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

Satoko MORIGUCHI* Akiyoshi SHIOURA† Nobuyuki TSUCHIMURA‡

January 17, 2011

Abstract

The concept of M-convexity for functions in integer variables, introduced by Murota (1996), 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. By establishing proximity theorems we develop a new algorithm based on continuous relaxation. We apply the approach to some special cases of the separable convex quadratic resource allocation problem and the convex quadratic tree resource allocation problem to obtain faster algorithms.

*School of Industrial Technology, Advanced Institute of Industrial Technology, Tokyo 140-0011, Japan, satoko@aiit.ac.jp

†Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan, shioura@dais.is.tohoku.ac.jp, tel/fax:+81-22-795-4753

‡Department of Informatics, Kwansei Gakuin University, Sanda 669-1337, Japan, tutimura@kwansei.ac.jp

1 Introduction

The concept of M-convexity for functions in integer variables, introduced by Murota [21, 22], plays a primary role in the theory of discrete convex analysis [24]. 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 the most fundamental optimization problem concerning M-convex functions. For this problem, various approaches have been proposed to develop efficient algorithms [18, 34, 35, 37]. In this paper, we discuss a continuous relaxation approach for M-convex function minimization.

M-convex function minimization Let n be a positive integer with $n \geq 2$, and put $N = \{1, 2, \ldots, n\}$. 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 $\mathrm{dom}_{\mathbf{Z}}\, g$ of an M-convex function g is the set of integral points in the base polyhedron of some integral submodular system; in particular, $\mathrm{dom}_{\mathbf{Z}}\, g$ lies on a hyperplane $\{x \in \mathbf{Z}^n \mid x(N) = r\}$ for some integer r (see [24, Section 6.1]). In this sense, the concept of M-convex function can be seen as a generalization of the concept of integral base polyhedron. M-convex function is also an extension of valuated matroid introduced by Dress and Wenzel [3]; a function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ is a valuated matroid if and only if $-g$ is an M-convex function with $\mathrm{dom}_{\mathbf{Z}}\, g \subseteq \{0, 1\}^n$.

In this paper, we consider the minimization of an M-convex function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$:

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

Below we give two 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 nonnegative-valued submodular function, i.e., ρ satisfies $\rho(X) + \rho(Y) \geq \rho(X \cap Y) + \rho(X \cup Y)$ for every $X, Y \in 2^N$. We assume $\rho(\emptyset) = 0$ and $\rho(N) < +\infty$. The *(separable convex) resource allocation problem under a submodular constraint* [6, 11, 14, 16] is formulated as follows:

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

where $x(Y) = \sum_{i \in Y} x(i)$ for $Y \subseteq N$ and $\mathbf{0} = (0, 0, \dots, 0) \in \mathbf{Z}^n$. A very special case of (SC) is the *simple resource allocation problem* [6, 14, 16] formulated as follows:

$$\textsf{(Simple)} \quad \text{Minimize} \quad \sum_{i=1}^{n} f_i(x(i)) \quad \text{subject to} \quad x(N) = K,\ \mathbf{0} \le x \le u,\ x \in \mathbf{Z}^n,$$

where $K \in \mathbf{Z}_+$ and $u \in \mathbf{Z}^n_+$. The problem (SC) is extensively discussed in the literature; see [6, 14, 16] for comprehensive review of this problem and [7, 10, 11, 12] for efficient algorithms.

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

$$g_{\textsf{SC}}(x) = \begin{cases} \sum_{i=1}^{n} f_i(x(i)) & \text{(if } x \in \mathbf{Z}^n \text{ is a feasible solution to (SC)),} \\ +\infty & \text{(otherwise)} \end{cases}$$

satisfies (M-EXC[**Z**]) (see [22, Example 2.2], [24, Section 6.3]). □

Example 1.2 (Laminar convex resource allocation problem). Let $\mathcal{F} \subseteq 2^N$ be a laminar family, i.e., for every $X, Y \in \mathcal{F}$ either $X \subseteq Y$, $X \supseteq Y$, or $X \cap Y = \emptyset$ holds. Note that $|\mathcal{F}| = \mathrm{O}(n)$. We consider the following minimization problem with a nonseparable convex objective function:

$$\textsf{(Laminar)} \left| \begin{array}{ll} \text{Minimize} & \displaystyle\sum_{Y \in \mathcal{F}} f_Y(x(Y)) \\ \text{subject to} & x(N) = K,\ \ell_Y \le x(Y) \le u_Y\ (Y \in \mathcal{F}), \\ & x \ge \mathbf{0},\ x \in \mathbf{Z}^n, \end{array} \right.$$

where $f_Y : \mathbf{R} \to \mathbf{R}$ $(Y \in \mathcal{F})$ is a family of univariate convex functions, $K \in \mathbf{Z}_+$, and $\ell_Y, u_Y \in \mathbf{Z}_+$ for every $Y \in \mathcal{F}$. We call this problem the *laminar convex resource allocation problem.* The problem (Laminar) is a special case of (MC) since the function $g_{\textsf{Laminar}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ defined by

$$g_{\textsf{Laminar}}(x) = \begin{cases} \sum_{Y \in \mathcal{F}} f_Y(x(Y)) & (x \in \mathbf{Z}^n \text{ is a feasible solution to (Laminar)}), \\ +\infty & \text{(otherwise)} \end{cases}$$

satisfies (M-EXC[**Z**]) (see [18, Example 2.3], [24, Section 6.3]). The problem (Laminar) is an important special case of (MC) since the objective function is nonseparable and continuous relaxation approach proposed in this paper can be applied in a natural way. □

For the problem of minimizing an M-convex function, various approaches have been proposed to develop efficient algorithms [18, 34, 35, 37], and the best time complexity bounds are $\mathrm{O}((n^3 + n^2 \log(L/n))(\log(L/n)/\log n)F)$ and $\mathrm{O}(n^3 \log(L/n)F)$, where L is an upper bound on the L_∞ distance between two vectors in $\mathrm{dom}_{\mathbf{Z}}\, g$, i.e.,

$$L = \max\{\|x - y\|_\infty \mid x, y \in \mathrm{dom}_{\mathbf{Z}}\, g\}, \tag{1.1}$$

and F denotes the time to evaluate the function value of the M-convex function g. In this paper, we consider a new approach for M-convex function minimization based on continuous relaxation.

Continuous relaxation Continuous relaxations of (SC) and $(\mathsf{Laminar})$ can be naturally obtained by removing the integrality constraint "$x \in \mathbf{Z}^n$." Indeed, a continuous relaxation approach for solving (SC) is proposed by Hochbaum [11] (see also Hochbaum and Hong [12]), and the approach is applied to obtain efficient algorithms for some important special cases of (SC) with quadratic objective functions by Hochbaum and Hong [12].

The continuous relaxation approach consists of the following three major steps:

Step 1: Compute an optimal solution $x_* \in \mathbf{R}^n$ to the continuous relaxation problem.
Step 2: Round the optimal solution x_* to an integral vector $y^\circ \in \mathbf{Z}^n$.
Step 3: Using y° as an initial solution, compute an optimal solution $y_* \in \mathbf{Z}^n$ to the original problem by a greedy algorithm.

In this paper, we extend such approach to M-convex function minimization. Although this approach cannot possibly provide a faster algorithm for the general case of (MC), it is shown that faster algorithms can be obtained for some important special cases of (MC).

To extend the continuous relaxation approach to the problem (MC), we need to define a continuous relaxation of (MC) in an appropriate way. In this paper, we define a continuous relaxation of (MC) by using the concept of *M-convex function in real variables* introduced by Murota and Shioura [29]. A function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ in real 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 $\mathrm{dom}_{\mathbf{R}} f = \{x \in \mathbf{R}^n \mid f(x) < +\infty\}$. An M-convex function is said to be a *closed proper M-convex function* if it is a closed proper convex function, in addition (the definition of closed proper convex functions is given in Section 2). M-convex functions in real variables constitute a subclass of convex functions with additional combinatorial properties such as supermodularity and local polyhedral structure (see, e.g., [24, 28, 29, 30, 31]). Fundamental properties of M-convex functions are investigated in [30], such as equivalent axioms, subgradients, directional derivatives, etc.

It is known (see, e.g, [24, Section 6.11]) that for every M-convex function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ in integer variables, there exists a closed proper M-convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ in real variables such that $f(x) = g(x)$ $(\forall x \in \mathbf{Z}^n)$. Using such an M-convex function f in real variables, we define a continuous relaxation of (MC) as follows:

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

We note that continuous relaxations of (SC) and $(\mathsf{Laminar})$ can be also formulated in this form by using functions $f_{\mathsf{SC}} : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ and $f_{\mathsf{Laminar}} : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ which are defined in the same way as g_{SC} and g_{Laminar} 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_{SC} and f_{Laminar} are M-convex functions in real variables satisfying $f_{\mathsf{SC}}(x) = g_{\mathsf{SC}}(x)$ and $f_{\mathsf{Laminar}}(x) = g_{\mathsf{Laminar}}(x)$ for all $x \in \mathbf{Z}^n$.

Our results Efficiency of algorithms based on continuous relaxation depends on the distance between optimal solutions to the original problem and to its continuous relaxation, and so-called "proximity theorem" provides a theoretical guarantee for the closeness of these two kinds of optimal solutions. The main result in this paper is a proximity theorem for the problem (MC), stating that the L_∞ distance between optimal solutions to (MC) and to its continuous relaxation ($\overline{\mathsf{MC}}$) is bounded by $n-1$.

Theorem 1.3 (Proximity theorem for (MC) with respect to the L_∞ distance).
(i) *For every optimal solution* $y_* \in \mathbf{Z}^n$ *to* (MC)*, there exists an optimal solution* $x_* \in \mathbf{R}^n$ *to* ($\overline{\mathsf{MC}}$) *such that* $\|x_* - y_*\|_\infty < n-1$*.*
(ii) *For every optimal solution* $x_* \in \mathbf{R}^n$ *to* ($\overline{\mathsf{MC}}$)*, there exists an optimal solution* $y_* \in \mathbf{Z}^n$ *to* (MC) *such that* $\|y_* - x_*\|_\infty < n-1$*.*

The proof of this theorem is given in Section 4. This theorem implies, in particular, that there exists an optimal solution to (MC) if and only if there exists an optimal solution to ($\overline{\mathsf{MC}}$). In Section 4 we also give an example to show that the bound $n-1$ in Theorem 1.3 is the best possible, even for a very simple special case.

We then give some proximity results with respect to the L_1 distance. As an immediate corollary of Theorem 1.3, we obtain the following result for (MC):

Corollary 1.4 (Proximity theorem for (MC) with respect to the L_1 distance).
(i) *For every optimal solution* $y_* \in \mathbf{Z}^n$ *to* (MC)*, there exists an optimal solution* $x_* \in \mathbf{R}^n$ *to* ($\overline{\mathsf{MC}}$) *such that* $\|x_* - y_*\|_1 < n(n-1)$*.*
(ii) *For every optimal solution* $x_* \in \mathbf{R}^n$ *to* ($\overline{\mathsf{MC}}$)*, there exists an optimal solution* $y_* \in \mathbf{Z}^n$ *to* (MC) *such that* $\|y_* - x_*\|_1 < n(n-1)$*.*

We show that the bound $n(n-1)$ for the general case can be reduced to $2(n-1)$ for the two special cases (SC) and ($\mathsf{Laminar}$). We denote by ($\overline{\mathsf{SC}}$) and ($\overline{\mathsf{Laminar}}$) the continuous relaxation of (SC) and ($\mathsf{Laminar}$), respectively.

Theorem 1.5 (Proximity theorem for (SC) with respect to the L_1 distance).
(i) *For every optimal solution* $y_* \in \mathbf{Z}^n$ *to* (SC)*, there exists an optimal solution* $x_* \in \mathbf{R}^n$ *to* ($\overline{\mathsf{SC}}$) *such that* $\|x_* - y_*\|_1 < 2(n-1)$*.*
(ii) *For every optimal solution* $x_* \in \mathbf{R}^n$ *to* ($\overline{\mathsf{SC}}$)*, there exists an optimal solution* $y_* \in \mathbf{Z}^n$ *to* (SC) *such that* $\|y_* - x_*\|_1 < 2(n-1)$*.*

Theorem 1.6 (Proximity theorem for ($\mathsf{Laminar}$) with respect to the L_1 distance).
(i) *For every optimal solution* $y_* \in \mathbf{Z}^n$ *to* ($\mathsf{Laminar}$)*, there exists an optimal solution* $x_* \in \mathbf{R}^n$ *to* ($\overline{\mathsf{Laminar}}$) *such that* $\|x_* - y_*\|_1 < 2(n-1)$*.*
(ii) *For every optimal solution* $x_* \in \mathbf{R}^n$ *to* ($\overline{\mathsf{Laminar}}$)*, there exists an optimal solution* $y_* \in \mathbf{Z}^n$ *to* ($\mathsf{Laminar}$) *such that* $\|y_* - x_*\|_1 < 2(n-1)$*.*

The proofs of Theorems 1.5 and 1.6 are given in Section 4. We also show in Section 4 the tightness of the bound $2(n-1)$ in Theorems 1.5 and 1.6 by using a simple example. A proximity statement for the problem (SC) similar to Theorems 1.5 is already presented in [11]; we point out the incorrectness of the proximity statement in [11] by giving a counterexample in Section 2.

Finally, we propose an algorithm for (MC) based on continuous relaxation in Section 3. Our algorithm is a natural extension of the continuous relaxation algorithm of Hochbaum [11], where we use a new greedy-type algorithm for (MC). We analyze the time complexity of the proposed algorithm by using the proximity theorems shown in this paper and obtain the following result, where T_{relax} denotes the time required for solving the continuous relaxation $(\overline{\mathsf{MC}})$, and T_{round} denotes the time to round a given feasible solution $x \in \mathbf{R}^n$ of $(\overline{\mathsf{MC}})$ to a feasible solution $y \in \mathbf{Z}^n$ of (MC) satisfying $\|y - x\|_1 < n$. We note that $T_{\text{round}} = \mathrm{O}(n^2 \log L)$ (see [34]), where L is given by (1.1).

Theorem 1.7. *Our algorithm based on continuous relaxation finds an optimal solution to* (MC) *in* $\mathrm{O}(T_{\text{relax}} + T_{\text{round}} + n^3 F)$ *time.*

Remark 1.8. In fact, we do not need an exact optimal solution of the continuous relaxation $(\overline{\mathsf{MC}})$; it suffices to compute an "approximate" optimal solution $x_a \in \mathbf{R}^n$ of $(\overline{\mathsf{MC}})$ in the sense that $\|x_a - x_*\|_\infty \leq n$ holds for some optimal solution $x_* \in \mathbf{R}^n$ of $(\overline{\mathsf{MC}})$. We denote by $T_{\text{relax-apx}}$ the time required for computing such x. Usually, $T_{\text{relax-apx}}$ is smaller than T_{relax}. Then, Theorem 1.3 implies that there exists an optimal solution $y_* \in \mathbf{Z}^n$ of (MC) satisfying $\|y_* - x_a\|_\infty \leq \|y_* - x_*\|_\infty + \|x_* - x_a\|_\infty \leq 2n$. By using this bound, we can show in a similar way as Theorem 1.7 that (MC) can be solved in $\mathrm{O}(T_{\text{relax-apx}} + T_{\text{round}} + n^3 F)$ time.

Computation of an (approximate) optimal solution x_a of $(\overline{\mathsf{MC}})$ can be done by using similar algorithmic approaches as (MC). For example, we can apply the scaling approach for (MC) used in [35] to $(\overline{\mathsf{MC}})$ to obtain an $\mathrm{O}((n^3 + n^2 \log(L/n))(\log(L/n)/\log n)F)$-time algorithm for computing an approximate optimal solution, provided that the directional derivative of the objective function f can be computed in $\mathrm{O}(F)$ time. This implies that our continuous relaxation algorithm runs in $\mathrm{O}((n^3 + n^2 \log(L/n))(\log(L/n)/\log n)F)$ time. This time complexity bound is the same as the one of the previous best time complexity bound for (MC), i.e., the continuous relaxation approach does not lead to the reduction of the time complexity for the general case of (MC).

Although the continuous relaxation approach cannot possibly provide a faster algorithm for the general case of (MC), faster algorithms can be obtained for some special cases of (MC) by using the continuous relaxation approach. Indeed, we apply the continuous relaxation approach to (Laminar) and some special cases of (SC) with quadratic objective functions, in a similar way as in Hochbaum and Hong [12]. It is known that various classes of convex quadratic optimization problems in real variables can be solved in strongly polynomial time (see, e.g., [2, 12, 36]). Using this fact and also devising efficient implementations of the continuous relaxation approach, we show that (Laminar) and some special cases of (SC) with quadratic objective functions can be solved efficiently in strongly polynomial time.

The previous best time complexity bound for (Laminar) is $\mathrm{O}(n^3)$ by Tamir [36], which is also based on a continuous relaxation approach with a weaker proximity theorem in [9]. We present a better time complexity by using a refined proximity theorem (Theorem 1.6) and some algorithmic techniques.

Theorem 1.9. *The problem* (Laminar) *can be solved in* $\mathrm{O}(n^2)$ *time if the objective function is quadratic.*

Hochbaum and Hong [12] develop efficient algorithms for some special cases of ($\overline{\mathsf{SC}}$) with quadratic objective functions. Then, they state in their paper [12] that the corresponding special cases of (SC) can be solved efficiently in strongly polynomial time by using a proximity result in [11]. The proximity result, however, is incorrect, as pointed out in Section 2.4, and therefore the time complexity results for the special cases of (SC) are no longer valid.[1] We show that by using our proximity theorem (Theorem 1.5), the special cases of (SC) discussed in [12] can be solved in (almost) the same time complexity as stated in [12] (see Section 3.4 for details). In particular, our algorithms are the fastest for the special cases of (SC).

