MATHEMATICAL ENGINEERING TECHNICAL REPORTS

On Some Properties of Polyhedral L-concave Maximization Algorithm

Kazuo MUROTA and Akiyoshi SHIOURA

METR 2013–18 September 2013

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.keisu.t.u-tokyo.ac.jp/research/techrep/index.html

On Some Properties of Polyhedral L-concave Maximization Algorithm*

Kazuo MUROTA

Graduate School of Information Science and Technology
University of Tokyo, Tokyo 113-8656, Japan

and

Akiyoshi SHIOURA

Graduate School of Information Sciences
Tohoku University, Sendai 980-8579, Japan

September 2013

Abstract

Hassin (1983) proposed a dual algorithm for the minimum cost flow problem, which iteratively updates dual variables only in a steepest ascent manner. This algorithm is generalized to the minimum cost submodular flow problem by Chung and Tcha (1991). It is known in discrete convex analysis that the dual of the minimum cost flow problem is formulated as the maximization of a polyhedral L-concave function. It is recently pointed out that Hassin's algorithm can be recognized as a steepest ascent algorithm for polyhedral L-concave functions. The objective of this paper is to show some nice properties of the steepest ascent algorithm for polyhedral L-concave functions. We show that the algorithm is endowed with the monotonicity property of Hassin's algorithm. Moreover, the algorithm finds the "nearest" optimal solution to a given initial solution, and the trajectory of the solutions generated by the algorithm is a "shortest" path from the initial solution to the "nearest" optimal solution.

*This research is partially supported by KAKENHI (21360045, 21740060, 24500002) and the Aihara Project, the FIRST program from JSPS.

1 Introduction

Among many algorithms for the minimum cost flow problem (see, e.g., [1, 11]), Hassin's dual algorithm [5] is unique in that it maintains only dual variables, while most of the other algorithms use primal (i.e., flow) variables. Hassin's algorithm iteratively chooses a subset of dual variables which corresponds to a graph cut and increments them so that the dual objective function increases strictly. It is shown in [5] that the sequence of solutions generated by the algorithm has a certain monotonicity property, from which follows the finite termination of the algorithm. Hassin's algorithm is later generalized to the minimum cost submodular flow problem by Chung and Tcha [2].

It is known in discrete convex analysis ([6], [7]) that the dual of the minimum cost (submodular) flow problem is formulated as the maximization of a polyhedral L-concave function. The concept of polyhedral L-concave functions in real variables was introduced by Murota and Shioura [9] as a variant of L-concave functions originally defined for functions on integer lattice points. It is pointed out in our recent paper [10] that Hassin's algorithm as well as Chung–Tcha's algorithm can be recognized as a steepest ascent algorithm for polyhedral L-concave functions, where the steepest ascent direction is chosen from a finite set of 0-1 vectors. This observation indicates that the steepest ascent algorithm for polyhedral L-concave functions is fundamental in combinatorial optimization.

In this paper, we investigate the behavior of the steepest ascent algorithm for polyhedral L-concave function maximization and show its nice properties. First, it is endowed with the same monotonicity property as that of Hassin's algorithm, which guarantees its finite termination. Second, for any initial solution, the algorithm finds the "nearest" optimal solution in the ℓ_∞-distance from the initial solution. Third, the trajectory of the solutions generated by the algorithm is a "shortest" path from the initial solution to the "nearest" optimal solution in the sense that the total sum of the step lengths is equal to the ℓ_∞-distance from the initial solution to the nearest optimal solution. Our second and third results imply, in particular, that Hassin's and Chung–Tcha's algorithms are "efficient" algorithms in some sense. The steepest ascent algorithm for polyhedral L-concave functions can naturally be adapted to polyhedral $\mathrm{L}^\natural$-concave functions. The algorithm outputs the optimal solution that is "nearest" with respect to a variant of the ℓ_∞-distance.

2 Preliminaries on L-concave Functions

We review the concept of polyhedral L-concave functions. Throughout this paper, let V be a finite set. For a function $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$, its *effective* *domain* is defined as

$$\mathrm{dom}\, g = \{p \in \mathbb{R}^V \mid g(p) > -\infty\}.$$

A function $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$ is said to be a *polyhedral concave function* if the set

$$\{(p, \alpha) \in \mathbb{R}^V \times \mathbb{R} \mid p \in \mathrm{dom}\, g,\ \alpha \leq g(p)\}$$

is a (nonempty) polyhedron. Equivalently, $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$ is a polyhedral concave function if there exist a nonempty polyhedron $S \subseteq \mathbb{R}^V$, $a_1, \ldots, a_m \in \mathbb{R}^V$, and $b_1, \ldots, b_m \in \mathbb{R}$ such that

$$\mathrm{dom}\, g = S, \qquad g(p) = \min_{1 \leq i \leq m} \{a_i^{\mathrm{T}} p + b_i\} \qquad (p \in S).$$

For $p \in \mathrm{dom}\, g$ and $q \in \mathbb{R}^V$, we define

$$g'(p; q) = \lim_{\lambda \downarrow 0} \frac{g(p + \lambda q) - g(p)}{\lambda},$$

which is called the *directional derivative* of g at p in direction q. We also define

$$\bar{c}(p; q) = \max\{\lambda \in \mathbb{R}_+ \mid g(p + \lambda q) - g(p) = \lambda\, g'(p; q)\}. \qquad (2.1)$$

Note that $\bar{c}(p; q) > 0$ and $g(p + \lambda q) - g(p) = \lambda\, g'(p; q)$ holds for every λ with $0 \leq \lambda \leq \bar{c}(p; q)$.

A polyhedral concave function $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$ is said to be *polyhedral L-concave* [9] if it satisfies the following conditions:

(LF1) $g(p) + g(q) \leq g(p \wedge q) + g(p \vee q) \quad (\forall p, q \in \mathrm{dom}\, g)$,
(LF2) $\exists r \in \mathbb{R}$ s.t. $g(p + \lambda \mathbf{1}) = g(p) + \lambda r \ (\forall p \in \mathrm{dom}\, g,\ \forall \lambda \in \mathbb{R})$,

where $p \wedge q, p \vee q\ (\in \mathbb{R}^V)$ denote the vectors with

$$(p \wedge q)(v) = \min\{p(v), q(v)\}, \quad (p \vee q)(v) = \max\{p(v), q(v)\} \quad (v \in V),$$

and $\mathbf{1}\ (\in \mathbb{R}^V)$ is the vector with each component being equal to one. Note that $r = 0$ is assumed in (LF2) whenever we consider maximization of a polyhedral L-concave function since otherwise there exists no maximizer.

A typical example of polyhedral L-concave functions arises from the maximum weight tension problem. For a directed graph $G = (V, E)$, let $\varphi_{uv} : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ be an edge weight function for $(u, v) \in E$. We assume that functions φ_{uv} ($(u, v) \in E$) are polyhedral (or piecewise-linear) concave functions. The maximum weight tension problem is formulated as follows:

$$\textbf{(MWT)} \quad \left| \begin{array}{ll} \text{Maximize} & \displaystyle\sum_{(u,v) \in E} \varphi_{uv}(p(u) - p(v)) \\ \text{subject to} & p \in \mathbb{R}^V. \end{array} \right.$$

We denote by $g_{\rm T} : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$ the objective function of the problem (MWT). Note that $\mathrm{dom}\, g_{\rm T} \neq \emptyset$ if and only if (MWT) has a feasible solution, i.e., there exists some $p \in \mathbb{R}^V$ such that $\varphi_{uv}(p(u) - p(v)) > -\infty$ for all $(u, v) \in E$.

Proposition 2.1 ([9, Example 2.4]). *Suppose that* $\mathrm{dom}\, g_{\rm T} \neq \emptyset$. *Then, function* $g_{\rm T}$ *is polyhedral L-concave with* $r = 0$ *in* (LF2).

Another example of polyhedral L-concave function comes from the so-called Lovász extension of a submodular function. A set function $\rho : 2^V \to \mathbb{R}$ is said to be *submodular* if it satisfies

$$\rho(X) + \rho(Y) \geq \rho(X \cap Y) + \rho(X \cup Y) \qquad (\forall X, Y \subseteq V).$$

Given a set function $\rho : 2^V \to \mathbb{R}$ with $\rho(\emptyset) = 0$, define a function $\hat{\rho} : \mathbb{R}^V \to \mathbb{R}$ by

$$\hat{\rho}(p) = \sum_{i=1}^{h-1} (\tilde{p}_i - \tilde{p}_{i+1})\rho(L_i) + \tilde{p}_h \rho(L_h) \qquad (p \in \mathbb{R}^V), \tag{2.2}$$

where $\tilde{p}_1 > \tilde{p}_2 > \cdots > \tilde{p}_h$ are distinct values of components of p and

$$L_i \;=\; \{v \in V \mid p(v) \geq \tilde{p}_i\} \qquad (i = 1, 2, \ldots, h).$$

The function $\hat{\rho}$ is called the *Lovász extension* of ρ.

