Egerváry Research Group
on Combinatorial Optimization

Technical reportS

TR-2018-14. 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

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

András Frank⋆ and Kazuo Murota⋆⋆

**Abstract**

We consider discrete decreasing minimization problem on an integral base-polyhedron. 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.

In contrast to the constructive and algorithmic approach in Part I, Part II offers structural views from discrete convex analysis (DCA) by making full use of the fundamental results on M-convex sets and M-convex functions. 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 duality in DCA; this approach also yields a novel min-max formula that shows a natural link to majorization. A general result on the Fenchel-type 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 with a tight bound, which leads to a continuous relaxation algorithm for finding a decreasingly minimal element of an M-convex set. Thus the relationship between the continuous and discrete cases is completely clarified.

**Keywords:** base-polyhedra, discrete convex analysis, lexicographically optimal, majorization, min-max formulas, principal partition

⋆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 Hungarian Scientific Research Fund - OTKA, No K109240.

⋆⋆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.

# 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 [10] around 1980 under the name of lexicographically optimal bases of a base-polyhedron, as a generalization of lexicographically optimal maximal flows considered by Megiddo [27]. Our study can be regarded as a discrete counter-part of the work by Fujishige [10], [11, Section 9] on the lexicographically optimal base and the principal partition of a base-polyhedron.

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.

In Section 2 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 duality in DCA; this approach also yields a novel min-max formula that shows a natural link to majorization. In Section 5 we use a general result on the Fenchel-type 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 with a tight bound, which leads to a continuous relaxation algorithm for finding a decreasingly minimal element of an M-convex set. In Appendix A we show an additional property of a dec-min element that it minimizes non-separable symmetric convex functions. In Appendix B 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 [8].

### 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 $y{\downarrow} = (5,5,5,1,1)$.

A vector $x$ is **decreasingly smaller than or equal to** vector $y$, in notation $x \leq_{\rm dec} y$, if they are either value-equivalent or $x <_{\rm dec} y$. 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. Therefore 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 paper, 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 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 $\chi_Z$, that is, $\chi_Z(v) = 1$ if $v \in Z$ and $\chi_Z(v) = 0$ otherwise. For a polyhedron $B$, notation $\ddot{B}$ (pronounced: dotted $B$) means the set of integral members (elements, vectors, points) of $B$.

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

$$b(X) + b(Y) \geq b(X \cap Y) + b(X \cup Y).$$

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

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

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:

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

We call the set $\ddot{B}$ of integral elements of an integral base-polyhedron $B$ an **M-convex set**. Originally, this basic notion of discrete convex analysis is defined as a set of integral points in $\mathbf{R}^S$ satisfying certain exchange axioms, and it is known that the two properties are equivalent ([32],Theorem 4.15).

### Discrete convex functions

For a function $\varphi : \mathbf{Z} \to \mathbf{R} \cup \{+\infty\}$ the **effective domain** of $\varphi$ is denoted as $\mathrm{dom}\,\varphi = \{k \in \mathbf{Z} : \varphi(k) < +\infty\}$. A function $\varphi : \mathbf{Z} \to \mathbf{R} \cup \{+\infty\}$ is called **discrete convex** (or simply **convex**) if $\varphi(k-1)+\varphi(k+1) \geq 2\varphi(k)$ for all $k \in \mathrm{dom}\,\varphi$, and **strictly convex** if $\varphi(k-1)+\varphi(k+1) > 2\varphi(k)$ for all $k \in \mathrm{dom}\,\varphi$.

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

$$\Phi(x) = \sum[\varphi_s(x(s)) : s \in S]$$

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 $\Phi$ a **symmetric separable convex function** if $\varphi_s$ does not depend on $s$, that is, if $\varphi_s = \varphi$ for all $s \in S$ for some discrete convex function $\varphi$. We call $\Phi$ a **symmetric separable strictly convex function** if $\varphi$ 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 [3] and

Marshall–Olkin–Arnold [25]. 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 [8]. 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 [14] 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 of 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 $j$-th component $\overline{x}(j)$ is equal to the sum of the first $j$ 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 [25, p.13] that

$$x \prec y \iff -x \prec -y. \tag{1}$$

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 [25, 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,*

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

(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* $(z)^{+} = \max(z, 0)$ *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 3.3 of Part I [8]). •

**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 [35]. A vector $x$ is said to be **weakly submajorized** by another vector $y$ if $\overline{x} \leq \overline{y}$, with the standard notation $x \prec_{\mathrm{w}} y$. 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 (33) in Section 4.2 when the underlying set $D$ is an M-convex set (= the integer points of an integral base-polyhedron), and by (53) in Section 4.8 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 in Theorem 4.6 in Section 4.6. •

## 2.2 Majorization and decreasing-minimality

Majorization and decreasing-minimality are closely related, as is explicit in Tamir [35].

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

*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 $x \leq_{\mathrm{dec}} y$. Since $x \prec y$, we have $-x \prec -y$ by (1). By the above argument applied to $(-x, -y)$, we obtain $-x \leq_{\mathrm{dec}} -y$, which is equivalent to $x \geq_{\mathrm{inc}} y$. □

**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 $x \not\prec y$. •

**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 (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 [14], which is already mentioned in Section 6 of Part I [8].

**Proposition 2.5** (Groenevelt [14]; cf. [11, Theorem 8.1])**.** *Let $B$ be an integral base-polyhedron, $\ddot{B}$ 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 $\ddot{B}$ 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 $m + \chi_s - \chi_t \in \ddot{B}$.* •

Theorem 2.4 can be derived from the combination of Proposition 2.5 with Proposition 2.1. Let $m \in \ddot{B}$ be a minimizer of the square-sum $\sum[x(s)^2 : s \in S]$ over $\ddot{B}$; 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 $m + \chi_s - \chi_t \in \ddot{B}$. 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 [8] we have shown, without using the terminology of majorization, 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 Theorem 2.4. •

**Remark 2.5.** A variant of majorization concept, "weak submajorization" (cf., Remark 2.1), is investigated for integral g-polymatroids by Tamir [35] 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 [8]. More specifically, we derive these results from the optimality criterion for M-convex functions, which is described in Theorem 3.6 in Section 3.2.

**THEOREM 3.1** ([8, 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 $m$.* •

**Proposition 3.2** ([8, Proposition 6.1]). *Let $\Phi(x) = \sum[\varphi(x(s)) : s \in S]$ be a symmetric separable convex function with $\varphi : \mathbf{Z} \to \mathbf{R}$. Each dec-min element of $\overset{....}{B}$ is a minimizer of $\Phi$.* •

**THEOREM 3.3** ([8, Theorem 6.2]). *Let $\Phi(x) = \sum[\varphi(x(s)) : s \in S]$ be a symmetric separable strictly convex function with $\varphi : \mathbf{Z} \to \mathbf{R}$. An element $m$ of $\overset{....}{B}$ is a minimizer of $\Phi$ if and only if $m$ is a dec-min element of $\overset{....}{B}$.* •

It should be clear in the above that $\overset{....}{B}$ denotes an M-convex set (the set of integral points of an integral base-polyhedron), and a **1-tightening step** for $m \in \overset{....}{B}$ 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 $m+\chi_s-\chi_t \in \overset{....}{B}$.

## 3.1 Convex formulation of decreasing minimality

As already mentioned at the end of Section 6.1 of Part I [8], a dec-min element can be characterized as a minimizer of 'rapidly' increasing convex function. This characterization enables us to make use of DCA for the analysis of decreasing minimality.

We say that $\varphi : \mathbf{Z} \to \mathbf{R}$ is **rapidly increasing** if

$$\cdots \ll \varphi(k-1) \ll \varphi(k) \ll \varphi(k+1) \ll \cdots , \tag{2}$$

from which the strict convexity of $\varphi$ follows. We can formulate a quantitative version of this notion as

$$\varphi(k+1) \geq (n+1)\, \varphi(k) > 0 \qquad (k \in \mathbf{Z}), \tag{3}$$

where $n = |S|$. For example, $\varphi(k) = (n+1)^k$ $(k \in \mathbf{Z})$ satisfies this condition. As is easily expected, $x <_{\rm dec} y$ is equivalent to $\Phi(x) < \Phi(y)$.

**Proposition 3.4.** *Assume rapid increase* (3)*. 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.* (The proof is straightforward and easy.) 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

$$\Phi(x) = \sum_k \Theta(x, k)\varphi(k). \tag{4}$$

Obviously, $\Phi(x) = \Phi(y)$ if and only if $x$ is value-equivalent to $y$.

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 $\Theta(x, \hat{k}) < \Theta(y, \hat{k})$. On the other hand, we have

$$\begin{aligned}\Phi(x) - \Phi(y) &= \sum_k (\Theta(x, k) - \Theta(y, k))\varphi(k) \\ &= (\Theta(x, \hat{k}) - \Theta(y, \hat{k}))\varphi(\hat{k}) + \sum_{k<\hat{k}} (\Theta(x, k) - \Theta(y, k))\varphi(k),\end{aligned}$$

where

$$0 \leq \sum_{k<\hat{k}} \Theta(x, k)\varphi(k) \leq \varphi(\hat{k}-1) \sum_{k<\hat{k}} \Theta(x, k) \leq n\varphi(\hat{k}-1) \leq \frac{n}{n+1}\varphi(\hat{k})$$

and similarly for $y$. Therefore we have $|\sum_{k<\hat{k}}(\Theta(x, k) - \Theta(y, k))\varphi(k)| \leq \frac{n}{n+1}\varphi(\hat{k})$, and hence

$$(\Theta(x, \hat{k}) - \Theta(y, \hat{k}) - \frac{n}{n+1})\varphi(\hat{k}) \ \leq\ \Phi(x) - \Phi(y) \ \leq\ (\Theta(x, \hat{k}) - \Theta(y, \hat{k}) + \frac{n}{n+1})\varphi(\hat{k}).$$

This shows that $\Phi(x) < \Phi(y)$ if and only if $\Theta(x, \hat{k}) < \Theta(y, \hat{k})$. □

By Proposition 3.4 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 3.5.** *Assume rapid increase* (3) *and let $D$ be an arbitrary subset of* $\mathbf{Z}^S$*. 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 3.5 is not particularly new. Similar ideas are scattered in the literature of related topics such as majorization (Marshall–Olkin–Arnold [25]) and shifted optimization (Levin–Onn [24]). •

**Remark 3.2.** The relations of being majorized ($\prec$), weakly submajorized ($\prec_{\mathrm{w}}$), and decreasingly-smaller ($\leq_{\mathrm{dec}}$) are characterized with reference to different classes of symmetric separable convex functions as follows (Proposition 2.1, [25, 4.B.2], and Proposition 3.4):

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

## 3.2 M-convex function minimization in discrete convex analysis

In this section we introduce a fundamental concept in DCA, M-convex functions, 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 Theorem 3.1, Proposition 3.2, and Theorem 3.3 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

$$\mathrm{supp}^+(z) = \{s \in S : z(s) > 0\}, \qquad \mathrm{supp}^-(z) = \{t \in S : z(t) < 0\}. \tag{5}$$

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

$$f(x) + f(y) \geq f(x - \chi_s + \chi_t) + f(y + \chi_s - \chi_t). \tag{6}$$

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 $y \notin \mathrm{dom} f$, (6) trivially holds with $f(x) + f(y) = +\infty$. 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.

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)

$$f(x) + f(y) \geq f(x - \chi_s) + f(y + \chi_s) \tag{7}$$

or (ii) there exists some $t \in \mathrm{supp}^-(x - y)$ for which (6) 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 (M$^\natural$-convex set). 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 3.6** ([32, 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*

$$f(x^*) \le f(x^* + \chi_s - \chi_t) \quad \text{for all } s, t \in S, \tag{8}$$

$$f(x^*) \le f(x^* + \chi_s) \quad \text{for all } s \in S, \tag{9}$$

$$f(x^*) \le f(x^* - \chi_t) \quad \text{for all } t \in S. \tag{10}$$

*If $f$ is M-convex, $x^*$ is a minimizer of $f$ if and only if* (8) *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

$$\Phi(x) = \sum [\varphi_s(x(s)) : s \in S], \tag{11}$$

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

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

$$\text{Minimize } \ \Phi(x) = \sum [\varphi_s(x(s)) : s \in S] \ \text{ subject to } \ x \in \ddot{B}. \tag{12}$$

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

$$\delta(x) = \begin{cases} 0 & (x \in \ddot{B}), \\ +\infty & (\text{otherwise}), \end{cases} \tag{13}$$

we can rewrite (12) as

$$\text{Minimize } \ \Phi(x) + \delta(x). \tag{14}$$

This problem is amenable to discrete convex analysis, since the separable convex function $\Phi$ is M$^\natural$-convex, the indicator function $\delta$ 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 4.1 in Section 4.1).

An application of the M-optimality criterion (Theorem 3.6) to our function $\Phi + \delta$ gives the following important result due to Groenevelt [14], which was shown as Proposition 2.5 and stated again for its relevance here.

