MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

# M-convex Function Minimization by Continuous Relaxation Approach —Proximity Theorem and Algorithm—

Satoko MORIGUCHI, Akiyoshi SHIOURA, and
Nobuyuki TSUCHIMURA

(Communicated by Kazuo MUROTA)

METR 2008–38 September 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**

# M-convex Function Minimization by Continuous Relaxation Approach —Proximity Theorem and Algorithm—

Satoko MORIGUCHI

School of Industrial Technology
Advanced Institute of Industrial Technology
satoko@aiit.ac.jp

Akiyoshi SHIOURA

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

Nobuyuki TSUCHIMURA

Department of Informatics
Kwansei Gakuin University
tutimura@kwansei.ac.jp

September 2008

**Abstract**

The concept of M-convexity for functions in integer variables, introduced by Murota (1995), plays a primary role in the theory of discrete convex analysis. In this paper, we consider the problem of minimizing an M-convex function, which is a natural generalization of the separable convex resource allocation problem under a submodular constraint and contains some classes of nonseparable convex function minimization on integer lattice points. We propose a new approach for M-convex function minimization based on continuous relaxation. We show proximity theorems for M-convex function minimization and its continuous relaxation, and develop a new algorithm based on continuous relaxation by using the proximity theorems. The practical performance of the proposed algorithm is evaluated by computational experiments.

## 1 Introduction

The concept of M-convexity for functions in integer variables, introduced by Murota [11, 12], plays a primary role in the theory of discrete convex analysis [13]. M-convex functions enjoy various nice properties as "discrete convexity" such as a local characterization for global minimality, extensibility to ordinary convex functions, conjugacy, duality, etc. We consider the problem of minimizing an M-convex function, which is fundamental in discrete optimization. For this problem, various approaches have been proposed to develop efficient algorithms [9, 20, 21, 22]. In this paper, we propose a new approach for M-convex function minimization based on continuous relaxation.

**M-convex Function Minimization** Let $n$ be a positive integer and $N = \{1, 2, \ldots, n\}$. A function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ in integer variables is said to be *M-convex* if it satisfies (M-EXC[**Z**]):

**(M-EXC[Z])**
$\forall x, y \in \mathrm{dom}_{\mathbf{Z}}\, g$, $\forall i \in \mathrm{supp}^+(x - y)$, $\exists j \in \mathrm{supp}^-(x - y)$:

$$g(x) + g(y) \geq g(x - \chi_i + \chi_j) + g(y + \chi_i - \chi_j),$$

where the effective domain of $g$ is given by $\mathrm{dom}_{\mathbf{Z}}\, g = \{x \in \mathbf{Z}^n \mid g(x) < +\infty\}$, $\mathrm{supp}^+(x) = \{i \in N \mid x(i) > 0\}$, $\mathrm{supp}^-(x) = \{i \in N \mid x(i) < 0\}$, and $\chi_i \in \{0, 1\}^n$ $(i \in N)$ denotes the characteristic vector of $i \in N$, i.e., $\chi_i(i) = 1$ and $\chi_i(j) = 0$ for $j \in N \setminus \{i\}$. By definition, the effective domain $\mathrm{dom}_{\mathbf{Z}}\, g$ lies on a hyperplane $\{x \in \mathbf{Z}^n \mid x(N) = r\}$ for some integer $r$.

Minimization of an M-convex function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ is formulated as

$$\text{(MC)} \qquad \text{Minimize} \quad g(x) \qquad \text{subject to} \quad x \in \mathrm{dom}_{\mathbf{Z}}\, g.$$

Below we give some important special cases of the problem (MC).

**Example 1.1 (Resource Allocation Problem under a Submodular Constraint).** Let $f_i : \mathbf{R} \to \mathbf{R}$ $(i \in N)$ be a family of univariate convex functions. Also, let $\rho : 2^N \to \mathbf{Z} \cup \{+\infty\}$ be a submodular function, i.e., $\rho$ satisfies $\rho(X) + \rho(Y) \geq \rho(X \cap Y) + \rho(X \cup Y)$ for every $X, Y \in 2^N$. We assume $\rho(\emptyset) = 0$, $\rho(Y) \geq 0$ $(\forall Y \subseteq N)$, and $\rho(N) < +\infty$. The *(separable convex) resource allocation problem under a submodular constraint* [1, 6, 7] is formulated as follows:

$$\text{(SC)} \quad \left| \begin{array}{ll} \text{Minimize} & \displaystyle\sum_{i=1}^{n} f_i(x(i)) \\ \text{subject to} & x(N) = \rho(N),\ x(Y) \leq \rho(Y)\ (Y \in 2^N), \\ & x \geq \mathbf{0},\ x \in \mathbf{Z}^n, \end{array} \right.$$

where $x(Y) = \sum_{i \in Y} x(i)$ for $Y \subseteq N$ and $\mathbf{0} = (0, 0, \ldots, 0) \in \mathbf{Z}^n$. One of the simplest special cases of (SC) is the *simple (separable convex) resource allocation problem* [1, 6, 7]:

$$\text{(SIMPLE)} \left| \begin{array}{ll} \text{Minimize} & \displaystyle\sum_{i=1}^{n} f_i(x(i)) \\ \text{subject to} & x(N) = K, \ \mathbf{0} \leq x \leq u, \ x \in \mathbf{Z}^n, \end{array} \right.$$

where $K \in \mathbf{Z}_+$ and $u \in (\mathbf{Z}_+ \cup \{+\infty\})^n$. See [1, 6, 7] for comprehensive review of (SC) and [2, 3, 4, 5, 10] for efficient algorithms.

The problem (SC) is a special case of (MC) since the function $g_{\mathrm{SC}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ defined by

$$g_{\mathrm{SC}}(x) = \begin{cases} \displaystyle\sum_{i=1}^{n} f_i(x(i)) & (\text{if } x \in \mathbf{Z}^n \text{ satisfies } x(N) = \rho(N), \\ & \qquad x(Y) \leq \rho(Y) \ (Y \in 2^N), \ x \geq \mathbf{0}), \\ +\infty & (\text{otherwise}) \end{cases}$$

satisfies (M-EXC[**Z**]) (see [12, Example 2.2], [13, Section 6.3]). □

**Example 1.2 (Extension of Resource Allocation Problem under a Tree Constraint).** Let $\mathcal{F} \subseteq 2^N$ be a laminar family, i.e., for every $X, Y \in \mathcal{F}$ either of $X \subseteq Y$, $X \supseteq Y$, and $X \cap Y = \emptyset$ holds. The *resource allocation problem under a tree constraint* [6, 7] is formulated as

$$\begin{array}{ll} \text{Minimize} & \displaystyle\sum_{i=1}^{n} f_i(x(i)) \\ \text{subject to} & x(N) = K, \ x(Y) \leq u_X \ (Y \in \mathcal{F}), \\ & x \geq \mathbf{0}, \ x \in \mathbf{Z}^n, \end{array}$$

where $f_i : \mathbf{R} \to \mathbf{R}$ is a univariate convex function for $i \in N$, $K \in \mathbf{Z}_+$, and $u_X \in \mathbf{Z}_+$ for $X \in \mathcal{F}$. We consider an extension of this problem with a nonseparable convex objective function, which we call the *laminar convex resource allocation problem*:

$$\text{(LC)} \left| \begin{array}{ll} \text{Minimize} & \displaystyle\sum_{Y \in \mathcal{F}}^{n} f_Y(x(Y)) \\ \text{subject to} & x(N) = K, \ x(Y) \leq u_X \ (Y \in \mathcal{F}), \\ & x \geq \mathbf{0}, \ x \in \mathbf{Z}^n, \end{array} \right.$$

where $f_X : \mathbf{R} \to \mathbf{R}$ is a univariate convex function for $X \in \mathcal{F}$. The problem (LC) is a special case of (MC) since the function $g_{\mathrm{LC}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ defined by

$$g_{\mathrm{LC}}(x) = \begin{cases} \displaystyle\sum_{Y \in \mathcal{F}} f_Y(x(Y)) & (x \in \mathbf{Z}^n \text{ satisfies } x(N) = K, \\ & \qquad x(Y) \leq u_Y \ (Y \in \mathcal{F}), \ x \geq \mathbf{0}), \\ +\infty & (\text{otherwise}) \end{cases}$$

satisfies (M-EXC[**Z**]) (see [9, Example 2.3], [13, Section 6.3]). □

**Continuous Relaxation** Continuous relaxations of (SC) and (LC) can be naturally obtained by replacing the integrality constraint "$x \in \mathbf{Z}^n$" with "$x \in \mathbf{R}^n$." This motivates us to consider continuous relaxation of the problem (MC). Our continuous relaxation of (MC) is associated with the concept of *M-convex function in real variables*, which is introduced by Murota–Shioura [17] as an extension of M-convex function in integer variables. A function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ in continuous variables is said to be *M-convex* if it is convex and satisfies (M-EXC[$\mathbf{R}$]):

