EGERVÁRY RESEARCH GROUP
ON COMBINATORIAL OPTIMIZATION

TECHNICAL REPORTS

TR-2019-08. Published by the Egerváry Research Group, Pázmány P. sétány 1/C, H–1117, Budapest, Hungary. Web site: www.cs.elte.hu/egres . ISSN 1587–4451.

Discrete Decreasing Minimization, Part II: Views from Discrete Convex Analysis

András Frank and Kazuo Murota

August 2018 (Version July 15, 2019)

Discrete Decreasing Minimization, Part II: Views from Discrete Convex Analysis

András Frank⋆ and Kazuo Murota⋆⋆

Abstract

We continue to consider the discrete decreasing minimization problem on an integral base-polyhedron treated in Part I. The problem is to find a lexicographically minimal integral vector in an integral base-polyhedron, where the components of a vector are arranged in a decreasing order. This study can be regarded as a discrete counter-part of the work by Fujishige (1980) on the lexicographically optimal base and the principal partition of a base-polyhedron in continuous variables. The objective of Part II is two-fold. The first is to offer structural views from discrete convex analysis (DCA) on the results of Part I obtained by the constructive and algorithmic approach. The second objective is to pave the way of DCA approach to discrete decreasing minimization on other discrete structures (the intersection of M-convex sets, flows, submodular flows) that we consider in Parts III and IV.

We derive the structural results in Part I from fundamental facts on M-convex sets and M-convex functions in DCA. The characterization of decreasing minimality in terms of 1-tightening steps (exchange operations) is derived from the local condition of global minimality for M-convex functions, known as M-optimality criterion in DCA. The min-max formulas, including the one for the square-sum of components, are derived as special cases of the Fenchel-type discrete duality in DCA. A general result on the Fenchel-type discrete duality in DCA offers a short alternative proof to the statement that the decreasingly minimal elements of an M-convex set form a matroidal M-convex set.

A direct characterization is given to the canonical partition, which was constructed by an iterative procedure in Part I. This reveals the precise relationship between the canonical partition for the discrete case and the principal partition for the continuous case. Moreover, this result entails a proximity theorem, stating that every decreasingly minimal element is contained in the small box containing the (unique) fractional decreasingly minimal element (the minimum-norm point), leading further to a continuous relaxation algorithm for finding a decreasingly minimal element of an M-convex

⋆MTA-ELTE Egerváry Research Group, Department of Operations Research, Eötvös University, Pázmány P. s. 1/c, Budapest, Hungary, H-1117. e-mail: `frank@cs.elte.hu`. The research was partially supported by the National Research, Development and Innovation Fund of Hungary (FK_18) – No. NKFI-128673.

⋆⋆Department of Economics and Business Administration, Tokyo Metropolitan University, Tokyo 192-0397, Japan, e-mail: `murota@tmu.ac.jp`. The research was supported by CREST, JST, Grant Number JPMJCR14D2, Japan, and JSPS KAKENHI Grant Number 26280004.

set. Thus the relationship between the continuous and discrete cases is completely clarified.

Furthermore, we present DCA min-max formulas to be needed in Parts III and IV, where the discrete decreasing minimization problem is considered for network flows, the intersection of two M-convex sets, and submodular flows.

Keywords: base-polyhedra, discrete convex analysis, Fenchel-type min-max formula, lexicographically optimal, majorization, principal partition.

Contents

1 Introduction

We continue to consider discrete decreasing minimization on an integral base-polyhedron studied in Part I. The problem is to find a lexicographically minimal (dec-min) integral vector in an integral base-polyhedron, where the components of a vector are arranged in a decreasing order (see Section 1.1 for precise description of the problem). While our present study deals with the discrete case, the continuous case was investigated by Fujishige [13] around 1980 under the name of lexicographically optimal bases of a base-polyhedron, as a generalization of lexicographically optimal maximal flows considered by Megiddo [29]. Our study can be regarded as a discrete counter-part of the work by Fujishige [13], [14, Section 9] on the lexicographically optimal base and the principal partition of a base-polyhedron. The objective of Part II is two-fold. The first is to offer structural views from discrete convex analysis (DCA) on the results of Part I obtained by the constructive and algorithmic approach. The second objective is to pave the way of DCA approach to discrete decreasing minimization on other discrete structures (the intersection of M-convex sets, flows, submodular flows) that we consider in Parts III and IV.

In Part I of this paper, we have shown the following:

- A characterization of decreasing minimality by 1-tightening steps (exchange operations),
- A (dual) characterization of decreasing minimality by the canonical chain,
- The structure of the dec-min elements as a matroidal M-convex set,
- A characterization of a dec-min element as a minimizer of square-sum of components,
- A min-max formula for the square-sum of components,
- A strongly polynomial algorithm for finding a dec-min element and the canonical chain,
- Applications.

In contrast to the constructive and algorithmic approach in Part I, Part II offers structural views from discrete convex analysis (DCA) as well as from majorization. The concept of majorization ordering offers a useful general framework to discuss decreasing minimality. The relevance of DCA to decreasing minimization is not surprising, since an M-convex set is nothing but the set of integral points of an integral base-polyhedron and a separable convex function on an M-convex set is an M-convex function. In particular, the square-sum of components of a vector in an M-convex set is an M-convex function. It will be shown that most of the important structural results obtained in Part I can be derived from the Fenchel-type discrete duality theorem, which is a main characteristic of DCA as compared with other theories of discrete functions such as [37].

In Section 2 of this paper the basic facts about majorization are described. In Section 3 we derive the characterization of decreasing minimality in terms of 1-tightening steps (exchange operations) from the local characterization of global minimality for M-convex functions, known as M-optimality criterion in DCA. In Section 4, the min-max formulas, including the one for the square-sum of components, are derived as special cases of the Fenchel-type discrete duality in DCA. We also show a novel min-max formula, which reinforces the link between the present study and the theory of majorization. In Section 5 we use a general result on the Fenchel-type discrete duality in DCA for a short alternative proof to the statement that the decreasingly minimal elements of an M-convex set form a matroidal M-convex set. The relationship between the continuous and discrete cases is clarified in Section 6. We reveal the precise relation between the canonical partition and the principal partition by establishing an alternative direct characterization of the canonical partition, which was constructed by an iterative procedure in Part I. The obtained result provides a proximity theorem, stating that every dec-min element is contained in the small box containing the (unique) fractional dec-min element (the minimum-norm point), and hence a continuous relaxation algorithm for finding a decreasingly minimal element of an M-convex set. In Section 7 we present DCA results relevant to Parts III and IV, where discrete decreasing minimization is considered for the set of integral feasible flows, the intersection of two M-convex sets, and the set of integral members of an integral submodular flow polyhedron. In Appendix A we offer a brief survey of early papers and books related to decreasing minimization on base-polyhedra.

1.1 Definition and notation

We review some definitions and notations introduced in Part I [9].

Decreasing minimality

For a vector x, let $x{\downarrow}$ denote the vector obtained from x by rearranging its components in a decreasing order. For example, $x{\downarrow} = (5,5,4,2,1)$ when $x = (2,5,5,1,4)$. We call two vectors x and y (of same dimension) **value-equivalent** if $x{\downarrow} = y{\downarrow}$. For example, $(2,5,5,1,4)$ and $(1,4,5,2,5)$ are value-equivalent while the vectors $(3,5,5,3,4)$ and $(3,4,5,4,4)$ are not.

A vector x is **decreasingly smaller** than vector y, in notation $x <_{\rm dec} y$, if $x{\downarrow}$ is lexicographically smaller than $y{\downarrow}$ in the sense that they are not value-equivalent and $x{\downarrow}(j) < y{\downarrow}(j)$ for the smallest subscript j for which $x{\downarrow}(j)$ and $y{\downarrow}(j)$ differ. For example, $x = (2,5,5,1,4)$ is decreasingly smaller than $y = (1,5,5,5,1)$ since $x{\downarrow} = (5,5,4,2,1)$ is lexicographically smaller than $y{\downarrow} = (5,5,5,1,1)$. We write $x \leq_{\rm dec} y$ to mean that x is decreasingly smaller than or value-equivalent to y.

For a set Q of vectors, $x \in Q$ is **decreasingly minimal** (**dec-min**, for short) if $x \leq_{\rm dec} y$ for every $y \in Q$. Note that the dec-min elements of Q are value-equivalent. An element m of Q is dec-min if its largest component is as small as possible, within this, its second largest component (with the same or smaller value than the largest one) is as small as possible, and so on. An element x of Q is said to be a **max-minimized** element (a **max-minimizer**, for short) if its largest component is as small as possible.

In an analogous way, for a vector x, we let $x{\uparrow}$ denote the vector obtained from x by rearranging its components in an increasing order. A vector y is **increasingly larger** than vector x, in notation $y >_{\rm inc} x$, if they are not value-equivalent and $y{\uparrow}(j) > x{\uparrow}(j)$ holds for the smallest subscript j for which $y{\uparrow}(j)$ and $x{\uparrow}(j)$ differ. We write $y \geq_{\rm inc} x$ if either $y >_{\rm inc} x$ or x

and y are value-equivalent. Furthermore, we call an element m of Q **increasingly maximal** (**inc-max** for short) if its smallest component is as large as possible over the elements of Q, within this its second smallest component is as large as possible, and so on.

The **decreasing minimization problem** is to find a dec-min element of a given set Q of vectors. When the set Q consists of integral vectors, we speak of discrete decreasing minimization. In Parts I and II of this series of papers, we deal with the case where the set Q is an M-convex set, i.e., the set of integral members of an integral base-polyhedron. In Part III, the set Q will be the integral feasible flows. In Part IV, the set Q will be the intersection of two M-convex sets, or more generally, the set of integral members of an integral submodular flow polyhedron.

Base polyhedra

Throughout the paper, S denotes a finite non-empty ground-set. For a vector $m \in \mathbf{R}^S$ (or function $m : S \to \mathbf{R}$) and a subset $X \subseteq S$, we use the notation $\widetilde{m}(X) = \sum[m(v) : v \in X]$. The characteristic (or incidence) vector of a subset $Z \subseteq S$ is denoted by χ_Z, that is, $\chi_Z(v) = 1$ if $v \in Z$ and $\chi_Z(v) = 0$ otherwise. For a polyhedron B, notation $\dddot{B}$ (pronounced: dotted B) means the set of integral members (elements, vectors, points) of B.

Let b be a set-function for which $b(\emptyset) = 0$ and $b(X) = +\infty$ is allowed but $b(X) = -\infty$ is not. The submodular inequality for subsets $X, Y \subseteq S$ is defined by

$$b(X) + b(Y) \geq b(X \cap Y) + b(X \cup Y). \tag{1.1}$$

We say that b is submodular if the submodular inequality holds for every pair of subsets $X, Y \subseteq S$ with finite b-values. A set-function p is supermodular if $-p$ is submodular. A (possibly unbounded) **base-polyhedron** B in $\mathbf{R}^S$ is defined by

$$B = B(b) = \{x \in \mathbf{R}^S : \widetilde{x}(S) = b(S),\ \widetilde{x}(Z) \leq b(Z) \text{ for every } Z \subset S\}. \tag{1.2}$$

A non-empty base-polyhedron B can also be defined by a supermodular function p for which $p(\emptyset) = 0$ and $p(S)$ is finite as follows:

$$B = B'(p) = \{x \in \mathbf{R}^S : \widetilde{x}(S) = p(S),\ \widetilde{x}(Z) \geq p(Z) \text{ for every } Z \subset S\}. \tag{1.3}$$

We call the set $\dddot{B}$ of integral elements of an integral base-polyhedron B an **M-convex set**. Originally, this basic notion of discrete convex analysis was defined as a set of integral points in $\mathbf{R}^S$ satisfying certain exchange axioms, and it has been known that the two properties are equivalent ([35, Theorem 4.15]).

Discrete convex functions

For a function $\varphi : \mathbf{Z} \to \mathbf{R} \cup \{-\infty, +\infty\}$ the **effective domain** of φ is denoted as $\mathrm{dom}\,\varphi = \{k \in \mathbf{Z} : -\infty < \varphi(k) < +\infty\}$. A function $\varphi : \mathbf{Z} \to \mathbf{R} \cup \{+\infty\}$ is called **discrete convex** (or simply **convex**) if

$$\varphi(k-1) + \varphi(k+1) \geq 2\varphi(k) \tag{1.4}$$

for all $k \in \mathrm{dom}\,\varphi$, and **strictly convex** if $\mathrm{dom}\,\varphi = \mathbf{Z}$ and $\varphi(k-1) + \varphi(k+1) > 2\varphi(k)$ for all $k \in \mathbf{Z}$.

A function $\Phi : \mathbf{Z}^S \to \mathbf{R} \cup \{+\infty\}$ of the form

$$\Phi(x) = \sum [\varphi_s(x(s)) : s \in S] \tag{1.5}$$

is called a **separable (discrete) convex function** if, for each $s \in S$, $\varphi_s : \mathbf{Z} \to \mathbf{R} \cup \{+\infty\}$ is a discrete convex function. We call Φ a **symmetric separable convex function** if φ_s does not depend on s, that is, if $\varphi_s = \varphi$ for all $s \in S$ for some discrete convex function φ. We call Φ a **symmetric separable strictly convex function** if φ is strictly convex.

2 Connection to majorization

Majorization ordering (or dominance ordering) is a well-established notion studied in diverse contexts including statistics and economics, as described in Arnold–Sarabia [4] and Marshall–Olkin–Arnold [28]. In this section we describe the relevant results known in the literature of majorization, and indicate a close relationship to decreasing minimality investigated in our series of papers.

We have dual objectives in this section. First, we intend to reinforce the connection between majorization and combinatorial optimization. It is also hoped that this will lead to future applications of our results in areas like statistics and economics, in addition to those areas related to graphs, networks, and matroids mentioned in the introduction of Part I [9]. In economics, for example, egalitarian allocation for indivisible goods can possibly be formulated and analyzed by means of discrete decreasing minimization.

Second, we point out substantial technical connections between majorization and our results in Part I. We argue that some of our results can be derived from the combination of the classical results about majorization and the results of Groenevelt [17] for the minimization of separable convex functions over the integer points in an integral base-polyhedron. We also point out that some of the standard characterizations for least majorization are associated with min-max duality relations in the case where the underlying set is the integer points of an integral base-polyhedron or the intersection of two integral base-polyhedra.

2.1 Majorization ordering

We review standard results known in the literature of majorization in a way suitable for our discussion.

Recall that $x{\downarrow}$ denotes the vector obtained from a vector $x \in \mathbf{R}^n$ by rearranging its components in a decreasing order. Let $\overline{x}$ denote the vector whose k-th component $\overline{x}(k)$ is equal to the sum of the first k components of $x{\downarrow}$. A vector x is said to be **majorized** by another vector y, in notation $x \prec y$, if $\overline{x} \leq \overline{y}$ and $\overline{x}(n) = \overline{y}(n)$. It is easy to see [28, p.13] that

$$x \prec y \iff -x \prec -y. \tag{2.1}$$

(At first glance, the equivalence in (2.1) may look strange, but observe that $x \prec y$ means that x is more uniform than y, which is equivalent to saying that $-x$ is more uniform than $-y$.)

Majorization is discussed more often for real vectors, but here we are primarily interested in integer vectors.

As an immediate adaptation of the standard results [28, 1.A.3 in p.14], the following proposition gives equivalent conditions for majorization for integer vectors. A T-transform (also called a Robin Hood operation) means a linear transformation of the form $T = (1-\lambda)I + \lambda Q$, where $0 \leq \lambda \leq 1$ and Q is a permutation matrix that interchanges just two elements (transposition). In other words, a T-transform is a mapping of the form $x \mapsto x + \hat{\lambda}(\chi_s - \chi_t)$ with $0 \leq \hat{\lambda} \leq x(t) - x(s)$. It is noteworthy that this operation with $\hat{\lambda} = 1$ corresponds to the basis exchange in an integral base-polyhedron.

Proposition 2.1. *The following conditions are equivalent for $x, y \in \mathbf{Z}^n$:*

(i) *$x \prec y$ (x is majorized by y), that is,*

$$\sum_{i=1}^{k} x{\downarrow}(i) \leq \sum_{i=1}^{k} y{\downarrow}(i) \quad (k = 1, \ldots, n-1), \qquad \sum_{i=1}^{n} x{\downarrow}(i) = \sum_{i=1}^{n} y{\downarrow}(i). \tag{2.2}$$

(ii) *$x = yP$ for some doubly stochastic matrix P, where x and y are regarded as row vectors.*

(iii) *x can be derived from y by successive applications of a finite number of T-transforms.*

(iv) $\sum_{i=1}^{n} \varphi(x(i)) \leq \sum_{i=1}^{n} \varphi(y(i))$ *for all discrete convex functions $\varphi : \mathbf{Z} \to \mathbf{R}$.*

(v) $\sum_{i=1}^{n} x(i) = \sum_{i=1}^{n} y(i)$ *and* $\sum_{i=1}^{n} (x(i) - a)^+ \leq \sum_{i=1}^{n} (y(i) - a)^+$ *for all $a \in \mathbf{Z}$. where $(z)^+ = \max\{0, z\}$ for any $z \in \mathbf{Z}$.* ■

Let D be an arbitrary subset of $\mathbf{Z}^n$. An element x of D is said to be **least majorized** in D if x is majorized by all $y \in D$. A least majorized element may not exist in general, as the following example shows.

Example 2.1. Let $D = \{(2,0,0,0),\ (1,-1,1,1)\}$. For $x = (2,0,0,0)$ and $y = (1,-1,1,1)$ we have $x{\downarrow} = (2,0,0,0)$ and $y{\downarrow} = (1,1,1,-1)$. Therefore, $x = (2,0,0,0)$ is increasingly maximal in D and $y = (1,-1,1,1)$ is decreasingly minimal in D. However, there exists no least majorized element in D, since $\overline{x} = (2,2,2,2)$ and $\overline{y} = (1,2,3,2)$, for which neither $\overline{x} \leq \overline{y}$ nor $\overline{y} \leq \overline{x}$ holds. We note that D here arises from the intersection of two integral base-polyhedra (see Section 3.4 of Part I [9]). ■

Remark 2.1. In discussing the existence and properties of a least majorized element, we are primarily concerned with a subset D of $\mathbf{Z}^n$ whose elements have a constant component-sum. If the component-sum is not constant on D, we need to introduce a more general notion [40]. A vector x is said to be **weakly submajorized** by another vector y, denoted $x \prec_{\mathrm{w}} y$, if $\overline{x} \leq \overline{y}$. An element x of D is said to be **least weakly submajorized** in D if x is weakly submajorized by all $y \in D$. The distinction of "weakly submajorized" and "majorized" is not necessary for a base-polyhedron or the intersection of base-polyhedra, whereas we have to distinguish these concepts for a g-polymatroid and a submodular flow polyhedron. ■

Remark 2.2. The characterization of a least majorized element in (iv) in Proposition 2.1 can be associated with a min-max duality relation, which is given by (4.22) in Section 4.3 when the underlying set D is an M-convex set (= the integer points of an integral base-polyhedron), and by (7.46) in Section 7.2 when D is the intersection of two M-convex sets. For an M-convex set, the min-max formula associated with (v) in Proposition 2.1 is given by (4.32) in Theorem 4.7 in Section 4.7. ■

2.2 Majorization and decreasing-minimality

Majorization and decreasing-minimality are closely related, as is explicit in Tamir [40].

Proposition 2.2. *If $x \prec y$, then $x \leq_{\rm dec} y$ and $x \geq_{\rm inc} y$.*

Proof. Suppose that $x \prec y$. If $\overline{x} = \overline{y}$, then $x{\downarrow} = y{\downarrow}$, and hence x and y are value-equivalent. If $\overline{x} < \overline{y}$, then there exists an index k with $1 \leq k \leq n$ such that $x{\downarrow}(i) = y{\downarrow}(i)$ for $i = 1, \ldots, k-1$ and $x{\downarrow}(k) < y{\downarrow}(k)$. This shows that x is decreasingly smaller than y. In either case, we have $x \leq_{\rm dec} y$. Since $x \prec y$, we have $-x \prec -y$ by (2.1). By the above argument applied to $(-x, -y)$, we obtain $-x \leq_{\rm dec} -y$, which is equivalent to $x \geq_{\rm inc} y$. ■

Remark 2.3. The converse of Proposition 2.2 is not true. That is, $x \prec y$ does not follow from $x \leq_{\rm dec} y$ and $x \geq_{\rm inc} y$. For instance, for $x = (2, 2, -2, -2)$ and $y = (3, 0, 0, -3)$ we have $x \leq_{\rm dec} y$ and $x \geq_{\rm inc} y$, but $x \not\prec y$ since $\overline{x} = (2, 4, 2, 0)$ and $\overline{y} = (3, 3, 3, 0)$. ■

Proposition 2.3. *Let D be an arbitrary subset of $\mathbf{Z}^n$ and assume that D admits a least majorized element. For any $x \in D$ the following three conditions are equivalent.*

(A) *x is least majorized in D.*

(B) *x is decreasingly minimal in D.*

(C) *x is increasingly maximal in D.*

Proof. (A)→(B) By Proposition 2.2, a least majorized element is decreasingly minimal.

(B)→(A) Take a least majorized element y, which exists by the assumption. By definition we have $\overline{y} \leq \overline{x}$. Since $x \leq_{\rm dec} y$, we have either $x{\downarrow} = y{\downarrow}$ or there exists an index k with $1 \leq k \leq n$ such that $x{\downarrow}(i) = y{\downarrow}(i)$ for $i = 1, \ldots, k-1$ and $x{\downarrow}(k) < y{\downarrow}(k)$. In the latter case we have $\overline{x}(k) < \overline{y}(k)$, which contradicts $\overline{y} \leq \overline{x}$. Therefore we have $x{\downarrow} = y{\downarrow}$, which implies that x is a least majorized element.

(A)↔(C) For any $y \in D$, we have

$$x \prec y \iff -x \prec -y \iff -x \leq_{\rm dec} -y \iff x \geq_{\rm inc} y$$

by (2.1) and (A)↔(B) for $(-x, -y)$. ■

2.3 Majorization in integral base-polyhedra

In this section we consider majorization ordering for integer points in an integral base-polyhedron. In discrete convex analysis, the set of the integer points of an integral base-polyhedron is called an M-convex set.

The following fundamental fact has long been recognized by experts, though it was difficult for the present authors to identify its origin in the literature (see Remark 2.5).

Theorem 2.4. *The set of the integer points of an integral base-polyhedron admits a least majorized element.* ■

This fact can be regarded as a corollary of the following fundamental result of Groenevelt [17], which is already mentioned in Section 6 of Part I [9].

Proposition 2.5 (Groenevelt [17]; cf. [14, Theorem 8.1])**.** *Let B be an integral base-polyhedron, $\overset{....}{B}$ be the set of its integral elements, and $\Phi(x) = \sum[\varphi_s(x(s)) : s \in S]$ for $x \in \mathbf{Z}^S$, where $\varphi_s : \mathbf{Z} \to \mathbf{R} \cup \{+\infty\}$ is a discrete convex function for each $s \in S$. An element m of $\overset{....}{B}$ is a minimizer of $\Phi(x)$ if and only if $\varphi_s(m(s)+1) + \varphi_t(m(t)-1) \geq \varphi_s(m(s)) + \varphi_t(m(t))$ whenever $m + \chi_s - \chi_t \in \overset{....}{B}$.* ■

Theorem 2.4 can be derived from the combination of Proposition 2.5 with Proposition 2.1. Let $m \in \overset{....}{B}$ be a minimizer of the square-sum $\sum[x(s)^2 : s \in S]$ over $\overset{....}{B}$; note that such m exists. Then, by Proposition 2.5 (only-if part), we have $(m(s)+1)^2 + (m(t)-1)^2 \geq m(s)^2 + m(t)^2$ whenever $m+\chi_s-\chi_t \in \overset{....}{B}$. Here the inequality $(m(s)+1)^2+(m(t)-1)^2 \geq m(s)^2+m(t)^2$ is equivalent to $m(s)-m(t)+1 \geq 0$, which implies $\varphi(m(s)+1)+\varphi(m(t)-1) \geq \varphi(m(s))+\varphi(m(t))$ for any discrete convex function $\varphi : \mathbf{Z} \to \mathbf{R}$. Therefore, by Proposition 2.5 (if part), m is a minimizer of any symmetric separable convex function $\sum[\varphi(x(s)) : s \in S]$ over $\overset{....}{B}$. By the equivalence of (i) and (iv) in Proposition 2.1, this element m is a least majorized element of $\overset{....}{B}$.

The combination of Theorem 2.4 and Proposition 2.3 implies the following.

Theorem 2.6. *Let B be an integral base-polyhedron and $\overset{....}{B}$ be the set of its integral elements. An element m of $\overset{....}{B}$ is decreasingly minimal if and only if m is least majorized in $\overset{....}{B}$.* ■

Remark 2.4. In Theorem 3.5 of Part I [9] we have shown that a dec-min element of $\overset{....}{B}$ has the property (2.2), which is referred to as "min k-largest-sum" in [9]. This implies that any dec-min element of $\overset{....}{B}$ is a least majorized element of $\overset{....}{B}$. Since a dec-min element always exists, this theorem also implies the existence of a least majorized element in $\overset{....}{B}$. ■

Remark 2.5. A variant of majorization concept, "weak submajorization" (cf., Remark 2.1), is investigated for integral g-polymatroids by Tamir [40] and for jump systems by Ando [2]. These results are a direct extension of Theorem 2.4. Therefore, we may safely say that Theorem 2.4 with the above proof was known to experts before 1995. ■

3 Convex minimization and decreasing minimality

In this section we shed the light of discrete convex analysis on the following results obtained in Part I [9]. More specifically, we derive these results from the optimality criterion for M-convex functions, which is described in Section 3.2.

Theorem 3.1 ([9, Theorem 3.3, (A) & (C1)])**.** *An element m of $\overset{....}{B}$ is a dec-min element of $\overset{....}{B}$ if and only if there is no 1-tightening step for m.* ■

Theorem 3.2 ([9, Corollary 6.3]). *Let $\Phi(x) = \sum[\varphi(x(s)) : s \in S]$ be a symmetric separable convex function with $\varphi : \mathbf{Z} \to \mathbf{R}$. An element m of $\ddot{B}$ is a minimizer of Φ if m is a dec-min element of $\ddot{B}$, and the converse is also true if, in addition, Φ is strictly convex.* ∎

It should be clear in the above that $\ddot{B}$ denotes an M-convex set (the set of integral points of an integral base-polyhedron), and a **1-tightening step** for $m \in \ddot{B}$ means the operation of replacing m to $m + \chi_s - \chi_t$ for some $s, t \in S$ such that $m(t) \geq m(s) + 2$ and $m + \chi_s - \chi_t \in \ddot{B}$.

3.1 Convex formulation of decreasing minimality

A dec-min element can be characterized as a minimizer of 'rapidly increasing' convex function. This characterization enables us to make use of discrete convex analysis in investigating decreasing minimality.

We say that a positive-valued function $\varphi : \mathbf{Z} \to \mathbf{R}$ is N-increasing, where $N > 0$, if

$$\varphi(k+1) \geq N\,\varphi(k) > 0 \qquad (k \in \mathbf{Z}). \tag{3.1}$$

With the choice of a sufficiently large N, this concept formulates the intuitive notion that φ is "rapidly increasing." An N-increasing function φ with $N \geq 2$ is strictly convex, since $\varphi(k-1) + \varphi(k+1) > \varphi(k+1) \geq N\varphi(k) \geq 2\varphi(k)$.

As is easily expected, $x <_{\mathrm{dec}} y$ is equivalent to $\Phi(x) < \Phi(y)$ defined by such φ, as follows.

Proposition 3.3. *Assume $|S| \geq 2$ and that φ is $|S|$-increasing. A vector $x \in \mathbf{Z}^S$ is decreasingly-smaller than a vector $y \in \mathbf{Z}^S$ if and only if $\Phi(x) < \Phi(y)$.*

Proof. For $x \in \mathbf{Z}^S$ and $k \in \mathbf{Z}$, let $\Theta(x,k)$ denote the number of elements s of S with $x(s) = k$, i.e., $\Theta(x,k) = |\{s \in S : x(s) = k\}|$. Then we have

$$\Phi(x) = \sum_k \Theta(x,k)\varphi(k). \tag{3.2}$$

Obviously, $\Phi(x) = \Phi(y)$ if x and y are value-equivalent. Suppose that x is not value-equivalent to y, and let $\hat{k}$ be the largest k with $\Theta(x,k) \neq \Theta(y,k)$. By definition, x is decreasingly-smaller than y if and only if $\Theta(x,\hat{k}) < \Theta(y,\hat{k})$.

We show that $\Theta(x,\hat{k}) < \Theta(y,\hat{k})$ implies $\Phi(x) < \Phi(y)$. Then the converse also follows from this (by exchaning the roles of x and y). Let $T := \sum_{k>\hat{k}} \Theta(x,k)\varphi(k) = \sum_{k>\hat{k}} \Theta(y,k)\varphi(k)$. It follows from

$$\begin{aligned}
\Phi(x) &= T + \Theta(x,\hat{k})\varphi(\hat{k}) + \sum_{k<\hat{k}} \Theta(x,k)\varphi(k) \\
&\leq T + \Theta(x,\hat{k})\varphi(\hat{k}) + \varphi(\hat{k}-1) \sum_{k<\hat{k}} \Theta(x,k) \\
&\leq T + \Theta(x,\hat{k})\varphi(\hat{k}) + \varphi(\hat{k}) \frac{1}{|S|} \sum_{k<\hat{k}} \Theta(x,k) \\
&\leq T + (\Theta(x,\hat{k}) + 1)\varphi(\hat{k}), \qquad (3.3) \\
\Phi(y) &= T + \Theta(y,\hat{k})\varphi(\hat{k}) + \sum_{k<\hat{k}} \Theta(y,k)\varphi(k) \\
&\geq T + \Theta(y,\hat{k})\varphi(\hat{k}) \qquad (3.4)
\end{aligned}$$

that

$$\Phi(y) - \Phi(x) \geq (\Theta(y, \hat{k}) - \Theta(x, \hat{k}) - 1)\varphi(\hat{k}) \geq 0. \tag{3.5}$$

Here we can exclude the possibility of equality. Suppose we have equalities in (3.5). This implies that $\Theta(y, \hat{k}) = \Theta(x, \hat{k}) + 1$ and that we have equalities throughout (3.3) and (3.3). From (3.3) we obtain $\sum_{k<\hat{k}} \Theta(x, k) = |S|$, from which $\Theta(x, k) = 0$ for all $k \geq \hat{k}$. Therefore we have $\Theta(y, k) = 0$ for all $k > \hat{k}$ and $\Theta(y, \hat{k}) = 1$. From (3.4), on the other hand, we obtain $\Theta(y, k) = 0$ for all $k < \hat{k}$. This contradicts the relation $\sum_k \Theta(y, k) = |S| \geq 2$. ■

By Proposition 3.3 above, the problem of finding a dec-min element can be recast into a convex minimization problem. It is emphasized that for this equivalence, the underlying set may be any subset of $\mathbf{Z}^S$ (not necessarily an M-convex set).