**Proposition 3.7** (Groenevelt [14]; cf. [11, Theorem 8.1]). *Let $B$ be an integral base-polyhedron and $\ddot{B}$ be the set of its integral elements. An element $m$ of $\ddot{B}$ is a minimizer of $\Phi(x) = \sum[\varphi_s(x(s)) : s \in S]$ over $\ddot{B}$ if and only if $\varphi_s(m(s)+1) + \varphi_t(m(t)-1) \geq \varphi_s(m(s)) + \varphi_t(m(t))$ whenever $m + \chi_s - \chi_t \in \ddot{B}$.* •

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 the 1-tightening step introduced in Part I [8]. Recall that 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 $m + \chi_s - \chi_t \in \ddot{B}$.

**Proposition 3.8.** *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 $\ddot{B}$ is a minimizer of $\Phi$ over $\ddot{B}$ if there is no 1-tightening step for $m$. The converse is also true if $\varphi$ is strictly convex.*

*Proof.* By Proposition 3.7, $m$ is a minimizer of $\Phi$ 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 $m + \chi_s - \chi_t \in \ddot{B}$. By the convexity of $\varphi$, we have this inequality if $m(t) \leq m(s) + 1$, and the converse is also true when $\varphi$ 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 $m + \chi_s - \chi_t \in \ddot{B}$. □

## 3.4 DCA-based proofs of the theorems

The combination of Proposition 3.8 with Proposition 3.5 provides alternative proofs of Theorem 3.1, Proposition 3.2, and Theorem 3.3.

**Proof of Theorem 3.1:** Let $\Phi$ be a symmetric separable convex function with rapidly increasing $\varphi$. By Proposition 3.5, $m$ is dec-min if and only if $m$ is a minimizer of $\Phi$. On the other hand, since $\Phi$ is strictly convex, Proposition 3.8 shows that $m$ is a minimizer of $\Phi$ if and only if there is no 1-tightening step for $m$. Therefore, $m$ is a dec-min element of $\ddot{B}$ if and only if there is no 1-tightening step for $m$.

**Proof of Proposition 3.2 and Theorem 3.3:** Let $\Phi$ be a symmetric separable convex function. By Proposition 3.8, $m$ is a minimizer of $\Phi$ if there is no 1-tightening step for $m$; and the converse is also true for strictly convex $\Phi$. 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 $\Phi$ if $m$ is a dec-min element of $\ddot{B}$; and the converse is also true for strictly convex $\Phi$.

## 3.5 Extension to generalized polymatroids

In this section we shed the light of DCA on the result of Tamir [35] 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 Part I [8] for the definition of g-polymatroids and [32] for more about M$^\natural$-convexity.

The M-optimality criterion (Theorem 3.6) immediately implies the following generalization of Proposition 3.7.

**Proposition 3.9.** *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 3.8 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 $\varphi$. 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 3.10.** *Let $\Phi$ be a symmetric separable convex function represented as $\Phi(x) = \sum[\varphi(x(s)) : s \in S]$ with monotone non-decreasing discrete convex $\varphi$. An element $m$ of $\overset{....}{Q}$ is a minimizer of $\Phi$ 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 $\varphi$ is strictly convex and strictly monotone increasing.* •

On the basis of Proposition 3.10 we can show the existence of a least weakly submajorized element in $\overset{....}{Q}$, which is the result of Tamir [35].

# 4 Min-max formulas

We derive the following formulas, established by constructive methods in Part I [8], from the Fenchel-type duality in discrete convex analysis. Recall that $p$ is an integer-valued supermodular function on the ground-set $S$, $B$ 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)[1] of the function $p$, i.e., $\hat{p}(\pi) = \min\{\pi x : x \in \overset{....}{B}\}$.

- [8, Theorem 6.3] For the square-sum we have

$$\min\{\sum_{s\in S} m(s)^2 : m \in \overset{....}{B}\} = \max\{\hat{p}(\pi) - \sum_{s\in S} \left\lfloor \frac{\pi(s)}{2} \right\rfloor \left\lceil \frac{\pi(s)}{2} \right\rceil : \pi \in \mathbf{Z}^S\}. \quad (15)$$

[1] See Section 6.2 of Part I [8] for the linear extension $\hat{p}$.

- [8, Theorem 4.1] For the largest component $\beta_1$ of a max-minimizer of $\overset{....}{B}$, we have

$$\beta_1 = \max\{\left\lceil \frac{p(X)}{|X|} \right\rceil : \emptyset \neq X \subseteq S\}. \tag{16}$$

  Recall that $\beta_1$ is equal to the largest component $m(s)$ of any dec-min element $m$ of $\overset{....}{B}$.

- [8, Theorem 4.3] For the minimum number $r_1$ of $\beta_1$-valued components of a $\beta_1$-covered member of $\overset{....}{B}$, we have

$$r_1 = \max\{p(X) - (\beta_1 - 1)|X| : X \subseteq S\}. \tag{17}$$

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

Moreover, the following new formula will be established in Section 4.6 also from the Fenchel-type duality in DCA.

- For each integer $a$, we have

$$\min\{\sum_{s \in S}(m(s) - a)^+ : m \in \overset{....}{B}\} = \max\{p(X) - a|X| : X \subseteq S\}. \tag{18}$$

This formula (18) for $a = \beta_1 - 1$ coincides with the formula (17) for $r_1$. Thus (18) generalizes (17). It will be shown in Theorem 4.6 that an element of $\overset{....}{B}$ is decreasingly minimal if and only if it is a minimizer of the left-hand side of (18) universally for all $a \in \mathbf{Z}$. Recall that we have encountered the expression $\sum[(x(s) - a)^+ : s \in S]$ in Proposition 2.1 about majorization.

## 4.1 Fenchel-type duality in discrete convex analysis

In this section we describe an important result in DCA, the Fenchel-type duality theorem, which we use to derive the min-max formulas related to dec-min elements. The Fenchel-type duality theorem in DCA originates in Murota [30] and is formulated for integer-valued functions in [31, 32].

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

$$f^{\bullet}(\pi) = \sup\{\langle \pi, x \rangle - f(x) : x \in \mathbf{Z}^S\} \qquad (\pi \in \mathbf{Z}^S), \tag{19}$$

$$h^{\circ}(\pi) = \inf\{\langle \pi, x \rangle - h(x) : x \in \mathbf{Z}^S\} \qquad (\pi \in \mathbf{Z}^S), \tag{20}$$

where $\langle \pi, x \rangle$ means the (standard) inner product of vectors $\pi$ and $x$. Since the functions are integer-valued, the supremum in (19) is attained if it is finite-valued. Similarly for the infimum in (20). Accordingly, we henceforth write "max" and "min" in place of "sup" in (19) and "inf" in (20), respectively.

The Fenchel-type 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 $\pi \in \mathbf{Z}^S$. It is known as the **weak duality** that the minimum of $f - h$ is larger than or equal to the maximum of $h^{\circ} - f^{\bullet}$.

Therefore, 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 $\pi$ since the functions are integer-valued. If $\mathrm{dom}\, f \cap \mathrm{dom}\, h = \emptyset$, we say understand (by convention) that the minimum of $f - h$ is equal to $+\infty$ and write "$\inf\{f(x) - h(x) : x \in \mathbf{Z}^S\} = +\infty$." Similarly, if $\mathrm{dom}\, f^{\bullet} \cap \mathrm{dom}\, h^{\circ} = \emptyset$, we write "$\sup\{h^{\circ}(\pi) - f^{\bullet}(\pi) : \pi \in \mathbf{Z}^S\} = -\infty$."

The following theorem shows the **strong duality** under the assumption that either (i) $\mathrm{dom}\, f \cap \mathrm{dom}\, h \neq \emptyset$, that is, there exists $x$ for which both $f(x)$ and $h(x)$ are finite (**primal feasibility**) or (ii) $\mathrm{dom}\, f^{\bullet} \cap \mathrm{dom}\, h^{\circ} \neq \emptyset$, that is, there exists $\pi$ for which both $f^{\bullet}(\pi)$ and $h^{\circ}(\pi)$ are finite (**dual feasibility**).

**THEOREM 4.1** (Fenchel-type duality theorem [31, 32]). *Let* $f : \mathbf{Z}^S \to \mathbf{Z} \cup \{+\infty\}$ *be an integer-valued* M$^{\natural}$*-convex function and* $h : \mathbf{Z}^S \to \mathbf{Z} \cup \{-\infty\}$ *be an integer-valued* 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*

$$\inf\{f(x) - h(x) : x \in \mathbf{Z}^S\} = \sup\{h^{\circ}(\pi) - f^{\bullet}(\pi) : \pi \in \mathbf{Z}^S\}. \tag{21}$$

*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 infimum and the supremum are attained.* •

To emphasize that the infimum and the supremum are attained in the finite-valued case, we henceforth express (21) as

$$\min\{f(x) - h(x) : x \in \mathbf{Z}^S\} = \max\{h^{\circ}(\pi) - f^{\bullet}(\pi) : \pi \in \mathbf{Z}^S\}. \tag{22}$$

It is noted, however, that we do not exclude the possibility of the unbounded case where both sides are equal to $-\infty$ or $+\infty$.

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

$$g(\pi) + g(\tau) \geq g\left(\left\lceil \frac{\pi + \tau}{2} \right\rceil\right) + g\left(\left\lfloor \frac{\pi + \tau}{2} \right\rfloor\right) \qquad (\pi, \tau \in \mathbf{Z}^S), \tag{23}$$

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. We refer to (23) 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** ([32, 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* $\pi^*$ *is a maximizer of* $g$ *if and only if it is locally maximal in the sense that*

$$g(\pi^*) \geq g(\pi^* - \chi_Y) \quad \text{for all } Y \subseteq S, \tag{24}$$

$$g(\pi^*) \geq g(\pi^* + \chi_Y) \quad \text{for all } Y \subseteq S. \tag{25}$$

•

The reader is referred to [32, Chapter 7] for more properties of L$^\natural$-convex functions and [32, Chapter 8] for the conjugacy between M$^\natural$-convexity and L$^\natural$-convexity. In particular, [32, Figure 8.1] offers the whole picture of conjugacy relationship.

**Remark 4.1.** 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 (21) 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 $\overset{....}{B}_1$ and $\overset{....}{B}_2$ associated with integral base-polyhedra $B_1$ and $B_2$, respectively, and for $i = 1, 2$, let $f_i$ be the indicator function of $\overset{....}{B}_i$ (i.e., $f_i(x) = 0$ if $x \in \overset{....}{B}_i$, and $f_i(x) = +\infty$ if $x \in \mathbf{Z}^S \setminus \overset{....}{B}_i$). 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 (21) is a nonlinear generalization of the weighted polymatroid intersection problem; see [32, Section 8.2.3] for details. •

**Remark 4.2.** 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 (21) is an L$^\natural$-concave function. In contrast, the function $f(x) - h(x)$ to be minimized on the left-hand side of (21) is not an M$^\natural$-convex function, as explained in Remark 4.1 above. In this sense, the left-hand side (minimization) and the right-hand side (maximization) are not symmetric. •

## 4.2 Min-max formula for separable convex functions on a base-polyhedron

We consider the problem of minimizing $\Phi(x) + \delta(x)$, where

$$\Phi(x) = \sum [\varphi_s(x(s)) : s \in S] \tag{26}$$

is an integer-valued separable convex function defined by integer-valued discrete convex functions $\varphi_s : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ and $\delta$ is the indicator function of $\overset{....}{B}$ defined in (13).

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 [8]. In contrast, we are now going to apply the Fenchel-type duality theorem to the minimization of the function $\Phi + \delta = \Phi - (-\delta)$. In so doing we can separate the constraint term $\delta(x)$ for the base-polyhedron from the objective function $\Phi(x)$.

With the choice of $f = \Phi$ and $h = -\delta$ in the min-max relation (22), the left-hand side of (22) represents minimization of $\Phi$ over the M-convex set $\overset{....}{B}$. We denote the conjugate function of $\varphi_s$ by $\psi_s$, which is a function $\psi_s : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ defined by

$$\psi_s(\ell) = \max\{k\ell - \varphi_s(k) : k \in \mathbf{Z}\} \qquad (\ell \in \mathbf{Z}). \tag{27}$$

Then the conjugate function of $f$ is given by

$$f^\bullet(\pi) = \sum [\psi_s(\pi(s)) : s \in S] \qquad (\pi \in \mathbf{Z}^S). \tag{28}$$

On the other hand, the conjugate function $h^\circ$ of $h$ is given by

$$h^\circ(\pi) = \min\{\langle \pi, x\rangle + \delta(x) : x \in \mathbf{Z}^S\} = \min\{\langle \pi, x\rangle : x \in \overset{....}{B}\} = \hat{p}(\pi) \quad (\pi \in \mathbf{Z}^S), \tag{29}$$

where $\hat{p}$ is the linear extension (Lovász extension)[2] of the supermodular function $p$ describing $B$.

Substituting (28) and (29) into (22) we obtain (30) below.