Proposition 2.2 ([9, Theorem 4.36]). *For a submodular set function* $\rho : 2^V \to \mathbb{R}$ *with* $\rho(\emptyset) = 0$, *the function* $-\hat{\rho}$ *is a polyhedral L-concave function. If* $\rho(V) = 0$, *then* $-\hat{\rho}$ *satisfies the property* (LF2) *with* $r = 0$.

3 Hassin's and Chung–Tcha's Algorithms

As the motivation of the present paper, we review the dual algorithms for the minimum cost flow problem by Hassin [5] and for the minimum cost submodular flow problem by Chung and Tcha [2]. We also observe the polyhedral L-concavity of the dual objective functions of the problems in [5] and in [2]. In the following, we denote by $\chi_X \in \{0, 1\}^V$ the characteristic vector of $X \subseteq V$, i.e., $\chi_X(v) = 1$ if $v \in X$ and $\chi_X(v) = 0$ if $v \in V \setminus X$.

3.1 Hassin's Algorithm

For a directed graph $G = (V, E)$, nonnegative edge capacity $c(e)$, and edge cost $\gamma(e)$ for $e \in E$, the minimum cost flow problem is formulated as follows:

$$\begin{array}{lll} \text{Minimize} & \displaystyle\sum_{(u,v) \in E} \gamma(u, v) x(u, v) & \\ \text{subject to} & \partial x(u) = 0 \quad (u \in V), & \\ & 0 \leq x(u, v) \leq c(u, v) \quad ((u, v) \in E), & \end{array}$$

where

$$\partial x(u) = \sum_{v:(u,v)\in E} x(u,v) - \sum_{v:(v,u)\in E} x(v,u) \qquad (u \in V).$$

The dual problem is given as

$$\begin{array}{ll} \text{Maximize} & g_{\rm H}(p) \equiv \displaystyle\sum_{(u,v)\in E} c(u,v)\min\{0, p(u)-p(v)+\gamma(u,v)\} \\ \text{subject to} & p(v) \in \mathbb{R} \quad (v \in V). \end{array}$$

The function $g_{\rm H}$ is polyhedral L-concave since it is a special case of the function $g_{\rm T}$ in Proposition 2.1.

Hassin's algorithm is described as follows. For $p \in \mathbb{R}^V$ and $X \subseteq V$, we define

$$I(p,X) \;=\; \sum_{(u,v)\in E^{<}_{\rm out}(p,X)} c(u,v) - \sum_{(u,v)\in E^{\leq}_{\rm in}(p,X)} c(u,v), \qquad (3.1)$$

where

$$\begin{aligned} E^{<}_{\rm out}(p,X) &= \{(u,v)\in E \mid p(u)-p(v)+\gamma(u,v) < 0,\ u \in X,\ v \in V\setminus X\}, \\ E^{\leq}_{\rm in}(p,X) &= \{(u,v)\in E \mid p(u)-p(v)+\gamma(u,v) \leq 0,\ u \in V\setminus X,\ v \in X\}. \end{aligned}$$

We also define $\lambda(p,X)$ by

$$\begin{aligned} \lambda(p,X) \;=\; \min\bigl\{&|p(u)-p(v)+\gamma(u,v)| \\ &\bigm| (u,v) \in E^{<}_{\rm out}(p,X) \cup E^{>}_{\rm in}(p,X)\bigr\}, \end{aligned} \qquad (3.2)$$

where

$$E^{>}_{\rm in}(p,X) = \{(u,v)\in E \mid p(u)-p(v)+\gamma(u,v) > 0,\ u \in V\setminus X, v \in X\}.$$

Then, it holds that

$$g_{\rm H}(p+\alpha\chi_X) - g_{\rm H}(p) = \alpha I(p,X) \qquad (0 \leq \forall\alpha \leq \lambda(p,X)).$$

Hassin's Algorithm

Step 0: Set $p := p^\circ$, where p° is an initial vector chosen from $\mathbb{R}^V$.

Step 1: Find $X \subseteq V$ that maximizes $I(p,X)$; if there exist more than one such X, then take a (unique) minimal one.

Step 2: If $I(p,X) \leq 0$, then stop; p is a maximizer of $g_{\rm H}$.

Step 3: Set $p := p + \lambda(p,X)\,\chi_X$. Go to Step 1. □

For each positive integer k, we denote by X_k and p_k, respectively, the set X and the vector p just after Step 1 in the k-th iteration. The next property shows that the value $I(p_k, X_k)$ is monotone nonincreasing.

Proposition 3.1 ([5]). *For $k = 1, 2, \ldots$, $I(p_k, X_k) \geq I(p_{k+1}, X_{k+1})$ holds. Moreover, if $I(p_k, X_k) = I(p_{k+1}, X_{k+1})$, then we have $X_k \subsetneq X_{k+1}$.*

It is observed in [5] that the set of possible values of $I(p, X)$ is finite, and hence the algorithm terminates in a finite number of iterations by Proposition 3.1 (see [5] for details).

3.2 Chung–Tcha's Algorithm

Suppose now that a submodular function $\rho : 2^V \to \mathbb{R}$ with $\rho(\emptyset) = \rho(V) = 0$ is given, in addition to a directed graph $G = (V, E)$, nonnegative edge capacity $c(e)$ and edge cost $\gamma(e)$ for $e \in E$. Then, the minimum cost submodular flow problem is formulated as follows:

$$\begin{array}{ll} \text{Minimize} & \displaystyle\sum_{(u,v)\in E} \gamma(u,v)x(u,v) \\ \text{subject to} & \displaystyle\sum_{u\in Y} \partial x(u) \leq \rho(Y) \quad (Y \subsetneq V), \\ & \displaystyle\sum_{u\in V} \partial x(u) = \rho(V), \\ & 0 \leq x(u,v) \leq c(u,v) \quad ((u,v) \in E). \end{array}$$

The linear programming dual is given as

$$\begin{array}{ll} \text{Maximize} & \displaystyle -\sum_{(u,v)\in E} c(u,v)s(u,v) - \sum_{Y\subseteq V} \rho(Y)t(Y) \\ \text{subject to} & \displaystyle -s(u,v) - \sum_{Y:u\in Y} t(Y) + \sum_{Y:v\in Y} t(Y) \leq \gamma(u,v) \quad ((u,v) \in E), \\ & s(u,v) \geq 0 \qquad ((u,v) \in E), \\ & t(Y) \geq 0 \quad (Y \subsetneq V), \qquad t(V) \in \mathbb{R}. \end{array}$$

It is known that for every vector $p \in \mathbb{R}^V$, the real numbers $s_p(u,v)$ $((u,v) \in E)$ and $t_p(Y)$ $(Y \subseteq V)$ defined by

$$\begin{aligned} s_p(u,v) &= -\min\{0, p(u) - p(v) + \gamma(u,v)\} \qquad ((u,v) \in E), \\ t_p(Y) &= \begin{cases} \tilde{p}_i - \tilde{p}_{i+1} & (\text{if } Y = L_i,\ 1 \leq i \leq h-1), \\ \tilde{p}_h & (\text{if } Y = L_h), \\ 0 & (\text{otherwise}) \end{cases} \end{aligned} \tag{3.3}$$

provide a feasible solution of the dual problem, where

$$\begin{array}{l} \tilde{p}_1 > \tilde{p}_2 > \cdots > \tilde{p}_h \text{ are the distinct values of components of } p, \\ L_i = \{v \in V \mid p(v) \geq \tilde{p}_i\} \qquad (i = 1, 2, \ldots, h). \end{array}$$

Moreover, some optimal solution of the dual problem can be represented in the form of (3.3) for some p (see [2, 3]; see also Theorem 5.6 and its proof in [4]). Hence, the dual problem is rewritten as follows:

$$\begin{array}{ll} \text{Maximize} & g_{\mathrm{CT}}(p) \equiv \displaystyle\sum_{(u,v)\in E} c(u,v)\min\{0, p(u)-p(v)+\gamma(u,v)\} - \hat{\rho}(p) \\ \text{subject to} & p(v) \in \mathbb{R} \quad (v \in V), \end{array}$$

where $\hat{\rho} : \mathbb{R}^V \to \mathbb{R}$ is the Lovász extension of ρ given by (2.2). It is observed that the objective function g_{CT} is expressed as $g_{\mathrm{CT}} = g_{\mathrm{H}} - \hat{\rho}$, which implies that g_{CT} is polyhedral L-concave since both of g_{H} and $-\hat{\rho}$ are polyhedral L-concave functions and polyhedral L-concavity is closed under addition.