**(M-EXC[R])**
$\forall x, y \in \mathrm{dom}_{\mathbf{R}} f$, $\forall i \in \mathrm{supp}^+(x-y)$, $\exists j \in \mathrm{supp}^-(x-y)$, $\exists \alpha_0 > 0$:

$$f(x)+f(y) \geq f(x-\alpha(\chi_i-\chi_j))+f(y+\alpha(\chi_i-\chi_j)) \quad (\forall \alpha \in [0, \alpha_0]),$$

where $\mathrm{dom}_{\mathbf{R}} f = \{x \in \mathbf{R}^n \mid f(x) < +\infty\}$. M-convex functions in continuous variables constitute a subclass of convex functions with additional combinatorial properties such as supermodularity and local polyhedral structure (see, e.g., [13, 16, 17, 18, 19]). Fundamental properties of M-convex functions are investigated in [18], such as equivalent axioms, subgradients, directional derivatives, etc.

The following relationship holds between the two kinds of M-convex functions. An M-convex function is said to be *closed proper M-convex* if it is closed proper convex, in addition (the definition of closed proper convex functions is given at the end of this section).

**Theorem 1.3 (cf. [13, Section 6.11]).** *For any M-convex function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ in integer variables, there exists some closed proper M-convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ in continuous variables such that $f(x) = g(x)$ $(x \in \mathbf{Z}^n)$.*

Based on this fact, we consider in this paper the following continuous relaxation of (MC):

$$(\overline{\mathrm{MC}}) \text{ Minimize} \quad f(x) \qquad \text{subject to} \quad x \in \mathrm{dom}_{\mathbf{R}} f,$$

where $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is a closed proper M-convex function satisfying the condition $f(x) = g(x)$ $(x \in \mathbf{Z}^n)$. We note that continuous relaxations of (SC) and (LC) can be also formulated in the form ($\overline{\mathrm{MC}}$) by using functions $f_{\mathrm{SC}} : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ and $f_{\mathrm{LC}} : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ which are defined in a similar way as $g_{\mathrm{SC}}$ and $g_{\mathrm{LC}}$ in Examples 1.1 and 1.2, where "$x \in \mathbf{Z}^n$" is replaced with "$x \in \mathbf{R}^n$." It should be mentioned that $f_{\mathrm{SC}}$ and $f_{\mathrm{LC}}$ are M-convex functions in continuous variables satisfying $f_{\mathrm{SC}}(x) = g_{\mathrm{SC}}(x)$ and $f_{\mathrm{LC}}(x) = g_{\mathrm{LC}}(x)$ for $x \in \mathbf{Z}^n$.

**Our Results** An optimal solution of an optimization problem is expected to be close to an optimal solution of its continuous relaxation. Therefore, an optimal solution of the continuous relaxation can be used as a good initial solution of algorithms for the optimization problem. Efficiency of algorithms based on continuous relaxation depends on the distance between optimal solutions of the optimization problem and its continuous relaxation, and so-called "proximity theorem" provides a theoretical guarantee for the closeness of these two kinds of optimal solutions. For example, an algorithm based on continuous relaxation is proposed for (SIMPLE), where a proximity theorem in terms of the $L_1$ distance is used (see [6, Section 4.6]; see also Remark 5.1). Continuous relaxation is also used in [5] to show the strongly polynomial-time solvability of several special cases of (SC) with quadratic objective function (see Remark 5.3).

The main result in this paper is to show proximity theorems for the problem (MC), stating that the $L_\infty$ distance between optimal solutions of (MC) and its continuous relaxation is bounded by $n-1$ (see Corollary 2.2 and Theorem 2.3). Since the problem (SC) is a special case of (MC), the bound $n-1$ also applies to (SC), which slightly improves the previous bound $n$ for (SC) shown in [4] (see Remark 5.3). We also give an example to show that the bound $n-1$ is the best possible, even for the special case of (SIMPLE).

We then apply the proximity theorems to develop an efficient algorithm for (MC). It is known that (MC) can be solved by a greedy-type algorithm in pseudo-polynomial time [13, 14]. We propose a new algorithm by combining the greedy-type algorithm with continuous relaxation. It is shown by using the proximity theorems that our algorithm terminates in $O(n^2)$ iterations. Therefore, our algorithm can be faster than the existing polynomial-time algorithms [20, 21, 22] if continuous relaxation can be solved quickly.

To evaluate the practical performance of our algorithm, we implement our algorithm and some existing algorithms and perform computational experiments with randomly generated instances of (MC). It is observed from the experimental results that our algorithm is much faster than the existing algorithms for the tested instances.

The organization of this paper is as follows. Proximity theorems for (MC) are presented in Section 2, while the proofs are given later in Section 4. In Section 3, we apply the proximity theorems to develop an efficient algorithm for (MC), and show the results of computational experiments. Finally, we give some concluding remarks in Section 5.

**Definitions and Notation** We denote by $\mathbf{R}_+$ (resp., by $\mathbf{Z}_+$) the sets of nonnegative real numbers (resp., nonnegative integers). Inequalities and equalities for vectors $x, y \in \mathbf{R}^n$ mean component-wise inequalities and equalities; for example, $x \leq y$ reads $x(i) \leq y(i)$ $(i \in N)$. We also define

$$||x||_\infty = \max_{i \in N} |x(i)|, \qquad ||x||_1 = \sum_{i \in N} |x(i)|.$$

Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a function. A function $f$ is said to be *convex* if its epigraph $\{(x, \alpha) \in \mathbf{R}^n \times \mathbf{R} \mid \alpha \geq f(x)\}$ is a convex set. A convex function $f$ is said to be *proper* if the effective domain $\mathrm{dom}_{\mathbf{R}} f$ is nonempty, and *closed* if its epigraph is a closed set.

## 2 Proximity Theorems

We show proximity theorems for M-convex function minimization (MC) and its continuous relaxation. More precisely, we mainly consider the following problem instead of (MC):

$$\text{Minimize} \quad f(x) \qquad \text{subject to} \quad x \in \mathrm{dom}_{\mathbf{R}} f \cap \mathbf{Z}^n, \tag{2.1}$$

where $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is a closed proper M-convex function. We see from Theorem 1.3 that this problem is more general than (MC) (see also Remark 2.4).

**Theorem 2.1.** *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper M-convex function. Suppose that $y_* \in \mathrm{dom}_{\mathbf{R}} f$ satisfies*

$$f(y_* - \chi_i + \chi_j) \geq f(y_*) \qquad (\forall i, j \in N). \tag{2.2}$$

*Then,* $\arg\min f \neq \emptyset$ *and there exists some* $x_* \in \arg\min f$ *such that*

$$||x_* - y_*||_\infty < n - 1.$$

*Proof.* Proof is given in Section 4.1. □

For a function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$, we define the *discretization* $f_{\mathbf{Z}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ of $f$ by

$$f_{\mathbf{Z}}(x) = \begin{cases} f(x) & (x \in \mathbf{Z}^n), \\ +\infty & (\text{otherwise}). \end{cases}$$

For the discretization $f_{\mathbf{Z}}$ of $f$, any $y_* \in \arg\min f_{\mathbf{Z}}$ satisfies the condition (2.2). Hence, we obtain the following corollary.