**THEOREM 4.3.** *Assume that either* (i) *there exists* $x \in \overset{....}{B}$ *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:*

$$\min\{\sum_{s\in S} \varphi_s(x(s)) : x \in \overset{....}{B}\} = \max\{\hat{p}(\pi) - \sum_{s\in S} \psi_s(\pi(s)) : \pi \in \mathbf{Z}^S\}. \tag{30}$$

*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 (30) is an L$^\natural$-concave function (cf. Remark 4.2). We state this as a proposition for later references.

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

In applications of (30) with concrete functions $\varphi_s$, it is often the case that the conjugate functions $\psi_s$ can be computed to explicit forms. This is illustrated in the following examples.

**Example 4.1.** Let $B$ be an integral base-polyhedron and $c \in \mathbf{Z}^S$ be an integer vector. We consider the minimum $\ell_1$-distance from an integer point $m$ in $B$ to the given point $c$. As a special case of the min-max formula (30) we can obtain the following min-max relation:

$$\begin{aligned} &\min\{\sum_{s\in S} |m(s) - c(s)| : m \in \overset{....}{B}\} \\ &= \max\{p(X) - b(Y) - \widetilde{c}(X) + \widetilde{c}(Y) : X, Y \subseteq S;\ X \cap Y = \emptyset\}, \end{aligned} \tag{31}$$

where $p$ and $b$ are the supermodular and submodular functions associated with $B$, and $\widetilde{c}(X) = \sum\{c(s) : s \in X\}$ for $X \subseteq S$. To derive (31) from (30), we choose $\varphi_s(k) = |k - c(s)|$, for which the left-hand side of (30) coincides with that of (31). The conjugate functions can be computed as

$$\psi_s(\ell) = \max\{k\ell - |k - c(s)| : k \in \mathbf{Z}\} = \begin{cases} c(s)\ell & (\ell = -1, 0, +1), \\ +\infty & (\text{otherwise}), \end{cases}$$

and the right-hand side of (30) reads

$$\max\{\hat{p}(\pi) - \sum_{s\in S} c(s)\pi(s) : \pi(s) \in \{-1, 0, +1\}\ (s \in S)\}.$$

On representing $\pi = \chi_X - \chi_Y$ with disjoint subsets $X$ and $Y$, we obtain $\hat{p}(\pi) = p(X) - b(Y)$ and $\sum_{s\in S} c(s)\pi(s) = \widetilde{c}(X) - \widetilde{c}(Y)$. Thus the right-hand side of (30) coincides with that of (31). •

[2] See Section 6.2 of Part I [8] for the linear extension $\hat{p}$.

**Example 4.2.** Let $B$ be an integral base-polyhedron and $c, d \in \mathbf{Z}^S$ be integer vectors with $c \leq d$. We consider the minimum $\ell_1$-distance from an integer point $m$ in $B$ to the interval (box) specified by $c$ and $d$. The distance is represented as $\sum[D_s(m(s)) : s \in S]$ with

$$D_s(k) = \min\{|k - z| : c(s) \leq z \leq d(s)\} = \max(c(s) - k, 0, k - d(s)).$$

As a special case of the min-max formula (30) we can obtain the following min-max relation:

$$\begin{aligned} &\min\{\sum[D_s(m(s)) : s \in S] : m \in \overset{....}{B}\} \\ &= \max\{p(X) - b(Y) - \widetilde{d}(X) + \widetilde{c}(Y) : X, Y \subseteq S;\ X \cap Y = \emptyset\}, \end{aligned} \tag{32}$$

where $p$ and $b$ are the supermodular and submodular functions associated with $B$. To derive (32) from (30), we choose $\varphi_s(k) = D_s(k)$, for which the left-hand side of (30) coincides with that of (32). The conjugate functions can be computed as

$$\psi_s(\ell) = \max\{k\ell - D_s(k) : k \in \mathbf{Z}\} = \begin{cases} -c(s) & (\ell = -1), \\ 0 & (\ell = 0), \\ d(s) & (\ell = +1), \\ +\infty & (\text{otherwise}), \end{cases}$$

and the right-hand side of (30) reads

$$\max\{\hat{p}(\pi) - \sum_{s:\pi(s)=+1} d(s) + \sum_{s:\pi(s)=-1} c(s) : \pi(s) \in \{-1, 0, +1\}\ (s \in S)\}.$$

On representing $\pi = \chi_X - \chi_Y$ with disjoint subsets $X$ and $Y$, we obtain

$$\hat{p}(\pi) = p(X) - b(Y), \qquad \sum_{s:\pi(s)=+1} d(s) = \widetilde{d}(X), \qquad \sum_{s:\pi(s)=-1} c(s) = \widetilde{c}(Y).$$

Thus the right-hand side of (30) coincides with that of (32). •

When specialized to a symmetric function $\Phi$, the min-max formula (30) is simplified to

$$\min\{\sum_{s \in S} \varphi(x(s)) : x \in \overset{....}{B}\} = \max\{\hat{p}(\pi) - \sum_{s \in S} \psi(\pi(s)) : \pi \in \mathbf{Z}^S\}, \tag{33}$$

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 $\varphi$ defined as $\psi(\ell) = \max\{k\ell - \varphi(k) : k \in \mathbf{Z}\}$ for $\ell \in \mathbf{Z}$. With appropriate choices of $\varphi$ in (33) we derive the formulas (15), (16), and (17) in Sections 4.3, 4.4, and 4.5, respectively.

## 4.3 DCA-based proof of the min-max formula for square-sum minimization

The min-max formula (15) for the square-sum can be derived immediately from our duality formula (33). For $\varphi(k) = k^2$, the conjugate function $\psi(\ell)$ for $\ell \in \mathbf{Z}$ is given explicitly as

$$\psi(\ell) = \max\{k\ell - k^2 : k \in \mathbf{Z}\} = \max\{k\ell - k^2 : k \in \{\lfloor \ell/2 \rfloor, \lceil \ell/2 \rceil\}\} = \left\lfloor \frac{\ell}{2} \right\rfloor \cdot \left\lceil \frac{\ell}{2} \right\rceil. \tag{34}$$

The substitution of (34) into (33) yields (15). Note that the primal feasibility is satisfied since $\varphi(k)$ is finite for all $k$.

**Remark 4.3.** We can also formulate a min-max formula for a nonsymmetric quadratic function $\sum_{s\in S} c(s)m(s)^2$, where $c(s)$ is a positive integer for each $s \in S$. For $\varphi(k) = ck^2$ with a positive integer $c$, the conjugate function $\psi(\ell)$ is given as

$$\begin{aligned}\psi(\ell) &= \max\{k\ell - ck^2 : k \in \mathbf{Z}\} = \max\{k\ell - ck^2 : k \in \{\lfloor \ell/2c \rfloor, \lceil \ell/2c \rceil\}\,\} \\ &= \max\left( \left\lfloor \frac{\ell}{2c} \right\rfloor \left(\ell - c\left\lfloor \frac{\ell}{2c} \right\rfloor\right), \left\lceil \frac{\ell}{2c} \right\rceil \left(\ell - c\left\lceil \frac{\ell}{2c} \right\rceil\right)\right). \end{aligned} \quad (35)$$

Therefore, the min-max formula reads:

$$\begin{aligned}&\min\{\sum_{s\in S} c(s)m(s)^2 : m \in \ddot{\ddot{B}}\} \\ &= \max\{\hat{p}(\pi) - \sum_{s\in S} \max\left( \left\lfloor \frac{\pi(s)}{2c(s)} \right\rfloor \left(\pi(s) - c(s)\left\lfloor \frac{\pi(s)}{2c(s)} \right\rfloor\right), \left\lceil \frac{\pi(s)}{2c(s)} \right\rceil \left(\pi(s) - c(s)\left\lceil \frac{\pi(s)}{2c(s)} \right\rceil\right)\right) : \pi \in \mathbf{Z}^S\}.\end{aligned}$$

•

## 4.4 DCA-based proof of the formula for $\beta_1$

The formula (16) for the largest component $\beta_1$ of a max-minimizer of $\ddot{\ddot{B}}$ can also be derived from our duality formula (33). With a nonnegative integer $\alpha$ as a parameter, we choose

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

in (33). By the definition of $\beta_1$, the left-hand side of (33) is equal to zero if $\alpha \ge \beta_1$, and equal to $+\infty$ if $\alpha \le \beta_1 - 1$. Hence $\beta_1$ is equal to the minimum of $\alpha$ for which the left-hand side is equal to zero.

The conjugate function $\psi$ of $\varphi$ is given by

$$\psi(\ell) = \max\{k\ell : k \le \alpha\} = \begin{cases} +\infty & (\ell \le -1), \\ 0 & (\ell = 0), \\ \alpha\ell & (\ell \ge 1). \end{cases} \quad (36)$$

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 $\lambda$). This implies, in particular, that the maximization problem on the right-hand side of (33) is feasible for all $\alpha$ and hence the identity (33) holds, which reads either $0 = 0$ or $+\infty = +\infty$. Since $\beta_1$ is the minimum of $\alpha$ for which the left-hand side is equal to 0, we can say that $\beta_1$ is the minimum of $\alpha$ 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 (24) in Theorem 4.2 is satisfied trivially by (36), whereas the second condition (25) reads $g(\pi^* + \chi_Y) = p(X) - \alpha|X| \le 0$. Therefore, the right-hand side of (33) is equal to zero if and only if $\max\{p(X) - \alpha|X| : X \subseteq S\} = 0$, from which follows the formula (16).

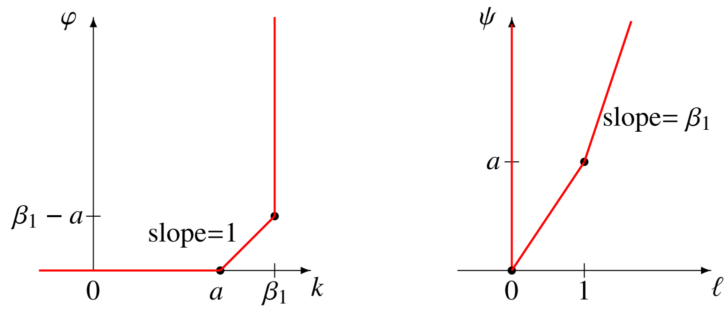


Figure 1: Mutually conjugate discrete convex functions $\varphi$ and $\psi$ in (37) and (38)

## 4.5 DCA-based proof of the formula for $r_1$

The formula (17) for the minimum number $r_1$ of $\beta_1$-valued components of a $\beta_1$-covered member of $\ddot{B}$ can also be derived from our duality formula (33). We choose

$$\varphi(k) = \begin{cases} 0 & (k \leq \beta_1 - 1), \\ 1 & (k = \beta_1), \\ +\infty & (k \geq \beta_1 + 1), \end{cases} \tag{37}$$

whose graph is given by the left of Fig. 1, where $a = \beta_1 - 1$. By the definitions of $\beta_1$ and $r_1$, the minimum in (33) is equal to $r_1$. In particular, the primal problem is feasible, and hence the identity (33) holds.

The conjugate function $\psi$ of $\varphi$ is given by

$$\begin{aligned} \psi(\ell) &= \max\left(\max\{k\ell : k \leq \beta_1 - 1\},\ \beta_1\ell - 1\right) \\ &= \begin{cases} +\infty & (\ell \leq -1), \\ 0 & (\ell = 0), \\ \beta_1\ell - 1 & (\ell \geq 1), \end{cases} \end{aligned} \tag{38}$$

whose graph is given by the right of Fig. 1, where $a = \beta_1 - 1$. 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 $\pi$ to $\{0,1\}$-vectors, as shown in Lemma 4.5 below. For $\pi = \chi_X$ 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 (33) is equal to $\max\{p(X) - (\beta_1 - 1)|X| : X \subseteq S\}$. Thus the formula (17) is derived.

**Lemma 4.5.** *There exists a $\{0,1\}$-vector $\pi$ 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 (24) and (25) in the L-optimality criterion (Theorem 4.2) are satisfied.

Proof of (24): $g(\pi^*) \geq g(\pi^* - \chi_Y)$. We may assume $Y \subseteq A$, since, otherwise, $\pi^* - \chi_Y \notin \operatorname{dom} g$ by (38). 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 (25): $g(\pi^*) \geq g(\pi^* + \chi_Y)$. Since

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

the definitions of $g$ and $\psi$ show

$$\begin{aligned} g(\pi^* + \chi_Y) &= \hat{p}(\chi_A + \chi_Y) - (\beta_1 + a)|A \cap Y| - a|(A \cup Y) \setminus (A \cap Y)| \\ &= \hat{p}(\chi_A + \chi_Y) - (\beta_1 + a)|A \cap Y| - a(|A \cup Y| - |A \cap Y|) \\ &= \hat{p}(\chi_A + \chi_Y) - \beta_1|A \cap Y| - a|A \cup Y|. \end{aligned} \tag{39}$$

By the definition of $\hat{p}$ we have

$$\hat{p}(\chi_A + \chi_Y) = 2p(A \cap Y) + [p(A \cup Y) - p(A \cap Y)] = p(A \cap Y) + p(A \cup Y), \tag{40}$$

where

$$p(A \cap Y) \leq \beta_1|A \cap Y|, \tag{41}$$

$$p(A \cup Y) \leq (p(A) - a|A|) + a|A \cup Y|. \tag{42}$$

It should be clear that (41) holds since $(\beta_1, \beta_1, \ldots, \beta_1)$ belongs to the supermodular polyhedra defined by $p$, and (42) holds since $A$ is a maximizer of $p(X) - a|X|$. It follows from (39)–(42) that

$$g(\pi^* + \chi_Y) \leq p(A) - a|A| = g(\pi^*),$$

proving (25). □

## 4.6 Another min-max formula

In this section we establish another min-max formula, which is announced in (18) at the beginning of Section 4.