Chung–Tcha's algorithm is described as follows. Recall the definitions of $I(p,X)$ and $\lambda(p,X)$ in (3.1) and in (3.2), respectively. We also define

$$\mu(p,X) = \min\{\tilde{p}_i - \tilde{p}_{i+1} \mid 1 \le i \le h-1,\ (L_{i+1} \setminus L_i) \cap X \neq \emptyset, \\ (L_i \setminus L_{i-1}) \setminus X \neq \emptyset\},$$

where L_0 is defined to be the empty set. Then, it holds that

$$g_{\mathrm{CT}}(p + \alpha\chi_X) - g_{\mathrm{CT}}(p) = \alpha(I(p,X) - \hat{\rho}'(p;\chi_X))$$

for every $\alpha \in \mathbb{R}$ with $0 \le \alpha \le \min\{\lambda(p,X), \mu(p,X)\}$, where $\hat{\rho}'(p;\chi_X)$ is the directional derivative[1] of $\hat{\rho}$ at p in direction χ_X.

Chung–Tcha's Algorithm

Step 0: Set $p := p^\circ$, where p° is an initial vector chosen from $\mathbb{R}^V$.

Step 1: Find $X \subseteq V$ that maximizes $I(p,X) - \hat{\rho}'(p;\chi_X)$.

Step 2: If $I(p,X) \le \hat{\rho}'(p;\chi_X)$, then stop; p is a maximizer of g_{CT}.

Step 3: Set $p := p + \min\{\lambda(p,X), \mu(p,X)\}\,\chi_X$. Go to Step 1. □

Chung and Tcha derive a pseudo-polynomial bound on the number of iterations of the algorithm by assuming that the edge costs $\gamma(e)$ are all integer-valued [2].

4 Steepest Ascent Algorithms

4.1 Algorithm for Polyhedral L-concave Functions

We consider the following steepest ascent algorithm for the maximization of a polyhedral L-concave function $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$, where it is assumed that $\mathrm{dom}\, g$ is bounded and g satisfies the property (LF2) with $r = 0$. Whereas a standard steepest ascent algorithm iteratively updates a current solution p by using a direction $q \in \mathbb{R}^V$ which maximizes the value of directional derivative $g'(p;q)$, our algorithm uses a restricted class of directions given by 0-1 vectors. Recall the definition of $\bar{c}(p;q)$ in (2.1).

[1] $\hat{\rho}'(p;\chi_X)$ admits an explicit formula [2], which is omitted here.

Steepest Ascent Algorithm for Polyhedral L-concave Function

Step 0: Set $p := p^\circ$, where p° is an initial vector chosen from $\mathrm{dom}\, g$.

Step 1: Let $X \subseteq V$ be a set maximizing the value $g'(p; \chi_X)$; if there exist more than one such X, then take a (unique) minimal one.

Step 2: If $g'(p; \chi_X) \leq 0$, then output the current vector $p = p^*$ and stop (p^* is a maximizer of g).

Step 3: Set $\lambda := \bar{c}(p; \chi_X)$ and $p := p + \lambda\chi_X$. Go to Step 1. □

We note that the minimal X that maximizes $g'(p; \chi_X)$ in Step 1 is uniquely determined by the following property:

Proposition 4.1. *Let $p \in \mathrm{dom}\, g$. If $X, Z \in \arg\max\{g'(p; \chi_Y) \mid Y \subseteq V\}$, then it holds that $X \cap Z, X \cup Z \in \arg\max\{g'(p; \chi_Y) \mid Y \subseteq V\}$.*

Proof. By the property (LF1) of g, we have

$$g'(p; \chi_X) + g'(p; \chi_Z) \leq g'(p; \chi_{X\cap Z}) + g'(p; \chi_{X\cup Z}).$$

Since $X, Z \in \arg\max\{g'(p; \chi_Y) \mid Y \subseteq V\}$, this implies $X \cap Z, X \cup Z \in \arg\max\{g'(p; \chi_Y) \mid Y \subseteq V\}$. □

The validity of the steepest ascent algorithm follows immediately from the following proposition, stating that a maximizer of a polyhedral L-concave function is characterized by a local property.

Proposition 4.2 ([9, Theorem 4.29])**.** *Let $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$ be a polyhedral L-concave function, and $p \in \mathrm{dom}\, g$. Then, p is a maximizer of g if and only if $g'(p; \chi_X) \leq 0$ for every $X \subseteq V$.*

Hence, the output of the algorithm is a maximizer of function g.

Remark 4.3. It is easy to see that the steepest ascent algorithm above coincides with Hassin's and Chung–Tcha's algorithms when applied to polyhedral L-concave functions g_{H} and g_{CT}, respectively. Our algorithm is different from Chung–Tcha's algorithm in the choice of X in Step 1. The unique minimal maximizer X of $g'(p; \chi_X)$ is chosen in our algorithm to guarantee the finite termination (see Theorem 4.5 and Corollary 4.6), whereas Chung–Tcha's algorithm takes arbitrary maximizer and imposes integrality assumption on the input to show the the finite termination. □

Remark 4.4. The steepest ascent algorithm presented in this section can also be seen as a natural generalization of the steepest ascent algorithm for L-concave functions defined on integer lattice points (see [7, 8]). □

We now present three theorems as the main results of the paper. They show nice properties which are peculiar to the steepest ascent algorithm above for polyhedral L-concave functions and are not shared by the ordinary steepest ascent algorithm for general concave functions.

It is shown first that the monotonicity property of Hassin's algorithm (Proposition 3.1) extends to the steepest ascent algorithm. Let p_k and X_k be the vector p and the set X in Step 1 of the k-th iteration. We denote by m the total number of iterations executed in the algorithm.

Theorem 4.5. *Suppose that the steepest ascent algorithm is applied to a polyhedral L-concave function g. For each $k = 1, 2, \ldots, m$, we have $g'(p_k; \chi_{X_k}) \geq g'(p_{k+1}; \chi_{X_{k+1}})$. Moreover, if $g'(p_k; \chi_{X_k}) = g'(p_{k+1}; \chi_{X_{k+1}})$, then $X_k \subsetneq X_{k+1}$.*

Proof. Proof is given in Section 5.1. □

From this property follows the finite termination of the algorithm.

Corollary 4.6. *The steepest ascent algorithm finds a maximizer of a polyhedral L-concave function $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$ with bounded* $\mathrm{dom}\, g$ *in a finite number of iterations.*

Proof. By Theorem 4.5, it suffices to show that $g'(p_k; \chi_{X_k})$ takes a value in a finite set of real numbers. Let

$$D = \{g'(p; \chi_X) \mid p \in \mathrm{dom}\, g,\ X \subseteq V,\ g'(p; \chi_X) > -\infty\}.$$

Since g is a polyhedral concave function, it can be represented as

$$g(p) = \min_{1 \leq i \leq t} \{a_i^{\mathrm{T}} p + b_i\} \qquad (p \in \mathrm{dom}\, g)$$

for some $a_i \in \mathbb{R}^V$ and $b_i \in \mathbb{R}$ $(i = 1, 2, \ldots, t)$. Hence, if $g'(p; \chi_X) > -\infty$, then we have $g'(p; \chi_X) = a_i^{\mathrm{T}} \chi_X$ for some i. This implies that D is a finite set. □

Next we show that for every initial solution p°, the maximizer p^* of g found by the algorithm is "nearest" to p° in the following sense. Note that there always exists a maximizer p of g satisfying $p \geq p^\circ$ by the property (LF2). Note also that by the property (LF1), a minimal maximizer p of g under the condition $p \geq p^\circ$ is uniquely determined.

Theorem 4.7. *Suppose that the steepest ascent algorithm is applied to a polyhedral L-concave function g with the initial solution $p^\circ \in \mathrm{dom}\, g$. Then, the output p^* of the algorithm is the unique minimal maximizer of g under the condition $p^* \geq p^\circ$; in particular, p^* satisfies*

$$\|p^* - p^\circ\|_\infty = \min\{\|p - p^\circ\|_\infty \mid p \in \arg\max g,\ p \geq p^\circ\}. \tag{4.1}$$

Proof. Proof is given in Section 5.2. □

We finally show that the trajectory of the solutions generated by the algorithm is a "shortest" path from the initial solution p° to the "nearest" maximizer p^* in the sense that the total sum of the step lengths is equal to the ℓ_∞-distance $\|p^* - p^\circ\|_\infty$ from the initial solution to the nearest optimal solution. Let λ_k be the step size λ computed in Step 3 of the k-th iteration.

