

Decreasing Minimization on Base-Polyhedra: Relation Between Discrete and Continuous Cases

András Frank* and Kazuo Murota†

February 2022 / April 2022

Abstract

This paper is concerned with the relationship between the discrete and the continuous decreasing minimization problem on base-polyhedra. The continuous version (under the name of lexicographically optimal base of a polymatroid) was solved by Fujishige in 1980, with subsequent elaborations described in his book (1991). The discrete counterpart of the dec-min problem (concerning M-convex sets) was settled only recently by the present authors, with a strongly polynomial algorithm to compute not only a single decreasing minimal element but also the matroidal structure of all decreasing minimal elements and the dual object called the canonical partition. The objective of this paper is to offer a complete picture on the relationship between the continuous and discrete dec-min problems on base-polyhedra by establishing novel technical results and integrating known results. In particular, we derive proximity results, asserting the geometric closeness of the decreasingly minimal elements in the continuous and discrete cases, by revealing the relation between the principal partition and the canonical partition. We also describe decomposition-type algorithms for the discrete case following the approach of Fujishige and Groenevelt.

Keywords: Base-polyhedron, M-convex set, Decreasing minimization, Lexicographic optimization, Principal partition, Decomposition algorithm.

Mathematics Subject Classification (2010): 90C27, 05C, 68R10

*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. ORCID: 0000-0001-6161-4848. The research was partially supported by the National Research, Development and Innovation Fund of Hungary (FK_18) – No. NKFI-128673.

†The Institute of Statistical Mathematics, Tokyo 190-8562, Japan. Also Faculty of Economics and Business Administration, Tokyo Metropolitan University, Tokyo 192-0397, Japan, e-mail: murota@tmu.ac.jp. ORCID: 0000-0003-1518-9152. The research was supported by JSPS KAKENHI Grant Number JP20K11697.

Contents

1	Introduction	3
2	Comparison of continuous and discrete cases	5
2.1	Underlying set	5
2.2	Decreasing minimality and increasing maximality	6
2.3	Characterizations	6
2.4	Uniqueness	7
2.5	Square-sum minimization	7
2.6	Principal and canonical partitions	8
2.7	Proximity	8
2.8	Algorithm	9
2.9	Weighting	10
3	Principal partition and canonical partition	10
3.1	Review of the principal partition	11
3.2	Review of the canonical partition	13
3.3	An alternative characterization of the canonical partition	13
3.4	Canonical partition from the principal partition	16
4	Proximity results and continuous relaxation algorithm	18
4.1	Proximity theorem using the fractional dec-min element	18
4.2	Proximity theorem using a piecewise-linear square-sum minimizer	19
4.3	Continuous relaxation algorithm	21
5	Decomposition algorithms for square-sum minimization on an M-convex set	23
5.1	General remarks	23
5.2	Groenevelt-type decomposition algorithm	24
5.3	Fujishige-type decomposition algorithm	28

1 Introduction

An element of a set of vectors, in general, is called **decreasingly minimal** (dec-min) if its largest component is as small as possible, within this, its second largest component is as small as possible, and so on. The term “decreasing minimization” means the problem of finding a dec-min element of a given set of vectors (or a minimum cost dec-min element with respect to a given linear cost-function). When the given set of vectors consists of integral vectors, this problem is also referred to as discrete decreasing minimization. In the literature, typically the term lexicographic optimization is used, but we prefer “decreasing minimization” because we also consider its natural counterpart “increasing maximization,” and the use of these two symmetric terms seems more appropriate to distinguish the two related notions. An element of a set of vectors is called **increasingly maximal** (inc-max) if its smallest component is as large as possible, within this, its second smallest component is as large as possible, and so on. For example, the vector $(2, 2, 1, 1)$ is both dec-min and inc-max in the set $\{(2, 2, 1, 1), (2, 2, 2, 0), (2, 1, 3, 0)\}$. In another set $\{(2, 0, 0, 0), (1, -1, 1, 1)\}$, the dec-min element is $(1, -1, 1, 1)$ and the inc-max element is $(2, 0, 0, 0)$.

The decreasing minimization (even its weighted form) on a base-polyhedron B was investigated by Fujishige [13] around 1980 under the name of lexicographically optimal bases, as a generalization of lexicographically optimal maximal flows considered by Megiddo [28]. (A lexicographically optimal base in [13] means an inc-max member of B , and a lexicographically optimal maximal flow in [28] means a maximum flow that is inc-max on the set of source-edges.) Lexicographically optimal bases are discussed in detail in the book of Fujishige [14, Sections 8 and 9]. Among others, two important features of base-polyhedra are discovered. The first is that the unique dec-min element of a base-polyhedron B is the unique inc-max element of B , while the second is that the unique dec-min element of B is the unique square-sum minimizer (that is, the minimum ℓ_2 -norm element) of B . These coincidences are surprising in the light that neither of the analogous statements hold for the intersection of two base-polyhedra [9]. Furthermore, the concept of principal partition of the ground-set defined by a base-polyhedron [15, 24] plays a crucial role, and the critical values associated with the partition characterizes the lexicographically optimal base (see Section 3.1). Fujishige developed two algorithms for finding the dec-min element of base-polyhedron B . His “monotone algorithm” [14, Section 9.2] is not polynomial in its original form but it can immediately be made strongly polynomial with the aid of the Newton–Dinkelbach algorithm [17, 34]. The other algorithm [14, Section 8.2], called “decomposition algorithm,” is strongly polynomial and does not rely on Newton–Dinkelbach. In game theory, the lexicographically optimal base was discussed under the name of egalitarian allocation by Dutta and Ray [5] with reference to the framework of majorization [1, 27]; see also Dutta [4].

The discrete version of the problem considers dec-min elements of the set of integral elements of an integral base-polyhedron, which set is called an M-convex set in discrete convex analysis [30, 31, 32]; the M-convex set arising from an integral base-polyhedron B is denoted as $\overset{\dots}{B}$. Discrete decreasing minimization on an M-convex set has been studied recently by the present authors [9, 10]. It was shown in [9] that there are interesting coincidences with the continuous case. For example, an element z of an M-convex set $\overset{\dots}{B}$ is dec-min if and only if z is inc-max and if and only if z is a square-sum minimizer. On the other hand, there are fundamental differences between the discrete and continuous cases. For example, it was shown in [9] that there may be several dec-min elements of an M-convex set $\overset{\dots}{B}$, and the set of dec-min elements of $\overset{\dots}{B}$ is itself an M-convex set, which can be obtained from a

matroid by translating the set of characteristic vectors of its bases by an integral vector. This property made it possible to solve algorithmically even the minimum cost dec-min problem for M -convex sets (a problem which makes no sense in the continuous case where the dec-min element is unique). Furthermore, as a discrete counterpart of the principal partition, the concept of “canonical partition” of the ground-set was introduced to characterize the set of all dec-min elements of an M -convex set (see Section 3.2). The subsequent paper [10] presented a strongly polynomial algorithm for computing a dec-min element of an M -convex set together with the canonical partition, and discussed applications to a variety of problems including graph orientations [7] and resource allocation problems [19, 23, 26]. This algorithm relied on a discrete variant of the Newton–Dinkelbach algorithm, and it may be considered a discrete counterpart of the “monotone algorithm” of Fujishige. While the dec-min problem on an M -convex set contains the discrete version of Megiddo’s problem as a special case, it does not capture an extension of Megiddo’s problem in which we seek a feasible integral flow that is inc-max (or dec-min) on an arbitrarily specified set of edges. This more general problem for network flows was investigated recently in [11], and a further generalization to submodular integral flows in [12].

