SIAM J. OPTIM.
Vol. 14, No. 3, pp. 699–707

ON STEEPEST DESCENT ALGORITHMS FOR DISCRETE CONVEX FUNCTIONS*

KAZUO MUROTA†

Abstract. This paper investigates the complexity of steepest descent algorithms for two classes of discrete convex functions: M-convex functions and L-convex functions. Simple tie-breaking rules yield complexity bounds that are polynomials in the dimension of the variables and the size of the effective domain. Combining the present results with a standard scaling approach leads to an efficient algorithm for L-convex function minimization.

Key words. discrete optimization, discrete convex function, steepest descent algorithm, M-convex function, L-convex function

AMS subject classifications. 90C10, 90C25, 90C35, 90C27

DOI. 10.1137/S1052623402419005

1. Introduction.

Discrete convex functions have long been attracting research interest in the area of discrete optimization. Miller [15] was a forerunner in the early 1970s. The relationship between submodularity and convexity was discussed in Edmonds [3], and deeper understanding of this relationship was gained in the 1980s by Frank [5], Fujishige [6], and Lovász [13] (see also [7]). Favati and Tardella [4] introduced integrally convex functions to show a local characterization for global minimality, and Dress and Wenzel [2] considered valuated matroids in terms of a greedy algorithm. Recently, Murota [17, 18, 20, 21] advocated "discrete convex analysis," where M-convex and L-convex functions play central roles. M$^\natural$-convex and L$^\natural$-convex functions,[1] introduced, respectively, by Murota and Shioura [22] and Fujishige and Murota [8], are variants of M-convex and L-convex functions. It was shown in [8] that L$^\natural$-convex functions are the same as the submodular integrally convex functions considered in [4].

Minimization of discrete convex functions is most fundamental in discrete optimization. In fact, we have recently witnessed dramatic progress of algorithms for submodular set-function minimization; see, e.g., Iwata [10], Iwata, Fleischer, and Fujishige [11], Schrijver [23], and a survey by McCormick [14].

M-convex function minimization contains the minimum-weight matroid-base problem (see, e.g., [1]) as a very special case. Minimization of an M-convex function on $\{0, 1\}$-vectors is equivalent to maximization of a matroid valuation, for which the greedy algorithm of Dress and Wenzel [2] works. The first polynomial time algorithm for general M-convex functions was given by Shioura [24], and scaling algorithms were considered by Moriguchi, Murota, and Shioura [16], Tamura [26], and Shioura [25].

For L-convex function minimization the algorithm of Favati and Tardella [4], originally meant for submodular integrally convex functions, works with slight modi-

*Received by the editors November 30, 2002; accepted for publication (in revised form) August 6, 2003; published electronically December 19, 2003. This work was supported by the Kayamori Foundation of Informational Science Advancement, a Grant-in-Aid of the Ministry of Education, Culture, Sports, Science and Technology of Japan, and the 21st Century COE Program on Information Science and Technology Strategic Core.

http://www.siam.org/journals/siopt/14-3/41900.html

†Graduate School of Information Science and Technology, University of Tokyo, and PRESTO, JST, Tokyo 113-8656, Japan (murota@mist.i.u-tokyo.ac.jp).

[1] "M$^\natural$-convex" should be read "M-natural-convex," and similarly for "L$^\natural$-convex."

fications. It is the first polynomial time algorithm for L-convex function minimization, but it is not practical, being based on the ellipsoid method. A steepest descent algorithm was proposed by Murota [19], with a subsequent improvement by Iwata [9] using a scaling technique. The steepest descent algorithm heavily depends on algorithms for submodular set-function minimization.

In this paper we investigate the complexity of steepest descent algorithms for M-convex functions and L-convex functions. With certain simple tie-breaking rules we can obtain complexity bounds that are polynomials in the dimension n of the variables and the size K of the effective domain. Combining the present complexity bound with a standard scaling approach results in an efficient algorithm for L-convex function minimization of complexity bounded by polynomials in n and $\log K$. This is faster than any other known algorithms for L-convex function minimization.