Theorem 4.8. *Suppose that the steepest ascent algorithm is applied to a polyhedral L-concave function g with the initial solution $p^\circ \in \operatorname{dom} g$. Then, the output p^* of the algorithm satisfies*

$$\|p^* - p^\circ\|_\infty = \sum_{k=1}^{m-1} \lambda_k.$$

Proof. Proof is given in Section 5.2. □

Remark 4.9. In Step 1 of the steepest ascent algorithm presented above, we assume that the set X is the unique *minimal* maximizer of $g'(p; \chi_X)$. In fact, this minimality assumption is not needed to prove Theorem 4.8 and the equation (4.1) in Theorem 4.7, whereas it is required in the proofs of Theorem 4.5 and the minimality of p^* in Theorem 4.7. See Section 5.2 for details. □

4.2 Algorithm for Polyhedral L$^\natural$-concave Functions

We consider a variant of polyhedral L-concave functions, called polyhedral L$^\natural$-concave functions, and show that the steepest ascent algorithm for maximization of polyhedral L-concave functions is naturally adapted to polyhedral L$^\natural$-concave functions.

A polyhedral concave function $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$ is said to be *L$^\natural$-concave* if the function $\tilde{g} : \mathbb{R}^{\tilde{V}} \to \mathbb{R} \cup \{-\infty\}$ defined by

$$\tilde{g}(\eta, p) = g(p - \eta \mathbf{1}) \quad ((\eta, p) \in \mathbb{R} \times \mathbb{R}^V = \mathbb{R}^{\tilde{V}}) \tag{4.2}$$

is a polyhedral L-concave function, where $\tilde{V} = \{v_0\} \cup V$. Polyhedral L$^\natural$-concavity of g is characterized by the following "translation-supermodularity" [9, Theorem 4.39]:

$$g(p) + g(q) \leq g(p \vee (q - \lambda \mathbf{1})) + g((p + \lambda \mathbf{1}) \wedge q) \qquad (\forall p, q \in \operatorname{dom} g,\ \forall \lambda \geq 0).$$

We now consider the maximization of a polyhedral L$^\natural$-concave function $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$. The steepest ascent algorithm in Section 4.1 applied to the polyhedral L-concave function $\tilde{g}$ given by (4.2) yields the following algorithm for the polyhedral L$^\natural$-concave function g through the following correspondence (see also [7, Section 10.3.1]):

$$\begin{array}{ccc} \tilde{g} & & g \\ \hline (\eta, p) & \Longleftrightarrow & q = p - \eta \mathbf{1} \\ (\eta, p) + \lambda(0, \chi_X) & \Longleftrightarrow & q + \lambda \chi_X \\ (\eta, p) + \lambda(1, \chi_X) & \Longleftrightarrow & q - \lambda \chi_{V \setminus X} \\ \hline \end{array} \tag{4.3}$$

Steepest Ascent Algorithm for Polyhedral L$^\natural$-concave Function

Step 0: Set $p := p^\circ$, where p° is an initial vector chosen from $\mathrm{dom}\, g$.

Step 1: Let $\sigma \in \{+1, -1\}$ and $X \subseteq V$ be a pair of a sign and a set maximizing the value $g'(p; \sigma \chi_X)$; if there exist more than one such pair, then choose σ and X according to the following rule:
(i) if there exists such (σ, X) with $\sigma = +1$, then set $\sigma = +1$ and take a (unique) minimal X.
(ii) otherwise, set $\sigma = -1$ and take a (unique) maximal X.

Step 2: If $g'(p; \sigma \chi_X) \leq 0$, then output the current vector $p = p^*$ and stop (p^* is a maximizer of g).

Step 3: Set $\lambda := \bar{c}(p; \sigma \chi_X)$ and $p := p + \lambda \sigma \chi_X$. Go to Step 1. □

The properties of the steepest ascent algorithm for polyhedral L-concave functions in Section 4.1 can be restated in terms of polyhedral L$^\natural$-concave functions as follows. Let p_k, X_k, and σ_k be the vector p, the set X, and $\sigma \in \{+1, -1\}$ in Step 1 of the k-th iteration. Denote by m the total number of iterations executed in the algorithm.

Theorem 4.10. *Suppose that the steepest ascent algorithm is applied to a polyhedral L$^\natural$-concave function g. For each $k = 1, 2, \ldots, m$, the following properties hold:*
(i) $g'(p_k; \sigma_k \chi_{X_k}) \geq g'(p_{k+1}; \sigma_{k+1} \chi_{X_{k+1}})$.
(ii) *Suppose that* $g'(p_k; \sigma_k \chi_{X_k}) = g'(p_{k+1}; \sigma_{k+1} \chi_{X_{k+1}})$.
(ii-a) *If* $\sigma_k = -1$, *then* $\sigma_{k+1} = -1$.
(ii-b) *If* $\sigma_k = \sigma_{k+1} = +1$, *then* $X_k \subsetneq X_{k+1}$.
(ii-c) *If* $\sigma_k = \sigma_{k+1} = -1$, *then* $X_k \supsetneq X_{k+1}$.

Proof. Let $\tilde{g} : \mathbb{R}^{\tilde{V}} \to \mathbb{R} \cup \{-\infty\}$ be the L-concave function associated with g by (4.2), and apply the steepest ascent algorithm for the polyhedral L-concave function to $\tilde{g}$ with the initial vector $(0, p^\circ)$. Let $\tilde{p}_k$ and $\tilde{X}_k$ be the vector $\tilde{p}$ and the set $\tilde{X}$ in Step 1 of the k-th iteration. By the correspondence (4.3) between g and $\tilde{g}$, we have the following relations:

$$X_k = \begin{cases} \tilde{X}_k & (\text{if } \sigma_k = +1), \\ \tilde{V} \setminus \tilde{X}_k & (\text{if } \sigma_k = -1), \end{cases} \qquad g'(p_k; \sigma_k \chi_{X_k}) = \tilde{g}'(\tilde{p}_k; \chi_{\tilde{X}_k}).$$

Therefore, the statement follows immediately from Theorem 4.5. □

For a vector $q \in \mathbb{R}^V$, denote

$$\|q\|_\infty^+ = \max_{i \in V} \max(0, q(i)), \qquad \|q\|_\infty^- = \max_{i \in V} \max(0, -q(i)).$$

Note that

$$\|q\|_\infty = \max(\|q\|_\infty^+, \|q\|_\infty^-)$$

holds, and $\|q\|_\infty^+ + \|q\|_\infty^-$ serves as a norm of q (satisfying the axioms of norms). Accordingly, the value $\|p^* - p\|_\infty^+ + \|p^* - p\|_\infty^-$ represents a "distance" between two vectors p^* and p.

Theorems 4.7 and 4.8 for L-concave functions are translated to L$^\natural$-concave functions as follows.

Theorem 4.11. *Suppose that the steepest ascent algorithm is applied to a polyhedral L$^\natural$-concave function g with the initial solution $p^\circ \in \mathrm{dom}\, g$. Then, the output p^* of the algorithm satisfies*

$$\begin{aligned} &\|p^* - p^\circ\|_\infty^+ + \|p^* - p^\circ\|_\infty^- \\ &= \min\{\|p - p^\circ\|_\infty^+ + \|p - p^\circ\|_\infty^- \mid p \in \arg\max g\}. \end{aligned} \tag{4.4}$$

Let λ_k be the step size λ computed in Step 3 of the k-th iteration of the steepest ascent algorithm for L$^\natural$-concave functions.

Theorem 4.12. *Suppose that the steepest ascent algorithm is applied to a polyhedral L$^\natural$-concave function g with the initial solution $p^\circ \in \mathrm{dom}\, g$. Then, the output p^* of the algorithm satisfies*

$$\|p^* - p^\circ\|_\infty^+ + \|p^* - p^\circ\|_\infty^- = \sum_{k=1}^{m-1} \lambda_k.$$

Proof of Theorems 4.11 and 4.12.