The objective of this paper is to offer a complete picture on the relationship between the continuous and discrete dec-min problems on base-polyhedra by establishing novel technical results and integrating known results from the literature. As such this paper is partly a research paper and partly a survey paper.

Our novel technical results include a theorem that reveals a precise relation between the dual objects, namely, the principal partition of the ground-set in the continuous case and the canonical partition in the discrete case. This theorem is used to prove proximity results, asserting the geometric closeness of the decreasingly minimal elements in the continuous and discrete cases. The obtained proximity results, in turn, give rise to continuous relaxation algorithms for computing a dec-min element of an M -convex set in strongly polynomial time using the (fractional) dec-min element of a base-polyhedron.

Another major topic of this paper is concerned with decomposition-type algorithms for computing a dec-min element of an M -convex set. The decomposition method was initiated by Fujishige [13] for computing a dec-min element of a base-polyhedron, or equivalently for minimizing a quadratic function on a base-polyhedron, and was extended by Groenevelt [18] to separable convex functions on a base-polyhedron and also on the set of integral points of an integral base-polyhedron (namely, an M -convex set). With a scrutiny of what is known in the literature about these decomposition algorithms, we give precise description of two decomposition algorithms for an M -convex set in a self-contained manner, one based on Groenevelt [18] and the other based on Fujishige [14, Section 8.2].

The paper is organized as follows. Major ingredients and properties of decreasing minimization on base-polyhedra and M -convex sets are compared in Section 2. In Section 3 we reveal the precise relation between the principal partition and the canonical partition through a novel characterization of the canonical partition. Using these results, we derive, in Section 4, proximity results and a continuous relaxation algorithm for computing a dec-min element of an M -convex set. Finally, in Section 5, we deal with decomposition algorithms to compute a dec-min element of an M -convex set.

Notation We basically follow notation in [9]. Let S be a finite ground-set. For a vector $x \in \mathbf{R}^S$ or a function $x : S \rightarrow \mathbf{R}$, we define the set-function $\tilde{x} : 2^S \rightarrow \mathbf{R}$ by $\tilde{x}(Z) := \sum_{s \in Z} [x(s) : s \in Z]$ ($Z \subseteq S$). The characteristic (or incidence) vector of a subset $Z \subseteq S$ is denoted by

χ_Z , that is, $\chi_Z(s) = 1$ if $s \in Z$ and $\chi_Z(s) = 0$ otherwise. The vector with all components 1 is denoted by $\mathbf{1}$, while $\mathbf{0}$ is the zero vector. For any real number $\alpha \in \mathbf{R}$, $\lfloor \alpha \rfloor$ denotes the largest integer not larger than α , and $\lceil \alpha \rceil$ the smallest integer not smaller than α . This notation is extended to vectors by componentwise applications. For any integral polyhedron $P \subseteq \mathbf{R}^S$, we use the notation $\overset{\dots}{P}$ to denote the set of integral elements of P , that is, $\overset{\dots}{P} := P \cap \mathbf{Z}^S$, where $\overset{\dots}{P}$ may be pronounced ‘‘dotted P .’’ The notation is intended to refer intuitively to the set of lattice points of P .

2 Comparison of continuous and discrete cases

In this section we compare the decreasing minimization problems on a base-polyhedron and on an M-convex set in terms of various aspects. Most of them are based on our present knowledge from [9, 10, 13, 14], while some others serve as motivations for the investigations to be made in the present paper. From these comparisons, it may safely be said that the discrete case is, in spite of some important similarities, 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. The continuous case is referred to as Case \mathbf{R} and the discrete case as Case \mathbf{Z} . We use notation $m_{\mathbf{R}}$ and $m_{\mathbf{Z}}$ for a dec-min element in Case \mathbf{R} and Case \mathbf{Z} , respectively.

2.1 Underlying set

Let b be a set-function on a ground-set S with $b(\emptyset) = 0$, for which $b(X) = +\infty$ is allowed but $b(X) = -\infty$ is not. We say that b is **submodular** (or fully submodular) if the submodular inequality

$$b(X) + b(Y) \geq b(X \cap Y) + b(X \cup Y) \quad (2.1)$$

holds for every pair of subsets $X, Y \subseteq S$ with finite b -values. A set-function p with $p(\emptyset) = 0$, for which $p(X) = -\infty$ is allowed but $p(X) = +\infty$ is not, is said to be **supermodular** if $-p$ is submodular.

For a submodular set-function b on S with $b(S)$ finite, the **base-polyhedron** B is defined by

$$B = B(b) := \{x \in \mathbf{R}^S : \bar{x}(S) = b(S), \bar{x}(Z) \leq b(Z) \text{ for every } Z \subset S\}, \quad (2.2)$$

which is possibly unbounded but never empty. The empty set, however, is also considered a base-polyhedron by convention. When b is integer-valued, $B(b)$ is an integral polyhedron, which is referred to as **integral base-polyhedron**. A non-empty base-polyhedron B can also be defined by a supermodular function p with $p(S)$ finite as follows:

$$B = B'(p) := \{x \in \mathbf{R}^S : \bar{x}(S) = p(S), \bar{x}(Z) \geq p(Z) \text{ for every } Z \subset S\}. \quad (2.3)$$

We have $B'(p) = B(b)$ if p is a complementary function of b , that is, if $p(X) = b(S) - b(S - X)$ for all $X \subseteq S$.