**Corollary 2.2.** *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper M-convex function and $f_{\mathbf{Z}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ the discretization of $f$. For every $y_* \in \arg\min f_{\mathbf{Z}}$, there exists some $x_* \in \arg\min f$ such that*

$$||x_* - y_*||_\infty < n - 1.$$

*In particular,* $\arg\min f_{\mathbf{Z}} \neq \emptyset$ *implies* $\arg\min f \neq \emptyset$.

**Theorem 2.3.** *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper M-convex function, and $f_{\mathbf{Z}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ the discretization of $f$. For every $x_* \in \arg\min f$, there exists some $y_* \in \arg\min f_{\mathbf{Z}}$ such that*

$$||y_* - x_*||_\infty < n - 1.$$

*In particular,* $\arg\min f \neq \emptyset$ *implies* $\arg\min f_{\mathbf{Z}} \neq \emptyset$.

*Proof.* Proof is given in Section 4.2. □

It should be mentioned that the proximity theorems above do not assume (M-EXC[**Z**]) for the discretization $f_{\mathbf{Z}}$.

**Remark 2.4.** As mentioned at the beginning of this section, the problem (2.1) is more general than the problem (MC). To illustrate that the problem (2.1) properly contains (MC), we show an example of M-convex function in continuous variables for which the discretization does not satisfy (M-EXC[**Z**]).

Let $S \subseteq \mathbf{R}^4$ be a set defined by

$$S = \{\sum_{i=1}^{2}\sum_{j=3}^{4} \alpha_{ij}(\chi_i - \chi_j) \mid 0 \leq \alpha_{ij} \leq 1/2 \ (i = 1, 2, \ j = 3, 4)\}.$$

We consider a function $f : \mathbf{R}^4 \to \mathbf{R} \cup \{+\infty\}$ such that $\mathrm{dom}_{\mathbf{R}} f = S$ and $f(x) = 0$ ($\forall x \in \mathrm{dom}_{\mathbf{R}} f$), which is an M-convex function in continuous variables. The discretization $f_{\mathbf{Z}} : \mathbf{Z}^4 \to \mathbf{R} \cup \{+\infty\}$ of $f$ is a function such that $\mathrm{dom}_{\mathbf{Z}} f_{\mathbf{Z}} = S \cap \mathbf{Z}^4 = \{(0, 0, 0, 0), (1, 1, -1, -1)\}$ and $f_{\mathbf{Z}}(x) = 0$ ($\forall x \in \mathrm{dom}_{\mathbf{Z}} f_{\mathbf{Z}}$), which does not satisfy (M-EXC[**Z**]). □

**Remark 2.5.** Minimizers of a closed proper M-convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ can be characterized by the condition $f'(x; i, j) \geq 0$ ($\forall i, j \in N$) (see [13, 18]), where for $x \in \mathrm{dom}_{\mathbf{R}} f$ and $i, j \in N$ the directional derivative $f'(x; i, j)$ is defined by

$$f'(x; i, j) = \lim_{\alpha \downarrow 0} \frac{f(x + \alpha(\chi_i - \chi_j)) - f(x)}{\alpha}.$$

Hence, the statement of Theorem 2.3 can be rewritten as follows (cf. Theorem 2.1):

> *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper M-convex function, and $f_{\mathbf{Z}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ the discretization of $f$. Suppose that $x_* \in \mathrm{dom}_{\mathbf{R}} f$ satisfies $f'(x; i, j) \geq 0$ ($\forall i, j \in N$). Then,* $\arg\min f_{\mathbf{Z}} \neq \emptyset$ *and there exists some $y_* \in \arg\min f_{\mathbf{Z}}$ such that $||y_* - x_*||_\infty < n - 1$.*

□

The following examples show that the bound $n-1$ in Corollary 2.2 and Theorem 2.3 is tight, even for the special case of the simple resource allocation problem (SIMPLE).

**Example 2.6.** For an arbitrarily chosen small positive number $\delta$, we consider the problem (SIMPLE), where $K = n-1$, $u(i) = +\infty$ $(i \in N)$, and convex functions $f_i : \mathbf{R} \to \mathbf{R}$ $(i \in N)$ are given as

$$\begin{aligned} f_1(\alpha) &= \alpha \quad (\alpha \in \mathbf{R}), \\ f_i(\alpha) &= \max\left\{0, \left(1+\frac{1}{\delta}\right)\alpha - \frac{1-\delta^2}{\delta}\right\} \quad (\alpha \in \mathbf{R},\ i = 2,3,\ldots,n). \end{aligned}$$

It is noted that $f_i(\alpha) = 0$ for $\alpha \in [0, 1-\delta]$ and $f_i(1) - f_i(0) = 1 + \delta > 1$ for $i = 2,3,\ldots,n$. Hence, optimal solutions $y_* \in \mathbf{Z}^n$ and $x_* \in \mathbf{R}^n$ of the problem (SIMPLE) and its continuous relaxation, respectively, are uniquely given as

$$y_* = (n-1, 0, \ldots, 0), \qquad x_* = ((n-1)\delta, 1-\delta, \ldots, 1-\delta).$$

It is easy to see that $||y_* - x_*||_\infty = (n-1)(1-\delta)$, which can be arbitrarily close to $n-1$. □

**Example 2.7.** Let $\delta$ be an arbitrarily chosen small positive number and put $\eta = 3\delta(1-\delta) - \delta$. We again consider the problem (SIMPLE), where $K = n-1$, $u(i) = +\infty$ $(i \in N)$, and convex functions $f_i : \mathbf{R} \to \mathbf{R}$ $(i \in N)$ are given as

$$\begin{aligned} f_1(\alpha) &= 2\delta\alpha \quad (\alpha \in \mathbf{R}), \\ f_i(\alpha) &= \max\left\{-\frac{\eta}{\delta}(\alpha-\delta), 3\delta(\alpha-\delta)\right\} \quad (\alpha \in \mathbf{R},\ i = 2,3,\ldots,n). \end{aligned}$$

It is noted that $f_i(1) - f_i(0) = \delta$ for $i = 2,3,\ldots,n$. Hence, optimal solutions $y_* \in \mathbf{Z}^n$ and $x_* \in \mathbf{R}^n$ of (SIMPLE) and its continuous relaxation, respectively, are uniquely given as

$$y_* = (0, 1, \ldots, 1), \qquad x_* = ((n-1)(1-\delta), \delta, \ldots, \delta).$$

It is easy to see that $||y_* - x_*||_\infty = (n-1)(1-\delta)$, which can be arbitrarily close to $n-1$. □

**Remark 2.8.** The following proximity theorem is known for M-convex functions in integer variables (see, e.g., [13, Theorem 6.37]), although it is not useful in proving the proximity theorems in this paper.

**Theorem 2.9.** *Let $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ be an M-convex function in integer variables and $\alpha$ a positive integer. Suppose that $y_\alpha \in \mathrm{dom}_{\mathbf{Z}}\, g$ satisfies*

$$g(y_\alpha - \alpha\chi_i + \alpha\chi_j) \geq g(y_\alpha) \qquad (\forall i, j \in N).$$

*Then,* $\arg\min g \neq \emptyset$ *and there exists some $y_* \in \arg\min g$ such that*

$$||y_* - y_\alpha||_\infty \leq (n-1)(\alpha-1).$$

□

**Remark 2.10.** For any closed proper M-convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ and $\alpha > 0$, we define a function $f_\alpha : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ by

$$f_\alpha(x) = f(\alpha x) \qquad (x \in \mathbf{R}^n).$$

Then, $f_\alpha$ is a closed proper M-convex function as well [13, Theorem 6.49 (2)]. Corollary 2.2 and Theorem 2.3 applied to $f_\alpha$ can be restated in terms of $f$ as follows, which are seemingly more general but equivalent.