We first prove Theorem 4.11. Let $\tilde{g} : \mathbb{R}^{\tilde{V}} \to \mathbb{R} \cup \{-\infty\}$ be the L-concave function associated with g by (4.2), and apply the steepest ascent algorithm for polyhedral L-concave functions to $\tilde{g}$ with the initial vector $(0, p^\circ)$. Let $(\zeta, q) \in \mathbb{R}^{\tilde{V}}$ be the output of the algorithm. Then, we have $(\zeta, q) \geq (0, p^\circ)$ and $p^* = q - \zeta \mathbf{1}$ by the correspondence (4.3). Theorem 4.7 for $\tilde{g}$ implies that

$$\begin{aligned} &\|(\zeta, q) - (0, p^\circ)\|_\infty \\ &= \min\{\|(\eta, p) - (0, p^\circ)\|_\infty \mid (\eta, p) \in \arg\max \tilde{g},\ (\eta, p) \geq (0, p^\circ)\}. \end{aligned} \tag{4.5}$$

The desired equality (4.4) can be derived from this by rewriting the both sides of (4.5) as follows:

$$\|(\zeta, q) - (0, p^\circ)\|_\infty = \|p^* - p^\circ\|_\infty^+ + \|p^* - p^\circ\|_\infty^-, \tag{4.6}$$

$$\begin{aligned} &\min\{\|(\eta, p) - (0, p^\circ)\|_\infty \mid (\eta, p) \in \arg\max \tilde{g},\ (\eta, p) \geq (0, p^\circ)\} \\ &= \min\{\|p - p^\circ\|_\infty^+ + \|p - p^\circ\|_\infty^- \mid p \in \arg\max g\}. \end{aligned} \tag{4.7}$$

We prove (4.6) only since (4.7) can be shown quite similarly.

We have

$$\min(\zeta, \min_{i\in V}\{q(i) - p^\circ(i)\}) = 0 \tag{4.8}$$

since otherwise for a sufficiently small positive ε, the vector $(\zeta - \varepsilon, q - \varepsilon\mathbf{1})$ is a maximizer of $\tilde{g}$ satisfying

$$(\zeta - \varepsilon, q - \varepsilon\mathbf{1}) \geq (0, p^\circ), \quad \|(\zeta - \varepsilon, q - \varepsilon\mathbf{1}) - (0, p^\circ)\|_\infty < \|(\zeta, q) - (0, p^\circ)\|_\infty,$$

a contradiction to (4.5). We also have

$$\|(\zeta, q) - (0, p^\circ)\|_\infty = \max(\zeta, \max_{i\in V}\{q(i) - p^\circ(i)\}), \tag{4.9}$$

since $(\zeta, q) \geq (0, p^\circ)$.

For the terms on the right-hand side of (4.6) we have

$$\begin{aligned}
\|p^* - p^\circ\|_\infty^+ &= \|(q - \zeta\mathbf{1}) - p^\circ\|_\infty^+ \\
&= \max(0, \max_{i\in V}\{q(i) - p^\circ(i)\} - \zeta) \\
&= \|(\zeta, q) - (0, p^\circ)\|_\infty - \zeta, \qquad (4.10)\\
\|p^* - p^\circ\|_\infty^- &= \|(q - \zeta\mathbf{1}) - p^\circ\|_\infty^- \\
&= \max(0, \max_{i\in V}\{p^\circ(i) - q(i)\} + \zeta) \\
&= -\min(\zeta, \min_{i\in V}\{q(i) - p^\circ(i)\}) + \zeta \;=\; \zeta, \qquad (4.11)
\end{aligned}$$

where the last equality of (4.10) is by (4.9) and the last equality of (4.11) is by (4.8). From (4.10) and (4.11) follows (4.6).

We next prove Theorem 4.12. Note that the value $\sum_{k=1}^{m-1} \lambda_k$ is the same for both of the L$^\natural$-concave function g and the L-concave function $\tilde{g}$. Hence, Theorem 4.8 for the L-concave function $\tilde{g}$ implies that

$$\|(\zeta, q) - (0, p^\circ)\|_\infty = \sum_{k=1}^{m-1} \lambda_k,$$

which, together with (4.6), gives the statement of Theorem 4.12. □

The theorems for polyhedral L$^\natural$-concave functions in Section 4.2 (Theorems 4.10, 4.11 and 4.12) have thus been derived from the corresponding theorems for polyhedral L-concave functions in Section 4.1 (Theorems 4.5, 4.7 and 4.8) through rather mechanical translations based on the correspondence (4.3). It is emphasized, however, that the class of polyhedral L$^\natural$-concave functions contains that of polyhedral L-concave functions as a special case, and accordingly, the theorems in Section 4.2 are more general than the theorems in Section 4.1.

5 Proofs

In this section we give proofs of Theorems 4.5, 4.7, and 4.8, where the following property of polyhedral L-concave functions is used.

Lemma 5.1 ([9, Lemma 4.28]). *Let $g : \mathbb{R}^V \to \mathbb{R} \cup \{-\infty\}$ be a polyhedral L-concave function. Then, we have*

$$g(p) + g(q) \leq g(p + \lambda\chi_X) + g(q - \lambda\chi_X)$$

for every $p, q \in \operatorname{dom} g$ with $\{i \in V \mid p(i) < q(i)\} \neq \emptyset$ and $\lambda \in \mathbb{R}$ with $0 \leq \lambda \leq \lambda' - \lambda''$, where

$$\lambda' = \max_{i \in V}\{q(i) - p(i)\}, \quad X = \{i \in V \mid q(i) - p(i) = \lambda'\},$$
$$\lambda'' = \max_{i \in V \setminus X}\{q(i) - p(i)\}.$$

5.1 Proof of Theorem 4.5

The following is the key property for the proof of Theorem 4.5. We say that $X \subseteq V$ is a *steepest ascent direction* of function g at $p \in \operatorname{dom} g$ if

$$g'(p; \chi_X) = \max\{g'(p; \chi_Y) \mid Y \subseteq V\}.$$

By Proposition 4.2 and the property (LF2) with $r = 0$, every steepest ascent direction X satisfies $\emptyset \subsetneq X \subsetneq V$ if p is not a maximizer of g.

Lemma 5.2. *Let $p \in \operatorname{dom} g$ be a vector with $p \notin \arg\max g$, $X \subseteq V$ be a steepest ascent direction of g at p, and $\lambda \in \mathbb{R}$ be a real number with $0 < \lambda \leq \bar{c}(p; \chi_X)$. Put $q = p + \lambda\chi_X$, and let $Y \subseteq V$ be a steepest ascent direction of g at q.*
(i) *It holds that $g'(q; \chi_Y) \leq g'(p; \chi_X)$.*
(ii) *If $g'(p; \chi_X) = g'(q; \chi_Y)$, then $X \cap Y$ is also a steepest ascent direction at p.*
(iii) *Suppose that $\lambda < \bar{c}(p; \chi_X)$. Then, X is a steepest ascent direction at q.*

Proof. By the choice of λ and concavity of g, we have

$$g(p + \lambda\chi_X) - g(p) = \lambda\, g'(p; \chi_X). \tag{5.1}$$

Let $\varepsilon \in \mathbb{R}$ be a positive real number with $\varepsilon < \lambda$ such that

$$g(q + \varepsilon\chi_Y) - g(q) = \varepsilon g'(q; \chi_Y). \tag{5.2}$$

Put $\hat{q} = q + \varepsilon\chi_Y$. By (5.1) and (5.2), we have

$$g(\hat{q}) - g(p) = \lambda g'(p; \chi_X) + \varepsilon g'(q; \chi_Y). \tag{5.3}$$

Note that $\hat{q}$ can be represented as

$$\hat{q} = p + \varepsilon\chi_{X\cup Y} + (\lambda - \varepsilon)\chi_X + \varepsilon\chi_{X\cap Y}.$$

Claim: The following inequalities hold:

$$g(\hat{q}) - g(\hat{q} - \varepsilon\chi_{X\cap Y}) \le \varepsilon g'(p; \chi_X), \tag{5.4}$$

$$g(p + \varepsilon\chi_{X\cup Y} + (\lambda - \varepsilon)\chi_X) - g(p + \varepsilon\chi_{X\cup Y}) \le (\lambda - \varepsilon)g'(p; \chi_X), \tag{5.5}$$

$$g(p + \varepsilon\chi_{X\cup Y}) - g(p) \le \varepsilon g'(p; \chi_X). \tag{5.6}$$

[Proof of Claim] The inequality (5.6) can be shown as follows:

$$g(p + \varepsilon\chi_{X\cup Y}) - g(p) \le \varepsilon g'(p; \chi_{X\cup Y}) \le \varepsilon g'(p; \chi_X),$$

where the first inequality is by the concavity of g and the second inequality follows from the fact that X is a steepest ascent direction of g at p.