Proposition 3.4. *Let D be an arbitrary subset of $\mathbf{Z}^S$, where $|S| \geq 2$, and assume that φ is $|S|$-increasing. An element m of D is decreasingly-minimal in D if and only if it minimizes $\Phi(x) = \sum_{s \in S} \varphi(x(s))$ among all members of D.* ■

Remark 3.1. The characterization of a decreasingly-minimal elements as a minimizer of a rapidly increasing convex function in Proposition 3.4 is not particularly new. Similar ideas are scattered in the literature of related topics such as majorization (Marshall–Olkin–Arnold [28]) and shifted optimization (Levin–Onn [27]). ■

Remark 3.2. The relations of being majorized ($\prec$), weakly submajorized ($\prec_{\rm w}$), and decreasingly-smaller ($\leq_{\rm dec}$) are characterized with reference to different classes of symmetric separable convex functions as follows (Proposition 2.1, [28, 4.B.2], and Proposition 3.3):

- $x \prec y \iff \sum_{i=1}^{n} \varphi(x(i)) \leq \sum_{i=1}^{n} \varphi(y(i))$ for all convex φ,
- $x \prec_{\rm w} y \iff \sum_{i=1}^{n} \varphi(x(i)) \leq \sum_{i=1}^{n} \varphi(y(i))$ for all increasing (nondecreasing) convex φ,
- $x \leq_{\rm dec} y \iff \sum_{i=1}^{n} \varphi(x(i)) \leq \sum_{i=1}^{n} \varphi(y(i))$ for all rapidly increasing convex φ. ■

3.2 M-convex function minimization in discrete convex analysis

In this section we introduce M-convex functions, a fundamental concept in discrete convex analysis [35], along with a local optimality condition for a minimizer of an M-convex function. Since a separable convex function on an M-convex set is an M-convex function (cf. Section 3.3), this optimality criterion renders alternative proofs of Theorems 3.1 and 3.2 about the dec-min elements of an M-convex set (cf. Section 3.4).

For a vector $z \in \mathbf{R}^S$ in general, we define the positive and negative supports of z as

$$\mathrm{supp}^+(z) = \{s \in S : z(s) > 0\}, \qquad \mathrm{supp}^-(z) = \{t \in S : z(t) < 0\}. \tag{3.6}$$

For a function $f : \mathbf{Z}^S \to \mathbf{R} \cup \{-\infty, +\infty\}$, the effective domain is defined as $\mathrm{dom}\, f = \{x \in \mathbf{Z}^S : -\infty < f(x) < +\infty\}$.

A function $f : \mathbf{Z}^S \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom}\, f \neq \emptyset$ is called **M-convex** if, for any $x, y \in \mathbf{Z}^S$ and $s \in \mathrm{supp}^+(x - y)$, there exists some $t \in \mathrm{supp}^-(x - y)$ such that

$$f(x) + f(y) \geq f(x - \chi_s + \chi_t) + f(y + \chi_s - \chi_t). \tag{3.7}$$

In the above statement we may change "for any $x, y \in \mathbf{Z}^S$" to "for any $x, y \in \mathrm{dom}\, f$" since if $x \notin \mathrm{dom}\, f$ or $y \notin \mathrm{dom}\, f$, (3.7) trivially holds with $f(x) + f(y) = +\infty$. We often refer to this defining property as the **exchange property** of an M-convex function. It follows from this definition that $\mathrm{dom}\, f$ consists of the integer points of an integral base-polyhedron (an M-convex set). A function f is called **M-concave** if $-f$ is M-convex. We remark that the exchange property (3.7) of an M-convex function is a quantitative extension of the symmetric exchange property of matroid bases.

A function $f : \mathbf{Z}^S \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom}\, f \neq \emptyset$ is called M$^\natural$**-convex** if, for any $x, y \in \mathbf{Z}^S$ and $s \in \mathrm{supp}^+(x - y)$, we have (i)

$$f(x) + f(y) \geq f(x - \chi_s) + f(y + \chi_s) \tag{3.8}$$

or (ii) there exists some $t \in \mathrm{supp}^-(x - y)$ for which (3.7) holds. It follows from this definition that the effective domain of an M$^\natural$-convex function consists of the integer points of an integral g-polymatroid [8]; such a set is called M$^\natural$**-convex set** in DCA. An M-convex function is M$^\natural$-convex. A function f is called **M**$^\natural$**-concave** if $-f$ is M$^\natural$-convex.

The following is a local characterization of global minimality for M- or M$^\natural$-convex functions, called the M-optimality criterion.

Theorem 3.5 ([35, Theorem 6.26])**.** *Let $f : \mathbf{Z}^S \to \mathbf{R} \cup \{+\infty\}$ be an M$^\natural$-convex function, and $x^* \in \mathrm{dom}\, f$. Then x^* is a minimizer of f if and only if it is locally minimal in the sense that*

$$f(x^*) \leq f(x^* + \chi_s - \chi_t) \quad \text{for all } s, t \in S, \tag{3.9}$$

$$f(x^*) \leq f(x^* + \chi_s) \quad \text{for all } s \in S, \tag{3.10}$$

$$f(x^*) \leq f(x^* - \chi_t) \quad \text{for all } t \in S. \tag{3.11}$$

If f is M-convex, x^* is a minimizer of f if and only if (3.9) *holds.* ■

3.3 Separable convex function minimization in discrete convex analysis

Minimization of a separable convex function over the set of integral points of an integral base-polyhedron can be treated successfully as a special case of M-convex function minimization presented in Section 3.2.

We consider a function $\Phi : \mathbf{Z}^S \to \mathbf{R} \cup \{+\infty\}$ of the form

$$\Phi(x) = \sum [\varphi_s(x(s)) : s \in S], \tag{3.12}$$

where, for each $s \in S$, the function $\varphi_s : \mathbf{Z} \to \mathbf{R} \cup \{+\infty\}$ is discrete convex (i.e., $\varphi_s(k - 1) + \varphi_s(k + 1) \geq 2\varphi_s(k)$ for all $k \in \mathrm{dom}\, \varphi_s$). Such function Φ is called a separable (discrete) convex function. We call Φ symmetric if $\varphi_s = \varphi$ for all $s \in S$.

Let $\ddddot{B}$ be the set of integral points of an integral base-polyhedron B. The problem we consider is:

$$\text{Minimize } \ \Phi(x) = \sum[\varphi_s(x(s)) : s \in S] \ \text{ subject to } \ x \in \ddddot{B}. \tag{3.13}$$

Using the indicator function $\delta : \mathbf{Z}^S \to \mathbf{R} \cup \{+\infty\}$ of $\ddddot{B}$ defined as

$$\delta(x) = \begin{cases} 0 & (x \in \ddddot{B}), \\ +\infty & (\text{otherwise}), \end{cases} \tag{3.14}$$

we can rewrite (3.13) as

$$\text{Minimize } \ \Phi(x) + \delta(x). \tag{3.15}$$

This problem is amenable to discrete convex analysis, since the separable convex function Φ is M$^\natural$-convex, the indicator function δ of an M-convex set is M-convex, and moreover, the function $\Phi + \delta$ is M-convex. Indeed it is easy to verify that these functions satisfy the defining exchange property. In this connection it is noted that the sum of an M-convex function and an M$^\natural$-convex function is not necessarily M$^\natural$-convex, but the sum of an M-convex function and a separable convex function is always M-convex (cf. Remark 4.3 in Section 4.2).

An application of the M-optimality criterion (Theorem 3.5) to our function $\Phi + \delta$ gives the following important result due to Groenevelt [17], which was shown as Proposition 2.5 and stated again for its relevance here.

Proposition 3.6 (Groenevelt [17]; cf. [14, Theorem 8.1])**.** *Let B be an integral base-polyhedron and $\ddddot{B}$ be the set of its integral elements. An element m of $\ddddot{B}$ is a minimizer of $\Phi(x) = \sum[\varphi_s(x(s)) : s \in S]$ over $\ddddot{B}$ if and only if $\varphi_s(m(s) + 1) + \varphi_t(m(t) - 1) \geq \varphi_s(m(s)) + \varphi_t(m(t))$ whenever $m + \chi_s - \chi_t \in \ddddot{B}$.* ■

In the special case of symmetric separable convex functions, with $\varphi_s = \varphi$ for all $s \in S$, we can relate the above condition to 1-tightening steps. Recall that a 1-tightening step for $m \in \ddddot{B}$ means the operation of replacing m to $m + \chi_s - \chi_t$ for some $s, t \in S$ such that $m(t) \geq m(s) + 2$ and $m + \chi_s - \chi_t \in \ddddot{B}$.

Proposition 3.7. *For any symmetric separable discrete convex function $\Phi(x) = \sum[\varphi(x(s)) : s \in S]$ with $\varphi : \mathbf{Z} \to \mathbf{R} \cup \{+\infty\}$, an element m of $\ddddot{B}$ is a minimizer of Φ over $\ddddot{B}$ if there is no 1-tightening step for m. The converse is also true if φ is strictly convex.*

Proof. By Proposition 3.6, m is a minimizer of Φ if and only if

$$\varphi(m(s) + 1) + \varphi(m(t) - 1) \geq \varphi(m(s)) + \varphi(m(t))$$

for all $s, t \in S$ such that $m + \chi_s - \chi_t \in \ddddot{B}$. By the convexity of φ, we have this inequality if $m(t) \leq m(s) + 1$, and the converse is also true when φ is strictly convex. Finally we note that there is no 1-tightening step for m if and only if $m(t) \leq m(s) + 1$ for all $s, t \in S$ such that $m + \chi_s - \chi_t \in \ddddot{B}$. ■

3.4 DCA-based proofs of the theorems

The combination of Proposition 3.7 with Proposition 3.4 provides alternative proofs of Theorems 3.1 and 3.2.

Proof of Theorem 3.1: Let Φ be a symmetric separable convex function with rapidly increasing φ. By Proposition 3.4, m is dec-min if and only if m is a minimizer of Φ. On the other hand, since Φ is strictly convex, Proposition 3.7 shows that m is a minimizer of Φ if and only if there is no 1-tightening step for m. Therefore, m is a dec-min element of $\overset{....}{B}$ if and only if there is no 1-tightening step for m.

Proof of Theorem 3.2: Let Φ be a symmetric separable convex function. By Proposition 3.7, m is a minimizer of Φ if there is no 1-tightening step for m; and the converse is also true for strictly convex Φ. Theorem 3.1, on the other hand, shows that there is no 1-tightening step for m if and only if m is a dec-min element. Therefore, m is a minimizer of Φ if m is a dec-min element of $\overset{....}{B}$; and the converse is also true for strictly convex Φ.

3.5 Extension to generalized polymatroids

In this section we shed the light of DCA on the result of Tamir [40] about the majorization ordering in generalized polymatroids (g-polymatroids). This is based on the fact that the set $\overset{....}{Q}$ of integral points of an integral g-polymatroid Q is an M$^\natural$-convex set, and accordingly, the indicator function of $\overset{....}{Q}$ is an M$^\natural$-convex function. See [8] for the definition of g-polymatroids and [35] for more about M$^\natural$-convexity.

The M-optimality criterion (Theorem 3.5) immediately implies the following generalization of Proposition 3.6.

Proposition 3.8. *Let Q be an integral g-polymatroid and $\overset{....}{Q}$ be the set of its integral elements. An element m of $\overset{....}{Q}$ is a minimizer of a separable convex function $\Phi(x) = \sum[\varphi_s(x(s)) : s \in S]$ over $\overset{....}{Q}$ if and only if*

- $\varphi_s(m(s) + 1) + \varphi_t(m(t) - 1) \geq \varphi_s(m(s)) + \varphi_t(m(t))$ *whenever* $m + \chi_s - \chi_t \in \overset{....}{Q}$,
- $\varphi_s(m(s) + 1) \geq \varphi_s(m(s))$ *whenever* $m + \chi_s \in \overset{....}{Q}$, *and*
- $\varphi_t(m(t) - 1) \geq \varphi_t(m(t))$ *whenever* $m - \chi_t \in \overset{....}{Q}$. ■

Proposition 3.7 for a symmetric separable convex function $\Phi(x) = \sum[\varphi(x(s)) : s \in S]$ can be adapted to g-polymatroids under the additional assumption of monotonicity of φ. Let B denote the set of minimal elements of an integral g-polymatroid Q, and $\overset{....}{B}$ the set of integral members of B. As is well known, B is an integral base-polyhedron and $\overset{....}{B}$ is an M-convex set.

Proposition 3.9. *Let Φ be a symmetric separable convex function represented as $\Phi(x) = \sum[\varphi(x(s)) : s \in S]$ with monotone non-decreasing discrete convex φ. An element m of $\overset{....}{Q}$ is a minimizer of Φ over $\overset{....}{Q}$ if m belongs to $\overset{....}{B}$ and $m(t) \leq m(s) + 1$ whenever $m + \chi_s - \chi_t$ is in $\overset{....}{B}$. The converse is also true if φ is strictly convex and strictly monotone increasing.* ■

On the basis of Proposition 3.9 we can show the existence of a least weakly submajorized element in $\overset{....}{Q}$, which is the result of Tamir [40].

4 Min-max formulas

Key min-max formulas on discrete decreasing minimization, established by constructive methods in Part I [9], are derived here from the Fenchel-type discrete duality in discrete convex analysis. These formulas can in fact be derived from a special case of the Fenchel-type discrete duality where a separable convex function is minimized over an M-convex set. This special case often provides interesting min-max relations in applications and deserves particular attention. The (general) Fenchel-type discrete duality is described in Section 4.2 and its special case for separable convex functions in Section 4.3.

4.1 Min-max formulas for decreasing minimization

In this section we treat the formulas (4.1), (4.2), (4.3), and (4.4) below. Recall that p is an integer-valued (fully) supermodular function on the ground-set S describing a base-polyhedron B and $\hat{p}$ is the linear extension (Lovász extension) of p, whose definition is given in (4.19) in Section 4.3.

- [9, Theorem 6.6] For the square-sum we have

$$\min\{\sum_{s\in S} m(s)^2 : m \in \overset{....}{B}\} = \max\{\hat{p}(\pi) - \sum_{s\in S} \left\lfloor \frac{\pi(s)}{2} \right\rfloor \left\lceil \frac{\pi(s)}{2} \right\rceil : \pi \in \mathbf{Z}^S\}. \tag{4.1}$$

- [9, Theorem 4.1] For the largest component β_1 of a max-minimizer of $\overset{....}{B}$, we have

$$\beta_1 = \max\{\left\lceil \frac{p(X)}{|X|} \right\rceil : \emptyset \neq X \subseteq S\}. \tag{4.2}$$

  Recall that β_1 is equal to the largest component of any dec-min element of $\overset{....}{B}$.

- [9, Theorem 4.3] For the minimum number r_1 of β_1-valued components of a β_1-covered member of $\overset{....}{B}$, we have

$$r_1 = \max\{p(X) - (\beta_1 - 1)|X| : X \subseteq S\}. \tag{4.3}$$

  Recall that $r_1 = |\{s \in S : m(s) = \beta_1\}|$ for any dec-min element m of $\overset{....}{B}$.

Moreover, the following min-max formula will be established in Section 4.7 as a generalization of (4.3). We refer to $\sum_{s\in S}(m(s) - a)^+$ in the minimization below as the **total a-excess** of m.

- For each integer a, we have

$$\min\{\sum_{s\in S}(m(s) - a)^+ : m \in \overset{....}{B}\} = \max\{p(X) - a|X| : X \subseteq S\}. \tag{4.4}$$

Note that this formula (4.4) for $a = \beta_1 - 1$ reduces to the formula (4.3) for r_1. It will be shown in Theorem 4.7 that an element of $\ddot{B}$ is decreasingly minimal if and only if it is a minimizer of the left-hand side of (4.4) universally for all $a \in \mathbf{Z}$. We remark that the minimization problem above is known to be most fundamental in the literature of majorization, whereas the function $p(X) - a|X|$ to be maximized plays the pivotal role in characterizing the canonical partition and the essential value-sequence (cf., Section 6.3). Thus the min-max formula (4.4) reinforces the link between the present study and the theory of majorization.

4.2 Fenchel-type discrete duality in discrete convex analysis

In this section we describe an important result in DCA, the Fenchel-type discrete duality theorem, which we use to derive the min-max formulas related to dec-min elements. The Fenchel-type discrete duality theorem in DCA originates in Murota [32] and is formulated for integer-valued functions in [33, 35].

For any integer-valued functions $f : \mathbf{Z}^S \to \mathbf{Z} \cup \{+\infty\}$ and $h : \mathbf{Z}^S \to \mathbf{Z} \cup \{-\infty\}$, we define their (convex and concave) **conjugate functions** by

$$f^{\bullet}(\pi) = \sup\{\langle \pi, x \rangle - f(x) : x \in \mathbf{Z}^S\} \qquad (\pi \in \mathbf{Z}^S), \tag{4.5}$$

$$h^{\circ}(\pi) = \inf\{\langle \pi, x \rangle - h(x) : x \in \mathbf{Z}^S\} \qquad (\pi \in \mathbf{Z}^S), \tag{4.6}$$

where $\langle \pi, x \rangle$ means the (standard) inner product of vectors π and x. Note that both x and π are integer vectors. Since the functions are integer-valued, the supremum in (4.5) is attained if it is finite-valued. Similarly for the infimum in (4.6). Accordingly, we henceforth write "max" and "min" in place of "sup" in (4.5) and "inf" in (4.6), respectively.

The Fenchel-type discrete duality is concerned with the relationship between the minimum of $f(x) - h(x)$ over $x \in \mathbf{Z}^S$ and the maximum of $h^{\circ}(\pi) - f^{\bullet}(\pi)$ over $\pi \in \mathbf{Z}^S$. By the definition of the conjugate functions in (4.5) and (4.6) we have inequalities (called the Fenchel–Young inequalities)

$$f(x) + f^{\bullet}(\pi) \geq \langle \pi, x \rangle, \tag{4.7}$$

$$h(x) + h^{\circ}(\pi) \leq \langle \pi, x \rangle \tag{4.8}$$

for any x and π, and hence

$$f(x) - h(x) \geq h^{\circ}(\pi) - f^{\bullet}(\pi) \tag{4.9}$$

for any x and π. Therefore we have **weak duality**:

$$\min\{f(x) - h(x) : x \in \mathbf{Z}^S\} \geq \max\{h^{\circ}(\pi) - f^{\bullet}(\pi) : \pi \in \mathbf{Z}^S\}. \tag{4.10}$$

It is noted, however, that in this expression using "min" and "max" we do not exclude the possibility of the unbounded case where $\min\{\cdots\}$ and/or $\max\{\cdots\}$ are equal to $-\infty$ or $+\infty$ (we avoid using "inf" and "sup" for wider audience). Here we note the following.

1. If $\mathrm{dom}\, f \cap \mathrm{dom}\, h \neq \emptyset$ and $\mathrm{dom}\, f^{\bullet} \cap \mathrm{dom}\, h^{\circ} \neq \emptyset$, both $\min\{f(x) - h(x) : x \in \mathbf{Z}^S\}$ and $\max\{h^{\circ}(\pi) - f^{\bullet}(\pi) : \pi \in \mathbf{Z}^S\}$ are finite integers and the minimum and the maximum are attained by some x and π since the functions are integer-valued.

2. If $\mathrm{dom}\, f \cap \mathrm{dom}\, h = \emptyset$, we understand (by convention) that the minimum of $f - h$ is equal to $+\infty$, that is, $\min\{f(x) - h(x) : x \in \mathbf{Z}^S\} = +\infty$.

3. If $\mathrm{dom}\, f^\bullet \cap \mathrm{dom}\, h^\circ = \emptyset$, we understand (by convention) that the maximum of $h^\circ - f^\bullet$ is equal to $-\infty$, that is, $\max\{h^\circ(\pi) - f^\bullet(\pi) : \pi \in \mathbf{Z}^S\} = -\infty$.

We say that **strong duality** holds if equality holds in (4.10).

The strong duality does hold for a pair of an $\mathrm{M}^\natural$-convex function f and an $\mathrm{M}^\natural$-concave function h, as the following theorem shows. This is called the **Fenchel-type discrete duality theorem** [33, 35]. To be more precise, we need to assume that at lease one of the following two conditions is satisfied:
(i) there exists x for which both $f(x)$ and $h(x)$ are finite (**primal feasibility**, $\mathrm{dom}\, f \cap \mathrm{dom}\, h \neq \emptyset$),
(ii) there exists π for which both $f^\bullet(\pi)$ and $h^\circ(\pi)$ are finite (**dual feasibility**, $\mathrm{dom}\, f^\bullet \cap \mathrm{dom}\, h^\circ \neq \emptyset$).
Note that these two feasibility conditions, (i) and (ii), are mutually independent, and there is an example for which both conditions fail simultaneously [35, p.220, Note 8.18].

Theorem 4.1 (Fenchel-type discrete duality theorem [33, 35])**.** *Let* $f : \mathbf{Z}^S \to \mathbf{Z} \cup \{+\infty\}$ *be an integer-valued* $\mathrm{M}^\natural$*-convex function and* $h : \mathbf{Z}^S \to \mathbf{Z} \cup \{-\infty\}$ *be an integer-valued* $\mathrm{M}^\natural$*-concave function such that* $\mathrm{dom}\, f \cap \mathrm{dom}\, h \neq \emptyset$ *or* $\mathrm{dom}\, f^\bullet \cap \mathrm{dom}\, h^\circ \neq \emptyset$*. Then we have*

$$\min\{f(x) - h(x) : x \in \mathbf{Z}^S\} = \max\{h^\circ(\pi) - f^\bullet(\pi) : \pi \in \mathbf{Z}^S\}. \quad (4.11)$$

This common value is finite if and only if $\mathrm{dom}\, f \cap \mathrm{dom}\, h \neq \emptyset$ *and* $\mathrm{dom}\, f^\bullet \cap \mathrm{dom}\, h^\circ \neq \emptyset$*, and then the minimum and the maximum are attained.* ■

The essential content of the above theorem may be expressed as follows: If $f(x) - h(x)$ is bounded from below, then $h^\circ(\pi) - f^\bullet(\pi)$ is bounded from above, and the minimum of $f(x) - h(x)$ and the maximum of $h^\circ(\pi) - f^\bullet(\pi)$ coincide.

Remark 4.1. The Fenchel-type duality theorem is the central duality theorem in discrete convex analysis. The duality phenomenon captured by this theorem can be formulated in several different, mutually equivalent, forms including the M-separation theorem [35, Theorem 8.15], the L-separation theorem [35, Theorem 8.16], and the M-convex intersection theorem [35, Theorem 8.17]. These duality theorems include a number of important results as special cases such as Edmonds' intersection theorem, Fujishige's Fenchel-type duality theorem for submodular set functions, Frank's discrete separation theorem for submodular/supermodular functions, and Frank's weight splitting theorem for the weighted matroid intersection problem. See [35, Section 8.2, Fig.8.2] for this relationship. ■

Remark 4.2. The Fenchel-type discrete duality theorem offers an optimality certificate for the minimization problem of $f(x) - h(x)$. Two cases are to be distinguished.

1. If the explicit forms of the conjugate functions $f^\bullet(\pi)$ and $h^\circ(\pi)$ are known, we can easily evaluate the value of $h^\circ(\pi) - f^\bullet(\pi)$ for any integer vector π. Given an integral vector π as a certificate of optimality for an allegedly optimal x, we only have to compute the values of $f(x) - h(x)$ and $h^\circ(\pi) - f^\bullet(\pi)$ and compare the two values

(integers) for their equality. Thus the availability of explicit forms of the conjugate functions is computationally convenient as well as intuitively appealing. The min-max formula (4.1) for the square-sum minimization over an M-convex set falls into this case.

2. Even if explicit forms of the conjugate functions are not available, the Fenchel-type discrete duality theorem offers a computationally efficient (polynomial-time) method for verifying the optimality if it is combined with the M-optimality criterion (Theorem 3.5). We shall discuss this method in Section 5.1; see Remark 5.1. ■

The conjugate of an M$^\natural$-convex function is endowed with another kind of discrete convexity, called L$^\natural$-convexity. A function $g : \mathbf{Z}^S \to \mathbf{R} \cup \{+\infty\}$ with $\mathrm{dom}\, g \neq \emptyset$ is called **L$^\natural$-convex** if it satisfies the inequality

$$g(\pi) + g(\tau) \geq g\left(\left\lceil \frac{\pi+\tau}{2} \right\rceil\right) + g\left(\left\lfloor \frac{\pi+\tau}{2} \right\rfloor\right) \qquad (\pi, \tau \in \mathbf{Z}^S), \tag{4.12}$$

where, for $z \in \mathbf{R}$ in general, $\lceil z \rceil$ denotes the smallest integer not smaller than z (rounding-up to the nearest integer) and $\lfloor z \rfloor$ the largest integer not larger than z (rounding-down to the nearest integer), and this operation is extended to a vector by componentwise applications. The property (4.12) is referred to as **discrete midpoint convexity**. A function g is called **L$^\natural$-concave** if $-g$ is L$^\natural$-convex.

The following is a local characterization of global maximality for L$^\natural$-concave functions, called the L-optimality criterion (concave version).

Theorem 4.2 ([35, Theorem 7.14]). *Let $g : \mathbf{Z}^S \to \mathbf{R} \cup \{-\infty\}$ be an L$^\natural$-concave function, and $\pi^* \in \mathrm{dom}\, g$. Then π^* is a maximizer of g if and only if it is locally maximal in the sense that*

$$g(\pi^*) \geq g(\pi^* - \chi_Y) \quad \text{for all } \ Y \subseteq S, \tag{4.13}$$

$$g(\pi^*) \geq g(\pi^* + \chi_Y) \quad \text{for all } \ Y \subseteq S. \tag{4.14}$$

■

The reader is referred to [35, Chapter 7] for more properties of L$^\natural$-convex functions and [35, Chapter 8] for the conjugacy between M$^\natural$-convexity and L$^\natural$-convexity. In particular, [35, Figure 8.1] offers the whole picture of conjugacy relationship.

Remark 4.3. In Theorem 4.1 the functions $f(x)$ and $-h(x)$ are both M$^\natural$-convex, but the function $f(x) - h(x)$ to be minimized on the left-hand side of (4.11) is not necessarily M$^\natural$-convex, since the sum of M$^\natural$-convex functions may not be M$^\natural$-convex. To see this, consider two M-convex sets $\ddot{B}_1$ and $\ddot{B}_2$ associated with integral base-polyhedra B_1 and B_2, respectively, and for $i = 1, 2$, let f_i be the indicator function of $\ddot{B}_i$ (i.e., $f_i(x) = 0$ if $x \in \ddot{B}_i$, and $f_i(x) = +\infty$ if $x \in \mathbf{Z}^S \setminus \ddot{B}_i$). The function $f_1 + f_2$ is the indicator function of the set of integer points in the intersection $B_1 \cap B_2$, which is not a base-polyhedron in general. This argument also shows that the left-hand side of (4.11) is a nonlinear generalization of the weighted polymatroid intersection problem; see [35, Section 8.2.3] for details. ■

Remark 4.4. Functions $h^{\circ}(\pi)$ and $f^{\bullet}(\pi)$ in Theorem 4.1 are L$^{\natural}$-concave and L$^{\natural}$-convex, respectively. Since the sum of L$^{\natural}$-concave functions is L$^{\natural}$-concave, the function $h^{\circ}(\pi)-f^{\bullet}(\pi)$ to be maximized on the right-hand side of (4.11) is an L$^{\natural}$-concave function. In contrast, the function $f(x) - h(x)$ to be minimized on the left-hand side of (4.11) is not an M$^{\natural}$-convex function, as explained in Remark 4.3 above. In this sense, the left-hand side (minimization) and the right-hand side (maximization) are not symmetric. ■

4.3 Min-max formula for separable convex functions on an M-convex set

In this section the Fenchel-type discrete duality theorem is tailored to the problem of minimizing a separable convex function over an M-convex set. This special case deserves particular attention as it is suitable and sufficient for our use in decreasing minimization.

Consider the problem of minimizing an integer-valued separable convex function

$$\Phi(x) = \sum [\varphi_s(x(s)) : s \in S] \tag{4.15}$$

over an M-convex set $\ddddot{B}$, where each $\varphi_s : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ is an integer-valued discrete convex function in a single integer variable. This problem is equivalent to minimizing $\Phi(x) + \delta(x)$, where δ denotes the indicator function of $\ddddot{B}$ defined in (3.14).

In Section 3.3 we have regarded the function $\Phi + \delta$ as an M-convex function and applied the M-optimality criterion to derive some results obtained in Part I [9]. In contrast, we are now going to apply the Fenchel-type discrete duality theorem to the minimization of the function $\Phi + \delta = \Phi - (-\delta)$. In so doing we can separate the roles of the constraining M-convex set and the objective function $\Phi(x)$ itself.

With the choice of $f = \Phi$ and $h = -\delta$ in the min-max relation $\min\{f(x) - h(x)\} = \max\{h^{\circ}(\pi) - f^{\bullet}(\pi)\}$ in (4.11), the left-hand side represents minimization of Φ over the M-convex set $\ddddot{B}$. We denote the conjugate function of φ_s by ψ_s, which is a function $\psi_s : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ defined by

$$\psi_s(\ell) = \max\{k\ell - \varphi_s(k) : k \in \mathbf{Z}\} \qquad (\ell \in \mathbf{Z}). \tag{4.16}$$

Then the conjugate function of f is given by

$$f^{\bullet}(\pi) = \sum [\psi_s(\pi(s)) : s \in S] \qquad (\pi \in \mathbf{Z}^S). \tag{4.17}$$

On the other hand, the conjugate function h° of h is given by

$$h^{\circ}(\pi) = \min\{\langle \pi, x\rangle + \delta(x) : x \in \mathbf{Z}^S\} = \min\{\langle \pi, x\rangle : x \in \ddddot{B}\} = \hat{p}(\pi) \quad (\pi \in \mathbf{Z}^S) \tag{4.18}$$

in terms of the linear extension (Lovász extension) $\hat{p}$ of p. Recall that, for any set function p, $\hat{p}$ is defined [9, Part I, Section 6.2] as

$$\hat{p}(\pi) = p(I_n)\pi(s_n) + \sum_{j=1}^{n-1} p(I_j)[\pi(s_j) - \pi(s_{j+1})], \tag{4.19}$$

where $n = |S|$, the elements of S are indexed in such a way that $\pi(s_1) \geq \cdots \geq \pi(s_n)$, and $I_j = \{s_1, \ldots, s_j\}$ for $j = 1, \ldots, n$. If p is supermodular, we have

$$\hat{p}(\pi) = \min\{\pi x : x \in \ddddot{B}\}. \tag{4.20}$$

Substituting (4.17) and (4.18) into (4.11) we obtain (4.21) below.