The set $\overset{\dots}{B}$ of integral elements of an integral base-polyhedron B is called an **M-convex set** in discrete convex analysis [30, 31, 32]. To be more precise, an M-convex set is defined as a set of integral points satisfying a certain exchange axiom, and it is known that these two properties are equivalent ([31, Theorem 4.15]). It should be mentioned, however, that the set of integer elements of an integral base-polyhedron itself has long been recognized as a combinatorially nice object (although no particular name was coined). Indeed, this recognition is

already evident in Edmonds' classic paper [6], and Fujishige's book [14] presents the theory of submodular systems over an arbitrary totally additive group, of which the set \mathbf{Z} of integers is a special case. It is noted that the set of integral elements of a non-integral base-polyhedron is not necessarily an M-convex set.

In Case **R** of the dec-min problem, we seek a dec-min element of a base-polyhedron B described by a real-valued supermodular function p or a submodular function b . In Case **Z**, the dec-min problem is defined on an M-convex set $\overset{\dots}{B}$, which is the set of integral members of an integral base-polyhedron B described by an integer-valued p or b .

2.2 Decreasing minimality and increasing maximality

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

In both Case **R** and Case **Z**, decreasing minimality is equivalent to increasing maximality.

Theorem 2.1 ([13, 14]). *The unique decreasingly minimal element of B is the unique increasingly maximal element of B .* ■

Theorem 2.2 ([9, Theorem 3.3]). *An element of an M-convex set $\overset{\dots}{B}$ is decreasingly minimal in $\overset{\dots}{B}$ if and only if it is increasingly maximal in $\overset{\dots}{B}$.* ■

2.3 Characterizations

Let m be an element of an M-convex set $\overset{\dots}{B}$. A **1-tightening step** replaces m by $m' := m + \chi_s - \chi_t$, where s and t are elements of S for which $m(t) \geq m(s) + 2$ and m' belongs to $\overset{\dots}{B}$. A subset $X \subseteq S$ is called **m -tight** (with respect to p) if $\widetilde{m}(X) = p(X)$. A subset $X \subseteq S$ is called an **m -top** set if $m(s) \geq m(t)$ holds whenever $s \in X$ and $t \in S - X$. (For example, for $m = (3, 2, 2)$ indexed by $S = \{s_1, s_2, s_3\}$, there are five m -top sets: the empty set, $\{s_1\}$, $\{s_1, s_2\}$, $\{s_1, s_3\}$, and S .) We call an integral vector $x \in \mathbf{Z}^S$ **near-uniform** on a subset S' of S if its largest and smallest components on S' differ by at most 1, that is, if there exists some integer ℓ for which $x(s) \in \{\ell, \ell + 1\}$ for every $s \in S'$.

The following theorem from [9] gives fundamental characterizations of a dec-min element of an M-convex set.

Theorem 2.3 ([9, Theorem 3.3]). *For an element m of an M-convex set $\overset{\dots}{B} = \overset{\dots}{B}(p)$, the following three conditions are pairwise equivalent.*

- (A) m is decreasingly minimal in $\overset{\dots}{B}$.
- (B) There is no 1-tightening step for m . That is,

$$m(t) \geq m(s) + 2 \implies m + \chi_s - \chi_t \notin \overset{\dots}{B}. \quad (2.4)$$

- (C) There is a chain $(\emptyset \subset) C_1 \subset C_2 \subset \dots \subset C_\ell (= S)$ such that each C_i is an m -top and m -tight set (with respect to p) and m is near-uniform on each $S_i := C_i - C_{i-1}$ ($i = 1, 2, \dots, \ell$), where $C_0 := \emptyset$. ■

The corresponding theorem for Case **R** can be formulated from known results [13, 14] as follows, where a direct proof can also be obtained from an easy adaptation of the proof of [9, Theorem 3.3]. For an element m of a base-polyhedron $B = B'(p)$ defined by a real-valued

supermodular function p , a subset $X \subseteq S$ is called m -tight (with respect to p) if $\widetilde{m}(X) = p(X)$, and an m -top set if $m(s) \geq m(t)$ holds whenever $s \in X$ and $t \in S - X$. We call a vector $x \in \mathbf{R}^S$ **uniform** on a subset S' of S if $x(s) = x(t)$ for all $s, t \in S'$.

Theorem 2.4. *For an element m of a base-polyhedron $B = B'(p)$, the following three conditions are pairwise equivalent.*

- (A) m is decreasingly minimal in B .
- (B) m satisfies the following condition:

$$m(t) > m(s), \alpha > 0 \implies m + \alpha(\chi_s - \chi_t) \notin B. \quad (2.5)$$

- (C) *There is a chain $(\emptyset \subset) C_1 \subset C_2 \subset \dots \subset C_\ell (= S)$ such that each C_i is an m -top and m -tight set (with respect to p) and m is uniform on each $S_i := C_i - C_{i-1}$ ($i = 1, 2, \dots, \ell$), where $C_0 := \emptyset$.* ■

2.4 Uniqueness

The structures of dec-min elements have a striking difference in Case **R** and Case **Z**. In Case **R** the dec-min element $m_{\mathbf{R}}$ of B is uniquely determined. (The dec-min element, if any, is uniquely determined in an arbitrary polyhedron.) In Case **Z** the dec-min elements $m_{\mathbf{Z}}$ of $\overline{\overline{B}}$ are endowed with the structure of basis family of a matroid. This is stated in Theorem 2.5 below, where a **matroidal M-convex set** means an M-convex set in which the ℓ_∞ -distance of any two distinct members is equal to one.

Theorem 2.5 ([9, Theorem 5.7]). *The set of dec-min elements of an M-convex set $\overline{\overline{B}}$ is a matroidal M-convex set. In other words, there exist a matroid M^* and an integral vector Δ^* such that an element m of $\overline{\overline{B}}$ is decreasingly minimal if and only if m can be obtained as $m = \chi_L + \Delta^*$ with a basis L of M^* .* ■

The minimum cost dec-min element problem is to compute a dec-min element that has the smallest cost with respect to a given cost-function $c : S \rightarrow \mathbf{R}$ on the ground-set S . In Case **Z**, this problem is meaningful and interesting, and was solved in [9, Section 5.3] on the basis of Theorem 2.5; see [10] for its instances in graph orientation problems. In Case **R**, in contrast, this problem does not make any sense because of the uniqueness of the dec-min element of B .

2.5 Square-sum minimization

In both Case **R** and Case **Z**, a dec-min element is characterized as a minimizer of square-sum of the components:

$$W(x) := \sum [x(s)^2 : s \in S]. \quad (2.6)$$

Theorem 2.6 ([13, Theorem 3.3]). *An element m of a base-polyhedron B is a square-sum minimizer on B if and only if m is a dec-min element of B .* ■

Theorem 2.7 ([9, Corollary 6.4]). *An element m of an M-convex set $\overline{\overline{B}}$ is a square-sum minimizer on $\overline{\overline{B}}$ if and only if m is a dec-min element of $\overline{\overline{B}}$.* ■

In Case **R**, the variable x for minimization is a real vector, $x \in \mathbf{R}^S$, whereas in Case **Z** it is an integer vector, $x \in \mathbf{Z}^S$. In Case **R**, the minimizer of $W(x)$ over B is unique, and is often referred to as the **minimum norm point** of B . (Actually, the minimizer of $W(x)$ is unique for an arbitrary polyhedron, and is a rational vector for a rational polyhedron. However, the minimum norm point may not be the dec-min element.)