We then prove (5.4). It may be assumed that $X \cap Y \ne \emptyset$ since otherwise

$$g(\hat{q}) - g(\hat{q} - \varepsilon\chi_{X\cap Y}) = g(\hat{q}) - g(\hat{q}) = 0 < \varepsilon g'(p; \chi_X)$$

holds, where the inequality follows from $p \notin \arg\max g$ and Proposition 4.2. Since $\arg\max\{\hat{q}(i) - p(i) \mid i \in V\} = X \cap Y$, Lemma 5.1 implies that

$$g(p) + g(\hat{q}) \le g(p + \varepsilon\chi_{X\cap Y}) + g(\hat{q} - \varepsilon\chi_{X\cap Y}),$$

from which follows that

$$g(\hat{q}) - g(\hat{q} - \varepsilon\chi_{X\cap Y}) \le g(p + \varepsilon\chi_{X\cap Y}) - g(p) \le \varepsilon g'(p; \chi_{X\cap Y}) \le \varepsilon g'(p; \chi_X). \tag{5.7}$$

The inequality (5.5) can be shown in a similar way as (5.4) as follows. Lemma 5.1 implies that

$$\begin{aligned} & g(p) + g(p + \varepsilon\chi_{X\cup Y} + (\lambda - \varepsilon)\chi_X) \\ \le\ & g(p + (\lambda - \varepsilon)\chi_X) + g((p + \varepsilon\chi_{X\cup Y} + (\lambda - \varepsilon)\chi_X) - (\lambda - \varepsilon)\chi_X) \\ =\ & g(p + (\lambda - \varepsilon)\chi_X) + g(p + \varepsilon\chi_{X\cup Y}), \end{aligned}$$

from which follows that

$$\begin{aligned} & g((p + \varepsilon\chi_{X\cup Y} + (\lambda - \varepsilon)\chi_X) - g(p + \varepsilon\chi_{X\cup Y}) \\ \le\ & g(p + (\lambda - \varepsilon)\chi_X) - g(p) \le (\lambda - \varepsilon)g'(p; \chi_X). \end{aligned}$$

[End of Claim]

The inequalities (5.4), (5.5), and (5.6) imply

$$g(\hat{q}) - g(p) \le (\lambda + \varepsilon)g'(p; \chi_X). \tag{5.8}$$

Then, the claim (i) follows from (5.3) and (5.8).

To prove the claim (ii), assume that $g'(p;\chi_X) = g'(q;\chi_Y)$. It follows from (5.3) that $g(\hat{q}) - g(p) = (\lambda + \varepsilon)g'(p;\chi_X)$, which, together with the inequalities (5.4), (5.5), and (5.6), implies that all the inequalities (5.4), (5.5), and (5.6) hold with equality. In particular, we have

$$g(\hat{q}) - g(\hat{q} - \varepsilon\chi_{X\cap Y}) = \varepsilon g'(p;\chi_X), \tag{5.9}$$

from which follows that $X \cap Y \neq \emptyset$ since $g'(p;\chi_X) > 0$ by Proposition 4.2. By (5.7) and (5.9), we have

$$\varepsilon g'(p;\chi_X) = g(\hat{q}) - g(\hat{q} - \varepsilon\chi_{X\cap Y}) \leq g(p + \varepsilon\chi_{X\cap Y}) - g(p) \leq \varepsilon g'(p;\chi_{X\cap Y}).$$

This shows that $X \cap Y$ is also a steepest ascent direction of g at p.

We then prove the claim (iii). For $\lambda < \bar{c}(p;\chi_X)$, we have $g'(q;\chi_X) = g'(p;\chi_X)$, which, together with (i), implies

$$g'(q;\chi_X) = g'(p;\chi_X) \geq g'(q;\chi_Y),$$

i.e., X is a steepest ascent direction at q. □

We now prove Theorem 4.5. The inequality $g'(p_k;\chi_{X_k}) \geq g'(p_{k+1};\chi_{X_{k+1}})$ follows immediately from Lemma 5.2 (i). To prove the latter part of Theorem 4.5, suppose that $g'(p_k;\chi_{X_k}) = g'(p_{k+1};\chi_{X_{k+1}})$ holds. By the choice of p_{k+1}, we have $X_k \neq X_{k+1}$. Lemma 5.2 (ii) implies that $X_k \cap X_{k+1}$ is also a steepest ascent direction of g at p_k. Since X_k is the minimal steepest ascent direction at p_k, we have $X_k \cap X_{k+1} = X_k$, i.e., $X_k \subsetneq X_{k+1}$ holds.

5.2 Proof of Theorems 4.7 and 4.8

To prove Theorems 4.7 and 4.8, we first present some properties of steepest ascent directions of a polyhedral L-concave function.

Lemma 5.3. *Let $p \in \operatorname{dom} g$ be a vector with $p \notin \arg\max g$, and X be a (not necessarily minimal) steepest ascent direction of g at p. Also, let $\hat{p} \in \arg\max g$ be the unique minimal maximizer of g under the condition $\hat{p} \geq p$.*
(i) $\arg\max_{i\in V}\{\hat{p}(i) - p(i)\} \subseteq X$ *holds.*
(ii) *Let $B = \{i \in V \mid \hat{p}(i) = p(i)\}$. Then, $X \setminus B$ is also a steepest ascent direction at p; in particular, $X \cap B = \emptyset$ holds if X is the unique minimal steepest ascent direction at p.*

Proof. Denote

$$A = \arg\max_{i\in V}\{\hat{p}(i) - p(i)\}.$$

Since $\hat{p} \neq p$ and g satisfies the property (LF2) with $r = 0$, we have $\max_{i\in V}\{\hat{p}(i) - p(i)\} > 0$ and $B \neq \emptyset$. Let ε be a sufficiently small positive real number such that

$$g(p + \varepsilon\chi_X) - g(p) = \varepsilon g'(p;\chi_X). \tag{5.10}$$

Assume, to the contrary, that $A \setminus X \neq \emptyset$ holds. Then, we have

$$\arg\max_{i \in V}\{\hat{p}(i) - (p + \varepsilon\chi_X)(i)\} = A \setminus X \neq \emptyset.$$

Hence, Lemma 5.1 implies that

$$\begin{aligned} g(\hat{p}) + g(p + \varepsilon\chi_X) &\leq g(\hat{p} - \varepsilon\chi_{A\setminus X}) + g(p + \varepsilon\chi_X + \varepsilon\chi_{A\setminus X}) \\ &= g(\hat{p} - \varepsilon\chi_{A\setminus X}) + g(p + \varepsilon\chi_{X\cup A}). \end{aligned} \tag{5.11}$$

Since $A \setminus X \subseteq \{i \in V \mid \hat{p}(i) > p(i)\}$ and ε is sufficiently small, we have $\hat{p} \geq \hat{p} - \varepsilon\chi_{A\setminus X} \geq p$, implying that $g(\hat{p}) > g(\hat{p} - \varepsilon\chi_{A\setminus X})$ by the choice of $\hat{p}$. This inequality, together with (5.11), implies that $g(p+\varepsilon\chi_{X\cup A}) > g(p+\varepsilon\chi_X)$. From this inequality and (5.10) follows that

$$\varepsilon g'(p; \chi_{X\cup A}) \geq g(p + \varepsilon\chi_{X\cup A}) - g(p) > g(p + \varepsilon\chi_X) - g(p) = \varepsilon g'(p; \chi_X),$$

where the first inequality is by the concavity of g. This, however, is a contradiction to the choice of X. Hence, we have $A \subseteq X$.