Theorem 4.3. *Assume that* (i) *there exists $x \in \ddddot{B}$ such that $\varphi_s(x(s)) < +\infty$ for all $s \in S$ (primal feasibility) or* (ii) *there exists $\pi \in \mathbf{Z}^S$ such that $\hat{p}(\pi) > -\infty$ and $\psi_s(\pi(s)) < +\infty$ for all $s \in S$ (dual feasibility). Then we have the min-max relation:*

$$\min\{\sum_{s \in S} \varphi_s(x(s)) : x \in \ddddot{B}\} = \max\{\hat{p}(\pi) - \sum_{s \in S} \psi_s(\pi(s)) : \pi \in \mathbf{Z}^S\}. \tag{4.21}$$

The unbounded case with both sides being equal to $-\infty$ or $+\infty$ is also a possibility. ■

Since $\hat{p}(\pi)$ is an L$^\natural$-concave function and $\sum[\psi_s(\pi(s)) : s \in S]$ is an L$^\natural$-convex function, the function $g(\pi) := \hat{p}(\pi) - \sum[\psi_s(\pi(s)) : s \in S]$ to be maximized on the right-hand side of (4.21) is an L$^\natural$-concave function (cf. Remark 4.4). We state this as a proposition for later reference.

Proposition 4.4. *The function $g(\pi) = \hat{p}(\pi) - \sum[\psi_s(\pi(s)) : s \in S]$ is L$^\natural$-concave.* ■

When specialized to a symmetric function Φ, the min-max formula (4.21) is simplified to

$$\min\{\sum_{s \in S} \varphi(x(s)) : x \in \ddddot{B}\} = \max\{\hat{p}(\pi) - \sum_{s \in S} \psi(\pi(s)) : \pi \in \mathbf{Z}^S\}, \tag{4.22}$$

where $\varphi : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ is any integer-valued discrete convex function and $\psi : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ is the conjugate of φ defined as $\psi(\ell) = \max\{k\ell - \varphi(k) : k \in \mathbf{Z}\}$ for $\ell \in \mathbf{Z}$. With appropriate choices of φ in (4.22) we shall derive the formulas (4.1), (4.2), and (4.3).

In applications of (4.21) (resp., (4.22)) with concrete functions φ_s (resp., φ), it is often the case that the conjugate functions ψ_s (resp., ψ) can be given explicitly. This is illustrated in Section 7.1.

4.4 DCA-based proof of the min-max formula for the square-sum

The min-max formula (4.1) for the square-sum can be derived immediately from our duality formula (4.22). For $\varphi(k) = k^2$, the conjugate function $\psi(\ell)$ for $\ell \in \mathbf{Z}$ is given explicitly as

$$\psi(\ell) = \max\{k\ell - k^2 : k \in \mathbf{Z}\} = \max\{k\ell - k^2 : k \in \{\lfloor \ell/2 \rfloor, \lceil \ell/2 \rceil\}\} = \left\lfloor \frac{\ell}{2} \right\rfloor \cdot \left\lceil \frac{\ell}{2} \right\rceil. \tag{4.23}$$

The substitution of (4.23) into (4.22) yields (4.1). Note that the primal feasibility is satisfied since $\varphi(k)$ is finite for all k.

Remark 4.5. In Part I [9, Section 6.2] we provided a relatively simple algorithmic proof for the min-max formula (4.1), which did not use any tool from DCA. However, to figure out the min-max formula itself without the DCA background seems rather difficult. Indeed, the present authors first identified the formula (4.1) via DCA as above, and then came up with the algorithmic proof. This example demonstrates the role and effectiveness of DCA. ■

Remark 4.6. In Part I [9, Section 6.1] we have characterized a dec-min element as a square-sum minimizer and also as a difference-sum minimizer. Whereas a min-max formula can be obtained by DCA for the square-sum, this is not the case with the difference-sum. This is because the difference-sum is not M$^{\natural}$-convex (though it is L-convex), and therefore difference-sum minimization over an M-convex set does not fit into the framework of the Fenchel-type discrete duality in DCA. ■

4.5 DCA-based proof of the formula for β_1

The formula (4.2) for the largest component β_1 of a max-minimizer of $\ddot{B}$ can also be derived from our duality formula (4.22). With an integer parameter α we choose

$$\varphi(k) = \begin{cases} 0 & (k \leq \alpha), \\ +\infty & (k \geq \alpha + 1) \end{cases}$$

in (4.22). By the definition of β_1, the left-hand side of (4.22) is equal to zero if $\alpha \geq \beta_1$, and equal to $+\infty$ if $\alpha \leq \beta_1 - 1$. Hence β_1 is equal to the minimum of α for which the left-hand side is equal to zero.

The conjugate function ψ of φ is given by

$$\psi(\ell) = \max\{k\ell : k \leq \alpha\} = \begin{cases} +\infty & (\ell \leq -1), \\ 0 & (\ell = 0), \\ \alpha\ell & (\ell \geq 1). \end{cases} \tag{4.24}$$

Both $\hat{p}(\pi)$ and $\psi(\ell)$ are positively homogeneous (i.e., $\hat{p}(\lambda\pi) = \lambda\hat{p}(\pi)$ and $\psi(\lambda\ell) = \lambda\psi(\ell)$ for nonnegative integers λ). This implies, in particular, that the maximization problem on the right-hand side of (4.22) is feasible for all α and hence the identity (4.22) holds, which reads either $0 = 0$ or $+\infty = +\infty$. Since β_1 is the minimum of α for which the left-hand side is equal to 0, we can say that β_1 is the minimum of α for which the right-hand side is equal to 0.

Finally, we consider the condition that ensures $\pi^* = \mathbf{0}$ to be a maximizer of the function $g(\pi) := \hat{p}(\pi) - \sum_{s \in S} \psi(\pi(s))$. By the L$^{\natural}$-concavity of this function we can make use of Theorem 4.2 (L-optimality criterion). The first condition (4.13) in Theorem 4.2 is satisfied trivially by (4.24), whereas the second condition (4.14) reads $g(\pi^* + \chi_Y) = p(Y) - \alpha|Y| \leq 0$. Therefore, the right-hand side of (4.22) is equal to zero if and only if $\max\{p(Y) - \alpha|Y| : Y \subseteq S\} = 0$, from which follows the formula (4.2).

4.6 DCA-based proof of the formula for r_1

The formula (4.3) for the minimum number r_1 of β_1-valued components of a β_1-covered member of $\ddot{B}$ can also be derived from our duality formula (4.22). We choose

$$\varphi(k) = \begin{cases} 0 & (k \leq \beta_1 - 1), \\ 1 & (k = \beta_1), \\ +\infty & (k \geq \beta_1 + 1), \end{cases} \tag{4.25}$$

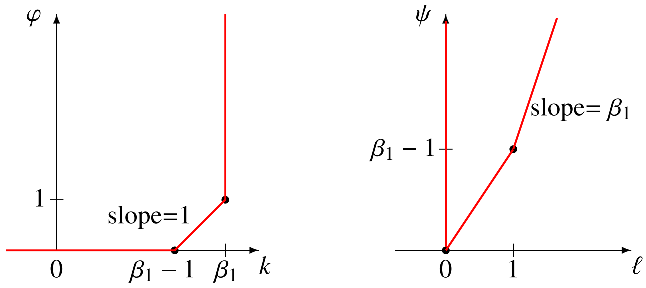

Figure 1: Mutually conjugate discrete convex functions φ and ψ in (4.25) and (4.26)

whose graph is given by the left of Fig. 1. By the definitions of β_1 and r_1, the minimum in (4.22) is equal to r_1. In particular, the primal problem is feasible, and hence the identity (4.22) holds.

The conjugate function ψ of φ is given by

$$\begin{aligned}\psi(\ell) &= \max\{\ \max\{k\ell : k \leq \beta_1 - 1\},\ \ \beta_1\ell - 1\ \} \\ &= \begin{cases} +\infty & (\ell \leq -1), \\ 0 & (\ell = 0), \\ \beta_1\ell - 1 & (\ell \geq 1), \end{cases} \end{aligned} \tag{4.26}$$

whose graph is given by the right of Fig. 1. In considering the maximum of $g(\pi) := \hat{p}(\pi) - \sum_{s\in S} \psi(\pi(s))$ over $\pi \in \mathbf{Z}^S$, we may restrict π to $\{0,1\}$-vectors, as shown in Lemma 4.5 below. For $\pi = \chi_X \in \{0,1\}^S$ with $X \subseteq S$, we have $\hat{p}(\pi) = \hat{p}(\chi_X) = p(X)$ and $\sum_{s\in S} \psi(\pi(s)) = \sum_{s\in S} \psi(\chi_X(s)) = \sum_{s\in X} \psi(1) = (\beta_1 - 1)|X|$, and therefore, the right-hand side of (4.22) is equal to $\max\{p(X) - (\beta_1 - 1)|X| : X \subseteq S\}$. Thus the formula (4.3) is derived.

Lemma 4.5. *There exists a* $\{0,1\}$*-vector* π *that attains the maximum of* $g(\pi)$ *over* $\pi \in \mathbf{Z}^S$.

Proof. Note first that g is an L$^\natural$-concave function, and define $a = \beta_1 - 1$. Let $A \subseteq S$ be a maximizer of $p(X) - a|X|$ over all subsets of S, and $\pi^* = \chi_A$. Then $g(\pi^*) = p(A) - a|A|$. We will show that the conditions (4.13) and (4.14) in the L-optimality criterion (Theorem 4.2) are satisfied.

Proof of $g(\pi^*) \geq g(\pi^* - \chi_Y)$ in (4.13): We may assume $Y \subseteq A$, since, otherwise, $\pi^* - \chi_Y \notin \mathrm{dom}\, g$ by (4.26). If $Y \subseteq A$, we have $\pi^* - \chi_Y = \chi_{A\setminus Y} = \chi_Z$, where $Z = A \setminus Y$. Hence,

$$g(\pi^* - \chi_Y) = g(\chi_Z) = p(Z) - a|Z| \leq p(A) - a|A| = g(\pi^*).$$

Proof of $g(\pi^*) \geq g(\pi^* + \chi_Y)$ in (4.14): Since

$$(\pi^* + \chi_Y)(s) = (\chi_A + \chi_Y)(s) = \begin{cases} 2 & (s \in A \cap Y), \\ 1 & (s \in (A \cup Y) \setminus (A \cap Y)), \\ 0 & (s \in S \setminus (A \cup Y)), \end{cases}$$

we have

$$\hat{p}(\pi^* + \chi_Y) = p(A \cap Y) + p(A \cup Y),$$
$$\sum_{s\in S} \psi((\pi^* + \chi_Y)(s)) = (2\beta_1 - 1)|A \cap Y| + (\beta_1 - 1)|(A \cup Y) \setminus (A \cap Y)| = \beta_1|A \cap Y| + a|A \cup Y|$$

by the definition (4.19) of $\hat{p}$ and the expression (4.26) of the conjugate function ψ. Hence

$$g(\pi^* + \chi_Y) = (p(A \cap Y) + p(A \cup Y)) - (\beta_1|A \cap Y| + a|A \cup Y|)$$
$$= (p(A \cap Y) - \beta_1|A \cap Y|) + (p(A \cup Y) - a|A \cup Y|).$$

Here we have

$$p(A \cap Y) - \beta_1|A \cap Y| \leq 0,$$
$$p(A \cup Y) - a|A \cup Y| \leq p(A) - a|A| = g(\pi^*),$$

since $(\beta_1, \beta_1, \ldots, \beta_1)$ belongs to the supermodular polyhedra defined by p, A is a maximizer of $p(X) - a|X|$, and $p(A) - a|A| = g(\pi^*)$. Therefore, $g(\pi^* + \chi_Y) \leq g(\pi^*)$. ■

4.7 Total a-excess and decreasing minimality

In this section, we explore a link between decreasing minimality and the total a-excess announced at the beginning of Section 4. The minimization problem in (4.4) (or (4.32) below) is most fundamental in the literature of majorization. Indeed, a least majorized element is characterized as a universal minimizer for all $a \in \mathbf{Z}$ (Proposition 2.1). On the other hand, the function $p(X) - a|X|$ to be maximized plays the pivotal role in characterizing the canonical partition and the essential value-sequence (cf., Section 6.3).

As a preparation, we recall ([8], [39]) that, for a nonnegative and (fully) supermodular function p_0, the polyhedron $C = \{x : \widetilde{x} \geq p_0\}$ is called a contra-polymatroid. Note that the nonnegativity and supermodularity of p_0 imply that p_0 is monotone non-decreasing and that $C \subseteq \mathbf{R}^S_+$, that is, every member of C is a nonnegative vector. When p_0 is integer-valued, C is an integer polyhedron. The corresponding version of Edmonds' greedy algorithm for polymatroids implies that C uniquely determines p_0, namely,

$$p_0(X) = \min\{\widetilde{z}(X) : z \in C\}. \tag{4.27}$$

It is known that, for a supermodular function p_1 with possibly negative values, the polyhedron

$$C(p_1) := \{x : x \geq \mathbf{0},\ \widetilde{x} \geq p_1\} \tag{4.28}$$

is a contra-polymatroid[1], for which the unique nonnegative supermodular bounding function p_0 is given by

$$p_0(X) = \max\{p_1(Y) : Y \subseteq X\}. \tag{4.29}$$

It follows from (4.27), (4.29), and the integrality of the polyhedron $C(p_1)$ that

$$\min\{\widetilde{z}(S) : z \in \ddot{C}(p_1)\} = \max\{p_1(X) : X \subseteq S\}, \tag{4.30}$$

where $\ddot{C}(p_1)$ denotes the set of the integral members of $C(p_1)$.

Lemma 4.6. *Let $B = B'(p)$ be an (integral) base-polyhedron defined by an integer-valued supermodular function p. For a vector $g : S \to \mathbf{Z}$,*

$$\min\{\sum[(m(s) - g(s))^+ : s \in S] : m \in \ddot{B}\} = \max\{p(X) - \widetilde{g}(X) : X \subseteq S\}. \tag{4.31}$$

[1]In the literature, (4.28) is used sometimes as the definition of a contra-polymatroid.

Proof. It is known (for example, from the discrete separation theorem for submodular set functions or from a version of Edmonds' polymatroid intersection theorem) that, for a function $g' : S \to \mathbf{Z}$, there is an element $m \in \overset{....}{B}$ for which $m \leq g'$ if and only if $p \leq \widetilde{g}'$. Therefore the minimization problem on the left-hand side of (4.31) is equivalent to finding a lowest lifting $g' := g + z$ of g with $z \geq \mathbf{0}$ such that $p \leq \widetilde{g}'$. That is, the minimum on the left-hand side of (4.31) is equal to $\min\{\, \widetilde{z}(S) : z \geq \mathbf{0},\ \widetilde{z} \geq p - \widetilde{g}\,\}$. By applying (4.30) to $p_1 := p - \widetilde{g}$, we obtain that this latter minimum is indeed equal to the right-hand side of (4.31). ■

The following theorem reinforces the link between the present study and the theory of majorization.

Theorem 4.7. *Let B be a base-polyhedron described by an integer-valued supermodular function p and $\overset{....}{B}$ the set of integral elements of B. For each integer a, we have the following min-max relation for the minimum of the total a-excess of the members of $\overset{....}{B}$:*

$$\min\{\sum_{s\in S}(m(s)-a)^+ : m \in \overset{....}{B}\} = \max\{p(X) - a|X| : X \subseteq S\}. \tag{4.32}$$

Moreover, an element of $\overset{....}{B}$ is a dec-min element of $\overset{....}{B}$ if and only if it is a minimizer on the left-hand side for every $a \in \mathbf{Z}$.

Proof. The min-max formula (4.32) follows from Lemma 4.6 as it is a special case of (4.31) when $g = (a, a, \ldots, a)$. Theorem 3.2 shows that any dec-min element of $\overset{....}{B}$ is a minimizer in (4.32) for every $a \in \mathbf{Z}$. The converse is also true, since $\sum[(x(s) - a)^+ : s \in S] = \sum[(y(s) - a)^+ : s \in S]$ for every $a \in \mathbf{Z}$ implies $x{\downarrow} = y{\downarrow}$. Therefore, an element of $\overset{....}{B}$ is dec-min if and only if it is a universal minimizer for every $a \in \mathbf{Z}$. ■

The established formula (4.32) generalizes the formula (4.3) for r_1. Indeed, the total a-excess for $a = \beta_1 - 1$ is given as

$$\sum_{s\in S}(m(s)-a)^+ = \sum_{s\in S}(m(s)-(\beta_1-1))^+ = |\{s \in S : m(s) = \beta_1\}| = r_1$$

for any dec-min element m of $\overset{....}{B}$.

For any dec-min element m of $\overset{....}{B}$ and for $k = \beta_1, \beta_1 - 1, \beta_1 - 2, \ldots$, let $\Theta(m, k)$ denote the number of components of m whose value are equal to k, that is,

$$\Theta(m, k) = |\{s \in S : m(s) = k\}|.$$

Note that $\Theta(m, \beta_1) = r_1$ and $\Theta(m, k)$ does not depend on the choice of m. Since

$$\sum_{s\in S}(m(s) - (\beta_1 - i - 1))^+ = \sum_{j=0}^{i}(i + 1 - j)\,\Theta(m, \beta_1 - j) \qquad (i = 0, 1, 2, \ldots),$$

the formula (4.32) implies

$$\sum_{j=0}^{i}(i - j + 1)\,\Theta(m, \beta_1 - j) = \max\{p(X) - (\beta_1 - i - 1)|X| : X \subseteq S\} \qquad (i = 0, 1, 2, \ldots). \tag{4.33}$$

This formula gives a recurrence formula for $\Theta(m,\beta_1), \Theta(m,\beta_1-1), \Theta(m,\beta_1-2), \ldots$ as

$$\begin{aligned}\Theta(m,\beta_1) &= \max\{p(X)-(\beta_1-1)|X| : X \subseteq S\},\\ \Theta(m,\beta_1-1) &= \max\{p(X)-(\beta_1-2)|X| : X \subseteq S\} - 2\,\Theta(m,\beta_1),\\ \Theta(m,\beta_1-2) &= \max\{p(X)-(\beta_1-3)|X| : X \subseteq S\} - 3\,\Theta(m,\beta_1) - 2\,\Theta(m,\beta_1-1),\\ &\cdots\cdots\cdots\cdots\cdots\cdots\end{aligned} \tag{4.34}$$

Remark 4.7. A DCA-based proof of the formula (4.32) is as follows. In (4.22) we choose

$$\varphi(k) = (k-a)^+ = \begin{cases} 0 & (k \leq a), \\ k-a & (k \geq a+1). \end{cases}$$

The left-hand side of (4.22) coincides with that of (4.32). The conjugate function ψ is given by

$$\psi(\ell) = \begin{cases} 0 & (\ell = 0), \\ a & (\ell = 1), \\ +\infty & (\ell \notin \{0,1\}). \end{cases}$$

Therefore, we may restrict π to $\{0,1\}$-vectors in considering the maximum of $g(\pi) := \hat{p}(\pi) - \sum_{s\in S} \psi(\pi(s))$ over $\pi \in \mathbf{Z}^S$. For $\pi = \chi_X \in \{0,1\}^S$ with $X \subseteq S$, we have $\hat{p}(\pi) = \hat{p}(\chi_X) = p(X)$ and $\sum_{s\in S} \psi(\pi(s)) = \sum_{s\in S} \psi(\chi_X(s)) = \sum_{s\in X} \psi(1) = a|X|$, and therefore, the right-hand side of (4.22) is equal to $\max\{p(X) - a|X| : X \subseteq S\}$. Thus the formula (4.32) is derived. ■

5 Structure of optimal solutions to square-sum minimization

In this section we offer the DCA view on the structure of optimal solutions of the min-max formula:

$$\min\{\sum[m(s)^2 : s \in S] : m \in \ddot{B}\} = \max\{\hat{p}(\pi) - \sum_{s\in S} \left\lfloor \frac{\pi(s)}{2} \right\rfloor \left\lceil \frac{\pi(s)}{2} \right\rceil : \pi \in \mathbf{Z}^S\}, \tag{5.1}$$

to which a DCA-based proof has been given in Section 4.4.

Concerning the optimal solutions to (5.1) the following results were obtained in Part I [9]. Recall that $\beta_1 > \beta_2 > \cdots > \beta_q$ denotes the essential value-sequence, $C_1 \subset C_2 \subset \cdots \subset C_q$ is the canonical chain, $\{S_1, S_2, \ldots, S_q\}$ is the canonical partition ($S_i = C_i - C_{i-1}$ and $C_0 = \emptyset$), π^* and Δ^* are integral vectors defined by

$$\pi^*(s) = 2\beta_i - 1, \quad \Delta^*(s) = \beta_i - 1 \qquad (s \in S_i;\ i = 1,2,\ldots,q),$$

and M^* denotes the direct sum of matroids $M_1, M_2, \ldots, M_q$ constructed in Section 5.3 of Part I [9].

Proposition 5.1 ([9, Corollary 6.15])**.** *The set* Π *of dual optimal integral vectors* π *in* (5.1) *is an* $\mathrm{L}^\natural$*-convex set. The unique smallest element of* Π *is* π^**.* ■

Theorem 5.2 ([9, Theorem 6.13]). *An integral vector π is a dual optimal solution in* (5.1) *if and only if the following three conditions hold for each $i = 1, 2, \ldots, q$:*

$$\pi(s) = 2\beta_i - 1 \ \text{ for every } s \in S_i - F_i, \tag{5.2}$$

$$2\beta_i - 1 \le \pi(s) \le 2\beta_i + 1 \ \text{ for every } s \in F_i, \tag{5.3}$$

$$\pi(s) - \pi(t) \ge 0 \ \text{ whenever } s, t \in F_i \text{ and } (s, t) \in A_i, \tag{5.4}$$

where F_i is the largest member of $\mathcal{F}_i = \{X \subseteq S_i : \beta_i|X| = p(C_{i-1} \cup X) - p(C_{i-1})\}$ and A_i is the set of pairs (s, t) such that $s, t \in F_i$ and there is no set in $\mathcal{F}_i$ which contains t and not s. ■

Theorem 5.3 ([9, Theorem 5.7]). *The set of dec-min elements of $\ddot{\overline{B}}$ is a matroidal M-convex set.*[2] *More precisely, an element m of $\ddot{\overline{B}}$ is decreasingly minimal if and only if m can be obtained in the form $m = \chi_L + \Delta^*$, where L is a basis of the matroid M^*.* ■

The objective of this section is to shed the light of DCA on these results. It will turn out that the general results in DCA capture the structural essence of the above statements, but do not provide the full statements with specific details. We first present a summary of the relevant results from DCA in Sections 5.1 and 5.2.

5.1 General results on the optimal solutions in the Fenchel-type discrete duality

We summarize the fundamental facts about the optimal solutions in the Fenchel-type min-max relation

$$\min\{f(x) - h(x) : x \in \mathbf{Z}^S\} = \max\{h^\circ(\pi) - f^\bullet(\pi) : \pi \in \mathbf{Z}^S\}, \tag{5.5}$$

where f is an integer-valued M$^\natural$-convex function and h is an integer-valued M$^\natural$-concave function. We assume that both $\operatorname{dom} f \cap \operatorname{dom} h$ and $\operatorname{dom} f^\bullet \cap \operatorname{dom} h^\circ$ are nonempty, in which case the common value in (5.5) is finite. We denote the set of the minimizers by $\mathcal{P}$ and the set of the maximizers by $\mathcal{D}$.

To derive the optimality criteria we recall the Fenchel–Young inequalities

$$f(x) + f^\bullet(\pi) \ge \langle \pi, x \rangle, \tag{5.6}$$

$$h(x) + h^\circ(\pi) \le \langle \pi, x \rangle, \tag{5.7}$$

which hold for any $x \in \mathbf{Z}^S$ and $\pi \in \mathbf{Z}^S$. These inequalities immediately imply the weak duality

$$f(x) - h(x) \ge h^\circ(\pi) - f^\bullet(\pi). \tag{5.8}$$

The inequality in (5.8) turns into an equality if and only if the inequalities in (5.6) and (5.7) are satisfied in equalities. The former condition is equivalent to saying that $x \in \mathcal{P}$ and $\pi \in \mathcal{D}$. The equality condition for (5.6) can be rewritten as

$$f(x) - \langle \pi, x \rangle = -f^\bullet(\pi) = -\max\{\langle \pi, y \rangle - f(y) : y \in \mathbf{Z}^S\} = \min\{f(y) - \langle \pi, y \rangle : y \in \mathbf{Z}^S\}. \tag{5.9}$$

[2] In Part I, we have defined a **matroidal M-convex set** as the set of integral elements of a translated matroid base-polyhedron. In other words, a matroidal M-convex set is an M-convex set in which the ℓ_∞-distance of any two distinct members is equal to one.

Similarly, the equality condition for (5.7) can be rewritten as

$$h(x) - \langle \pi, x \rangle = -h^{\circ}(\pi) = -\min\{\langle \pi, y \rangle - h(y) : y \in \mathbf{Z}^S\} = \max\{h(y) - \langle \pi, y \rangle : y \in \mathbf{Z}^S\}. \tag{5.10}$$

Therefore we have

$$x \in \mathcal{P} \text{ and } \pi \in \mathcal{D} \iff x \in \arg\min_y \{f(y) - \langle \pi, y \rangle\} \cap \arg\max_y \{h(y) - \langle \pi, y \rangle\}. \tag{5.11}$$

Furthermore, by the M-optimality criterion (Theorem 3.5) applied to $f(y) - \langle \pi, y \rangle$, we have $x \in \arg\min\{f(y) - \langle \pi, y \rangle\}$ if and only if

$$\begin{aligned} f(x) - \langle \pi, x \rangle &\leq f(x + \chi_s - \chi_t) - \langle \pi, x + \chi_s - \chi_t \rangle \qquad (\forall s, t \in S), \\ f(x) - \langle \pi, x \rangle &\leq f(x + \chi_s) - \langle \pi, x + \chi_s \rangle \qquad (\forall s \in S), \\ f(x) - \langle \pi, x \rangle &\leq f(x - \chi_t) - \langle \pi, x - \chi_t \rangle \qquad (\forall t \in S), \end{aligned}$$

that is, if and only if

$$\pi(s) - \pi(t) \leq f(x + \chi_s - \chi_t) - f(x) \qquad (\forall s, t \in S), \tag{5.12}$$

$$f(x) - f(x - \chi_s) \leq \pi(s) \leq f(x + \chi_s) - f(x) \qquad (\forall s \in S). \tag{5.13}$$

Similarly, we have $x \in \arg\max\{h(y) - \langle \pi, y \rangle\}$ if and only if

$$\pi(s) - \pi(t) \geq h(x + \chi_s - \chi_t) - h(x) \qquad (\forall s, t \in S), \tag{5.14}$$

$$h(x) - h(x - \chi_s) \geq \pi(s) \geq h(x + \chi_s) - h(x) \qquad (\forall s \in S). \tag{5.15}$$

Therefore,

$$x \in \mathcal{P} \text{ and } \pi \in \mathcal{D} \iff \text{(5.12), (5.13), (5.14), (5.15) hold.} \tag{5.16}$$

Using the integer biconjugacy $f^{\bullet\bullet} = f$ and $h^{\circ\circ} = h$ for M$^{\natural}$-convex/concave functions with respect to the discrete conjugates in (4.5) and (4.6) (cf. [35, Theorem 8.12]), we can rewrite (5.9) and (5.10), respectively, as

$$\begin{aligned} f^{\bullet}(\pi) - \langle \pi, x \rangle &= -f(x) = -f^{\bullet\bullet}(x) = \min\{f^{\bullet}(\tau) - \langle \tau, x \rangle : \tau \in \mathbf{Z}^S\}, \\ h^{\circ}(\pi) - \langle \pi, x \rangle &= -h(x) = -h^{\circ\circ}(x) = \max\{h^{\circ}(\tau) - \langle \tau, x \rangle : \tau \in \mathbf{Z}^S\}. \end{aligned}$$

Hence the equivalence in (5.11) can be rephrased in terms of the conjugate functions as

$$x \in \mathcal{P} \text{ and } \pi \in \mathcal{D} \iff \pi \in \arg\min_\tau \{f^{\bullet}(\tau) - \langle \tau, x \rangle\} \cap \arg\max_\tau \{h^{\circ}(\tau) - \langle \tau, x \rangle\}. \tag{5.17}$$

Furthermore, by the L-optimality criterion (Theorem 4.2) applied to the L$^{\natural}$-convex function $f^{\bullet}(\tau) - \langle \tau, x \rangle$, we have $\pi \in \arg\min\{f^{\bullet}(\tau) - \langle \tau, x \rangle\}$ if and only if

$$\begin{aligned} f^{\bullet}(\pi) - \langle \pi, x \rangle &\leq f^{\bullet}(\pi + \chi_Y) - \langle \pi + \chi_Y, x \rangle \qquad (\forall Y \subseteq S), \\ f^{\bullet}(\pi) - \langle \pi, x \rangle &\leq f^{\bullet}(\pi - \chi_Y) - \langle \pi - \chi_Y, x \rangle \qquad (\forall Y \subseteq S), \end{aligned}$$

that is, if and only if

$$f^{\bullet}(\pi) - f^{\bullet}(\pi - \chi_Y) \leq \sum_{s \in Y} x(s) \leq f^{\bullet}(\pi + \chi_Y) - f^{\bullet}(\pi) \qquad (\forall Y \subseteq S). \tag{5.18}$$

Similarly, we have $\pi \in \arg\max\{h^{\circ}(\tau) - \langle \tau, x \rangle\}$ if and only if

$$h^{\circ}(\pi) - h^{\circ}(\pi - \chi_Y) \geq \sum_{s \in Y} x(s) \geq h^{\circ}(\pi + \chi_Y) - h^{\circ}(\pi) \qquad (\forall Y \subseteq S). \tag{5.19}$$

Therefore,

$$x \in \mathcal{P} \text{ and } \pi \in \mathcal{D} \iff \text{(5.18), (5.19) hold.} \tag{5.20}$$

From the above argument we can obtain the following optimality criteria.