**Corollary 2.11.** *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper M-convex function. Also, let $\alpha > 0$ and $f_{\alpha\mathbf{Z}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ be a function defined by*

$$f_{\alpha\mathbf{Z}}(x) = \begin{cases} f(\alpha x) & (x \in \mathbf{Z}^n), \\ +\infty & (\textit{otherwise}), \end{cases}$$

(i) *For every $y_* \in \arg\min f_{\alpha\mathbf{Z}}$, there exists some $x_* \in \arg\min f$ such that*

$$||x_* - \alpha y_*||_\infty < \alpha(n-1).$$

(ii) *For every $x_* \in \arg\min f$, there exists some $y_* \in \arg\min f_{\alpha\mathbf{Z}}$ such that*

$$||\alpha y_* - x_*||_\infty < \alpha(n-1).$$

(iii) *We have $\arg\min f \neq \emptyset$ if and only if $\arg\min f_{\alpha\mathbf{Z}} \neq \emptyset$.*

□

# 3 New Algorithm Based on Continuous Relaxation

In this section, we propose a new algorithm for the problem (MC) using continuous relaxation.

## 3.1 Greedy Algorithm

Our algorithm uses the following greedy-type algorithm called "modified steepest descent algorithm" [9] (also called "greedy algorithm" in [21]) as a subroutine. The main idea of the modified steepest descent algorithm is to iteratively reduce a set containing a minimizer of an M-convex function by using the following property:

**Theorem 3.1 ([20, Theorem 2.2]).** *Let $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ be an M-convex function with $\arg\min g \neq \emptyset$. For $x \in \mathrm{dom}_{\mathbf{Z}}\, g$ and $i \in N$, suppose that $j \in N$ satisfies the condition*

$$g(x + \chi_j - \chi_i) = \min_{h \in N} g(x + \chi_h - \chi_i).$$

*Then, there exists $x_* \in \arg\min g$ such that $x_*(j) \geq x(j) + 1 - \chi_i(j)$.*

Below we describe a slightly modified version of the modified steepest descent algorithm. The vector $\ell \in (\mathbf{R} \cup \{-\infty\})^n$ is used to represent a set $\{x \in \mathbf{Z}^n \mid x \geq \ell\}$ containing a minimizer. We assume that an initial vector $x^\circ \in \mathrm{dom}_{\mathbf{Z}}\, g$ and a bound $L \in \mathbf{Z}_+$ satisfying

$$\arg\min g \cap \{x \in \mathrm{dom}_{\mathbf{Z}}\, g \mid ||x - x^\circ||_\infty \leq L\} \neq \emptyset$$

are given in advance. For example, we can use $L = \max\{||x - y||_\infty \mid x, y \in \mathrm{dom}_{\mathbf{Z}}\, g\}$ if $\mathrm{dom}_{\mathbf{Z}}\, g$ is bounded.

**Modified Steepest Descent Algorithm:**
**Step 0:** Put $x := x^\circ$ and $\ell(i) := x^\circ(i) - L$ for all $i \in N$.
**Step 1:** If $x = \ell$, then return $x$ ($x$ is a minimizer of $g$).
**Step 2:** Choose any $i \in N$ with $x(i) > \ell(i)$.
**Step 3:** Find $j \in N$ that minimizes $g(x + \chi_j - \chi_i)$.
**Step 4:** Set $\ell(j) := x(j) + 1 - \chi_i(j)$ and $x := x + \chi_j - \chi_i$. Go to Step 1.

It is shown that the modified steepest descent algorithm finds a minimizer of an M-convex function in a pseudo-polynomial number of iterations.

**Theorem 3.2 ([21]).** *Let $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ be an M-convex function. Then, the modified steepest descent algorithm finds a minimizer $x_*$ of $g$ in* $\mathrm{O}(nL)$ *iterations.*

## 3.2 Proposed Algorithm

We consider speed-up of the modified steepest descent algorithm by using an optimal solution of continuous relaxation. In the following, we assume that a closed proper M-convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ with $f(x) = g(x)$ ($x \in \mathbf{Z}^n$) is readily available and satisfies the following conditions:

**(A1):** Minimization of $f$ can be solved efficiently.
**(A2):** $\mathrm{dom}_{\mathbf{R}}\, f$ coincides with the convex closure of $\mathrm{dom}_{\mathbf{Z}}\, g$.

For example, the problems (SC) and (LC) with quadratic convex objective functions satisfy the conditions above. Our algorithm is described as follows.

**Continuous Relaxation Algorithm:**
**Step 1:** Compute a minimizer $x_* \in \mathrm{dom}_{\mathbf{R}}\, f$ of the function $f$.
**Step 2:** Compute an integral vector $y \in \mathrm{dom}_{\mathbf{Z}}\, g$ with $||y - x_*||_\infty \leq n$.
**Step 3:** Apply the modified steepest descent algorithm to the M-convex function $g$ with the initial vector $x^\circ = y$ and $L = 2n - 1$.

We see from Theorem 2.3 that there exists some $y_* \in \arg\min g$ such that $||y_* - x_*||_\infty < n - 1$. Hence, the vector $y$ computed in Step 2 satisfies

$$||y_* - y||_\infty \leq ||y_* - x_*||_\infty + ||x_* - y||_\infty \leq L.$$

This together with Theorem 3.2 implies that the modified steepest descent algorithm with $x^\circ = y$ and $L = 2n - 1$ terminates in $\mathrm{O}(n^2)$ iterations. We denote by $T_1$ the time required by Step 1, i.e., the time to compute a minimizer of $f$, by $T_2$ the time required by Step 2, and by $F$ the time to evaluate the function value of $g$.

**Theorem 3.3.** *The continuous relaxation algorithm finds a minimizer of an M-convex function* $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ *in* $\mathrm{O}(T_1 + T_2 + n^3 F)$ *time.*

The time complexity $\mathrm{O}(T_1 + T_2 + n^3 F)$ for the continuous relaxation algorithm is better than that for the original modified steepest descent algorithm if $T_1$ and $T_2$ are not so big.

Step 2 can be done in (weakly) polynomial time by using a similar technique as in [20, Theorem 2.5] since $\mathrm{dom}_{\mathbf{R}}\, f$ is an integral base polyhedron under the assumption (A2). In some special cases of (MC), Step 2 can be done more easily and efficiently; indeed, for the problem (LC), any feasible solution $x \in \mathbf{R}^n$ of the continuous relaxation can be rounded to a feasible solution $y \in \mathbf{Z}^n$ of (LC) with $||y - x||_\infty < 1$ in $\mathrm{O}(n)$ time.

### 3.3 Computational Experiments

We compare the performance of our continuous relaxation algorithm with those of the previously proposed algorithms by computational experiments. We implemented the following four algorithms for the problem (MC) in the C language:

| symbol | algorithm |
|---|---|
| SD | steepest descent algorithm [14] |
| SD2 | modified steepest descent algorithm |
| SCALING | steepest descent scaling algorithm [9], [13, Sec. 10.1.2] |
| RELAX | our continuous relaxation algorithm |

We use the following libraries:

- "L-BFGS" by Nocedal[1] with its C++ wrapper by Kudo[2], which is an implementation of quasi-Newton method for unconstrained nonlinear function optimization [8]. As the routine requires the gradient of the objective function, we use a finite-difference approximation by $n + 1$ times of the function evaluation. This is only used in RELAX.

- pseudo-random number generator "SIMD-oriented Fast Mersenne Twister" by Saito and Matsumoto[3]. This is used to generate test instances.

