MATHEMATICAL ENGINEERING TECHNICAL REPORTS

On the Pipage Rounding Algorithm for Submodular Function Maximization —A View from Discrete Convex Analysis—

Akiyoshi SHIOURA

(Communicated by Kazuo MUROTA)

METR 2008–03 January 2008

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 the Pipage Rounding Algorithm for Submodular Function Maximization —A View from Discrete Convex Analysis—

Akiyoshi SHIOURA

Graduate School of Information Sciences
Tohoku University
shioura@dais.is.tohoku.ac.jp

January 2008

Abstract

We consider the problem of maximizing a nondecreasing submodular set function under a matroid constraint. Recently, Calinescu et al. (2007) proposed an elegant framework for the approximation of this problem, which is based on the pipage rounding technique by Ageev and Sviridenko (2004), and showed that this framework indeed yields a $(1 - 1/e)$-approximation algorithm for the class of submodular functions which are represented as the sum of weighted rank functions of matroids. This paper sheds a new light on this result from the viewpoint of discrete convex analysis by extending it to the class of submodular functions which are the sum of M$^\natural$-concave functions. M$^\natural$-concave functions are a class of discrete concave functions introduced by Murota and Shioura (1999), and contain the class of the sum of weighted rank functions as a proper subclass. Our result provides a better understanding for why the pipage rounding algorithm works for the sum of weighted rank functions.

1 Introduction

We consider the maximization of a nondecreasing submodular function under a matroid constraint. In the area of mathematical programming, the maximization of a concave function is recognized as a tractable problem while the maximization of a convex function is hard to solve. In discrete optimization, submodular function is often regarded as discrete convexity, and indeed the maximization of a submodular function is known to be NP-hard. On the other hand, some classes of submodular functions are deeply related to discrete concavity (cf. [6, 10, 14]). For example, a set function

$f(X) = \varphi(|X|)$ given by a univariate concave function φ is a submodular function, and it is natural that such a function has discrete concavity. The objective of this paper is to shed a new light on the pipage rounding algorithm from the viewpoint of discrete convex analysis by pointing out that discrete concavity plays an essential role in computing an approximate solution in the maximization of a submodular function.

Our problem is formulated as follows:

$$\text{(P)} \quad \text{Maximize } f(X) \quad \text{subject to } X \in \mathcal{F},$$

where $f : 2^N \to \mathbb{R}$ is a nondecreasing submodular set function on a finite set N with $f(\emptyset) = 0$, and $\mathcal{M} = (N, \mathcal{F})$ is a matroid with the family of independence sets $\mathcal{F}$. We assume that the membership oracle for $\mathcal{M}$ is available. A set function $f : 2^N \to \mathbb{R}$ is said to be *submodular* if it satisfies $f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$ for any $X, Y \in 2^N$, and *nondecreasing* if $f(X) \leq f(Y)$ for any $X, Y \in 2^N$ with $X \subseteq Y$.

In the literature, various problems related to (P) have been discussed over decades [3, 4, 7, 17, 20]. Recently, Calinescu et al. [2] proposed an elegant framework for the approximation of the problem (P), which is based on the pipage rounding technique developed by Ageev and Sviridenko [1]. In their framework, they firstly consider a relaxation of the problem (P):

$$\text{(RP)} \quad \text{Maximize } \tilde{f}(x) \quad \text{subject to } x \in P(\mathcal{M}),$$

where $P(\mathcal{M})$ $(\subseteq \mathbb{R}^N)$ is the matroid polytope of $\mathcal{M}$ and $\tilde{f} : [0,1]^N \to \mathbb{R}$ is an *extension* of f, i.e., a nondecreasing concave function such that $\tilde{f}(\chi_X) = f(X)$ for any $X \in 2^N$ and its characteristic vector $\chi_X \in \{0,1\}^N$. Then, an optimal (fractional) solution $x \in [0,1]^N$ of the relaxed problem (RP) is computed and rounded to a $\{0,1\}$-vector that corresponds to an independent set of $\mathcal{M}$ by using a potential function defined over $[0,1]^N$. The main result of Calinescu et al. [2] is described as follows, where e denotes the base of natural logarithm, and for a matroid $\mathcal{M}' = (N, \mathcal{F}')$ and a nonnegative vector $w \in \mathbb{R}^N_+$, a *weighted rank function* $f : 2^N \to \mathbb{R}$ is defined by

$$f(X) = \max\{w(Y) \mid Y \in \mathcal{F}',\ Y \subseteq X\} \qquad (X \in 2^N). \tag{1}$$

Theorem 1.1 ([2]). *Let $f : 2^N \to \mathbb{R}$ be a nondecreasing submodular function with $f(\emptyset) = 0$. Suppose that there exists an extension $\tilde{f}$ of f such that the relaxed problem* (RP) *can be solved in polynomial time. Then, the pipage rounding algorithm (see Section 2.2) outputs a $(1 - 1/e)$-approximate solution of the problem* (P) *in polynomial time. In particular, if f is given as the sum of weighted rank functions, the pipage rounding algorithm outputs a $(1 - 1/e)$-approximate solution in polynomial time.*

A connection of this result to discrete concavity is made by the observation that a weighted rank function has discrete concavity called *$M^\natural$-concavity* (see Section 2.3 for the definition).

Theorem 1.2. *Any weighted rank function is an M$^\natural$-concave function.*

The concepts of M$^\natural$-concavity/M$^\natural$-convexity are introduced by Murota and Shioura [15] as discrete concavity/convexity for functions defined over the integer lattice, and are variants of M-concavity/M-convexity due to Murota [13]. These concepts play primary roles in the theory of discrete convex analysis [14]. The class of M$^\natural$-concave functions properly contains that of weighted rank functions; for example, the set function $f(X) = \varphi(|X|)$ with concave φ is an M$^\natural$-concave function and not a weighted rank function. Therefore, the class of the sum of M$^\natural$-concave functions contains the class of the sum of weighted rank functions, but so far we do not know whether this is a proper inclusion or not.