In both Case **R** and Case **Z**, there are min-max formulas for the square-sum. The min-max formulas refer to the well-known concept of linear extension (or Lovász extension) \hat{p} of p , which is defined for $\pi \in \mathbf{R}^S$ by

$$\hat{p}(\pi) := p(S_n)\pi(s_n) + \sum_{j=1}^{n-1} p(S_j)[\pi(s_j) - \pi(s_{j+1})], \quad (2.7)$$

where $n = |S|$, the elements of S are indexed in such a way that $\pi(s_1) \geq \pi(s_2) \geq \dots \geq \pi(s_n)$, and $S_j := \{s_1, s_2, \dots, s_j\}$ for $j = 1, 2, \dots, n$. Here $p(S_j)[\pi(s_j) - \pi(s_{j+1})]$ is defined to be 0 when $\pi(s_j) - \pi(s_{j+1}) = 0$ even if $p(S_j) = -\infty$. In Case **Z** we have the min-max identity [9, Theorem 6.10]:

$$\begin{aligned} & \min\left\{\sum [m(s)^2 : s \in S] : m \in \overset{\dots}{B}\right\} \\ &= \max\left\{\hat{p}(\pi) - \sum_{s \in S} \left\lfloor \frac{\pi(s)}{2} \right\rfloor \left\lceil \frac{\pi(s)}{2} \right\rceil : \pi \in \mathbf{Z}^S\right\}. \end{aligned} \quad (2.8)$$

In Case **R**, the corresponding formula is

$$\begin{aligned} & \min\left\{\sum [m(s)^2 : s \in S] : m \in B\right\} \\ &= \max\left\{\hat{p}(\pi) - \sum_{s \in S} \left(\frac{\pi(s)}{2}\right)^2 : \pi \in \mathbf{R}^S\right\}, \end{aligned} \quad (2.9)$$

which may be regarded as an adaptation of the standard quadratic programming duality to the case where the feasible region is a base-polyhedron.

2.6 Principal and canonical partitions

In either of the continuous and discrete problems, a certain partition of the ground-set S is known to play an essential role as the dual object to characterize dec-min elements. In Case **R**, the partition is called the principal partition, which is used to characterize the (unique) dec-min element $m_{\mathbf{R}}$ of B . In Case **Z**, another partition of S , called the canonical partition, characterizes the set of all dec-min elements $m_{\mathbf{Z}}$ of $\overset{\dots}{B}$. So far these two notions are introduced independently and nothing is known about their mutual relation. In Section 3, we shall reveal the precise relation between these partitions as well as the associated numbers called critical values and essential values.

2.7 Proximity

In general terms, proximity results refer to statements that the solutions to discrete and continuous versions of an optimization problem are geometrically close to each other. In Section 4 we shall obtain proximity results, showing that dec-min elements $m_{\mathbf{Z}}$ of $\overset{\dots}{B}$ are located near

the dec-min element $m_{\mathbf{R}}$ of B . The obtained proximity results, in turn, give rise to continuous relaxation algorithms for computing a dec-min element $m_{\mathbf{Z}}$ of an M-convex set in strongly polynomial time when the (fractional) dec-min element $m_{\mathbf{R}}$ of a base-polyhedron is given.

2.8 Algorithm

In Case **R**, the decomposition algorithm of Fujishige [13] in 1980 was already a strongly polynomial algorithm to find the (unique) dec-min element. This decomposition algorithm for computing the dec-min element, or equivalently for minimizing a quadratic function on a base-polyhedron, was extended by Groenevelt [18] to separable convex functions on a polymatroid.

Fujishige gave, in his book [14], two algorithms for finding the dec-min element for Case **R**. The first algorithm called “decomposition algorithm” [14, Section 8.2] is not completely the same as, but a variant of, the (original) decomposition algorithm of [13]. This variant is also strongly polynomial. The other algorithm called “monotone algorithm” [14, Section 9.2] is a procedure that computes iteratively the members of the principal partition belonging to B , as well as the critical values (which define the dec-min element $m_{\mathbf{R}}$ immediately). The algorithm is quite simple and natural but it relies on a subroutine for computing a certain number λ^* which is, in the present context, equivalent to computing $\max\{p(X)/|X| : \emptyset \neq X \subseteq S\}$. Though not mentioned explicitly in [14], this λ^* can be computed in strongly polynomial time with the help of the Newton–Dinkelbach algorithm, which follows, for example, from a result of Iwata et al. [25] for a more general problem. See Radzik [34] and Goemans et al. [17] for the Newton–Dinkelbach algorithm.

For Case **Z**, a strongly polynomial algorithm was devised in [10]. The algorithm in [10] relies on a discrete version of the Newton–Dinkelbach algorithm, and may be viewed as a discrete counterpart of Fujishige’s monotone algorithm. In particular, the algorithm of [10] computes iteratively the canonical chain and partition belonging to $\overset{\dots}{B}$, along with the essential value-sequence and a dec-min element itself. It is emphasized that the canonical chain and the essential value-sequence provide us with a structural description (Theorem 2.5) of the set of all dec-min elements of $\overset{\dots}{B}$, which, in turn, enables us to compute a minimum cost dec-min element of $\overset{\dots}{B}$ with respect to a linear cost-function. It is noted here that, if a single dec-min element of $\overset{\dots}{B}$ is available, the canonical chain and the essential value-sequence can be computed quite easily [10, Algorithm 2.3].

By the equivalence of dec-minimization and square-sum minimization (Theorem 2.7), any algorithm for square-sum minimization solves the dec-min problem. Groenevelt’s decomposition algorithm [18] for separable convex minimization on the set of integral points of an integral polymatroid can be adapted to minimizing the square-sum on an M-convex set in strongly polynomial time, which we describe in Section 5.2. It is natural to expect that Fujishige’s decomposition algorithm [14, Section 8.2] for Case **R** can be adapted to Case **Z** through a piecewise-linear extension of a given separable convex function in integer variables, as is suggested in [14, Section 8.3]. This approach leads indeed to another decomposition algorithm for minimizing the square-sum on an M-convex set, which we describe in Section 5.3. To realize strong polynomiality we need to devise a non-trivial gadget to cope with complications arising from integrality.