**THEOREM 4.6.** *Let $B$ be the base-polyhedron described by an integer-valued supermodular function $p$, and $\ddot{B}$ be the set of the integral elements of $B$. For each integer $a$, we have the min-max relation*

$$\min\{\sum_{s \in S}(m(s) - a)^+ : m \in \ddot{B}\} = \max\{p(X) - a|X| : X \subseteq S\}. \tag{43}$$

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

*Proof.* In our duality formula (33) we choose

$$\varphi(k) = (k - a)^+ = \begin{cases} 0 & (k \leq a), \\ k - a & (k \geq a + 1). \end{cases} \tag{44}$$

The left-hand side of (33) coincides with the left-hand side of (43). Proposition 3.2 shows that any dec-min element of $\ddot{B}$ is a minimizer in (43) 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 $x{\downarrow} = y{\downarrow}$.

The conjugate function $\psi$ of $\varphi$ is given by

$$\begin{aligned}\psi(\ell) &= \max\left(\max\{k\ell : k \leq a\},\ \max\{k\ell - (k-a) : k \geq a+1\}\right) \\ &= \begin{cases} 0 & (\ell = 0), \\ a & (\ell = 1), \\ +\infty & (\ell \notin \{0,1\}). \end{cases}\end{aligned} \tag{45}$$

Therefore, we may restrict $\pi$ 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 $\pi = \chi_X$ 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 (33) is equal to $\max\{p(X) - a|X| : X \subseteq S\}$. Thus the formula (43) is derived. □

The established formula (43) generalizes the formula (17) for $r_1$. Indeed, the formula (43) for $a = \beta_1 - 1$ coincides with the formula (17), since $\sum_{s\in S}(m(s) - a)^+ = \sum_{s\in S}(m(s) - (\beta_1 - 1))^+ = |\{s \in S : m(s) = \beta_1\}| = r_1$ for any dec-min element $m$ of $\ddot{B}$.

For $i = 1, 2, \ldots$, define[3] $\hat{r}_i$ by $\hat{r}_i = |\{s \in S : m(s) = \beta_1 - i + 1\}|$ with a dec-min element $m$ of $\ddot{B}$. Note that $\hat{r}_1 = r_1$ and $\hat{r}_i$ does not depend on the choice of $m$. Since

$$\sum[(m(s) - (\beta_1 - i))^+ : s \in S] = \sum_{j=1}^{i}(i - j + 1)\hat{r}_j \qquad (i = 1, 2, \ldots),$$

the formula (43) implies

$$\sum_{j=1}^{i}(i - j + 1)\hat{r}_j = \max\{p(X) - (\beta_1 - i)|X| : X \subseteq S\} \qquad (i = 1, 2, \ldots). \tag{46}$$

This formula gives a recurrence formula for $\hat{r}_1, \hat{r}_2, \hat{r}_3, \ldots$ as

$$\begin{aligned}\hat{r}_1 &= \max\{p(X) - (\beta_1 - 1)|X| : X \subseteq S\}, \\ \hat{r}_2 &= \max\{p(X) - (\beta_1 - 2)|X| : X \subseteq S\} - 2\hat{r}_1, \\ \hat{r}_3 &= \max\{p(X) - (\beta_1 - 3)|X| : X \subseteq S\} - 3\hat{r}_1 - 2\hat{r}_2, \\ &\cdots\cdots\cdots\end{aligned}$$

## 4.7 Min-max formula for separable convex functions on a g-polymatroid

Min-max formulas for base-polyhedra can also be adapted to g-polymatroids. Let $Q$ denote an integral g-polymatroid, and $\ddot{Q}$ be the set of the integral points of $Q$, where $\ddot{Q}$ is an M$^\natural$-convex set.

[3] In Section 5.2 of Part I [8], we have introduced notation $r_i$ for the number of elements $s$ of $S - C_i$ with $m(s) = \beta_i$, i.e., $r_i = |\{s \in S - C_i : m(s) = \beta_i\}|$. For $i \geq 2$, the number $r_i$ is (generally) not equal to $\hat{r}_i$ here.

The min-max formula (30) for a separable convex function can be adapted to a g-polymatroid as follows:

$$\min\{\sum[\varphi_s(x(s)) : s \in S] : x \in \ddot{Q}\} = \max\{\mu^Q_{\min}(\pi) - \sum[\psi_s(\pi(s)) : s \in S] : \pi \in \mathbf{Z}^S\}, \quad (47)$$

where $\mu^Q_{\min}(\pi) = \min\{\pi x : x \in Q\}$, and $\mu^Q_{\min}(\pi)$ can be expressed as a certain "linear extension" of the pair $(p, b)$ describing $Q$, though we do not enter into the details here.

As a special case of the above, the min-max formula for the square-sum is given as

$$\min\{\sum[m(s)^2 : s \in S] : m \in \ddot{Q}\} = \max\{\mu^Q_{\min}(\pi) - \sum_{s\in S} \left\lfloor \frac{\pi(s)}{2} \right\rfloor \left\lceil \frac{\pi(s)}{2} \right\rceil : \pi \in \mathbf{Z}^S\}. \quad (48)$$

It is noted, however, the square-sum minimizers are not directly related to dec-min elements of $\ddot{Q}$ unless $Q$ is contained in the nonnegative orthant.

## 4.8 Min-max formula for separable convex functions on the intersection of base-polyhedra

The duality formula for separable discrete convex functions on a single integral base-polyhedron, given in Theorem 4.3, admits an extension to separable discrete convex functions on the intersection of two integral base-polyhedra. In Part III [9] this extension serves as a basis of the study of decreasing-minimality in the intersection of two 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 of $p_i$ as $\hat{p}_i$, i.e., $\hat{p}_i(\pi) = \min\{\pi x : x \in \ddot{B}_i\}$ for $\pi \in \mathbf{Z}^S$.

Theorem 4.7 below gives a duality formula for separable discrete convex functions on the intersection of two integral base-polyhedra (=the intersection of two M-convex sets). 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 $\varphi_s$ by $\psi_s : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$, which is defined by (27).

**THEOREM 4.7.** *Assume that either* (i) *there exists* $x \in \ddot{B}_1 \cap \ddot{B}_2$ *such that* $\varphi_s(x(s)) < +\infty$ *for all* $s \in S$ *(primal feasibility) or* (ii) *there exists* $\pi_1, \pi_2 \in \mathbf{Z}^S$ *such that* $\hat{p}_1(\pi_1) > -\infty$, $\hat{p}_2(\pi_2) > -\infty$, *and* $\psi_s(\pi_1(s) + \pi_2(s)) < +\infty$ *for all* $s \in S$ *(dual feasibility). Then we have the min-max relation:*[4]

$$\min\{\sum_{s\in S} \varphi_s(x(s)) : x \in \ddot{B}_1 \cap \ddot{B}_2\}$$
$$= \max\{\hat{p}_1(\pi_1) + \hat{p}_2(\pi_2) - \sum_{s\in S} \psi_s(\pi_1(s) + \pi_2(s)) : \pi_1, \pi_2 \in \mathbf{Z}^S\}. \quad (49)$$

*Proof.* We denote the indicator functions of $\ddot{B}_1$ and $\ddot{B}_2$ by $\delta_1$ and $\delta_2$, respectively, and continue to use the notation $\Phi(x) = \sum[\varphi_s(x(s)) : s \in S]$ introduced in (26).

[4]The unbounded case with both sides of (49) being equal to $-\infty$ or $+\infty$ is also a possibility.

In the Fenchel-type duality

$$\min\{f(x) - h(x) : x \in \mathbf{Z}^S\} = \max\{h^{\circ}(\pi) - f^{\bullet}(\pi) : \pi \in \mathbf{Z}^S\} \tag{50}$$

in (22), we choose $f = \delta_2 + \Phi$ and $h = -\delta_1$. Since $f - h = \Phi + \delta_1 + \delta_2$, the left-hand side of (50) coincides with the left-hand side of (49).

The conjugate function $f^{\bullet}$ can be computed as follows. For $\pi \in \mathbf{Z}^S$ we define a function $\varphi_s^{\pi} : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ by $\varphi_s^{\pi}(k) = \varphi_s(k) - \pi(s)k$ for $k \in \mathbf{Z}$. Then the conjugate function $\psi_s^{\pi}$ of this function is given as

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

Using this expression and the min-max formula (30) for $B_2$ and $\varphi_s^{\pi}$, we obtain

$$\begin{aligned}f^{\bullet}(\pi) &= \max\{\langle \pi, x\rangle - \delta_2(x) - \sum_{s\in S}\varphi_s(x(s)) : x \in \mathbf{Z}^S\} \\ &= -\min\{\delta_2(x) + \sum_{s\in S}\varphi_s^{\pi}(x(s)) : x \in \mathbf{Z}^S\} \\ &= -\max\{\hat{p}_2(\pi') - \sum_{s\in S}\psi_s^{\pi}(\pi'(s)) : \pi' \in \mathbf{Z}^S\} \\ &= -\max\{\hat{p}_2(\pi') - \sum_{s\in S}\psi_s(\pi(s) + \pi'(s)) : \pi' \in \mathbf{Z}^S\} \qquad (\pi \in \mathbf{Z}^S).\end{aligned} \tag{51}$$

On the other hand, the conjugate function $h^{\circ}$ of $h = -\delta_1$ is equal to $\hat{p}_1$ by (29), i.e.,

$$h^{\circ}(\pi) = \hat{p}_1(\pi) \qquad (\pi \in \mathbf{Z}^S). \tag{52}$$

The substitution of (51) and (52) into $h^{\circ} - f^{\bullet}$ shows that the right-hand side of (50) coincides with the right-hand side of (49). □

When specialized to a symmetric function, the min-max formula (49) is simplified to

$$\begin{aligned}&\min\{\sum_{s\in S}\varphi(x(s)) : x \in \ddot{B}_1 \cap \ddot{B}_2\} \\ &= \max\{\hat{p}_1(\pi_1) + \hat{p}_2(\pi_2) - \sum_{s\in S}\psi(\pi_1(s) + \pi_2(s)) : \pi_1, \pi_2 \in \mathbf{Z}^S\},\end{aligned} \tag{53}$$

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 $\varphi$ defined as $\psi(\ell) = \max\{k\ell - \varphi(k) : k \in \mathbf{Z}\}$ for $\ell \in \mathbf{Z}$. The identity (53) will play a key role in the study of discrete decreasing minimization on the intersection of two base-polyhedra, just as (33) did for a single base-polyhedron.

As an example of (53) we mention a min-max identity for the minimum square-sum of components on the intersection of two integral base-polyhedra, which is an extension of (15) for a single integral base-polyhedron.

**Proposition 4.8.**

$$\min\{\sum_{s\in S} m(s)^2 : m \in \overset{\cdot\cdot\cdot\cdot}{B_1} \cap \overset{\cdot\cdot\cdot\cdot}{B_2}\}$$
$$= \max\{\hat{p}_1(\pi_1) + \hat{p}_2(\pi_2) - \sum_{s\in S} \left\lfloor \frac{\pi_1(s)+\pi_2(s)}{2} \right\rfloor \cdot \left\lceil \frac{\pi_1(s)+\pi_2(s)}{2} \right\rceil : \pi_1, \pi_2 \in \mathbf{Z}^S\}. \tag{54}$$

*Proof.* This is a special case of (53) with $\varphi(k) = k^2$ and $\psi(\ell) = \lfloor \ell/2 \rfloor \cdot \lceil \ell/2 \rceil$ (cf., (34)). □

If $\overset{\cdot\cdot\cdot\cdot}{B_1} \cap \overset{\cdot\cdot\cdot\cdot}{B_2} \neq \emptyset$, both sides of (54) are finite-valued, and the minimum and the maximum are attained. If $\overset{\cdot\cdot\cdot\cdot}{B_1} \cap \overset{\cdot\cdot\cdot\cdot}{B_2} = \emptyset$, the left-hand side of (54) is equal to $+\infty$ by convention and the right-hand side is unbounded above (hence equal to $+\infty$). Note also that $\overset{\cdot\cdot\cdot\cdot}{B_1} \cap \overset{\cdot\cdot\cdot\cdot}{B_2} \neq \emptyset$ if and only if $B_1 \cap B_2 \neq \emptyset$.