[1] `http://www.ece.northwestern.edu/~nocedal/lbfgs.html`
[2] `http://chasen.org/~taku/software/misc/lbfgs/`
[3] `http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/SFMT/`

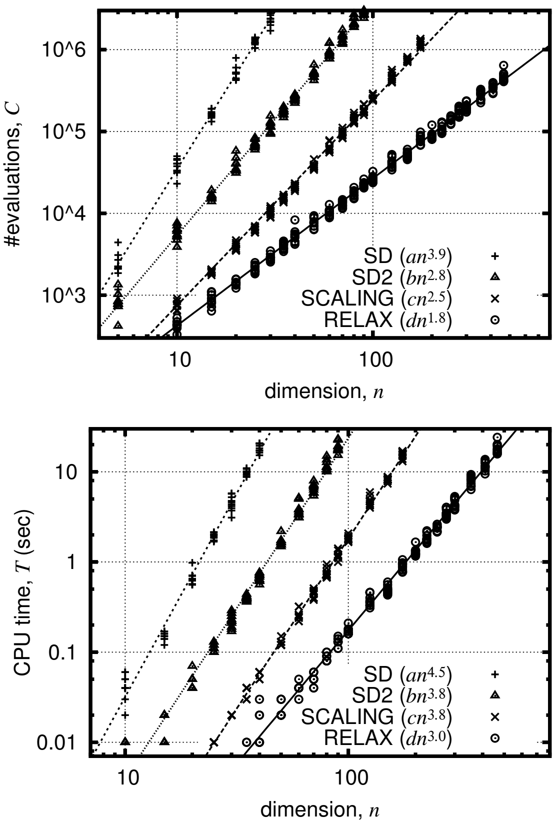


Figure 1: The number of function value evaluations and CPU time

We consider a special case of the problem (LC) as test instances for the problem (MC). More precisely, we consider an M-convex function $g : \mathbf{Z}^{n+1} \to \mathbf{R} \cup \{+\infty\}$ given as

$$g(x(0), x(1), \ldots, x(n)) = \begin{cases} \displaystyle\sum_{Y \in \mathcal{F}} \{a_Y x(Y)^2 + b_Y x(Y) + c_Y\} & (\text{if } \displaystyle\sum_{i=0}^{n} x(i) = 0), \\ +\infty & (\text{otherwise}), \end{cases}$$

where $\mathcal{F}$ is a laminar family of subsets of $\{1, 2, \ldots, n\}$. For each $n$, we generate ten instances with randomly chosen rational numbers $0 < a_X \leq 1000$, $-1000 \leq b_X, c_X \leq 1000$ $(X \in \mathcal{F})$. In addition, initial vectors $x^\circ \in \mathrm{dom}_{\mathbf{Z}}\, g$ used in the algorithms SD, SD2, and SCALING are also randomly generated under the condition $||x^\circ||_\infty \leq 10n$.

Our computational environment is described as follows:

> HP dx5150 SF/CT, AMD Athlon 64 3200+ processor (2.0GHz, 512KB L2 cache), 4GB memory, Vine Linux 4.1 (kernel 2.6.16), gcc 3.3.6.

We measure the number of function value evaluations and CPU time for each instance. Our experimental results are summarized in Figure 1, where the graph on the top (resp., on the bottom) shows the relationship between the number $C$ of function value evaluations (resp., CPU time $T$) and dimension $n$. It is easy to observe that in all of the implemented algorithms, it holds that $C = \mathrm{O}(n^h)$ for some $h$ and $T = \mathrm{O}(n^k)$ for some $k$. Actual numbers $h$ and $k$ for each algorithm are summarized below:

| algorithm | SD | SD2 | SCALING | RELAX |
|---|---|---|---|---|
| function value evaluations $C$ | $n^{3.9}$ | $n^{2.8}$ | $n^{2.5}$ | $n^{1.8}$ |
| CPU time $T$ | $n^{4.5}$ | $n^{3.8}$ | $n^{3.8}$ | $n^{3.0}$ |

These computational experiments show that our continuous relaxation algorithm is faster than the previously proposed algorithms, at least for the tested instances.

# 4 Proofs

Proofs of Theorems 2.1 and 2.3 use the following property, stating that the projection of a closed proper M-convex function along an arbitrarily chosen coordinate axis $i \in N$ is a supermodular function.

**Theorem 4.1 ([17, Proposition 3.12]).** *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper M-convex function. For every $x, y \in \mathbf{R}^n$ and $i \in N$, we have $f(x) + f(y) \le f(\hat{x}) + f(\check{y})$, where $\hat{x}$ and $\check{y}$ are given as*

$$\hat{x}(j) = \begin{cases} \min\{x(j), y(j)\} & (j \in N \setminus \{i\}), \\ x(N) - \sum_{k \in N \setminus \{i\}} \min\{x(k), y(k)\} & (j = i), \end{cases}$$

$$\check{y}(j) = \begin{cases} \max\{x(j), y(j)\} & (j \in N \setminus \{i\}), \\ y(N) - \sum_{k \in N \setminus \{i\}} \max\{x(k), y(k)\} & (j = i). \end{cases}$$

## 4.1 Proof of Theorem 2.1

Recall that $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is a closed proper M-convex function and the vector $y_* \in \mathrm{dom}_{\mathbf{R}} f$ satisfies

$$f(y_* - \chi_i + \chi_j) \ge f(y_*) \qquad (\forall i, j \in N). \tag{4.1}$$

To prove Theorem 2.1, it suffices to show the following property:

**(P1)** for every $x' \in \mathrm{dom}_{\mathbf{R}} f$, there exists some $x \in \mathrm{dom}_{\mathbf{R}} f$ such that $f(x) \le f(x')$ and $||x - y_*||_\infty < n - 1$.

Since $f$ is closed proper convex and the set $\{x \in \mathrm{dom}_{\mathbf{R}} f \mid ||x-y_*||_\infty < n-1\}$ is bounded, the property (P1) implies that there exists a minimizer of $f$ such that $||x-y_*||_\infty \leq n-1$. This shows that $\arg\min f \neq \emptyset$, in particular. Hence, (P1) immediately implies that there exists some $x_* \in \arg\min f$ such that $||x_* - y_*||_\infty < n-1$.

We now prove the property (P1). Let $x' \in \mathrm{dom}_{\mathbf{R}} f$ be any vector. Also, let $\hat{x} \in \mathrm{dom}_{\mathbf{R}} f$ be a vector satisfying $f(\hat{x}) \leq f(x')$, and suppose that $\hat{x}$ minimizes the $\mathrm{L}_1$ distance $||\hat{x} - y_*||_1$ from the vector $y_*$ among all such vectors. We show that for every $k \in N$ it holds that

$$|\hat{x}(k) - y_*(k)| < n-1. \tag{4.2}$$

In the following, we fix $k \in N$ and assume, without loss of generality, that $\hat{x}(k) > y_*(k)$ since the case $\hat{x}(k) < y_*(k)$ can be dealt with in a similar way and the case $\hat{x}(k) = y_*(k)$ immediately implies (4.2).

By the choice of $\hat{x}$, we have

$$\begin{array}{l} f(\hat{x} - \varepsilon(\chi_k - \chi_j)) > f(\hat{x}) \\ \qquad (\forall j \in \mathrm{supp}^-(\hat{x} - y_*),\ 0 < \forall \varepsilon \leq \min(\hat{x}(k) - y_*(k), y_*(j) - \hat{x}(j))). \end{array} \tag{4.3}$$

Let $\mathrm{supp}^-(\hat{x} - y_*) = \{j_1, j_2, \cdots, j_t\}$, where $t = |\mathrm{supp}^-(\hat{x} - y_*)|$ $(\leq n-1)$. Put $y_0 = y_*$, and we iteratively define $\lambda_h \in \mathbf{R}_+$ and $y_h \in \mathbf{R}^n$ for each $h = 1, 2, \cdots, t$ by

$$\begin{aligned} \lambda_h \;=\; \sup\{\lambda \mid\ & y_{h-1} + \lambda(\chi_k - \chi_{j_h}) \in \mathrm{dom}_{\mathbf{R}} f, \\ & \lambda \leq \min(\hat{x}(k) - y_{h-1}(k), y_{h-1}(j_h) - \hat{x}(j_h)), \\ & f(y_{h-1} + \lambda'(\chi_k - \chi_{j_h})) \text{ is strictly decreasing in } \lambda' \in [0, \lambda]\} \end{aligned}$$

and $y_h = y_{h-1} + \lambda_h(\chi_k - \chi_{j_h})$. By the definition of $y_h$ and closed convexity of $f$, we have

$$f(y_h) < f(y_{h-1}) \qquad \text{if } \lambda_h > 0, \tag{4.4}$$

$$\begin{array}{l} f(y_h + \lambda(\chi_k - \chi_{j_h})) \geq f(y_h) \quad (\forall \lambda > 0) \\ \qquad\qquad \text{if } \hat{x}(k) > y_h(k) \text{ and } y_h(j_h) > \hat{x}(j_h). \end{array} \tag{4.5}$$