Some conventions are introduced. We consider functions defined on integer lattice points that may possibly take $+\infty$, i.e., $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ with a finite set V of cardinality n. The effective domain of f is denoted by

$$\mathrm{dom}\, f = \{x \in \mathbf{Z}^V \mid f(x) < +\infty\}, \tag{1.1}$$

and the ℓ_1-size of $\mathrm{dom}\, f$ by

$$K_f = \max\{||x - y||_1 \mid x, y \in \mathrm{dom}\, f\}, \tag{1.2}$$

where the ℓ_1-norm of a vector $x = (x(v) \mid v \in V)$ with components indexed by V is designated by

$$||x||_1 = \sum_{v \in V} |x(v)|.$$

For a subset X of V we denote by χ_X the characteristic vector of X; $\chi_X(v)$ equals one or zero according to whether v belongs to X or not. For $u \in V$ we denote $\chi_{\{u\}}$ by χ_u.

2. M-convex function minimization. M-convex functions are defined in terms of a generalization of the exchange axiom for matroids. We say that a function $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom}\, f \neq \emptyset$ is M-convex if it satisfies the exchange axiom

(M-EXC) For $x, y \in \mathrm{dom}\, f$ and $u \in \mathrm{supp}^+(x - y)$, there exists $v \in \mathrm{supp}^-(x - y)$ such that

$$f(x) + f(y) \geq f(x - \chi_u + \chi_v) + f(y + \chi_u - \chi_v). \tag{2.1}$$

The inequality (2.1) implicitly imposes the condition that $x - \chi_u + \chi_v \in \mathrm{dom}\, f$ and $y + \chi_u - \chi_v \in \mathrm{dom}\, f$ for the finiteness of the right-hand side. It follows from (M-EXC) that the effective domain of an M-convex function lies on a hyperplane $\{x \in \mathbf{R}^V \mid \sum_{v \in V} x(v) = r\}$ for some integer r.

Global optimality for an M-convex function is characterized by local optimality.

LEMMA 2.1 (see [17, 20, 21]). *For an M-convex function f and $x \in \mathrm{dom}\, f$, we have*

$$f(x) \leq f(y) \quad (\forall\, y \in \mathbf{Z}^V) \iff f(x) \leq f(x - \chi_u + \chi_v) \quad (\forall\, u, v \in V).$$

This local characterization of global minimality naturally suggests the following algorithm of steepest descent type [16, 19, 24].

STEEPEST DESCENT ALGORITHM FOR AN M-CONVEX FUNCTION f.
S0: Find a vector $x \in \mathrm{dom}\, f$.
S1: Find $u, v \in V$ $(u \neq v)$ that minimize $f(x - \chi_u + \chi_v)$.
S2: If $f(x) \leq f(x - \chi_u + \chi_v)$, then stop ($x$ is a minimizer of f).
S3: Set $x := x - \chi_u + \chi_v$ and go to S1.

Step S1 can be done with n^2 evaluations of function f. At the termination of the algorithm in step S2, x is a global optimum by Lemma 2.1. The function value f decreases monotonically with iterations. This property alone does not ensure finite termination in general, although it does if f is integer-valued and bounded from below.

The following is a key property of the steepest descent algorithm for M-convex functions, showing an upper bound on the number of iterations in terms of the distance to the optimal solution rather than in terms of the function value. We denote by x° the initial vector found in step S0.

LEMMA 2.2. *If f has a unique minimizer, say x^*, the number of iterations is bounded by $||x^\circ - x^*||_1/2$.*

Proof. Put $x' = x - \chi_u + \chi_v$ in step S2. By Lemma 2.3 below we have $x^*(u) \leq x(u)-1 = x'(u)$ and $x^*(v) \geq x(v)+1 = x'(v)$, which implies $||x'-x^*||_1 = ||x-x^*||_1-2$. Note that $||x^\circ - x^*||_1$ is an even integer. $\square$

LEMMA 2.3 (see [24]; see also [21]). *Let $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ be an M-convex function with $\arg\min f \neq \emptyset$. For $x \in \mathrm{dom}\, f \setminus \arg\min f$, let $u, v \in V$ be such that*

$$f(x - \chi_u + \chi_v) = \min_{s,t \in V} f(x - \chi_s + \chi_t).$$

Then $u \neq v$ and there exists $x^* \in \arg\min f$ with

$$x^*(u) \leq x(u) - 1, \qquad x^*(v) \geq x(v) + 1.$$

When given an M-convex function f, which may have multiple minimizers, we consider a perturbation of the function so that we can use Lemma 2.2. Assume now that f has a bounded effective domain of ℓ_1-size K_f in (1.2). We arbitrarily fix a bijection $\varphi : V \to \{1, 2, \ldots, n\}$ to represent an ordering of the elements of V, put $v_i = \varphi^{-1}(i)$ for $i = 1, \ldots, n$, and define a function f_ε by

