1 > 0$, we have $E_{i^*} \neq D_{i^*}$. Also, by the assumption we have

$$f^*(X, Y) > -\delta_1 \qquad (\forall (X, Y) \in 3^{V^*}),$$

which is rewritten as

$$f((X, Y) \cup D_{i^*}) - f(D_{i^*}) > f(E_{i^*}) - f(D_{i^*}) \qquad (\forall (X, Y) \in V^*).$$

Hence $f((X, Y) \cup D_{i^*}) > f(E_{i^*})$ for all $(X, Y) \in 3^{V^*}$. This implies that any $(X, Y) \in 3^V$ with $(X, Y) \supseteq D_{i^*} \supseteq W_{i^*}$ is not a minimizer of f. ◻

In order to prove Lemma 16, we need one more extra lemma.

Lemma 21 *For a given $x \in \mathrm{P}^\alpha(f)$ suppose that we have $\|x\|_\alpha < -f(X, Y) + \gamma$ for some $(X, Y) \in 3^V$ and $\gamma > 0$. If $\tau\alpha^\tau(u)x(u) \leq -\gamma$ holds for some $u \in V$ and $\tau \in \{+, -\}$, then every minimizer of f contains u^τ.*

Proof Let (S, T) be any minimizer of f, and suppose that u^τ is not contained in (S, T). Let $Z = \{v^\sigma \in V^+ \cup V^- : \sigma x(v) < 0\}$. Then $u^\tau \in Z \setminus (S^+ \cup T^-)$. By using the assumption of the present lemma, we have

$$\begin{aligned}
\gamma &\geq \|x\|_\alpha + f(X, Y) \geq \|x\|_\alpha + f(S, T) \geq \|x\|_\alpha + \langle x, \chi^\alpha_{(S,T)} \rangle \\
&= -\sum_{v^\sigma \in Z} \sigma\alpha^{\sigma(v)} x(v) + \sum_{v^\sigma \in S^+ \cup T^-} \sigma\alpha^{\sigma(v)} x(v) \\
&= \sum_{v^\sigma \in Z \setminus (S^+ \cup T^-)} -\sigma\alpha^{\sigma(v)} x(v) + \sum_{v^\sigma \in (S^+ \cup T^-) \setminus Z} \sigma\alpha^{\sigma(v)} x(v).
\end{aligned}$$

By the definition of Z, each term in the summations is nonnegative. Therefore, by $u^\tau \in Z \setminus (S^+ \cup T^-)$, we get $\gamma \geq -\tau\alpha^\tau x(u)$, contradicting the assumption of the present lemma. ◻

Proof of Lemma 16 By Theorem 2, REFINE$(f^*, \hat{y}, \delta_1, \frac{\delta_1}{12\beta n^3})$ outputs $(A, B) \in 3^{V^*}$ and $x \in \mathbb{R}^{V^*}$ with $\|x\|_\alpha \leq 6\beta n^2 (\frac{\delta_1}{12\beta n^3}) - f^*(A, B) < \frac{\delta_1}{n} - f^*(A, B)$.

Let $(X, Y) \in 3^{V^*}$ be a δ_1-highly negative element for f^*. Since $x \in \mathrm{P}^\alpha(f^*)$, we have

$$\sum_{u^\tau \in X^+ \cup Y^-} \tau\alpha^\tau(u) x(u) = \left\langle x, \chi^\alpha_{(X,Y)} \right\rangle \leq f^*(X, Y) \leq -\delta_1.$$

Hence there is $u^\tau \in X^+ \cup Y^-$ such that $\tau\alpha^\tau(u)x(u) \leq \frac{\delta_1}{n}$.

Now, since $\|x\|_\alpha < \frac{\delta_1}{n} - f^*(A, B)$, Lemma 21 implies that u^τ is contained in every minimizer of f^*. ◻

This completes the proofs of all the lemmas stated in Sect. 4.3.

Acknowledgements We thank an anonymous referee for careful reading of an earlier version of this paper and for providing helpful comments. The present research was supported by JSPS KAKENHI Grant No. JP25280004. The second author was supported by JSPS Postdoctoral Fellowships for Research Abroad. The authors would like to acknowledge the support by the Hausdorff Trimester Program on Combinatorial Optimization of the Hausdorff Research Institute, the University of Bonn, where this work was partially carried out.