**Claim 1:** $\sum_{h=1}^{t} \lambda_h = \hat{x}(k) - y_0(k)$.

[Proof of Claim 1] Assume, to the contrary, that $\sum_{h=1}^{t} \lambda_h < \hat{x}(k) - y_0(k)$. Since $k \in \mathrm{supp}^+(\hat{x} - y_t)$, (M-EXC[$\mathbf{R}$]) implies that there exist $j_h \in \mathrm{supp}^-(\hat{x} - y_t) \subseteq \mathrm{supp}^-(\hat{x} - y_0)$ and a sufficiently small $\lambda > 0$ such that

$$f(\hat{x}) + f(y_t) \geq f(\hat{x} - \lambda(\chi_k - \chi_{j_h})) + f(y_t + \lambda(\chi_k - \chi_{j_h})).$$

By Theorem 4.1, we obtain

$$f(y_h + \lambda(\chi_k - \chi_{j_h})) + f(y_t) \leq f(y_t + \lambda(\chi_k - \chi_{j_h})) + f(y_h).$$

Combining the two inequalities, we have

$$f(y_h + \lambda(\chi_k - \chi_{j_h})) - f(y_h) \leq f(\hat{x}) - f(\hat{x} - \lambda(\chi_k - \chi_{j_h})) < 0,$$

where the last inequality is by (4.3). This, however, contradicts (4.5).

[End of Claim 1]

**Claim 2:** For $h = 1, 2, \ldots, t$, if $\lambda_h > 0$ then $f(y_* + \lambda_h(\chi_k - \chi_{j_h})) < f(y_*)$.
[Proof of Claim 2] Let $h$ be any integer in $\{1, 2, \ldots, t\}$ with $\lambda_h > 0$. By Theorem 4.1, we have

$$f(y_* + \lambda_h(\chi_k - \chi_{j_h})) + f(y_{h-1}) \leq f(y_h) + f(y_*),$$

which implies

$$f(y_* + \lambda_h(\chi_k - \chi_{j_h})) - f(y_*) \leq f(y_h) - f(y_{h-1}) < 0,$$

where the last inequality is by (4.4). [End of Claim 2]

By the inequality (4.1) and convexity of $f$, we have

$$f(y_* + \beta(\chi_k - \chi_j)) \geq f(y_*) \qquad (\forall \beta \geq 1, \ \forall j \in N).$$

Therefore, it follows from Claim 2 that $\lambda_h < 1$ for all $h = 1, 2, \ldots, t$, which, together with Claim 1, implies the desired inequality (4.2) as follows:

$$\hat{x}(k) - y_*(k) = \hat{x}(k) - y_0(k) = \sum_{h=1}^{t} \lambda_h < t \leq n - 1.$$

This concludes the proof of Theorem 2.1.

**Remark 4.2.** The proof of Theorem 2.1 above essentially shows the following properties:

**Theorem 4.3.** *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper M-convex function.*
(i) *Suppose that $y_* \in \mathrm{dom}_{\mathbf{R}} f$ and $k \in N$ satisfy the condition $f(y_* + \chi_k - \chi_j) \geq f(y_*)$ $(\forall j \in N)$. Then, for every $x' \in \mathrm{dom}_{\mathbf{R}} f$ there exists some $x \in \mathrm{dom}_{\mathbf{R}} f$ such that $f(x) \leq f(x')$ and $x(k) < y_*(k) + (n-1)$. In particular, it holds that*

$$\inf\{f(x) \mid x \in \mathrm{dom}_{\mathbf{R}} f, \ x(k) < y_*(k) + (n-1)\} = \inf f.$$

(ii) *Suppose that $y_* \in \mathrm{dom}_{\mathbf{R}} f$ and $k \in N$ satisfy the condition $f(y_* - \chi_k + \chi_j) \geq f(y_*)$ $(\forall j \in N)$. Then, for every $x' \in \mathrm{dom}_{\mathbf{R}} f$ there exists some $x \in \mathrm{dom}_{\mathbf{R}} f$ such that $f(x) \leq f(x')$ and $x(k) > y_*(k) + (n-1)$. In particular, it holds that*

$$\inf\{f(x) \mid x \in \mathrm{dom}_{\mathbf{R}} f, \ x(k) > y_*(k) - (n-1)\} = \inf f.$$

□

## 4.2 Proof of Theorem 2.3

The proof of Theorem 2.3 given below is similar to that for Theorem 2.1. Recall that $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is a closed proper M-convex function, $f_{\mathbf{Z}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ is the discretization of $f$, and $x_* \in \arg\min f$. To prove Theorem 2.3, we will show the following property holds:

> **(P2)** for every $y' \in \mathrm{dom}_{\mathbf{Z}} f_{\mathbf{Z}}$, there exists some $y \in \mathrm{dom}_{\mathbf{Z}} f_{\mathbf{Z}}$ such that $f_{\mathbf{Z}}(y) \le f_{\mathbf{Z}}(y')$ and $||y - x_*||_\infty < n - 1$.

Since $\{y \in \mathrm{dom}_{\mathbf{Z}} f_{\mathbf{Z}} \mid ||y - x_*||_\infty < n - 1\}$ is a finite set, the property (P2) immediately implies that there exists some $y_* \in \arg\min f_{\mathbf{Z}}$ such that $||y_* - x_*||_\infty < n - 1$.

We now prove the property (P2). Let $y' \in \mathrm{dom}_{\mathbf{Z}} f_{\mathbf{Z}}$ be any vector. Also, let $\hat{y} \in \mathrm{dom}_{\mathbf{Z}} f_{\mathbf{Z}}$ be a vector satisfying $f_{\mathbf{Z}}(\hat{y}) \le f_{\mathbf{Z}}(y')$, and suppose that $\hat{y}$ minimizes $||\hat{y} - x_*||_1$ among all such vectors. We show that for every $k \in N$ it holds that

$$|\hat{y}(k) - x_*(k)| < n - 1. \tag{4.6}$$

In the following, we fix $k \in N$ and assume, without loss of generality, that $\hat{y}(k) > x_*(k)$ since the case $\hat{y}(k) < x_*(k)$ can be dealt with in a similar way and the case $\hat{y}(k) = x_*(k)$ immediately implies (4.6).

By the choice of $\hat{y}$, we have

$$f(\hat{y} - \chi_k + \chi_j) > f(\hat{y}). \qquad (\forall j \in \mathrm{supp}^-(\hat{y} - x_*)). \tag{4.7}$$

Let $\mathrm{supp}^-(\hat{y} - x_*) = \{j_1, j_2, \cdots, j_t\}$, where $t = |\mathrm{supp}^-(\hat{y} - x_*)|$ $(\le n - 1)$. Put $y_0 = \hat{y}$, and for each $h = 1, 2, \cdots, t$, we iteratively define $\lambda_h \in \mathbf{R}_+$ and $y_h \in \mathbf{R}^n$ by

$$\begin{aligned} \lambda_h = \sup\{\lambda \mid\ & y_{h-1} - \lambda(\chi_k - \chi_{j_h}) \in \mathrm{dom}_{\mathbf{R}} f, \\ & \lambda \le \min(y_{h-1}(k) - x_*(k),\ x_*(j_h) - y_{h-1}(j_h)), \\ & f(y_{h-1} - \lambda(\chi_k - \chi_{j_h})) \le f(y_{h-1})\} \end{aligned}$$

and $y_h = y_{h-1} - \lambda_h(\chi_k - \chi_{j_h})$. By the definition of $y_h$ and closed convexity of $f$, we have

$$f(y_h) \le f(y_{h-1}) \qquad (h = 1, 2, \ldots, t), \tag{4.8}$$

$$f(y_h - \lambda(\chi_k - \chi_{j_h})) > f(y_h) \quad (\forall \lambda > 0) \quad \text{if } x_*(k) > y_h(k) \text{ and } y_h(j_h) > x_*(j_h). \tag{4.9}$$