Theorem 5.4. *Let f be an integer-valued M$^{\natural}$-convex function and h be an integer-valued M$^{\natural}$-concave function such that both $\mathcal{P}_0 := \mathrm{dom}\, f \cap \mathrm{dom}\, h$ and $\mathcal{D}_0 := \mathrm{dom}\, f^{\bullet} \cap \mathrm{dom}\, h^{\circ}$ are nonempty.*

(1) *Let $x \in \mathcal{P}_0$ and $\pi \in \mathcal{D}_0$. Then the following three conditions are pairwise equivalent.*

(a) *x and π are both optimal, that is, $x \in \mathcal{P}$ and $\pi \in \mathcal{D}$.*

(b) *The inequalities* (5.12), (5.13), (5.14), *and* (5.15) *are satisfied by x and π.*

(c) *The inequalities* (5.18) *and* (5.19) *are satisfied by x and π.*

(2) *Let $\hat{\pi} \in \mathcal{D}$ be an arbitrary dual optimal solution. Then $x^* \in \mathcal{P}_0$ is a minimizer of $f(x) - h(x)$ if and only if it is a minimizer of $f(x) - \langle \hat{\pi}, x \rangle$ and simultaneously a maximizer of $h(x) - \langle \hat{\pi}, x \rangle$, or equivalently, x^* satisfies* (5.18) *and* (5.19) *for $\pi = \hat{\pi}$. Namely,*

$$\mathcal{P} = \arg\min\{f(x) - \langle \hat{\pi}, x \rangle\} \cap \arg\max\{h(x) - \langle \hat{\pi}, x \rangle\} \tag{5.21}$$

$$= \{x \in \mathbf{Z}^S : \text{(5.12), (5.13), (5.14), (5.15) hold with } \pi = \hat{\pi}\} \tag{5.22}$$

$$= \{x \in \mathbf{Z}^S : \text{(5.18) and (5.19) hold with } \pi = \hat{\pi}\}. \tag{5.23}$$

(3) *Let $\hat{x} \in \mathcal{P}$ be an arbitrary primal optimal solution. Then $\pi^* \in \mathcal{D}_0$ is a maximizer of $h^{\circ}(\pi) - f^{\bullet}(\pi)$ if and only if it is a minimizer of $f^{\bullet}(\pi) - \langle \pi, \hat{x} \rangle$ and simultaneously a maximizer of $h^{\circ}(\pi) - \langle \pi, \hat{x} \rangle$, or equivalently, π^* satisfies the inequalities* (5.12), (5.13), (5.14), *and* (5.15) *for $x = \hat{x}$. Namely,*

$$\mathcal{D} = \arg\min\{f^{\bullet}(\pi) - \langle \pi, \hat{x} \rangle\} \cap \arg\max\{h^{\circ}(\pi) - \langle \pi, \hat{x} \rangle\} \tag{5.24}$$

$$= \{\pi \in \mathbf{Z}^S : \text{(5.12), (5.13), (5.14), (5.15) hold with } x = \hat{x}\} \tag{5.25}$$

$$= \{\pi \in \mathbf{Z}^S : \text{(5.18) and (5.19) hold with } x = \hat{x}\}. \tag{5.26}$$

■

It is emphasized that in the representation of $\mathcal{P}$, each of $\arg\min\{f(x) - \langle \hat{\pi}, x \rangle\}$ and $\arg\max\{h(x) - \langle \hat{\pi}, x \rangle\}$ depends on the choice of $\hat{\pi}$, but their intersection is uniquely determined and equal to $\mathcal{P}$. Similarly, in the representation of $\mathcal{D}$, each of $\arg\min\{f^{\bullet}(\pi) - \langle \pi, \hat{x} \rangle\}$ and $\arg\max\{h^{\circ}(\pi) - \langle \pi, \hat{x} \rangle\}$ depends on the choice of $\hat{x}$, but their intersection is uniquely determined and equal to $\mathcal{D}$.

The representation of $\mathcal{P}$ in (5.21) (or (5.23)) shows that $\mathcal{P}$ is the intersection of two M$^{\natural}$-convex sets. Such a set is called an **M$_2^{\natural}$-convex set** [35, Section 4.7]. Note that the intersection of M$^{\natural}$-convex sets is not always M$^{\natural}$-convex. The representation of $\mathcal{D}$ in (5.24) (or (5.25)) shows that $\mathcal{D}$ is the intersection of two L$^{\natural}$-convex sets. Since the intersection of two (or more) L$^{\natural}$-convex sets is again L$^{\natural}$-convex, $\mathcal{D}$ is an L$^{\natural}$-convex set.

Proposition 5.5. *In the Fenchel-type min-max relation* (5.5) *for $M^{\natural}$-convex/concave functions, the set $\mathcal{P}$ of the minimizers is an $M_2^{\natural}$-convex set and the set $\mathcal{D}$ of the maximizers is an $L^{\natural}$-convex set.* ■

Remark 5.1. In Remark 4.2 we have discussed the role of the Fenchel-type discrete duality theorem for the certificate of optimality in minimizing $f(x) - h(x)$. We have distinguished two cases according to wheter the explicit forms of the conjugate functions $f^{\bullet}(\pi)$ and $h^{\circ}(\pi)$ are available or not. If their explicit forms are known, we can verify the optimality of x by simply computing the values of $f(x)-h(x)$ for x and $h^{\circ}(\pi)-f^{\bullet}(\pi)$ for a given dual optimal π. Even if the explicit forms of the conjugate functions are not known, Theorem 5.4 (2) above enables us to verify the optimality of x by checking the inequalities (5.12), (5.13), (5.14), and (5.15) for a given dual optimal π. Note that we have $O(|S|^2)$ inequalities in total. We emphasize that Theorem 5.4 (2) is derived from a combination of the Fenchel-type discrete duality theorem (Theorem 4.1) with the M-optimality criterion (Theorem 3.5). ■

Remark 5.2. In convex analysis, as well as in discrete convex analysis, the optimality conditions such as those in Theorem 5.4 are expressed usually in terms of subgradients and subdifferentials. In this paper, however, we have intentionally avoided using these concepts for the sake of the audience from combinatorial optimization. In this remark we will briefly indicate how the results in Theorem 5.4 can be described and interpreted in terms of subgradients and subdifferentials.

Let $f : \mathbf{Z}^S \to \mathbf{Z} \cup \{+\infty\}$ and $h : \mathbf{Z}^S \to \mathbf{Z} \cup \{-\infty\}$ be integer-valued functions defined on $\mathbf{Z}^S$. The integral subdifferential of f at $x \in \mathrm{dom}\, f$ and its concave version for h at $x \in \mathrm{dom}\, h$ are the sets of integer vectors defined as

$$\partial f(x) := \{\pi \in \mathbf{Z}^S : f(y) - f(x) \geq \langle \pi, y - x \rangle \ (\forall y \in \mathbf{Z}^S)\}, \tag{5.27}$$

$$\partial h(x) := \{\pi \in \mathbf{Z}^S : h(y) - h(x) \leq \langle \pi, y - x \rangle \ (\forall y \in \mathbf{Z}^S)\}. \tag{5.28}$$

A member of $\partial f(x)$ is called a subgradient of f at x. Accordingly, the integral subdifferentials of $f^{\bullet}$ and h° at π are defined as

$$\partial f^{\bullet}(\pi) := \{x \in \mathbf{Z}^S : f^{\bullet}(\tau) - f^{\bullet}(\pi) \geq \langle \tau - \pi, x \rangle \ (\forall \tau \in \mathbf{Z}^S)\}, \tag{5.29}$$

$$\partial h^{\circ}(\pi) := \{x \in \mathbf{Z}^S : h^{\circ}(\tau) - h^{\circ}(\pi) \leq \langle \tau - \pi, x \rangle \ (\forall \tau \in \mathbf{Z}^S)\}, \tag{5.30}$$

where $\partial f^{\bullet}(\pi)$ is defined for $\pi \in \mathrm{dom}\, f^{\bullet}$ and $\partial h^{\circ}(\pi)$ for $\pi \in \mathrm{dom}\, h^{\circ}$. The following relations are straightforward translations of the corresponding results in (ordinary) convex analysis to the discrete setting (cf., [33], [35]):

$$\pi \in \partial f(x) \iff \text{equality holds in (5.6)} \iff x \in \partial f^{\bullet}(\pi), \tag{5.31}$$

$$\pi \in \partial h(x) \iff \text{equality holds in (5.7)} \iff x \in \partial h^{\circ}(\pi), \tag{5.32}$$

$$\partial f(x) = \arg\min_{\pi}\{f^{\bullet}(\pi) - \langle \pi, x \rangle\}, \tag{5.33}$$

$$\partial h(x) = \arg\max_{\pi}\{h^{\circ}(\pi) - \langle \pi, x \rangle\}, \tag{5.34}$$

$$\partial f^{\bullet}(\pi) = \arg\min_{x}\{f(x) - \langle \pi, x \rangle\}, \tag{5.35}$$

$$\partial h^{\circ}(\pi) = \arg\max_{x}\{h(x) - \langle \pi, x \rangle\}, \tag{5.36}$$

where the integer biconjugacy ($f^{\bullet\bullet} = f$, $h^{\circ\circ} = h$) is assumed, which is true for M$^\natural$-convex/concave functions. By using (5.35)–(5.36) in (5.21), and (5.33)–(5.34) in (5.24), respectively, we obtain the following representations of optimal solutions

$$\mathcal{P} = \partial f^{\bullet}(\hat{\pi}) \cap \partial h^{\circ}(\hat{\pi}), \tag{5.37}$$

$$\mathcal{D} = \partial f(\hat{x}) \cap \partial h(\hat{x}) \tag{5.38}$$

for any $\hat{\pi} \in \mathcal{D}$ and $\hat{x} \in \mathcal{P}$. We also have optimality criteria

$$x \in \mathcal{P} \iff \partial f(x) \cap \partial h(x) \neq \emptyset, \tag{5.39}$$

$$\pi \in \mathcal{D} \iff \partial f^{\bullet}(\pi) \cap \partial h^{\circ}(\pi) \neq \emptyset. \tag{5.40}$$

Finally it is worth mentioning that, by the M-L conjugacy [35, Chapter 8], the subdifferential of an M$^\natural$-convex function f (resp., an M$^\natural$-concave function h) is an L$^\natural$-convex set and the subdifferential of an L$^\natural$-convex function $f^{\bullet}$ (resp., an L$^\natural$-concave function h°) is an M$^\natural$-convex set. ■

5.2 Separable convex functions on an M-convex set

In Theorem 4.3 we have shown a min-max formula

$$\min\{\sum_{s \in S} \varphi_s(x(s)) : x \in \overset{....}{B}\} = \max\{\hat{p}(\pi) - \sum_{s \in S} \psi_s(\pi(s)) : \pi \in \mathbf{Z}^S\} \tag{5.41}$$

for an integer-valued separable convex function

$$\Phi(x) = \sum [\varphi_s(x(s)) : s \in S] \tag{5.42}$$

on an M-convex set $\overset{....}{B}$. Here we introduce notations for the set of feasible points:

$$\mathrm{dom}\, \Phi = \{x \in \mathbf{Z}^S : x(s) \in \mathrm{dom}\, \varphi_s \text{ for each } s \in S\}, \tag{5.43}$$

$$\mathcal{P}_0 = \overset{....}{B} \cap \mathrm{dom}\, \Phi = \{x \in \overset{....}{B} : x(s) \in \mathrm{dom}\, \varphi_s \text{ for each } s \in S\}, \tag{5.44}$$

$$\mathcal{D}_0 = \{\pi \in \mathbf{Z}^S : \pi \in \mathrm{dom}\, \hat{p},\ \pi(s) \in \mathrm{dom}\, \psi_s \text{ for each } s \in S\}. \tag{5.45}$$

The min-max formula (5.41) holds under the assumption of primal feasibility ($\mathcal{P}_0 \neq \emptyset$) or dual feasibility ($\mathcal{D}_0 \neq \emptyset$). The unbounded case with both sides of (5.41) being equal to $-\infty$ or $+\infty$ is also a possibility in general, but in this section we assume that the both sides are finite-valued and denote the set of the minimizers x by $\mathcal{P}$ and the set of the maximizers π by $\mathcal{D}$.

We can obtain the optimality conditions for (5.41) by applying Theorem 5.4 with

$$f(x) = \sum [\varphi_s(x(s)) : s \in S], \qquad h(x) = -\delta(x),$$

$$f^{\bullet}(\pi) = \sum [\psi_s(\pi(s)) : s \in S], \qquad h^{\circ}(\pi) = \hat{p}(\pi),$$

where δ is the indicator function of $\overset{....}{B}$ defined in (3.14). However, we present a direct derivation from (5.41) via weak duality (min ≥ max), as it should be more informative and convenient for readers.

For each conjugate pair (φ_s, ψ_s), it follows from the definition (4.16) that

$$\varphi_s(k) + \psi_s(\ell) \geq k\ell \qquad (k, \ell \in \mathbf{Z}), \tag{5.46}$$

which is known as the Fenchel–Young inequality, where the equality holds if and only if

$$\varphi_s(k) - \varphi_s(k-1) \leq \ell \leq \varphi_s(k+1) - \varphi_s(k). \tag{5.47}$$

Let $x \in \mathcal{P}_0$ and $\pi \in \mathcal{D}_0$. Then, using the Fenchel–Young inequality (5.46) as well as (4.19) for p, we obtain the weak duality:

$$\begin{aligned}
\sum_{s \in S} \varphi_s(x(s)) - \left(\hat{p}(\pi) - \sum_{s \in S} \psi_s(\pi(s)) \right) &= \sum_{s \in S} [\varphi_s(x(s)) + \psi_s(\pi(s))] \; - \hat{p}(\pi) \\
&\geq \sum_{s \in S} x(s)\pi(s) \; - \hat{p}(\pi) && (5.48)\\
&\geq \min\{\pi z : z \in \ddddot{B}\} - \hat{p}(\pi) \; = 0. && (5.49)
\end{aligned}$$

The optimality conditions can be obtained as the conditions for the inequalities in (5.48) and (5.49) to be equalities, as follows.

Proposition 5.6. *Assume that both $\mathcal{P}_0$ and $\mathcal{D}_0$ in* (5.44)–(5.45) *are nonempty.*

(1) *Let $x \in \mathcal{P}_0$ and $\pi \in \mathcal{D}_0$. Then $x \in \mathcal{P}$ and $\pi \in \mathcal{D}$ (that is, x and π are both optimal) if and only if the following two conditions are satisfied:*

$$\varphi_s(x(s)) - \varphi_s(x(s) - 1) \leq \pi(s) \leq \varphi_s(x(s) + 1) - \varphi_s(x(s)) \qquad (s \in S), \tag{5.50}$$

$$\pi(s) \geq \pi(t) \quad \text{for every } (s, t) \text{ with } x + \chi_s - \chi_t \in \ddddot{B}. \tag{5.51}$$

(2) *Let $\hat{\pi} \in \mathcal{D}$ be an arbitrary dual optimal solution. Then $x^* \in \mathcal{P}_0$ is a minimizer of $\Phi(x)$ over $\ddddot{B}$ if and only if it satisfies* (5.50) *and* (5.51) *for $\pi = \hat{\pi}$, or equivalently, it is a minimizer of $\sum[\varphi_s(x(s)) - \hat{\pi}(s)x(s) : s \in S]$ and simultaneously a $\hat{\pi}$-minimizer in $\ddddot{B}$. Namely,*

$$\begin{aligned}
\mathcal{P} &= \{x \in \mathcal{P}_0 : (5.50), (5.51) \text{ hold with } \pi = \hat{\pi}\} && (5.52)\\
&= \{x \in \mathrm{dom}\,\Phi : (5.50) \text{ holds with } \pi = \hat{\pi}\} \cap \{x \in \ddddot{B} : x \text{ is a } \hat{\pi}\text{-minimizer in } \ddddot{B}\}. && (5.53)
\end{aligned}$$

(3) *Let $\hat{x} \in \mathcal{P}$ be an arbitrary primal optimal solution. Then $\pi^* \in \mathcal{D}_0$ is a maximizer of $\hat{p}(\pi) - \sum_{s \in S} \psi_s(\pi(s))$ if and only if it satisfies the inequalities* (5.50) *and* (5.51) *for $x = \hat{x}$. Namely,*

$$\mathcal{D} = \{\pi \in \mathcal{D}_0 : (5.50), (5.51) \text{ hold with } x = \hat{x}\}. \tag{5.54}$$

Proof. The inequality (5.48) turns into an equality if and only if, for each $s \in S$, we have $\varphi_s(k) + \psi_s(\ell) = k\ell$ for $k = x(s)$ and $\ell = \pi(s)$. The latter condition is equivalent to (5.50) by (5.47). The other inequality (5.49) turns into an equality if and only if x is a π-minimizer in $\ddddot{B}$, which is equivalent to (5.51). Finally, we see from (5.41) that the two inequalities in (5.48) and (5.49) simultaneously turn into equality if $x \in \mathcal{P}$ and $\pi \in \mathcal{D}$. ■

Proposition 5.7. *In the min-max relation* (5.41) *for a separable convex function on an M-convex set, the set $\mathcal{D}$ of the maximizers is an $L^\natural$-convex set and the set $\mathcal{P}$ of the minimizers is an M-convex set.*

Proof. The representation (5.54) shows that $\mathcal{D}$ is described by the inequalities in (5.50) and (5.51). Hence $\mathcal{D}$ is L$^\natural$-convex. (The L$^\natural$-convexity of $\mathcal{D}$ can also be obtained from Proposition 5.5.) In the representation (5.53) of $\mathcal{P}$, the first set $\{x \in \mathrm{dom}\,\Phi : (5.50) \text{ holds with } \pi = \hat{\pi}\}$ is a box, while the set of $\hat{\pi}$-minimizers in $\ddot{B}$ is an M-convex set. Therefore, $\mathcal{P}$ is an M-convex set. ■

5.3 Dual optimal solutions to square-sum minimization

The min-max formula (5.1) for the square-sum minimization is a special case of the min-max formula (5.41) with $\varphi_s(k) = \varphi(k) = k^2$ and $\psi_s(\ell) = \psi(\ell) = \lfloor \ell/2 \rfloor \cdot \lceil \ell/2 \rceil$ for $k, \ell \in \mathbf{Z}$ (cf., (4.23)). Accordingly, we can apply the general results (Proposition 5.6, in particular) for the analysis of the optimal solutions in the min-max formula (5.1). In this section we consider the dual solutions, whereas the primal solutions are treated in Section 5.4.

The function $g(\pi) = \hat{p}(\pi) - \sum[\psi(\pi(s)) : s \in S]$ to be maximized in (5.1) is L$^\natural$-concave by Proposition 4.4, and the maximizers of an L$^\natural$-concave function form an L$^\natural$-convex set [35, Theorem 7.17]. Therefore, the set Π of dual optimal solutions is an L$^\natural$-convex set, which is the first statement of Proposition 5.1. The L$^\natural$-convexity of Π implies that there exists a unique smallest element of Π. The second statement of Proposition 5.1 shows that this smallest element is given by π^*, but this fact is not easily shown by general arguments from discrete convex analysis.

Next we consider Theorem 5.2, which gives a representation of Π. According to the general result stated in Proposition 5.6 (3), we can obtain another representation of Π by choosing any dec-min element $\hat{x}$ of $\ddot{B}$, which is a primal optimal solution for (5.1). In this case the condition (5.50) reads

$$2x(s) - 1 \leq \pi(s) \leq 2x(s) + 1 \qquad (s \in S), \tag{5.55}$$

since $\varphi(k) - \varphi(k-1) = k^2 - (k-1)^2 = 2k - 1$ and $\varphi(k+1) - \varphi(k) = (k+1)^2 - k^2 = 2k + 1$.

Proposition 5.8. *Let m be any dec-min element of $\ddot{B}$. The set Π of dual optimal solutions to* (5.1) *is represented as $\Pi = I(m) \cap P(m)$, where*

$$I(m) = \{\pi \in \mathbf{Z}^S : 2m(s) - 1 \leq \pi(s) \leq 2m(s) + 1 \text{ for all } s \in S\},$$
$$P(m) = \{\pi \in \mathbf{Z}^S : \pi(s) \geq \pi(t) \ \ \text{for every } (s,t) \ \text{ with } \ x + \chi_s - \chi_t \in \ddot{B}\}.$$

Hence Π is an L$^\natural$*-convex set.* ■

Let us compare the representations of Π in Proposition 5.8 and Theorem 5.2. Roughly speaking, $I(m)$ corresponds to the first two conditions (5.2) and (5.3) in Theorem 5.2 and $P(m)$ to the third condition (5.4). However, there is an essential difference between Proposition 5.8 and Theorem 5.2. Namely, each of $I(m)$ and $P(m)$ varies with the choice of m, while their intersection is uniquely determined and equal to Π. In this sense, the description of Π in Proposition 5.8 is not canonical. Theorem 5.2 is a much stronger statement, giving a canonical description of Π without reference to a particular primal optimal solution.

Remark 5.3. Proposition 5.8 above is equivalent to Proposition 6.11 of Part I [9], though in a slightly different form. Recall the optimality criteria there:[3]

$$\begin{array}{ll} \text{(O1)} & m(s) \in \{\lfloor \pi(s)/2 \rfloor, \lceil \pi(s)/2 \rceil\} \text{ for each } \ s \in S, \\ \text{(O2)} & \text{each strict } \pi\text{-top-set is } m\text{-tight with respect to } p. \end{array}$$

The set $I(m)$ corresponds to the first optimality criterion (O1), since $2m(s) - 1 \leq \pi(s) \leq 2m(s) + 1$ if and only if $m(s) \in \{\lfloor \pi(s)/2 \rfloor, \lceil \pi(s)/2 \rceil\}$. The equivalence of $P(m)$ to the second criterion (O2) is a well-known characterization of a minimum weight base. ■

5.4 Primal optimal solutions to square-sum minimization

We now turn to the primal problem of (5.1), namely, the square-sum minimization.

Let $\mathrm{dm}(\ddddot{B})$ denote the set of the dec-min elements of $\ddddot{B}$. By Theorem 3.2, $\mathrm{dm}(\ddddot{B})$ coincides with the set of primal optimal solutions for (5.1). According to the general result in Proposition 5.6 (2), a representation of $\mathrm{dm}(\ddddot{B})$ can be obtained by choosing any dual optimal solution $\hat{\pi}$. In this case the condition (5.50) is simplified to (5.55), which can be rewritten as

$$x(s) \in \{\lfloor \pi(s)/2 \rfloor, \lceil \pi(s)/2 \rceil\} \qquad (s \in S). \tag{5.56}$$

Thus the following representation of the set of dec-min elements is obtained.

Proposition 5.9. *Let $\hat{\pi}$ be any dual optimal solution to* (5.1). *The set* $\mathrm{dm}(\ddddot{B})$ *of dec-min elements of $\ddddot{B}$ is represented as* $\mathrm{dm}(\ddddot{B}) = T(\hat{\pi}) \cap \ddddot{B}^{\circ}(\hat{\pi})$*, where*

$$\begin{aligned} T(\hat{\pi}) &= \{m \in \mathbf{Z}^S : m(s) \in \{\lfloor \hat{\pi}(s)/2 \rfloor, \lceil \hat{\pi}(s)/2 \rceil\} \ (s \in S)\}, \\ \ddddot{B}^{\circ}(\hat{\pi}) &= \{m \in \ddddot{B} : m \text{ is a minimum } \hat{\pi}\text{-weight element of } \ddddot{B}\}. \end{aligned}$$

Hence $\mathrm{dm}(\ddddot{B})$ *is a matroidal M-convex set.* ■

Again, each of $T(\hat{\pi})$ and $\ddddot{B}^{\circ}(\hat{\pi})$ varies with the choice of $\hat{\pi}$, but their intersection is uniquely determined and is equal to $\mathrm{dm}(\ddddot{B})$. Here, $\ddddot{B}^{\circ}(\hat{\pi})$ is the integral elements of a face of B, and is an M-convex set. As for $T(\hat{\pi})$, note that, for each $s \in S$, the two numbers $\lfloor \hat{\pi}(s)/2 \rfloor$ and $\lceil \hat{\pi}(s)/2 \rceil$ are the same integer or consecutive integers. Therefore, $\mathrm{dm}(\ddddot{B})$ is a matroidal M-convex set. In other words, there exist a matroid $\hat{M}$ and a translation vector $\hat{\Delta} \in \mathbf{Z}^S$ such that

$$\mathrm{dm}(\ddddot{B}) = T(\hat{\pi}) \cap \ddddot{B}^{\circ}(\hat{\pi}) = \{\chi_L + \hat{\Delta} : L \text{ is a basis of } \hat{M}\}.$$

In this construction both $\hat{M}$ and $\hat{\Delta}$ depend on the chosen $\hat{\pi}$; in particular, $\hat{\Delta} = \lfloor \hat{\pi}/2 \rfloor$.

Theorem 5.3 is significantly stronger than Proposition 5.9, in that it gives a concrete description of the matroid $\hat{M}$ by referring to the canonical chain. The translation vector Δ^* in Theorem 5.3 corresponds to the choice of $\hat{\pi} = \pi^*$; note that we indeed have the relation $\Delta^* = \lfloor \pi^*/2 \rfloor$.

[3] For a given vector π in $\mathbf{R}^S$, we call a non-empty set $X \subseteq S$ a **π-top set** if $\pi(u) \geq \pi(v)$ holds whenever $u \in X$ and $v \in S - X$. If $\pi(u) > \pi(v)$ holds whenever $u \in X$ and $v \in S - X$, we speak of a **strict π-top set**. We call a subset $X \subseteq S$ **m-tight** with respect to p if $\widetilde{m}(X) = p(X)$.

Remark 5.4. Proposition 5.9 implies, in particular, that the dec-min elements of an M-convex set is contained in a small (unit-sized) box. Note that such a property does not hold for an arbitrary integral polyhedron. To see this, consider the line segment P in $\mathbf{R}^3$ connecting two points $(2,1,0)$ and $(1,0,2)$. This P is an integral polyhedron, $\overset{....}{P} = \{(2,1,0),(1,0,2)\}$, and $\overset{....}{P}$ is not an M-convex set. Both $(2,1,0)$ and $(1,0,2)$ are dec-min in $\overset{....}{P}$, but there exists no small box containing them, since their third components differ by 2. In Part III prove that this small box property also holds for network flows. ■

6 Comparison of continuous and discrete cases

While our present study is focused on the discrete case for an M-convex set $\overset{....}{B}$, the continuous case for a base-polyhedron B was investigated by Fujishige [13] around 1980 under the name of lexicographically optimal bases, as a generalization of lexicographically optimal maximal flows considered by Megiddo [29]. Lexicographically optimal bases are discussed in detail in [14, Section 9]. Later in game theory Dutta–Ray [6] treated majorization ordering in the continuous case under the name of egalitarian allocation; see also Dutta [5]. See also the survey of related papers in Appendix A.

Section 6.1 offers comparisons of major ingredients in discrete and continuous cases. These comparisons show that the discrete case is significantly different from the continuous case, being endowed with a number of intriguing combinatorial structures on top of the geometric structures known in the continuous case. Section 6.2 is devoted to a review of the principal partition (adapted to a supermodular function), Section 6.3 gives an alternative characterization of the canonical partition, and Section 6.4 clarifies their relationship. Algorithmic implications are discussed in Section 6.5.

6.1 Summary of comparisons

The continuous case is referred to as Case **R** and the discrete case as Case **Z**. We use notation $m_{\mathbf{R}}$ and $m_{\mathbf{Z}}$ for the dec-min element in Case **R** and Case **Z**, respectively.

Underlying set In Case **R** we consider a base-polyhedron B described by a real-valued supermodular function p or a submodular function b. In Case **Z** we consider the set $\overset{....}{B}$ of integral members of an integral base-polyhedron B described by an integer-valued p or b.

Terminology In Case **R** the terminology of "lexicographically optimal base" (or "lexico-optimal base") is used in [13, 14]. A lexico-optimal base is the same as an inc-max element in our terminology, whereas a dec-min element is called a "co-lexicographically optimal base" in [14].

Weighting In Case **R** a weight vector is introduced to define and analyze lexico-optimality, while this is not the case in this paper for Case **Z**. In the following comparisons we always assume that no weighting is introduced in Cases **R** and **Z**. In a forthcoming paper [12], however, we consider discrete decreasing minimality with respect to a weight vector.

Decreasing minimality and increasing maximality In Case $\mathbf{Z}$ decreasing minimality in $\overset{....}{B}$ is equivalent to increasing maximality. This statement is also true in Case $\mathbf{R}$. That is, an element of B is dec-min in B if and only if it is inc-max in B. Moreover, a least majorized element exists in $\overset{....}{B}$ (in Case $\mathbf{Z}$) and in B (in Case $\mathbf{R}$).

Square-sum minimization In both Cases $\mathbf{Z}$ and $\mathbf{R}$, a dec-min element is characterized as a minimizer of square-sum of the components $W(x) = \sum[x(s)^2 : s \in S]$. In Case $\mathbf{R}$, the minimizer is unique, and is often referred to as the minimum norm point.

Uniqueness The structures of dec-min elements have a striking difference in Cases $\mathbf{R}$ and $\mathbf{Z}$. In Case $\mathbf{R}$ the dec-min element of B is uniquely determined, and is given by the minimum norm point of B. In Case $\mathbf{Z}$ the dec-min elements of $\overset{....}{B}$ are endowed with the structure of basis family of a matroid, as formulated in Theorem 5.3. The minimum norm point of B can be expressed as a convex combination of the dec-min elements of $\overset{....}{B}$ (cf., Theorem 6.7).

Proximity Every dec-min element $m_{\mathbf{Z}}$ of $\overset{....}{B}$ is located near the minimum norm point $m_{\mathbf{R}}$ of B, satisfying $\lfloor m_{\mathbf{R}} \rfloor \leq m_{\mathbf{Z}} \leq \lceil m_{\mathbf{R}} \rceil$ (cf., Theorem 6.6). However, not every integer vector $m_{\mathbf{Z}}$ in B satisfying $\lfloor m_{\mathbf{R}} \rfloor \leq m_{\mathbf{Z}} \leq \lceil m_{\mathbf{R}} \rceil$ is a dec-min element of $\overset{....}{B}$, which is demonstrated by the following example.