An M$^\natural$-concave function has a natural extension called the *concave closure* (see Section 2 for the precise definition). This enables us to solve the maximization of the sum of the concave closures of M$^\natural$-concave functions (almost) optimally in polynomial time. We assume that the membership oracle for $\mathcal{M}$ and the function evaluation oracles for M$^\natural$-concave functions are available. We denote by n the cardinality of N.

Theorem 1.3. *Let $f_k : 2^N \to \mathbb{R}$ $(k = 1, 2, \ldots, m)$ be a family of nondecreasing M$^\natural$-concave functions with $f_k(\emptyset) = 0$, and denote by $\overline{f}_k : [0,1]^N \to \mathbb{R}$ the concave closure of f_k for $k = 1, 2, \ldots, m$. Suppose that the function $\tilde{f}$ in the problem* (RP) *is given as $\tilde{f}(x) = \sum_{k=1}^{m} \overline{f}_k(x)$.*
(i) *For any $\varepsilon > 0$, a $(1 - \varepsilon)$-approximate solution of* (RP) *can be computed in time polynomial in n, m, Γ, and $\log(1/\varepsilon)$, where*

$$\Gamma = \sum_{k=1}^{m} \max_{X \in 2^N} |\log f_k(X)|. \tag{2}$$

(ii) *If each f_k is an integer-valued function, then an optimal solution of* (RP) *can be computed in time polynomial in n, m, and Γ.*

Our algorithm used in the proof of Theorem 1.3 is based on the ellipsoid method combined with an algorithm for computing a subgradient of the concave function $\tilde{f}$. Since $\tilde{f}(x) = \sum_{k=1}^{m} \overline{f}_k(x)$, a subgradient of $\tilde{f}$ is given as the sum of subgradients of the functions $\overline{f}_k$ $(k = 1, 2, \ldots, m)$, and subgradients of each $\overline{f}_k$ are computed in polynomial time by using the combinatorial structure of M$^\natural$-concave functions.

As a corollary of Theorem 1.3, we see that the pipage rounding algorithm of Calinescu et al. [2] also works for the sum of M$^\natural$-concave functions.

Corollary 1.4. *Suppose that the function f is given as $f(X) = \sum_{k=1}^{m} f_k(X)$ with a family of nondecreasing M$^\natural$-concave functions $f_k : 2^N \to \mathbb{R}$ $(k = 1, 2, \ldots, m)$.*
(i) *For any $\varepsilon > 0$, a $(1 - 1/e - \varepsilon)$-approximate solution of the problem* (P)

can be obtained in time polynomial in n*,* m*,* Γ*, and* $\log(1/\varepsilon)$*.*
(ii) *If each* f_k *is an integer-valued function, then a* $(1-1/e)$*-approximate solution of* (P) *can be obtained in time polynomial in* n*,* m*, and* Γ*.*

Our results show that the success of the pipage rounding algorithm for the sum of weighted rank functions can be understood as a special case of Corollary 1.4.

The organization of this paper is as follows. In Section 2, we review the pipage rounding framework of Calinescu et al. [2] as well as the definition and some fundamental properties of $\mathrm{M}^\natural$-concavity. In Section 3, we present an algorithm for computing a subgradient of the concave closure of an $\mathrm{M}^\natural$-concave function. Finally, we propose a polynomial-time algorithm for the maximization of the sum of concave closures in Section 4.

2 Preliminaries

2.1 Matroids

Throughout this paper, we assume that $\mathcal{M} = (N, \mathcal{F})$ is a matroid with the family of independent sets $\mathcal{F}$, which gives a constraint in the problem (P). We denote by $\mathcal{B}$ the base family of $\mathcal{M}$, and let $r_\mathcal{M} : 2^N \to \{0, 1, \ldots, n\}$ be the rank function of $\mathcal{M}$. For any $X \in 2^N$, we denote by $\chi_X \in \{0,1\}^N$ the characteristic vector of X, i.e., $(\chi_X)(j) = 1$ $(j \in X)$ and $(\chi_X)(j) = 0$ $(j \in N \setminus X)$. The matroid polytope $P(\mathcal{M})$ (resp., the base polytope $B(\mathcal{M})$) is defined as the convex hull of the set of $\{0,1\}$-vectors $\{\chi_X \mid X \in \mathcal{F}\}$ (resp., $\{\chi_X \mid X \in \mathcal{B}\}$). They are also given as

$$\begin{aligned} P(\mathcal{M}) &= \{x \in \mathbb{R}^N \mid x(X) \leq r_\mathcal{M}(X)\ (X \subseteq N)\}, \\ B(\mathcal{M}) &= \{x \in \mathbb{R}^N \mid x \in P(\mathcal{M}),\ x(N) = r_\mathcal{M}(N)\}. \end{aligned}$$

Given a vector $x \in P(\mathcal{M})$, we say that a subset X of N is *tight* if $x(X) = r_\mathcal{M}(X)$.

In the following, we assume that the membership oracle for $\mathcal{F}$ is available. Since the function value of the matroid rank function $r_\mathcal{M}$ for $\mathcal{M}$ can be computed by using the membership oracle at most n times, the following problems concerning $\mathcal{M}$ can be solved in polynomial time as well by using submodular function minimization algorithms [9, 18].

- [membership] check whether a given vector $x \in \mathbb{R}^N$ is contained in $P(\mathcal{M})$ (or $B(\mathcal{M})$) or not,
- [separation] given a vector $x \in \mathbb{R}^N$ not contained in $P(\mathcal{M})$ (or $B(\mathcal{M})$), find a set $X \in 2^N$ such that $x(X) > r_\mathcal{M}(X)$,
- [saturation capacity] for $x \in P(\mathcal{M})$ and $i \in N$, compute the value

$$\hat{c}(x, i) = \max\{\eta \mid \eta \in \mathbb{R},\ x + \eta\chi_i \in P(\mathcal{M})\},$$

• [exchange capacity] for $x \in B(\mathcal{M})$ and $i, j \in N$, compute the value

$$\hat{c}(x, i, j) = \max\{\eta \mid \eta \in \mathbb{R},\ x + \eta(\chi_i - \chi_j) \in B(\mathcal{M})\}.$$

2.2 Pipage Rounding Algorithm

The pipage rounding algorithm [2] for the problem (P) consists of the following three steps:

1. Define a relaxed problem (RP) of the original problem (P).
2. Compute an (approximately) optimal solution x^* of the relaxed problem (RP).
3. Round the fractional vector x^* to obtain a $\{0, 1\}$-vector $\hat{x}$.