**Remark 4.4.** We can also formulate a min-max formula for a nonsymmetric quadratic function $\sum_{s\in S} c(s)m(s)^2$, where $c(s)$ is a positive integer for each $s \in S$ (cf. Remark 4.3). On recalling the conjugate function in (35), we obtain the min-max formula

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

•

# 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:

$$\min\{\sum[m(s)^2 : s \in S] : m \in \overset{\cdot\cdot\cdot\cdot}{B}\} = \max\{\hat{p}(\pi) - \sum_{s\in S} \left\lfloor \frac{\pi(s)}{2} \right\rfloor \left\lceil \frac{\pi(s)}{2} \right\rceil : \pi \in \mathbf{Z}^S\}, \tag{55}$$

to which a DCA-based proof has been given in Section 4.3.

Concerning the optimal solutions to (55) the following results were obtained in Part I [8]. 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, $\pi^*$ and $\Delta^*$ 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 5.2 of Part I [8].

**Proposition 5.1** ([8, Corollary 6.11]). *The set* $\Pi$ *of dual optimal integral vectors* $\pi$ *in* (55) *is an* $\mathrm{L}^\natural$*-convex set. The unique smallest element of* $\Pi$ *is* $\pi^*$. •

**THEOREM 5.2** ([8, Theorem 6.9]). *An integral vector* $\pi$ *is a dual optimal solution in* (55) *if and only if the following three conditions hold for each* $i = 1, 2, \ldots, q$ :

$$\pi(s) = 2\beta_i - 1 \ \text{ for every } s \in S_i - F_i, \tag{56}$$
$$2\beta_i - 1 \leq \pi(s) \leq 2\beta_i + 1 \ \text{ for every } s \in F_i, \tag{57}$$
$$\pi(s) - \pi(t) \geq 0 \ \text{ whenever } s, t \in F_i \text{ and } (s, t) \in A_i, \tag{58}$$

*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** ([8, Theorem 5.5]). *The set of dec-min elements of* $\ddot{\ddot{B}}$ *is a matroidal M-convex set.*[5] *More precisely, an element* $m$ *of* $\ddot{\ddot{B}}$ *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 Section 5.1.

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

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

$$\min\{f(x) - h(x) : x \in \mathbf{Z}^S\} = \max\{h^\circ(\pi) - f^\bullet(\pi) : \pi \in \mathbf{Z}^S\}, \tag{59}$$

where $f$ is an integer-valued $\mathrm{M}^\natural$-convex function and $h$ is an integer-valued $\mathrm{M}^\natural$-concave function. We assume that the common value in (59) is finite and denote the set of the minimizers by $\mathcal{P}$ and the set of the maximizers by $\mathcal{D}$.

The (convex/concave) integral subdifferentials of $f$ and $h$ at $x \in \mathbf{Z}^S$ are the sets of vectors defined as:

$$\begin{aligned}
\partial f(x) &= \{\pi \in \mathbf{Z}^S : f(y) - f(x) \geq \langle \pi, y - x \rangle \ (\forall y \in \mathbf{Z}^S)\},\\
\partial h(x) &= \{\pi \in \mathbf{Z}^S : h(y) - h(x) \leq \langle \pi, y - x \rangle \ (\forall y \in \mathbf{Z}^S)\},
\end{aligned}$$

where $\partial f(x)$ is defined for $x \in \mathrm{dom}\, f$ and $\partial h(x)$ for $x \in \mathrm{dom}\, h$. Similarly, the (convex/concave) integral subdifferentials of $f^\bullet$ and $h^\circ$ at $\pi \in \mathbf{Z}^S$ are defined as

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

[5] 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 $\ell_\infty$-distance of any two distinct members is equal to one.

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 representations of the sets of optimal solutions are immediate consequences of Theorem 4.1; we include the proof for completeness.

**Proposition 5.4.**

(1) *Assume* $x \in \mathrm{dom}\, f \cap \mathrm{dom}\, h$ *and* $\pi \in \mathrm{dom}\, f^{\bullet} \cap \mathrm{dom}\, h^{\circ}$. *We have* $x \in \mathcal{P}$ *and* $\pi \in \mathcal{D}$ *if and only if* $\pi \in \partial f(x) \cap \partial h(x)$, *or equivalently,* $x \in \partial f^{\bullet}(\pi) \cap \partial h^{\circ}(\pi)$.

(2) *For any* $\hat{x} \in \mathcal{P}$, *we have* $\mathcal{D} = \partial f(\hat{x}) \cap \partial h(\hat{x})$.

(3) *For any* $\hat{\pi} \in \mathcal{D}$, *we have* $\mathcal{P} = \partial f^{\bullet}(\hat{\pi}) \cap \partial h^{\circ}(\hat{\pi})$.

*Proof.* (1) By the definition of the conjugate functions in (19) and (20) we have

$$f(x) + f^{\bullet}(\pi) \geq \langle \pi, x \rangle, \tag{60}$$
$$h(x) + h^{\circ}(\pi) \leq \langle \pi, x \rangle \tag{61}$$

for any $x$ and $\pi$, from which follows the weak duality:

$$f(x) - h(x) \geq h^{\circ}(\pi) - f^{\bullet}(\pi). \tag{62}$$

The equality holds in (62) if and only if the two inequalities in (60) and (61) are satisfied in equalities. We have equality in (60) if and only if $\pi \in \partial f(x)$, or equivalently, $x \in \partial f^{\bullet}(\pi)$ by the biconjugacy $f^{\bullet\bullet} = f$ (cf. [32, Theorem 8.12]). Similarly, we have equality in (61) if and only if $\pi \in \partial h(x)$, or equivalently, $x \in \partial h^{\circ}(\pi)$.

(2) and (3) follow immediately from (1). □

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

The subdifferential of an M$^{\natural}$-convex function $f$ admits a more concrete representation as a consequence of the M-optimality criterion (Theorem 3.6), which gives a local characterization of global minimality for M$^{\natural}$-convex functions. Namely, for $x \in \mathrm{dom}\, f$, we have

$$\begin{aligned}
\partial f(x) &= \{\pi \in \mathbf{Z}^S : f(y) - f(x) \geq \langle \pi, y - x \rangle \ (\forall y \in \mathbf{Z}^S)\} \\
&= \{\pi \in \mathbf{Z}^S : x \in \arg\min(f(y) - \langle \pi, y \rangle)\} \\
&= \{\pi \in \mathbf{Z}^S : f(x) - \langle \pi, x \rangle \leq f(x + \chi_s - \chi_t) - \langle \pi, x + \chi_s - \chi_t \rangle \ (\forall s, t \in S), \\
&\qquad f(x) - \langle \pi, x \rangle \leq f(x + \chi_s) - \langle \pi, x + \chi_s \rangle \ (\forall s \in S), \\
&\qquad f(x) - \langle \pi, x \rangle \leq f(x - \chi_t) - \langle \pi, x - \chi_t \rangle \ (\forall t \in S)\} \\
&= \{\pi \in \mathbf{Z}^S : \pi(s) - \pi(t) \leq f(x + \chi_s - \chi_t) - f(x) \ (\forall s, t \in S), \\
&\qquad f(x) - f(x - \chi_s) \leq \pi(s) \leq f(x + \chi_s) - f(x) \ (\forall s \in S)\}.
\end{aligned} \tag{63}$$

This expression shows that $\partial f(x)$ is an L$^{\natural}$-convex set. Similarly,

$$\begin{aligned}
\partial h(x) &= \{\pi \in \mathbf{Z}^S : \pi(s) - \pi(t) \geq h(x + \chi_s - \chi_t) - h(x) \ (\forall s, t \in S), \\
&\qquad h(x) - h(x - \chi_s) \geq \pi(s) \geq h(x + \chi_s) - h(x) \ (\forall s \in S)\},
\end{aligned} \tag{64}$$

which is also an L$^\natural$-convex set.

On the other hand, the conjugate function $f^\bullet$ is an L$^\natural$-convex function, and by Theorem 4.2 (L-optimality criterion), the subdifferential of $f^\bullet$ is given as

$$\begin{aligned}
\partial f^\bullet(\pi) &= \{x \in \mathbf{Z}^S : f^\bullet(\tau) - f^\bullet(\pi) \geq \langle \tau - \pi, x\rangle \ (\forall \tau \in \mathbf{Z}^S)\} \\
&= \{x \in \mathbf{Z}^S : \pi \in \arg\min(f^\bullet(\tau) - \langle \tau, x\rangle)\} \\
&= \{x \in \mathbf{Z}^S : f^\bullet(\pi) - \langle \pi, x\rangle \leq f^\bullet(\pi + \chi_Y) - \langle \pi + \chi_Y, x\rangle \ (\forall Y \subseteq S), \\
&\qquad f^\bullet(\pi) - \langle \pi, x\rangle \leq f^\bullet(\pi - \chi_Y) - \langle \pi - \chi_Y, x\rangle \ (\forall Y \subseteq S)\} \\
&= \{x \in \mathbf{Z}^S : f^\bullet(\pi) - f^\bullet(\pi - \chi_Y) \leq \sum_{s \in Y} x(s) \leq f^\bullet(\pi + \chi_Y) - f^\bullet(\pi) \ (\forall Y \subseteq S)\}. \qquad (65)
\end{aligned}$$

By L$^\natural$-convexity, $f^\bullet(\pi+\chi_Y)-f^\bullet(\pi)$ is submodular in $Y$ and $f^\bullet(\pi)-f^\bullet(\pi-\chi_Y)$ is supermodular in $Y$, and moreover, they are paramodular in the sense of [7, Section 14.1] (or a strong pair). Therefore, the set $\partial f^\bullet(\pi)$ is an M$^\natural$-convex set (the set of integer points in an integral g-polymatroid). Similarly,

$$\partial h^\circ(\pi) = \{x \in \mathbf{Z}^S : h^\circ(\pi) - h^\circ(\pi - \chi_Y) \geq \sum_{s \in Y} x(s) \geq h^\circ(\pi + \chi_Y) - h^\circ(\pi) \ (\forall Y \subseteq S)\} \qquad (66)$$

is an M$^\natural$-convex set.

The reader is referred to [32, Chapter 8] for the conjugacy between M$^\natural$-convexity and L$^\natural$-convexity. In particular, [32, Figure 8.1] shows the whole picture of conjugacy relationship.

## 5.2 Structure of dual optimal solutions to square-sum minimization

The min-max formula (55) for the square-sum minimization is a special case of the Fenchel-type duality (59) with

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

where $\varphi(k) = k^2$ and $\psi(\ell) = \lfloor \ell/2 \rfloor \cdot \lceil \ell/2 \rceil$ for $k, \ell \in \mathbf{Z}$. Accordingly, we can apply the general results (Proposition 5.4, in particular) summarized in Section 5.1 for the analysis of the optimal solutions in the min-max formula (55). In this section we consider the dual solutions, whereas the primal solutions are treated in Section 5.3.

The function $g(\pi) = \hat{p}(\pi) - \sum[\psi(\pi(s)) : s \in S]$ to be maximized in (55) is L$^\natural$-concave by Proposition 4.4, and the maximizers of an L$^\natural$-concave function form an L$^\natural$-convex set [32, Theorem 7.17]. Therefore, the set $\Pi$ of dual optimal solutions is an L$^\natural$-convex set, which is the first statement of Proposition 5.1. The L$^\natural$-convexity of $\Pi$ implies that there exists a unique smallest element of $\Pi$. The second statement of Proposition 5.1 shows that this smallest element is given by $\pi^*$, 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 $\Pi$. According to the general result stated in Proposition 5.4 (2), we can obtain another representation of $\Pi$ of the form $\Pi = \partial f(\hat{x}) \cap \partial h(\hat{x})$ by choosing any dec-min element $\hat{x}$ of $\ddot{B}$, which is a primal optimal solution for (55). In the following we compare the two representations of $\Pi$.

The subdifferential $\partial f(x)$ at $x \in \mathbf{Z}^S$ can be computed as

$$\partial f(x) = \{\pi \in \mathbf{Z}^S : 2x(s) - 1 \leq \pi(s) \leq 2x(s) + 1 \ (s \in S)\}. \tag{67}$$

This follows from the expression (63), in which $f(x) = \sum[\varphi(x(s)) : s \in S]$, $\varphi(k+1) - \varphi(k) = (k+1)^2 - k^2 = 2k + 1$, and $\varphi(k) - \varphi(k-1) = k^2 - (k-1)^2 = 2k - 1$. For $h(x) = -\delta(x)$, on the other hand, we observe from the definition of the subdifferential that $\pi$ belongs to $\partial h(x)$ if and only if $x$ is a minimum $\pi$-weight base. Then Proposition 5.4 (2) gives the following representation of $\Pi$.