Example 6.1. Let $\overset{....}{B}$ be an M-convex set consisting of five vectors[4]

$$m_1 = (2,1,1,0), \quad m_2 = (2,1,0,1), \quad m_3 = (1,2,1,0), \quad m_4 = (1,2,0,1), \quad m_5 = (2,2,0,0)$$

and B be its convex hull. The dec-min elements of $\overset{....}{B}$ are m_1, m_2, m_3, and m_4, whereas $m_5 = (2,2,0,0)$ is not dec-min. The minimum norm point of the base-polyhedron B is $m_{\mathbf{R}} = (3/2, 3/2, 1/2, 1/2)$, for which $\lfloor m_{\mathbf{R}} \rfloor = (1,1,0,0)$ and $\lceil m_{\mathbf{R}} \rceil = (2,2,1,1)$. The point $m_5 = (2,2,0,0)$ satisfies $\lfloor m_{\mathbf{R}} \rfloor \leq m_5 \leq \lceil m_{\mathbf{R}} \rceil$ but it is not a dec-min element. ■

Min-max formula In Case $\mathbf{Z}$ we have the min-max identity (4.1):

$$\min\{\sum[m(s)^2 : s \in S] : m \in \overset{....}{B}\} = \max\{\hat{p}(\pi) - \sum_{s \in S} \left\lfloor \frac{\pi(s)}{2} \right\rfloor \left\lceil \frac{\pi(s)}{2} \right\rceil : \pi \in \mathbf{Z}^S\}.$$

In Case $\mathbf{R}$ the corresponding formula is

$$\min\{\sum[x(s)^2 : s \in S] : x \in B\} = \max\{\hat{p}(\pi) - \sum_{s \in S} \left(\frac{\pi(s)}{2}\right)^2 : \pi \in \mathbf{R}^S\}, \qquad (6.1)$$

which may be regarded as an adaptation of the standard quadratic programming duality to the case where the feasible region is a base-polyhedron. To the best knowledge of the authors, the formula (6.1) has never been shown in the literature.

4 $\overset{....}{B}$ is obtained from $\{(1,0,1,0), (1,0,0,1), (0,1,1,0), (0,1,0,1), (1,1,0,0)\}$ (basis family of rank 2 matroid) by a translation with $(1,1,0,0)$.

Principal partition vs canonical partition The canonical partition for Case **Z** is closely related to the principal partition for Case **R**. The principal partition (adapted to a supermodular function) is described in Section 6.2 and the following relations are established in Sections 6.3 and 6.4. We denote the canonical partition by $\{S_1, S_2, \ldots, S_q\}$ and the principal partition by $\{\hat{S}_1, \hat{S}_2, \ldots, \hat{S}_r\}$. They are constructed from the canonical chain $C_1 \subset C_2 \subset \cdots \subset C_q$ and the principal chain $\hat{C}_1 \subset \hat{C}_2 \subset \cdots \subset \hat{C}_r$, respectively, as the families of difference sets: $S_j = C_j - C_{j-1}$ for $j = 1, 2, \ldots, q$ and $\hat{S}_i = \hat{C}_i - \hat{C}_{i-1}$ for $i = 1, 2, \ldots, r$, where $C_0 = \hat{C}_0 = \emptyset$. We denote the essential values by $\beta_1 > \beta_2 > \cdots > \beta_q$ and the critical values by $\lambda_1 > \lambda_2 > \cdots > \lambda_r$.

- An integer β is an essential value for Case **Z** if and only if there exists a critical value λ for Case **R** satisfying $\beta \geq \lambda > \beta - 1$. The essential values $\beta_1 > \beta_2 > \cdots > \beta_q$ are obtained from the critical values $\lambda_1 > \lambda_2 > \cdots > \lambda_r$ as the distinct members of the rounded-up integers $\lceil \lambda_1 \rceil \geq \lceil \lambda_2 \rceil \geq \cdots \geq \lceil \lambda_r \rceil$.
- The canonical partition $\{S_1, S_2, \ldots, S_q\}$ is obtained as an aggregation from the principal partition $\{\hat{S}_1, \hat{S}_2, \ldots, \hat{S}_r\}$; we have $S_j = \bigcup_{i \in I(j)} \hat{S}_i$, where $I(j) = \{i : \lceil \lambda_i \rceil = \beta_j\}$.
- The canonical chain $\{C_j\}$ is a subchain of the principal chain $\{\hat{C}_i\}$; we have $C_j = \hat{C}_i$ for $i = \max I(j)$.
- In Case **R**, the dec-min element $m_{\mathbf{R}}$ of B is uniform on each member $\hat{S}_i$ of the principal partition, i.e., $m_{\mathbf{R}}(s) = \lambda_i$ if $s \in \hat{S}_i$, where $i = 1, 2, \ldots, r$ (cf., Proposition 6.2). In Case **Z**, the dec-min element $m_{\mathbf{Z}}$ of $\ddddot{B}$ is near-uniform on each member S_j of the canonical partition, i.e., $m_{\mathbf{Z}}(s) \in \{\beta_j, \beta_j - 1\}$ if $s \in S_j$, where $j = 1, 2, \ldots, q$ (cf., Theorem 5.1 of Part I [9]).

Algorithm In Case **Z** we have developed a strongly polynomial algorithm for finding a dec-min element of $\ddddot{B}$ (Section 7 of Part I [9]). In Case **R** the decomposition algorithm of Fujishige [13] finds the minimum norm point $m_{\mathbf{R}}$ in strongly polynomial time. Our proximity result (Theorem 6.6) leads to the following "continuous relaxation" approach. Let $\ell = \lfloor m_{\mathbf{R}} \rfloor$ and $u = \lceil m_{\mathbf{R}} \rceil$ and denote the intersection of $\ddddot{B}$ with the box (interval) $[\ell, u]$ by $\ddddot{B}_{\ell}^{u}$. The dec-min element of $\ddddot{B}_{\ell}^{u}$ is also a dec-min element of $\ddddot{B}$, since the box $[\ell, u]$ contains all dec-min elements of $\ddddot{B}$ by Theorem 6.6. Since $0 \leq u(s) - \ell(s) \leq 1$ for all $s \in S$, $\ddddot{B}_{\ell}^{u}$ can be regarded as a matroid translated by ℓ, i.e., $\ddddot{B}_{\ell}^{u} = \{\ell + \chi_L : L \text{ is a base of } M\}$, where M is a matroid. Therefore, the dec-min element of $\ddddot{B}_{\ell}^{u}$ can be computed as the minimum weight base of matroid M with respect to the weight vector w defined by $w(s) = u(s)^2 - \ell(s)^2$ $(s \in S)$. By the greedy algorithm we can find the minimum weight base of M in strongly polynomial time. Thus the total running time of this algorithm is bounded by strongly polynomial time. Variants of such continuous relaxation algorithm are given in Section 6.5. In the literature [14, 17, 20, 26] we can find continuous relaxation algorithms that are strongly polynomial for special classes of base-polyhedra; see Appendix A for details.

6.2 Review of the principal partition

As is pointed out by Fujishige [13], the dec-min element in the continuous case is closely related to the principal partition. The principal partition is the central concept in a structural theory for submodular functions developed mainly in Japan; Iri gives an early survey in [24] and Fujishige provides a comprehensive historical and technical account in [15]. In this section we summarize the results that are relevant to the analysis of the dec-min element in the continuous case. Originally [13], the results are stated for a real-valued submodular function, and the present version is a translation for a real-valued supermodular function $p : 2^S \to \mathbf{R} \cup \{-\infty\}$.

For any real number λ, let $\mathcal{L}(\lambda)$ denote the family of all maximizers of $p(X) - \lambda|X|$. Then $\mathcal{L}(\lambda)$ is a ring family (lattice), and we denote its smallest member by $L(\lambda)$. That is, $L(\lambda)$ denotes the smallest maximizer of $p(X) - \lambda|X|$.

The following is a well-known basic fact. The proof is included for completeness.

Proposition 6.1. *Let $\lambda > \lambda'$. If $X \in \mathcal{L}(\lambda)$ and $Y \in \mathcal{L}(\lambda')$, then $X \subseteq Y$. In particular, $L(\lambda) \subseteq L(\lambda')$.*

Proof. Let $X \in \mathcal{L}(\lambda)$ and $Y \in \mathcal{L}(\lambda')$. We have

$$\begin{aligned} p(X) + p(Y) &\leq p(X \cap Y) + p(X \cup Y), \\ \lambda|X| + \lambda'|Y| &= \lambda|X \cap Y| + \lambda'|X \cup Y| + (\lambda - \lambda')|X - Y| \\ &\geq \lambda|X \cap Y| + \lambda'|X \cup Y|. \end{aligned} \tag{6.2}$$

It follows from these inequalities that

$$(p(X) - \lambda|X|) + (p(Y) - \lambda'|Y|) \leq (p(X \cap Y) - \lambda|X \cap Y|) + (p(X \cup Y) - \lambda'|X \cup Y|).$$

Here the reverse inequality $\geq$ is also true by $X \in \mathcal{L}(\lambda)$ and $Y \in \mathcal{L}(\lambda')$. Therefore, we have equality in (6.2), which implies $|X - Y| = 0$, i.e., $X \subseteq Y$. ■

There are finitely many numbers λ for which $|\mathcal{L}(\lambda)| \geq 2$. We denote such numbers as $\lambda_1 > \lambda_2 > \cdots > \lambda_r$, which are called the **critical values**. It is easy to see that λ is a critical value if and only if $L(\lambda) \neq L(\lambda - \varepsilon)$ for any $\varepsilon > 0$.

The **principal partition** $\{\hat{S}_1, \hat{S}_2, \ldots, \hat{S}_r\}$ is defined by

$$\hat{S}_i = \max \mathcal{L}(\lambda_i) - \min \mathcal{L}(\lambda_i) \qquad (i = 1, 2, \ldots, r), \tag{6.3}$$

which says that $\hat{S}_i$ is the difference of the largest and the smallest element of $\mathcal{L}(\lambda_i)$. Alternatively,

$$\hat{S}_i = L(\lambda_i - \varepsilon) - L(\lambda_i) \tag{6.4}$$

for a sufficiently small $\varepsilon > 0$.

By defining $\hat{C}_i = \hat{S}_1 \cup \hat{S}_2 \cup \cdots \cup \hat{S}_i$ for $i = 1, 2, \ldots, r$ we obtain a chain: $\hat{C}_1 \subset \hat{C}_2 \subset \cdots \subset \hat{C}_r$, where $\hat{C}_1 \neq \emptyset$ and $\hat{C}_r = S$; we also define $\hat{C}_0 = \emptyset$. Then the chain $(\emptyset =)\hat{C}_0 \subset \hat{C}_1 \subset \hat{C}_2 \subset \cdots \subset \hat{C}_r \ (= S)$ is a maximal chain of the lattice $\bigcup_{\lambda \in \mathbf{R}} \mathcal{L}(\lambda)$. In this paper we call this chain the **principal chain**. By slight abuse of terminology the principal chain sometime means the chain $\hat{C}_1 \subset \hat{C}_2 \subset \cdots \subset \hat{C}_r \ (= S)$ without $\hat{C}_0 \ (= \emptyset)$.

Let $m_{\mathbf{R}} \in \mathbf{R}^S$ be the minimum norm point of B, which is the unique dec-min element of B. The critical values are exactly those numbers that appear as component values of $m_{\mathbf{R}}$. Moreover, the vector $m_{\mathbf{R}}$ is uniform on each member $\hat{S}_i$.

Proposition 6.2 (Fujishige [13]). *$m_{\mathbf{R}}(s) = \lambda_i$ if $s \in \hat{S}_i$, where $i = 1, 2, \dots, r$.* ∎

6.3 New characterization of the canonical partition

For the discrete case, the canonical partition describes the structure of dec-min elements. In particular, a dec-min element is near-uniform on each member of the canonical partition.[5] In Part I [9], the canonical partition has been defined iteratively using contractions. In this section we give a non-iterative construction of this canonical partition, which reflects the underlying structure more directly. This alternative construction enables us to reveal the precise relation between the discrete and continuous cases in Section 6.4.

We first recall the iterative construction from Section 5 of Part I [9]. Let $p : 2^S \to \mathbf{Z} \cup \{-\infty\}$ be an integer-valued supermodular function with $p(\emptyset) = 0$ and $p(S) > -\infty$, and $C_0 = \emptyset$. For $j = 1, 2, \dots, q$, define

$$\beta_j = \max \left\{ \left\lceil \frac{p(X \cup C_{j-1}) - p(C_{j-1})}{|X|} \right\rceil : \emptyset \neq X \subseteq \overline{C_{j-1}} \right\}, \tag{6.5}$$

$$h_j(X) = p(X \cup C_{j-1}) - (\beta_j - 1)|X| - p(C_{j-1}) \qquad (X \subseteq \overline{C_{j-1}}), \tag{6.6}$$

$$S_j = \text{smallest subset of } \overline{C_{j-1}} \text{ maximizing } h_j, \tag{6.7}$$

$$C_j = C_{j-1} \cup S_j, \tag{6.8}$$

where $\overline{C_{j-1}} = S - C_{j-1}$ and the index q is determined by the condition that $C_{q-1} \neq S$ and $C_q = S$.

According to the above definitions, we have that

$$C_j \text{ is the smallest maximizer of } p(X) - (\beta_j - 1)|X| \text{ among all } Z \supseteq C_{j-1}. \tag{6.9}$$

We will show in Proposition 6.3 below that C_j is, in fact, the smallest maximizer of $p(X) - (\beta_j - 1)|X|$ among all subsets X of S.

For any integer β, let $\mathcal{L}(\beta)$ denote the family of all maximizers of $p(X) - \beta|X|$, and $L(\beta)$ be the smallest element of $\mathcal{L}(\beta)$, where the smallest element exists in $\mathcal{L}(\beta)$ since $\mathcal{L}(\beta)$ is a lattice (ring family). (These notations are consistent with the ones introduced in Section 6.2.)

Proposition 6.3.
(1) $\beta_1 > \beta_2 > \cdots > \beta_q$.
(2) *For each j with $1 \leq j \leq q$, C_j is the smallest maximizer of $p(X) - (\beta_j - 1)|X|$ among all subsets X of S.*

Proof. (1) The monotonicity of the β-values is already shown in Section 5 of Part I [9], but we give an alternative proof here. Let $j \geq 2$. By (6.5), we have $\beta_{j-1} > \beta_j$ if and only if

$$\beta_{j-1} > \left\lceil \frac{p(X \cup C_{j-1}) - p(C_{j-1})}{|X|} \right\rceil \tag{6.10}$$

[5] That is, $|m_{\mathbf{Z}}(s) - m_{\mathbf{Z}}(t)| \leq 1$ if $\{s, t\} \subseteq S_j$ for some S_j (cf., Theorem 5.1 of Part I [9]).

for every X with $\emptyset \neq X \subseteq \overline{C_{j-1}}$. Furthermore,

$$\begin{aligned}(6.10) &\iff \beta_{j-1} - 1 \geq \frac{p(X \cup C_{j-1}) - p(C_{j-1})}{|X|} \\ &\iff p(X \cup C_{j-1}) - p(C_{j-1}) \leq (\beta_{j-1} - 1)|X| \\ &\iff p(X \cup C_{j-1}) - (\beta_{j-1} - 1)|X \cup C_{j-1}| \leq p(C_{j-1}) - (\beta_{j-1} - 1)|C_{j-1}|.\end{aligned}$$

The last inequality holds, since the set $X \cup C_{j-1}$ contains C_{j-2}, whereas C_{j-1} is the (smallest) maximizer of $p(X) - (\beta_{j-1} - 1)|X|$ among all X containing C_{j-2}. We have thus shown $\beta_{j-1} > \beta_j$.

(2) We prove $C_j = L(\beta_j - 1)$ for $j = 1, 2, \ldots, q$ by induction on j. This holds for $j = 1$ by definition. Let $j \geq 2$. By Proposition 6.1 for $\lambda = \beta_{j-1} - 1$ and $\lambda' = \beta_j - 1$, the smallest maximizer of $p(X) - (\beta_j - 1)|X|$ is a superset of $L(\beta_{j-1} - 1)$, where $L(\beta_{j-1} - 1) = C_{j-1}$ by the induction hypothesis. Combining this with (6.9), we obtain $C_j = L(\beta_j - 1)$. ■

We now give an alternative characterization of the essential value-sequence $\beta_1 > \beta_2 > \cdots > \beta_q$ defined by (6.5)–(6.8). We consider the family $\{L(\beta) : \beta \in \mathbf{Z}\}$ of the smallest maximizers of $p(X) - \beta|X|$ for all integers β. Each C_j is a member of this family, since $C_j = L(\beta_j - 1)$ $(j = 1, 2, \ldots, q)$ by Proposition 6.3(2).

Proposition 6.4. *As β is decreased from $+\infty$ to $-\infty$ (or from β_1 to $\beta_q - 1$), the smallest maximizer $L(\beta)$ is monotone nondecreasing. We have $L(\beta) \neq L(\beta - 1)$ if and only if β is equal to an essential value. Therefore, the essential value-sequence $\beta_1 > \beta_2 > \cdots > \beta_q$ is characterized by the property*[6]

$$\emptyset = L(\beta_1) \subset L(\beta_1 - 1) = \cdots = L(\beta_2) \subset L(\beta_2 - 1) = \cdots = L(\beta_q) \subset L(\beta_q - 1) = S. \quad (6.11)$$

Proof. The monotonicity of $L(\beta)$ follows from Proposition 6.1. We will show (i) $L(\beta_1) = \emptyset$, (ii) $L(\beta_{j-1} - 1) = L(\beta_j)$ for $j = 2, \ldots, q$, and (iii) $L(\beta_j) \subset L(\beta_j - 1)$ for $j = 1, 2, \ldots, q$.

(i) Since $\beta_1 = \max\{\lceil p(X)/|X| \rceil : X \neq \emptyset\}$, we have $p(X) - \beta_1|X| \leq 0$ for all $X \neq \emptyset$, whereas $p(X) - \beta_1|X| = 0$ for $X = \emptyset$. Therefore, $L(\beta_1) = \emptyset$.

(ii) Let $2 \leq j \leq q$. For short we write $C = C_{j-1}$. Define $h(Y) = p(Y) - \beta_j|Y|$ for any subset Y of S, and let A be the smallest maximizer of h, which means $A = L(\beta_j)$. For any nonempty subset X of $\overline{C}$ $(= S - C)$ we have

$$\beta_j \geq \left\lceil \frac{p(X \cup C) - p(C)}{|X|} \right\rceil \geq \frac{p(X \cup C) - p(C)}{|X|},$$

which implies $p(X \cup C) - \beta_j|X \cup C| \leq p(C) - \beta_j|C|$, that is,

$$h(Y) \leq h(C) \qquad \text{for all } \ Y \supseteq C. \quad (6.12)$$

By supermodularity of p we have $h(A) + h(C) \leq h(A \cup C) + h(A \cap C)$, whereas $h(C) \geq h(A \cup C)$ by (6.12). Therefore, $h(A) \leq h(A \cap C)$. Since A is the smallest maximizer of h, this implies that $A = A \cap C$, i.e., $A \subseteq C$. Recalling $A = L(\beta_j)$ and $C = C_{j-1} = L(\beta_{j-1} - 1)$, we obtain

[6] Recall that "$\subset$" means "$\subseteq$ and $\neq$."

$L(\beta_j) \subseteq L(\beta_{j-1} - 1)$. We also have $L(\beta_j) \supseteq L(\beta_{j-1} - 1)$ by the monotonicity. Therefore, $L(\beta_j) = L(\beta_{j-1} - 1)$.

(iii) Let $1 \leq j \leq q$. We continue to write $C = C_{j-1}$. Take a nonempty subset Z of $\overline{C}$ which gives the maximum in the definition of β_j, i.e.,

$$\beta_j = \max \left\{ \left\lceil \frac{p(X \cup C) - p(C)}{|X|} \right\rceil : \emptyset \neq X \subseteq \overline{C} \right\} = \left\lceil \frac{p(Z \cup C) - p(C)}{|Z|} \right\rceil .$$

Then we have

$$\frac{p(Z \cup C) - p(C)}{|Z|} > \beta_j - 1,$$

which implies

$$p(Z \cup C) - (\beta_j - 1)|Z \cup C| > p(C) - (\beta_j - 1)|C|.$$

This shows that $C = C_{j-1} = L(\beta_{j-1} - 1)$ is not a maximizer of $p(Y) - (\beta_j - 1)|Y|$, and hence $L(\beta_{j-1} - 1) \neq L(\beta_j - 1)$. On the other hand, we have $L(\beta_{j-1} - 1) = L(\beta_j)$ by (ii) and $L(\beta_j) \subseteq L(\beta_j - 1)$ by the monotonicity in Proposition 6.1. Therefore, $L(\beta_j) \subset L(\beta_j - 1)$. ■

Proposition 6.4 justifies the following alternative definition of the essential value-sequence, the canonical chain, and the canonical partition:

> Consider the smallest maximizer $L(\beta)$ of $p(X) - \beta|X|$ for all integers β. There are finitely many β for which $L(\beta) \neq L(\beta - 1)$. Denote such integers as $\beta_1 > \beta_2 > \cdots > \beta_q$ and call them the **essential value-sequence**. Furthermore, define $C_j = L(\beta_j - 1)$ for $j = 1, 2, \ldots, q$ to obtain a chain: $C_1 \subset C_2 \subset \cdots \subset C_q$. Call this the **canonical chain**. Finally define a partition $\{S_1, S_2, \ldots, S_q\}$ of S by $S_j = C_j - C_{j-1}$ for $j = 1, 2, \ldots, q$, where $C_0 = \emptyset$, and call this the **canonical partition**.

This alternative construction clearly exhibits the parallelism between the canonical partition in Case **Z** and the principal partition in Case **R**. In particular, the essential value-sequence is exactly the discrete counterpart of the critical values. This is discussed in the next section.

6.4 Canonical partition from the principal partition

The characterization of the canonical partition shown in Section 6.3 enables us to obtain the canonical partition for Case **Z** from the principal partition for Case **R** as follows.

Theorem 6.5.
(1) *An integer β is an essential value if and only if there exists a critical value λ satisfying $\beta \geq \lambda > \beta - 1$.*
(2) *The essential values $\beta_1 > \beta_2 > \cdots > \beta_q$ are obtained from the critical values $\lambda_1 > \lambda_2 > \cdots > \lambda_r$ as the distinct members of the rounded-up integers $\lceil \lambda_1 \rceil \geq \lceil \lambda_2 \rceil \geq \cdots \geq \lceil \lambda_r \rceil$. Let $I(j) = \{i : \lceil \lambda_i \rceil = \beta_j\}$ for $j = 1, 2, \ldots, q$.*

(3) *The canonical partition* $\{S_1, S_2, \ldots, S_q\}$ *is obtained as an aggregation from the principal partition* $\{\hat{S}_1, \hat{S}_2, \ldots, \hat{S}_r\}$*; it is given as*

$$S_j = \bigcup_{i \in I(j)} \hat{S}_i \qquad (j = 1, 2, \ldots, q). \tag{6.13}$$

(4) *The canonical chain* $\{C_j\}$ *is a subchain of the principal chain* $\{\hat{C}_i\}$*; it is given as* $C_j = \hat{C}_i$ *for* $i = \max I(j)$. ■

In Case **R**, the dec-min element $m_{\mathbf{R}}$ of B is uniform on each member $\hat{S}_i$ of the principal partition, i.e., $m_{\mathbf{R}}(s) = \lambda_i$ if $s \in \hat{S}_i$, where $i = 1, 2, \ldots, r$ (cf., Proposition 6.2). In Case **Z**, the dec-min element $m_{\mathbf{Z}}$ of $\overset{....}{B}$ is near-uniform on each member S_j of the canonical partition, i.e., $m_{\mathbf{Z}}(s) \in \{\beta_j, \beta_j - 1\}$ if $s \in S_j$, where $j = 1, 2, \ldots, q$ (cf., Theorem 5.1 of Part I [9]). Combining these results with Theorem 6.5 above we can obtain a (strong) proximity theorem for dec-min elements.

Theorem 6.6 (Proximity)**.** *Let* $m_{\mathbf{R}}$ *be the minimum norm point of* B*. Then every dec-min element* $m_{\mathbf{Z}}$ *of* $\overset{....}{B}$ *satisfies* $\lfloor m_{\mathbf{R}} \rfloor \leq m_{\mathbf{Z}} \leq \lceil m_{\mathbf{R}} \rceil$.

Proof. For $s \in S$ let $\hat{S}_i$ denote the member of the principal partition containing s, and λ_i be the associated critical value. We have $m_{\mathbf{R}}(s) = \lambda_i$ by Proposition 6.2. Let $\beta_j = \lceil \lambda_i \rceil$. This is an essential value, and the corresponding member S_j of the canonical partition contains the element s by Theorem 6.5. We have $m_{\mathbf{Z}}(s) \in \{\beta_j, \beta_j - 1\}$ by Theorem 5.1 of Part I [9]. Therefore, $m_{\mathbf{Z}} \leq \lceil m_{\mathbf{R}} \rceil$.

Next we apply the above argument to $-B$, which is an integral base-polyhedron. Since $-m_{\mathbf{R}}$ is the minimum norm point of $-B$ and $-m_{\mathbf{Z}}$ is a dec-min (=inc-max) element for $-\overset{....}{B}$, we obtain $-m_{\mathbf{Z}} \leq \lceil -m_{\mathbf{R}} \rceil$, which is equivalent to $m_{\mathbf{Z}} \geq \lfloor m_{\mathbf{R}} \rfloor$. ■

Remark 6.1. Theorem 6.6 implies a weaker statement that

$$\text{There exists a dec-min element } m_{\mathbf{Z}} \text{ of } \overset{....}{B} \text{ satisfying } \lfloor m_{\mathbf{R}} \rfloor \leq m_{\mathbf{Z}} \leq \lceil m_{\mathbf{R}} \rceil, \tag{6.14}$$

where $m_{\mathbf{R}}$ is the minimum norm point of B. This statement (6.14) should not be confused with Proposition 6.8 in Section 6.5, which is another proximity statement referring to a minimizer of the piecewise extension of the quadratic function, not to the minimum norm point (minimizer of the quadratic function itself). ■

Theorem 6.7. *The minimum norm point of* B *can be represented as a convex combination of the dec-min elements of* $\overset{....}{B}$.

Proof. On one hand, it was shown in Section 5.1 of Part I [9] that the dec-min elements of $\overset{....}{B}$ lie on the face $B^{\oplus}$ of B defined by the canonical chain $C_1 \subset C_2 \subset \cdots \subset C_q$. This face is the intersection of B with the hyperplanes $\{x \in \mathbf{R}^S : \widetilde{x}(C_j) = p(C_j)\}$ $(j = 1, 2, \ldots, q)$. On the other hand, it is known ([13], [14, Section 9.2]) that the minimum norm point $m_{\mathbf{R}}$ of B lies on the face of B defined by the principal chain $\hat{C}_1 \subset \hat{C}_2 \subset \cdots \subset \hat{C}_r$, which is the intersection of B with the hyperplanes $\{x \in \mathbf{R}^S : \widetilde{x}(\hat{C}_i) = p(\hat{C}_i)\}$ $(i = 1, 2, \ldots, r)$. Since the principal chain is a refinement of the canonical chain (Theorem 6.5), the latter face is a face of $B^{\oplus}$. Therefore, $m_{\mathbf{R}}$ belongs to $B^{\oplus}$. The point $m_{\mathbf{R}}$ also belongs to $T^* = \{x \in \mathbf{R}^S : \beta_j - 1 \leq x(s) \leq \beta_j$ whenever $s \in S_j$ $(j = 1, 2, \ldots, q)\}$, since $m_{\mathbf{R}}(s) = \lambda_i$ for $s \in \hat{S}_i$ (Proposition 6.2) and $S_j = \bigcup\{\hat{S}_i : \lceil \lambda_i \rceil = \beta_j\}$ (Theorem 6.5). Therefore, $m_{\mathbf{R}}$ is a member of $B^{\bullet} = B^{\oplus} \cap T^*$. By recalling that $B^{\bullet}$ is an integral base-polyhedron whose vertices are precisely the dec-min elements of $\overset{....}{B}$, we conclude that $m_{\mathbf{R}}$ can be represented as a convex combination of the dec-min elements of $\overset{....}{B}$. ■

The following two examples illustrate Theorem 6.5.