Organization The organization of this paper is as follows. In Section 2, we explain fundamental concepts related to submodular functions and M-convex functions, and give formulations of discrete convex optimization problems discussed in this paper. In Section 2, we also review the continuous relaxation approach for discrete convex optimization problems, including the resource allocation problems. In Section 3, we propose an algorithm for (MC) based on continuous relaxation, and apply it to (Laminar) and some special cases of (SC) with quadratic objective functions to obtain efficient algorithms. Finally, proofs of the proximity theorems are given in Section 4.

2 Preliminaries

2.1 Definitions and notation

Throughout the paper, let n be a positive integer with $n \geq 2$ and put $N = \{1, 2, \ldots, n\}$. We denote by $\mathbf{R}$ (resp., by $\mathbf{R}_+$) the sets of real numbers (resp., nonnegative real numbers). Similarly, we denote by $\mathbf{Z}$ (resp., by $\mathbf{Z}_+$) the sets of integers (resp., nonnegative integers).

Let $x = (x(1), x(2), \ldots, x(n)) \in \mathbf{R}^n$ be a vector. We denote

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

For a subset $Y \subseteq N$, we denote $x(Y) = \sum_{i \in Y} x(i)$. We define

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

The vectors $\lceil x \rceil, \lfloor x \rfloor \in \mathbf{Z}^n$ are given by

$$(\lceil x \rceil)(i) = \lceil x(i) \rceil, \qquad (\lfloor x \rfloor)(i) = \lfloor x(i) \rfloor \qquad (i \in N).$$

We define $\mathbf{1} = (1, 1, \ldots, 1) \in \mathbf{Z}^n$ and $\mathbf{0} = (0, 0, \ldots, 0) \in \mathbf{Z}^n$. For $Y \subseteq N$, we denote by $\chi_Y \in \{0, 1\}^n$ the characteristic vector of Y, i.e., $\chi_Y(i) = 1$ if $i \in Y$ and $\chi_Y(i) = 0$ otherwise. In particular, we denote $\chi_i = \chi_{\{i\}}$ for every $i \in N$. Inequalities and equalities for vectors $x, y \in \mathbf{R}^n$ mean component-wise inequalities and equalities; for example, $x \leq y$ reads $x(i) \leq y(i)$ for all $i \in N$. For a nonempty set $S \subseteq \mathbf{R}^n$, the *closed convex hull* (or *convex closure*) of S is the (uniquely determined) smallest closed convex set containing S.

[1]The time complexity results for special cases of ($\overline{\mathsf{SC}}$) in [12] are independent of the incorrect proximity result in [11] and therefore remains valid.

Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a function. The *effective domain* $\mathrm{dom}_{\mathbf{R}} f$ of f is defined by $\mathrm{dom}_{\mathbf{R}} f = \{x \in \mathbf{R}^n \mid f(x) < +\infty\}$. 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. We denote the set of minimizers of f by

$$\arg\min_{\mathbf{R}} f = \{x \in \mathbf{R}^n \mid f(x) \leq f(y) \ (\forall y \in \mathbf{R}^n)\}.$$

Note that for a closed proper convex function f, the set $\arg\min_{\mathbf{R}} f$ is a closed set.

For a function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ defined on the integer lattice points, we define the effective domain $\mathrm{dom}_{\mathbf{Z}} g$ and the set of minimizers $\arg\min_{\mathbf{Z}} g$ by

$$\begin{aligned} \mathrm{dom}_{\mathbf{Z}} g &= \{x \in \mathbf{Z}^n \mid g(x) < +\infty\}, \\ \arg\min_{\mathbf{Z}} g &= \{x \in \mathbf{Z}^n \mid g(x) \leq g(y) \ (\forall y \in \mathbf{Z}^n)\}. \end{aligned}$$

2.2 Base polyhedra, M-convex sets, and M-convex functions

In this section we explain the concepts of base polyhedra, M-convex sets, and M-convex functions. We refer to [6, 24] for comprehensive treatment of these concepts.

A set function $\rho : 2^N \to \mathbf{R} \cup \{+\infty\}$ is said to be *submodular* if it satisfies the *submodular inequality*:

$$\rho(X) + \rho(Y) \geq \rho(X \cap Y) + \rho(X \cup Y) \qquad (\forall X, Y \in 2^N).$$

For a submodular function $\rho : 2^N \to \mathbf{R} \cup \{+\infty\}$ with $\rho(\emptyset) = 0$ and $\rho(N) < +\infty$, we consider a polyhedron

$$B(\rho) = \{x \in \mathbf{R}^n \mid x(Y) \leq \rho(Y) \ (\forall Y \in 2^N), \ x(N) = \rho(N)\},$$

which is called the *base polyhedron* associated with ρ. It is known that if ρ is an integer-valued function, then $B(\rho)$ is an integral polyhedron, i.e., the closed convex hull of the set of integral points in $B(\rho)$ coincides with $B(\rho)$. For an integer-valued submodular function $\rho : 2^N \to \mathbf{Z} \cup \{+\infty\}$ with $\rho(\emptyset) = 0$ and $\rho(N) < +\infty$, we consider the set $S = B(\rho) \cap \mathbf{Z}^n$ of integral vectors in a base polyhedron, which we call an *M-convex set*.

Base polyhedra and M-convex sets can be characterized by the following exchange properties (see, e.g., [4], [6, Corollary 20.6], [24, Section 4], [30, Theorem 3.1]).

Theorem 2.1.
(i) *A nonempty closed set $S \subseteq \mathbf{R}^n$ is a base polyhedron if and only if it satisfies* (B-EXC[**R**])*:*

(B-EXC[R]) $\forall x, y \in S$, $\forall i \in \mathrm{supp}^+(x-y)$, $\exists j \in \mathrm{supp}^-(x-y)$, $\exists \alpha_0 > 0$:

$$x - \alpha(\chi_i - \chi_j) \in S, \quad y + \alpha(\chi_i - \chi_j) \in S \qquad (\forall \alpha \in [0, \alpha_0]).$$

(ii) *A nonempty set $S \subseteq \mathbf{Z}^n$ is M-convex if and only if it satisfies* (B-EXC[**Z**])*:*

(B-EXC[Z]) $\forall x, y \in S$, $\forall i \in \mathrm{supp}^+(x - y)$, $\exists j \in \mathrm{supp}^-(x - y)$:

$$x - \chi_i + \chi_j \in S, \qquad y + \chi_i - \chi_j \in S.$$

An M-convex set S with $S \subseteq \{0,1\}^n$ is essentially equivalent to the concept of matroid in the following sense: a nonempty set $S \subseteq \{0,1\}^n$ is M-convex if and only if $S = \{\chi_Y \in \{0,1\}^n \mid Y \in \mathcal{B}\}$ holds for the base family $\mathcal{B} \subseteq 2^N$ of some matroid $\mathcal{M} = (N, \mathcal{B})$.

We consider a function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ defined on integer lattice points. A function g with nonempty $\mathrm{dom}_{\mathbf{Z}}\, g$ is said to be *M-convex* if it satisfies (M-EXC[$\mathbf{Z}$]):

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

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

From the definition and Theorem 2.1 (ii), the effective domain $\mathrm{dom}_{\mathbf{Z}}\, g$ of an M-convex function g is an M-convex set. In particular, there exists some $r \in \mathbf{Z}$ such that $\mathrm{dom}_{\mathbf{Z}}\, g$ is contained in the hyperplane $\{x \in \mathbf{Z}^n \mid x(N) = r\}$. A set $S \subseteq \mathbf{Z}^n$ is M-convex if and only if its indicator function $\delta_S : \mathbf{Z}^n \to \{0, +\infty\}$ defined by

$$\delta_S(x) = \begin{cases} 0 & (x \in S), \\ +\infty & (\text{otherwise}), \end{cases}$$

is M-convex.

Remark 2.2. The concept of M-convex function is deeply related to submodularity/supermodularity in the following sense (see also Lemma 4.2 in Section 4).

Since the effective domain of an M-convex function is contained in a hyperplane $\{x \in \mathbf{Z}^n \mid x(N) = r\}$ for some $r \in \mathbf{Z}$, we may consider the projection $g' : \mathbf{Z}^{n-1} \to \mathbf{R} \cup \{+\infty\}$ of an M-convex function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ along an arbitrarily chosen coordinate axis $k \in N$, where the projection g' of g is a function in $n-1$ variables defined by

$$\begin{aligned} & g'(x(1), \ldots, x(k-1), x(k+1), \ldots, x(n)) \\ = \quad & g(x(1), \ldots, x(k-1), r - x(N \setminus \{k\}), x(k+1), \ldots, x(n)). \end{aligned}$$

A function g' obtained by the projection of an M-convex function is called an *M$^\natural$-convex function* [27]. It is known that an M$^\natural$-convex function is a supermodular function on the integer lattice points, i.e., it satisfies the following inequality [24, Theorem 6.19]:

$$g(x) + g(y) \leq g(x \vee y) + g(x \wedge y) \qquad (\forall x, y \in \mathbf{Z}^n),$$

where for $x, y \in \mathbf{R}^n$ the vectors $x \vee y \in \mathbf{R}^n$ and $x \wedge y \in \mathbf{R}^n$ are defined by

$$(x \vee y)(i) = \max\{x(i), y(i)\}, \quad (x \wedge y)(i) = \min\{x(i), y(i)\} \quad (i \in N).$$

We note that the supermodular inequality is void for an M-convex function g because $x \vee y, x \wedge y \in \mathrm{dom}_{\mathbf{Z}}\, g$ occurs only when $x = y \in \mathrm{dom}_{\mathbf{Z}}\, g$. $\square$

The concept of M-convex function is extended to functions defined on the real space $\mathbf{R}^n$. A function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ with nonempty $\mathrm{dom}_{\mathbf{R}}\, f$ is said to be *M-convex* if it is convex and satisfies (M-EXC[$\mathbf{R}$]):

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

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

An M-convex function is said to be a *closed proper M-convex function* if it is a closed proper convex function, in addition. From the definition and Theorem 2.1 (i), the effective domain $\mathrm{dom}_{\mathbf{R}}\, f$ of an M-convex function f is a base polyhedron if $\mathrm{dom}_{\mathbf{R}}\, f$ is a closed set. In particular, there exists some $r \in \mathbf{R}$ such that $\mathrm{dom}_{\mathbf{R}}\, f$ is contained in the hyperplane $\{x \in \mathbf{R}^n \mid x(N) = r\}$. A set $S \subseteq \mathbf{R}^n$ is a base polyhedron if and only if its indicator function $\delta_S : \mathbf{R}^n \to \{0, +\infty\}$ defined by

$$\delta_S(x) = \begin{cases} 0 & (x \in S), \\ +\infty & (\text{otherwise}), \end{cases}$$

is a closed proper M-convex function.

2.3 M-convex function minimization and resource allocation problems

We explain the formulations of the problem of M-convex function minimization and resource allocation problems discussed in this paper. In addition, we explain the continuous relaxation of each problem.

M-convex function minimization The problem of M-convex function minimization is formulated as follows:

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

where $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ is an M-convex function. The following property is useful in defining a continuous relaxation of (MC).

Theorem 2.3 (cf. [24, Section 6.11])**.** *For every M-convex function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$, there exists a closed proper M-convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ such that $f(x) = g(x)$ for all $x \in \mathbf{Z}^n$ and* $\mathrm{dom}_{\mathbf{R}}\, f$ *is the closed convex hull of* $\mathrm{dom}_{\mathbf{Z}}\, g$.

Based on this property, we define a continuous relaxation $(\overline{\mathsf{MC}})$ of (MC) by

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

where $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is a closed proper M-convex function such that $f(x) = g(x)$ for all $x \in \mathbf{Z}^n$ and $\mathrm{dom}_{\mathbf{R}}\, f$ is the closed convex hull of $\mathrm{dom}_{\mathbf{Z}}\, g$.

We then explain two important special cases of (MC): the laminar convex resource allocation problem $(\mathsf{Laminar})$ and the resource allocation problem under a submodular constraint (SC).

Laminar convex resource allocation problem The laminar convex resource allocation problem is formulated as follows (see [18], [24, Section 6.3]):

$$(\mathsf{Laminar}) \quad \left| \begin{array}{ll} \text{Minimize} & \displaystyle\sum_{Y \in \mathcal{F}} f_Y(x(Y)) \\ \text{subject to} & x(N) = K, \ \ell_Y \le x(Y) \le u_Y \ (Y \in \mathcal{F}), \\ & x \ge \mathbf{0}, \ x \in \mathbf{Z}^n, \end{array} \right.$$

where $\mathcal{F} \subseteq 2^N$ is a laminar family, $f_Y : \mathbf{R} \to \mathbf{R}$ $(Y \in \mathcal{F})$ is a family of univariate convex functions, $K \in \mathbf{Z}_+$, and $\ell_Y, u_Y \in \mathbf{Z}_+$ $(Y \in \mathcal{F})$. Note that $|\mathcal{F}| = \mathrm{O}(n)$. Throughout the paper we assume that the problem $(\mathsf{Laminar})$ has a feasible solution. The continuous relaxation $(\overline{\mathsf{Laminar}})$

of (Laminar) is obtained by removing the integrality constraint "$x \in \mathbf{Z}^n$" from (Laminar). We assume, without loss of generality, that

$$\emptyset \notin \mathcal{F}, \quad N \in \mathcal{F}, \quad \{i\} \in \mathcal{F} \ (\forall i \in N). \tag{2.2}$$

It is known that the problem (Laminar) can be reformulated as a convex cost flow problem on a tree network, as explained below.

For $Y \in \mathcal{F} \setminus \{N\}$, we call $X \in \mathcal{F}$ the *parent* of Y if X is the unique minimal set in $\mathcal{F}$ which properly contains Y (i.e., $X \supset Y$ and $X \neq Y$), and denote by $p(Y)$ the parent of Y. For $X \in \mathcal{F}$ with $|X| \geq 2$, we call $Y \in \mathcal{F}$ a *child* of X if $X = p(Y)$. The condition (2.2) implies that every $X \in \mathcal{F}$ with $|X| \geq 2$ has at least one child. We consider a (directed) tree graph $T = (V, A)$ defined by

$$V = \{v_Y \mid Y \in \mathcal{F}\}, \qquad A = \{(v_X, v_Y) \mid X, Y \in \mathcal{F},\ Y \text{ is a child of } X\}.$$

Hence, the node v_N is the root node of T, while nodes corresponding to the singleton sets are leaf nodes of T. In the convex cost flow formulation, we send flow from the root node v_N to leaf nodes. That is, v_N is the source node and leaf nodes are sink nodes. We use a flow variable $\varphi(i, j)$ for each arc $(i, j) \in A$, and consider the integrality constraint on flow variables

$$\varphi(i, j) \in \mathbf{Z} \qquad (\forall (i, j) \in A)$$

and flow balance constraints

$$\begin{aligned} \sum \{\varphi(v_N, v_Y) \mid v_Y \in V,\ (v_N, v_Y) \in A\} &= K, \\ \sum \{\varphi(v_X, v_Y) \mid v_Y \in V, (v_X, v_Y) \in A\} &= \varphi(v_{p(X)}, v_X) \qquad (\forall v_X \in V,\ v_X \neq v_N). \end{aligned}$$

We see from the flow balance constraints that the value of the flow variable $\varphi(v_{p(Y)}, v_Y)$ corresponds to the value $x(Y)$ for every $Y \in \mathcal{F} \setminus \{N\}$. For each arc $(v_{p(Y)}, v_Y) \in A$, the convex cost function of the arc is given by f_Y and the lower and upper capacity of the arc are ℓ_Y and u_Y, respectively. It is easy to see that the problem (Laminar) is equivalent to the convex cost flow problem defined above, and ($\overline{\text{Laminar}}$) is equivalent to the continuous relaxation of the convex cost flow problem obtained by removing the integrality constraint on flow variables.

Resource allocation problem under a submodular constraint The resource allocation problem under a submodular constraint (SC) is formulated as follows [6, 11, 14, 16]:

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

where $f_i : \mathbf{R} \to \mathbf{R}$ $(i \in N)$ is a family of univariate convex functions, and $\rho : 2^N \to \mathbf{Z}_+ \cup \{+\infty\}$ is a nonnegative-valued submodular function satisfying $\rho(\emptyset) = 0$ and $\rho(N) < +\infty$.

We also consider special cases of (SC) discussed in [11, 12]. The simple resource allocation problem (Simple) reads as

$$\text{(Simple)} \quad \text{Minimize} \quad \sum_{i=1}^{n} f_i(x(i)) \quad \text{subject to} \quad x(N) = K,\ \mathbf{0} \leq x \leq u,\ x \in \mathbf{Z}^n,$$

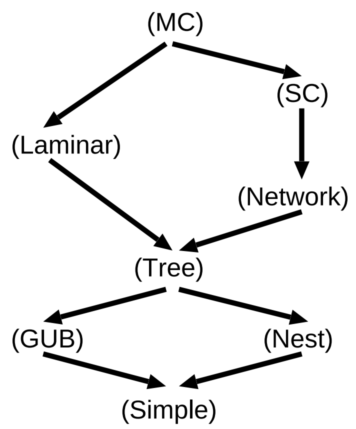

Figure 1: Relationship among the discrete convex optimization problems

where $f_i : \mathbf{R} \to \mathbf{R}$ $(i \in N)$ is a family of univariate convex functions, $K \in \mathbf{Z}_+$, and $u \in \mathbf{Z}^n_+$.

The other special cases of (SC) in [11, 12] can be obtained by adding some constraints to (Simple). The generalized upper bound resource allocation problem (GUB) is obtained by adding to (Simple) the constraints $x(S_t) = u_{S_t}$ $(t = 1, 2, \ldots, m)$, where $\{S_1, S_2, \ldots, S_m\}$ is a partition of N, and $u_{S_t} \in \mathbf{Z}_+$ $(t = 1, 2, \ldots, m)$. The nested resource allocation problem (Nest) is obtained by adding to (Simple) the constraints $x(S_t) = u_{S_t}$ $(t = 1, 2, \ldots, m)$, where $\{S_1, S_2, \ldots, S_m\}$ is a family of nested subsets of N, i.e., it satisfies the condition $\emptyset \neq S_1 \subset S_2 \subset \cdots \subset S_m \subset N$, and $u_{S_t} \in \mathbf{Z}_+$ $(t = 1, 2, \ldots, m)$. The tree resource allocation problem (Tree) is obtained by adding to (Simple) the constraints $x(Y) = u_Y$ $(Y \in \mathcal{F})$, where $\mathcal{F}(\subseteq 2^N)$ is a laminar family and $u_Y \in \mathbf{Z}_+$ $(Y \in \mathcal{F})$. Hence, (Tree) is also a special case of (Laminar).