**Proposition 5.5.** *Let $m$ be any dec-min element of $\ddot{B}$. The set $\Pi$ of dual optimal solutions to* (55) *is represented as $\Pi = I(m) \cap W(m)$, where*

$$I(m) = \{\pi \in \mathbf{Z}^S : 2m(s) - 1 \leq \pi(s) \leq 2m(s) + 1 \text{ for all } s \in S\},$$
$$W(m) = \{\pi \in \mathbf{Z}^S : m \text{ is a minimum } \pi\text{-weight element of } \ddot{B}\}.$$

•

Roughly speaking, $I(m)$ corresponds to the first two conditions (56) and (57) in Theorem 5.2 and $W(m)$ to the third condition (58). However, there is an essential difference between Proposition 5.5 and Theorem 5.2. As already mentioned right after Proposition 5.4, each of $I(m)$ and $W(m)$ varies with the choice of $m$, while their intersection is uniquely determined and equal to $\Pi$. In this sense, the description of $\Pi$ in Proposition 5.5 is not canonical. Theorem 5.2 is a much stronger statement, giving a canonical description of $\Pi$ without reference to a particular primal optimal solution.

**Remark 5.1.** Proposition 5.5 above is equivalent to Proposition 6.7 of Part I [8], though in a slightly different form. Recall the optimality criteria there:[6]

(O1) $m(s) \in \{\lfloor \pi(s)/2 \rfloor, \lceil \pi(s)/2 \rceil\}$ for each $s \in S$,

(O2) each strict $\pi$-top-set is $m$-tight with respect to $p$.

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 $W(m)$ to the second criterion (O2) is a well-known characterization of a minimum weight base. •

## 5.3 Structure of primal optimal solutions to square-sum minimization

We now turn to the primal (minimization) problem of (55).

Let $\mathrm{dm}(\ddot{B})$ denote the set of the dec-min elements of $\ddot{B}$. By Theorem 3.3, $\mathrm{dm}(\ddot{B})$ coincides with the set of primal optimal solutions for (55). According to the general result in Proposition 5.4 (3), a representation of $\mathrm{dm}(\ddot{B})$ in the form of $\mathrm{dm}(\ddot{B}) = \partial f^{\bullet}(\hat{\pi}) \cap \partial h^{\circ}(\hat{\pi})$ is obtained by choosing any dual optimal solution $\hat{\pi}$.

[6]For a given vector $\pi$ in $\mathbf{R}^S$, we call a non-empty set $X \subseteq S$ a **$\pi$-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 $\pi$-top set**. We call a subset $X \subseteq S$ **$m$-tight** with respect to $p$ if $\widetilde{m}(X) = p(X)$.

For $\psi(\ell) = \lfloor \ell/2 \rfloor \cdot \lceil \ell/2 \rceil$ we have $\partial\psi(\ell) = \{\lfloor \ell/2 \rfloor, \lceil \ell/2 \rceil\}$. Therefore,

$$\partial f^{\bullet}(\pi) = \{x \in \mathbf{Z}^S : x(s) \in \{\lfloor \pi(s)/2 \rfloor, \lceil \pi(s)/2 \rceil\} \ (s \in S)\}. \tag{68}$$

Since $x \in \partial h^{\circ}(\pi)$ if and only if $\pi \in \partial h(x)$, we have

$$\partial h^{\circ}(\pi) = \{x \in \mathbf{Z}^S : x \text{ is a minimum } \pi\text{-weight element of } \ddddot{B}\}. \tag{69}$$

Then Proposition 5.4 (3) gives the following representation of the set of dec-min elements $\mathrm{dm}(\ddddot{B})$.

**Proposition 5.6.** *Let $\hat{\pi}$ be any dual optimal solution to* (55)*. 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*

$$T(\hat{\pi}) = \{m \in \mathbf{Z}^S : m(s) \in \{\lfloor \hat{\pi}(s)/2 \rfloor, \lceil \hat{\pi}(s)/2 \rceil\} \ (s \in S)\},$$

$$\ddddot{B}^{\circ}(\hat{\pi}) = \{m \in \ddddot{B} : x \text{ is a minimum } \hat{\pi}\text{-weight element of } \ddddot{B}\}.$$

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

# 6 Comparison of continuous and discrete cases

While our present study is focused on the discrete case for an M-convex set $\ddddot{B}$, the continuous case for a base-polyhedron $B$ was investigated by Fujishige [10] around 1980 under the name of lexicographically optimal bases, as a generalization of lexicographically optimal maximal flows considered by Megiddo [27]. Lexicographically optimal bases are discussed in detail in [11, Section 9]. Later in game theory Dutta–Ray [5] treated majorization ordering in the continuous case under the name of egalitarian allocation; see also Dutta [4]. See also the survey of related papers in Appendix B.

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 [10, 11]. 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 [11].

**Weighting** In Case **R** a weight vector can be introduced to define and analyze lexico-optimality, while this is not the case with Case **Z** treated in this paper. In the following comparisons we always assume that no weighting is introduced in Cases **R** and **Z**.

**Decreasing minimality and increasing maximality** In Case **Z** decreasing minimality in $\overset{....}{B}$ is equivalent to increasing maximality. This statement is also true in Case **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 **Z**) and in $B$ (in Case **R**).

**Square-sum minimization** In both Cases **Z** and **R**, a dec-min element is characterized as a minimizer of square-sum of the components $\Phi(x) = \sum[x(s)^2 : s \in S]$. In Case **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 **R** and **Z**. In Case **R** the dec-min element of $B$ is uniquely determined, and is given by the minimum norm point of $B$. In Case **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.

**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[7]

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

[7] $\overset{....}{B}$ 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 $(1,1,0,0)$.

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 a dec-min element. 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 **Z** we have the min-max identity (15):

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

In Case **R** the corresponding formula is

$$\min\{\sum[m(s)^2 : s \in S] : m \in B\} = \max\{\hat{p}(\pi) - \sum_{s\in S} \left(\frac{\pi(s)}{2}\right)^2 : \pi \in \mathbf{R}^S\}, \tag{70}$$

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 (70) has never been shown in the literature.

**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 $\beta$ is an essential value for Case **Z** if and only if there exists a critical value $\lambda$ 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 from the principal partition $\{\hat{S}_1, \hat{S}_2, \ldots, \hat{S}_r\}$ as an aggregation; 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 $\overset{....}{B}$ 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 [8]).

**Algorithm** In Case **Z** we have developed a strongly polynomial algorithm for finding a dec-min element of $\overset{....}{B}$ (Section 7 of Part I [8]). In Case **R** the decomposition algorithm of Fujishige [10] 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 $\overset{....}{B}$ with the box (interval) $[\ell, u]$ by $\overset{....}{B}{}^{u}_{\ell}$. The dec-min element of $\overset{....}{B}{}^{u}_{\ell}$ is also a dec-min element of $\overset{....}{B}$, since the box $[\ell, u]$ contains all dec-min elements of $\overset{....}{B}$ by Theorem 6.6. Since $0 \leq u(s) - \ell(s) \leq 1$ for all $s \in S$, $\overset{....}{B}{}^{u}_{\ell}$ can be regarded as a matroid translated by $\ell$, i.e., $\overset{....}{B}{}^{u}_{\ell} = \{\ell + \chi_L : L \text{ is a base of } M\}$, where $M$ is a matroid. Therefore, the dec-min element of $\overset{....}{B}{}^{u}_{\ell}$ 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 [11, 14, 18, 23] we can find continuous relaxation algorithms that are strongly polynomial for special classes of base-polyhedra; see Appendix B for details.

## 6.2 Review of the principal partition

As is pointed out by Fujishige [10], the dec-min element in the continuous case is closely related to the principal partition. The principal partition is a structural theory for submodular functions developed mainly in Japan; Iri [21] is an early survey and Fujishige [12] provides a comprehensive historical and technical account. In this section we summarize the results that are relevant to the analysis of the dec-min element in the continuous case. Originally [10], 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 $\lambda$, 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

$$\begin{aligned} p(X) + p(Y) &\leq p(X \cap Y) + p(X \cup Y), \\ \lambda|X| + \lambda'|Y| &= \lambda|X \cap Y| + \lambda'|X \cup Y| + (\lambda - \lambda')|X - Y| \\ &\geq \lambda|X \cap Y| + \lambda'|X \cup Y|. \end{aligned} \tag{71}$$

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 (71), which implies $|X - Y| = 0$, i.e., $X \subseteq Y$. □

There are finitely many numbers $\lambda$ 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 $\lambda$ 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

$$\hat{S}_i = \max \mathcal{L}(\lambda_i) - \min \mathcal{L}(\lambda_i) \qquad (i = 1, 2, \ldots, r), \tag{72}$$

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

$$\hat{S}_i = L(\lambda_i - \varepsilon) - L(\lambda_i) \tag{73}$$

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 [10]). *$m_{\mathbf{R}}(s) = \lambda_i$ if $s \in \hat{S}_i$, where $i = 1, 2, \ldots, r$.* •

## 6.3 New characterization of the canonical partition

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.[8]

In Part I [8], 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 [8]. Let $p : 2^S \to \mathbf{Z} \cup \{-\infty\}$ be an integer-valued supermodular function, and $C_0 = \emptyset$. For $j = 1, 2, \ldots, q$, define

$$\beta_j = \max \left\{ \left\lceil \frac{p(X \cup C_{j-1}) - p(C_{j-1})}{|X|} \right\rceil : \emptyset \neq X \subseteq \overline{C_{j-1}} \right\}, \tag{74}$$

$$h_j(X) = p(X \cup C_{j-1}) - (\beta_j - 1)|X| - p(C_{j-1}) \qquad (X \subseteq \overline{C_{j-1}}), \tag{75}$$

$$S_j = \text{smallest subset of } \overline{C_{j-1}} \text{ maximizing } h_j, \tag{76}$$

$$C_j = C_{j-1} \cup S_j, \tag{77}$$

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 $C_q = S$. The subset $C_j$ defined by the above recurrence relations admits in fact a direct characterization as the smallest maximizer of $p(X) - (\beta_j - 1)|X|$.

[8] 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 [8]).

**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 $p(X) - (\beta_j - 1)|X|$ over all subsets $X$ of $S$.*

*Proof.* The monotonicity of the $\beta$-values in (1) is already shown in Section 5 of Part I [8], but the following proof includes an alternative proof of (1) as well. For $j = 2, 3, \ldots, q$ we prove (i) $\beta_{j-1} > \beta_j$ and (ii) $C_j$ is the smallest maximizer of $p(X) - (\beta_j - 1)|X|$. First note that (ii) holds for $j = 1$ (by definition). Let $j \geq 2$.

(i) We prove $\beta_{j-1} > \beta_j$ assuming (ii) for $j - 1$, i.e., under the assumption that $C_{j-1}$ is the smallest maximizer of $p(X) - (\beta_{j-1} - 1)|X|$. By (74), $\beta_{j-1} > \beta_j$ if and only if $\beta_{j-1} > \left\lceil \frac{p(X \cup C_{j-1}) - p(C_{j-1})}{|X|} \right\rceil$ for all $X$ with $\emptyset \neq X \subseteq \overline{C_{j-1}}$. Let $X$ be such a subset. We use a simpler notation $C = C_{j-1}$. Then

$$
\begin{aligned}
& \beta_{j-1} > \left\lceil \frac{p(X \cup C) - p(C)}{|X|} \right\rceil \\
\iff \; & \beta_{j-1} - 1 \geq \frac{p(X \cup C) - p(C)}{|X|} \\
\iff \; & p(X \cup C) - p(C) \leq (\beta_{j-1} - 1)|X| \\
\iff \; & p(X \cup C) - (\beta_{j-1} - 1)|X \cup C| \leq p(C) - (\beta_{j-1} - 1)|C|.
\end{aligned}
$$

The last inequality holds by the assumption of (ii) for $j - 1$. We have thus shown $\beta_{j-1} > \beta_j$.

(ii) We next prove that $C_j$ is the smallest maximizer of $p(X) - (\beta_j - 1)|X|$. For notational simplicity, we define $\beta = \beta_{j-1} - 1$, $\beta' = \beta_j - 1$, $C = C_{j-1}$, and $C' = C_j$. By definition, $C' = C \cup S_j$ is the smallest maximizer of $p(X) - \beta'|X|$ among all subsets $X$ containing $C$. In the following we prove that any maximizer $Z$ of $p(X) - \beta'|X|$ is a superset of $C$. By supermodularity we have $p(Z) + p(C) \leq p(Z \cup C) + p(Z \cap C)$, which implies

$$
p(Z) - \beta'|Z| \leq \big(p(Z \cup C) - \beta'|Z \cup C|\big) + \big(p(Z \cap C) - p(C) + \beta'|C - Z|\big). \tag{78}
$$

For the first term on the right-hand side we have

$$
p(Z \cup C) - \beta'|Z \cup C| \leq p(C') - \beta'|C'|, \tag{79}
$$

since $Z \cup C \supseteq C$ and $C'$ is a maximizer of $p(X) - \beta'|X|$ among $X$ containing $C$. For the second term on the right-hand side of (78) we have

$$
\begin{aligned}
& p(Z \cap C) - p(C) + \beta'|C - Z| \\
& \leq p(Z \cap C) - p(C) + \beta|C - Z| \\
& = \big(p(Z \cap C) - \beta|Z \cap C|\big) - \big(p(C) - \beta|C|\big) \leq 0,
\end{aligned} \tag{80}
$$

since $\beta' < \beta$ by (i) and $C$ is a maximizer of $p(X) - \beta|X|$. Combining (78), (79), and (80) we obtain $p(Z) - \beta'|Z| \leq p(C') - \beta'|C'|$. By the choice of $Z$ we have equality here, which occurs only when $\beta'|C - Z| = \beta|C - Z|$, i.e., $Z \supseteq C$. Therefore, any maximizer of $p(X) - \beta'|X|$ is a superset of $C$. □

For any integer $\beta$, 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.)