We explain the details of each step below.

To define a relaxation (RP) of the problem (P), we use an *extension* $\tilde{f} : [0, 1]^N \to \mathbb{R}$ of f which is a nondecreasing concave function satisfying $\tilde{f}(\chi_X) = f(X)$ $(X \in 2^N)$. For example, the *concave closure*[1] $\overline{f}$ of f given by

$$\overline{f}(x) \quad = \quad \max\left\{ \sum_{X \subseteq N} \lambda_X f(X) \;\middle|\; \sum_{X \subseteq N} \lambda_X = 1,\ \lambda_X \geq 0,\ \sum_{X \subseteq N} \lambda_X \chi_X = x \right\}$$

can be used as an extension of f. If the function f is given as $f(x) = \sum_{k=1}^{m} f_k(x)$ with a family of set functions $f_k : 2^N \to \mathbb{R}$ $(k = 1, 2, \ldots, m)$, then we can also use the sum of the concave closures $\sum_{k=1}^{m} \overline{f}_k(x)$ as an extension of f.

In the second step, we compute an (approximately) optimal solution x^* of the relaxed problem (RP). We may assume that $x^* \in B(\mathcal{M})$, since otherwise we can find $x \in B(\mathcal{M})$ with $\tilde{f}(x) \geq \tilde{f}(x^*)$ by computing the saturation capacity $\hat{c}(x^*, i)$ at most n times.

In the third step, we round the fractional vector $x^* \in B(\mathcal{M})$ to a $\{0, 1\}$-vector χ_X with $X \in \mathcal{B}$ by using a *potential function* $F : [0, 1]^N \to \mathbb{R}$ defined by

$$F(x) = \sum_{X \subseteq N} \left(\prod_{j \in X} x(j) \right) \left(\prod_{j \in N \setminus X} (1 - x(j)) \right) f(X)$$

for $x = (x(1), x(2), \ldots, x(n)) \in [0, 1]^N$. Note that $F(\chi_X) = f(X)$ for any $X \in 2^N$. We assume that the function evaluation oracle for $F(x)$ is available, as in [2]. We note that the function value of F can be evaluated to any desired accuracy in polynomial time by taking sufficiently many independent samples (see [2]).

Rounding of a fractional vector is done by using the following procedure.

[1] The concave closure of f is denoted by f^+ in [2].

Procedure ROUNDING(x)
Input: a vector $x \in B(\mathcal{M})$
Output: a subset $X \in \mathcal{B}$ such that $F(\chi_X) \geq F(x)$
Step 1: If $x \in \{0,1\}^N$, then output the set $X \in 2^N$ with $\chi_X = x$, and stop.
Step 2: Let Y be a minimal tight set (w.r.t. $r_{\mathcal{M}}$) with $|Y \cap \{j \in N \mid 0 < x(j) < 1\}| > 0$.
Step 3: Choose any distinct elements i, i' in $Y \cap \{j \in N \mid 0 < x(j) < 1\}$.
Step 4: Put $x' = x + \hat{c}(x,i,i')(\chi_i - \chi_{i'})$ and $x'' = x + \hat{c}(x,i',i)(\chi_{i'} - \chi_i)$. If $F(x') \geq F(x'')$, then put $x := x'$; otherwise put $x := x''$.
Step 5: Go to Step 1.

Theorem 2.1 ([2]). *The procedure* ROUNDING *terminates in* $\mathrm{O}(n^2)$ *iterations. Given a function evaluation oracle for F and a membership oracle for $B(\mathcal{M})$, the procedure can be implemented to run in polynomial time.*

The correctness of the procedure ROUNDING follows from the following property of F.

Proposition 2.2 ([2]). *For any $x \in B(\mathcal{M})$ and distinct $i, j \in N$, the function $\varphi(\eta) = F(x + \eta(\chi_i - \chi_j))$ is a convex function in the interval $\eta \in [-\hat{c}(x,j,i), \hat{c}(x,i,j)]$.*

The quality of the solution obtained by the procedure ROUNDING depends on the choice of the extension $\tilde{f}$. We denote by OPT the optimal value of the problem (P).

Theorem 2.3 (cf. [2]). *Suppose that $F(y) \geq \alpha \tilde{f}(y)$ holds for all $y \in [0,1]^N$. Given a β-approximate solution $x \in [0,1]^N$ of the problem* (RP)*, the procedure* ROUNDING *outputs a subset $X \in 2^N$ satisfying $f(X) \geq \alpha\beta\,\mathrm{OPT}$.*

The following properties show that if we use the function $\overline{f}(x)$ (or $\sum_{k=1}^m \overline{f}_k(x)$) as an extension of f and we can solve the problem (RP) exactly (i.e., $\beta = 1$ in Theorem 2.3) in polynomial time, then the pipage rounding algorithm is a $(1-1/e)$-approximation algorithm for the problem (P).

Theorem 2.4 ([2]). *For any nondecreasing submodular function $f : 2^N \to \mathbb{R}$ with $f(\emptyset) = 0$, we have $F(x) \geq (1-1/e)\overline{f}(x)$ $(x \in [0,1]^N)$.*

Corollary 2.5 (cf. [2]). *Suppose that the function f is given as $f(X) = \sum_{k=1}^m f_k(X)$ with a family of nondecreasing submodular functions $f_k : 2^N \to \mathbb{R}$ with $f_k(\emptyset) = 0$ $(k = 1, 2, \ldots, m)$. Define $\tilde{f} : [0,1]^N \to \mathbb{R}$ by $\tilde{f}(x) = \sum_{k=1}^m \overline{f}_k(x)$. Then, we have $F(x) \geq (1-1/e)\tilde{f}(x)$.*

2.3 M$^\natural$-concave Functions

We review the definition of M$^\natural$-concavity and show some fundamental properties.