The network resource allocation problem (Network) is defined by using a network with a single source and multiple sinks. Given a directed graph $G = (V, A)$ with node set V and arc set A, let $s \in V$ be the unique source node and $N = \{1, 2, \ldots, n\}$ $(\subseteq V)$ be the set of sink nodes. The supply of the source is given by $K \in \mathbf{Z}_+$, and the capacity of each arc $(i, j) \in A$ is given by $c(i, j) \in \mathbf{Z}_+$. Using flow variables $\{\varphi(i, j) \mid (i, j) \in A\}$ in addition to variables $\{x(i) \mid i \in N\}$, the constraints of the problem (Network) is described as follows:

$$
\begin{aligned}
&\sum\{\varphi(i,j) \mid j \in V,\ (i,j) \in A\} - \sum\{\varphi(j,i) \mid j \in V,\ (j,i) \in A\} = 0 \quad (\forall i \in V \setminus (\{s\} \cup N)),\\
&\sum\{\varphi(s,j) \mid j \in V,\ (s,j) \in A\} - \sum\{\varphi(j,s) \mid j \in V,\ (j,s) \in A\} = K,\\
&\sum\{\varphi(i,j) \mid j \in V,\ (i,j) \in A\} - \sum\{\varphi(j,i) \mid j \in V,\ (j,i) \in A\} = -x(i) \quad (\forall i \in N),\\
&0 \leq \varphi(i,j) \leq c(i,j)\ (\forall (i,j) \in A),\ \mathbf{0} \leq x \leq u,\ x \in \mathbf{Z}^n.
\end{aligned}
$$

It is easy to see that the constraints above imply $x(N) = K$. The relationship among the discrete convex optimization problems explained above is summarized in Figure 1.

For the problems (SC), (Simple), (GUB), (Nest), (Tree), and (Network), we denote by $(\overline{\mathsf{SC}})$, $(\overline{\mathsf{Simple}})$, $(\overline{\mathsf{GUB}})$, $(\overline{\mathsf{Nest}})$, $(\overline{\mathsf{Tree}})$, and $(\overline{\mathsf{Network}})$, respectively, the continuous relaxation which can be obtained by removing the integrality constraint "$x \in \mathbf{Z}^n$."

2.4 Review of continuous relaxation approach for discrete convex optimization problems

In this section we review the existing results on continuous relaxation approach for M-convex function minimization, resource allocation problems, and related problems.

Simple resource allocation problem Weinsten and Yu [38] (see also [14, Section 4.6]) propose an algorithm for (Simple) based on continuous relaxation, and the time complexity of the algorithm is $\mathrm{O}(T_{\overline{\mathsf{Simple}}}(n) + n \log n)$, where $T_{\overline{\mathsf{Simple}}}(n)$ denotes the time complexity for solving the continuous relaxation ($\overline{\mathsf{Simple}}$) with n variables. A proximity theorem for (Simple) is implicit in [38], which states that for every optimal solution $x_* \in \mathbf{R}^n$ of ($\overline{\mathsf{Simple}}$), there exists an optimal solution $y_* \in \mathbf{Z}^n$ of (Simple) satisfying $\|y_* - x_*\|_1 = \mathrm{O}(n)$. Later, Ibaraki and Katoh [14, Section 4.6] improve the algorithm of Weinsten and Yu [38] by using the technique of Frederickson and Johnson [5] so that it runs in $\mathrm{O}(T_{\overline{\mathsf{Simple}}}(n) + n)$ time. If the objective function is quadratic, then ($\overline{\mathsf{Simple}}$) can be solved in linear time by the algorithm by Brucker [2], and therefore (Simple) can be solved in $\mathrm{O}(n)$ time by the algorithm of Ibaraki and Katoh [14, Section 4.6].

Hochbaum [11] shows that the problem (GUB) can be reduced to the problem (Simple) by solving a "disjoint" family of the problems (Simple) which can be obtained from the given instance of (GUB). It follows from this observation that the problem (GUB) with quadratic objective function can be solved in $\mathrm{O}(n)$ time as well.

Separable convex function minimization under linear constraints Hochbaum and Shanthikumar [13] apply continuous relaxation approach to the following separable convex function minimization problem under linear constraints and integrality constraint:

$$(\mathsf{IP}) \quad \text{Minimize} \quad \sum_{i=1}^{n} f_i(x(i)) \quad \text{subject to} \quad Ax \geq b, \quad x \in \mathbf{Z}^n,$$

where $f_i : \mathbf{R} \to \mathbf{R}$ ($i \in N$) is a family of univariate convex functions, A is an integral $m \times n$ matrix, and $b \in \mathbf{Z}^m$. We denote by $\Delta \in \mathbf{Z}_+$ the maximum absolute value of a subdeterminant of the matrix A. For this problem the continuous relaxation ($\overline{\mathsf{IP}}$) can be easily obtained by removing the integrality constraint "$x \in \mathbf{Z}^n$." It is easy to see that each of the problems (Simple), (GUB), (Nest), (Tree), (Network), and (Laminar) can be formulated as the problem (IP) with $\Delta = 1$, while $\Delta = \mathrm{O}(2^n)$ for the problem (SC).

For the problem (IP) the following proximity theorem is known:

Theorem 2.4 ([13, Theorem 3.3]).
(i) *For every optimal solution $y_* \in \mathbf{Z}^n$ to* (IP)*, there exists an optimal solution $x_* \in \mathbf{R}^n$ to* ($\overline{\mathsf{IP}}$) *such that $\|x_* - y_*\|_\infty \leq n\Delta$.*
(ii) *For every optimal solution $x_* \in \mathbf{R}^n$ to* ($\overline{\mathsf{IP}}$)*, there exists an optimal solution $y_* \in \mathbf{Z}^n$ to* (IP) *such that $\|y_* - x_*\|_\infty \leq n\Delta$.*

Using this proximity theorem an efficient algorithm for (IP) is devised in [13]. We note that Granot and Skorin-Kapov [9] discuss the special case of (IP) with a separable quadratic objective function and obtain a similar proximity result.

Resource allocation problem under a submodular constraint A continuous relaxation approach for the problem (SC) is proposed by Hochbaum [11], where the following "proximity theorem" is stated:

Statement A (Corollary 4.3 in [11])
(i) For every optimal solution $y_* \in \mathbf{Z}^n$ to (SC), there exists some optimal solution $x_* \in \mathbf{R}^n$ to $(\overline{\mathsf{SC}})$ such that $y_* - \mathbf{1} < x_* < y_* + n\mathbf{1}$.
(ii) For every optimal solution $x_* \in \mathbf{R}^n$ to $(\overline{\mathsf{SC}})$, there exists some optimal solution $y_* \in \mathbf{Z}^n$ to (SC) such that $y_* - \mathbf{1} < x_* < y_* + n\mathbf{1}$.

In particular, the claim (ii) in Statement A implies that the vector $u = x_* + \mathbf{1}$ is an upper bound of some optimal solution to (SC). This observation is used in [12] to devise efficient algorithms for special cases of (SC) with quadratic objective functions.

Statement A, however, is incorrect; indeed, Example 2.5 below shows that Statement A does not hold even for the problem (Simple) with a quadratic objective function, which is a very special case of (SC).

Example 2.5. Let δ be a sufficiently small positive number with $\delta < 1$. We consider the problem (Simple), where $K = n - 1$, $u(i) = +\infty$ $(i \in N)$, and

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

It is noted that $f_i(1) - f_i(0) = 2\delta > \delta$ for $i = 2, 3, \ldots, n$. Hence, an optimal solution $y_* \in \mathbf{Z}^n$ to the problem (Simple) is uniquely given by $y_* = (n-1, 0, \ldots, 0)$. On the other hand, for $i = 2, 3, \ldots, n$ the slope of the function f_i is equal to δ if $\alpha = (1-\delta)/2$ and more than δ if $\alpha > (1-\delta)/2$. Therefore, an optimal solution $x_* \in \mathbf{R}^n$ to the continuous relaxation $(\overline{\mathsf{Simple}})$ is uniquely given by

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

Since δ is a sufficiently small positive number, we have

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

This implies that if $n \geq 4$ then $y_*(1) - x_*(1) > 1$ holds, and therefore the inequality $y_* - \mathbf{1} < x_*$ does not hold. $\square$

The continuous relaxation approach in [11] is applied to special cases of (SC) with quadratic objective functions in Hochbaum and Hong [12]. To develop efficient algorithms based on continuous relaxation, Hochbaum and Hong [12] propose fast algorithms for the problems $(\overline{\mathsf{Nest}})$, $(\overline{\mathsf{Tree}})$, and $(\overline{\mathsf{Network}})$ with quadratic objective functions.

Theorem 2.6 ([12]).
(i) $(\overline{\mathsf{Nest}})$ *and* $(\overline{\mathsf{Tree}})$ *can be solved in* $\mathrm{O}(n \log n)$ *time if objective functions are quadratic.*
(ii) $(\overline{\mathsf{Network}})$ *can be solved in* $\mathrm{O}(|V||A| \log(|V|^2/|A|))$ *time if the objective function is quadratic.*

Based on these results and Statement A above, efficient algorithms for (Nest), (Tree), and (Network) with quadratic objective functions are developed [12], and the time complexity of each algorithm is $\mathrm{O}(n \log n)$ for (Nest), $\mathrm{O}(n \log n)$ for (Tree), and $\mathrm{O}(|V||A| \log(|V|^2/|A|))$ for (Network). These algorithms for (Nest), (Tree), and (Network) are, however, no longer valid since Statement A is incorrect.

Instead of Statement A, we can use the following property which is an immediate corollary of Theorem 2.4 (ii). Recall that $\Delta = 1$ holds in Theorem 2.4 for the problems (Nest), (Tree), and (Network).

Proposition 2.7. *Let $x_* \in \mathbf{R}^n$ be an optimal solution to $(\overline{\mathsf{Nest}})$ ($(\overline{\mathsf{Tree}})$ and $(\overline{\mathsf{Network}})$, respectively). Then, there exists an optimal solution $y_* \in \mathbf{Z}^n$ to (Nest) ((Tree) and (Network), respectively) satisfying $y_* \geq x_* - n\mathbf{1}$ and $\|y_* - (x_* - n\mathbf{1})\|_1 = \mathrm{O}(n^2)$.*

This property shows that the vector $\ell = x_* - n\mathbf{1}$ can be used as a lower bound of an optimal solution to (Nest), (Tree), and (Network). Using this fact and a similar technique as in [12] we can show that (Nest), (Tree), and (Network) can be solved in time $\mathrm{O}(n^2 \log n)$, $\mathrm{O}(n^2 \log n)$, and $\mathrm{O}(n^2|A| + |V||A| \log(|V|^2/|A|))$, respectively, which are worse than the bounds shown in [12] by a factor of n.

In this paper, we will prove an alternative proximity theorem for (SC) (Theorem 1.5), and devise efficient algorithms for (Nest), (Tree), and (Network) based on continuous relaxation which have better time complexity than those mentioned above.

Remark 2.8. Example 2.5 shows that for an optimal solution x_* to the continuous relaxation $(\overline{\mathsf{Simple}})$, the vector $\lceil x_* \rceil$ cannot be an upper bound of any optimal solution to the the original problem (Simple). The following example shows that the vector $\lfloor x_* \rfloor$ cannot be a lower bound of any optimal solution to (Simple).

Example 2.9. Let δ be an arbitrarily chosen small positive number with $\delta < 1$ and put $\eta = 3\delta(1-\delta) - \delta = 2\delta - 3\delta^2$ (> 0). We consider the problem (Simple), where $K = n - 1$, $u(i) = +\infty$ ($i \in N$), and

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

For $i = 2, 3, \ldots, n$, it holds that

$$\begin{aligned} f_i(\alpha + 2) - f_i(\alpha + 1) = 3\delta > \delta = f_i(1) - f_i(0) \qquad (\forall \alpha \in \mathbf{Z}_+), \\ f_i(1) - f_i(0) = \delta < 2\delta = f_1(\alpha + 1) - f_1(\alpha) \qquad (\forall \alpha \in \mathbf{Z}_+). \end{aligned}$$

These inequalities imply that the vector $y_* = (0, 1, \ldots, 1)$ is the unique optimal solution to (Simple). On the other hand, for $i = 2, 3, \ldots, n$, the slope $-\eta/\delta$ of the function f_i in the interval $[0, \delta]$ is strictly smaller than the slope 2δ of the function f_1, and the slope 3δ of the function f_i in the interval $[\delta, +\infty)$ is strictly larger than 2δ. Hence, an optimal solution $x_* \in \mathbf{R}^n$ to the continuous relaxation $(\overline{\mathsf{Simple}})$ is uniquely given as $x_* = ((n-1)(1-\delta), \delta, \ldots, \delta)$. If $n \geq 3$ then the inequality $y_* \geq \lfloor x_* \rfloor$ does not hold.

On the other hand, the following weaker statement holds for the problem (Simple) [17, Theorem 2] (see also [14, Section 4.6]):

For every optimal solution $x_* \in \mathbf{R}^n$ to ($\overline{\mathsf{Simple}}$), there exists some optimal solution $y_* \in \mathbf{Z}^n$ to (Simple) satisfying either $y_* \leq \lceil x_* \rceil$ or $y_* \geq \lfloor x_* \rfloor$ (or both).

$\square$

L-convex function minimization Continuous relaxation approach is also applied to L-convex function minimization. The concept of L-convex function is introduced by Murota [22] as a class of discrete convex functions on integer lattice points. L-convex functions and M-convex functions form two distinct classes of discrete convex functions that are conjugate to each other under the Legendre-Fenchel transformation (see [24] for details).

A function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ is said to be *L-convex* if it satisfies the following properties:

(LF1[Z]) g is submodular, i.e., $g(x)+g(y) \geq g(x \vee y)+g(x \wedge y)$ for all $x, y \in \mathrm{dom}_{\mathbf{Z}}\, g$,
(LF2[Z]) $\exists r \in \mathbf{R}$ such that $g(x + \mathbf{1}) = g(x) + r$ for all $x \in \mathrm{dom}_{\mathbf{Z}}\, g$.

When we discuss the minimization of an L-convex function, we consider the case where the value r in (LF2[**Z**]) is zero, since otherwise a minimizer of g does not exist. A typical example of an L-convex function is a function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ defined by

$$g(x) = \sum_{i,j \in N,\ i \neq j} g_{ij}(x(i) - x(j)) \qquad (x \in \mathbf{Z}^n),$$

where $g_{ij} : \mathbf{R} \to \mathbf{R} \cup \{+\infty\}$ $(i, j \in N,\ i \neq j)$ is a family of univariate convex functions.

We consider the minimization of an L-convex function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$. To define a continuous relaxation of L-convex function minimization, we use the concept of closed proper L-convex function; a closed proper convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is said to be *L-convex* if it satisfies the following properties:

(LF1[R]) f is submodular, i.e., $f(x) + f(y) \geq f(x \vee y) + f(x \wedge y)$ for all $x, y \in \mathrm{dom}_{\mathbf{R}}\, f$,
(LF2[R]) $\exists r \in \mathbf{R}$ such that $f(x + \lambda \mathbf{1}) = f(x) + \lambda r$ for all $x \in \mathrm{dom}_{\mathbf{R}}\, f$ and $\lambda \in \mathbf{R}$.

The restriction of a closed proper L-convex function to integer lattice points is an L-convex function.

Proposition 2.10 ([20]). *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper L-convex function. For any positive real number α, define a function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ by $g(x) = f(\alpha x)$ $(x \in \mathbf{Z}^n)$. Then, g is an L-convex function on the integer lattice points, i.e., satisfies* (LF1[**Z**]) *and* (LF2[**Z**]).

On the other hand, it is known that for every L-convex function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ on integer lattice points, there exists a closed proper L-convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ such that $f(x) = g(x)$ for all $x \in \mathbf{Z}^n$ [24, 28]. Hence, the minimization of such a closed proper L-convex function f can be used as a continuous relaxation for the minimization of an L-convex function g.

The following proximity result is shown for L-convex function minimization.

Theorem 2.11 ([20]). *Let $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ be an L-convex function, and $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper L-convex function such that $f(x) = g(x)$ for all $x \in \mathbf{Z}^n$.*
(i) *For every $y_* \in \arg\min_{\mathbf{Z}} g$, there exists some $x_* \in \arg\min_{\mathbf{R}} f$ such that $\|x_* - y_*\|_\infty < n-1$.*
(ii) *For every $x_* \in \arg\min_{\mathbf{R}} f$, there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ such that $\|y_* - x_*\|_\infty < n-1$.*

This theorem is shown by using Proposition 2.10 and the following proximity result concerning a scaled problem (see [24, Theorem 7.18]).

Theorem 2.12. *Let $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ be an L-convex function. Let α be a positive integer, and define $g_\alpha : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ by $g_\alpha(x) = g(\alpha x)$ $(x \in \mathbf{Z}^n)$. Assume that $\mathrm{dom}_{\mathbf{Z}}\, g_\alpha \neq \emptyset$.*
(i) *For every $y_* \in \arg\min_{\mathbf{Z}} g$, there exists some $x_* \in \arg\min_{\mathbf{Z}} g_\alpha$ such that $\|\alpha x_* - y_*\|_\infty \leq (n-1)(\alpha-1)$.*
(ii) *For every $x_* \in \arg\min_{\mathbf{Z}} g_\alpha$, there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ such that $\|y_* - \alpha x_*\|_\infty \leq (n-1)(\alpha-1)$.*

Note that neither (MC) nor (SC) is a special case of L-convex function minimization, and therefore Theorem 2.11 cannot be applied.