**Claim 1:** $\sum_{h=1}^{t} \lambda_h = y_0(k) - x_*(k)$.
[Proof of Claim 1] Assume, to the contrary, that $\sum_{h=1}^{t} \lambda_h < y_0(k) - x_*(k)$. Since $k \in \mathrm{supp}^+(y_t - x_*)$, there exist $j_h \in \mathrm{supp}^-(y_t - x_*) \subseteq \mathrm{supp}^-(y_0 - x_*)$ and a sufficiently small $\lambda > 0$ such that

$$f(y_t) + f(x_*) \ge f(y_t - \lambda(\chi_k - \chi_{j_h})) + f(x_* + \lambda(\chi_k - \chi_{j_h})).$$

By Theorem 4.1, we obtain

$$f(y_h - \lambda(\chi_k - \chi_{j_h})) + f(y_t) \leq f(y_t - \lambda(\chi_k - \chi_{j_h})) + f(y_h).$$

Combining the two inequalities, we have

$$f(y_h - \lambda(\chi_k - \chi_{j_h})) - f(y_h) \leq f(x_*) - f(x_* + \lambda(\chi_k - \chi_{j_h})) \leq 0,$$

where the last inequality follows from $x_* \in \arg\min f$. This, however, contradicts (4.9). [End of Claim 1]

**Claim 2:** For $h = 1, 2, \ldots, t$, we have $f(\hat{y} - \lambda_h(\chi_k - \chi_{j_h})) \leq f(\hat{y})$.
[Proof of Claim 2] Let $h \in \{1, 2, \ldots, t\}$. By Theorem 4.1, we have

$$f(\hat{y} - \lambda_h(\chi_k - \chi_{j_h})) + f(y_{h-1}) \leq f(y_h) + f(\hat{y}),$$

which implies

$$f(\hat{y} - \lambda_h(\chi_k - \chi_{j_h})) - f(\hat{y}) \leq f(y_h) - f(y_{h-1}) \leq 0,$$

where the last inequality is by (4.8). [End of Claim 2]

By the convexity and Claim 2, it holds that

$$f(\hat{y} - \lambda(\chi_k - \chi_{j_h})) \leq f(\hat{y}) \qquad (\forall \lambda \in [0, \lambda_h],\ \forall h = 1, 2, \ldots, t).$$

Hence, it follows from (4.7) that $\lambda_h < 1$ for all $h = 1, 2, \ldots, t$, which, together with Claim 1, implies the desired inequality (4.6) as follows:

$$\hat{y}(k) - x_*(k) = y_0(k) - x_*(k) = \sum_{h=1}^{t} \lambda_h < t \leq n - 1.$$

This concludes the proof of Theorem 2.3.

**Remark 4.4.** The proof of Theorem 2.3 above essentially shows the following properties:

**Theorem 4.5.** *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper M-convex function and $f_{\mathbf{Z}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ the discretization of $f$.*
(i) *Suppose that $x_* \in \mathrm{dom}_{\mathbf{R}} f$ and $k \in N$ satisfy the condition $f'(x_*; k, j) \geq 0$ $(\forall j \in N)$. Then, it holds that*

$$\inf\{f_{\mathbf{Z}}(y) \mid y \in \mathrm{dom}_{\mathbf{Z}} f_{\mathbf{Z}},\ y(k) < x_*(k) + (n-1)\} = \inf f_{\mathbf{Z}}.$$

(ii) *Suppose that $x_* \in \mathrm{dom}_{\mathbf{R}} f$ and $k \in N$ satisfy the condition $f'(x_*; j, k) \geq 0$ $(\forall j \in N)$. Then, it holds that*

$$\inf\{f_{\mathbf{Z}}(y) \mid y \in \mathrm{dom}_{\mathbf{Z}} f_{\mathbf{Z}},\ y(k) > y_*(k) - (n-1)\} = \inf f_{\mathbf{Z}}.$$

□

## 5 Concluding Remarks

**Remark 5.1.** We consider the following statements for M-convex function minimization (MC) and its continuous relaxation ($\overline{\text{MC}}$).

> (i) for every optimal solution $x_* \in \mathbf{R}^n$ of ($\overline{\text{MC}}$), there exists some optimal solution $y_* \in \mathbf{Z}^n$ of (MC) such that $y_* \leq \lceil x_* \rceil$,
> (ii) for every optimal solution $x_* \in \mathbf{R}^n$ of ($\overline{\text{MC}}$), there exists some optimal solution $y_* \in \mathbf{Z}^n$ of (MC) such that $y_* \geq \lfloor x_* \rfloor$,

where vectors $\lceil x_* \rceil, \lfloor x_* \rfloor \in \mathbf{Z}^n$ are given by $(\lceil x_* \rceil)(i) = \lceil x_*(i) \rceil$, $(\lfloor x_* \rfloor)(i) = \lfloor x_*(i) \rfloor$ $(i \in N)$. Examples 2.6 and 2.7 show that neither of these two statements hold, even for the simple resource allocation problem (SIMPLE) (see also Remark 5.3).

On the other hand, (SIMPLE) and its continuous relaxation ($\overline{\text{SIMPLE}}$) satisfies the following weaker statement (see, e.g., [6, Section 4.6] for a proof):

> (iii) for every optimal solution $x_* \in \mathbf{R}^n$ of ($\overline{\text{SIMPLE}}$), there exists some optimal solution $y_* \in \mathbf{Z}^n$ of (SIMPLE) satisfying either $y_* \leq \lceil x_* \rceil$ or $y_* \geq \lfloor x_* \rfloor$ (or both).

□

**Remark 5.2.** Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper convex function and $f_{\mathbf{Z}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ the discretization of $f$. Corollary 2.2 and Theorem 2.3 show that for any M-convex $f$, we have $\arg\min f \neq \emptyset$ if and only if $\arg\min f_{\mathbf{Z}} \neq \emptyset$. In the general case where $f$ is not necessarily M-convex, however, the properties $\arg\min f \neq \emptyset$ and $\arg\min f_{\mathbf{Z}} \neq \emptyset$ are independent of each other, as shown in the following two examples.

Let $f_0 : \mathbf{R}^2 \to \mathbf{R} \cup \{+\infty\}$ be a function defined by

$$f_0(x_1, x_2) = \begin{cases} -1 & (\text{if } (x_1, x_2) = (0, 0.5)), \\ 1/(x_1 + 1) & (\text{if } x_1 \in \mathbf{Z}_+, x_2 = 0), \\ +\infty & (\text{otherwise}). \end{cases}$$

Let $f : \mathbf{R}^2 \to \mathbf{R} \cup \{+\infty\}$ be the convex closure of $f_0$ (see, e.g., [13] for the definition of the convex closure) and $f_{\mathbf{Z}} : \mathbf{Z}^2 \to \mathbf{R} \cup \{+\infty\}$ the discretization of $f$. By definition, the function $f$ is a closed proper convex function such that $\mathrm{dom}_{\mathbf{R}} f = \{(x_1, x_2) \in \mathbf{R}^2 \mid x_1 \geq 0,\ 0 \leq x_2 \leq 0.5\}$. We have $f(x_1, x_2) = f_0(x_1, x_2)$ for every $(x_1, x_2) \in \mathrm{dom}_{\mathbf{R}} f_0$, and the function $f$ is linear on the convex hull of the set $\{(1/n, 0), (1/(n+1), 0), (0, 0.5)\}$ for each $n \in \mathbf{Z}_+$. Moreover, we have $f(x_1, x_2) = 0$ if $x_1 > 0$ and $x_2 = 0.5$. Hence, $\arg\min f = \{(0, 0.5)\} \neq \emptyset$ holds. On the other hand, $\inf f_{\mathbf{Z}} = \inf\{1/(x_1 + 1) \mid x_1 \in \mathbf{Z}_+\} = 0$ and $\arg\min f_{\mathbf{Z}} = \emptyset$.