A function $h : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ defined over the integer lattice is said to be *M$^\natural$-concave* if it satisfies the following property:

$$\forall x, y \in \mathrm{dom}_{\mathbb{Z}} h,\ \forall i \in \mathrm{supp}^+(x-y),\ \exists j \in \mathrm{supp}^-(x-y) \cup \{0\}:$$

$$h(x) + h(y) \le h(x - \chi_i + \chi_j) + h(y + \chi_i - \chi_j),$$

where $\mathrm{dom}_{\mathbb{Z}} h = \{x \in \mathbb{Z}^N \mid h(x) > -\infty\}$, for a vector $x \in \mathbb{R}^N$ we define

$$\mathrm{supp}^+(x) = \{i \in N \mid x(i) > 0\}, \qquad \mathrm{supp}^-(x) = \{i \in N \mid x(i) < 0\},$$

and $\chi_0 = \mathbf{0} \in \mathbb{R}^N$. We note that for any M$^\natural$-concave function h and any $p \in \mathbb{R}^N$, the function $h(x) + p^\top x$ is also M$^\natural$-concave in x.

The following property shows that M$^\natural$-concave functions constitute a subclass of submodular functions. For any vectors $x, y \in \mathbb{R}^N$ we define $x \vee y, x \wedge y \in \mathbb{R}^N$ by $(x \vee y)(i) = \max\{x(i), y(i)\}$ and $(x \wedge y)(i) = \min\{x(i), y(i)\}$ for $i \in N$.

Theorem 2.6 ([14]). *An M$^\natural$-concave function $h : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ is a submodular function, i.e., $h(x) + h(y) \ge h(x \vee y) + h(x \wedge y)$ holds for any $x, y \in \mathrm{dom}_{\mathbb{Z}} h$.*

M$^\natural$-concavity for set functions can be naturally defined through the one-to-one correspondence between set functions $f : 2^N \to \mathbb{R}$ and functions $h : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ with $\mathrm{dom}_{\mathbb{Z}} h = \{0,1\}^N$. That is, a set function $f : 2^N \to \mathbb{R} \cup \{-\infty\}$ is said to be *M$^\natural$-concave* if f satisfies the following property:

$\forall X, Y \in 2^N$ with $f(X) > -\infty$, $f(Y) > -\infty$, $\forall i \in X \setminus Y$, it holds that

$$f(X) + f(Y) \le \max\Big[f(X \setminus \{i\}) + f(Y \cup \{i\}), \max_{j \in Y \setminus X} \Big\{ f((X \setminus \{i\}) \cup \{j\}) + f((Y \cup \{i\}) \setminus \{j\}) \Big\} \Big].$$

Maximization of an M$^\natural$-concave set function can be done efficiently.

Theorem 2.7 (cf. [14, 19]). *Let $f : 2^N \to \mathbb{R}$ be an M$^\natural$-concave set function. Then, a maximizer of f can be obtained by evaluating the function value of f at most n^2 times.*

We give a proof of Theorem 1.2 stating that any weighted rank function is an M$^\natural$-concave set function. As shown in Theorem 2.6, the class of M$^\natural$-concave set functions is (properly) contained in the class of submodular set functions.

Proof of Theorem 1.2. Let $f : 2^N \to \mathbb{R}$ be a weighted rank function represented as (1) with a matroid $\mathcal{M}' = (N, \mathcal{F}')$ and a nonnegative vector $w \in \mathbb{R}^N_+$. Define $f_1, f_2 : 2^N \to \mathbb{R} \cup \{-\infty\}$ by

$$f_1(X) = \begin{cases} w(X) & (X \in \mathcal{F}'), \\ -\infty & (\text{otherwise}), \end{cases} \qquad f_2(X) = 0 \quad (X \in 2^N).$$

Then, both f_1 and f_2 are M$^\natural$-concave functions. Moreover, we have

$$f(X) = \max\{f_1(Y) + f_2(X \setminus Y) \mid Y \in 2^N,\ Y \subseteq X\} \qquad (X \in 2^N),$$

which implies that f is M$^\natural$-concave as well (see [14, Theorem 6.13]). □

We give some other examples of M$^\natural$-concave set functions.

Example 2.8 (laminar concave function). Let $\mathcal{F} \subseteq 2^N$ be a laminar family, i.e., for any $X, Y \in \mathcal{F}$ we have $X \setminus Y = \emptyset$, $Y \setminus X = \emptyset$, or $X \cap Y = \emptyset$. For a family of univariate concave functions $\varphi_Y : \mathbb{Z} \to \mathbb{R}$ $(Y \in \mathcal{F})$, the function $f : 2^N \to \mathbb{R}$ defined by

$$f(X) = \sum_{Y \in \mathcal{F}} \varphi_Y(|X \cap Y|) \qquad (X \in 2^N)$$

is an M$^\natural$-concave function. In particular, f is nondecreasing if φ_Y is nondecreasing for all $Y \in \mathcal{F}$.

Example 2.9. Let $G = (U, V; E)$ be a complete bipartite graph with vertex set $U \cup V$ and edge set E, and let $w_e \in \mathbb{R}_+$ be the weight of edge $e \in E$. We define a function $f : 2^U \to \mathbb{R}$ by

$$f(X) = \max\{\sum_{e \in F} w_e \mid F : \text{matching of } G,\ \{\partial^+ e \mid e \in F\} = X\},$$

where $\partial^+ e \in U$ denotes the end vertex of the edge $e \in E$ contained in U. Then, f is a nondecreasing M$^\natural$-concave function.