Example 6.2. Let $S = \{s_1, s_2\}$ and $\overset{....}{B} = \{(0,3), (1,2), (2,1)\}$, where B is the line segment connecting $(0,3)$ and $(2,1)$. For $\overset{....}{B}$ there are two dec-min elements: $m_{\mathbf{Z}}^{(1)} = (1,2)$ and $m_{\mathbf{Z}}^{(2)} = (2,1)$. The minimum norm point (dec-min element) of B is $m_{\mathbf{R}} = (3/2, 3/2)$. The supermodular function p is given by

$$p(\emptyset) = 0, \quad p(\{s_1\}) = 0, \quad p(\{s_2\}) = 1, \quad p(\{s_1, s_2\}) = 3,$$

and we have

$$p(X) - \lambda|X| = \begin{cases} 0 & (X = \emptyset), \\ -\lambda & (X = \{s_1\}), \\ 1 - \lambda & (X = \{s_2\}), \\ 3 - 2\lambda & (X = \{s_1, s_2\}). \end{cases}$$

There is only one ($r = 1$) critical value $\lambda_1 = 3/2$ and the associated sublattice is $\mathcal{L}(\lambda_1) = \{\emptyset, S\}$. The principal partition is a trivial partition $\{S\}$. Since $\lceil \lambda_1 \rceil = 2$, we have $\beta_1 = 2$ with $q = 1$, and the (only) member S_1 in the canonical partition is given by $S_1 = L(\beta_1 - 1) = L(1) = S$. Accordingly, the canonical chain consists of only one member $C_1 = S$. ■

Example 6.3. We consider Example 6.1 again. We have $S = \{s_1, s_2, s_3, s_4\}$ and $\overset{....}{B}$ consists of five vectors: $m_1 = (2,1,1,0)$, $m_2 = (2,1,0,1)$, $m_3 = (1,2,1,0)$, $m_4 = (1,2,0,1)$, and $m_5 = (2,2,0,0)$, of which the first four members, m_1 to m_4, are the dec-min elements. The supermodular function p is given by

$$\begin{aligned} &p(\emptyset) = 0, \quad p(\{s_1\}) = p(\{s_2\}) = 1, \quad p(\{s_3\}) = p(\{s_4\}) = 0, \\ &p(\{s_1, s_2\}) = 3, \quad p(\{s_3, s_4\}) = 0, \quad p(\{s_1, s_3\}) = p(\{s_2, s_3\}) = p(\{s_1, s_4\}) = p(\{s_2, s_4\}) = 1, \\ &p(\{s_1, s_2, s_3\}) = p(\{s_1, s_2, s_4\}) = 3, \quad p(\{s_1, s_3, s_4\}) = p(\{s_2, s_3, s_4\}) = 2, \\ &p(\{s_1, s_2, s_3, s_4\}) = 4. \end{aligned}$$

We have

$$\max\{p(X) - \lambda|X| : X \subseteq S\} = \max\{0,\ 1 - \lambda,\ 3 - 2\lambda,\ 3 - 3\lambda,\ 4 - 4\lambda\}.$$

There are two ($r = 2$) critical values $\lambda_1 = 3/2$ and $\lambda_2 = 1/2$, with the associated sublattices $\mathcal{L}(\lambda_1) = \{\emptyset, \{s_1, s_2\}\}$ and $\mathcal{L}(\lambda_2) = \{\{s_1, s_2\}, S\}$. The principal chain is given by $\emptyset \subset \{s_1, s_2\} \subset S$, and the principal partition is a bipartition with $\hat{S}_1 = \{s_1, s_2\}$ and $\hat{S}_2 = \{s_3, s_4\}$. The minimum norm point of the base-polyhedron B is given by $m_{\mathbf{R}} = (3/2, 3/2, 1/2, 1/2)$ by Proposition 6.2. Since $\lceil \lambda_1 \rceil = 2$ and $\lceil \lambda_2 \rceil = 1$, we have $\beta_1 = 2$ and $\beta_2 = 1$ with $q = 2$. The canonical chain consists of two members $C_1 = L(\beta_1 - 1) = L(1) = \{s_1, s_2\}$ and $C_2 = L(\beta_2 - 1) = L(0) = S$. Accordingly, the canonical partition is given by $S_1 = \{s_1, s_2\}$ and $S_2 = \{s_3, s_4\}$. ■

6.5 Continuous relaxation algorithms

In Section 7 of Part I [9], we have presented a strongly polynomial algorithm for finding a dec-min element of $\overset{....}{B}$ as well as for finding the canonical partition. This is based on an iterative approach to construct a dec-min element along the canonical chain.

By making use of the relation between Case **R** and Case **Z**, we can construct continuous relaxation algorithms, which first compute a real (fractional) vector that is guaranteed to be close to an integral dec-min element, and then find the integral dec-min element by solving a linearly weighted matroid optimization problem.

In our continuous relaxation algorithms, we first apply some algorithm for Case **R** to find two integer vectors ℓ and u such that $\mathbf{0} \leq u - \ell \leq \mathbf{1}$, (i.e., $0 \leq u(s) - \ell(s) \leq 1$ for all $s \in S$) and the box $[\ell, u]$ contains at least one dec-min element of $\overset{....}{B}$, i.e.,

$$\ell \leq m_{\mathbf{Z}} \leq u \tag{6.15}$$

for some dec-min element $m_{\mathbf{Z}}$ of $\overset{....}{B}$. We denote the intersection of $\overset{....}{B}$ and $[\ell, u]$ by $\overset{....}{B}{}_{\ell}^{u}$. Then the dec-min element of $\overset{....}{B}{}_{\ell}^{u}$ is a dec-min element of $\overset{....}{B}$. Since $\mathbf{0} \leq u - \ell \leq \mathbf{1}$, $\overset{....}{B}{}_{\ell}^{u}$ can be regarded as a matroid translated by ℓ, i.e., $\overset{....}{B}{}_{\ell}^{u} = \{\ell + \chi_L : L \text{ is a base of } M\}$ for some matroid M. Therefore, the dec-min element of $\overset{....}{B}{}_{\ell}^{u}$ can be computed as the minimum weight base of matroid M with respect to the weight vector w defined by $w(s) = u(s)^2 - \ell(s)^2$ $(s \in S)$. By the greedy algorithm we can find the minimum weight base of M in strongly polynomial time.

We can conceive two different algorithms for finding vectors ℓ and u.

(a) Using the minimum norm point

In Theorem 6.6 we have shown that every dec-min element $m_{\mathbf{Z}}$ of $\overset{....}{B}$ satisfies $\lfloor m_{\mathbf{R}} \rfloor \leq m_{\mathbf{Z}} \leq \lceil m_{\mathbf{R}} \rceil$ for the minimum norm point $m_{\mathbf{R}}$ of B. Therefore, we can choose $\ell = \lfloor m_{\mathbf{R}} \rfloor$ and $u = \lceil m_{\mathbf{R}} \rceil$ in (6.15). With this choice of (ℓ, u), $\overset{....}{B}{}_{\ell}^{u}$ contains all dec-min elements of $\overset{....}{B}$. The decomposition algorithm of Fujishige [13] (see also [14, Section 8.2]) finds the minimum norm point $m_{\mathbf{R}}$ in strongly polynomial time. Therefore, the continuous relaxation algorithm using the minimum norm point is a strongly polynomial algorithm.

Example 6.4. We continue with Example 6.3, where $\overset{....}{B}$ consists of five vectors: $m_1 = (2, 1, 1, 0)$, $m_2 = (2, 1, 0, 1)$, $m_3 = (1, 2, 1, 0)$, $m_4 = (1, 2, 0, 1)$, and $m_5 = (2, 2, 0, 0)$. From the minimum norm point $m_{\mathbf{R}} = (3/2, 3/2, 1/2, 1/2)$, we obtain $\ell = (1, 1, 0, 0)$ and $u = (2, 2, 1, 1)$, and hence $w = (3, 3, 1, 1)$. Since $W(m_i) = 10$ for $i = 1, \ldots, 4$ and $W(m_5) = 12$, the dec-min elements are given by m_1 to m_4. ■

(b) Using the piecewise-linear extension

The algorithm of Groenevelt [17] (see also [14, Section 8.3]) employs a piecewise-linear extension of the objective function. For the quadratic function $\varphi(k) = k^2$, the piecewise-linear extension $\overline{\varphi} : \mathbf{R} \to \mathbf{R}$ is given by: $\overline{\varphi}(t) = (2k - 1)t - k(k - 1)$ if $k - 1 \leq |t| \leq k$ for $k \in \mathbf{Z}$.

The following proximity property is a special case of an observation of Groenevelt [17] (see also [14, Theorem 8.3]).

Proposition 6.8 (Groenevelt [17]). For any minimizer $\overline{m}_{\mathbf{R}} \in \mathbf{R}^S$ of the function $\overline{\Phi}(x) = \sum_{s\in S} \overline{\varphi}(x(s))$ over B, there exists a minimizer $m_{\mathbf{Z}} \in \mathbf{Z}^S$ of $\Phi(x) = \sum_{s\in S} x(s)^2$ over $\overset{....}{B}$ satisfying $\lfloor \overline{m}_{\mathbf{R}} \rfloor \leq m_{\mathbf{Z}} \leq \lceil \overline{m}_{\mathbf{R}} \rceil$.

Proof. (We give a proof for completeness, though it is easy and standard.) By the integrality of B, we can express $\overline{m}_{\mathbf{R}}$ as a convex combination of integral member $z_1, z_2, \ldots, z_k$ of B satisfying $\lfloor \overline{m}_{\mathbf{R}} \rfloor \leq z_i \leq \lceil \overline{m}_{\mathbf{R}} \rceil$ $(i = 1, 2, \ldots, k)$, where $\overline{m}_{\mathbf{R}} = \sum_{i=1}^{k} \lambda_i z_i$ with $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_i > 0$ $(i = 1, 2, \ldots, k)$. Since $\overline{\Phi}$ is piecewise-linear, we have $\overline{\Phi}(\overline{m}_{\mathbf{R}}) = \sum_{i=1}^{k} \lambda_i \Phi(z_i)$, in which $\Phi(z_i) = \overline{\Phi}(z_i) \geq \overline{\Phi}(\overline{m}_{\mathbf{R}})$. Therefore, $z_1, z_2, \ldots, z_k$ are the minimizers of Φ on $\overset{....}{B}$. We can take any z_i as $m_{\mathbf{Z}}$. ■

By Proposition 6.8 we can take $\ell = \lfloor \overline{m}_{\mathbf{R}} \rfloor$ and $u = \lceil \overline{m}_{\mathbf{R}} \rceil$ in (6.15). In this case, however, $\overset{....}{B}{}_{\ell}^{u}$ may not contain all dec-min elements of $\overset{....}{B}$. The complexity of computing $\overline{m}_{\mathbf{R}}$ is not fully analyzed in the literature [14, 17, 26]. See also Remark 6.1.

Remark 6.2. Minimization of a separable convex function on a base-polyhedron has been investigated in the literature of resource allocation under the name of "resource allocation problems under submodular constraints" (Hochbaum [20], Ibaraki–Katoh [22], Katoh–Ibaraki [25], Katoh–Shioura–Ibaraki [26]). The continuous relaxation approach for the case of discrete variables is considered, e.g., by Hochbaum [19] and Hochbaum–Hong [21]. A more recent paper by Moriguchi–Shioura–Tsuchimura [31] discusses this approach in a more general context of M-convex function minimization in discrete convex analysis. It is known ([21, 31], [26, Theorem 23]) that a convex quadratic function $\sum a_i x_i^2$ in discrete variables can be minimized over an integral base-polyhedron in strongly polynomial time if the base-polyhedron has a special structure like "Nested", "Tree," or "Network" in the terminology of [26]. ■

7 Min-max formulas for separable convex functions in DCA

The objective of this section is to pave the way of DCA approach to discrete decreasing minimization on other discrete structures (the intersection of M-convex sets, network flows, submodular flows) that we consider in Parts III and IV [10, 12]. Min-max formulas for separable convex functions on the intersection of M-convex sets and ordinary/submodular flows are presented in a way suitable for their use in Parts III and IV.

In Section 4.3 we have considered the min-max formula

$$\min\{\sum_{s\in S} \varphi_s(x(s)) : x \in \overset{....}{B}\} = \max\{\hat{p}(\pi) - \sum_{s\in S} \psi_s(\pi(s)) : \pi \in \mathbf{Z}^S\} \tag{7.1}$$

for a separable convex function on an M-convex set. Here, p is an integer-valued (fully) supermodular function on S, B is the base-polyhedron defined by p, $\overset{....}{B}$ is the set of integral points of B, and $\hat{p}$ is the linear extension (Lovász extension) of p. For each $s \in S$, $\varphi_s : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ is an integer-valued (discrete) convex function and ψ_s is the conjugate function of φ_s. Furthermore, the sets of primal and dual optimal solutions of (7.1) are described in Section 5.2. These results have been used for the DCA-based proofs of some key results on decreasing minimization on an M-convex set in Sections 4.4, 5.3, and 5.4.

The min-max formula (4.1) for the square-sum has been obtained as a special case of (7.1) where the conjugate functions can be given explicitly. To emphasize the role of explicit forms of conjugate functions, we offer in Section 7.1 several examples of (discrete) convex functions that admit explicit expressions of conjugate functions. These worked-out examples of conjugate functions and min-max formulas will hopefully trigger other applications of discrete convex analysis.

7.1 Examples of explicit conjugate functions

In this section we offer several examples of (discrete) convex functions whose conjugate functions can be given explicitly. An explicit representation of the conjugate function renders an easily checkable certificate of optimality in the min-max formulas such as (7.1).

For an integer-valued discrete convex function $\varphi : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$, we denote its conjugate function $\varphi^{\bullet}$ by ψ. That is, function $\psi : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ is defined by

$$\psi(\ell) = \max\{k\ell - \varphi(k) : k \in \mathbf{Z}\} \qquad (\ell \in \mathbf{Z}). \tag{7.2}$$

Obviously, we have

$$\varphi(k) + \psi(\ell) \geq k\ell \qquad (k, \ell \in \mathbf{Z}), \tag{7.3}$$

which is known as the Fenchel–Young inequality, and the equality holds in (7.3) if and only if

$$\varphi(k) - \varphi(k-1) \leq \ell \leq \varphi(k+1) - \varphi(k). \tag{7.4}$$

It is worth noting that, for $a, b, c \in \mathbf{Z}$, the conjugate of the function $\varphi_{a,b,c}(k) = \varphi(k-a)+bk+c$ is given by $\psi(\ell - b) + a(\ell - b) - c$. With abuse of notation we express this as

$$(\varphi(k-a) + bk + c)^{\bullet} = \psi(\ell - b) + a(\ell - b) - c. \tag{7.5}$$

In what follows we demonstrate how to calculate the conjugate functions for piecewise-linear functions, ℓ_1-distances, quadratic functions, power products, and exponential functions.

7.1.1 Piecewise-linear functions

Let a be an integer. For a piecewise-linear function φ defined by

$$\varphi(k) = (k-a)^{+} = \max\{0, k-a\} \qquad (k \in \mathbf{Z}), \tag{7.6}$$

the conjugate function ψ is given by

$$\psi(\ell) = \begin{cases} 0 & (\ell = 0), \\ a & (\ell = 1), \\ +\infty & (\ell \notin \{0, 1\}). \end{cases} \tag{7.7}$$

This explicit form can be used in the DCA-based proof of Theorem 4.7; see Remark 4.7.

For another piecewise-linear function φ defined by

$$\varphi(k) = \begin{cases} 0 & (0 \leq k \leq a), \\ \lambda(k-a) & (a \leq k \leq b), \\ +\infty & (k \leq -1 \text{ or } k \geq b+1) \end{cases} \tag{7.8}$$

for $a, b, \lambda \in \mathbf{Z}$ with $0 \leq a \leq b$ and $\lambda \geq 0$, the conjugate function ψ is given by

$$\psi(\ell) \quad = \quad \begin{cases} 0 & (\ell \leq 0), \\ a\ell & (0 \leq \ell \leq \lambda), \\ b\ell - (b-a)\lambda & (\ell \geq \lambda). \end{cases} \tag{7.9}$$

7.1.2 ℓ_1-distances

Let a be an integer. For function φ defined by

$$\varphi(k) = |k-a| \qquad (k \in \mathbf{Z}), \tag{7.10}$$

the conjugate function ψ is given by

$$\psi(\ell) = \begin{cases} a\ell & (\ell = -1, 0, +1), \\ +\infty & (\text{otherwise}). \end{cases} \tag{7.11}$$

A min-max relation for the minimum ℓ_1-distance between an integer point of B and a given integer point c can be obtained from the min-max formula (7.1). Recall that p and b are, respectively, the supermodular and submodular functions associated with B, and our convention $\widetilde{c}(X) = \sum\{c(s) : s \in X\}$.

Proposition 7.1. *For $c \in \mathbf{Z}^S$,*

$$\begin{aligned} &\min\{\sum_{s \in S} |x(s) - c(s)| : x \in \overset{....}{B}\} \\ &= \max\{p(X) - b(Y) - \widetilde{c}(X) + \widetilde{c}(Y) : X, Y \subseteq S;\ X \cap Y = \emptyset\}. \end{aligned} \tag{7.12}$$

Proof. We choose $\varphi_s(k) = |k - c(s)|$ in (7.1). By (7.11), we may assume $\pi \in \{-1, 0, +1\}^S$ on the right-hand side of (7.1). On representing $\pi = \chi_X - \chi_Y$ with disjoint subsets X and Y, we obtain $\hat{p}(\pi) = p(X) - b(Y)$ and $\sum_{s \in S} \psi_s(\pi(s)) = \widetilde{c}(X) - \widetilde{c}(Y)$. Therefore the right-hand side of (7.1) coincides with that of (7.12). ∎

Let a and b be integers with $a \leq b$, and define φ by

$$\varphi(k) = \min\{|k - z| : a \leq z \leq b\} = \max\{a - k, 0, k - b\} \qquad (k \in \mathbf{Z}). \tag{7.13}$$

This function represents the distance from an integer k to the integer interval $[a, b]_{\mathbf{Z}} := \{z \in \mathbf{Z} : a \leq z \leq b\}$. The conjugate function ψ is given by

$$\psi(\ell) = \begin{cases} -a & (\ell = -1), \\ 0 & (\ell = 0), \\ b & (\ell = +1), \\ +\infty & (\text{otherwise}). \end{cases} \tag{7.14}$$

A min-max relation for the minimum ℓ_1-distance between an integer point of B and a given integer interval $[c, d]_{\mathbf{Z}} := \{y \in \mathbf{Z}^S : c(s) \leq y(s) \leq d(s)\ (s \in S)\}$ can be obtained from the min-max formula (7.1), where $c, d \in \mathbf{Z}^S$ and $c \leq d$.

Proposition 7.2. *For $c, d \in \mathbf{Z}^S$ with $c \leq d$,*

$$\begin{aligned} &\min\{\|x - y\|_1 : x \in \ddddot{B},\ y \in [c,d]_{\mathbf{Z}}\} \\ &= \max\{p(X) - b(Y) - \widetilde{d}(X) + \widetilde{c}(Y) : X, Y \subseteq S\,;\ X \cap Y = \emptyset\}. \end{aligned} \tag{7.15}$$

Proof. With reference to (7.13), we define $\varphi_s(k) = \min\{|k - z| : c(d) \leq z \leq d(s)\}$. Then

$$\begin{aligned} &\min\{\ \|x - y\|_1 : x \in \ddddot{B},\ y \in [c,d]_{\mathbf{Z}}\} \\ &= \min\{\ \sum_{s \in S} |x(s) - y(s)| : x \in \ddddot{B},\ y \in [c,d]_{\mathbf{Z}}\} \\ &= \min\{\ \min\{\sum_{s \in S} |x(s) - y(s)| : c(d) \leq y(s) \leq d(s)\ (s \in S)\} : x \in \ddddot{B}\} \\ &= \min\{\ \sum_{s \in S} \min\{\ |x(s) - y(s)| : c(d) \leq y(s) \leq d(s)\} : x \in \ddddot{B}\} \\ &= \min\{\ \sum_{s \in S} \varphi_s(x(s)) : x \in \ddddot{B}\}. \end{aligned}$$

Thus the left-hand side of (7.15) is in the form of the left-hand side of the min-max formula (7.1). By (7.14), we may assume $\pi \in \{-1, 0, +1\}^S$ on the right-hand side of (7.1). On representing $\pi = \chi_X - \chi_Y$ with disjoint subsets X and Y, we obtain $\hat{p}(\pi) = p(X) - b(Y)$ and $\sum_{s \in S} \psi_s(\pi(s)) = \widetilde{d}(X) - \widetilde{c}(Y)$. Therefore the right-hand side of (7.1) coincides with that of (7.15). ■

7.1.3 Quadratic functions

For a quadratic function φ defined by

$$\varphi(k) = ak^2 \qquad (k \in \mathbf{Z}) \tag{7.16}$$

with a positive integer a, the conjugate function ψ is given (cf., Remark 7.1) by

$$\psi(\ell) = \left\lfloor \frac{1}{2}\left(\frac{\ell}{a} + 1\right) \right\rfloor \left(\ell - a \left\lfloor \frac{1}{2}\left(\frac{\ell}{a} + 1\right) \right\rfloor\right), \tag{7.17}$$

which admits the following alternative expressions:

$$\psi(\ell) = \left\lceil \frac{1}{2}\left(\frac{\ell}{a} - 1\right) \right\rceil \left(\ell - a \left\lceil \frac{1}{2}\left(\frac{\ell}{a} - 1\right) \right\rceil\right), \tag{7.18}$$

$$\psi(\ell) = \max\left\{ \left\lfloor \frac{\ell}{2a} \right\rfloor \left(\ell - a \left\lfloor \frac{\ell}{2a} \right\rfloor\right),\ \left\lceil \frac{\ell}{2a} \right\rceil \left(\ell - a \left\lceil \frac{\ell}{2a} \right\rceil\right) \right\}. \tag{7.19}$$

If $a = 1$, these expressions reduce to $\psi(\ell) = \lfloor \ell/2 \rfloor \cdot \lceil \ell/2 \rceil$ in (4.23).

The min-max formula (4.1) for the square-sum can be extended for a nonsymmetric quadratic function $\sum_{s \in S} c(s)x(s)^2$, where $c(s)$ is a positive integer for each $s \in S$.

Theorem 7.3. *For an integer vector $c \in \mathbf{Z}^S$ with $c(s) \geq 1$ for every $s \in S$,*

$$\min\{\sum_{s \in S} c(s)x(s)^2 : x \in \ddot{\ddot{B}}\}$$
$$= \max\{\hat{p}(\pi) - \sum_{s \in S} \left\lfloor \frac{1}{2}\left(\frac{\pi(s)}{c(s)} + 1\right)\right\rfloor \left(\pi(s) - c(s)\left\lfloor \frac{1}{2}\left(\frac{\pi(s)}{c(s)} + 1\right)\right\rfloor\right) : \pi \in \mathbf{Z}^S\}. \tag{7.20}$$

■

In the basic case where $c(s) = 1$ for all $s \in S$, we had a combinatorial constructive proof in Part I [9]. Such a direct combinatorial proof, not relying on the Fenchel-type discrete duality in DCA, for the general case of (7.20) would be an interesting topic.

Remark 7.1. We derive (7.17), (7.18), and (7.19). Since φ is discrete convex, the maximum in the definition (7.2) of $\psi(\ell)$ is attained by k satisfying

$$\varphi(k) - \varphi(k-1) \leq \ell \leq \varphi(k+1) - \varphi(k). \tag{7.21}$$

For $\varphi(k) = ak^2$ this condition reads $a(2k-1) \leq \ell \leq a(2k+1)$, or equivalently

$$\frac{1}{2}\left(\frac{\ell}{a} - 1\right) \leq k \leq \frac{1}{2}\left(\frac{\ell}{a} + 1\right).$$

Therefore, the maximum in (7.2) is attained by $k = \left\lfloor \frac{1}{2}\left(\frac{\ell}{a} + 1\right)\right\rfloor$ and also by $k = \left\lceil \frac{1}{2}\left(\frac{\ell}{a} - 1\right)\right\rceil$. This gives (7.17) and (7.18), respectively.

To derive (7.19) we consider $\varphi(t) = at^2$ in $t \in \mathbf{R}$ and its derivative $\varphi'(t) = 2at$. Let k_ℓ be the integer satisfying

$$\varphi'(k_\ell) \leq \ell < \varphi'(k_\ell + 1). \tag{7.22}$$

Then the maximum in the definition (7.2) of $\psi(\ell)$ is attained by $k = k_\ell$ if $\varphi'(k_\ell) = \ell$, and otherwise by $k = k_\ell$ or $k_\ell + 1$. For $\varphi(k) = ak^2$, we have $k_\ell = \left\lfloor \frac{\ell}{2a} \right\rfloor$ and the maximum is attained by $k = \left\lfloor \frac{\ell}{2a} \right\rfloor$ or $k = \left\lceil \frac{\ell}{2a} \right\rceil$. Hence we have (7.19). ■

7.1.4 Power products

For function φ defined by

$$\varphi(k) = a\, k^{2b} \qquad (k \in \mathbf{Z}) \tag{7.23}$$

with positive integers a and b, the conjugate function ψ is given (cf., Remark 7.2) by

$$\psi(\ell) = \max\left\{ \ell \lfloor K(\ell) \rfloor - a \lfloor K(\ell) \rfloor^{2b},\ \ \ell \lceil K(\ell) \rceil - a \lceil K(\ell) \rceil^{2b} \right\}, \tag{7.24}$$

where

$$K(\ell) = \left(\frac{\ell}{2ab}\right)^{1/(2b-1)}.$$

By choosing $a = 1$ and $b = 2$, for example, we obtain a min-max formula

$$\min\{\sum_{s \in S} x(s)^4 : x \in \ddot{\ddot{B}}\}$$
$$= \max\{\hat{p}(\pi) - \sum_{s \in S} \max\left\{ \pi(s) \left\lfloor (\pi(s)/4)^{1/3} \right\rfloor - \left\lfloor (\pi(s)/4)^{1/3} \right\rfloor^4, \right.$$
$$\left. \pi(s) \left\lceil (\pi(s)/4)^{1/3} \right\rceil - \left\lceil (\pi(s)/4)^{1/3} \right\rceil^4 \right\} : \pi \in \mathbf{Z}^S\}. \tag{7.25}$$

Remark 7.2. We derive (7.24) on the basis of (7.22) for $\varphi(t) = a\, t^{2b}$ and $\varphi'(t) = 2ab\, t^{2b-1}$. We have $k_\ell = \left\lfloor \left(\frac{\ell}{2ab}\right)^{1/(2b-1)} \right\rfloor$, and the maximum in (7.2) is attained by $k = \left\lfloor \left(\frac{\ell}{2ab}\right)^{1/(2b-1)} \right\rfloor$ or $k = \left\lceil \left(\frac{\ell}{2ab}\right)^{1/(2b-1)} \right\rceil$. Hence follows (7.24). ■

7.1.5 Exponential functions

For an exponential function φ defined by

$$\varphi(k) = \begin{cases} 2^k & (k \geq 0), \\ +\infty & (\text{otherwise}), \end{cases} \tag{7.26}$$

the conjugate function ψ is given (cf., Remark 7.3) by

$$\psi(\ell) = \ell \lceil \log_2 \ell \rceil - 2^{\lceil \log_2 \ell \rceil}. \tag{7.27}$$

More generally, for function φ defined by

$$\varphi(k) = \begin{cases} a\, b^k & (k \geq 0), \\ +\infty & (\text{otherwise}) \end{cases} \tag{7.28}$$

with integers $a \geq 1$ and $b \geq 2$, the conjugate function ψ is given (cf., Remark 7.3) by

$$\psi(\ell) = \ell \left\lceil \log_b \left(\frac{\ell}{a(b-1)} \right) \right\rceil - a\, b^{\left\lceil \log_b \left(\frac{\ell}{a(b-1)} \right) \right\rceil}. \tag{7.29}$$

Theorem 7.4. *Assume that B is contained in the nonnegative orthant $\mathbf{Z}^S_+$. Then*

$$\begin{aligned} &\min\{\sum_{s \in S} 2^{x(s)} : x \in \overline{\overline{B}}\} \\ &= \max\{\hat{p}(\pi) - \sum_{s \in S} \left(\pi(s) \lceil \log_2 \pi(s) \rceil - 2^{\lceil \log_2 \pi(s) \rceil} \right) : \pi \in \mathbf{Z}^S \}. \end{aligned} \tag{7.30}$$

More generally, for an integer vector $c, d \in \mathbf{Z}^S$ with $c(s) \geq 1$, $d(s) \geq 2$ $(s \in S)$,

$$\begin{aligned} &\min\{\sum_{s \in S} c(s)\, d(s)^{x(s)} : x \in \overline{\overline{B}}\} \\ &= \max\{\hat{p}(\pi) - \sum_{s \in S} \left(\pi(s) \lceil K(\ell) \rceil - c(s)\, d(s)^{\lceil K(\ell) \rceil} \right) : \pi \in \mathbf{Z}^S \}, \end{aligned} \tag{7.31}$$

where

$$K(\ell) = \log_{d(s)} \left(\frac{\pi(s)}{c(s)(d(s)-1)} \right).$$

■

Remark 7.3. We derive (7.29) on the basis of (7.21). For $\varphi(k) = a\, b^k$, the condition (7.21) reads $a(b-1)b^{k-1} \leq \ell \leq a(b-1)b^k$, or equivalently

$$k - 1 \;\leq\; \log_b \left(\frac{\ell}{a(b-1)} \right) \;\leq\; k.$$

Therefore, the maximum in (7.2) is attained by $k = \left\lceil \log_b \left(\frac{\ell}{a(b-1)} \right) \right\rceil$. Hence follows (7.29). By setting $a = 1$ and $b = 2$ in (7.29), we obtain (7.27). ■

7.2 Separable convex functions on the intersection of M-convex sets

The duality formula (7.1) for separable convex functions on an M-convex set admits an extension to separable convex functions on the intersection of two M-convex sets. In Part IV [11] this extension serves as a basis of the study of decreasing-minimality in the intersection of two M-convex sets (integral base polyhedra).

Let B_1 and B_2 be two integral base-polyhedra, and p_1 and p_2 be the associated (integer-valued) supermodular functions. For $i = 1, 2$, the set of integer points of B_i is denoted as $\ddot{B}_i$, and the linear extension (Lovász extension) of p_i as $\hat{p}_i$. For each $s \in S$, let $\varphi_s : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ be an integer-valued discrete convex function. As before we denote the conjugate function of φ_s by $\psi_s : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$, which is defined by (4.16). Recall notation $\mathrm{dom}\,\Phi = \{x \in \mathbf{Z}^S : x(s) \in \mathrm{dom}\,\varphi_s \text{ for each } s \in S\}$.

The following theorem gives a duality formula for separable discrete convex functions on the intersection of two M-convex sets. We introduce notations for feasible vectors:

$$\mathcal{P}_0 = \{x \in \ddot{B}_1 \cap \ddot{B}_2 : x(s) \in \mathrm{dom}\,\varphi_s \text{ for each } s \in S\} \ = \ddot{B}_1 \cap \ddot{B}_2 \cap \mathrm{dom}\,\Phi, \tag{7.32}$$

$$\mathcal{D}_0 = \{(\pi_1, \pi_2) \in \mathbf{Z}^S \times \mathbf{Z}^S : \pi_i \in \mathrm{dom}\,\hat{p}_i \ (i = 1, 2),\ \pi_1(s) + \pi_2(s) \in \mathrm{dom}\,\psi_s \text{ for each } s \in S\}. \tag{7.33}$$

Theorem 7.5. *Assume that $\mathcal{P}_0 \neq \emptyset$ (primal feasibility) or $\mathcal{D}_0 \neq \emptyset$ (dual feasibility) holds. Then we have the min-max relation:*[7]

$$\begin{aligned}&\min\{\sum_{s \in S} \varphi_s(x(s)) : x \in \ddot{B}_1 \cap \ddot{B}_2\} \\ &= \max\{\hat{p}_1(\pi_1) + \hat{p}_2(\pi_2) - \sum_{s \in S} \psi_s(\pi_1(s) + \pi_2(s)) : \pi_1, \pi_2 \in \mathbf{Z}^S\}.\end{aligned} \tag{7.34}$$

Proof. We give a proof based on an iterative application of the Fenchel duality theorem (Theorem 4.1), while the weak duality (min ≥ max) is demonstrated in Remark 7.4.