M-convex function minimization Although there is no proximity result concerning the continuous relaxation of (MC) so far, the following proximity theorem concerning the scaled problem of (MC) is shown in [18, Theorem 3.4] (see also [24, Theorem 6.37]):

Theorem 2.13. *Let $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ be an M-convex function. Let α be a positive integer, and define $g_\alpha : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ by $g_\alpha(x) = g(\alpha x)$ $(x \in \mathbf{Z}^n)$. Assume that $\mathrm{dom}_{\mathbf{Z}}\, g_\alpha \neq \emptyset$.*
(i) *For every $y_* \in \arg\min_{\mathbf{Z}} g$, there exists some $x_* \in \arg\min_{\mathbf{Z}} g_\alpha$ such that $\|\alpha x_* - y_*\|_\infty \leq (n-1)(\alpha-1)$.*
(ii) *For every $x_* \in \arg\min_{\mathbf{Z}} g_\alpha$, there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ such that $\|y_* - \alpha x_*\|_\infty \leq (n-1)(\alpha-1)$.*

It seems that a proximity theorem concerning the continuous relaxation of (MC) can be easily shown by using Theorem 2.13, in a similar way as Theorem 2.11 for L-convex function minimization. This approach, however, does not work since a statement similar to Proposition 2.10 does not hold for closed proper M-convex functions; it is also noted that g_α in Theorem 2.13 above is not necessarily M-convex.

3 Algorithms based on continuous relaxation

In this section we propose a new algorithm for the problem (MC) based on continuous relaxation. For this, we firstly propose in Section 3.1 a new greedy-type algorithm for (MC), which will be used within the continuous relaxation algorithm proposed in Section 3.2. The continuous relaxation algorithm is then applied to some special cases of (MC) such as (Laminar), (Nest), (Tree), and (Network) in Sections 3.3 and 3.4, which yields the best time complexity bounds for these special cases.

Throughout this section, we assume that $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ is an M-convex function.

3.1 New greedy algorithm for M-convex function minimization

We propose a new greedy-type algorithm for the problem (MC). This greedy algorithm will be used in the algorithm based on continuous relaxation. The algorithm is similar to but runs faster than the existing greedy algorithms proposed in [18, 35]

The algorithm is based on the following optimality condition of the problem (MC).

Lemma 3.1 ([21, Theorem 2.4], [23, Theorem 2.2]). *For a vector $y \in \arg\min_{\mathbf{Z}} g$, it holds that $y \in \arg\min_{\mathbf{Z}} g$ if and only if $g(y - \chi_i + \chi_j) \geq g(y)$ holds for all $i, j \in N$.*

By using a local information of a given vector, we can get a useful information of the area containing a minimizer of an M-convex function.

Lemma 3.2 ([34, Theorem 2.2]). *Suppose that $\arg\min_{\mathbf{Z}} g \neq \emptyset$. Let $y \in \mathrm{dom}_{\mathbf{Z}}\, g$ and $h \in N$.*
(i) *Suppose that element $i \in N$ satisfies the condition*

$$g(y + \chi_i - \chi_h) = \min_{i' \in N} g(y + \chi_{i'} - \chi_h).$$

Then, there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ satisfying $y_*(i) \geq (y + \chi_i - \chi_h)(i) = y(i) + 1 - \chi_h(i)$.
(ii) *Suppose that element $j \in N$ satisfies the condition*

$$g(y + \chi_h - \chi_j) = \min_{j' \in N} g(y + \chi_h - \chi_{j'}).$$

Then, there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ satisfying $y_*(j) \leq (y + \chi_h - \chi_j)(j) = y(j) - 1 + \chi_h(j)$.

Lemma 3.3. *Let $y \in \mathrm{dom}_{\mathbf{Z}}\, g$ and $h \in N$.*
(i) *Suppose that there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ satisfying $y_*(h) \leq y(h) - 1$. Then, there exists $i \in N$ satisfying $i \neq h$ and $g(y + \chi_i - \chi_h) = \min_{i' \in N} g(y + \chi_{i'} - \chi_h)$.*
(ii) *Suppose that there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ satisfying $y_*(h) \geq y(h) + 1$. Then, there exists $j \in N$ satisfying $j \neq h$ and $g(y + \chi_h - \chi_j) = \min_{j' \in N} g(y + \chi_h - \chi_{j'})$.*

Proof. We prove (i) only since (ii) can be shown similarly. It suffices to show that there exists $i \in N \setminus \{h\}$ satisfying $g(y + \chi_i - \chi_h) \leq g(y)$. Since $h \in \mathrm{supp}^+(y - y_*)$, the property (M-EXC[$\mathbf{Z}$]) for g implies that there exists $i \in \mathrm{supp}^-(y - y_*)$ such that

$$g(y) + g(y_*) \geq g(y - \chi_h + \chi_i) + g(y_* + \chi_h - \chi_i). \tag{3.3}$$

Since $y_* \in \arg\min_{\mathbf{Z}} g$, we have $g(y_* + \chi_h - \chi_i) \geq g(y_*)$, which, together with (3.3), implies $g(y) \geq g(y - \chi_h + \chi_i)$. Note that $i \neq h$ holds. □

Lemma 3.4. *Let $i, j \in N$, and $\alpha, \beta \in \mathbf{Z}$. Suppose that there exist some $y', y'' \in \arg\min_{\mathbf{Z}} g$ satisfying $y'(i) \geq \alpha$ and $y''(j) \leq \beta$.*
(i) *If $j \neq i$, then there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ satisfying both of $y_*(i) \geq \alpha$ and $y_*(j) \leq \beta$.*
(ii) *If $j = i$ and $\alpha \leq \beta$, then there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ satisfying $\alpha \leq y_*(i) \leq \beta$.*

Proof. Let y' be a minimizer of g satisfying $y'(i) \geq \alpha$. Let y'' be a minimizer of g satisfying $y''(j) \leq \beta$, and suppose that y'' has the maximum value of $y''(i)$ among all minimizers with $y''(j) \leq \beta$. Assume, to the contrary, that $y''(i) < \alpha$ holds. Since $i \in \mathrm{supp}^+(y' - y'')$, the property (M-EXC[$\mathbf{Z}$]) for g implies that there exists $h \in \mathrm{supp}^-(y' - y'')$ such that

$$g(y') + g(y'') \geq g(y' - \chi_i + \chi_h) + g(y'' + \chi_i - \chi_h).$$

Since $y', y'' \in \arg\min_{\mathbf{Z}} g$, this inequality implies that $y' - \chi_i + \chi_h, y'' + \chi_i - \chi_h \in \arg\min_{\mathbf{Z}} g$. Put $y_* = y'' + \chi_i - \chi_h$. If $j \neq i$, then we have $y_* \in \arg\min_{\mathbf{Z}} g$, $y_*(j) \leq y''(j) \leq \beta$, and $y_*(i) > y''(i)$, a contradiction to the choice of y''. If $j = i$ and $\alpha \leq \beta$, then we have $y_* \in \arg\min_{\mathbf{Z}} g$ and $y''(i) < y_*(i) \leq \alpha \leq \beta$, a contradiction to the choice of y''. Hence, $y''(i) \geq \alpha$ holds for each case. □

For vectors $\ell \in (\mathbf{Z} \cup \{-\infty\})^n$ and $u \in (\mathbf{Z} \cup \{+\infty\})^n$, we define a function $g_\ell^u : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ by

$$g_\ell^u(y) = \begin{cases} g(y) & (\text{if } \ell \leq y \leq u), \\ +\infty & (\text{otherwise}), \end{cases} \qquad (y \in \mathbf{Z}^n). \tag{3.4}$$

M-convexity of a function is preserved under this restriction operation.

Lemma 3.5 ([21, Lemma 2.5]). *For vectors $\ell \in (\mathbf{Z} \cup \{-\infty\})^n$ and $u \in (\mathbf{Z} \cup \{+\infty\})^n$, the function $g_\ell^u : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ defined by (3.4) is M-convex if $\mathrm{dom}_{\mathbf{Z}}\, g_\ell^u = \{y \in \mathbf{Z}^n \mid y \in \mathrm{dom}_{\mathbf{Z}}\, g,\ \ell \leq y \leq u\} \neq \emptyset$.*

We now describe an algorithm for (MC). We assume that an initial vector $y^\circ \in \mathrm{dom}_{\mathbf{Z}}\, g$ is given in advance.

Algorithm NewGreedy
Step 0: Set $y := y^\circ$, and $\ell(k) := -\infty$ and $u(k) := +\infty$ for all $k \in N$.
Step 1: Let $h \in N$ be an arbitrarily chosen element with $\ell(h) < u(h)$.
Step 2: Find $i_1 \in N$ that minimizes $g(y + \chi_{i_1} - \chi_h)$ under the constraint $\ell \leq y + \chi_{i_1} - \chi_h \leq u$.
Step 3: Find $i_2 \in N$ that minimizes $g(y - \chi_{i_2} + \chi_h)$ under the constraint $\ell \leq y - \chi_{i_2} + \chi_h \leq u$.
Step 4: If $h = i_1 = i_2$, then go to Step 5; if $h \neq i_1$, then go to Step 6;
otherwise (i.e., $h \neq i_2$), go to Step 7.
Step 5: Set $\ell(h) := y(h)$, $u(h) := y(h)$, and go to Step 8.
Step 6: Find $j_1 \in N \setminus \{i_1\}$ that minimizes $g(y + \chi_{i_1} - \chi_{j_1})$ under the constraint
$\ell \leq y + \chi_{i_1} - \chi_{j_1} \leq u$. Set $\ell(i_1) := y(i_1) + 1$, $u(j_1) := y(j_1) - 1$, and $y := y + \chi_{i_1} - \chi_{j_1}$.
Go to Step 8.
Step 7: Find $j_2 \in N \setminus \{i_2\}$ that minimizes $g(y - \chi_{i_2} + \chi_{j_2})$ under the constraint
$\ell \leq y - \chi_{i_2} + \chi_{j_2} \leq u$. Set $u(i_2) := y(i_2) - 1$, $\ell(j_2) := y(j_2) + 1$, and $y := y - \chi_{i_2} + \chi_{j_2}$.
Go to Step 8.
Step 8: If every element $k \in N$ satisfies $\ell(k) = u(k)$, then output y and stop
(y is a minimizer of g); otherwise, go to Step 1.

We first prove the correctness of the algorithm.

Lemma 3.6. *The interval $[\ell, u]$ always contains a minimizer of g.*

Proof. We prove the statement of the lemma by induction on the number of iterations. It is obvious that $[\ell, u]$ contains a minimizer of g at the beginning of the algorithm.

First of all, assume that $h = i_1 = i_2$ holds in Step 4. We consider the function g_ℓ^u defined by (3.4), which is M-convex by Lemma 3.5. By induction hypothesis, it holds that $\arg\min_{\mathbf{Z}} g_\ell^u \subseteq \arg\min_{\mathbf{Z}} g$. We see from Lemma 3.2 applied to g_ℓ^u that there exist $y', y'' \in \arg\min_{\mathbf{Z}} g_\ell^u$ satisfying $y'(h) \geq y(h)$ and $y''(h) \leq y(h)$. Then, Lemma 3.4 (ii) implies that there exists $y_* \in \arg\min_{\mathbf{Z}} g_\ell^u$

with $y(h) \leq y_*(h) \leq y(h)$. Hence, $[\ell, u]$ contains a minimizer of g after the update of $\ell(h)$ and $u(h)$ in Step 5.

We then consider the case where $h \neq i_1$ holds in Step 4. We see from Lemma 3.2 (i) applied to g_ℓ^u that there exists $y' \in \arg\min_{\mathbf{Z}} g_\ell^u$ satisfying $y'(i_1) \geq y(i_1) + 1$. By Lemma 3.2 (ii) and Lemma 3.3 (ii), there exists $y'' \in \arg\min_{\mathbf{Z}} g_\ell^u$ satisfying $y''(j_1) \leq y(j_1) - 1$. Then, Lemma 3.4 (i) implies that there exists $y_* \in \arg\min_{\mathbf{Z}} g_\ell^u$ satisfying both of $y_*(i_1) \geq y(i_1) + 1$ and $y_*(j_1) \leq y(j_1) - 1$. Hence, $[\ell, u]$ contains a minimizer of g after the update of $\ell(i_1)$ and $u(j_1)$ in Step 6.

The case where $h \neq i_2$ holds in Step 4 can be dealt with in a similar way as the previous case. This concludes the proof. □

We denote by $y_{\rm out}$ the output of the algorithm. The vector $y_{\rm out}$ satisfies $y_{\rm out} = \ell = u$, i.e., $y_{\rm out}$ is the unique vector in the interval $[\ell, u]$ when the algorithm terminates. Hence, $y_{\rm out}$ is a minimizer of g by Lemma 3.6.

We then analyze the number of iterations.

Lemma 3.7. *If Step 6 or Step 7 is executed, then $\|y - y_{\rm out}\|_1$ reduces by two.*

Proof. The vector $y_{\rm out}$ is always contained in the interval $[\ell, u]$ since the vector ℓ (resp., u) is nondecreasing (resp., nonincreasing) in each iteration. We consider the case where Step 6 is executed since the case of Step 7 can be dealt with similarly. Let $y_{\rm old}$ (resp., $y_{\rm new}$) the vector y before the update (resp., after the update) in Step 6 . We also denote by $\ell_{\rm new}$ and $u_{\rm new}$ the vectors ℓ and u after the update in Step 6. Then, we have

$$\begin{aligned} y_{\rm out}(i_1) &\geq \ell_{\rm new}(i_1) = y_{\rm old}(i_1) + 1 = y_{\rm new}(i_1) > y_{\rm old}(i_1), \\ y_{\rm out}(j_1) &\leq u_{\rm new}(j_1) = y_{\rm old}(j_1) - 1 = y_{\rm new}(j_1) < y_{\rm old}(j_1). \end{aligned}$$

Hence, we have $\|y_{\rm new} - y_{\rm out}\|_1 = \|y_{\rm old} - y_{\rm out}\|_1 - 2$. □

Lemma 3.8. *The algorithm* NewGreedy *terminates in* $\mathrm{O}(n + \|y^\circ - y_{\rm out}\|_1)$ *iterations.*

Proof. Once the equation $\ell(h) = u(h)$ is satisfied for some $h \in N$, this equation is always satisfied in the following iterations since the vector ℓ (resp., u) is nondecreasing (resp., nonincreasing) in each iteration. Hence, Step 5 is executed at most n times. In addition, the value $\|y - y_{\rm out}\|_1$ remains the same after the execution of Step 5. By Lemma 3.7, the value $\|y - y_{\rm out}\|_1$ decreases by two whenever Step 6 or 7 is executed. Since the initial value of $\|y - y_{\rm out}\|_1$ is $\|y^\circ - y_{\rm out}\|_1$, Steps 6 and 7 are executed $\|y^\circ - y_{\rm out}\|_1/2$ times. Hence, the algorithm NewGreedy terminates in $\mathrm{O}(n + \|y^\circ - y_{\rm out}\|_1)$ iterations. □

It is easy to see that each iteration of the algorithm can be done in $\mathrm{O}(nF)$ time, where F denotes the time to evaluate the value $g(y)$ of the given M-convex function g and a given vector $y \in \mathbf{Z}^n$. Hence, we obtain the following result.

Theorem 3.9. *The algorithm* NewGreedy *outputs a minimizer of an M-convex function g in $\mathrm{O}(nF(n + \|y^\circ - y_{\rm out}\|_1))$ time.*

Note that in the special case of (SC), F represents the time to check whether a given vector is feasible, plus the time to evaluate the value of the objective function if a given vector is feasible.

We see from Theorem 3.9 that the running time of the algorithm NewGreedy depends on the distance between the initial vector y° and the minimizer $y_{\rm out}$ of the M-convex function g computed by the algorithm. With a slight modification as in [25], the algorithm always finds a minimizer of g having the smallest L_1 distance from y°.

Our idea is to consider a perturbed function

$$g_\varepsilon(y) = g(y) + \varepsilon\|y^\circ - y\|_1 \qquad (y \in \mathbf{Z}^n)$$

instead of the original function g, where ε is a sufficiently small positive number. Note that $\varepsilon\|y^\circ - y\|_1$ is a separable convex function in y, and M-convexity of a function is preserved by the addition of a separable convex function [24, Theorem 6.13]. Hence, g_ε is an M-convex function. Due to the choice of ε, we have $y_* \in \arg\min_{\mathbf{Z}} g_\varepsilon$ if and only if $y_* \in \arg\min_{\mathbf{Z}} g$ and $\|y^\circ - y_*\|_1 = \min\{\|y^\circ - y\|_1 \mid y \in \arg\min_{\mathbf{Z}} g\}$. Hence, it suffices to find a minimizer of g_ε.

Suppose that we apply the algorithm NewGreedy to the perturbed function g_ε. This amounts to using the following rules in Steps 2, 3, 6, and 7:

> **(Rule 1)** Suppose that we want to find an element i that minimizes $g(y + \chi_i - \chi_h)$ under the constraint $\ell \leq y+\chi_i-\chi_h \leq u$. If there exists such i satisfying $y(i) < y^\circ(i)$, then we take it; otherwise we take any such element.
> **(Rule 2)** Suppose that we want to find an element i that minimizes $g(y - \chi_i + \chi_h)$ under the constraint $\ell \leq y-\chi_i+\chi_h \leq u$. If there exists such i satisfying $y(i) > y^\circ(i)$, then we take it; otherwise we take any such element.

Theorem 3.10. *The algorithm* NewGreedy *with the rules* (Rule 1) *and* (Rule 2) *finds a minimizer y_* of an M-convex function g satisfying $\|y^\circ - y_*\|_1 = \min\{\|y^\circ - y\|_1 \mid y \in \arg\min_{\mathbf{Z}} g\}$ in $\mathrm{O}(nF(n + \min\{\|y^\circ - y\|_1 \mid y \in \arg\min_{\mathbf{Z}} g\}))$ time.*

3.2 Algorithm based on continuous relaxation

We propose a new algorithm for (MC) based on continuous relaxation. Assume that we are given a closed proper M-convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ such that $f(y) = g(y)$ $(\forall y \in \mathbf{Z}^n)$ and $\mathrm{dom}_{\mathbf{R}} f$ is the closed convex hull of $\mathrm{dom}_{\mathbf{Z}}\, g$. The algorithm is described as follows.