We also consider M$^\natural$-concavity for polyhedral concave functions. A polyhedral concave function $h : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty\}$ is said to be *M$^\natural$-concave* if it satisfies the following property:

> $\forall x, y \in \mathrm{dom}_{\mathbb{R}} h$, $\forall i \in \mathrm{supp}^+(x - y)$, $\exists j \in \mathrm{supp}^-(x - y) \cup \{0\}$, $\exists \eta_0 > 0$:
>
> $$h(x) + h(y) \le h(x - \eta(\chi_i - \chi_j)) + h(y + \eta(\chi_i - \chi_j)) \qquad (\forall \eta \in [0, \eta_0]),$$

where $\mathrm{dom}_{\mathbb{R}} h = \{x \in \mathbb{R}^N \mid h(x) > -\infty\}$.

Theorem 2.10 ([14, 16]). *For any M$^\natural$-concave function $h : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ with bounded $\mathrm{dom}_{\mathbb{Z}} h$, its concave closure $\overline{h} : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty\}$ is a polyhedral M$^\natural$-concave function.*

A nonempty set $S \subseteq \mathbb{R}^N$ is called a *g-polymatroid* [5] if there exists a pair of a submodular set function $\rho : 2^N \to \mathbb{R} \cup \{+\infty\}$ and a supermodular set function $\mu : 2^N \to \mathbb{R} \cup \{-\infty\}$ such that $\rho(\emptyset) = \mu(\emptyset) = 0$, $\rho(X) - \rho(X \setminus Y) \ge \mu(Y) - \mu(Y \setminus X)$ $(X, Y \subseteq N)$, and

$$S = \{x \in \mathbb{R}^N \mid \mu(X) \le x(X) \le \rho(X)\ (X \in 2^N)\}.$$

Theorem 2.11 ([14, 16]). *Let $h : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ be an M$^\natural$-concave function over the integer lattice, and let $\overline{h} : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty\}$ be its concave closure. For any $p \in \mathbb{R}^N$, the set $\arg\max\{\overline{h}(x) - p^\top x \mid x \in \mathbb{R}^N\}$ is an integral g-polymatroid if it is not empty.*

Finally, we explain the concept of L$^\natural$-concavity, which is deeply related to the concept of M$^\natural$-concavity. A polyhedral concave function $g : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty\}$ is said to be *L$^\natural$-concave* if it satisfies the following property:

$$g(p)+g(q) \leq g((p-\lambda \mathbf{1}) \vee q)+g(p \wedge (q+\lambda \mathbf{1})) \quad (\forall p, q \in \mathbb{R}^N,\ \forall \lambda \in \mathbb{R}_+), \quad (3)$$

where $\mathbf{1} \in \mathbb{Z}^N$ is the vector with all components equal to one.

Theorem 2.12 ([14, 16]). *Let $h : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ be an M$^\natural$-concave function with bounded $\mathrm{dom}_{\mathbb{Z}} h$, and define a function $h^\circ : \mathbb{R}^N \to \mathbb{R}$ by $h^\circ(p) = \min\{p^\top x - h(x) \mid x \in \mathbb{Z}^N\}$ $(p \in \mathbb{R}^N)$. Then, h° is a polyhedral L$^\natural$-concave function.*

L$^\natural$-concavity is also defined for functions over the integer lattice. A function $g : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ is said to be *L$^\natural$-concave* if it satisfies (3), where $p, q \in \mathbb{Z}^N$ and $\lambda \in \mathbb{Z}_+$. The maximization of an L$^\natural$-concave function over the integer lattice can be solved efficiently.

Theorem 2.13 ([14]). *Let $g : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ be an L$^\natural$-concave function with bounded $\mathrm{dom}_{\mathbb{Z}} g$. Then, a maximizer of g can be computed in time polynomial in n and Ψ, where $\Psi = \max_{i \in N}[\log \max\{p(i) - q(i) \mid p, q \in \mathrm{dom}_{\mathbb{Z}} g\}]$.*

3 Approximation Algorithms for Concave Closure

For a concave function $h : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty\}$, a vector $p \in \mathbb{R}^N$ is called a *subgradient* of h at x if p satisfies $h(y) - h(x) \leq p^\top (y - x)$ $(y \in \mathbb{R}^N)$, and the set of subgradients of h at x is denoted by $\partial h(x)$ $(\subseteq \mathbb{R}^N)$. In this section, we show that for any nondecreasing M$^\natural$-concave set function f, an approximate subgradient of the concave closure $\overline{f}$ of f can be computed efficiently. Recall the definition of Γ in (2).

Theorem 3.1. *Let $f : 2^N \to \mathbb{R}$ be a nondecreasing M$^\natural$-concave function with $f(\emptyset) = 0$.*
(i) *For any $\delta > 0$, we can compute a vector $p \in \mathbb{R}^N$ and a real number $\alpha \in \mathbb{R}$ satisfying*

$$\overline{f}(y) - \overline{f}(x) \leq p^\top (y - x) + \delta, \qquad \overline{f}(x) \leq \alpha \leq \overline{f}(x) + \delta \quad (4)$$

in time polynomial in n, Γ, and $\log(1/\delta)$.
(ii) *Suppose f is an integer-valued function. Then, we can compute a subgradient $p \in \partial \overline{f}(x) \cap \mathbb{Z}^N$ and the exact value of $\overline{f}(x)$ in time polynomial in n and Γ.*

In the following, we give a sketch of the proof of Theorem 3.1.

Define a vector $u \in \mathbb{R}^N$ by $u(i) = f(\{i\})$ $(i \in N)$. We first show that for any $x \in [0,1]^N$ there exists a subgradient p of $\overline{f}$ at x such that $\mathbf{0} \leq p \leq u$.