$$f_\varepsilon(x) = f(x) + \sum_{i=1}^{n} \varepsilon^i \, x(v_i),$$

where $\varepsilon > 0$. This function is M-convex, and, for a sufficiently small ε, it has a unique minimizer that is also a minimizer of f. Suppose that the steepest descent algorithm is applied to the perturbed function f_ε. Since

$$f_\varepsilon(x - \chi_u + \chi_v) = f(x - \chi_u + \chi_v) + \sum_{i=1}^{n} \varepsilon^i x(v_i) - \varepsilon^{\varphi(u)} + \varepsilon^{\varphi(v)}$$

this amounts to employing a tie-breaking rule:

$$\text{Take } (u, v) \text{ that lexicographically minimizes } \Phi(u, v), \tag{2.2}$$

where

$$\Phi(u, v) = \begin{cases} (-1, \varphi(u), -\varphi(v)) & \text{if } \varphi(u) < \varphi(v), \\ (+1, -\varphi(v), \varphi(u)) & \text{if } \varphi(u) > \varphi(v), \end{cases}$$

in case of multiple candidates in step S1 of the steepest descent algorithm applied to f. Combining this observation with Lemma 2.1 yields the following complexity bound, where F_f denotes an upper bound on the time to evaluate f.

THEOREM 2.4. *For an M-convex function f with finite K_f, the number of iterations in the steepest descent algorithm with tie-breaking rule* (2.2) *is bounded by $K_f/2$. Hence, if a vector in* $\mathrm{dom}\, f$ *is given, the algorithm finds a minimizer of f in* $\mathrm{O}(F_f \cdot n^2 K_f)$ *time.*

Although a number of algorithms of smaller theoretical complexity are already known for M-convex function minimization [24, 25, 26], the present analysis is intended to reveal the most fundamental fact about M-convex function minimization. The tie-breaking rule (2.2) as well as the steepest descent algorithm can be adapted to M$^\natural$-convex function minimization.

3. L-convex function minimization. L-convex functions are defined in terms of submodularity on integer lattice points. For integer vectors $p, q \in \mathbf{Z}^V$ we denote by $p \vee q$ and $p \wedge q$ the vectors of componentwise maximum and minimum of p and q, i.e.,

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

We say that a function $g : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom}\, g \neq \emptyset$ is L-convex if it satisfies

$$\text{(SBF)} \qquad g(p) + g(q) \geq g(p \vee q) + g(p \wedge q) \qquad (\forall p, q \in \mathbf{Z}^V),$$

$$\text{(TRF)} \qquad \exists r \in \mathbf{R} \text{ such that } g(p + \mathbf{1}) = g(p) + r \qquad (\forall p \in \mathbf{Z}^V),$$

where $\mathbf{1} = (1, 1, \ldots, 1) \in \mathbf{Z}^V$. In this paper we assume $r = 0$, since otherwise g is not bounded from below and does not have a minimum.

Global optimality for an L-convex function is characterized by local optimality.

LEMMA 3.1 (see [18, 20, 21]). *For an L-convex function g with $r = 0$ in* (TRF) *and $p \in \mathrm{dom}\, g$, we have*

$$g(p) \leq g(q) \quad (\forall\, q \in \mathbf{Z}^V) \iff g(p) \leq g(p + \chi_X) \quad (\forall\, X \subseteq V).$$

This local characterization of global minimality naturally suggests the following algorithm of steepest descent type [19]. Recall our assumption $r = 0$ in (TRF).

STEEPEST DESCENT ALGORITHM FOR AN L-CONVEX FUNCTION g.

S0: Find a vector $p \in \mathrm{dom}\, g$.
S1: Find $X \subseteq V$ that minimizes $g(p + \chi_X)$.
S2: If $g(p) \leq g(p + \chi_X)$, then stop (p is a minimizer of g).
S3: Set $p := p + \chi_X$ and go to S1.