We then show that $X \setminus B$ is also a steepest ascent direction at p. We may assume that $X \cap B \neq \emptyset$. Then, we have

$$\arg\max_{i \in V}\{(p + \varepsilon\chi_X)(i) - \hat{p}(i)\} = X \cap B \neq \emptyset.$$

It follows from Lemma 5.1 that

$$\begin{aligned} g(p + \varepsilon\chi_X) + g(\hat{p}) &\leq g(p + \varepsilon\chi_X - \varepsilon\chi_{X\cap B}) + g(\hat{p} + \varepsilon\chi_{X\cap B}) \\ &= g(p + \varepsilon\chi_{X\setminus B}) + g(\hat{p} + \varepsilon\chi_{X\cap B}). \end{aligned} \tag{5.12}$$

Since $\hat{p}$ is a maximizer of g, we have $g(\hat{p}) \geq g(\hat{p} + \varepsilon\chi_{X\cap B})$, which, together with (5.12), implies $g(p + \varepsilon\chi_X) \leq g(p + \varepsilon\chi_{X\setminus B})$. Hence, it follows that

$$\begin{aligned} \varepsilon g'(p, \chi_X) &= g(p + \varepsilon\chi_X) - g(p) \\ &\leq g(p + \varepsilon\chi_{X\setminus B}) - g(p) \\ &\leq \varepsilon g'(p, \chi_{X\setminus B}), \end{aligned}$$

i.e., $X \setminus B$ is also a steepest ascent direction at p. □

Lemma 5.4. *Let p, X, and $\hat{p}$ be as in Lemma 5.3, and put*

$$B = \{i \in V \mid \hat{p}(i) = p(i)\}, \quad \alpha = \min\{\hat{p}(i) - p(i) \mid i \in V,\ \hat{p}(i) - p(i) > 0\}.$$

For every λ with $0 \leq \lambda \leq \min\{\alpha, \bar{c}(p; \chi_X)\}$, the vector $\hat{q}_\lambda = \hat{p} + \lambda\chi_{X\cap B}$ satisfies the following properties:
(i) *$\hat{q}_\lambda$ is the unique minimal maximizer of g under the condition $\hat{q}_\lambda \geq p + \lambda\chi_X$.*
(ii) $\|\hat{q}_\lambda - (p + \lambda\chi_X)\|_\infty = \|\hat{p} - p\|_\infty - \lambda$.

Proof. [Proof of (i)] Since $0 \leq \lambda \leq \alpha$, the inequality $\hat{p}+\lambda\chi_{X\cap B} \geq p+\lambda\chi_X$ holds.

We then prove that $\hat{q}_\lambda$ is a maximizer of g for every λ with $0 \leq \lambda \leq \min\{\alpha, \bar{c}(p;\chi_X)\}$. By the concavity of g and $\hat{p} \in \arg\max g$, it suffices to prove $\hat{q}_\lambda \in \arg\max g$ for $\lambda = \min\{\alpha, \bar{c}(p;\chi_X)\}$.

We may assume $X \cap B \neq \emptyset$ since otherwise the claim holds immediately. Let λ^* be the maximum real number with $0 \leq \lambda^* \leq \min\{\alpha, \bar{c}(p;\chi_X)\}$ such that $\hat{p} + \lambda^*\chi_{X\cap B} \in \arg\max g$. In the following, we assume, to the contrary, that $\lambda^* < \min\{\alpha, \bar{c}(p;\chi_X)\}$ and derive a contradiction.

Since $\lambda^* < \min\{\alpha, \bar{c}(p;\chi_X)\}$, there exists a sufficiently small positive ε with $\lambda^* + \varepsilon < \min\{\alpha, \bar{c}(p;\chi_X)\}$ such that

$$\begin{aligned} & q' \equiv p + (\lambda^* + \varepsilon)\chi_X \in \operatorname{dom} g, \\ & g(q') - g(p + \lambda^*\chi_X) = \varepsilon g'(p + \lambda^*\chi_X; \chi_X). \end{aligned} \tag{5.13}$$

Let $q'' = \hat{p} + \lambda^*\chi_{X\cap B} \in \arg\max g$. Then, it holds that

$$\arg\max_{i\in V}\{q'(i) - q''(i)\} = X \cap B \neq \emptyset.$$

It follows from Lemma 5.1 that

$$\begin{aligned} g(q') + g(q'') & \leq g(q' - \varepsilon\chi_{X\cap B}) + g(q'' + \varepsilon\chi_{X\cap B}) \\ & = g(p + \lambda^*\chi_X + \varepsilon\chi_{X\setminus B}) + g(\hat{p} + (\lambda^* + \varepsilon)\chi_{X\cap B}). \end{aligned} \tag{5.14}$$

Note that the set X is a steepest ascent direction at $p + \lambda^*\chi_X$ by Lemma 5.2 (iii) since $\lambda^* < \bar{c}(p;\chi_X)$. Hence, we have $g'(p + \lambda^*\chi_X;\chi_X) \geq g'(p + \lambda^*\chi_X;\chi_{X\setminus B})$, from which follows that

$$\begin{aligned} g(q') - g(p + \lambda^*\chi_X) & = \varepsilon g'(p + \lambda^*\chi_X;\chi_X) \\ & \geq \varepsilon g'(p + \lambda^*\chi_X;\chi_{X\setminus B}) \\ & \geq g(p + \lambda^*\chi_X + \varepsilon\chi_{X\setminus B}) - g(p + \lambda^*\chi_X), \end{aligned} \tag{5.15}$$

where the equality follows from (5.13) and the last inequality is by the concavity of g. By (5.14) and (5.15), it holds that $g(q'') \leq g(\hat{p} + (\lambda^* + \varepsilon)\chi_{X\cap B})$, i.e., $\hat{p} + (\lambda^* + \varepsilon)\chi_{X\cap B}$ is a maximizer of g, a contradiction to the definition of λ^*. Hence, we have $\hat{q}_\lambda \in \arg\max g$ for $\lambda = \min\{\alpha, \bar{c}(p;\chi_X)\}$.

We finally show the minimality of the maximizer $\hat{q}_\lambda$ under the condition $\hat{q}_\lambda \geq p + \lambda\chi_X$. Let q^* be a maximizer of g satisfying $q^* \geq p + \lambda\chi_X$. Then, the property (LF1) for g implies that the vector $\hat{p} \wedge q^*$ is a maximizer of g and satisfies $\hat{p} \wedge q^* \geq p$. Hence, we have $\hat{p} \wedge q^* \geq \hat{p}$ by the definition of $\hat{p}$. From this inequality follows that $q^* \geq \hat{p}$. Since $q^* \geq p + \lambda\chi_X$ holds by assumption, it holds that

$$q^* \geq \hat{p} \vee (p + \lambda\chi_X) = \hat{q}_\lambda,$$

from which follows that $\hat{q}_\lambda$ is the unique minimal maximizer of g under the condition $\hat{q}_\lambda \geq p + \lambda\chi_X$.

[Proof of (ii)] We show that

$$\|(\hat{p} + \mu\chi_{X\cap B}) - (p + \mu\chi_X)\|_\infty = \|\hat{p} - p\|_\infty - \mu \tag{5.16}$$

for every μ with $0 \leq \mu < \min\{\alpha, \bar{c}(p; \chi_X)\}$.

Put $\hat{q}_\mu = \hat{p} + \mu\chi_{X\cap B}$ and $q_\mu = p + \mu\chi_X$. By the claim (i) of this proposition, $\hat{q}_\mu$ is the unique minimal maximizer of g under the condition $\hat{q}_\mu \geq q_\mu$. In addition, the set X is a steepest ascent direction at q by Lemma 5.2 (iii) since $0 \leq \mu < \bar{c}(p; \chi_X)$. Hence, Lemma 5.3 applied to q_μ, X, and $\hat{q}_\mu$ implies that

$$\arg\max_{i\in V}\{\hat{q}_\mu(i) - q_\mu(i)\} \subseteq X.$$

It follows that

$$\|\hat{q}_\mu - q_\mu\|_\infty = \max_{i\in X}\{\hat{q}_\mu(i) - q_\mu(i)\}. \tag{5.17}$$

In the same way, by using Lemma 5.3 we can show that

$$\|\hat{p} - p\|_\infty = \max_{i\in X}\{\hat{p}(i) - p(i)\}. \tag{5.18}$$

We also have $\hat{p}(i) - p(i) = \hat{q}_\mu(i) - q_\mu(i) = 0$ for $i \in B$. Hence, it holds that

$$\begin{aligned}\max_{i\in X}\{\hat{q}_\mu(i) - q_\mu(i)\} &= \max_{i\in X\setminus B}\{\hat{q}_\mu(i) - q_\mu(i)\} \\ &= \max_{i\in X\setminus B}\{\hat{p}(i) - p(i)\} - \mu \\ &= \max_{i\in X}\{\hat{p}(i) - p(i)\} - \mu. \end{aligned} \tag{5.19}$$

From (5.17), (5.18), and (5.19) follows (5.16). □

Lemma 5.5. *Let p, X, and $\hat{p}$ be as in Lemma 5.3, and put $\lambda^* = \bar{c}(p; \chi_X)$. Then, the vector $\hat{q} = \hat{p} \vee (p + \lambda^*\chi_X)$ is the unique minimal maximizer of g under the condition $\hat{q} \geq p + \lambda^*\chi_X$ and satisfies*

$$\|\hat{q} - (p + \lambda^*\chi_X)\|_\infty = \|\hat{p} - p\|_\infty - \lambda^*. \tag{5.20}$$

Moreover, if X is the unique minimal steepest ascent direction at p, then $\hat{q} = \hat{p}$.