Lemma 3.2. *For any $x, y \in [0,1]^N$ with $x \geq y$ and for any $i \in \mathrm{supp}^+(x-y)$, there exists $\eta_0 > 0$ such that $\overline{f}(x) + \overline{f}(y) \leq \overline{f}(x - \eta\chi_i) + \overline{f}(y + \eta\chi_i)$ $(\forall \eta \in [0, \eta_0])$.*

Lemma 3.3. *For any $x \in \{0,1\}^N$ and $i \in N$ with $x(i) = 0$, we have $\overline{f}(x + \eta\chi_i) - \overline{f}(x) = \eta\{\overline{f}(x + \chi_i) - \overline{f}(x)\}$ $(\forall \eta \in [0,1])$.*

Lemma 3.4. *For any $x \in [0,1]^N$, there exists a subgradient $p \in \partial\overline{f}(x)$ such that $\mathbf{0} \leq p \leq u$.*

Proof. Let $x \in [0,1]^N$. Then, there exists a subgradient $p \in \partial\overline{f}(x)$ such that the set

$$S = \{y \in [0,1]^N \mid \overline{f}(y) - \overline{f}(x) = p^\top(y - x)\}$$

satisfies $|S \cap \{0,1\}^N| \geq n + 1$. We show that such a subgradient p satisfies $\mathbf{0} \leq p \leq u$.

We note that S is a full-dimensional polytope. Let $x_0 \in [0,1]^N$ be a vector in the interior of S. Then, there exists $\varepsilon > 0$ such that

$$\varepsilon' p(i) = \overline{f}(x_0 + \varepsilon'\chi_i) - \overline{f}(x_0) = \overline{f}(x_0) - \overline{f}(x_0 - \varepsilon'\chi_i) \qquad (\forall i \in N,\ 0 \leq \forall\varepsilon' \leq \varepsilon). \tag{5}$$

Since $\mathbf{0} < x_0 < \mathbf{1}$, Lemma 3.2 implies that

$$\overline{f}(\varepsilon'\chi_i) - \overline{f}(\mathbf{0}) \geq \overline{f}(x_0) - \overline{f}(x_0 - \varepsilon'\chi_i) \qquad (i \in N), \tag{6}$$

$$\overline{f}(x_0 + \varepsilon'\chi_i) - \overline{f}(x_0) \geq \overline{f}(\mathbf{1}) - \overline{f}(\mathbf{1} - \varepsilon'\chi_i) \qquad (i \in N) \tag{7}$$

for a sufficiently small $\varepsilon' > 0$. By Lemma 3.3, we have

$$\begin{aligned} \overline{f}(\varepsilon'\chi_i) - \overline{f}(\mathbf{0}) &= \varepsilon'\{\overline{f}(\chi_i) - \overline{f}(\mathbf{0})\} \\ &= \varepsilon'\{f(\{i\}) - f(\emptyset)\} = \varepsilon' u(i) \qquad (i \in N), \quad (8) \\ \overline{f}(\mathbf{1}) - \overline{f}(\mathbf{1} - \varepsilon'\chi_i) &= \varepsilon'\{\overline{f}(\mathbf{1}) - \overline{f}(\mathbf{1} - \chi_i)\} \\ &= \varepsilon'\{f(N) - f(N \setminus \{i\})\} \geq 0 \quad (i \in N), \quad (9) \end{aligned}$$

where the last inequality in (9) is due to the monotonicity of f. Combining (5), (6), (7), (8), and (9), we obtain $\mathbf{0} \leq p \leq u$. □

By the definition of the concave closure and LP duality, we have

$$\begin{aligned} \overline{f}(x) &= \min\{p^\top x + \gamma \mid p^\top \chi_X + \gamma \geq f(X)\ (X \in 2^N),\ p \in \mathbb{R}^N,\ \gamma \in \mathbb{R}\} \\ &= \min\{p^\top x - f^\circ(p) \mid p \in \mathbb{R}^N\} \qquad (x \in [0,1]^N), \end{aligned}$$

where $f^\circ(p) = \min\{p^\top \chi_X - f(X) \mid X \in 2^N\}$. We also define $g(p) = f^\circ(p) - p^\top x$ $(p \in \mathbb{R}^N)$. The following property, together with Lemma 3.4, implies that finding a subgradient of $\overline{f}$ can be reduced to the maximization of the function g over $p \in [\mathbf{0}, u]$.

Theorem 3.5 ([14, 16]). *For any $x \in [0,1]^N$, it holds that $\partial\overline{f}(x) = \arg\max\{g(p) \mid p \in \mathbb{R}^N\}$. If f is an integer-valued function, then $\partial\overline{f}(x)$ is an integral polyhedron.*

Let $p^* \in \mathbb{R}^N$ be a maximizer of g, and $p \in \mathbb{R}^N$ be any vector with $||p - p^*||_\infty \leq \delta/n$. We note that $\overline{f}(x) = -g(p^*)$. Since $p^* \in \partial\overline{f}(x)$ by Theorem 3.5, we have

$$\overline{f}(y) - \overline{f}(x) \leq (p^*)^\top(y-x) = p^\top(y-x) + (p^* - p)^\top(y-x) \leq p^\top(y-x) + \delta.$$

Since the function g is written as $g(p) = \min\{p^\top(\chi_X - x) - f(X) \mid X \in 2^N\}$, we have $g(p^*) - \delta \leq g(p) \leq g(p^*)$. This shows that a vector $p \in \mathbb{R}^N$ with $||p - p^*||_\infty \leq \delta/n$ and the real number $\alpha = -g(p)$ satisfy the condition (4). To prove Theorem 3.1 (i), it suffices to show that such p can be computed in polynomial time.

The function g is a polyhedral L$^\natural$-concave function by Theorem 2.12, and the function value of g can be computed in polynomial time by Theorem 2.7. Let $\delta' = \delta/n^2$, and define a function $g_\mathbb{Z} : \mathbb{Z}^N \to \mathbb{R} \cup \{-\infty\}$ by

$$g_\mathbb{Z}(p) = \begin{cases} g(\delta' p) & (\text{if } p \in \mathbb{Z}^N \text{ and } p(i) \in [0, u(i)/\delta'] \text{ for all } i \in N), \\ -\infty & (\text{otherwise}). \end{cases}$$

Then, $g_\mathbb{Z}$ is an L$^\natural$-concave function over the integer lattice. The L-proximity theorem in [14, Theorem 7.18] implies the following property, which states that any maximizer of $g_\mathbb{Z}$ is sufficiently close to a maximizer of g.