Step S1 amounts to minimizing a set-function

$$\rho_p(X) = g(p + \chi_X) - g(p)$$

over all subsets X of V. As a consequence of (SBF) this function is submodular, i.e.,

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

and can be minimized in strongly polynomial time (see, e.g., [10, 11, 14, 23]). At the termination of the algorithm in step S2, p is a global optimum by Lemma 3.1. The function value g decreases monotonically with iterations. This property alone does not ensure finite termination in general, although it does if g is integer-valued and bounded from below.

We can guarantee an upper bound on the number of iterations by introducing a tie-breaking rule in step S1:

$$\text{Take the (unique) minimal minimizer } X \text{ of } \rho_p. \tag{3.1}$$

Let p° be the initial vector found in step S0. If g has a minimizer at all, it has, by (TRF), a minimizer p^* satisfying $p^\circ \le p^*$. Let p^* denote the smallest of such minimizers, which exists since $p^* \wedge q^* \in \arg\min g$ for $p^*, q^* \in \arg\min g$.

Lemma 3.2. *In step* S1*, $p \le p^*$ implies $p + \chi_X \le p^*$. Hence the number of iterations is bounded by $||p^\circ - p^*||_1$.*

Proof. Put $Y = \{v \in V \mid p(v) = p^*(v)\}$ and $p' = p + \chi_X$. By submodularity we have

$$g(p^*) + g(p') \ge g(p^* \vee p') + g(p^* \wedge p'),$$

whereas $g(p^*) \le g(p^* \vee p')$ since p^* is a minimizer of g. Hence $g(p') \ge g(p^* \wedge p')$. Here we have $p' = p + \chi_X$ and $p^* \wedge p' = p + \chi_{X \setminus Y}$, whereas X is the minimal minimizer by the tie-breaking rule (3.1). This means that $X \setminus Y = X$, i.e., $X \cap Y = \emptyset$. Therefore, $p' = p + \chi_X \le p^*$. □

It is easy to find the minimal minimizer of ρ_p using the existing algorithms for submodular set-function minimization. For example, with Schrijver's algorithm [23] we can find the minimal minimizer with $\mathrm{O}(n^8)$ function evaluations and $\mathrm{O}(n^9)$ arithmetic operations. Assuming that the minimal minimizer of a submodular set-function can be computed with $\mathrm{O}(\sigma(n))$ function evaluations and $\mathrm{O}(\tau(n))$ arithmetic operations, and denoting by F_g an upper bound on the time to evaluate g, we can perform step S1 in $\mathrm{O}(\sigma(n)F_g + \tau(n))$ time, where $(\sigma(n), \tau(n)) = (n^8, n^9)$ is a valid choice. We measure the size of the effective domain of g by

$$\hat{K}_g = \max\{||p - q||_1 \mid p, q \in \mathrm{dom}\, g,\ p(v) = q(v) \text{ for some } v \in V\}, \tag{3.2}$$

where it is noted that $\mathrm{dom}\, g$ itself is unbounded by (TRF).

Theorem 3.3. *For an L-convex function g with finite $\hat{K}_g$, the number of iterations in the steepest descent algorithm with tie-breaking rule* (3.1) *is bounded by $\hat{K}_g$. Hence, if a vector in* $\mathrm{dom}\, g$ *is given, the algorithm finds a minimizer of g in $\mathrm{O}((\sigma(n)F_g + \tau(n))\hat{K}_g)$ time.*

Proof. We have $||p^\circ - p^*||_1 \le \hat{K}_g$ since $p^\circ(v) = p^*(v)$ for some $v \in V$. Then the claim follows from Lemma 3.2. □

A function $g : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ is called $\mathrm{L}^\natural$-convex if the function

$$\tilde{g}(p_0, p) = g(p - p_0\mathbf{1}) \qquad (p_0 \in \mathbf{Z}, p \in \mathbf{Z}^V) \tag{3.3}$$

is an L-convex function in $n+1$ variables. Whereas $\mathrm{L}^\natural$-convex functions are conceptually equivalent to L-convex functions by the relation (3.3), the class of $\mathrm{L}^\natural$-convex functions in n variables is strictly larger than that of L-convex functions in n variables. The steepest descent algorithm for L-convex functions can be adapted to $\mathrm{L}^\natural$-convex function minimization.

Steepest descent algorithm for an $\mathrm{L}^\natural$-convex function g.