We then consider a closed proper convex function $h : \mathbf{R}^2 \to \mathbf{R} \cup \{+\infty\}$ defined by

$$h(x_1, x_2) = \begin{cases} 1/(x_1 + 1) & (\text{if } x_1 \geq 0,\ x_2 = \sqrt{x_1}), \\ +\infty & (\text{otherwise}). \end{cases}$$

Let $h_{\mathbf{Z}} : \mathbf{Z}^2 \to \mathbf{R} \cup \{+\infty\}$ be the discretization of $h$. Then, $\mathrm{dom}_{\mathbf{Z}}\, h_{\mathbf{Z}} = \arg\min h_{\mathbf{Z}} = \{(0,0)\}$ and $\min h_{\mathbf{Z}} = 1$. On the other hand, $\inf h = \inf_{x_1 \geq 0} 1/(x_1 + 1) = 0$ and there exists no $(x_1, x_2) \in \mathrm{dom}_{\mathbf{R}}\, h$ with $h(x_1, x_2) = \inf h$, i.e., $\arg\min h = \emptyset$. □

**Remark 5.3.** In Hochbaum [4], the following "proximity theorem" is presented for the problem (SC) and its continuous relaxation $(\overline{\mathrm{SC}})$, where $\mathbf{1} = (1, 1, \ldots, 1) \in \mathbf{Z}^n$:

> **Statement A (Corollary 4.3 in [4])**
> (i) For every optimal solution $y_* \in \mathbf{Z}^n$ of (SC), there exists some optimal solution $x_* \in \mathbf{R}^n$ of $(\overline{\mathrm{SC}})$ such that
>
> $$y_* - \mathbf{1} < x_* < y_* + n\mathbf{1}.$$
>
> (ii) For every optimal solution $x_* \in \mathbf{R}^n$ of $(\overline{\mathrm{SC}})$, there exists some optimal solution $y_* \in \mathbf{Z}^n$ of (SC) such that
>
> $$y_* - \mathbf{1} < x_* < y_* + n\mathbf{1}.$$

This statement, however, is incorrect; indeed, Examples 2.6 and 2.7 show that Statement A does not hold even for the simple resource allocation problems (see Remark 5.1). Moreover, Example 5.4 below shows that Statement A does not hold even for the simple resource allocation problem with quadratic objective function.

Statement A is used in the paper [5] to show the strongly polynomial-time solvability of several special cases of (SC) with quadratic objective function. In particular, from Statement A follows the bound $\mathrm{O}(n)$ for the $\mathrm{L}_1$ distance between optimal solutions of (SC) and $(\overline{\mathrm{SC}})$, which is used in the paper [5] to analyze the time complexity of the proposed algorithms. We can still show the results of strongly polynomial-time solvability in [5] by using Theorem 2.3 instead of Statement A since Theorem 2.3 implies the bound $\mathrm{O}(n^2)$ for the $\mathrm{L}_1$ distance between optimal solutions of (SC) and $(\overline{\mathrm{SC}})$. It is not clear, however, whether the time complexity results in [5] still hold true without Statement A since our bound $\mathrm{O}(n^2)$ is worse than $\mathrm{O}(n)$ used in [5]. □

**Example 5.4.** For a sufficiently small positive number $\delta$, we consider the problem (SIMPLE), where $K = n - 1$, $u(i) = +\infty$ $(i \in N)$, and convex functions $f_i : \mathbf{R} \to \mathbf{R}$ $(i \in N)$ are given as

$$\begin{aligned} f_1(\alpha) &= \delta\alpha \qquad (\alpha \in \mathbf{R}), \\ f_i(\alpha) &= (\alpha - 0.5 + \delta)^2 \qquad (\alpha \in \mathbf{R},\ i = 2, 3, \ldots, n). \end{aligned}$$

It is noted that $f_i(1) - f_i(0) = 2\delta$ for $i = 2, 3, \ldots, n$. Then, the optimal solutions $y_* \in \mathbf{Z}^n$ and $x_* \in \mathbf{R}^n$ of the problem (SIMPLE) and its continuous relaxation, respectively, are uniquely given as follows:

$$y_* = (n-1, 0, \ldots, 0), \qquad x_* = \left( \frac{(n-1)(1+\delta)}{2}, \frac{1-\delta}{2}, \frac{1-\delta}{2}, \ldots, \frac{1-\delta}{2} \right).$$

Since $\delta$ is a sufficiently small positive number, we have

$$y_*(1) - x_*(1) = (n-1) - \frac{(n-1)(1+\delta)}{2} = \frac{(n-1)(1-\delta)}{2} > 1.$$

This shows that $y_* - \mathbf{1} < x_*$ does not hold. □

## Acknowledgements

The authors thank Kazuo Murota for his valuable comments. This work is supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

[1] S. Fujishige, Submodular Functions and Optimization, Second Edition, Elsevier, Amsterdam, 2005.

[2] E. Girlich, M. Kovalev, and A. Zaporozhets, A polynomial algorithm for resource allocation problems with polymatroid constraints, Optimization, 37 (1996), pp. 73–86.

[3] H. Groenevelt, Two algorithms for maximizing a separable concave function over a polymatroid feasible region, European J. Oper. Res., 54 (1991), pp. 227–236.

[4] D.S. Hochbaum, Lower and upper bounds for the allocation problem and other nonlinear optimization problems, Math. Oper. Res., 19 (1994), pp. 390–409.

[5] D.S. Hochbaum, and S.-P. Hong, About strongly polynomial time algorithms for quadratic optimization over submodular constraints, Math. Program., 69 (1995), pp. 269–309.

[6] T. Ibaraki and N. Katoh, Resource Allocation Problems: Algorithmic Approaches, MIT Press, Boston, MA, 1988.

[7] N. Katoh and T. Ibaraki, Resource allocation problems, in: Handbook of Combinatorial Optimization II, D. -Z. Du and P. M. Pardalos, eds., Kluwer, Boston, MA, 1998, pp. 159–260.

[8] D.C. Liu and J. Nocedal, On the limited memory method for large scale optimization, Math. Prog., 45 (1989), pp. 503–528,

[9] S. Moriguchi, K. Murota, and A. Shioura, Scaling algorithms for M-convex function minimization, IEICE Trans. Fund., E85-A (2002), pp. 922–929.

[10] S. Moriguchi and A. Shioura, On Hochbaum's scaling algorithm for the general resource allocation problem, Math. Oper. Res., 29 (2004), pp. 394–397.

[11] K. Murota, Convexity and Steinitz's exchange property, Adv. Math., 124 (1996), pp. 272–311.

[12] K. Murota, Discrete convex analysis, Math. Program., 83 (1998), pp. 313–371.

[13] K. Murota, Discrete Convex Analysis, Society for Industrial and Applied Mathematics, Philadelphia, 2003.

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

[15] K. Murota, Primer of Discrete Convex Analysis: Discrete versus Continuous Optimization, Kyoritsu Publishing Co., Tokyo, 2007. (In Japanese)

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

[17] K. Murota and A. Shioura, Conjugacy relationship between M-convex and L-convex functions in continuous variables, Math. Program., 101 (2004), pp. 415-433.

[18] K. Murota and A. Shioura, Fundamental properties of M-convex and L-convex functions in continuous variables, IEICE Trans. Fund., E87-A (2004), pp. 1042–1052.

[19] K. Murota and A. Shioura, Quadratic M-convex and L-convex functions, Adv. Appl. Math., 33 (2004), pp. 318–341.

[20] A. Shioura, Minimization of an M-convex function, Discrete Appl. Math., 84 (1998), pp. 215–220.

[21] A. Shioura, Fast scaling algorithms for M-convex function minimization with application to resource allocation problem, Discrete Appl. Math., 134 (2003), pp. 303–316.

[22] A. Tamura, Coordinatewise domain scaling algorithm for M-convex function minimization, Math. Prog., 102 (2005), pp. 339–354.