Algorithm Continuous
Step 1: Compute a minimizer $x_* \in \mathrm{dom}_{\mathbf{R}} f$ of the closed proper M-convex function f.
Step 2: Compute an integral vector $y^\circ \in \mathrm{dom}_{\mathbf{Z}}\, g$ with $\|y^\circ - x_*\|_1 < n$.
Step 3: Apply the algorithm NewGreedy to the M-convex function g, where we use y° as an initial vector.

We note that $\mathrm{dom}_{\mathbf{R}} f$ is an integral base polyhedron since it is the closed convex hull of the M-convex set $\mathrm{dom}_{\mathbf{Z}}\, g$ (see Section 2.2). Hence, for any $x \in \mathrm{dom}_{\mathbf{R}} f$ there exists an integral vector $y \in \mathrm{dom}_{\mathbf{R}} f \cap \mathbf{Z}^n = \mathrm{dom}_{\mathbf{Z}}\, g$ satisfying $\|y - x\|_\infty < 1$ (see [6, 24]), and such a vector y satisfies $\|y - x_*\|_1 < n$. For some special cases of (MC) such as (Laminar) and (Network), we can compute such an integral vector efficiently, as explained later.

We see from Corollary 1.4 that there exists some $y_* \in \arg\min_{\mathbf{Z}} g$ such that $\|y_* - x_*\|_1 < n(n-1)$. Hence, the vector y° computed in Step 2 satisfies

$$\|y_* - y^\circ\|_1 \leq \|y_* - x_*\|_1 + \|x_* - y^\circ\|_1 < n(n-1) + n = n^2.$$

This, together with Theorem 3.10, implies that the algorithm NewGreedy in Step 3 terminates in $\mathrm{O}(n^3 F)$ time.

We denote by T_{relax} the time required by Step 1, i.e., the time required for solving the continuous relaxation ($\overline{\mathsf{MC}}$), and by T_{round} the time required by Step 2. Then, the algorithm Continuous finds an optimal solution to (MC) in $\mathrm{O}(T_{\mathrm{relax}}+T_{\mathrm{round}}+n^3F)$ time. This concludes the proof of Theorem 1.7.

3.3 Application to laminar convex resource allocation problem

We apply the continuous relaxation algorithm Continuous proposed in Section 3.2 to the problem ($\mathsf{Laminar}$). In particular, we consider the case where the objective function is given by a quadratic function, and prove Theorem 1.9. In the following, it is assumed that we are given a formulation of ($\mathsf{Laminar}$) as a convex cost flow problem on a tree network (for the convex cost flow formulation, see Section 2.3). Since we have a tree network representing the laminar family $\mathcal{F}$, for every $X \in \mathcal{F}$ we can find all children of X in $\mathrm{O}(k)$ time, where k is the number of children of X; moreover, we can find the parent of X in $\mathrm{O}(1)$ time if $X \neq N$. We also assume that for each function f_X ($X \in \mathcal{F}$) used in the objective function of ($\mathsf{Laminar}$), the function value can be evaluated in constant time.

We firstly explain how to compute an optimal solution $x_* \in \mathbf{R}^n$ to ($\overline{\mathsf{Laminar}}$) with a quadratic objective function. Since this problem is equivalent to a quadratic convex cost real-valued flow problem on a tree network, it can be solved by the algorithm of Tamir [36] in $\mathrm{O}(n^2)$ time. Hence, Step 1 of the algorithm Continuous can be done in $\mathrm{O}(n^2)$ time, provided that the objective function is quadratic.

We next explain how to round an optimal solution $x_* \in \mathbf{R}^n$ of ($\overline{\mathsf{Laminar}}$) to a feasible solution of ($\mathsf{Laminar}$). We here use a network flow technique. Since the coefficient matrix defining the system of inequalities $\ell_Y \leq x(Y) \leq u_Y$ ($Y \in \mathcal{F}$) is totally unimodular, there exists an integral vector $y \in \mathbf{Z}^n$ satisfying

$$(\ell_Y \leq)\ \lfloor x_*(Y) \rfloor \leq y(Y) \leq \lceil x_*(Y) \rceil\ (\leq u_Y) \qquad (\forall Y \in \mathcal{F}) \tag{3.5}$$

(see, e.g., [1, 33]). Note that the condition (3.5) implies that y is a feasible solution to ($\mathsf{Laminar}$) with $\|y - x_*\|_1 < n$ since $\{i\} \in \mathcal{F}$ for all $i \in N$. An integral vector $y \in \mathbf{Z}^n$ satisfying (3.5) can be computed in $\mathrm{O}(n)$ time, as follows.

For $Y \in \mathcal{F}$, let $\ell'_Y = \lfloor x_*(Y) \rfloor$ and $u'_Y = \lceil x_*(Y) \rceil$. To obtain a vector $y \in \mathbf{Z}^n$ satisfying $\ell'_Y \leq y(Y) \leq u'_Y$ ($\forall Y \in \mathcal{F}$), we compute the value φ_Y ($Y \in \mathcal{F}$) satisfying the conditions

$$\begin{aligned}
&\ell'_Y \leq \varphi_Y \leq u'_Y \qquad (\forall Y \in \mathcal{F}),\\
&\varphi_N = K,\\
&\varphi_X = \sum\{\varphi_Y \mid Y \in \mathcal{F},\ Y \text{ is a child of } X\} \quad (\forall X \in \mathcal{F},\ |X| \geq 2)
\end{aligned}$$

in a top-down fashion:

Step 0: Put $\varphi_N = K$.

Step 1: If the value φ_Y is computed for all $Y \in \mathcal{F}$, then stop, and output the vector $y \in \mathbf{Z}^n$ defined by $y(i) = \varphi_{\{i\}}$ ($i \in N$).

Step 2: Let $X \in \mathcal{F}$ be a set with $|X| \geq 2$ such that φ_X is already computed but for some child Y of X the value φ_Y is not computed. Let $Y_1, Y_2, \ldots, Y_k \in \mathcal{F}$ be all children of X. For $t = 1, 2, \ldots, k$, we assign the value ℓ'_{Y_t} or u'_{Y_t} to φ_{Y_t} appropriately, so that $\sum_{t=1}^{k} \varphi_{Y_t} = \varphi_X$ holds. Go to Step 1.

Step 2 can be done in $\mathrm{O}(k)$ time since the values ℓ'_Y, u'_Y satisfy the following properties:

$$u'_Y = \ell'_Y \text{ or } u'_Y = \ell'_Y + 1 \quad (\forall Y \in \mathcal{F}),$$
$$\sum_{t=1}^{k} \ell'_{Y_t} \leq \ell'_X \leq u'_X \leq \sum_{t=1}^{k} u'_{Y_t} \quad (\forall X \in \mathcal{F},\ |X| \geq 2,\ Y_1, Y_2, \ldots, Y_k \in \mathcal{F} : \text{children of } X).$$

Since $|\mathcal{F}| = \mathrm{O}(n)$, the algorithm described above runs in $\mathrm{O}(n)$ time. This shows that Step 2 of the algorithm Continuous can be done in $\mathrm{O}(n)$ time.

We finally show that Step 3 of the algorithm Continuous can be done in $\mathrm{O}(n^2)$ time. We note that the vector $y^\circ \in \mathbf{Z}^n$ obtained in Step 2 satisfies $\|y^\circ - x_*\|_1 < n$, which, together with the proximity theorem for (Laminar) (Theorem 1.6), implies that there exists an optimal solution $y_* \in \mathbf{Z}^n$ to (Laminar) such that $\|y_* - y^\circ\|_1 < 3n$. Hence, the algorithm NewGreedy used in Step 3 of Continuous terminates in $\mathrm{O}(n)$ iterations by Lemma 3.8. Since $|\mathcal{F}| = \mathrm{O}(n)$, the evaluation of the objective function requires $\mathrm{O}(n)$ time. Therefore, each iteration of NewGreedy requires $\mathrm{O}(n^2)$ time if we use a naive implementation. This can be reduced to $\mathrm{O}(n)$ time by using a network flow technique, as follows.

We again use the reformulation of (Laminar) as a convex cost flow problem on a tree network. We will construct a so-called *residual network* (or *auxiliary network*) of the convex cost flow problem (see, e.g., [1]). Given a feasible solution $y \in \mathbf{Z}^n$ to (Laminar), we construct a directed graph $G_y = (V, A_y)$, where $V = \{v_Y \mid Y \in \mathcal{F}\}$ and the arc set A_y is given as

$$A_y = \{(v_{p(Y)}, v_Y) \mid Y \in \mathcal{F} \setminus \{N\},\ y(Y) < u_Y\} \cup \{(v_Y, v_{p(Y)}) \mid Y \in \mathcal{F} \setminus \{N\},\ y(Y) > \ell_Y\}.$$

We then define the length of each arc as follows:

for the arc of the form $(v_{p(Y)}, v_Y)$, its length is $f_Y(y(Y) + 1) - f_Y(y(Y))$;
for the arc of the form $(v_Y, v_{p(Y)})$, its length is $f_Y(y(Y) - 1) - f_Y(y(Y))$.

We note that $\{f_Y(y(Y)+1) - f_Y(y(Y))\} + \{f_Y(y(Y)-1) - f_Y(y(Y))\} \geq 0$ since f_Y is a convex function. This implies that the graph G_y does not contain a directed cycle of negative length.

We see that for $i, j \in N$, the vector $y - \chi_i + \chi_j$ is a feasible solution to (Laminar) if and only if there exists a directed path from the node $v_{\{i\}}$ to the node $v_{\{j\}}$. If $y - \chi_i + \chi_j$ is feasible for some $i, j \in N$, then it holds that

$$f_{\mathrm{sum}}(y - \chi_i + \chi_j) - f_{\mathrm{sum}}(y) = \text{the length of a shortest directed path from } v_{\{i\}} \text{ to } v_{\{j\}},$$

where $f_{\mathrm{sum}}(y') = \sum_{Y \in \mathcal{F}} f_Y(y'(Y))$ for $y' \in \mathbf{Z}^n$. Since the underlying (undirected) graph of G_y is a tree, a shortest path from some node to another node is uniquely determined. For a fixed $i \in N$, we can compute the length of the shortest directed path from $v_{\{i\}}$ to $v_{\{j\}}$ for all $j \in N$

in $\mathrm{O}(n)$ time by using a linear-time graph search algorithm. Similarly, for a fixed $i \in N$, we can also compute the length of the directed simple path from $v_{\{j\}}$ to $v_{\{i\}}$ for all $j \in N$ in $\mathrm{O}(n)$ time. This shows that each iteration of the algorithm NewGreedy can be done in $\mathrm{O}(n)$ time.

Summarizing the discussion above, we can solve the problem (Laminar) in $\mathrm{O}(n^2)$ time if the objective function is quadratic. This concludes the proof of Theorem 1.9. Since the problems (Nest) and (Tree) are special cases of (Laminar), we obtain the following results as a corollary of Theorem 1.9.

Corollary 3.11. *The problems* (Nest) *and* (Tree) *can be solved in* $\mathrm{O}(n^2)$ *time if the objective functions are quadratic.*

It should be mentioned that by Theorem 2.6 the continuous relaxation problems ($\overline{\mathsf{Nest}}$) and ($\overline{\mathsf{Tree}}$) of (Nest) and (Tree), respectively, can be solved in $\mathrm{O}(n \log n)$ time, which is smaller than the bound $\mathrm{O}(n^2)$ for ($\overline{\mathsf{Laminar}}$). This does not affect the running time of the algorithm Continuous since Step 3 requires $\mathrm{O}(n^2)$ time even for (Nest) and (Tree). It is an open question whether there exist $\mathrm{O}(n \log n)$-time algorithms for (Nest) and (Tree) with quadratic objective functions.

3.4 Application to network resource allocation problems

We apply the continuous relaxation algorithm Continuous proposed in Section 3.2 to the problem (Network), which is a special case of (SC). We consider the case where the objective function is given by a quadratic function. In the following, we assume that for each function f_i ($i \in N$) used in the objective function of (Network), the function value can be evaluated in constant time.

We firstly consider Step 1 of the algorithm Continuous, where the continuous relaxation problem ($\overline{\mathsf{Network}}$) is solved. By Theorem 2.6, ($\overline{\mathsf{Network}}$) can be solved in $\mathrm{O}(|V||A| \log(|V|^2/|A|))$ time if the objective function is quadratic, where V and A denote the node set and the arc set of the underlying directed graph.

We then explain how to round an optimal solution $x_* \in \mathbf{R}^n$ of the continuous relaxation problem to a feasible solution of the original problem. Let $(x_*, \varphi_*) \in \mathbf{R}^n \times \mathbf{R}^A$ be an optimal solution to ($\overline{\mathsf{Network}}$). Since the coefficient matrix defining the system of the constraints in ($\overline{\mathsf{Network}}$) is totally unimodular, there exists a pair of integral vectors $y \in \mathbf{Z}^n$ and $\psi \in \mathbf{Z}^A$ such that (y, ψ) is a feasible solution to (Network) satisfying

$$\lfloor x_*(i) \rfloor \leq y(i) \leq \lceil x_*(i) \rceil \quad (\forall i \in N), \qquad \lfloor \varphi_*(a) \rfloor \leq \psi(a) \leq \lceil \varphi_*(a) \rceil \quad (\forall a \in A) \quad (3.6)$$

(see, e.g., [1, 33]). Such (y, ψ) can be computed efficiently by using a max-flow algorithm; for example, it can be computed in $\mathrm{O}(|V||A| \log(|V|^2/|A|))$ time by the algorithm of Goldberg and Tarjan [8]. Note that the condition (3.6) implies $\|y - x_*\|_1 < n$, in particular. Hence, Step 2 of the algorithm Continuous can be done in $\mathrm{O}(|V||A| \log(|V|^2/|A|))$ time.

We finally show that Step 3 of the algorithm Continuous can be done in $\mathrm{O}(n(|V| + |A|))$ time. We note that the vector $y^\circ \in \mathbf{Z}^n$ obtained in Step 2 satisfies $\|y^\circ - x_*\|_1 < n$, which, together with the proximity theorem for (SC) (Theorem 1.5), implies that there exists an optimal solution $y_* \in \mathbf{Z}^n$ to (Network) such that $\|y_* - y^\circ\|_1 < 3n$. Hence, the algorithm NewGreedy used in Step 3 of Continuous terminates in $\mathrm{O}(n)$ iterations by Lemma 3.8.

In the following, we show that each iteration of NewGreedy used in Step 3 can be done in $\mathrm{O}(|V| + |A|)$ time by using a network flow technique.

We construct a so-called *residual network* (or *auxiliary network*) of the convex cost flow problem (see, e.g., [1]). Given a feasible solution $(y, \psi) \in \mathbf{Z}^n \times \mathbf{Z}^A$ to (Network), we construct a directed graph $G_\psi = (V, A_\psi)$, where

$$A_\psi = \{(h,k) \mid (h,k) \in A,\ \psi(h,k) < c(h,k)\} \cup \{(k,h) \mid (h,k) \in A,\ \psi(h,k) > 0\}.$$

We see that for $i, j \in N$, the vector $y - \chi_i + \chi_j$ is a feasible solution to (Network) if and only if $\mathbf{0} \le y - \chi_i + \chi_j \le u$ and there exists a directed path from the node i to the node j in G_ψ. We also note that

$$f_{\mathrm{sum}}(y - \chi_i + \chi_j) - f_{\mathrm{sum}}(y) = \{f_i(y(i) - 1) - f_i(y(i))\} + \{f_j(y(j) + 1) - f_j(y(j))\},$$

where $f_{\mathrm{sum}}(y') = \sum_{k \in N} f_k(y'(k))$ for $y' \in \mathbf{Z}^n$. For a fixed $i \in N$, we can compute the set of nodes in G_ψ reachable from the node i in $\mathrm{O}(|V|+|A|)$ time by using a linear-time graph search algorithm. Similarly, for a fixed $i \in N$, we can compute the set of nodes in G_ψ reachable to the node i in $\mathrm{O}(|V|+|A|)$ time. This fact implies that each iteration of the algorithm NewGreedy can be done in $\mathrm{O}(|V| + |A|)$ time.

Summarizing the discussion above, we obtain the following result for the problem (Network) with a quadratic objective function.

Theorem 3.12. *The problem* (Network) *can be solved in* $\mathrm{O}(|V||A| \log(|V|^2/|A|))$ *time if the objective function is quadratic.*

4 Proofs

We give proofs of proximity theorems; proofs of Theorems 1.3, 1.5, and 1.6 are given in Sections 4.1, 4.2, and 4.3, respectively. Before starting the proofs, we show the tightness of the bounds in the proximity theorems by using Example 2.9 in Section 2.4. It shows that the bounds in the proximity theorems are tight, even for the very special case of the simple resource allocation problem (Simple).

Example 2.9. (continued) The vector $y_* = (0, 1, \ldots, 1)$ is the unique optimal solution to (Simple), while $x_* = ((n-1)(1-\delta), \delta, \ldots, \delta)$ is the unique optimal solution to the continuous relaxation $(\overline{\mathsf{Simple}})$. It holds that

$$\|y_* - x_*\|_\infty = (n-1)(1-\delta), \qquad \|y_* - x_*\|_1 = 2(n-1)(1-\delta),$$

which can be arbitrarily close to $n - 1$ and $2(n-1)$, respectively. □

4.1 Proof of Theorem 1.3

We prove Theorem 1.3, a proximity theorem for (MC). To prove this, it is convenient to consider the problem (GMC) instead of (MC):

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

where $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is a closed proper M-convex function in real variables. The problem (GMC) is more general than (MC) since for every M-convex function $g : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$, there exists a closed proper M-convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ in real variables such that $f(x) = g(x)$ for all $x \in \mathbf{Z}^n$, as stated in Theorem 2.3.