S0: Find a vector $p \in \mathrm{dom}\, g$.
S1: Find $\varepsilon \in \{1, -1\}$ and $X \subseteq V$ that minimize $g(p + \varepsilon\chi_X)$.
S2: If $g(p) \le g(p + \varepsilon\chi_X)$, then stop ($p$ is a minimizer of g).
S3: Set $p := p + \varepsilon\chi_X$ and go to S1.

Step S1 amounts to minimizing a pair of submodular set functions

$$\rho_p^+(X) = g(p+\chi_X) - g(p), \qquad \rho_p^-(X) = g(p-\chi_X) - g(p).$$

Let X^+ be the minimal minimizer of ρ_p^+, and let X^- be the maximal minimizer of ρ_p^-. The tie-breaking rule for step S1 reads

$$(\varepsilon, X) = \begin{cases} (1, X^+) & \text{if } \min \rho_p^+ \leq \min \rho_p^-, \\ (-1, X^-) & \text{if } \min \rho_p^+ > \min \rho_p^-. \end{cases} \tag{3.4}$$

This is a translation of the tie-breaking rule (3.1) for $\tilde{g}$ in (3.3) through the correspondence

$$\begin{array}{ccc} g & & \tilde{g} \\ \hline p \to p + \chi_X & \Longleftrightarrow & \tilde{p} \to \tilde{p} + (0, \chi_X) \\ p \to p - \chi_X & \Longleftrightarrow & \tilde{p} \to \tilde{p} + (1, \chi_{V\setminus X}) \\ \hline \end{array},$$

where $\tilde{p} = (0, p) \in \mathbf{Z}^{1+n}$. Since $(1, \chi_{V\setminus X^-})$ cannot be minimal in the presence of $(0, \chi_{X^+})$, we choose $(1, X^+)$ in the case of $\min \rho_p^+ = \min \rho_p^-$.

In view of the complexity bound given in Theorem 3.3 we note that the size $\hat{K}_{\tilde{g}}$ of the effective domain of the associated L-convex function $\tilde{g}$ is bounded in terms of the size of $\mathrm{dom}\, g$. The ℓ_1-size and ℓ_∞-size of $\mathrm{dom}\, g$ are denoted, respectively, by K_g in (1.2) and

$$K_g^\infty = \max\{||p-q||_\infty \mid p, q \in \mathrm{dom}\, g\}.$$

LEMMA 3.4. $\hat{K}_{\tilde{g}} \leq K_g + nK_g^\infty \leq \min[\ (n+1)K_g,\ 2nK_g^\infty\ ]$.

Proof. Take $\tilde{p} = (p_0, p)$ and $\tilde{q} = (q_0, q)$ in $\mathrm{dom}\, \tilde{g}$ such that $\hat{K}_{\tilde{g}} = |p_0 - q_0| + ||p-q||_1$ and either (i) $p_0 = q_0$ or (ii) $p(v) = q(v)$ for some $v \in V$. We may assume $p_0 \geq q_0$ and $p \geq q$ since $\tilde{p} \vee \tilde{q}, \tilde{p} \wedge \tilde{q} \in \mathrm{dom}\, \tilde{g}$ and $||(\tilde{p} \vee \tilde{q}) - (\tilde{p} \wedge \tilde{q})||_1 = ||\tilde{p} - \tilde{q}||_1$. The vectors $p' = p - p_0\mathbf{1}$ and $q' = q - q_0\mathbf{1}$ belong to $\mathrm{dom}\, g$. In case (i), we have $\hat{K}_{\tilde{g}} = ||p-q||_1 = ||p'-q'||_1 \leq K_g$. In case (ii), we have $p_0 - q_0 = q'(v) - p'(v)$ and

$$\begin{aligned} \hat{K}_{\tilde{g}} &= |p_0 - q_0| + ||p-q||_1 \\ &= (p_0 - q_0) + \sum_{u \in V} (p(u) - q(u)) \\ &= (p_0 - q_0) + \sum_{u \in V} (p'(u) - q'(u)) + n(p_0 - q_0) \\ &= \sum_{u \neq v} (p'(u) - q'(u)) - n(p'(v) - q'(v)) \\ &\leq K_g + nK_g^\infty. \end{aligned}$$

Note finally that $K_g \leq nK_g^\infty$ and $K_g^\infty \leq K_g$. ∎

4. Discussion.

4.1. Scaling algorithm. Scaling is one of the common techniques in designing efficient algorithms. This is also the case with L- or M-convex function minimization. We deal with L-convex function minimization to demonstrate an implication of our result stated in Theorem 3.3.