References

1. Ando, K., Fujishige, S.: On structures of bisubmodular polyhedra. Math. Progr. **74**, 293–317 (1996)
2. Ando, K., Fujishige, S., Naitoh, T.: A characterization of bisubmodular functions. Discrete Math. **148**, 299–303 (1996)
3. Bouchet, A.: Greedy algorithm and symmetric matroids. Math. Progr. **38**, 147–159 (1987)
4. Bouchet, A., Cunningham, W.H.: Delta-matroids, jump systems, and bisubmodular polyhedra. SIAM J. Discrete Math. **8**, 17–32 (1995)
5. Chandrasekaran, R., Kabadi, S.N.: Pseudomatroids. Discrete Math. **71**, 205–217 (1988)
6. Dress, A., Havel, T.: Some combinatorial properties of discriminants in metric vector spaces. Adv. Math. **62**, 285–312 (1986)
7. Dunstan, F.D.J., Welsh, D.J.A.: A greedy algorithm for solving a certain class of linear programmes. Math. Progr. **5**, 338–353 (1973)
8. Fujishige, S.: Submodular Functions and Optimization, 2nd edn. Elsevier, Amsterdam (2015)
9. Fujishige, S., Iwata, S.: Bisubmodular function minimization. SIAM J. Discrete Math. **19**, 1065–1073 (2005)
10. Fujishige, S., Király, T., Makino, K., Takazawa, K., Tanigawa, S.: Minimizing submodular functions on diamonds via generalized fractional matroid matchings. EGRES Technical Reports, No. 2014-14 (2014)
11. Fujishige, S., Tanigawa, S.: A min–max theorem for transversal submodular functions and its implications. SIAM J. Discrete Math. **28**, 1855–1875 (2013)
12. Fujishige, S., Tanigawa, S., Yoshida, Y.: Generalized skew bisubmodularity: a characterization and a min–max theorem. Discrete Optim. **12**, 1–9 (2014)
13. Hirai, H.: Discrete convexity and polynomial solvability in minimum 0-extension problems. Math. Progr. Ser. A **155**, 1–55 (2016)
14. Huber, A., Kolmogorov, V.: Towards minimizing k-submodular functions. In: Proceedings of the 2nd International Symposium on Combinatorial Optimization, pp. 451–462 (2012)
15. Huber, A., Krokhin, A.: Oracle tractability of skew bisubmodular functions. SIAM J. Discrete Math. **28**, 1828–1837 (2014)
16. Huber, A., Krokhin, A., Powell, R.: Skew bisubmodularity and valued CSPs. SIAM J. Comput. **43**, 1064–1084 (2014)
17. Iwata, A., Fleischer, L., Fujishige, S.: A combinatorial strongly polynomial algorithm for minimizing submodular functions. J. ACM **48**, 761–777 (2001)
18. Kolmogorov, V.: Submodularity on a tree: unifying $L^\natural$-convex and bisubmodular functions. In: Proceedings of the 36th International Symposium on Mathematical Foundations of Computer Science. LNCS 6907, pp. 400–411 (2011)
19. Krokhin, A., Larose, B.: Maximizing supermodular functions on product lattices, with application to maximum constraint satisfaction. SIAM J. Discrete Math. **22**, 312–328 (2008)
20. Kuivinen, F.: On the complexity of submodular function minimisation on diamonds. Discrete Optim. **8**, 459–477 (2011)
21. McCormick, S.T., Fujishige, S.: Strongly polynomial and fully combinatorial algorithms for bisubmodular function minimization. Math. Progr. Ser. A **122**, 87–120 (2010)
22. Qi, L.: Directed submodularity, ditroids and directed submodular flows. Math. Progr. **42**, 579–599 (1988)
23. Schrijver, A.: A combinatorial algorithm minimizing submodular functions in strongly polynomial time. J. Combin. Theory Ser. B. **80**, 346–355 (2000)
24. Živný, S.: The Complexity of Valued Constraint Satisfaction Problems. Springer, Berlin (2012)