Continuous relaxation of (GMC) can be naturally defined as follows:

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

We will show a proximity result for the problem (GMC).

Theorem 4.1.
(i) *For every optimal solution $y_* \in \mathbf{Z}^n$ to* (GMC), *there exists an optimal solution $x_* \in \mathbf{R}^n$ to* ($\overline{\mathsf{GMC}}$) *such that $\|x_* - y_*\|_\infty < n - 1$.*
(ii) *For every optimal solution $x_* \in \mathbf{R}^n$ to* ($\overline{\mathsf{GMC}}$), *there exists an optimal solution $y_* \in \mathbf{Z}^n$ to* (GMC) *such that $\|y_* - x_*\|_\infty < n - 1$.*

Since the problem (MC) is a special case of (GMC), Theorem 1.3 follows immediately from Theorem 4.1. It should be mentioned that Theorem 4.1 implies that a closed proper M-convex function f satisfies $\arg\min_{\mathbf{R}} f \neq \emptyset$ if and only if $\arg\min\{f(y) \mid y \in \mathbf{Z}^n\} \neq \emptyset$ (see Remark 4.4 below). In the following, we prove Theorem 4.1.

4.1.1 Proof of Theorem 4.1 (i)

To prove the statement (i) of Theorem 4.1, we use the following two properties.

The next lemma states that the projection of a closed proper M-convex function f along an arbitrarily chosen coordinate axis $i \in N$ is a supermodular function.

Lemma 4.2 ([29, Proposition 3.12]). *For every $x, y \in \mathbf{R}^n$ and every $i \in N$, we have $f(x) + f(y) \leq f(\hat{x}) + f(\check{y})$, where $\hat{x}$ and $\check{y}$ are given as*

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

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

For $\gamma \in \mathbf{R}$, we define $\mathrm{level}(f, \gamma) = \{x \in \mathbf{R}^n \mid f(x) \leq \gamma\}$. We note that $\mathrm{level}(f, \gamma)$ is a closed set since f is a closed convex function (see, e.g., [32, Theorem 7.1]).

Lemma 4.3. *Let $y_* \in \mathrm{dom}_{\mathbf{R}}\, f$, and $\gamma \in \mathbf{R}$ be a real number such that $\mathrm{level}(f, \gamma) \neq \emptyset$. Suppose that $\tilde{x} \in \mathrm{level}(f, \gamma)$ be a vector minimizing the value $\|\tilde{x} - y_*\|_1$ among all vectors in $\mathrm{level}(f, \gamma)$.*
(i) *If $k \in N$ satisfies the condition*

$$f(y_* - \chi_i + \chi_k) \geq f(y_*) \qquad (\forall i \in N), \tag{4.1}$$

then it holds that $\tilde{x}(k) - y_*(k) < n - 1$.
(ii) *If $k \in N$ satisfies the condition*

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

then it holds that $\tilde{x}(k) - y_*(k) > -(n - 1)$.

Proof. We prove (i) only since (ii) can be shown in a similar way. If $\tilde{x}(k) \le y_*(k)$, then we have $\tilde{x}(k) - y_*(k) \le 0 < n-1$. Hence, we may assume that $\tilde{x}(k) > y_*(k)$. We have

$$f(\tilde{x}-\varepsilon(\chi_k-\chi_i)) > f(\tilde{x}) \quad (\forall i \in \mathrm{supp}^-(\tilde{x}-y_*),\ 0<\forall\varepsilon \le \min(\tilde{x}(k)-y_*(k), y_*(i)-\tilde{x}(i))) \tag{4.2}$$

since otherwise there exists a vector $x' \in \mathrm{dom}_{\mathbf{R}} f$ satisfying $f(x') \le f(\tilde{x}) \le \gamma$ and $\|x'-y_*\|_1 < \|\tilde{x}-y_*\|_1$, a contradiction to the choice of $\tilde{x}$. Let $\mathrm{supp}^-(\tilde{x}-y_*) = \{i_1, i_2, \ldots, i_t\}$, where $t = |\mathrm{supp}^-(\tilde{x}-y_*)|$ $(\le n-1)$. Put $y_0 = y_*$, and iteratively define $\lambda_h \in \mathbf{R}_+$ and $y_h \in \mathbf{R}^n$ for each $h = 1, 2, \ldots, t$ by

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

and $y_h = y_{h-1} + \lambda_h(\chi_k - \chi_{i_h})$. Note that the definition of λ_h allows for $\lambda_h = 0$. By the definition of y_h and closed convexity of f, we have

$$f(y_h) < f(y_{h-1}) \qquad \text{if } \lambda_h > 0, \tag{4.3}$$

$$f(y_h + \lambda(\chi_k - \chi_{i_h})) \ge f(y_h) \quad (\forall \lambda > 0) \quad \text{if } \tilde{x}(k) > y_h(k) \text{ and } y_h(i_h) > \tilde{x}(i_h). \tag{4.4}$$

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

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

$$f(\tilde{x}) + f(y_t) \ge f(\tilde{x} - \lambda(\chi_k - \chi_{i_h})) + f(y_t + \lambda(\chi_k - \chi_{i_h})).$$

By Lemma 4.2 with $i = k$, we obtain

$$f(y_h + \lambda(\chi_k - \chi_{i_h})) + f(y_t) \le f(y_t + \lambda(\chi_k - \chi_{i_h})) + f(y_h).$$

Combining the two inequalities, we have

$$f(y_h + \lambda(\chi_k - \chi_{i_h})) - f(y_h) \le f(\tilde{x}) - f(\tilde{x} - \lambda(\chi_k - \chi_{i_h})) < 0,$$

where the last inequality is by (4.2) since $i_h \in \mathrm{supp}^-(\tilde{x}-y_t) \subseteq \mathrm{supp}^-(\tilde{x}-y_*)$. This, however, contradicts (4.4). [End of the proof of Claim 1]

Claim 2: For $h = 1, 2, \ldots, t$, if $\lambda_h > 0$ then $f(y_* + \lambda_h(\chi_k - \chi_{i_h})) < f(y_*)$.

[Proof of Claim 2] Let h be any integer in $\{1, 2, \ldots, t\}$ with $\lambda_h > 0$. By Lemma 4.2 with $i = k$, we have

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

which implies

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

where the last inequality is by (4.3). [End of the proof of Claim 2]

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

$$f(y_* + \beta(\chi_k - \chi_i)) \ge f(y_*) \qquad (\forall \beta \ge 1,\ \forall i \in N).$$

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 as follows:

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

□

We are now ready to prove the statement (i) of Theorem 4.1. Let y_* be a minimizer of (GMC), i.e., y_* satisfies $f(y_*) = \min\{f(y) \mid y \in \mathbf{Z}^n\}$. Then, we have the following inequalities for every $k \in N$:

$$f(y_* - \chi_i + \chi_k) \geq f(y_*) \qquad (\forall i \in N), \tag{4.5}$$

$$f(y_* - \chi_k + \chi_j) \geq f(y_*) \qquad (\forall j \in N). \tag{4.6}$$

Assume that $\arg\min_{\mathbf{R}} f \neq \emptyset$, and let $\gamma = \min\{f(x) \mid x \in \mathbf{R}^n\}$. Then, $\mathrm{level}(f, \gamma) = \arg\min_{\mathbf{R}} f$ holds. Let $\tilde{x} \in \mathbf{R}^n$ be a vector in $\mathrm{level}(f, \gamma)$, and assume that $\tilde{x}$ minimizes the value $\|\tilde{x} - y_*\|_1$ among all vectors in $\mathrm{level}(f, \gamma)$. Lemma 4.3 (i) and (4.5) imply that $\tilde{x}(k) - y_*(k) < n-1$ for every $k \in N$. Similarly, Lemma 4.3 (ii) and (4.6) imply that $\tilde{x}(k) - y_*(k) > -(n-1)$ for every $k \in N$. This shows that $\tilde{x} \in \mathrm{level}(f, \gamma) = \arg\min_{\mathbf{R}} f$ satisfies $\|\tilde{x} - y_*\|_\infty < n - 1$.

It remains to prove that $\arg\min_{\mathbf{R}} f \neq \emptyset$. For this, we show the following property:

$$\forall \gamma \in \mathbf{R} \text{ with } \mathrm{level}(f, \gamma) \neq \emptyset,\ \exists x \in \mathrm{level}(f, \gamma) \text{ such that } \|x - y_*\|_\infty \leq n - 1. \tag{4.7}$$

Let $\tilde{x} \in \mathrm{dom}_{\mathbf{R}} f$ be a vector in $\mathrm{level}(f, \gamma)$, and assume that $\tilde{x}$ minimizes the value $\|\tilde{x} - y_*\|_1$ among all vectors in $\mathrm{level}(f, \gamma)$. Lemma 4.3 (i) and (4.5) imply that $\tilde{x}(k) - y_*(k) < n - 1$ for every $k \in N$. Similarly, Lemma 4.3 (ii) and (4.6) imply that $\tilde{x}(k) - y_*(k) > -(n-1)$ for every $k \in N$. This shows that (4.7) holds.

The property (4.7) implies that

$$\begin{aligned} \inf\{f(x) \mid x \in \mathrm{dom}_{\mathbf{R}} f,\ \|x - y_*\|_\infty \leq n - 1\} &= \inf\{f(x) \mid x \in \mathrm{dom}_{\mathbf{R}} f\}, \\ \arg\min\{f(x) \mid x \in \mathrm{dom}_{\mathbf{R}} f,\ \|x - y_*\|_\infty \leq n - 1\} &\subseteq \arg\min_{\mathbf{R}} f. \end{aligned}$$

We have $\arg\min\{f(x) \mid x \in \mathrm{dom}_{\mathbf{R}} f \mid \|x - y_*\|_\infty \leq n - 1\} \neq \emptyset$ since f is a closed proper convex function and the set $\{x \in \mathrm{dom}_{\mathbf{R}} f \mid \|x - y_*\|_\infty \leq n - 1\}$ is bounded and closed. Hence, it holds that $\arg\min_{\mathbf{R}} f \neq \emptyset$.

4.1.2 Proof of Theorem 4.1 (ii)

To prove the statement (ii) of Theorem 4.1, we use the statement (i) shown in Section 4.1.1. Let $x_* \in \mathbf{R}^n$ be an optimal solution to $(\overline{\mathsf{GMC}})$. Let $y_* \in \mathbf{Z}^n$ be an optimal solution to (GMC) minimizing the value $\|x_* - y_*\|_1$ among all optimal solutions to (GMC). Using a positive number δ, we define a new problem:

$$(\mathsf{GMC}^\delta) \qquad \text{Minimize} \quad f(y) + \delta \|y - x_*\|_1 \qquad \text{subject to} \quad y \in \mathbf{Z}^n.$$

The function $\delta\|y - x_*\|_1$ is a separable convex function in y, and therefore $f(y) + \delta\|y - x_*\|_1$ is a closed proper M-convex function in y since the addition of a separable convex function preserves M-convexity [24, Theorem 6.49]. It is easy to see that x_* is the unique optimal solution to the continuous relaxation of (GMC^{δ}). In addition, if δ is a sufficiently small positive number, then y_* is an optimal solution to (GMC^{δ}). Hence, by applying Theorem 4.1 (i) to (GMC^{δ}) and its continuous relaxation, we obtain $\|x_* - y_*\|_\infty < n - 1$.

Remark 4.4. Theorem 4.1 implies that a closed proper M-convex function f satisfies $\arg\min_{\mathbf{R}} f \neq \emptyset$ if and only if $\arg\min\{f(y) \mid y \in \mathbf{Z}^n\} \neq \emptyset$ holds. In the general case where f is not necessarily M-convex, however, the properties $\arg\min_{\mathbf{R}} f \neq \emptyset$ and $\arg\min\{f(y) \mid y \in \mathbf{Z}^n\} \neq \emptyset$ are independent of each other, as shown in the following two examples.

Let $f : \mathbf{R}^2 \to \mathbf{R} \cup \{+\infty\}$ be a closed proper convex function defined by

$$f(x_1, x_2) = \begin{cases} \dfrac{(2x_2 - 1)^2}{x_1 + 1} & (\text{if } x_1 \geq 0 \text{ and } 0 \leq x_2 \leq 1), \\ +\infty & (\text{otherwise}). \end{cases}$$

We have $\arg\min_{\mathbf{R}} f = \{(x_1, 0.5) \mid x_1 \in \mathbf{R},\ x_1 \geq 0\} \neq \emptyset$. On the other hand, $f(y) > 0$ for all integral vectors $y \in \mathbf{Z}^n$ and $\inf\{f(y) \mid y \in \mathbf{Z}^2\} = \inf\{1/(y_1 + 1) \mid y_1 \in \mathbf{Z}_+\} = 0$. This shows that $\arg\min\{f(y) \mid y \in \mathbf{Z}^2\} = \emptyset$ holds.

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

$$g(x_1, x_2) = \begin{cases} 1/(x_1 + 1) & (\text{if } x_1 \geq 0 \text{ and } x_2 = \sqrt{2} \cdot x_1), \\ +\infty & (\text{otherwise}). \end{cases}$$

Then, $\arg\min\{g(y) \mid y \in \mathbf{Z}^2\} = \{(0, 0)\}$ since $(0, 0)$ is the unique integral vector with finite function value of g. On the other hand, $g(x) > 0$ for all $x \in \mathbf{R}^2$ and $\inf g = \inf\{1/(x_1 + 1) \mid x_1 \geq 0\} = 0$. This shows that $\arg\min_{\mathbf{R}} g = \emptyset$ holds. $\square$

Remark 4.5. 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 $f_\alpha(x) = f(\alpha x)$ $(x \in \mathbf{R}^n)$. Then, f_α is a closed proper M-convex function as well [24, Theorem 6.49 (2)]. Theorem 4.1 applied to f_α can be restated in terms of f as follows, which are seemingly more general but equivalent. $\square$

Corollary 4.6. *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ be a closed proper M-convex function. For $\alpha > 0$, we define $f_{\alpha\mathbf{Z}} : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ by*

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

(i) *For every $y_* \in \arg\min_{\mathbf{Z}} f_{\alpha\mathbf{Z}}$, there exists some $x_* \in \arg\min_{\mathbf{R}} f$ such that $\|x_* - \alpha y_*\|_\infty < \alpha(n - 1)$.*
(ii) *For every $x_* \in \arg\min_{\mathbf{R}} f$, there exists some $y_* \in \arg\min_{\mathbf{Z}} f_{\alpha\mathbf{Z}}$ such that $\|\alpha y_* - x_*\|_\infty < \alpha(n - 1)$.*
(iii) *We have $\arg\min_{\mathbf{R}} f \neq \emptyset$ if and only if $\arg\min_{\mathbf{Z}} f_{\alpha\mathbf{Z}} \neq \emptyset$.*

4.2 Proof of Theorem 1.5

We prove Theorem 1.5, a proximity theorem for (SC). It is known that the set of feasible solutions of (SC) is represented as the M-convex set $B(\rho_+) \cap \mathbf{Z}^n$ with a submodular function $\rho_+ : 2^N \to \mathbf{Z} \cup \{+\infty\}$ given by $\rho_+(Y) = \min\{\rho(Z) \mid Z \supseteq Y\}$ $(Y \subseteq N)$ (see [6, Section 3.1 (b)]). Hence, the problem (SC) is rewritten as follows:

$$\text{Minimize} \quad \textstyle\sum_{i=1}^n f_i(x(i)) \qquad \text{subject to} \quad x \in B(\rho_+) \cap \mathbf{Z}^n.$$

Based on the observation above, we consider the problem of minimizing a separable convex function on an M-convex set, which is (slightly) more general than (SC):

$$(\mathsf{GSC}) \qquad \text{Minimize} \quad \textstyle\sum_{i=1}^n f_i(x(i)) \qquad \text{subject to} \quad x \in B(\rho) \cap \mathbf{Z}^n,$$

where $f_i : \mathbf{R} \to \mathbf{R}$ $(i \in N)$ is a family of univariate convex functions, and $\rho : 2^N \to \mathbf{Z} \cup \{+\infty\}$ is an integer-valued submodular function satisfying $\rho(\emptyset) = 0$ and $\rho(N) < +\infty$. Continuous relaxation of (GSC) can be naturally defined as follows:

$$(\overline{\mathsf{GSC}}) \qquad \text{Minimize} \quad \textstyle\sum_{i=1}^n f_i(x(i)) \qquad \text{subject to} \quad x \in B(\rho).$$

We will show a proximity result for the problem (GSC).

Theorem 4.7.
(i) *For every optimal solution $y_* \in \mathbf{Z}^n$ to* (GSC)*, there exists an optimal solution $x_* \in \mathbf{R}^n$ to* ($\overline{\mathsf{GSC}}$) *such that $\|x_* - y_*\|_1 < 2(n-1)$.*
(ii) *For every optimal solution $x_* \in \mathbf{R}^n$ to* ($\overline{\mathsf{GSC}}$)*, there exists an optimal solution $y_* \in \mathbf{Z}^n$ to* (GSC) *such that $\|y_* - x_*\|_1 < 2(n-1)$.*

Then, Theorem 1.5 is an immediate consequence of this.

In the following, we prove Theorem 4.7, where we use the following properties concerning the problem (GSC). For $x \in \mathbf{R}^n$ we denote

$$f_{\mathrm{sum}}(x) = \sum_{i \in N} f_i(x(i)).$$

Lemma 4.8. *Let $\varphi : \mathbf{R} \to \mathbf{R}$ be a univariate convex function. For $\alpha, \beta \in \mathbf{R}$ with $\alpha < \beta$ and $\varepsilon \in \mathbf{R}$ with $0 \le \varepsilon \le \beta - \alpha$, it holds that $\varphi(\alpha) + \varphi(\beta) \ge \varphi(\alpha + \varepsilon) + \varphi(\beta - \varepsilon)$.*

Lemma 4.9 (cf. [29])**.** *For $x, y \in \mathbf{R}^n$, $i \in \mathrm{supp}^+(x - y)$, and $j \in \mathrm{supp}^-(x - y)$, it holds that*

$$f_{\mathrm{sum}}(x) + f_{\mathrm{sum}}(y) \ge f_{\mathrm{sum}}(x - \alpha(\chi_i - \chi_j)) + f_{\mathrm{sum}}(y + \alpha(\chi_i - \chi_j))$$

for every $\alpha \in \mathbf{R}$ with $0 \le \alpha \le \min\{x(i) - y(i), y(j) - x(j)\}$.