A scaling algorithm to minimize an L-convex function g finds a minimizer of the scaled function $g_\alpha(q) = g(p^\circ + \alpha q)$ for $\alpha = \alpha^\circ, \alpha^\circ/2, \alpha^\circ/4, \alpha^\circ/8, \ldots$, starting with a sufficiently large α° (a power of 2) until reaching $\alpha = 1$, where p° is an initial solution. For each α, g_α is an L-convex function, which can be minimized, e.g., by the steepest descent algorithm. The scaling algorithm reads as follows, where

$$\hat{K}_g^\infty = \max\{||p-q||_\infty \mid p, q \in \mathrm{dom}\, g,\ p(v) = q(v) \text{ for some } v \in V\}$$

and $r = 0$ in (TRF).

Scaling algorithm for an L-convex function g.

S0: Find a vector $p \in \mathrm{dom}\, g$, and set $\alpha := 2^{\lceil \log_2(\hat{K}_g^\infty/2n)\rceil}$.

S1: Find an integer vector q that minimizes $g(p+\alpha q)$ and set $p := p + \alpha q$.

S2: If $\alpha = 1$, then stop (p is a minimizer of g).

S3: Set $\alpha := \alpha/2$ and go to S1.

The success of this scaling approach hinges on the efficiency of the minimization in step S1. By a proximity theorem due to [12] (see Proposition 8.9 in [20] or Theorem 7.18 in [21]) there exists a minimizer q of $g(p+\alpha q)$ such that $\mathbf{0} \le q \le (n-1)\mathbf{1}$. Our complexity bound (Lemma 3.2 or Theorem 3.3) guarantees that the steepest descent algorithm with tie-breaking rule (3.1) finds the minimizer in step S1 in $\mathrm{O}((\sigma(n)F_g + \tau(n))n^2)$ time. The number of executions of step S1 is bounded by $\lceil \log_2(\hat{K}_g^\infty/2n)\rceil$, and at the termination of the algorithm in step S2 with $\alpha = 1$, p is a minimizer of g by Lemma 3.1. Thus the result of the present paper guarantees the efficiency of the scaling approach based on steepest descent algorithm.

It is in order here to compare our algorithm with the scaling algorithm of [9], which is described in [20]. In [9] step S1 above is performed via submodular set-function minimization over a ring family on a ground set of cardinality $\le n^2$. This is based on the general fact (Birkhoff's representation theorem) that any distributive lattice can be represented as a boolean lattice over a ground set, and the size of the ground set is equal to the length of a maximal chain of the distributive lattice. Thus the minimization of the scaled function in step S1 can be carried out with $\mathrm{O}(\sigma(n^2))$ evaluations of g. Although the complexity of this algorithm for step S1 is bounded by a polynomial in n, the algorithm is not easy to implement and will be slow in practice. Our steepest descent algorithm above is much simpler, both conceptually and algorithmically, and will be faster in practice, performing the minimization of the scaled function in step S1 with $\mathrm{O}(\sigma(n)n^2)$ evaluations of g. Note that $\sigma(n)n^2$ is smaller in order than $\sigma(n^2)$ if $\sigma(n) = n^s$ with $s > 2$.

As for M-convex function minimization, a similar scaling approach works, provided that the scaled function $f_\alpha(y) = f(x + \alpha y)$ remains M-convex for any α and x, although this is not always the case; see [16]. See [25] and [26] for more sophisticated scaling algorithms for M-convex function minimization.

4.2. Integrally convex functions. Global optimality is characterized by local optimality also for integrally convex functions, of which M-convex and L-convex functions are special cases. Namely, it is known [4] that, for an integrally convex function f, a point x in $\mathrm{dom}\, f$ is a global minimizer of f if and only if $f(x) \le f(x+\chi_Y-\chi_Z)$ for all disjoint subsets $Y, Z \subseteq V$. This fact would naturally suggest the following generic scheme of steepest descent algorithms for minimizing an integrally convex function.

Steepest descent scheme for an integrally convex function f.

S0: Find a vector $x \in \mathrm{dom}\, f$.

S1: Find disjoint $Y, Z \subseteq V$ that minimize $f(x - \chi_Y + \chi_Z)$.