Table 1 summarizes the development of the decomposition algorithms for Case **R** and Case **Z**. In Section 5.1 we remark on the differences of these decomposition algorithms.

Table 1: Decomposition algorithms for minimization on a base-polyhedron

	Case R	Case Z
Quadratic (square-sum)	Fujishige [13]	(Sec. 5 of this paper)
Separable convex	↓	
	Groenevelt [18]	→ Groenevelt [18]
	↓	
	Fujishige [14, Sec. 8.2]	→ Fujishige [14, Sec. 8.3]

2.9 Weighting

In Case **R** a weight vector was introduced to define and analyze lexico-optimality in [13, 14, 38]. In Case **Z** the unweighted problem has been investigated in [9, 10] and the weighted case will be treated in a forthcoming paper. Although we restrict ourselves, in this paper, to unweighted problems in Case **R** and Case **Z**, some facts about the weighted dec-min problem are mentioned below. (The ‘weighted’ dec-min problem here should not be confused with the ‘minimum cost’ dec-min problem discussed after Theorem 2.5.)

Let w be a positive vector on S , which is assumed to be an integral vector in Case **Z**. The decreasing minimization problem with weighting w is to find an element m of B (resp., $\overset{\dots}{B}$ in Case **Z**) for which the vector $(w(s)m(s) : s \in S)$ is decreasingly minimal in B (resp., $\overset{\dots}{B}$ in Case **Z**). Such an element is called a w -dec-min element. We define w -inc-max in an obvious manner. The (original) decreasing minimization without weighting corresponds to the case of $w = (1, 1, \dots, 1)$. (It is noted that the weight vector w in [13, 14] corresponds to the (componentwise) reciprocal of the vector w here.) The w -dec-min problem should not be confused with the minimum cost dec-min problem with respect to a given linear cost-function on S . Even for the weighted dec-min problem, we can formulate its minimum cost version, which is to find the minimum cost w -dec-min element with respect to a given cost-function $c : S \rightarrow \mathbf{R}$ on S .

According to the results of [13, 14, 38], we may say, roughly, that there is not much difference between the weighted and unweighted problems in Case **R**. In Case **Z**, however, weighting causes substantial complications. For example, w -dec-min and w -inc-max coincide in Case **R**, but not in Case **Z**. In Case **R**, there is a unique w -dec-min element in B and is characterized as the unique minimizer of the weighted square-sum

$$W_w(x) := \sum [w(s)x(s)^2 : s \in S]. \quad (2.10)$$

In contrast, square-sum minimization does not characterize w -dec-minimality in Case **Z**. As a concrete example, consider the line segment B on the plane \mathbf{R}^2 connecting $(2, 0)$ and $(0, 2)$. This B is an integral base-polyhedron, and the corresponding M-convex set is $\overset{\dots}{B} = \{(2, 0), (1, 1), (0, 2)\}$. For the weight vector $w = (1, 2)$, $m_1 := (2, 0)$ is the (unique) w -dec-min element of $\overset{\dots}{B}$ and $m_2 := (1, 1)$ is the (unique) w -inc-max element of $\overset{\dots}{B}$. Since $W_w(m_1) = 4 > 3 = W_w(m_2)$, w -dec-minimality is not characterized by square-sum minimality.

3 Principal partition and canonical partition

A review of the principal partition is offered in Section 3.1 with emphasis on its role in decreasing minimization, while Section 3.2 is a review of the canonical partition for discrete

decreasing minimization. Section 3.3 gives a new characterization of the canonical partition, which is used in Section 3.4 to clarify the relationship between the principal and canonical partitions. It is mentioned that this section is based on our (unpublished) technical report [8].

3.1 Review of the principal partition

As is pointed out by Fujishige [13], the dec-min element in the continuous case is closely related to the principal partition. The principal partition is the central concept in a structural theory for submodular functions; see Iri [24] for an early survey and Fujishige [15] for 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 [13], the results are stated for a real-valued submodular function, and the description below is a translation for a real-valued supermodular function $p : 2^S \rightarrow \mathbf{R} \cup \{-\infty\}$ with $p(\emptyset) = 0$.

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

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

Proposition 3.1. (1) If $\lambda > \lambda'$, $X \in \mathcal{L}(\lambda)$, and $Y \in \mathcal{L}(\lambda')$, then $X \subseteq Y$.
(2) If $\lambda \geq \lambda'$, then $L(\lambda) \subseteq L(\lambda')$.

Proof. (1) Let $X \in \mathcal{L}(\lambda)$ and $Y \in \mathcal{L}(\lambda')$ for $\lambda > \lambda'$. We have

$$p(X) + p(Y) \leq p(X \cap Y) + p(X \cup Y), \quad (3.1)$$

$$\begin{aligned} \lambda|X| + \lambda'|Y| &= \lambda|X \cap Y| + \lambda'|X \cup Y| + (\lambda - \lambda')|X - Y| \\ &\geq \lambda|X \cap Y| + \lambda'|X \cup Y|. \end{aligned} \quad (3.2)$$

From these inequalities it follows that

$$\begin{aligned} (p(X) - \lambda|X|) + (p(Y) - \lambda'|Y|) \\ \leq (p(X \cap Y) - \lambda|X \cap Y|) + (p(X \cup Y) - \lambda'|X \cup Y|). \end{aligned}$$

Here we also have the reverse inequality \geq , since X is a maximizer for λ and Y is a maximizer for λ' . Therefore, we have equality in (3.2), from which follows $(\lambda - \lambda')|X - Y| = 0$. Since $\lambda - \lambda' > 0$, this implies $X \subseteq Y$.

(2) This follows immediately from (1). ■

There are finitely many numbers λ for which $|\mathcal{L}(\lambda)| \geq 2$. We denote such numbers as $\lambda_1 > \lambda_2 > \dots > \lambda_r$, which are called the **critical values**. Note that the condition $|\mathcal{L}(\lambda)| \geq 2$ for a critical value is equivalent to saying that the largest element of $\mathcal{L}(\lambda)$ is distinct from the smallest element of $\mathcal{L}(\lambda)$. Since the largest element of $\mathcal{L}(\lambda)$ is equal to $L(\lambda - \varepsilon)$ for a sufficiently small $\varepsilon > 0$, we can say that λ is a critical value if and only if $L(\lambda) \neq L(\lambda - \varepsilon)$ for any $\varepsilon > 0$. Thus we obtain:

$$\begin{aligned} \emptyset = L(\lambda_1) \subset L(\lambda_1 - \varepsilon) = \dots = L(\lambda_2) \subset L(\lambda_2 - \varepsilon) = \dots \\ \dots = L(\lambda_r) \subset L(\lambda_r - \varepsilon) = S \end{aligned} \quad (3.3)$$

for sufficiently small $\varepsilon > 0$.