Proof. Note that $\min_{i\in V}\{\hat{p}(i) - p(i)\} = 0$ holds by the property (LF2) and the choice of $\hat{p}$. Let $\delta_0 < \delta_1 < \cdots < \delta_{s-1} < \lambda^*$ be the distinct numbers in

$$\{\hat{p}(i) - p(i) \mid i \in V,\ \hat{p}(i) - p(i) < \lambda^*\},$$

and put $\delta_s = \lambda^*$. Note that $\delta_0 = 0$ since $\min_{i \in V}\{\hat{p}(i) - p(i)\} = 0$. For $h = 0, 1, \ldots, s$, we define

$$q_h = p + \delta_h \chi_X, \qquad X_h = \{i \in X \mid \hat{p}(i) - p(i) < \delta_h\},$$
$$\hat{q}_h = \hat{p} \vee (p + \delta_h \chi_X) = \hat{p} + \sum_{j=1}^{h} (\delta_j - \delta_{j-1}) \chi_{X_j}.$$

Then, it holds that

$$q_0 = p, \quad q_s = p + \lambda^* \chi_X, \quad q_{h+1} = q_h + (\delta_{h+1} - \delta_h)\chi_X \ (h = 0, 1, \ldots, s-1),$$
$$\hat{q}_0 = \hat{p}, \quad \hat{q}_s = \hat{q}, \quad \hat{q}_{h+1} = \hat{q}_h + (\delta_{h+1} - \delta_h)\chi_{X_{h+1}} \ (h = 0, 1, \ldots, s-1).$$

In addition, we have

$$X_{h+1} = \{i \in X \mid \hat{q}_h(i) = q_h(i)\}, \tag{5.21}$$
$$\min\{\hat{q}_h(i) - q_h(i) \mid i \in V,\ \hat{q}_h(i) - q_h(i) > 0\} = \delta_{h+1} - \delta_h$$

for $h = 0, 1, \ldots, s-1$. Hence, the repeated application of Lemma 5.4 implies that for each $h = 1, 2, \ldots, s$, the vector $\hat{q}_h$ is the unique minimal maximizer of g under the condition $\hat{q}_h \geq q_h$ and satisfies

$$\|\hat{q}_h - q_h\|_\infty = \|\hat{q}_{h-1} - q_{h-1}\|_\infty - (\delta_h - \delta_{h-1}). \tag{5.22}$$

This, in particular, implies that the vector $\hat{q}$ is the unique minimal maximizer of g under the condition $\hat{q} \geq p + \lambda^* \chi_X$ since $q_s = p + \lambda^* \chi_X$ and $\hat{q}_s = \hat{q}$. From (5.22) follows that

$$\begin{aligned} \|\hat{q} - (p + \lambda^* \chi_X)\|_\infty &= \|\hat{q}_s - q_s\|_\infty \\ &= \|\hat{q}_0 - q_0\|_\infty - \sum_{h=1}^{s} (\delta_h - \delta_{h-1}) \ = \ \|\hat{p} - p\|_\infty - \lambda^*, \end{aligned}$$

i.e., (5.20) holds.

We then assume that X is the unique minimal steepest ascent direction at p. For $h = 0, 1, \ldots, s-1$, the set X is also a steepest ascent direction at q_h by Lemma 5.2 (iii) since $\delta_h < \lambda^* = \bar{c}(p; \chi_X)$ holds. Moreover, X is the unique minimal steepest ascent direction at q_h by Lemma 5.2 (ii) since $g'(q_h; \chi_X) = g'(p; \chi_X)$. Therefore, Lemma 5.3 and the equation (5.21) imply $X_h = \emptyset$ for $h = 1, 2, \ldots, s$, from which follows that $\hat{q} = \hat{p}$. □

Theorems 4.7 and 4.8 can be proved as follows. Let p^* be a maximizer of g which is the output of the algorithm. For $k = 1, 2, \ldots, m$, we denote by $\hat{p}_k$ the unique minimal maximizer of g under the condition $\hat{p}_k \geq p_k$. Note that p_m is the output of the algorithm and therefore $p_m = \hat{p}_m = p^*$ holds. The repeated application of Lemma 5.5 implies that for $k = 1, 2, \ldots, m-1$, the vector $\check{q}_k = \hat{p}_k \vee (p_k + \lambda_k \chi_{X_k})$ is the unique minimal maximizer of g under the condition $\check{q}_k \geq p_k + \lambda_k \chi_{X_k} = p_{k+1}$ and satisfies

$$\|\check{q}_k - (p_k + \lambda_k \chi_X)\|_\infty = \|\hat{p}_k - p_k\|_\infty - \lambda_k.$$

This shows that $\check{q}_k = \hat{p}_{k+1}$ and

$$\|\hat{p}_{k+1} - p_{k+1}\|_\infty = \|\hat{p}_k - p_k\|_\infty - \lambda_k.$$

Hence, it follows that

$$\begin{aligned}\sum_{k=1}^{m-1} \lambda_k &= \sum_{k=1}^{m-1} \left[\|\hat{p}_k - p_k\|_\infty - \|\hat{p}_{k+1} - p_{k+1}\|_\infty \right] \\ &= \|\hat{p}_1 - p_1\|_\infty - \|\hat{p}_m - p_m\|_\infty \\ &= \|\hat{p}_1 - p^\circ\|_\infty \\ &= \min\{\|p - p^\circ\|_\infty \mid p \in \arg\max g,\ p \geq p^\circ\}, \end{aligned} \tag{5.23}$$

where the third equality is by $p_1 = p^\circ$ and $p_m = \hat{p}_m$. Since $p^* = p_m \geq p_1 = p^\circ$, we have

$$\|p^* - p^\circ\|_\infty \geq \min\{\|p - p^\circ\|_\infty \mid p \in \arg\max g,\ p \geq p^\circ\}. \tag{5.24}$$

We also have

$$\|p^* - p^\circ\|_\infty = \|p_m - p_1\|_\infty = \left\| \sum_{k=1}^{m-1} \lambda_k \chi_{X_k} \right\|_\infty \leq \sum_{k=1}^{m-1} \lambda_k. \tag{5.25}$$

From (5.23), (5.24), and (5.25) follows that the inequalities (5.24) and (5.25) hold with equality, i.e., the equation (4.1) in Theorem 4.7 and Theorem 4.8 hold.

We then assume that X_k is the unique minimal steepest ascent direction at p_k for each $k = 1, 2, \ldots, m-1$. Then, the repeated application of Lemma 5.5 implies that $\hat{p}_k = \hat{p}_1$ for all $k = 2, 3, \ldots, m$. In particular, we have $p^* = \hat{p}_m = \hat{p}_1$. Hence, p^* is the unique minimal maximizer of g under the condition $p^* \geq p^\circ$. This completes the proof of Theorem 4.7.

References

[1] Ahuja, R.K., Magnanti, T.L., Orlin, J.B.: Network Flows: Theory, Algorithms, and Applications. Prentice Hall, Upper Side River (1993).

[2] Chung, N.-k., Tcha, D.-w.: A dual algorithm for submodular flow problems. Oper. Res. Letters **10**, 489–495 (1991).

[3] Edmonds, J., Giles, R.: A min-max relation for submodular functions on graphs. Ann. Discrete Math. **1**, 185–204 (1977).

[4] Fujishige, F.: Submodular Functions and Optimization, Second Edition. Elsevier, Amsterdam (2005).

[5] Hassin, R.: The minimum cost flow problem: a unifying approach to dual algorithms and a new tree-search algorithm. Math. Program. **25**, 228–239 (1983).

[6] Murota, K.: Discrete convex analysis. Math. Program. **83**, 313–371 (1998).

[7] Murota, K.: Discrete Convex Analysis. SIAM, Philadelphia (2003).

[8] Murota, K.: On steepest descent algorithms for discrete convex functions. SIAM J. Optim. **14**, 699–707 (2003).

[9] Murota, K., Shioura, A: Extension of M-convexity and L-convexity to polyhedral convex functions. Adv. Appl. Math. **25**, 352–427 (2000).

[10] Murota, K., Shioura, A: Dijkstra's algorithm and L-concave function maximization. Math. Program., published online (2013).

[11] Schrijver, A.: Combinatorial Optimization: Polyhedra and Efficiency. Springer, Berlin (2003).