We denote the indicator functions of $\ddot{B}_1$ and $\ddot{B}_2$ by δ_1 and δ_2, respectively, and use the notation $\Phi(x) = \sum[\varphi_s(x(s)) : s \in S]$. In the Fenchel-type discrete duality

$$\min\{f(x) - h(x) : x \in \mathbf{Z}^S\} = \max\{h^{\circ}(\pi) - f^{\bullet}(\pi) : \pi \in \mathbf{Z}^S\} \tag{7.35}$$

in (4.11), we choose $f = \delta_2 + \Phi$ and $h = -\delta_1$. Since $f - h = \Phi + \delta_1 + \delta_2$, the left-hand side of (7.35) coincides with the left-hand side of (7.34).

The conjugate function $f^{\bullet}$ can be computed as follows. For $\pi \in \mathbf{Z}^S$ we define $\varphi_s^{\pi}(k) = \varphi_s(k) - \pi(s)k$ for $k \in \mathbf{Z}$ and $s \in S$. Then the conjugate function ψ_s^{π} of function φ_s^{π} is given as

$$\begin{aligned}\psi_s^{\pi}(\ell) &= \max\{k\ell - \varphi_s^{\pi}(k) : k \in \mathbf{Z}\} \\ &= \max\{k(\ell + \pi(s)) - \varphi_s(k) : k \in \mathbf{Z}\} \\ &= \psi_s(\ell + \pi(s)) \qquad (\ell \in \mathbf{Z}).\end{aligned}$$

[7] The unbounded case with both sides of (7.34) being equal to $-\infty$ or $+\infty$ is also a possibility.

Using this expression and the min-max formula (7.1) for B_2 and φ_s^π, we obtain

$$\begin{aligned}
f^\bullet(\pi) &= \max\{\langle \pi, x\rangle - \delta_2(x) - \sum_{s\in S} \varphi_s(x(s)) : x \in \mathbf{Z}^S\} \\
&= \max\{\sum_{s\in S} [\pi(s)x(s) - \varphi_s(x(s))] - \delta_2(x) : x \in \mathbf{Z}^S\} \\
&= -\min\{\sum_{s\in S} \varphi_s^\pi(x(s)) : x \in \ddot{B}_2\} \\
&= -\max\{\hat{p}_2(\pi') - \sum_{s\in S} \psi_s^\pi(\pi'(s)) : \pi' \in \mathbf{Z}^S\} \\
&= -\max\{\hat{p}_2(\pi') - \sum_{s\in S} \psi_s(\pi(s) + \pi'(s)) : \pi' \in \mathbf{Z}^S\} \qquad (\pi \in \mathbf{Z}^S).
\end{aligned} \tag{7.36}$$

On the other hand, the conjugate function h° of $h = -\delta_1$ is equal to $\hat{p}_1$ by (4.18), i.e.,

$$h^\circ(\pi) = \hat{p}_1(\pi) \qquad (\pi \in \mathbf{Z}^S). \tag{7.37}$$

The substitution of (7.36) and (7.37) into $h^\circ - f^\bullet$ shows that the right-hand side of (7.35) coincides with the right-hand side of (7.34). ■

Remark 7.4. The weak duality (min $\geq$ max) in (7.34) is shown here. Let $x \in \mathcal{P}_0$ and $(\pi_1, \pi_2) \in \mathcal{D}_0$. Then, using the Fenchel–Young inequality (7.3) for (φ_s, ψ_s) as well as (4.20) for $p = p_i$ $(i = 1, 2)$, we obtain

$$\begin{aligned}
&\sum_{s\in S} \varphi_s(x(s)) - \left(\hat{p}_1(\pi_1) + \hat{p}_2(\pi_2) - \sum_{s\in S} \psi_s(\pi_1(s) + \pi_2(s))\right) \\
&= \sum_{s\in S} [\varphi_s(x(s)) + \psi_s(\pi_1(s) + \pi_2(s))] - \hat{p}_1(\pi_1) - \hat{p}_2(\pi_2) \\
&\geq \sum_{s\in S} x(s)(\pi_1(s) + \pi_2(s)) - \hat{p}_1(\pi_1) - \hat{p}_2(\pi_2) \qquad (7.38) \\
&= \sum_{s\in S} \pi_1(s)x(s) + \sum_{s\in S} \pi_2(s)x(s) - \hat{p}_1(\pi_1) - \hat{p}_2(\pi_2) \\
&\geq \min\{\pi_1 z : z \in \ddot{B}_1\} + \min\{\pi_2 z : z \in \ddot{B}_2\} - \hat{p}_1(\pi_1) - \hat{p}_2(\pi_2) = 0, \qquad (7.39)
\end{aligned}$$

showing the weak duality. The optimality conditions can be obtained as the conditions for the inequalities in (7.38) and (7.39) to be equalities, as stated in Proposition 7.6 below. ■

In the min-max formula (7.34) we denote the set of the minimizers x by $\mathcal{P}$ and the set of the maximizers (π_1, π_2) by $\mathcal{D}$. The following proposition follows from the combination of Theorem 7.5 and Remark 7.4. We remark that this proposition is a special case of Theorem 5.4.

Proposition 7.6. *Assume that both $\mathcal{P}_0$ and $\mathcal{D}_0$ in* (7.32)–(7.33) *are nonempty.*

(1) *Let $x \in \mathcal{P}_0$ and $(\pi_1, \pi_2) \in \mathcal{D}_0$. Then $x \in \mathcal{P}$ and $(\pi_1, \pi_2) \in \mathcal{D}$ if and only if the following three conditions are satisfied:*

$$\varphi_s(x(s)) - \varphi_s(x(s) - 1) \leq \pi_1(s) + \pi_2(s) \leq \varphi_s(x(s) + 1) - \varphi_s(x(s)) \qquad (s \in S), \tag{7.40}$$

$$\pi_1(s) \geq \pi_1(t) \quad \text{for every } (s, t) \text{ with } x + \chi_s - \chi_t \in \ddot{B}_1, \tag{7.41}$$

$$\pi_2(s) \geq \pi_2(t) \quad \text{for every } (s, t) \text{ with } x + \chi_s - \chi_t \in \ddot{B}_2. \tag{7.42}$$

(2) *For any* $(\hat{\pi}_1, \hat{\pi}_2) \in \mathcal{D}$, *we have*

$$\begin{aligned}\mathcal{P} &= \{x \in \mathcal{P}_0 : (7.40), (7.41), (7.42) \text{ hold with } (\pi_1, \pi_2) = (\hat{\pi}_1, \hat{\pi}_2)\} &(7.43)\\ &= \{x \in \mathrm{dom}\,\Phi : (7.40) \text{ holds with } (\pi_1, \pi_2) = (\hat{\pi}_1, \hat{\pi}_2)\}\\ &\quad \cap \{x \in \overset{....}{B_1} : x \text{ is a } \hat{\pi}_1\text{-minimizer in } \overset{....}{B_1}\ \}\\ &\quad \cap \{x \in \overset{....}{B_2} : x \text{ is a } \hat{\pi}_2\text{-minimizer in } \overset{....}{B_2}\ \}. &(7.44)\end{aligned}$$

(3) *For any* $\hat{x} \in \mathcal{P}$, *we have*

$$\mathcal{D} = \{(\pi_1, \pi_2) \in \mathcal{D}_0 : (7.40), (7.41), (7.42) \text{ hold with } x = \hat{x}\}. \tag{7.45}$$

Proof. The inequality (7.38) turns into an equality if and only if, for each $s \in S$, we have $\varphi_s(k) + \psi_s(\ell) = k\ell$ for $k = x(s)$ and $\ell = \pi_1(s) + \pi_2(s)$. The latter condition is equivalent to (7.40) by (7.4). The other inequality (7.39) turns into an equality if and only if x is a π_i-minimizer in $\overset{....}{B_i}$ for $i = 1, 2$, that is, (7.41) and (7.42) hold. Finally, we see from Theorem 7.5 that the two inequalities in (7.38) and (7.39) simultaneously turn into equality for some x and (π_1, π_2). ■

Proposition 7.7. *In the min-max relation* (7.34) *for a separable convex function on the intersection of two M-convex sets, the set* $\mathcal{D}' := \{(\pi_1, -\pi_2) : (\pi_1, \pi_2) \in \mathcal{D}\}$ *corresponding to the maximizers is an* $\mathrm{L}^\natural$*-convex set and the set* $\mathcal{P}$ *of the minimizers is an* $\mathrm{M}_2^\natural$*-convex set.*

Proof. The representation (7.45) shows that $\mathcal{D}$ is described by the inequalities in (7.40), (7.41), and (7.42). Hence $\mathcal{D}'$ is $\mathrm{L}^\natural$-convex. In the representation (7.44) of $\mathcal{P}$, the first set $\{x \in \mathrm{dom}\,\Phi : (7.40) \text{ holds with } (\pi_1, \pi_2) = (\hat{\pi}_1, \hat{\pi}_2)\}$ is a box, while for each $i = 1, 2$, the set of $\hat{\pi}_i$-minimizers in $\overset{....}{B_i}$ is an M-convex set. Therefore, $\mathcal{P}$ is an $\mathrm{M}_2^\natural$-convex set. ■

When specialized to a symmetric function, the min-max formula (7.34) is simplified to

$$\begin{aligned}&\min\{\sum_{s \in S} \varphi(x(s)) : x \in \overset{....}{B_1} \cap \overset{....}{B_2}\}\\ &= \max\{\hat{p}_1(\pi_1) + \hat{p}_2(\pi_2) - \sum_{s \in S} \psi(\pi_1(s) + \pi_2(s)) : \pi_1, \pi_2 \in \mathbf{Z}^S\},\end{aligned} \tag{7.46}$$

where $\varphi : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ is any integer-valued discrete convex function and $\psi : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ is the conjugate of φ. The identity (7.46) will play a key role in the study of discrete decreasing minimization on the intersection of two M-convex sets, just as (4.22) did for an M-convex set.

As an example of (7.46) we mention a min-max identity for the minimum square-sum of components on the intersection of two M-convex sets, which is an extension of (4.1) for an M-convex set.

Theorem 7.8.

$$\begin{aligned}&\min\{\sum_{s \in S} x(s)^2 : x \in \overset{....}{B_1} \cap \overset{....}{B_2}\}\\ &= \max\{\hat{p}_1(\pi_1) + \hat{p}_2(\pi_2) - \sum_{s \in S} \left\lfloor \frac{\pi_1(s) + \pi_2(s)}{2} \right\rfloor \cdot \left\lceil \frac{\pi_1(s) + \pi_2(s)}{2} \right\rceil : \pi_1, \pi_2 \in \mathbf{Z}^S\}.\end{aligned} \tag{7.47}$$

Proof. This is a special case of (7.46) with $\varphi(k) = k^2$ and $\psi(\ell) = \lfloor \ell/2 \rfloor \cdot \lceil \ell/2 \rceil$ (cf., (4.23)). ■

If $\ddddot{B}_1 \cap \ddddot{B}_2 \neq \emptyset$, both sides of (7.47) are finite-valued, and the minimum and the maximum are attained. If $\ddddot{B}_1 \cap \ddddot{B}_2 = \emptyset$, the left-hand side of (7.47) is equal to $+\infty$ by convention and the right-hand side is unbounded above (hence equal to $+\infty$). Note also that $\ddddot{B}_1 \cap \ddddot{B}_2 \neq \emptyset$ if and only if $B_1 \cap B_2 \neq \emptyset$.

We can also formulate a min-max formula for a quadratic function $\sum_{s \in S} c(s)x(s)^2$, where $c(s)$ is a positive integer for each $s \in S$. On recalling the conjugate function in (7.17), we obtain the following min-max formula.

Theorem 7.9. *For an integer vector $c \in \mathbf{Z}^S$ with $c(s) \geq 1$ for all $s \in S$,*

$$\begin{aligned} &\min\{\sum_{s \in S} c(s)x(s)^2 : x \in \ddddot{B}_1 \cap \ddddot{B}_2\} \\ &= \max\{\hat{p}_1(\pi_1) + \hat{p}_2(\pi_2) - \sum_{s \in S} \left\lfloor \frac{1}{2}\left(\frac{\pi(s)}{c(s)} + 1\right) \right\rfloor \left(\pi(s) - c(s)\left\lfloor \frac{1}{2}\left(\frac{\pi(s)}{c(s)} + 1\right)\right\rfloor\right) : \\ &\qquad \pi = \pi_1 + \pi_2, \ \ \pi_1, \pi_2 \in \mathbf{Z}^S\}. \end{aligned} \tag{7.48}$$

■

We have obtained the min-max formulas (7.47) and (7.48) as special cases of the Fenchel-type discrete duality in DCA. Direct algorithmic proofs, not relying on the DCA machinery, would be an interesting research topic.

7.3 Separable convex functions on network flows

Let $D = (V, A)$ be a digraph, and suppose that we are given a finite integer-valued function m on V for which $\widetilde{m}(V) = 0$. A **flow** means simply a function on A, and we are interested in flow x that satisfies

$$\varrho_x(v) - \delta_x(v) = m(v) \qquad \text{for each node } v \in V, \tag{7.49}$$

where

$$\varrho_x(v) := \sum[x(uv) : uv \in A], \qquad \delta_x(v) := \sum[x(vu) : vu \in A].$$

We consider a convex cost integer flow problem. For each edge $e \in A$, an integer-valued (discrete) convex function $\varphi_e : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ is given, and we seek an integral flow x that minimizes the sum of the edge costs $\Phi(x) = \sum_{e \in A} \varphi_e(x(e))$ subject to the constraint (7.49). For the function value $\Phi(x)$ to be finite, we must have

$$x(e) \in \mathrm{dom}\, \varphi_e \qquad \text{for each edge } e \in A, \tag{7.50}$$

and therefore, capacity constraints, if any, can be represented (implicitly) in terms of the cost function φ_e. A flow x is called **feasible** if it satisfies the conditions (7.49) and (7.50).

Convex cost flow problem (1):

$$\text{Minimize} \qquad \Phi(x) = \sum_{e \in A} \varphi_e(x(e)) \tag{7.51}$$

$$\text{subject to} \qquad \varrho_x(v) - \delta_x(v) = m(v) \qquad (v \in V), \tag{7.52}$$

$$x(e) \in \mathbf{Z} \qquad (e \in A). \tag{7.53}$$

The dual problem, in its integer version, is as follows (cf., e.g., [1], [23], [35], [38]). For each $e \in A$, let $\psi_e : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ denote the conjugate of φ_e, that is,

$$\psi_e(\ell) = \max\{k\ell - \varphi_e(k) : k \in \mathbf{Z}\} \qquad (\ell \in \mathbf{Z}), \tag{7.54}$$

which is also an integer-valued (discrete) convex function. This function ψ_e represents the (dual) cost function associated with edge $e \in A$. The decision variable in the dual problem is an integer-valued potential $\pi : V \to \mathbf{Z}$ defined on the node-set V. Recall the notation $\pi m = \sum_{v \in V} \pi(v) m(v)$.

Dual to the convex cost flow problem (1):

$$\text{Maximize} \qquad \Psi(\pi) = \pi m - \sum_{e=uv \in A} \psi_e(\pi(v) - \pi(u)) \tag{7.55}$$

$$\text{subject to} \qquad \pi(v) \in \mathbf{Z} \qquad (v \in V). \tag{7.56}$$

We introduce notations for feasible flows and potentials:

$$\mathcal{P}_0 = \{x \in \mathbf{Z}^A : x \text{ satisfies (7.50) and (7.52)}\}, \tag{7.57}$$

$$\mathcal{D}_0 = \{\pi \in \mathbf{Z}^V : \pi(v) - \pi(u) \in \operatorname{dom} \psi_e \text{ for each } e = uv \in A\}. \tag{7.58}$$

Theorem 7.10. *Assume primal feasibility ($\mathcal{P}_0 \neq \emptyset$) or dual feasibility ($\mathcal{D}_0 \neq \emptyset$). Then we have the min-max relation:*

$$\min\{\Phi(x) : x \in \mathbf{Z}^A \text{ satisfies (7.52)}\} = \max\{\Psi(\pi) : \pi \in \mathbf{Z}^V\}. \tag{7.59}$$

The unbounded case with both sides being equal to $-\infty$ or $+\infty$ is also a possibility. ■

In using this min-max relation in Part III [10] it is convenient to introduce capacity constraints explicitly. We denote the integer-valued lower and upper bound functions on A by f and g, for which $f \leq g$ is assumed, and impose the capacity constraint

$$f(e) \leq x(e) \leq g(e) \qquad \text{for each edge } e \in A. \tag{7.60}$$

With this explicit form of capacity constraints, a flow x is called **feasible** if it satisfies the conditions (7.49), (7.50) and (7.60). The primal problem reads as follows.

Convex cost flow problem (2):

$$\text{Minimize} \qquad \Phi(x) = \sum_{e \in A} \varphi_e(x(e)) \tag{7.61}$$

$$\text{subject to} \qquad \varrho_x(v) - \delta_x(v) = m(v) \qquad (v \in V), \tag{7.62}$$

$$f(e) \leq x(e) \leq g(e) \qquad (e \in A), \tag{7.63}$$

$$x(e) \in \mathbf{Z} \qquad (e \in A). \tag{7.64}$$

The corresponding dual problem can be given as follows (cf., Remark 7.5), where the decision variables consist of an integer-valued potential $\pi : V \to \mathbf{Z}$ on V and integer-valued functions $\tau_1, \tau_2 : A \to \mathbf{Z}$ on A. The constraint (7.66) below says that the tension (potential difference) is split into two parts τ_1 and τ_2.

Dual to the convex cost flow problem (2):

$$\text{Maximize} \quad \Psi(\pi,\tau_1,\tau_2) = \pi m - \sum_{e\in A}\Big(\psi_e(\tau_1(e)) + \max\{f(e)\tau_2(e),\ g(e)\tau_2(e)\} \Big) \tag{7.65}$$

$$\text{subject to} \quad \pi(v) - \pi(u) = \tau_1(e) + \tau_2(e) \qquad (e = uv \in A), \tag{7.66}$$

$$\pi(v) \in \mathbf{Z} \qquad (v \in V), \tag{7.67}$$

$$\tau_1(e), \tau_2(e) \in \mathbf{Z} \qquad (e \in A). \tag{7.68}$$

We introduce notations for feasible flows and potentials/tensions:

$$\mathcal{P}_0 = \{x \in \mathbf{Z}^A : x \text{ satisfies (7.50), (7.62), (7.63) }\}, \tag{7.69}$$

$$\mathcal{D}_0 = \{(\pi,\tau_1,\tau_2) \in \mathbf{Z}^V \times \mathbf{Z}^A \times \mathbf{Z}^A : \text{(7.66)},\ \tau_1(e) \in \mathrm{dom}\,\psi_e \text{ for each } e \in A\}. \tag{7.70}$$

Theorem 7.11. *Assume primal feasibility ($\mathcal{P}_0 \neq \emptyset$) or dual feasibility ($\mathcal{D}_0 \neq \emptyset$). Then we have the min-max relation:*

$$\begin{aligned} &\min\{\Phi(x) : x \in \mathbf{Z}^A \text{ satisfies (7.62) and (7.63)}\} \\ &= \max\{\Psi(\pi,\tau_1,\tau_2) : \pi \in \mathbf{Z}^V \text{ and } \tau_1, \tau_2 \in \mathbf{Z}^A \text{ satisfy (7.66)}\}. \end{aligned} \tag{7.71}$$

The unbounded case with both sides being equal to $-\infty$ or $+\infty$ is also a possibility. ■

For the min-max formula (7.71) we can obtain the following optimality criterion, where we denote the set of the minimizers x by $\mathcal{P}$ and the set of the maximizers (π,τ_1,τ_2) by $\mathcal{D}$.

Proposition 7.12. *Assume that both $\mathcal{P}_0$ and $\mathcal{D}_0$ in* (7.69)–(7.70) *are nonempty.*

(1) *Let $x \in \mathcal{P}_0$ and $(\pi,\tau_1,\tau_2) \in \mathcal{D}_0$. Then $x \in \mathcal{P}$ and $(\pi,\tau_1,\tau_2) \in \mathcal{D}$ if and only if the following two conditions are satisfied:*

$$\varphi_e(x(e)) - \varphi_e(x(e)-1) \leq \tau_1(e) \leq \varphi_e(x(e)+1) - \varphi_e(x(e)) \qquad (e \in A), \tag{7.72}$$

$$\tau_2(e) \begin{cases} = 0 & \text{if } f(e)+1 \leq x(e) \leq g(e)-1, \\ \leq 0 & \text{if } x(e) = f(e), \\ \geq 0 & \text{if } x(e) = g(e) \end{cases} \qquad (e \in A). \tag{7.73}$$

(2) *For any $(\hat{\pi},\hat{\tau}_1,\hat{\tau}_2) \in \mathcal{D}$, we have*

$$\mathcal{P} = \{x \in \mathcal{P}_0 : \text{(7.72) and (7.73) hold with } (\pi,\tau_1,\tau_2) = (\hat{\pi},\hat{\tau}_1,\hat{\tau}_2)\}, \tag{7.74}$$

where the conditions in (7.72) *and* (7.73) *can be rewritten as*

$$x(e) \in \arg\min_k \{\varphi_e(k) - \tau_1(e)k\} \qquad (e \in A), \tag{7.75}$$

$$\begin{cases} x(e) = f(e) & \text{if } \tau_2(e) < 0, \\ f(e) \leq x(e) \leq g(e) & \text{if } \tau_2(e) = 0, \\ x(e) = g(e) & \text{if } \tau_2(e) > 0 \end{cases} \qquad (e \in A). \tag{7.76}$$

(3) *For any $\hat{x} \in \mathcal{P}$, we have*

$$\mathcal{D} = \{(\pi,\tau_1,\tau_2) \in \mathcal{D}_0 : \text{(7.72) and (7.73) hold with } x = \hat{x}\}. \tag{7.77}$$

Proof. This is a special case of Proposition 7.16 in Section 7.4. ■

The condition (7.73), or equivalently (7.76), expresses the so-called kilter condition for flow $x(e)$ and tension $\tau_2(e)$, whereas the condition (7.72), or equivalently (7.75), can be regarded as a nonlinear version thereof for flow $x(e)$ and tension $\tau_1(e)$.

Remark 7.5. Here we derive the dual problem (7.65)–(7.68) from the basic case in (7.55)–(7.56). For each $e \in A$, let δ_e denote the indicator function of the integer interval $[f(e), g(e)]_{\mathbf{Z}}$, define $\tilde{\varphi}_e := \varphi_e + \delta_e$, and let $\tilde{\psi}_e$ be the conjugate function of $\tilde{\varphi}_e$. By the claim below we obtain the following expression

$$\tilde{\psi}_e(\pi(v) - \pi(u)) = \min\left\{\psi_e(\ell_1) + \max\{f(e)\ell_2, g(e)\ell_2\} : \ell_1, \ell_2 \in \mathbf{Z},\ \ell_1 + \ell_2 = \pi(v) - \pi(u)\right\}.$$

The substitution of this expression into (7.55) results in (7.65)–(7.68).

Claim: Let $\varphi : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ be a (discrete) convex function, $\delta : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ the indicator function of an integer interval $[a, b]_{\mathbf{Z}}$ with $a \leq b$. Then the conjugate function $(\varphi + \delta)^{\bullet}$ of $\varphi + \delta$ is given by

$$(\varphi + \delta)^{\bullet}(\ell) = \min\left\{\varphi^{\bullet}(\ell_1) + \max\{a\ell_2, b\ell_2\} : \ell_1, \ell_2 \in \mathbf{Z},\ \ell_1 + \ell_2 = \ell\right\}. \tag{7.78}$$

(Proof) By Theorem 8.36 of [35], $(\varphi + \delta)^{\bullet}$ is equal to the infimum convolution of $\varphi^{\bullet}$ and $\delta^{\bullet}$, that is,

$$(\varphi + \delta)^{\bullet}(\ell) = \min\left\{\varphi^{\bullet}(\ell_1) + \delta^{\bullet}(\ell_2) : \ell_1, \ell_2 \in \mathbf{Z},\ \ell_1 + \ell_2 = \ell\right\}.$$

Here we have

$$\delta^{\bullet}(\ell) = \max\{k\ell - \delta(k)\} = \max\{k\ell : a \leq k \leq b\} = \max\{a\ell, b\ell\}.$$

Hence follows (7.78). ■

Remark 7.6. The feasibility of the primal problems can be expressed by a variant of the Hoffman-condition. Denote the integer interval of $\mathrm{dom}\,\varphi_e$ by $[f'(e), g'(e)]_{\mathbf{Z}}$ with $f'(e) \in \mathbf{Z} \cup \{-\infty\}$ and $g'(e) \in \mathbf{Z} \cup \{+\infty\}$. Then, by Hoffman's theorem, there exists a feasible flow for the basic problem (7.51)–(7.53) if and only if

$$\varrho_{g'}(Z) - \delta_{f'}(Z) \geq \widetilde{m}(Z) \qquad \text{for all} \quad Z \subseteq V \tag{7.79}$$

is satisfied. For the problem (7.61)–(7.64) with explicit capacity constraints, we replace $f'(e)$ and $g'(e)$ by $\max\{f(e), f'(e)\}$ and $\min\{g(e), g'(e)\}$, respectively. ■

7.4 Separable convex functions on submodular flows

Let $D = (V, A)$ be a digraph, and suppose that we are given an integral base-polyhedron B with ground-set V. We assume that B is described as $B = B'(p)$ in (1.3) by an integer-valued (fully) supermodular function $p : 2^V \to \mathbf{Z} \cup \{-\infty\}$ with $p(V) = 0$, which is equivalent to saying that B is described as $B = B(b)$ in (1.2) by an integer-valued (fully) submodular function $b : 2^V \to \mathbf{Z} \cup \{+\infty\}$ with $b(V) = 0$, where b is the complementary function of p.

Here we are interested in an integral flow $x : A \to \mathbf{Z}$ such that the net-in-flow vector $(\varrho_x(v) - \delta_x(v) : v \in V)$ belongs to B, which we express as

$$(\varrho_x(v) - \delta_x(v) : v \in V) \in \overset{....}{B}. \qquad (7.80)$$

Such a flow x is called a **submodular flow**. The constraint (7.49) for the ordinary flow problem in Section 7.3 is a (very) special case of (7.80) where the bounding submodular function b (or the supermodular function p) is a modular function $\widetilde{m}$ defined by the vector m.

We consider a convex cost integer submodular flow problem. For each edge $e \in A$, an integer-valued (discrete) convex function $\varphi_e : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ is given, and we seek an integral flow x that minimizes the sum of the edge costs $\Phi(x) = \sum_{e \in A} \varphi_e(x(e))$ subject to the submodular constraint (7.80). For the function value $\Phi(x)$ to be finite, we must have

$$x(e) \in \mathrm{dom}\, \varphi_e \quad \text{for each edge } e \in A, \qquad (7.81)$$

and therefore, capacity constraints, if any, can be represented (implicitly) in terms of the cost function φ_e. A **feasible submodular flow** means a flow x that satisfies the conditions (7.80) and (7.81).

Convex cost submodular flow problem (1):

$$\text{Minimize} \quad \Phi(x) = \sum_{e \in A} \varphi_e(x(e)) \qquad (7.82)$$

$$\text{subject to} \quad (\varrho_x(v) - \delta_x(v) : v \in V) \in \overset{....}{B}, \qquad (7.83)$$

$$x(e) \in \mathbf{Z} \qquad (e \in A). \qquad (7.84)$$

In discrete convex analysis, a systematic study of convex-cost submodular flows has been conducted in a more general framework called the **M-convex submodular flow problem**, where particular emphasis is laid on duality theorems (Murota [34, 35]).

The decision variable in the dual problem is an integer-valued potential $\pi : V \to \mathbf{Z}$. The objective function $\Psi(\pi)$ involves the linear extension (Lovász extension) $\hat{p}(\pi)$ of the supermodular function p defining B as well as the conjugate function ψ_e of φ_e for all $e \in A$. It is worth noting that πm in (7.55) is replaced by $\hat{p}(\pi)$ in (7.85).

Dual to the convex cost submodular flow problem (1):

$$\text{Maximize} \quad \Psi(\pi) = \hat{p}(\pi) - \sum_{e = uv \in A} \psi_e(\pi(v) - \pi(u)) \qquad (7.85)$$

$$\text{subject to} \quad \pi(v) \in \mathbf{Z} \qquad (v \in V). \qquad (7.86)$$

The following min-max formula can be derived as a special case of a min-max formula [35, (9.83), page 270] for M-convex submodular flows, while the weak duality (min ≥ max) is demonstrated in Remark 7.7 below. We also mention that the min-max formula (7.89) below can be regarded as being equivalent to the Fenchel-type discrete duality theorem (Theorem 4.1); see [35, Section 9.1.4] for the detail of this equivalence. We introduce notations for feasible flows and potentials:

$$\mathcal{P}_0 = \{x \in \mathbf{Z}^A : x \text{ satisfies (7.81) and (7.83)}\}, \qquad (7.87)$$

$$\mathcal{D}_0 = \{\pi \in \mathbf{Z}^V : \pi \in \mathrm{dom}\, \hat{p},\ \ \pi(v) - \pi(u) \in \mathrm{dom}\, \psi_e \text{ for each } e = uv \in A\}. \qquad (7.88)$$

Theorem 7.13. *Assume primal feasibility ($\mathcal{P}_0 \neq \emptyset$) or dual feasibility ($\mathcal{D}_0 \neq \emptyset$). Then we have the min-max relation:*

$$\min\{\Phi(x) : x \in \mathbf{Z}^A \text{ satisfies (7.83)}\} = \max\{\Psi(\pi) : \pi \in \mathbf{Z}^V\}. \tag{7.89}$$

The unbounded case with both sides being equal to $-\infty$ or $+\infty$ is also a possibility. ∎

Remark 7.7. The weak duality $\Phi(x) \geq \Psi(\pi)$ is shown here. Let x and π be primal and dual feasible solutions. Then, using the Fenchel–Young inequality (7.3) for (φ_e, ψ_e) and the feasibility condition (7.83) as well as the expression (4.20) for $\hat{p}(\pi)$, we obtain

$$\begin{aligned}
\Phi(x) - \Psi(\pi) &= \sum_{e=uv \in A} [\varphi_e(x(e)) + \psi_e(\pi(v) - \pi(u))] \ - \hat{p}(\pi) \\
&\geq \sum_{e=uv \in A} x(e)(\pi(v) - \pi(u)) \ - \hat{p}(\pi) \qquad (7.90) \\
&= \sum_{v \in V} \pi(v)(\varrho_x(v) - \delta_x(v)) \ - \hat{p}(\pi) \\
&\geq \min\{\pi z : z \in \ddot{B}\} \ - \hat{p}(\pi) \ = 0. \qquad (7.91)
\end{aligned}$$

This shows the weak duality. The optimality conditions can be obtained as the conditions for the inequalities in (7.90) and (7.91) to be equalities. See Proposition 7.14 below. ∎

In the min-max formula (7.89) we denote the set of the minimizers x by $\mathcal{P}$ and the set of the maximizers π by $\mathcal{D}$.