We consider the family $\{L(\beta) : \beta \in \mathbf{Z}\}$ of the smallest maximizers of $p(X) - \beta|X|$ for all integers $\beta$. The canonical chain is contained in this family, since $C_j = L(\beta_j - 1)$ ($j = 1, 2, \ldots, q$) by Proposition 6.3(2).

We now state the key property of the essential value-sequence $\beta_1 > \beta_2 > \cdots > \beta_q$ defined by (74)–(77).

**Proposition 6.4.** *As $\beta$ is decreased from $+\infty$ to $-\infty$ (or from $\beta_1$ to $\beta_q - 1$), the smallest maximizer $L(\beta)$ is monotone nondecreasing. We have $L(\beta) \neq L(\beta - 1)$ if and only if $\beta$ is equal to an essential value $\beta_j$. That is,*[9]

$$\emptyset = L(\beta_1) \subset L(\beta_1 - 1) = \cdots = L(\beta_2) \subset L(\beta_2 - 1) = \cdots = L(\beta_q) \subset L(\beta_q - 1) = S. \quad (81)$$

*Proof.* By the monotonicity shown in Proposition 6.1, it suffices to prove (i) $L(\beta_1) = \emptyset$, (ii) $L(\beta_{j-1} - 1) \supseteq L(\beta_j)$ for $j = 2, \ldots, q$, and (iii) $L(\beta_j) \neq L(\beta_j - 1)$ 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 $C = C_{j-1}$ and $\beta' = \beta_j$. Define $h(Y) = p(Y) - \beta'|Y|$ for any subset $Y$ of $S$, and let $A$ be the smallest maximizer of $h$, which means $A = L(\beta') = L(\beta_j)$. For any nonempty subset $X$ of $\overline{C}$ $(= S - C)$ we have

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

which implies $p(X \cup C) - \beta'|X \cup C| \leq p(C) - \beta'|C|$. In other words, we have

$$h(Y) \leq h(C) \qquad \text{for all } \ Y \supseteq C. \quad (82)$$

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 (82). 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., $A \subseteq C$. Finally we recall $A = L(\beta_j)$ and $C = C_{j-1} = L(\beta_{j-1} - 1)$, to obtain $L(\beta_j) \subseteq L(\beta_{j-1} - 1)$.

(iii) Let $1 \leq j \leq q$. We continue to write $C = C_{j-1}$. Take a nonempty subset $Z$ of $\overline{C}$ for which

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

Then we have

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

which implies

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

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

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

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 $\beta$. There are finitely many $\beta$ 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 $C_j = L(\beta_j)$ 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 $\beta$ is an essential value if and only if there exists a critical value $\lambda$ 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 from the principal partition $\{\hat{S}_1, \hat{S}_2, \ldots, \hat{S}_r\}$ as an aggregation; it is given as*

$$S_j = \bigcup_{i \in I(j)} \hat{S}_i \qquad (j = 1, 2, \ldots, q). \tag{83}$$

(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 [8]). 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 $\lambda_i$ 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 [8]. 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

$$\text{There exists a dec-min element } m_{\mathbf{Z}} \text{ of } \overset{....}{B} \text{ satisfying } \lfloor m_{\mathbf{R}} \rfloor \leq m_{\mathbf{Z}} \leq \lceil m_{\mathbf{R}} \rceil, \tag{84}$$

where $m_{\mathbf{R}}$ is the minimum norm point of $B$. This statement (84) should not be confused with Proposition 6.7 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). •

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 $(0,3)$ abd $(2,1)$. 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 critical value $\lambda_1 = 3/2$ and the associated sublattice is $\mathcal{L}(\lambda_1) = \{\emptyset, S\}$, where $r = 1$. 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

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

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 [8], 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 $\ell$ 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.,

$$\ell \leq m_{\mathbf{Z}} \leq u \tag{85}$$

for some dec-min element $m_{\mathbf{Z}}$ of $\overset{....}{B}$. We denote the intersection of $\overset{....}{B}$ and $[\ell, u]$ by $\overset{....}{B}{}^u_\ell$. Then the dec-min element of $\overset{....}{B}{}^u_\ell$ is a dec-min element of $\overset{....}{B}$. Since $\mathbf{0} \leq u - \ell \leq \mathbf{1}$, $\overset{....}{B}{}^u_\ell$ can be regarded as a matroid translated by $\ell$, i.e., $\overset{....}{B}{}^u_\ell = \{\ell + \chi_L : L \text{ is a base of } M\}$ for some matroid $M$. Therefore, the dec-min element of $\overset{....}{B}{}^u_\ell$ 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 $\ell$ 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 (85). With this choice of $(\ell, u)$, $\overset{....}{B}{}^u_\ell$ contains all dec-min elements of $\overset{....}{B}$. The decomposition algorithm of Fujishige [10] (see also [11, 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 $\ddot{B}$ 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 [14] (see also [11, 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 [14].

**Proposition 6.7** (Groenevelt [14])**.** 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 $\ddot{B}$ 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 $\Phi$ on $\ddot{B}$. We can take any $z_i$ as $m_{\mathbf{Z}}$. □

By Proposition 6.7 we can take $\ell = \lfloor \overline{m}_{\mathbf{R}} \rfloor$ and $u = \lceil \overline{m}_{\mathbf{R}} \rceil$ in (85). In this case, however, $\ddot{B}_{\ell}^{u}$ may not contain all dec-min elements of $\ddot{B}$. The complexity of computing $\overline{m}_{\mathbf{R}}$ is not fully analyzed in the literature [11, 14, 23]. 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 [18], Ibaraki–Katoh [20], Katoh–Ibaraki [22], Katoh–Shioura–Ibaraki [23]). The continuous relaxation approach for the case of discrete variables is considered, e.g., by Hochbaum [16] and Hochbaum–Hong [19]. A more recent paper by Moriguchi–Shioura–Tsuchimura [29] discusses this approach in a more general context of M-convex function minimization in discrete convex analysis. It is known ([19, 29], [23, 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 notation of [23]. •

# A Non-separable symmetric convex minimization

For any symmetric strictly convex function $\Phi$ (not necessarily separable), the integral dec-min elements of an integral base-polyhedron are characterized as the minimizers of $\Phi$. We establish this characterization in this appendix.

Let $S = \{1, 2, \ldots, n\}$ and $\Phi : \mathbf{Z}^S \to \mathbf{R} \cup \{+\infty\}$. We say that function $\Phi$ is *symmetric* if

$$\Phi(x(1), x(2), \ldots, x(n)) = \Phi(x(\sigma(1)), x(\sigma(2)), \ldots, x(\sigma(n))) \tag{86}$$

for all permutations $\sigma$ of $(1, 2, \ldots, n)$, and *strictly convex* if

$$t\Phi(x) + (1-t)\Phi(y) > \Phi(tx + (1-t)y) \tag{87}$$

whenever $x, y \in \mathrm{dom}\,\Phi$, $0 < t < 1$, and $tx + (1-t)y$ is an integral vector.

We consider minimizing $\Phi$ over $\overset{....}{B}$, the set of integral points of an integral base-polyhedron $B$. It is noted that the strict convexity (87) does not imply the uniqueness of minimizers. For example, the square sum $\Phi(x) = \sum_{s\in S} x(s)^2$ is strictly convex, but its minimizers on $\overset{....}{B} = \{x \in \mathbf{Z}^S : \sum_{s\in S} x(s) = 1\}$ are given by $\chi_s$ for all $s \in S$.

**THEOREM A.1.** *Let $\overset{....}{B}$ be the set of integral points of an integral base-polyhedron, $\Phi$ be a symmetric strictly convex function, and $x \in \overset{....}{B}$.*
(1) *$\Phi$ has a minimizer on $\overset{....}{B}$.*
(2) *$x$ is a dec-min element of $\overset{....}{B}$ if and only if it is a minimizer of $\Phi$ on $\overset{....}{B}$.*

*Proof.* We first assume that $\overset{....}{B}$ is bounded. In this case the existence of a minimizer in (1) is obvious.

(2) "minimizer of $\Phi \Rightarrow$ dec-min": Let $x$ be a minimizer of $\Phi$. To prove that $x$ is dec-min, it suffices, by Theorem 3.1, to show

$$x(t) \geq x(s) + 2 \implies x + \chi_s - \chi_t \notin \overset{....}{B} \tag{88}$$

for all $s, t \in S$. Assume, indirectly, that $x(t) \geq x(s) + 2$ and $z := x + \chi_s - \chi_t \in \overset{....}{B}$ for some $s, t \in S$. Let $\alpha = x(t) - x(s)$, where $\alpha \geq 2$. Define $y = x + \alpha(\chi_s - \chi_t)$, and note that $y$ is obtained from $x$ by interchanging the components at $s$ and $t$, and therefore, $\Phi(x) = \Phi(y)$ by symmetry (86). We have $y \in \mathrm{dom}\,\Phi$. However, we may or may not have $y \in \overset{....}{B}$.

We now consider the strict convexity (87). For $t = 1 - 1/\alpha$ we have

$$tx + (1-t)y = \left(1 - \frac{1}{\alpha}\right)x + \frac{1}{\alpha}(x + \alpha(\chi_s - \chi_t)) = x + \chi_s - \chi_t = z \in \mathbf{Z}^S.$$

Therefore, $\Phi(x) = t\Phi(x) + (1-t)\Phi(y) > \Phi(z)$, a contradiction to $x$ being a minimizer of $\Phi$ on $\overset{....}{B}$.

(2) "dec-min $\Rightarrow$ minimizer of $\Phi$": Let $x$ be a dec-min element of $\overset{....}{B}$. Let $y$ be a minimizer of $\Phi$, which is dec-min by the above argument. Then $x{\downarrow} = y{\downarrow}$. It follows from this and symmetry (86) that $\Phi(x) = \Phi(x{\downarrow}) = \Phi(y{\downarrow}) = \Phi(y)$. Therefore, $x$ is also a minimizer of $\Phi$.

Next we consider the case of unbounded $\overset{....}{B}$. Let $k_0$ be an integer such that $\{z \in \mathbf{Z}^S : \|z\|_\infty \leq k_0\}$ contains all dec-min elements of $\overset{....}{B}$. For $k \geq k_0$ we denote the intersection of $\overset{....}{B}$ and $\{z \in \mathbf{Z}^S : \|z\|_\infty \leq k\}$ by $\overset{....}{B}_k$. Obviously, $\overset{....}{B}$ and $\overset{....}{B}_k$ have the same dec-min elements. By the argument for the bounded case, the dec-min elements of $\overset{....}{B}_k$ are exactly the minimizers of $\Phi$ over $\overset{....}{B}_k$. Since this is true for all $k \geq k_0$, these dec-min elements are exactly the minimizers of $\Phi$ over $\overset{....}{B}$. □

We now consider (non-strict) convex function $\Phi$, which, by definition, satisfies

$$t\Phi(x) + (1-t)\Phi(y) \geq \Phi(tx + (1-t)y) \tag{89}$$

whenever $x, y \in \mathrm{dom}\,\Phi$, $0 < t < 1$, and $tx+(1-t)y$ is an integral vector. As easily imagined, a dec-min element is a minimizer of such function.

**Corollary A.2.** *Let $\ddot{B}$ be the set of integral points of an integral base-polyhedron, and $\Phi$ be a symmetric convex function. Then a dec-min element of $\ddot{B}$ is a minimizer of $\Phi$ on $\ddot{B}$.*

*Proof.* For $a > 0$ define $\Phi_a(x) = \Phi(x) + a \sum_{s \in S} x(s)^2$, which is symmetric strictly convex. Let $y$ be a dec-min element of $\ddot{B}$. By Theorem A.1, $y$ is a minimizer of $\Phi_a$ for any $a > 0$. If $y$ is not a minimizer of $\Phi$, there exists $z \in \ddot{B}$ with $\Phi(y) > \Phi(z)$. We can take a sufficiently small $a > 0$ for which $a \left| \sum_{s \in S} y(s)^2 - \sum_{s \in S} z(s)^2 \right| < \Phi(y) - \Phi(z)$. Then $\Phi_a(y) > \Phi_a(z)$, which is a contradiction to $y$ being a minimizer of $\Phi_a$. Therefore, $y$ must be a minimizer of $\Phi$. □

The following examples show the application of Corollary A.2.