Theorem 3.6 ([11]). *Let $p_\mathbb{Z} \in \mathbb{Z}^N$ be a maximizer of $g_\mathbb{Z}$. Then, there exists a maximizer p^* of g such that $||p^* - \delta' p_\mathbb{Z}||_\infty \leq n\delta' = \delta/n$.*

By Theorem 2.13, a maximizer of $g_\mathbb{Z}$ can be computed in time polynomial in n and $\max_{i \in N} \log(u(i)/\delta')$. This concludes the proof of Theorem 3.1 (i). We note that the algorithms in [14, Section 10.3] cannot be used for computing a maximizer of g exactly since the function g is a polyhedral L$^\natural$-concave function and its maximizer can be an irrational vector.

In the case where f is integer-valued, Lemma 3.4 and Theorem 3.5 imply that an optimal solution of the problem $\max\{g(p) \mid p \in \mathbb{Z}^N,\ p \in [\mathbf{0}, u]\}$ is a subgradient of $\overline{f}$ at x, and such an optimal solution can be obtained in polynomial time. Hence, Theorem 3.1 (ii) is proved.

4 Solving the Relaxed Problem

We prove Theorem 1.3 by giving a polynomial-time algorithm for the relaxed problem (RP).

We first prove Theorem 1.3 (i). Let α^* be the optimal value of the problem (RP), i.e., $\alpha^* = \max\{\tilde{f}(x) \mid x \in P(\mathcal{M})\}$. It suffices to show that for any $\varepsilon > 0$, we can find a vector $x \in P(\mathcal{M})$ with $\tilde{f}(x) \geq \alpha^* - \varepsilon$ in time polynomial in n, Γ, and $\log(1/\varepsilon)$. If we put $\varepsilon = \varepsilon' f(X)$ for $\varepsilon' > 0$ and an arbitrarily chosen $X \in \mathcal{F} \setminus \{\emptyset\}$ then we can obtain a $(1-\varepsilon')$-approximate solution of (RP) since $\tilde{f}(x)/\alpha^* \leq \tilde{f}(x)/f(X) \leq 1 - \varepsilon'$.

For any $\underline{\alpha} \in \mathbb{R}$, we define a set

$$\mathrm{L}(\underline{\alpha}) = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid \underline{\alpha} \leq \alpha \leq \tilde{f}(x)\}.$$

Note that $\mathrm{L}(\underline{\alpha}) \neq \emptyset$ if and only if $\underline{\alpha} \leq \alpha^*$. Given $\underline{\alpha}$, our algorithm described below either asserts $\mathrm{L}(\underline{\alpha}) = \emptyset$ or finds a point (x, α) such that $\underline{\alpha} \leq \alpha \leq \tilde{f}(x) + (\varepsilon/2)$. By combining this algorithm with binary search w.r.t. $\underline{\alpha}$, we can find an approximate solution x of (RP) satisfying $\tilde{f}(x) \geq \alpha^* - \varepsilon$.

Our algorithm for checking the nonemptyness of $\mathrm{L}(\underline{\alpha})$ is based on the ellipsoid method. Let $\delta > 0$ be a constant. In each iteration of the algorithm, we check whether the set $\mathrm{L}(\underline{\alpha})$ approximately contains the point (x_c, α_c) which is the center of the current ellipsoid, and computes a hyperplane which almost separates the point (x_c, α_c) and the set $\mathrm{L}(\underline{\alpha})$ in the following way:

Case 1: If $\alpha_c < \underline{\alpha}$, then we output $\alpha \geq \underline{\alpha}$ as a separating hyperplane.
Case 2: If $x_c \notin P(\mathcal{M})$, we compute a separating hyperplane for $P(\mathcal{M})$ and x_c and output it.
Case 3: Suppose that $\alpha_c \geq \underline{\alpha}$ and $x_c \in P(\mathcal{M})$. For each $k = 1, 2, \ldots, m$, we compute a real number β_k satisfying $\overline{f}_k(x_c) \leq \beta_k \leq \overline{f}_k(x_c) + \delta$ (see Theorem 3.1 (i)) and put $\beta = \sum_{k=1}^{m} \beta_k$.

Case 3-1: If $\alpha_c \leq \beta$, then we output the point (x_c, α_c) and stop.

Case 3-2: Suppose that $\alpha_c > \beta$. For each $k = 1, 2, \ldots, m$, we compute a vector $p_k \in \mathbb{R}^N$ satisfying $\overline{f}_k(x) - \overline{f}(x_c) \leq p_k^\top (x - x_c) + \delta$ for all $x \in [0,1]^N$ (see Theorem 3.1 (i)), and put $p = \sum_{k=1}^{m} p_k$. We output $\alpha - \beta \leq p^\top (x - x_c) + 2m\delta$ as a separating hyperplane. □

We note that the separating hyperplane obtained in Case 3-2 is satisfied by all $(x, \alpha) \in \mathrm{L}(\underline{\alpha})$. After a finite number of iterations, we can find a point (x_c, α_c) with $\underline{\alpha} \leq \alpha \leq \overline{f}(x) + m\delta$ or assert that the set $\mathrm{L}(\underline{\alpha})$ is empty, and the number of iteration is bounded by n, m, Γ, and $\log(1/\delta)$ (see, e.g., [8]). Therefore, we obtain a desired algorithm for $\mathrm{L}(\underline{\alpha})$ by putting $\delta = \varepsilon/2m$.

We then prove Theorem 1.3 (ii). When each f_k is integer-valued, we use the ellipsoid method in a different way, and apply it to find a vector in the set $S^* = \arg\max\{\tilde{f}(x) \mid x \in B(\mathcal{M})\}$. For the correctness and polynomial-time termination of the ellipsoid method, it suffices to prove the following (see [8]):

(a) S^* is a rational polytope such that the encoding length of each facet is bounded by a polynomial in the input size,
(b) a separating hyperplane for the set S^* and a given point $x \in [0,1]^N$ can be computed in time polynomial in the input size.