The **principal partition** $\{\hat{S}_1, \hat{S}_2, \dots, \hat{S}_r\}$ of the ground-set S is defined by

$$\hat{S}_i := L(\lambda_{i+1}) - L(\lambda_i) = L(\lambda_i - \varepsilon) - L(\lambda_i) \quad (i = 1, 2, \dots, r), \quad (3.4)$$

where $L(\lambda_{r+1}) := S$ by convention and ε is a sufficiently small positive number. By defining

$$\hat{C}_i := \hat{S}_1 \cup \hat{S}_2 \cup \dots \cup \hat{S}_i \quad (= L(\lambda_i - \varepsilon)) \quad (3.5)$$

for $i = 1, 2, \dots, r$, we obtain a chain:

$$\hat{C}_1 \subset \hat{C}_2 \subset \dots \subset \hat{C}_r, \quad (3.6)$$

where $\hat{C}_1 \neq \emptyset$ and $\hat{C}_r = S$. In this paper we call this chain the **principal chain**. We have $\hat{S}_i = \hat{C}_i - \hat{C}_{i-1}$ for $i = 1, 2, \dots, r$, where $\hat{C}_0 := \emptyset$.

Remark 3.1. The principal partition $\{\hat{S}_1, \hat{S}_2, \dots, \hat{S}_r\}$ defined here by (3.4) is, in fact, an aggregation of what is usually meant under the name of ‘‘principal partition’’ [15, 24]. The latter is defined as follows. Not only each $\mathcal{L}(\lambda)$ is a lattice, but their union $\mathcal{L}_{\text{all}} := \bigcup_{\lambda \in \mathbf{R}} \mathcal{L}(\lambda)$ is also a lattice by Proposition 3.1(1), and we have $\mathcal{L}_{\text{all}} = \bigcup_{i=1}^r \mathcal{L}(\lambda_i)$. A maximal chain of this lattice \mathcal{L}_{all} induces a partition of the ground-set S , and the induced partition is determined independently of the choice of a maximal chain. In [15, 24], this is called the principal partition of S . Furthermore, a partial order can be defined among the members of the partition. The chain

$$(\emptyset =) \hat{C}_0 \subset \hat{C}_1 \subset \hat{C}_2 \subset \dots \subset \hat{C}_r (= S) \quad (3.7)$$

associated with our partition $\{\hat{S}_1, \hat{S}_2, \dots, \hat{S}_r\}$ is certainly a chain of \mathcal{L}_{all} , which, however, may not be maximal. If the chain in (3.7) is not maximal, our partition is an aggregation of the principal partition in the sense of [15, 24]. ■

The following theorem shows the close relationship between the principal partition and the unique dec-min element (minimum norm point) of a base-polyhedron $B = B'(p)$.

Theorem 3.2 (Fujishige [13]). *Let $B = B'(p)$ be a base-polyhedron defined by a supermodular function p .*

(1) *Let $\{\hat{S}_1, \hat{S}_2, \dots, \hat{S}_r\}$ be the principal partition and $\lambda_1 > \lambda_2 > \dots > \lambda_r$ the critical values. The unique dec-min element $m_{\mathbf{R}}$ of B is given by $m_{\mathbf{R}}(s) = \lambda_i$ for $s \in \hat{S}_i$ and $i = 1, 2, \dots, r$. In particular, $m_{\mathbf{R}}$ is uniform on each member \hat{S}_i of the principal partition.*

(2) *Let $m_{\mathbf{R}}$ be the unique dec-min element of B . The critical values $\lambda_1 > \lambda_2 > \dots > \lambda_r$ are precisely those numbers that appear as component values of $m_{\mathbf{R}}$, and the principal partition $\{\hat{S}_1, \hat{S}_2, \dots, \hat{S}_r\}$ is given by $\hat{S}_i = \{s \in S : m_{\mathbf{R}}(s) = \lambda_i\}$ for $i = 1, 2, \dots, r$. ■*

The following characterization of dec-minimality can be formulated by combining Theorem 3.2 and Theorem 2.4.

Theorem 3.3. *Let $B = B'(p)$ be a base-polyhedron defined by a supermodular function p . An element $m \in B$ with distinct component values $\lambda'_1 > \lambda'_2 > \dots > \lambda'_\ell$ is the unique dec-min element of B if and only if each ‘‘level set’’ $\hat{C}_i := \{s \in S : m(s) \geq \lambda'_i\}$ is m -tight ($\bar{m}(\hat{C}_i) = p(\hat{C}_i)$) for $i = 1, 2, \dots, \ell$.*

Proof. The only-if part is immediate from Theorem 3.2(2). For the if-part, note that \hat{C}_i is an m -top set by definition, which is also m -tight by assumption. Then we can use (C) \Rightarrow (A) in Theorem 2.4. ■

3.2 Review of the canonical partition

In the discrete case, the canonical partition describes the structure of dec-min elements. The canonical partition is defined iteratively using contractions as follows [9].

Let $p : 2^S \rightarrow \mathbf{Z} \cup \{-\infty\}$ be an integer-valued supermodular function with $p(\emptyset) = 0$ and $p(S) > -\infty$, and define $C_0 := \emptyset$. For $j = 1, 2, \dots, q$, define

$$\beta_j := \max \left\{ \left\lfloor \frac{p(X \cup C_{j-1}) - p(C_{j-1})}{|X|} \right\rfloor : \emptyset \neq X \subseteq \overline{C_{j-1}} \right\}, \quad (3.8)$$

$$h_j(X) := p(X \cup C_{j-1}) - (\beta_j - 1)|X| - p(C_{j-1}) \quad (X \subseteq \overline{C_{j-1}}), \quad (3.9)$$

$$S_j := \text{smallest subset of } \overline{C_{j-1}} \text{ maximizing } h_j, \quad (3.10)$$

$$C_j := C_{j-1} \cup S_j, \quad (3.11)$$

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 resulting chain $C_1 \subset C_2 \subset \dots \subset C_q$ is called the **canonical chain**, and the partition $\{S_1, S_2, \dots, S_q\}$ is the **canonical partition**. The integers β_j , known [9, Theorem 5.5] to be decreasing, are called the **essential values**, and this decreasing sequence $\beta_1 > \beta_2 > \dots > \beta_q$ is named the **essential value-sequence**. We have $S_j = C_j - C_{j-1}$ for $j = 1, 2, \dots, q$.

The following theorem characterizes dec-min elements of $\overline{\overline{B}}$ using these concepts.