S2: If $f(x) \le f(x - \chi_Y + \chi_Z)$, then stop ($x$ is a minimizer of f).

S3: Set $x := x - \chi_Y + \chi_Z$ and go to S1.

The steepest descent algorithms for M-convex and L-convex functions in sections 2 and 3 both fit in this generic form. It is emphasized, however, that for a general integrally convex function no efficient algorithm for step S1 is available, whereas we do have polynomial time algorithms for M-convex and L-convex functions.

Acknowledgments. The author thanks Satoru Fujishige, Shiro Matuura, and Akihisa Tamura for helpful comments, and the anonymous referees for pointing out a flaw in the earlier version of the tie-breaking rule for M-convex function minimization.

REFERENCES

[1] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver, *Combinatorial Optimization*, John Wiley and Sons, New York, 1998.

[2] A. W. M. Dress and W. Wenzel, *Valuated matroid: A new look at the greedy algorithm*, Appl. Math. Lett., 3 (1990), pp. 33–35.

[3] J. Edmonds, *Submodular functions, matroids and certain polyhedra*, in Combinatorial Structures and Their Applications, R. Guy, H. Hanani, N. Sauer, and J. Schönheim, eds., Gordon and Breach, New York, 1970, pp. 69–87.

[4] P. Favati and F. Tardella, *Convexity in nonlinear integer programming*, Ricerca Operativa, 53 (1990), pp. 3–44.

[5] A. Frank, *An algorithm for submodular functions on graphs*, in Bonn Workshop on Combinatorial Optimization, Ann. Discrete Math. 16, North-Holland, Amsterdam, New York, 1982, pp. 97–120.

[6] S. Fujishige, *Theory of submodular programs: A Fenchel-type min-max theorem and subgradients of submodular functions*, Math. Program., 29 (1984), pp. 142–155.

[7] S. Fujishige, *Submodular Functions and Optimization*, Ann. Discrete Math. 47, North-Holland, Amsterdam, 1991.

[8] S. Fujishige and K. Murota, *Notes on L-/M-convex functions and the separation theorems*, Math. Program., 88 (2000), pp. 129–146.

[9] S. Iwata, Oral presentation at Workshop on Matroids, Matching, and Extensions, University of Waterloo, ON, Canada, 1999.

[10] S. Iwata, *A faster scaling algorithm for minimizing submodular functions*, in Integer Programming and Combinatorial Optimization, W. J. Cook and A. S. Schulz, eds., Lecture Notes in Comput. Sci. 2337, Springer-Verlag, New York, 2002, pp. 1–8.

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

[12] S. Iwata and M. Shigeno, *Conjugate scaling algorithm for Fenchel-type duality in discrete convex optimization*, SIAM J. Optim., 13 (2002), pp. 204–211.

[13] L. Lovász, *Submodular functions and convexity*, in Mathematical Programming—The State of the Art, A. Bachem, M. Grötschel, and B. Korte, eds., Springer-Verlag, Berlin, 1983, pp. 235–257.

[14] S. T. McCormick, *Submodular function minimization*, in Handbook on Discrete Optimization, K. Aardal, G. Nemhauser, and R. Weismantel, eds., Elsevier Science, Amsterdam, to appear.

[15] B. L. Miller, *On minimizing nonseparable functions defined on the integers with an inventory application*, SIAM J. Appl. Math., 21 (1971), pp. 166–185.

[16] S. Moriguchi, K. Murota, and A. Shioura, *Scaling algorithms for M-convex function minimization*, IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences, E85-A (2002), pp. 922–929.

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

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

[19] K. Murota, *Algorithms in discrete convex analysis*, IEICE Trans. Systems and Information, E83-D (2000), pp. 344–352.

[20] K. Murota, *Discrete Convex Analysis—An Introduction*, Kyoritsu, Tokyo, 2001 (in Japanese).

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

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

[23] A. Schrijver, *A combinatorial algorithm minimizing submodular functions in strongly polynomial time*, J. Combin. Theory Ser. B, 80 (2000), pp. 346–355.

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

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

[26] A. Tamura, *Coordinatewise domain scaling algorithm for M-convex function minimization*, in Integer Programming and Combinatorial Optimization, W. J. Cook and A. S. Schulz, eds., Lecture Notes in Comput. Sci. 2337, Springer-Verlag, New York, 2002, pp. 21–35.