Proof. The inequality follows from Lemma 4.8. □

4.2.1 Proof of Theorem 4.7 (i)

Let $y_* \in B(\rho) \cap \mathbf{Z}^n$ be an optimal solution to (GSC). Let $x_* \in B(\rho)$ be an optimal solution to ($\overline{\mathsf{GSC}}$) minimizing the value $\|x_* - y_*\|_1$ among all optimal solutions to ($\overline{\mathsf{GSC}}$). In the following, we show by induction on the cardinality of the set $\mathrm{supp}(x_* - y_*) = \{i \in N \mid x_*(i) \neq y_*(i)\}$ that

$$\|x_* - y_*\|_1 < 2(|\mathrm{supp}(x_* - y_*)| - 1) \qquad \text{if } x_* \neq y_*. \tag{4.8}$$

This implies the statement (i) since $|\mathrm{supp}(x_* - y_*)| \leq n$. We note that if $x_* \neq y_*$ then $|\mathrm{supp}(x_* - y_*)| \geq 2$ since $x_*(N) = y_*(N)$.

Assume that $x_* \neq y_*$. Put

$$L = \{i \in N \mid |x_*(i) - y_*(i)| \geq 1\}.$$

We consider the following three cases.

Case 1: $\mathrm{supp}^-(x_* - y_*) \cap L = \emptyset$,
Case 2: $\mathrm{supp}^-(x_* - y_*) \cap L \neq \emptyset$ and $\mathrm{supp}^+(x_* - y_*) \setminus L = \emptyset$,
Case 3: $\mathrm{supp}^+(x_* - y_*) \setminus L \neq \emptyset$.

(Case 1) By assumption, we have

$$|x_*(j) - y_*(j)| < 1 \qquad (\forall j \in \mathrm{supp}^-(x_* - y_*)). \tag{4.9}$$

Since $x_*(N) = y_*(N) = \rho(N)$, we have

$$\sum\{|x_*(i) - y_*(i)| \mid i \in \mathrm{supp}^+(x_* - y_*)\} = \sum\{|x_*(j) - y_*(j)| \mid j \in \mathrm{supp}^-(x_* - y_*)\}. \tag{4.10}$$

We also have

$$|\mathrm{supp}^-(x_* - y_*)| = |\mathrm{supp}(x_* - y_*)| - |\mathrm{supp}^+(x_* - y_*)| \leq |\mathrm{supp}(x_* - y_*)| - 1. \tag{4.11}$$

Therefore, it holds that

$$\|x_* - y_*\|_1 \;=\; 2 \sum_{j \in \mathrm{supp}^-(x_* - y_*)} |x_*(j) - y_*(j)| \;<\; 2|\mathrm{supp}^-(x_* - y_*)| \;\leq\; 2(|\mathrm{supp}(x_* - y_*)| - 1),$$

where the equality is by (4.10) and the first and the second inequality are by (4.9) and by (4.11), respectively. Hence, (4.8) holds.

(Case 2) Let $j \in \mathrm{supp}^-(x_* - y_*) \cap L$. Since $j \in \mathrm{supp}^+(y_* - x_*)$, Theorem 2.1 (i) implies that there exist some $i \in \mathrm{supp}^-(y_* - x_*)$ and a sufficiently small $\varepsilon > 0$ such that

$$y_* - \varepsilon(\chi_j - \chi_i) \in B(\rho), \qquad x_* + \varepsilon(\chi_j - \chi_i) \in B(\rho). \tag{4.12}$$

Since $y_* \in B(\rho) \cap \mathbf{Z}^n$ and ρ is an integer-valued function, the property $y_* - \varepsilon(\chi_j - \chi_i) \in B(\rho)$ implies $y_* - (\chi_j - \chi_i) \in B(\rho) \cap \mathbf{Z}^n$. In the following, we derive a contradiction by showing that $f_{\mathrm{sum}}(y_* - (\chi_j - \chi_i)) < f_{\mathrm{sum}}(y_*)$.

Note that $i \in \mathrm{supp}^-(y_* - x_*) = \mathrm{supp}^+(x_* - y_*) \subseteq L$ holds by the assumption of Case 2. Since $\|(x_* + \varepsilon(\chi_j - \chi_i)) - y_*\|_1 < \|x_* - y_*\|_1$, we have

$$f_{\mathrm{sum}}(x_* + \varepsilon(\chi_j - \chi_i)) > f_{\mathrm{sum}}(x_*) \tag{4.13}$$

by the choice of x_*. It follows from the convexity of $f_{\rm sum}$ and the inequality (4.13) that

$$f_{\rm sum}(x_* + (\chi_j - \chi_i)) > f_{\rm sum}(x_*). \tag{4.14}$$

By Lemma 4.9, it holds that

$$f_{\rm sum}(y_*) + f_{\rm sum}(x_*) \geq f_{\rm sum}(y_* - (\chi_j - \chi_i)) + f_{\rm sum}(x_* + (\chi_j - \chi_i)) \tag{4.15}$$

since $y_*(j) \geq x_*(j) + 1$ and $y_*(i) \leq x_*(i) - 1$. It follows from (4.14) and (4.15) that $f_{\rm sum}(y_* - (\chi_j - \chi_i)) < f_{\rm sum}(y_*)$, a contradiction to the optimality of y_* to (GSC). This means that Case 2 does not occur.

(Case 3) By assumption, there exists some $k \in {\rm supp}^+(x_* - y_*) \setminus L$. Let

$$B = \{x \in \mathbf{R}^n \mid x \in B(\rho),\ x(i) = y_*(i)\ (\forall i \in N \setminus {\rm supp}(x_* - y_*)),\ x(k) \leq y_*(k)\}.$$

Then, the set B is an integral base polyhedron and $B \cap \mathbf{Z}^n$ is an M-convex set (see, e.g., [6, Section 3.1 (b)]). Note that $y_* \in B$ and $x_* \notin B$. We consider the problem (GSC') and its continuous relaxation $(\overline{\mathsf{GSC}'})$:

$$\begin{array}{llll} (\mathsf{GSC}') & \text{Minimize} & \sum_{i=1}^n f_i(x(i)) & \text{subject to} \quad x \in B \cap \mathbf{Z}^n, \\ (\overline{\mathsf{GSC}'}) & \text{Minimize} & \sum_{i=1}^n f_i(x(i)) & \text{subject to} \quad x \in B. \end{array}$$

Then, y_* is an optimal solution to (GSC') since $y_* \in B$ and $B \subseteq B(\rho)$. This, together with Theorem 1.3, implies that there exists an optimal solution to $(\overline{\mathsf{GSC}'})$. Let S $(\subseteq B)$ be the set of optimal solutions to $(\overline{\mathsf{GSC}'})$, and S_* be the set of optimal solutions x to $(\overline{\mathsf{GSC}'})$ minimizing the value $\|x - y_*\|_1$, i.e.,

$$S_* = \{x \in S \mid \|x - y_*\|_1 \leq \|x' - y_*\|_1\ (\forall x' \in S)\}.$$

As shown later, there exists some $\tilde{x} \in S_*$ satisfying the following conditions:

$$\tilde{x}(k) = y_*(k), \tag{4.16}$$
$${\rm supp}^+(x_* - \tilde{x}) = \{k\}. \tag{4.17}$$

Since $x_*(N) = \tilde{x}(N) = \rho(N)$, it holds that

$$\sum\{|x_*(j) - \tilde{x}(j)| \mid j \in {\rm supp}^-(x_* - \tilde{x})\} = \sum\{|x_*(i) - \tilde{x}(i)| \mid i \in {\rm supp}^+(x_* - \tilde{x})\} = |x_*(k) - \tilde{x}(k)|,$$

where the last equality is by (4.17). This equation and (4.16) imply

$$\|x_* - \tilde{x}\|_1 = 2|x_*(k) - \tilde{x}(k)| = 2|x_*(k) - y_*(k)| < 2, \tag{4.18}$$

where the last inequality follows from $k \notin L$. If $\tilde{x} = y_*$, then the inequality (4.18) implies

$$\|x_* - y_*\|_1 = \|x_* - \tilde{x}\|_1 < 2 \leq 2(|{\rm supp}(x_* - y_*)| - 1),$$

i.e., (4.8) holds. Hence, we assume $\tilde{x} \neq y_*$. Since $\tilde{x} \in B$, we have ${\rm supp}(\tilde{x} - y_*) \subseteq {\rm supp}(x_* - y_*)$, which, together with (4.16), implies ${\rm supp}(\tilde{x} - y_*) \subseteq {\rm supp}(x_* - y_*) \setminus \{k\}$. Hence, we can apply the induction hypothesis to the problem (GSC') and vectors $\tilde{x}, y_*$ to obtain the inequality

$$\|\tilde{x} - y_*\|_1 < 2(|{\rm supp}(\tilde{x} - y_*)| - 1). \tag{4.19}$$

The inequality (4.8) follows from (4.18) and (4.19) as follows:

$$\begin{aligned}\|x_* - y_*\|_1 &\leq \|x_* - \tilde{x}\|_1 + \|\tilde{x} - y_*\|_1 \\ &< 2 + 2(|\mathrm{supp}(\tilde{x} - y_*)| - 1) \ \leq \ 2(|\mathrm{supp}(x_* - y_*)| - 1).\end{aligned}$$

It remains to show that there exists some $\tilde{x} \in S_*$ satisfying (4.16) and (4.17).

Let $\tilde{x} \in S_*$ be a vector maximizing the value $\tilde{x}(k)$ among all vectors in S_*, and assume, to the contrary, that $\tilde{x}(k) < y_*(k)$. Since $x_*(k) > y_*(k)$, we have $k \in \mathrm{supp}^+(x_* - \tilde{x})$. Therefore, Theorem 2.1 (i) and Lemma 4.9 imply that there exist some $j \in \mathrm{supp}^-(x_* - \tilde{x})$ and a sufficiently small positive number ε such that $x_* - \varepsilon(\chi_k - \chi_j) \in B(\rho)$, $\tilde{x} + \varepsilon(\chi_k - \chi_j) \in B(\rho)$, and

$$f_{\mathrm{sum}}(x_*) + f_{\mathrm{sum}}(\tilde{x}) \geq f_{\mathrm{sum}}(x_* - \varepsilon(\chi_k - \chi_j)) + f_{\mathrm{sum}}(\tilde{x} + \varepsilon(\chi_k - \chi_j)). \tag{4.20}$$

Since x_* is an optimal solution to $(\overline{\mathsf{GSC}})$, it holds that

$$f_{\mathrm{sum}}(x_*) \leq f_{\mathrm{sum}}(x_* - \varepsilon(\chi_k - \chi_j)). \tag{4.21}$$

Since $\tilde{x}(k) < y_*(k)$ and ε is sufficiently small, we have $\tilde{x} + \varepsilon(\chi_k - \chi_j) \in B$. This implies

$$f_{\mathrm{sum}}(\tilde{x}) \leq f_{\mathrm{sum}}(\tilde{x} + \varepsilon(\chi_k - \chi_j)) \tag{4.22}$$

since $\tilde{x}$ is an optimal solution to $(\overline{\mathsf{GSC}'})$. It follows from (4.20), (4.21), and (4.22) that the inequality (4.22) holds with equality, implying that $\tilde{x} + \varepsilon(\chi_k - \chi_j)$ is also an optimal solution to $(\overline{\mathsf{GSC}'})$. Since $\tilde{x}(k) < \tilde{x}(k) + \varepsilon \leq y_*(k)$, we have $\|(\tilde{x} + \varepsilon(\chi_k - \chi_j)) - y_*\|_1 \leq \|\tilde{x} - y_*\|_1$. This implies $\tilde{x} + \varepsilon(\chi_k - \chi_j) \in S_*$ since $\tilde{x} \in S_*$. This, however, is a contradiction since $\tilde{x}$ maximizes the value $\tilde{x}(k)$ among all vectors in S_*. Hence, $\tilde{x}$ satisfies (4.16).

Let $\tilde{x}$ be a vector in S_* satisfying (4.16), and assume that $\tilde{x}$ minimizes the value

$$\sum_{i \in N} \max\{x_*(i) - \tilde{x}(i), 0\}$$

among all such vectors. We will show that the vector $\tilde{x}$ chosen in this way satisfies (4.17).

Assume, to the contrary, that $\mathrm{supp}^+(x_* - \tilde{x}) \setminus \{k\} \neq \emptyset$, and let $h \in \mathrm{supp}^+(x_* - \tilde{x}) \setminus \{k\}$. Then, Theorem 2.1 (i) and Lemma 4.9 imply that there exist some $j \in \mathrm{supp}^-(x_* - \tilde{x})$ and a sufficiently small positive number ε such that $x_* - \varepsilon(\chi_h - \chi_j) \in B(\rho)$, $\tilde{x} + \varepsilon(\chi_h - \chi_j) \in B(\rho)$, and

$$f_{\mathrm{sum}}(x_*) + f_{\mathrm{sum}}(\tilde{x}) \geq f_{\mathrm{sum}}(x_* - \varepsilon(\chi_h - \chi_j)) + f_{\mathrm{sum}}(\tilde{x} + \varepsilon(\chi_h - \chi_j)). \tag{4.23}$$

Since x_* is an optimal solution to $(\overline{\mathsf{GSC}})$, we have

$$f_{\mathrm{sum}}(x_*) \leq f_{\mathrm{sum}}(x_* - \varepsilon(\chi_h - \chi_j)). \tag{4.24}$$

Since $\tilde{x}$ is an optimal solution to $(\overline{\mathsf{GSC}'})$ and $\tilde{x} + \varepsilon(\chi_h - \chi_j) \in B$, we have

$$f_{\mathrm{sum}}(\tilde{x}) \leq f_{\mathrm{sum}}(\tilde{x} + \varepsilon(\chi_h - \chi_j)). \tag{4.25}$$

It follows from (4.23), (4.24), and (4.25) that the inequalities (4.24) and (4.25) hold with equality. This implies that $x'_* = x_* - \varepsilon(\chi_h - \chi_j)$ is an optimal solution to $(\overline{\mathsf{GSC}})$, and $\tilde{x}' = \tilde{x} + \varepsilon(\chi_h - \chi_j)$ is an optimal solution to $(\overline{\mathsf{GSC}'})$. We have $x_*(h) \leq y_*(h)$ or $x_*(j) \geq y_*(j)$ since otherwise $\|x'_* - y_*\|_1 < \|x_* - y_*\|_1$ holds, contradicting the choice of x_*.

Obviously, $\tilde{x}'(k) = \tilde{x}(k) = y_*(k)$ holds. In the following, we show that $\tilde{x}'$ is a vector in S_* satisfying

$$\sum_{i \in N} \max\{x_*(i) - \tilde{x}'(i), 0\} < \sum_{i \in N} \max\{x_*(i) - \tilde{x}(i), 0\}, \tag{4.26}$$

which is a contradiction to the choice of $\tilde{x}$.

It holds that

$$\begin{aligned} |\tilde{x}'(i) - y_*(i)| &= |\tilde{x}(i) - y_*(i)| \quad (\forall i \in N \setminus \{h, j\}), \\ |\tilde{x}'(i) - y_*(i)| &\leq |\tilde{x}(i) - y_*(i)| + \varepsilon \quad (\forall i \in \{h, j\}). \end{aligned}$$

If $x_*(h) \leq y_*(h)$ holds, then we have $\tilde{x}(h) < \tilde{x}(h) + \varepsilon = \tilde{x}'(h) < x_*(h) \leq y_*(h)$, implying $|\tilde{x}'(h) - y_*(h)| = |\tilde{x}(h) - y_*(h)| - \varepsilon$. If $x_*(j) \geq y_*(j)$, then we have $\tilde{x}(j) > \tilde{x}(j) - \varepsilon = \tilde{x}'(j) > x_*(j) \geq y_*(j)$, implying $|\tilde{x}'(j) - y_*(j)| = |\tilde{x}(j) - y_*(j)| - \varepsilon$. Hence, it holds that $\|\tilde{x}' - y_*\|_1 \leq \|\tilde{x} - y_*\|_1$, i.e., $\tilde{x}' \in S_*$. Finally, the inequality (4.26) can be obtained as follows:

$$\begin{aligned} &\sum_{i \in N} \max\{x_*(i) - \tilde{x}'(i), 0\} - \sum_{i \in N} \max\{x_*(i) - \tilde{x}(i), 0\} \\ &\quad = \big[\max\{x_*(h) - \tilde{x}'(h), 0\} - \max\{x_*(h) - \tilde{x}(h), 0\}\big] \\ &\qquad\quad + \big[\max\{x_*(j) - \tilde{x}'(j), 0\} - \max\{x_*(j) - \tilde{x}(j), 0\}\big] \\ &\quad = \{x_*(h) - \tilde{x}(h) - \varepsilon\} - \{x_*(h) - \tilde{x}(h)\} \ = \ -\varepsilon \ < \ 0, \end{aligned}$$

where the second equality follows from $h \in \mathrm{supp}^+(x_* - \tilde{x})$, $j \in \mathrm{supp}^-(x_* - \tilde{x})$, and the fact that ε is a sufficiently small positive number.

Hence, $\tilde{x}$ satisfies (4.16) and (4.17). This concludes the proof of Theorem 4.7 (i).

4.2.2 Proof of Theorem 4.7 (ii)

To prove the statement (ii) of Theorem 4.7, we use the statement (i) which is already shown in Section 4.2.1. Let $x_* \in B(\rho)$ be an optimal solution to $(\overline{\mathsf{GSC}})$. Let $y_* \in B(\rho) \cap \mathbf{Z}^n$ be an optimal solution to (GSC) minimizing the value $\|x_* - y_*\|_1$ among all optimal solutions to (GSC). With a positive number δ we define a new problem:

(GSC^δ) Minimize $\sum_{i=1}^{n} \{f_i(x(i)) + \delta |x(i) - x_*(i)|\}$ subject to $x \in B(\rho) \cap \mathbf{Z}^n$.