Theorem 3.4 ([9, Corollary 5.2]). *Let $B = B'(p)$ be an integral base-polyhedron on ground-set S . Let $\{C_1, C_2, \dots, C_q\}$ be the canonical chain, $\{S_1, S_2, \dots, S_q\}$ the canonical partition of S , and $\beta_1 > \beta_2 > \dots > \beta_q$ the essential value-sequence belonging to $\overline{\overline{B}}$. Then an element $m \in \overline{\overline{B}}$ is decreasingly minimal in $\overline{\overline{B}}$ if and only if each C_j is m -tight (that is, $\overline{m}(C_j) = p(C_j)$) and $\beta_j - 1 \leq m(s) \leq \beta_j$ holds for each $s \in S_j$ ($j = 1, 2, \dots, q$). ■*

This theorem implies, in particular, that every dec-min element $m_{\mathbf{Z}}$ of $\overline{\overline{B}}$ is near-uniform on each member of the canonical partition. That is, $|m_{\mathbf{Z}}(s) - m_{\mathbf{Z}}(t)| \leq 1$ if $\{s, t\} \subseteq S_j$ for some member S_j of the canonical partition.

The following theorem shows that any dec-min element $m_{\mathbf{Z}}$ of $\overline{\overline{B}}$, in turn, determines the essential value-sequence, the canonical chain, and canonical partition.

Theorem 3.5 ([9, Corollary 5.4]). *Let $m_{\mathbf{Z}}$ be an arbitrary dec-min element of $\overline{\overline{B}}$. The first essential value β_1 is the largest $m_{\mathbf{Z}}$ -value and the first member C_1 of the canonical chain is the smallest $m_{\mathbf{Z}}$ -tight set containing all β_1 -valued elements. Moreover, for $j = 2, \dots, q$, the j -th essential value β_j is the largest value of $m_{\mathbf{Z}}(s)$ for $s \in S - C_{j-1}$ and the j -th member C_j of the canonical chain is the smallest $m_{\mathbf{Z}}$ -tight set (with respect to p) containing each element of $m_{\mathbf{Z}}$ -value at least β_j . ■*

These results (Theorems 3.4 and 3.5) may be viewed as a discrete counterpart of Theorem 3.2 for the continuous case.

3.3 An alternative characterization of the canonical partition

Originally [9], the canonical partition was defined iteratively using contractions, as described by (3.8)–(3.11) in Section 3.2. 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, in Section 3.4, the precise relation between the canonical partition and the principal partition.

By the definition given in (3.8)–(3.11), we have that

$$C_j \text{ is the smallest maximizer of } p(X) - (\beta_j - 1)|X| \text{ among all } X \supseteq C_{j-1}. \quad (3.12)$$

We will show, in Proposition 3.6 below, that C_j is, in fact, the smallest maximizer of $p(X) - (\beta_j - 1)|X|$ among all subsets X of S without the constraint of being a superset of C_{j-1} .

For any integer β , let $\mathcal{L}(\beta)$ denote the family of all maximizers of $p(X) - \beta|X|$, and $L(\beta)$ be the smallest element of $\mathcal{L}(\beta)$, where the smallest element exists in $\mathcal{L}(\beta)$ since $\mathcal{L}(\beta)$ is a lattice (ring family). (These notations are consistent with the ones introduced in Section 3.1.)

Proposition 3.6. *Let β_j and C_j ($j = 1, 2, \dots, q$) be defined by (3.8)–(3.11).*

(1) $\beta_1 > \beta_2 > \dots > \beta_q$.

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

Proof. (1) The monotonicity of the β -values is already shown in [9, Theorem 5.5], but we give an alternative proof here. Let $j \geq 2$. By (3.8), we have $\beta_{j-1} > \beta_j$ if and only if

$$\beta_{j-1} > \left\lceil \frac{p(X \cup C_{j-1}) - p(C_{j-1})}{|X|} \right\rceil \quad (3.13)$$

for every X with $\emptyset \neq X \subseteq \overline{C_{j-1}}$. Furthermore, we can rewrite the condition (3.13) as follows:

$$\begin{aligned} (3.13) &\Leftrightarrow \beta_{j-1} - 1 \geq \frac{p(X \cup C_{j-1}) - p(C_{j-1})}{|X|} \\ &\Leftrightarrow p(X \cup C_{j-1}) - p(C_{j-1}) \leq (\beta_{j-1} - 1)|X| \\ &\Leftrightarrow p(X \cup C_{j-1}) - (\beta_{j-1} - 1)|X \cup C_{j-1}| \leq p(C_{j-1}) - (\beta_{j-1} - 1)|C_{j-1}|. \end{aligned}$$

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

(2) We prove $C_j = L(\beta_j - 1)$ for $j = 1, 2, \dots, q$ by induction on j . This holds for $j = 1$ by definition. Let $j \geq 2$. By Proposition 3.1(2) for $\lambda = \beta_{j-1} - 1$ and $\lambda' = \beta_j - 1$, we have $L(\beta_j - 1) \supseteq L(\beta_{j-1} - 1)$, whereas $L(\beta_{j-1} - 1) = C_{j-1}$ by the induction hypothesis. Combining this with (3.12), we obtain $C_j = L(\beta_j - 1)$. ■

We now give an alternative characterization of the essential value-sequence $\beta_1 > \beta_2 > \dots > \beta_q$ defined by (3.8)–(3.11). We consider the family $\{L(\beta) : \beta \in \mathbf{Z}\}$ of the smallest maximizers of $p(X) - \beta|X|$ for all integers β . Each C_j is a member of this family, since $C_j = L(\beta_j - 1)$ ($j = 1, 2, \dots, q$) by Proposition 3.6(2).

Proposition 3.7. *As β is decreased from $+\infty$ to $-\infty$ (or from β_1 to $\beta_q - 1$), the smallest maximizer $L(\beta)$ is monotone non-decreasing. We have $L(\beta) \neq L(\beta - 1)$ if and only if β is equal to an essential value. Therefore, the essential value-sequence $\beta_1 > \beta_2 > \dots > \beta_q$ is characterized by the property:*

$$\begin{aligned} \emptyset = L(\beta_1) &\subset L(\beta_1 - 1) = \dots = L(\beta_2) \subset L(\beta_2 - 1) = \dots \\ &\dots = L(\beta_q) \subset L(\beta_q - 1) = S. \end{aligned} \quad (3.14)$$

Proof. The monotonicity of $L(\beta)$ follows from Proposition 3.1(2). We will show (i) $L(\beta_1) = \emptyset$, (ii) $L(\beta_{j-1} - 1) = L(\beta_j)$ for $j = 2, \dots, q$, and (iii) $L(\beta_j) \subset L(\beta_j - 1)$ for $j = 1, 2, \dots, q$.