For any $k = 1, 2, \ldots, m$ and any $p \in \mathbb{R}^N$, the set $\arg\max\{\overline{f}_k(x) - p^\top x \mid x \in [0,1]^N\}$ is an integral g-polymatroid by Theorem 2.11. Hence, S^* is given as the intersection of m integral g-polymatroids and a base polytope of a matroid. Therefore, S^* is represented by the inequalities of the form $x(X) \leq \gamma_X$ or $x(X) \geq \gamma_X$ with $X \in 2^N$ and an integer $\gamma_X \in \{0, 1, \ldots, n\}$. This fact shows that S^* is a rational polytope such that the encoding length of each facet is bounded by a polynomial in n.

We then explain how to compute a separating hyperplane for S^* and x. We first check whether $x \in B(\mathcal{M})$ or not. If $x \notin B(\mathcal{M})$, then we compute a separating hyperplane for $B(\mathcal{M})$ and x, and output it. If $x \in B(\mathcal{M})$, then we compute a subgradient $p_k \in \partial\overline{f}_k(x)$ in polynomial time, as shown in Theorem 3.1. Since $\tilde{f} = \sum_{k=1}^m \overline{f}_k$, the vector $p = \sum_{k=1}^m p_k$ is a subgradient of $\tilde{f}$ at x. Therefore, we have $0 \leq \tilde{f}(x^*) - \tilde{f}(x) \leq p^\top(x^* - x)$ for any $x^* \in S^*$, i.e., $p^\top x^* \leq p^\top x$ is a separating hyperplane for S^* and x. This concludes the proof of Theorem 1.3 (ii).

Remark Our result in this paper can be extended to functions defined over the integer lattice. Let $v \in \mathbb{Z}_+^N$, and consider a function $h : [\mathbf{0}, v] \cap \mathbb{Z}^N \to \mathbb{R}$. If h is a submodular function given as the sum of nondecreasing M$^\natural$-concave functions $h_k : [\mathbf{0}, v] \cap \mathbb{Z}^N \to \mathbb{R}$ $(k = 1, 2, \ldots, m)$ with $h_k(\mathbf{0}) = 0$, then we can compute a $(1 - 1/e)$-approximate solution of the problem $\max\{h(x) \mid x \in B \cap \mathbb{Z}^N\}$ in polynomial time, where B is an integral base polytope (of a submodular system) such that $B \subseteq [\mathbf{0}, v]$. We omit the details.

Acknowledgement

The author thanks Kazuo Murota for stimulating discussions and valuable comments on the manuscript. The author also thanks Satoru Iwata for helpful suggestions. This work is supported by Grant-in-Aid of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] A. Ageev and M. Sviridenko. Pipage rounding: a new method of constructing algorithm with proved performance guarantee. *J. Combinatorial Optimization*, 8:307–328, 2004.

[2] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák. Maximizing a submodular set function subject to a matroid constraint. *Proc. of IPCO 2007*, Springer LNCS 4513, 182–196, 2007.

[3] M. Conforti and G. Cornuéjols. Submodular set functions, matroids and the greedy algorithm: tight worst-case bounds and some generalizations

of the Rado–Edmonds theorem. *Discrete Applied Mathematics*, 7:251–274, 1984.

[4] M. L. Fisher, G.L. Nemhauser, and L. A. Wolsey. An analysis of approximations for maximizing submodular set functions II. *Mathematical Programming Study*, 8:73–87, 1978.

[5] A. Frank and É. Tardos. Generalized polymatroids and submodular flows. *Mathematical Programming*, 42:489–563, 1988.

[6] S. Fujishige. *Submodular Functions and Optimization, Second Edition*, Elsevier, 2005.

[7] P. R. Goundan and A. S. Schulz. Revisiting the greedy approach to submodular set function maximization. Working Paper, Massachusetts Institute of Technology, 2007.

[8] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization, Second Corrected Edition*, Springer, 1993.

[9] S. Iwata, L. Fleischer, and S. Fujishige. A combinatorial, strongly polynomial-time algorithm for minimizing submodular functions. *J. ACM*, 48:761–777, 2001.

[10] L. Lovász. Submodular functions and convexity. In A. Bachem, M. Grötschel, and B. Korte, eds., *Mathematical Programming: The State of the Art*, Springer, 234–257, 1983.

[11] S. Moriguchi and N. Tsuchimura. Discrete $\mathrm{L}^\natural$-/$\mathrm{M}^\natural$-convex function minimization based on continuous relaxation. Technical Report, Department of Mathematical Informatics, University of Tokyo, 2007.

[12] K. Murota. Valuated matroid intersection, I: optimality criteria, II: algorithms. *SIAM J. Discrete Mathematics*, 9:545-561, 562-576, 1996.

[13] K. Murota. Convexity and Steinitz's exchange property. *Advances in Mathematics*, 124:272-311, 1996.

[14] K. Murota. *Discrete Convex Analysis*, SIAM, 2003.

[15] K. Murota and A. Shioura. M-convex function on generalized polymatroid. *Mathematics of Operations Research*, 24:95–105, 1999.

[16] K. Murota and A. Shioura. Extension of M-convexity and L-convexity to polyhedral convex functions. *Advances in Applied Mathematics*, 25:352–427, 2000.

[17] G.L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions I. *Mathematical Programming*, 14:265–294, 1978.

[18] A. Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *J. Combinatorial Theory* (B), 80:346–355, 2000.

[19] A. Shioura. Fast scaling algorithms for M-convex function minimization with application to the resource allocation problem. *Discrete Applied Mathematics*, 134:303–316, 2004.

[20] M. Sviridenko. A note on maximizing a submodular set function subject to a knapsack constraint. *Operations Research Letters*, 32:41–43, 2004.