This problem is also the minimization of a separable convex function on an M-convex set. It is easy to see that x_* is the unique optimal solution to the continuous relaxation of (GSC^δ). In addition, if δ is a sufficiently small positive number, then y_* is an optimal solution to (GSC^δ). Hence, by applying Theorem 4.7 (i) to the problem (GSC^δ) and its continuous relaxation, we obtain $\|x_* - y_*\|_1 < 2(n-1)$.

4.3 Proof of Theorem 1.6

4.3.1 A key lemma

To prove Theorem 1.6, a proximity theorem for $(\mathsf{Laminar})$, we first show a useful lemma. Let $\mathcal{F} \subseteq 2^N$ be a laminar family satisfying the condition (2.2). Recall the definitions of the parent and a child of $X \in \mathcal{F}$ in Section 2.3. For distinct elements $i, j \in N$, a minimal set in $\mathcal{F}$

containing both of i, j is uniquely determined, and we call it the *lowest common ancestor* of i and j. For every distinct $i, j \in N$, a *(directed) path from i to j (in $\mathcal{F}$)* is defined as a sequence $S_0, S_1, \ldots, S_t$ $(t \geq 2)$ of sets in $\mathcal{F}$ satisfying the following conditions:

(i) $S_0 = \{i\}$, $S_t = \{j\}$,
(ii) there exists some k with $1 \leq k \leq t-1$ such that S_k is the lowest common ancestor of i and j,
(iii) for $h = 0, 1, \ldots, k-1$, the set S_{h+1} is the parent of S_h,
(iv) for $h = k+1, k+2, \ldots, t$, the set S_{h-1} is the parent of S_h.

Note that a path from i to j is uniquely determined.

Let $z \in \mathbf{R}^n$, and $i, j \in N$ be distinct elements. We define the *capacity* $\mathrm{cap}(i, j, \mathcal{F}, z)$ *with respect to i, j, $\mathcal{F}$, and z* by

$$\mathrm{cap}(i, j, \mathcal{F}, z) = \min\left[\min_{0 \leq h \leq k-1} \max(z(S_h), 0), \min_{k+1 \leq h \leq t} \max(-z(S_h), 0)\right],$$

where $S_0, S_1, \ldots, S_t$ is the path from i to j and S_k is the lowest common ancestor of i and j. Note that $\mathrm{cap}(i, j, \mathcal{F}, z) > 0$ if and only if

$$z(S_h) > 0 \ (h = 0, 1, \ldots, k-1), \qquad z(S_h) < 0 \ (h = k+1, k+2, \ldots, t).$$

We now give the statement of the key lemma.

Lemma 4.10. *Let $\mathcal{F} \subseteq 2^N$ be a laminar family satisfying the condition* (2.2). *Let $z \in \mathbf{R}^n$ be a vector with $z(N) = 0$, and suppose that* $\mathrm{cap}(i, j, \mathcal{F}, z) < 1$ *for every distinct elements $i, j \in N$. Then, $\|z\|_1 \leq 2(n-1)$ holds; the strict inequality $\|z\|_1 < 2(n-1)$ holds if $n \geq 2$.*

Proof. We show the statement of the lemma by induction on the cardinality n of the ground set N. If $n = 1$, then we have $z(1) = z(N) = 0$, implying $\|z\|_1 = 0 = 2(n-1)$. Hence, we consider the case with $n \geq 2$.

Claim 1: For every $i \in \mathrm{supp}^+(z)$, there exists some $j \in \mathrm{supp}^-(z)$ such that $\mathrm{cap}(i, j, \mathcal{F}, z) > 0$.

[Proof of Claim 1] Let $X \subseteq N$ be the unique minimal set in $\mathcal{F}$ satisfying $z(X) \leq 0$ and $i \in X$. Such X always exists since $z(N) = 0$ and $N \in \mathcal{F}$. Note that $|X| \geq 2$ since $z(i) > 0$. Let $S_0, S_1, \ldots, S_k$ $(k \geq 1)$ be the sequence of sets such that $S_0 = \{i\}$, $S_k = X$, and S_h is the parent of S_{h-1} $(h = 1, 2, \ldots, k)$. Then, we have $z(S_h) > 0$ $(h = 0, 1, \ldots, k-1)$ by the choice of X. Since $z(S_k) = z(X) \leq 0$ and $z(S_{k-1}) > 0$, there exists another child S_{k+1} of X such that $z(S_{k+1}) < 0$. If $|S_{k+1}| \geq 2$, then there exists some child $S_{k+2} \in \mathcal{F}$ of S_{k+1} such that $z(S_{k+2}) < 0$. Repeating this, we can obtain a sequence of sets $S_k, S_{k+1}, \ldots, S_{t-1}, S_t$ such that $|S_t| = 1$, $z(S_h) < 0$ $(h = k+1, k+2, \ldots, t)$, and S_h is a child of S_{h-1} $(h = k+1, k+2, \ldots, t)$. Let $j \in N$ be the unique element in S_t. Then, the observation above implies that $j \in \mathrm{supp}^-(z)$ and $\mathrm{cap}(i, j, \mathcal{F}, z) > 0$. [End of the proof of Claim 1]

We may assume that $z \neq \mathbf{0}$, and let $i_* \in \mathrm{supp}^+(z)$. By Claim 1, there exists some $j_* \in \mathrm{supp}^-(z)$ such that $\mathrm{cap}(i_*, j_*, \mathcal{F}, z) > 0$. Let $\tilde{S}_0, \tilde{S}_1, \ldots, \tilde{S}_t$ be the path from i_* to j_*, and $\tilde{S}_k$ $(1 \leq k \leq t-1)$ be the lowest common ancestor of i_* and j_*.

Define $\tilde{z} = z - \delta(\chi_{i_*} - \chi_{j_*})$ with $\delta = \mathrm{cap}(i_*, j_*, \mathcal{F}, z)$ (> 0). Then, we have $\|\tilde{z}\|_1 = \|z\|_1 - 2\delta$.

Claim 2: For every distinct $i, j \in N$, it holds that $\mathrm{cap}(i, j, \mathcal{F}, \tilde{z}) \leq \mathrm{cap}(i, j, \mathcal{F}, z)$ (< 1).

[Proof of Claim 2] It suffices to show that for every $X \in \mathcal{F}$, we have $0 \leq \tilde{z}(X) \leq z(X)$ if $z(X) \geq 0$ and $0 \geq \tilde{z}(X) \geq z(X)$ if $z(X) < 0$. We see from the definition of a path that $\tilde{S}_0, \ldots, \tilde{S}_{k-1}$ are those sets in $\mathcal{F}$ which contain i_* and not j_*. Therefore, for $X \in \{\tilde{S}_0, \ldots, \tilde{S}_{k-1}\}$, we have $z(X) \geq \delta$, and therefore $\tilde{z}(X) = z(X) - \delta \geq 0$. Similarly, each set $X \in \{\tilde{S}_{k+1}, \ldots, \tilde{S}_t\}$ contains j_* and not i_*, and therefore we have $z(X) \leq -\delta$ and $\tilde{z}(X) = z(X) + \delta \leq 0$. Finally, each $X \in \mathcal{F}$ not in $\{\tilde{S}_0, \ldots, \tilde{S}_{k-1}\} \cup \{\tilde{S}_{k+1}, \ldots, \tilde{S}_t\}$ contains both of i_* and j_*, or neither of i_* and j_*, implying that $\tilde{z}(X) = z(X)$. [End of the proof of Claim 2]

By the definition of $\delta = \mathrm{cap}(i_*, j_*, \mathcal{F}, z)$, there exists some $v \in \{0, \ldots, k-1\} \cup \{k+1, \ldots, t\}$ such that $|z(\tilde{S}_v)| = \delta$. This implies $\tilde{z}(\tilde{S}_v) = 0$. Let $N' = \tilde{S}_v$ and $N'' = N \setminus \tilde{S}_v$. Since $|\tilde{S}_v \cap \{i_*, j_*\}| = 1$, we have $1 \leq |N'| < n$ and $1 \leq |N''| < n$. We define new families $\mathcal{F}' \subseteq 2^{N'}$ and $\mathcal{F}'' \subseteq 2^{N''}$ as follows:

$$
\begin{aligned}
\mathcal{F}' &= \{Y \mid Y \in \mathcal{F},\ Y \subseteq \tilde{S}_v\}, \\
\mathcal{F}'' &= \{Y \mid Y \in \mathcal{F},\ Y \cap \tilde{S}_v = \emptyset\} \cup \{Y \setminus \tilde{S}_v \mid Y \in \mathcal{F},\ Y \supset \tilde{S}_v\}.
\end{aligned}
$$

Then, $\mathcal{F}'$ and $\mathcal{F}''$ are laminar families satisfying the condition (2.2). We also define vectors $z' \in \mathbf{R}^{N'}$ and $z'' \in \mathbf{R}^{N''}$ by $z'(i) = \tilde{z}(i)$ $(i \in N')$ and $z''(i) = \tilde{z}(i)$ $(i \in N'')$. The following claim can be shown easily from the observation that for every $i, j \in N'$ (resp., $i, j \in N''$) with $i \neq j$, the path in $\mathcal{F}'$ (resp., in $\mathcal{F}''$) from i to j corresponds to the path in $\mathcal{F}$ from i to j.

Claim 3: For every distinct $i, j \in N'$, it holds that $\mathrm{cap}(i, j, \mathcal{F}', z') = \mathrm{cap}(i, j, \mathcal{F}, \tilde{z})$. For every distinct $i, j \in N''$, it holds that $\mathrm{cap}(i, j, \mathcal{F}'', z'') = \mathrm{cap}(i, j, \mathcal{F}, \tilde{z})$.

Claims 2 and 3 imply that $\mathrm{cap}(i, j, \mathcal{F}', z') < 1$ for every distinct $i, j \in N'$ and $\mathrm{cap}(i, j, \mathcal{F}'', z'') < 1$ for every distinct $i, j \in N''$. Hence, we can apply the induction hypothesis to obtain $\|z'\|_1 \leq 2(|N'| - 1)$ and $\|z''\|_1 \leq 2(|N''| - 1)$. Since $\delta < 1$, we have

$$
\begin{aligned}
\|z\|_1 &= 2\delta + \|\tilde{z}\|_1 = 2\delta + \|z'\|_1 + \|z''\|_1 \\
&< 2 + 2(|N'| - 1) + 2(|N''| - 1) = 2(|N'| + |N''| - 1) = 2(n - 1).
\end{aligned}
$$

□

4.3.2 Proof of Theorem 1.6 (i)

Recall that $n \geq 2$ in the problem ($\mathsf{Laminar}$). Let $y_* \in \mathbf{Z}^n$ be an optimal solution to ($\mathsf{Laminar}$), and $x_* \in \mathbf{R}^n$ be an optimal solution to ($\overline{\mathsf{Laminar}}$) which minimizes the value $\|x_* - y_*\|_1$ among all optimal solutions to ($\overline{\mathsf{Laminar}}$). We may assume $x_* \neq y_*$ since otherwise the statement (i) holds immediately. Put $z = x_* - y_*$. Then, it holds that $z(N) = x_*(N) - y_*(N) = K - K = 0$. We will show that for every distinct $i, j \in N$ we have $\mathrm{cap}(i, j, \mathcal{F}, z) < 1$. Then, the inequality $\|x_* - y_*\|_1 < 2(n-1)$ follows from Lemma 4.10.

Let $S_0, S_1, S_2, \ldots, S_t$ be the path from i to j, and S_k $(1 \leq k \leq t-1)$ be the lowest common ancestor of i and j. Assume, to the contrary, that $\mathrm{cap}(i, j, \mathcal{F}, z) \geq 1$. Then, we have $z(S_h) \geq 1$ for $h = 0, 1, \ldots, k-1$ and $z(S_h) \leq -1$ for $h = k+1, k+2, \ldots, t$.

We define vectors $x' = x_* - \chi_i + \chi_j$ and $y' = y_* + \chi_i - \chi_j$. Then, we have

$$
\begin{aligned}
& u_X \geq x_*(X) > x'(X) = x_*(X) - 1 \geq y_*(X) \geq \ell_X \quad (\forall X \in \{S_0, \ldots, S_{k-1}\}), \\
& \ell_X \leq y_*(X) < y'(X) = y_*(X) + 1 \leq x_*(X) \leq u_X \quad (\forall X \in \{S_0, \ldots, S_{k-1}\}), \\
& \ell_X \leq x_*(X) < x'(X) = x_*(X) + 1 \leq y_*(X) \leq u_X \quad (\forall X \in \{S_{k+1}, \ldots, S_t\}), \\
& u_X \geq y_*(X) > y'(X) = y_*(X) - 1 \geq x_*(X) \geq \ell_X \quad (\forall X \in \{S_{k+1}, \ldots, S_t\}), \\
& x'(X) = x_*(X),\ y'(X) = y_*(X) \quad (\forall X \in \mathcal{F},\ X \notin \{S_0, \ldots, S_{k-1}\} \cup \{S_{k+1}, \ldots, S_t\}).
\end{aligned}
$$

Hence, y' and x' are feasible solutions to (Laminar) and ($\overline{\text{Laminar}}$), respectively. The inequalities above and Lemma 4.8 imply

$$
f_X(x_*(X)) + f_X(y_*(X)) \geq f_X(x'(X)) + f_X(y'(X)) \qquad (\forall X \in \mathcal{F}),
$$

from which follows

$$
\sum_{X \in \mathcal{F}} f_X(x_*(X)) + \sum_{X \in \mathcal{F}} f_X(y_*(X)) \geq \sum_{X \in \mathcal{F}} f_X(x'(X)) + \sum_{X \in \mathcal{F}} f_X(y'(X)). \tag{4.27}
$$

Since $\|x' - y_*\| < \|x_* - y_*\|_1$, we have $\sum_{X \in \mathcal{F}} f_X(x_*(X)) < \sum_{X \in \mathcal{F}} f_X(x'(X))$, which, together with (4.27), implies $\sum_{X \in \mathcal{F}} f_X(y_*(X)) > \sum_{X \in \mathcal{F}} f_X(y'(X))$, a contradiction to the optimality of y_* to (Laminar).

4.3.3 Proof of Theorem 1.6 (ii)

Let $x_* \in \mathbf{R}^n$ be an optimal solution to ($\overline{\text{Laminar}}$), and $y_* \in \mathbf{Z}^n$ be an optimal solution to (Laminar) which minimizes the value $\|x_* - y_*\|_1$ among all optimal solutions to (Laminar). We assume $x_* \neq y_*$, and put $z = x_* - y_*$. Then, we can show in a similar way as in Section 4.3.2 that $\mathrm{cap}(i, j, \mathcal{F}, z) < 1$ for every distinct $i, j \in N$, where we use the fact that $\sum_{X \in \mathcal{F}} f_X(y_*(X)) < \sum_{X \in \mathcal{F}} f_X(y'(X))$ holds for every $y' \in \mathbf{Z}^n$ with $\|x_* - y'\| < \|x_* - y_*\|_1$. Hence, the inequality $\|z\|_1 = \|x_* - y_*\|_1 < 2(n-1)$ follows from Lemma 4.10.

Acknowledgements

The authors thank Kazuo Murota and the anonymous referees for their valuable comments on the manuscript. In particular, the first example in Remark 4.4 is given by one of the referees, which is simpler and better than our original example. This work is partially supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

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

[2] P. Brucker, An $O(n)$ algorithm for quadratic knapsack problems, Oper. Res. Lett., 3 (1984), pp. 163–166.

[3] A. W. M. Dress and W. Wenzel, Valuated matroids, Adv. Math., 93 (1992), pp. 214–250.

[4] J. Edmonds, Submodular functions, matroids, and certain polyhedra, in Combinatorial Structures and their Applications, Gordon and Breach, New York, 1970, pp. 69–87.

[5] G. N. Frederickson and D. B. Johnson, The complexity of selection and ranking in $X+Y$ and matrices with sorted columns, J. Comput. Syst. Sci. 24 (1982), pp. 197–208.

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

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

[8] A. V. Goldberg and R. E. Tarjan, A new approach to the maximum-flow problem, J. ACM, 35 (1988), pp. 921–940.

[9] F. Granot and J. Skorin-Kapov, Some proximity and sensitivity results in quadratic integer programming, Math. Program., 47 (1990), pp. 259–268.

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

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

[12] 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.

[13] D. S. Hochbaum and J. G. Shanthikumar, Convex separable optimization is not much harder than linear optimization, J. ACM, 37 (1990), pp. 843–862.

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

[15] S. Iwata and M. Shigeno, Conjugate scaling algorithm for Fenchel-type duality in discrete convex optimization, SIAM J. Optim., 13 (2003), pp. 204–211.

[16] 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.

[17] I. Michaeli and M. A. Pollatschek, On some nonlinear knapsack problems, Ann. Discrete Math., 1 (1977), pp. 403–414.

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

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

[20] S. Moriguchi and N. Tsuchimura, Discrete L-convex function minimization based on continuous relaxation, Pacific J. Optimi., 5 (2009), pp. 227–236.

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

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

[23] K. Murota, Submodular flow problem with a nonseparable cost function, Combinatorica, 19 (1999), pp. 87–109.

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

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

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

[27] K. Murota and A. Shioura, M-convex function on generalized polymatroid., Math. Oper. Res., 24 (1999), pp. 95–105.

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

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

[30] 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.

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

[32] R. T. Rockafellar, Convex Analysis, Princeton Univ. Press, Princeton, NJ, 1970.

[33] A. Schrijver, Theory of Linear and Integer Programming, Wiley, New York, NY, 1986.

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

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

[36] A. Tamir, A strongly polynomial algorithm for minimum convex separable quadratic cost flow problems on series-parallel networks, Math. Program., 59 (1993), pp. 117–132.

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

[38] I. J. Weinsten and O. S. Yu, Comment on an integer maximization problem, Oper. Res., 21 (1973), pp. 648–650.