**Example A.1** (max-component)**.** The function $\Phi(x) = \max\{x(1), \ldots, x(n)\}$ is a symmetric convex function. Therefore, a dec-min element of $\ddot{B}$ is a max-minimizer of $\ddot{B}$. •

**Example A.2** (min-component)**.** The function $\Phi(x) = -\min\{x(1), \ldots, x(n)\}$ is a symmetric convex function. Therefore, a dec-min element of $\ddot{B}$ is a min-maximizer of $\ddot{B}$. •

**Example A.3** (range)**.** The function $\Phi(x) = \max\{x(1), \ldots, x(n)\} - \min\{x(1), \ldots, x(n)\}$ representing the range of the components of a vector is a symmetric convex function. Therefore, a dec-min element of $\ddot{B}$ minimizes the range of the components of a vector in $\ddot{B}$ •

**Example A.4** ($k$ largest component sum)**.** For an integer $k$ with $1 \leq k \leq n$, the function $\Phi(x)$ representing the sum of the $k$ largest components of a vector $x$ is a symmetric convex function. Using the notations introduced in Section 2.1, we can express $\Phi(x) = \overline{x}(k) = \sum_{i=1}^{k} x{\downarrow}(i)$. By Corollary A.2, a dec-min element of $\ddot{B}$ is a minimizer of this function, which is given already in Theorem 3.5 of Part I [8]. This fact means that every dec-min element of $\ddot{B}$ is a least majorized element of $\ddot{B}$, which is stated already in Theorem 2.6. •

**Example A.5** (2-separable convex)**.** For $a, b, c \geq 0$, the function defined by

$$\Phi(x) = a \sum_{i=1}^{n} |x(i)| + b \sum_{i \neq j} |x(i) - x(j)| + c \sum_{i \neq j} |x(i) + x(j)|$$

is a symmetric convex function. More generally, a symmetric *2-separable convex* function

$$\Phi(x) = \sum_{i=1}^{n} \varphi_0(x(i)) + \sum_{i \neq j} \varphi_-(|x(i) - x(j)|) + \sum_{i \neq j} \varphi_+(x(i) + x(j)),$$

where $\varphi_0, \varphi_-, \varphi_+ : \mathbf{Z} \to \mathbf{R}$ are discrete convex functions, satisfies (86) and (89). By Corollary A.2, a dec-min element of $\ddot{B}$ is a minimizer of such functions over $\ddot{B}$. The minimization of 2-separable convex functions is investigated in depth by Hochbaum and others [1, 17, 18] using network flow techniques. 2-separable diff-convex functions appear in the convex dual of the minimum cost network flow problem [1, 17, 18, 32]. •

**Remark A.1.** Theorem A.1 is a discrete counterpart of a result of Maruyama [26] for the continuous case. Our proof is an adaptation of the proof of [34, Corollary 13]. •

**Remark A.2.** Since non-separable symmetric convex functions are not necessarily M$^\natural$-convex, we cannot use the Fenchel-type duality theorem to obtain a min-max formula for the minimization of non-separable symmetric convex functions. •

**Remark A.3.** In connection to Examples A.1, A.2, and A.3 it is worth noting that more general nonlinear optimization problems in discrete variables are considered by Fujishige, Katoh, and Ichimori [13]; see also [11, Sections 10.2 and 11.2]. For each $i$ let $h_i : \mathbf{Z} \to \mathbf{R}$ be a monotone nondecreasing function on $\mathbf{Z}$ such that $\lim_{t\to+\infty} h_i(t) = +\infty$ and $\lim_{t\to-\infty} h_i(t) = -\infty$. Let $g : \mathbf{R}^2 \to \mathbf{R}$ be a function such that $g(u, v)$ is monotone nondecreasing in $u$ and monotone nonincreasing in $v$. Weakly polynomial algorithms are given for the following problems:

$$\text{Maximize} \quad \min_{1\leq i\leq n} h_i(x(i)) \quad \text{subject to} \quad x \in \overset{....}{B}; \tag{90}$$

$$\text{Minimize} \quad \max_{1\leq i\leq n} h_i(x(i)) \quad \text{subject to} \quad x \in \overset{....}{B}; \tag{91}$$

$$\text{Minimize} \quad g(\, \max_{1\leq i\leq n} h_i(x(i)),\ \min_{1\leq i\leq n} h_i(x(i))\,) \quad \text{subject to} \quad x \in \overset{....}{B}. \tag{92}$$

Note that the objective functions of these problems are not symmetric convex in general. Examples A.1, A.2, and A.3, respectively, are special cases of the above problems with $h_i(t) = t$ and $g(u, v) = u - v$. •

# B 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 [36] (1971), Megiddo [27] (1974), Fujishige [10] (1980), Groenevelt [14] (1985, 1991), Federgruen–Groenevelt [6] (1986), Ibaraki–Katoh [20] (1988), Dutta–Ray [5] (1989), Fujishige [11] (1991, 2005), Hochbaum [16] (1994), and Tamir [35] (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 [36] and "Least weakly submajorized" is used in Tamir [35]. These terms are not used in Marshall–Olkin–Arnold [25]. Dutta–Ray [5] uses "egalitarian" and does not use "majorization." The term "lexicographically optimal" in Veinott [36], Megiddo [27, 28], and Fujishige [10, 11] 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 $\mathbf{R}$ and $\mathbf{Z}$), 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) [36]**

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) [27]**

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 [36], but the difference 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 an inc-max element using a chain of cuts in the network (Theorem 4.6). Computational aspect is discussed in the companion paper [28], which gives an algorithm of complexity $O(n^5)$.

**Fujishige (1980) [10]**

This is the first paper that deals with base-polyhedra, beyond network flows. It considers Case **R** only. The lexicographically 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) [14]**

The technical report appeared in 1985, and the journal version in 1991. Already the technical report was influential, cited by [11, 1st ed.], [16], and [20].

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 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 [14, 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 $\tau_1$ = the number of operations needed to solve a single constraint problem, and $\tau_2$ = 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 $\chi$ 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 [15, 2nd edition].

**Federgruen–Groenevelt (1986) [6]**

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) [20]**

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 [10], Groenevelt [14], and Federgruen–Groenevelt [6]. In particular, Theorem 9.2.2 [20, p.156] states that the decomposition algorithm runs in polynomial time in $|E|$ and $\log M$, where $E$ is the ground set, $r$ is the submodular function for the submodular constraint, and $M$ is an upper bound on $r(E)$.

The contents of Chapter 9 are updated in a handbook chapter by Ibaraki–Katoh [22] in 1998. Its revised version by Katoh–Shioura–Ibaraki [23] in 2013 incorporates the views from discrete convex analysis.

**Dutta–Ray (1989) [5]**

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 [35], 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 simple combination of the results of Groenevelt [14] (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 [5] and Groenevelt [14] are unaware of each other; see Table 1 at the end of Appendix. We also note that Fujishige [10] deals with quadratic functions only, and hence the results of [10] do not entail the existence of a least majorized element.

**Fujishige (1st ed., 1991; 2nd ed. 2005) [11]**

This book offers a comprehensive exposition of the results of Fujishige [10] about the lexico-optimal (inc-max) element of a base-polyhedron in Case **R**. For Case **Z**, however, 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 [14] 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**, but no explicit statement for Case **Z** is found.

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). Majorization concept is not treated, either, though a reference to Dutta–Ray [5] is added in the second edition (Section 9.2).

**Hochbaum (1994) [16]**

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 [19]). The survey paper by Hochbaum [18] is informative and useful.

**Tamir (1995) [35]**

This paper 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 complexity in Case **Z**.

**Acknowledgement** The authors thank S. Fujishige and S. Iwata for discussion about the history of convex minimization over base-polyhedra. We also thank A. Shioura and A. Tamura for reading the manuscript and S. Moriguchi for checking examples.

Table 1: Referencing relations between papers

| | Vei [36] | Meg [27] | Fuj [10] | Gro [14] | F-G [6] | I-K [20] | D-R [5] | Fuj [11] | Hoc [16] | Tam [35] |
|---|---|---|---|---|---|---|---|---|---|---|
| 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)

# References

[1] R.K. Ahuja, D.S. Hochbaum, J. B. Orlin, A cut-based algorithm for the nonlinear dual of the minimum cost network flow problem. Algorithmica **39**, 189–208 (2004).

[2] K. Ando, Weak majorizations on finite jump systems. Mimeo (1996) Available from author's home page.

[3] B.C. Arnold, J.M. Sarabia, Majorization and the Lorenz Order with Applications in Applied Mathematics and Economics. Springer International Publishing, Cham (2018), (1st edn., 1987)

[4] B. Dutta, The egalitarian solution and reduced game properties in convex games. International Journal of Game Theory **19**, 153–169 (1990)

[5] B. Dutta, D. Ray, A concept of egalitarianism under participation constraints. Econometrica **57**, 615–635 (1989)

[6] A. Federgruen, H. Groenevelt, The greedy procedure for resource allocation problems: necessary and sufficient conditions for optimality. Operations Research **34**, 909–918 (1986)

[7] A. Frank, Connections in Combinatorial Optimization. Oxford University Press, Oxford (2011)

[8] A. Frank, K. Murota, Discrete decreasing minimization, Part I, Base-polyhedra with applications in network optimization. arXiv: 1808.07600 (August 2018)

[9] A. Frank, K. Murota, Discrete decreasing minimization, Part III, Submodular flows and the intersection of two base-polyhedra, in preparation

[10] S. Fujishige, Lexicographically optimal base of a polymatroid with respect to a weight vector. Mathematics of Operations Research **5**, 186–196 (1980)

[11] S. Fujishige, 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)

[12] S. Fujishige, Theory of principal partitions revisited. In: Cook, W., Lovász, L., Vygen, J. (eds.) Research Trends in Combinatorial Optimization, pp. 127–162. Springer, Berlin (2009)

[13] S. Fujishige, N. Katoh, N., T. Ichimori, The fair resource allocation problem with submodular constraints. Mathematics of Operations Research **13**, 164–173 (1988)

[14] H. Groenevelt, 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)

[15] M. Grötschel, L. Lovász, A. Schrijver, Geometric Algorithms and Combinatorial Optimization, 2nd edn. Springer, Berlin (1993)

[16] D.S. Hochbaum, Lower and upper bounds for the allocation problem and other nonlinear optimization problems. Mathematics of Operations Research **19**, 390–409 (1994)

[17] D.S. Hochbaum, Solving integer programs over monotone inequalities in three variables: A framework for half integrality and good approximations. European Journal of Operational Research **140**, 291–321 (2002)

[18] D.S. Hochbaum, Complexity and algorithms for nonlinear optimization problems. Annals of Operations Research **153**, 257–296 (2007)

[19] D.S. Hochbaum, S.-P. Hong, About strongly polynomial time algorithms for quadratic optimization over submodular constraints. Mathematical Programming **69**, 269–309 (1995)

[20] T. Ibaraki, N. Katoh, Resource Allocation Problems: Algorithmic Approaches. MIT Press, Boston (1988)

[21] M. Iri, 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)

[22] N. Katoh, T. Ibaraki, 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)

[23] N. Katoh, A. Shioura, T. Ibaraki, 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)

[24] A. Levin, S. Onn, Shifted matroid optimization. Operations Research Letters **44**, 535-539 (2016)

[25] A.W. Marshall, I. Olkin, B.C. Arnold, Inequalities: Theory of Majorization and Its Applications, 2nd edn. Springer, New York (2011), (1st edn., 1979)

[26] F. Maruyama, A unified study on problems in information theory via polymatroids. Graduation Thesis, University of Tokyo, Japan, 1978. (In Japanese.)

[27] N. Megiddo, Optimal flows in networks with multiple sources and sinks. Mathematical Programming **7**, 97–107 (1974)

[28] N. Megiddo, A good algorithm for lexicographically optimal flows in multi-terminal networks. Bulletin of the American Mathematical Society **83**, 407–409 (1977)

[29] S. Moriguchi, A. Shioura, N. Tsuchimura, M-convex function minimization by continuous relaxation approach—Proximity theorem and algorithm. SIAM Journal on Optimization **21**, 633–668 (2011)

[30] K. Murota, Convexity and Steinitz's exchange property. Advances in Mathematics **124**, 272–311 (1996)

[31] K. Murota, Discrete convex analysis. Mathematical Programming **83**, 313–371 (1998)

[32] K. Murota, Discrete Convex Analysis. SIAM Monographs on Discrete Mathematics and Applications, Vol. 10, Society for Industrial and Applied Mathematics, Philadelphia (2003)

[33] K. Murota, 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)

[34] K. Nagano, On convex minimization over base polytopes. In: Fischetti, M., Williamson, D.P. (eds.): Integer Programming and Combinatorial Optimization. Lecture Notes in Computer Science, vol. 4513, pp. 252–266 (2007)

[35] A. Tamir, Least majorized elements and generalized polymatroids. Mathematics of Operations Research **20**, 583–589 (1995)

[36] A.F. Veinott, Jr., Least $d$-majorized network flows with inventory and statistical applications. Management Science **17**, 547–567 (1971)