Proposition 7.14. *Assume that both $\mathcal{P}_0$ and $\mathcal{D}_0$ in* (7.87)–(7.88) *are nonempty.*

(1) *Let $x \in \mathcal{P}_0$ and $\pi \in \mathcal{D}_0$. Then $x \in \mathcal{P}$ and $\pi \in \mathcal{D}$ if and only if the following two conditions are satisfied:*

$$\varphi_e(x(e)) - \varphi_e(x(e) - 1) \leq \pi(v) - \pi(u) \leq \varphi_e(x(e) + 1) - \varphi_e(x(e)) \qquad (e = uv \in A), \tag{7.92}$$

Net-in-flow vector $(\varrho_x(v) - \delta_x(v) : v \in V)$ is a π-minimizer in $\ddot{B}$. (7.93)

(2) *For any $\hat{\pi} \in \mathcal{D}$, we have*

$$\mathcal{P} = \{x \in \mathcal{P}_0 : \text{(7.92) and (7.93) hold with } \pi = \hat{\pi}\}, \tag{7.94}$$

where the condition in (7.92) can be rewritten as

$$x(e) \in \arg\min_k \{\varphi_e(k) - (\pi(v) - \pi(u))k\} \qquad (e = (u, v) \in A). \tag{7.95}$$

(3) For any $\hat{x} \in \mathcal{P}$, we have

$$\mathcal{D} = \{\pi \in \mathcal{D}_0 : \text{(7.92) and (7.93) hold with } x = \hat{x}\}. \tag{7.96}$$

Proof. The inequality (7.90) turns into an equality if and only if, for each $e = uv \in A$, we have $\varphi_e(k) + \psi_e(\ell) = k\ell$ for $k = x(e)$ and $\ell = \pi(v) - \pi(u)$. The latter condition is equivalent to (7.92) by (7.4). The other inequality (7.91) is an equality if and only if (7.93) holds. ∎

In using the min-max relation in Part IV [11] it is convenient to introduce capacity constraints explicitly as

$$f(e) \leq x(e) \leq g(e) \qquad \text{for each edge } e \in A. \tag{7.97}$$

With this explicit form of capacity constraints, a flow x is called a **feasible submodular flow** if it satisfies the conditions (7.80) and (7.97) as well as (7.81). The primal problem reads as follows.

Convex cost submodular flow problem (2):

$$\begin{aligned} \text{Minimize} \quad & \Phi(x) = \sum_{e \in A} \varphi_e(x(e)) & (7.98)\\ \text{subject to} \quad & (\varrho_x(v) - \delta_x(v) : v \in V) \in \ddddot{B}, & (7.99)\\ & f(e) \leq x(e) \leq g(e) \qquad (e \in A), & (7.100)\\ & x(e) \in \mathbf{Z} \qquad (e \in A). & (7.101) \end{aligned}$$

The corresponding dual problem can be derived from (7.85)–(7.86) by the technique described in Remark 7.5. The decision variables of the resulting dual problem consist of an integer-valued potential $\pi : V \to \mathbf{Z}$ on V and integer-valued functions $\tau_1, \tau_2 : A \to \mathbf{Z}$ on A. The constraint (7.103) below says that the tension (potential difference) is split into two parts τ_1 and τ_2.

Dual to the convex cost submodular flow problem (2):

$$\begin{aligned} \text{Maximize} \quad & \Psi(\pi, \tau_1, \tau_2) = \hat{p}(\pi) - \sum_{e \in A} \Big(\psi_e(\tau_1(e)) + \max\{f(e)\tau_2(e),\ g(e)\tau_2(e)\} \Big) & (7.102)\\ \text{subject to} \quad & \pi(v) - \pi(u) = \tau_1(e) + \tau_2(e) \qquad (e = uv \in A), & (7.103)\\ & \pi(v) \in \mathbf{Z} \qquad (v \in V), & (7.104)\\ & \tau_1(e), \tau_2(e) \in \mathbf{Z} \qquad (e \in A). & (7.105) \end{aligned}$$

We introduce notations for feasible flows and potentials/tensions:

$$\mathcal{P}_0 = \{x \in \mathbf{Z}^A : x \text{ satisfies (7.81), (7.99), (7.100) }\}, \tag{7.106}$$

$$\mathcal{D}_0 = \{(\pi, \tau_1, \tau_2) \in \mathbf{Z}^V \times \mathbf{Z}^A \times \mathbf{Z}^A : \text{(7.103)},\ \pi \in \mathrm{dom}\, \hat{p},\ \ \tau_1(e) \in \mathrm{dom}\, \psi_e \text{ for each } \ e \in A\}. \tag{7.107}$$

Theorem 7.15. *Assume primal feasibility ($\mathcal{P}_0 \neq \emptyset$) or dual feasibility ($\mathcal{D}_0 \neq \emptyset$). Then we have the min-max relation:*

$$\begin{aligned} & \min\{\Phi(x) : x \in \mathbf{Z}^A \ \text{ satisfies (7.99) and (7.100)}\} \\ & = \max\{\Psi(\pi, \tau_1, \tau_2) : \pi \in \mathbf{Z}^V \text{ and } \tau_1, \tau_2 \in \mathbf{Z}^A \ \text{ satisfy (7.103)}\}. \end{aligned} \tag{7.108}$$

The unbounded case with both sides being equal to $-\infty$ or $+\infty$ is also a possibility. ■

In the min-max formula (7.108) we denote the set of the minimizers x by $\mathcal{P}$ and the set of the maximizers (π, τ_1, τ_2) by $\mathcal{D}$. The optimality criterion in Proposition 7.14 can be adapted for (7.108) as follows.

Proposition 7.16. *Assume that both $\mathcal{P}_0$ and $\mathcal{D}_0$ in* (7.106)–(7.107) *are nonempty.*

(1) *Let $x \in \mathcal{P}_0$ and $(\pi, \tau_1, \tau_2) \in \mathcal{D}_0$. Then $x \in \mathcal{P}$ and $(\pi, \tau_1, \tau_2) \in \mathcal{D}$ if and only if the following three conditions are satisfied:*

$$\varphi_e(x(e)) - \varphi_e(x(e)-1) \leq \tau_1(e) \leq \varphi_e(x(e)+1) - \varphi_e(x(e)) \qquad (e \in A), \tag{7.109}$$

$$\tau_2(e) \begin{cases} = 0 & \text{if } f(e)+1 \leq x(e) \leq g(e)-1, \\ \leq 0 & \text{if } x(e) = f(e), \\ \geq 0 & \text{if } x(e) = g(e) \end{cases} \qquad (e \in A), \tag{7.110}$$

$$\text{Net-in-flow vector } (\varrho_x(v) - \delta_x(v) : v \in V) \text{ is a } \pi\text{-minimizer in } \ddot{\ddot{B}}. \tag{7.111}$$

(2) *For any $(\hat{\pi}, \hat{\tau}_1, \hat{\tau}_2) \in \mathcal{D}$, we have*

$$\mathcal{P} = \{x \in \mathcal{P}_0 : \text{(7.109), (7.110), (7.111) hold with } (\pi, \tau_1, \tau_2) = (\hat{\pi}, \hat{\tau}_1, \hat{\tau}_2)\}, \tag{7.112}$$

where the conditions in (7.109) *and* (7.110) *can be rewritten as*

$$x(e) \in \arg\min_k \{\varphi_e(k) - \tau_1(e)k\} \qquad (e \in A), \tag{7.113}$$

$$\begin{cases} x(e) = f(e) & \text{if } \tau_2(e) < 0, \\ f(e) \leq x(e) \leq g(e) & \text{if } \tau_2(e) = 0, \\ x(e) = g(e) & \text{if } \tau_2(e) > 0 \end{cases} \qquad (e \in A). \tag{7.114}$$

(3) *For any $\hat{x} \in \mathcal{P}$, we have*

$$\mathcal{D} = \{(\pi, \tau_1, \tau_2) \in \mathcal{D}_0 : \text{(7.109), (7.110), (7.111) hold with } x = \hat{x}\}. \tag{7.115}$$

Proof. Rather than translating the conditions in Proposition 7.14 for the present case, we prove the claim by considering the weak duality $\Phi(x) \geq \Psi(\pi, \tau_1, \tau_2)$ directly for this case.

Let x and (π, τ_1, τ_2) be primal and dual feasible solutions. Then, using the Fenchel–Young inequality (7.3) for (φ_e, ψ_e) and (7.103), we obtain

$$\begin{aligned} &\Phi(x) - \Psi(\pi, \tau_1, \tau_2) \\ &= \sum_{e\in A} [\varphi_e(x(e)) + \psi_e(\tau_1(e))] + \sum_{e \in A} \max\{f(e)\tau_2(e),\ g(e)\tau_2(e)\} \ - \hat{p}(\pi) \\ &\geq \sum_{e \in A} x(e)\tau_1(e) + \sum_{e\in A} \max\{f(e)\tau_2(e),\ g(e)\tau_2(e)\} \ - \hat{p}(\pi) \qquad (7.116) \\ &= \sum_{e=uv\in A} x(e)(\pi(v) - \pi(u) - \tau_2(e)) + \sum_{e \in A} \max\{f(e)\tau_2(e),\ g(e)\tau_2(e)\} \ - \hat{p}(\pi) \\ &= \sum_{e=uv\in A} x(e)(\pi(v) - \pi(u)) + \sum_{e\in A} [\max\{f(e)\tau_2(e),\ g(e)\tau_2(e)\} - x(e)\tau_2(e)] \ - \hat{p}(\pi) \\ &= \left(\sum_{v \in V} \pi(v)(\varrho_x(v) - \delta_x(v)) \ - \hat{p}(\pi) \right) + \sum_{e\in A} [\max\{f(e)\tau_2(e),\ g(e)\tau_2(e)\} - x(e)\tau_2(e)]. \end{aligned}$$

For the former part of this expression we have

$$\sum_{v\in V} \pi(v)(\varrho_x(v) - \delta_x(v)) \ - \hat{p}(\pi) \geq \min\{\pi z : z \in \ddot{\ddot{B}}\} \ - \hat{p}(\pi) = 0 \tag{7.117}$$

by the feasibility condition (7.83) and the expression (4.20) for $\hat{p}(\pi)$, whereas, for each summand in the latter part we have

$$\max\{f(e)\tau_2(e),\ g(e)\tau_2(e)\} - x(e)\tau_2(e) \geq 0 \tag{7.118}$$

since $f(e) \leq x(e) \leq g(e)$ by the capacity constraint (7.100). Thus the weak duality is established. The optimality conditions can be obtained as the conditions for the inequalities in (7.116), (7.117) and (7.118) to be equalities, as in the proof of Proposition 7.14. ■

Remark 7.8. The feasibility of the primal problems can be expressed in terms of the submodular function b as follows, where we denote the integer interval of $\mathrm{dom}\,\varphi_e$ by $[f'(e), g'(e)]_{\mathbf{Z}}$ with $f'(e) \in \mathbf{Z} \cup \{-\infty\}$ and $g'(e) \in \mathbf{Z} \cup \{+\infty\}$. Then there exists a feasible flow for the basic problem (7.82)–(7.84) if and only if

$$\varrho_{f'}(Z) - \delta_{g'}(Z) \leq b(Z) \qquad \text{for all} \quad Z \subseteq V \tag{7.119}$$

is satisfied. For the problem (7.98)–(7.101) with explicit capacity constraints, we replace $f'(e)$ and $g'(e)$ by $\max\{f(e), f'(e)\}$ and $\min\{g(e), g'(e)\}$, respectively. ■

Acknowledgement We thank S. Fujishige and S. Iwata for discussion about the history of convex minimization over base-polyhedra. Thanks are also due to A. Shioura and A. Tamura for reading the manuscript and S. Moriguchi for checking examples. Insightful questions posed by Tamás Király led to Remark 5.4 and Theorem 6.7. This research was supported through the program "Research in Pairs" by the Mathematisches Forschungsinstitut Oberwolfach in 2019. The two weeks we could spend at Oberwolfach provided an exceptional opportunity to conduct particularly intensive research.

A Survey of early papers

This appendix offers a brief survey of earlier papers and books that deal with topics closely related to decreasing minimization on base-polyhdera. To be specific, we mention the following: Veinott [41] (1971), Megiddo [29] (1974), Fujishige [13] (1980), Groenevelt [17] (1985, 1991), Federgruen–Groenevelt [7] (1986), Ibaraki–Katoh [22] (1988), Dutta–Ray [6] (1989), Fujishige [14] (1991, 2005), Hochbaum [19] (1994), and Tamir [40] (1995).

Similar notions and terms are scattered in the literature such as "egalitarian," "lexicographically optimal," "least majorized," "least weakly submajorized," "decreasingly minimal (dec-min)," and "increasingly maximal (inc-max)." Unfortunately, these notions are discussed often independently in different context, without proper mutual recognition. The term "least majorized" is used in Veinott [41] and "Least weakly submajorized" is used in Tamir [40]. These terms are not used in Marshall–Olkin–Arnold [28]. Dutta–Ray [6] uses "egalitarian" and does not use "majorization." The term "lexicographically optimal" in Veinott [41], Megiddo [29, 30], and Fujishige [13, 14] means "increasingly maximal (inc-max)."

Three notions "dec-min", "inc-max", and "least majorized" are different in general. Generally, "least majorized" implies "dec-min" and "inc-max", but the converse is not true (see

Section 2.2). In base-polyhedron (in **R** and **Z**), however, the three notions coincide (see Section 2.3).

Another important aspect in majorization is minimization of symmetric separable convex functions. An element is least majorized if and only if it simultaneously minimizes all symmetric separable convex functions (see Proposition 2.1). Therefore, if a least majorized is known to exist, then it can be computed as a minimizer of the square-sum.

Veinott (1971) [41]

This paper deals with a network flow problem. The ground set is a star of arcs, i.e., the set of arcs incident to a single node. This amounts to considering a special case of a base-polyhedron. The main result is the unique existence of a least majorized element in Case **R**.

The computational aspect is also discussed. The problem is reduced to separable quadratic network flow problem. Then the paper describes an algorithm for nonlinear convex cost minimum flow problem. It also defines the dual problem using the conjugate function. Complexity of the algorithm is not discussed.

Case **Z** is also treated. Theorem 2 (1) shows the existence of an integral element that simultaneously minimizes all symmetric separable convex functions. The proof is based on rounding argument (continuous relaxation). That is, for a discrete convex function in integers, its piecewise-linear extension is considered and the integrality theorem is used to derive the existence of an integral minimizer. Thus the existence of a least majorized element is shown for the network flow in Case **Z**.

Megiddo (1974) [29]

This paper deals with a network flow problem. The ground set is the set of multi-terminals. This is more general than a star considered in Veinott [41], but the difference is not really essential. The paper defines the notions of "sink-optimality" and "source-optimality," which are increasing-maximality for vectors on the sink and source terminals, respectively. This paper considers Case **R** only. The main result is the characterization of an inc-max element using a chain of cuts in the network (Theorem 4.6). The computational aspect is discussed in the companion paper [30], which gives an algorithm of complexity $O(n^5)$.

Fujishige (1980) [13]

This is the first paper that deals with base-polyhedra, beyond network flows. It considers Case **R** only. Lexicographic optimality with respect to a weight vector is defined. The lexicographically optimal base with respect to a uniform weight coincides with the inc-max element of the base-polyhedron. The relation to weighted square-sum minimization is investigated in detail and the minimum norm point is highlighted. The principal partition for base-polyhedra is introduced, as a generalization of the known construction for matroids. The principal partition determines the lexico-optimal base. The proposed decomposition algorithm finds the lexico-optimal base as well as the principal partition in strongly polynomial time. While this paper covers various aspects of the lexico-optimal base, the majorization viewpoint is missing. In particular, it is not stated that the minimum norm point is actually a minimizer of all symmetric separable convex functions.

Groenevelt (1985, 1991) [17]

The technical report appeared in 1985, and the journal version in 1991. Already the technical report was influential, cited by [14, 1st ed.], [19], and [22].

The main concern of this paper is separable convex minimization (not restricted to symmetric separable convex functions) on base-polyhedra. Both continuous variables (Case **R**) and discrete variables (Case **Z**) are treated. In particular, this is the first paper that addressed minimization of separable convex functions on base polyhedra in discrete variables. One of the results says that, in any integral base-polyhedron, there exists an integral element that is a (simultaneous) minimizer of all symmetric separable convex functions. This paper does not discuss implications of this result to inc-maximality, dec-minimality, or majorization, though the result does imply the existence of a least majorized element by virtue of the well-known fact (Proposition 2.1) about majorization.

The paper presents two kinds of algorithms, the marginal allocation algorithm (of incremental type) and the decomposition algorithm (DA). Concerning complexity, the author argues that the algorithms are polynomial if the base-polyhedron are of some special types (tree-structured polymatroids, generalized symmetric polymatroids, network polymatroids). We quote the following statements from [17, p.234, journal version], where E denotes the ground set of a base-polyhedron and N is the associated submodular function, which is integer-valued in Case **Z**:

> The total complexity of DA is thus $\mathrm{O}(|E|(\tau_1 + \tau_2))$, where τ_1 = the number of operations needed to solve a single constraint problem, and τ_2 = the number of operations needed to perform one pass through Steps 2 and 3. It is well-known that in the discrete case $\tau_1 = \mathrm{O}(|E| \log(N(E)/|E|))$ (see Frederickson and Johnson (1982)), and in the continuous case $\tau_1 = \mathrm{O}(|E| \log |E| + \chi)$, where χ is the time needed to solve a certain type of non-linear equation (see Zipkin, 1980).

This paper was written in 1985 and at that time, no strongly polynomial algorithm for submodular function minimization was known; the strongly polynomial algorithm (using the ellipsoid method) first appeared in 1993 [18, 2nd edition].

Federgruen–Groenevelt (1986) [7]

This paper deals with base-polyhedra in Case **Z**. Main concern of this paper is to offer a general framework in which a greedy procedure called the marginal allocation algorithm (MAA) works. The concept of concave order is introduced as a class of admissible objective functions for which the greedy procedure works. The main result (Corollary 1 in Sec.3) states, roughly, that the MAA gives an optimal solution for every weakly concave order on polymatroids.

Ibaraki–Katoh (1988) [22]

This is the first comprehensive book for algorithmic aspects of the resource allocation problem and its extensions. Chapter 9, entitled "Resource allocation problems under submodular constrains" presents the fundamental and up-to-date results at that time, including those by Fujishige [13], Groenevelt [17], and Federgruen–Groenevelt [7]. In particular, Theorem 9.2.2 [22, p.156] states that the decomposition algorithm runs in polynomial time in $|E|$ and $\log M$, where E is the ground set and M is an upper bound on $r(E)$ for the submodular function r expressing the submodular constraint.

The contents of Chapter 9 of this book are updated in a handbook chapter by Ibaraki–Katoh [25] in 1998. Its revised version by Katoh–Shioura–Ibaraki [26] in 2013 incorporates the views from discrete convex analysis.

Dutta–Ray (1989) [6]

This paper deals with base-polyhedra in the context of game theory. Recall that the core of a convex game is nothing but the base-polyhedron. Naturally this paper deals exclusively with Case **R**. According to Tamir [40], this is the first paper proving the existence of a least majorized element in a base-polyhedron. Technically speaking, this result could be obtained from a combination of the results of Groenevelt [17] (which was written in 1985 and published in 1991) and a well-known fact "least majorized element ⇔ simultaneous minimizer of all symmetric separable convex functions" (see Proposition 2.1). However, Dutta–Ray [6] and Groenevelt [17] were unaware of each other; see Table 1 at the end of Appendix. We also note that Fujishige [13] deals with quadratic functions only, and hence the results of [13] do not imply the existence of a least majorized element.

Fujishige (1st ed., 1991; 2nd ed. 2005) [14]

This book offers a comprehensive exposition of the results of Fujishige [13] about the lexico-optimal (inc-max) element of a base-polyhedron in Case **R**. There is an explicit statement at the beginning of Section 9 that the argument is not applicable to Case **Z**.

For separable convex minimization, both Cases **R** and **Z** are treated. In particular, the results of Groenevelt [17] are described in a manner consistent with the other part of this book. It is stated that the decomposition algorithm works for Cases **R** and **Z**, but complexity analysis is explicit only for Case **R**. It is shown that the decomposition algorithm is strong polynomial for Case **R**. As a natural consequence of the fact that lexico-optimal bases in Case **Z** are not considered in this book, no connection is made between separable convex minimization and lexico-optimality (inc-max, dec-min) in Case **Z**.

Majorization concept is not treated in the first edition, whereas in the second edition the definition is given in Section 2.3 (p. 44) and a reference to Dutta–Ray [6] is added in Section 9.2 (p. 264).

Hochbaum (1994) [19]

This paper shows that there exist no strongly polynomial time algorithms to solve the resource allocation problem with a separable convex cost function. Subsequently, Hochbaum and her coworkers made significant contributions to resource allocation problems in discrete variables, dealing with important special cases and showing improved complexity bounds for the special cases (e.g., Hochbaum–Hong [21]). The survey paper by Hochbaum [20] is informative and useful.

Tamir (1995) [40]

This papers deals with g-polymatroids in Case **R** and Case **Z**. The relationship between majorization and decreasing-minimality is discussed explicitly.

The main result is the existence of a least weakly submajorized element in a g-polymatroid. The following sentences concerning Case **R** in pages 585–585 are informative:

> Fujishige (1980) extends the results of Megiddo to a general polymatroid and presents an algorithm to find a lexicographically optimal base of the polymatroid with respect to an arbitrary positive weight vector d. This weighted model is closely related to the concept of d-majorization introduced by Veinott (1971). Neither Megiddo nor Fujishige relate their results on lexicographically optimal bases to the stronger concept of majorization. (From Proposition 2.1 we note that if an arbitrary set has a least majorized element it is clearly lexicographically optimal. However, every convex and compact set S has a unique lexicographically maximum element, but might not have a least majorized element.) The fact that a polymatroid has a least majorized base is shown by Dutta and Ray (1989). They consider the core of a convex game as defined by Shapley (1971), which corresponds to a polymatroid. (Strictly speaking the former is defined as a contra-polymatroid; see next section.) We will extend and unify the above results by proving that a bounded generalized polymatroid contains both least submajorized and least supermajorized elements.

For the complexity of finding the unique minimizer $x^* \in \mathbf{R}^n$ of the square-sum over a g-polymatroid (Case **R**), the following statement can be found in page 587:

> x^* can be found in strongly polynomial time by modifying the procedure in Fujishige (1980) and Groenevelt (1991) which is applicable to polymatroids. The latter procedure can now be implemented to solve any convex separable quadratic over a polymatroid in a strongly polynomial time since its complexity is dominated by the efforts to minimize a (strongly) polynomial number of submodular functions.

There is no statement about the complexity in Case **Z**.

Table 1: Referencing relations between papers

	Vei [41]	Meg [29]	Fuj [13]	Gro [17]	F-G [7]	I-K [22]	D-R [6]	Fuj [14]	Hoc [19]	Tam [40]
Veinott 1971	·	–	–	–	–	–	–	–	–	–
Megiddo 1974	–	·	–	–	–	–	–	–	–	–
Fujishige 1980	–	R	·	–	–	–	–	–	–	–
Groenevelt 1985/91	–	R	R	·	R	–	–	–	–	–
Federgruen–Groenevelt 1986	–	R	R	–	·	–	–	–	–	–
Ibaraki–Katoh 1988	–	R	R	R	R	·	–	–	–	–
Dutta–Ray 1989	–	–	–	–	–	–	·	–	–	–
Fujishige 1991 (1st ed.)	–	R	R	R	–	R	R^{2nd}	·	–	–
Hochbaum 1994	–	–	–	R	R	R	–	–	·	–
Tamir 1995	R	R	R	R	–	–	R	R	–	·

Paper at the left refers to papers marked R in the same row
R^{2nd} means that reference is made in the 2nd edition (2005) only

References

[1] Ahuja, R.K., Hochbaum, D.S., Orlin, J.B.: Solving the convex cost integer dual network flow problem. Management Science **49**, 950–964 (2003)

[2] Ando, K.: Weak majorizations on finite jump systems. Mimeo (1996) Available from author's home page.

[3] Ando, K., Fujishige, S., Naitoh, T.: A greedy algorithm for minimizing a separable convex function over a finite jump system. Journal of the Operations Research Society of Japan **38**, 362–375 (1995)

[4] Arnold, B.C., Sarabia, J.M.: Majorization and the Lorenz Order with Applications in Applied Mathematics and Economics. Springer International Publishing, Cham (2018), (1st edn., 1987)

[5] Dutta, B.: The egalitarian solution and reduced game properties in convex games. International Journal of Game Theory **19**, 153–169 (1990)

[6] Dutta, B., Ray, D.: A concept of egalitarianism under participation constraints. Econometrica **57**, 615–635 (1989)

[7] Federgruen, A., Groenevelt, H.: The greedy procedure for resource allocation problems: necessary and sufficient conditions for optimality. Operations Research **34**, 909–918 (1986)

[8] Frank, A.: Connections in Combinatorial Optimization. Oxford University Press, Oxford (2011)

[9] Frank, A., Murota, K.: Discrete decreasing minimization, Part I: Base-polyhedra with applications in network optimization. arXiv: 1808.07600 (August 2018)

[10] Frank, A., Murota, K.: Discrete decreasing minimization, Part III: Network flows, arXiv: 1907.02673 (July 2019)

[11] Frank, A., Murota, K.: Discrete decreasing minimization, Part IV: Submodular flows and the intersection of two base-polyhedra, in preparation

[12] Frank, A., Murota, K.: Discrete decreasing minimization, Part V: Weighted cases, in preparation

[13] Fujishige, S.: Lexicographically optimal base of a polymatroid with respect to a weight vector. Mathematics of Operations Research **5**, 186–196 (1980)

[14] Fujishige, S.: Submodular Functions and Optimization, 1st edn. Annals of Discrete Mathematics **47**, North-Holland, Amsterdam (1991); 2nd edn. Annals of Discrete Mathematics **58**, Elsevier, Amsterdam (2005)

[15] Fujishige, S.: Theory of principal partitions revisited. In: Cook, W., Lovász, L., Vygen, J. (eds.) Research Trends in Combinatorial Optimization, pp. 127–162. Springer, Berlin (2009)

[16] Fujishige, S., Katoh, N., Ichimori, T.: The fair resource allocation problem with submodular constraints. Mathematics of Operations Research **13**, 164–173 (1988)

[17] Groenevelt, H.: Two algorithms for maximizing a separable concave function over a polymatroid feasible region. European Journal of Operational Research **54**, 227–236 (1991); The technical version appeared as Working Paper Series No. QM 8532, Graduate School of Management, University of Rochester (1985)

[18] Grötschel, M., Lovász, L., Schrijver, A.: Geometric Algorithms and Combinatorial Optimization, 2nd edn. Springer, Berlin (1993)

[19] Hochbaum, D.S.: Lower and upper bounds for the allocation problem and other nonlinear optimization problems. Mathematics of Operations Research **19**, 390–409 (1994)

[20] Hochbaum, D.S.: Complexity and algorithms for nonlinear optimization problems. Annals of Operations Research **153**, 257–296 (2007)

[21] Hochbaum, D.S., Hong, S.-P.: About strongly polynomial time algorithms for quadratic optimization over submodular constraints. Mathematical Programming **69**, 269–309 (1995)

[22] Ibaraki, T., Katoh, N.: Resource Allocation Problems: Algorithmic Approaches. MIT Press, Boston (1988)

[23] Iri, M.: Network Flow, Transportation and Scheduling—Theory and Algorithms. Academic Press, New York (1969)

[24] Iri, M. : A review of recent work in Japan on principal partitions of matroids and their applications. Annals of the New York Academy of Sciences **319**, 306–319 (1979)

[25] Katoh, N., Ibaraki, T.: Resource allocation problems. In: Du, D.-Z., Pardalos, P.M. (eds.) Handbook of Combinatorial Optimization, Vol.2, pp. 159–260. Kluwer Academic Publishers, Boston (1998)

[26] Katoh, N., Shioura, A., Ibaraki, T.: Resource allocation problems. In: Pardalos, P.M., Du, D.-Z., Graham, R.L. (eds.) Handbook of Combinatorial Optimization, 2nd ed., Vol. 5, pp. 2897-2988, Springer, Berlin (2013)

[27] Levin, A., Onn, S.: Shifted matroid optimization. Operations Research Letters **44**, 535–539 (2016)

[28] Marshall, A.W., Olkin, I., Arnold, B.C.: Inequalities: Theory of Majorization and Its Applications, 2nd edn. Springer, New York (2011), (1st edn., 1979)

[29] Megiddo, N.: Optimal flows in networks with multiple sources and sinks. Mathematical Programming **7**, 97–107 (1974)

[30] Megiddo, N.: A good algorithm for lexicographically optimal flows in multi-terminal networks. Bulletin of the American Mathematical Society **83**, 407–409 (1977)

[31] Moriguchi, S., Shioura, A., Tsuchimura, N.: M-convex function minimization by continuous relaxation approach—Proximity theorem and algorithm. SIAM Journal on Optimization **21**, 633–668 (2011)

[32] Murota, K.: Convexity and Steinitz's exchange property. Advances in Mathematics **124**, 272–311 (1996)

[33] Murota, K.: Discrete convex analysis. Mathematical Programming **83**, 313–371 (1998)

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

[35] Murota, K.: Discrete Convex Analysis. Society for Industrial and Applied Mathematics, Philadelphia (2003)

[36] Murota, K.: Recent developments in discrete convex analysis. In: Cook, W., Lovász, L., Vygen, J. (eds.) Research Trends in Combinatorial Optimization, Chapter 11, pp. 219–260. Springer, Berlin (2009)

[37] Onn, S.: Nonlinear Discrete Optimization: An Algorithmic Theory. European Mathematical Society, Zurich (2010)

[38] Rockafellar, R.T.: Network Flows and Monotropic Optimization. Wiley, New York (1984)

[39] Schrijver, A.: Combinatorial Optimization—Polyhedra and Efficiency. Springer, Heidelberg (2003)

[40] Tamir, A.: Least majorized elements and generalized polymatroids. Mathematics of Operations Research **20**, 583–589 (1995)

[41] Veinott, Jr., A.F.: Least d-majorized network flows with inventory and statistical applications. Management Science **17**, 547–567 (1971)