(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 put $C := C_{j-1}$ and define

$$h(Y) := p(Y) - \beta_j|Y|$$

for any subset Y of S . Let A be the smallest maximizer of h , which means $A = L(\beta_j)$. For any non-empty subset X of \overline{C} ($= S - C$) we have

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

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

$$h(Y) \leq h(C) \quad \text{for all } Y \supseteq C. \quad (3.15)$$

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 (3.15). 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., $A \subseteq C$. On recalling notations $A = L(\beta_j)$ and $C = C_{j-1} = L(\beta_{j-1} - 1)$, we obtain $L(\beta_j) \subseteq L(\beta_{j-1} - 1)$. We also have $L(\beta_j) \supseteq L(\beta_{j-1} - 1)$ by the monotonicity. Therefore, $L(\beta_j) = L(\beta_{j-1} - 1)$.

(iii) Let $1 \leq j \leq q$ and put $C := C_{j-1}$. Take a non-empty subset Z of \overline{C} which gives the maximum in the definition of β_j , i.e.,

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

Then we have

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

which implies

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

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

Proposition 3.7 justifies the following alternative definitions of the essential value-sequence, the canonical chain, and the canonical partition.

Proposition 3.8. *The essential value-sequence, the canonical chain, and the canonical partition can also be defined as follows:*

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

$$S_j := C_j - C_{j-1} = L(\beta_j - 1) - L(\beta_j) \quad (j = 1, 2, \dots, q), \quad (3.16)$$

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 values are the discrete counterpart of the critical values. This is discussed in the next section.

3.4 Canonical partition from the principal partition

The characterization of the canonical partition shown in Section 3.3 enables us to construct the canonical partition and essential values for Case **Z** from the principal partition and critical values for Case **R** as follows.

Theorem 3.9. *Let B be an integral base-polyhedron defined by an integer-valued supermodular function p .*

(1) *The essential values $\beta_1 > \beta_2 > \dots > \beta_q$ are obtained from the critical values $\lambda_1 > \lambda_2 > \dots > \lambda_r$ as the distinct members of the rounded-up integers $\lceil \lambda_1 \rceil \geq \lceil \lambda_2 \rceil \geq \dots \geq \lceil \lambda_r \rceil$. In particular, an integer β is an essential value if and only if there exists a critical value λ satisfying $\beta \geq \lambda > \beta - 1$.*

(2) *The canonical partition $\{S_1, S_2, \dots, S_q\}$ is an aggregation of the principal partition $\{\hat{S}_1, \hat{S}_2, \dots, \hat{S}_r\}$ given by*

$$S_j = \bigcup_{i \in I(j)} \hat{S}_i \quad (j = 1, 2, \dots, q), \quad (3.17)$$

where $I(j) := \{i : \lceil \lambda_i \rceil = \beta_j\}$ for $j = 1, 2, \dots, q$.

(3) *The canonical chain $\{C_j\}$ is a subchain of the principal chain $\{\hat{C}_i\}$; we have $C_j = \hat{C}_i$ for $i = \max I(j)$.*

Proof. (1) By Proposition 3.8, an integer β is an essential value if and only if $L(\beta) \neq L(\beta - 1)$, whereas a real number λ is a critical value if and only if $L(\lambda) \neq L(\lambda - \varepsilon)$ for any $\varepsilon > 0$. Hence follows the claim.

(2) For each $j = 1, 2, \dots, q$, let i_{\max} and i_{\min} denote the maximum and minimum elements of $I(j)$, respectively. By (3.4) and (3.16) we have

$$\begin{aligned} \bigcup_{i \in I(j)} \hat{S}_i &= \bigcup_{i \in I(j)} (L(\lambda_{i+1}) - L(\lambda_i)) = L(\lambda_{i_{\max}+1}) - L(\lambda_{i_{\min}}) \\ &= L(\lambda_{i_{\max}} - \varepsilon) - L(\lambda_{i_{\min}}) = L(\beta_j - 1) - L(\beta_j) = S_j. \end{aligned}$$

(3) By (3.17) we have

$$C_j = \bigcup_{k \leq j} S_k = \bigcup_{k \leq j} \bigcup_{i \in I(k)} \hat{S}_i = \bigcup_{i \leq i_{\max}} \hat{S}_i = \hat{C}_{i_{\max}}.$$

This completes the proof of Theorem 3.9. ■

The following two examples illustrate Theorem 3.9. The first example treats the simplest case to demonstrate the basic idea as well as the notation. The second is a running example, to be considered repeatedly to illustrate our subsequent arguments.

Example 3.1. Let $S = \{s_1, s_2\}$ and $\overset{\dots}{B} = \{(0, 3), (1, 2), (2, 1)\}$, where B is the line segment connecting $(0, 3)$ and $(2, 1)$. There are two dec-min elements of $\overset{\dots}{B}$, $m_{\mathbf{Z}}^{(1)} = (1, 2)$ 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 for $B = B'(p)$ is given by

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

and we have

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

There is only one ($r = 1$) critical value $\lambda_1 = 3/2$ and the associated sublattice is $\mathcal{L}(\lambda_1) = \{\emptyset, S\}$. The principal partition is a trivial partition $\{\hat{S}_1\}$ consisting of only one member $\hat{S}_1 = S$. 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 $S_1 = C_1 = L(\beta_1 - 1) = L(1) = S$. ■

Example 3.2. Let $S = \{s_1, s_2, s_3, s_4\}$ and consider five vectors

$$\begin{aligned} m_1 &= (2, 1, 1, 0), \quad m_2 = (2, 1, 0, 1), \quad m_3 = (1, 2, 1, 0), \\ m_4 &= (1, 2, 0, 1), \quad m_5 = (2, 2, 0, 0). \end{aligned}$$

These vectors are obtained by adding vector $(1, 1, 0, 0)$ to $(1, 0, 1, 0)$, $(1, 0, 0, 1)$, $(0, 1, 1, 0)$, $(0, 1, 0, 1)$, $(1, 1, 0, 0)$, which are the characteristic vectors of bases of a rank-2 matroid on S . Let B denote the convex hull of $\{m_1, m_2, \dots, m_5\}$. Then B is an integral base-polyhedron and $\overset{\dots}{B} = \{m_1, m_2, \dots, m_5\}$ is an M-convex set. The dec-min elements of $\overset{\dots}{B}$ are m_1, m_2, m_3 , and m_4 , whereas $m_5 = (2, 2, 0, 0)$ is not dec-min. The supermodular function p for $B = B'(p)$ is given by

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

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 $\{s_1, s_2\} \subset \{s_1, s_2, s_3, s_4\}$, 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 (unique dec-min element) of the base-polyhedron B is given by $m_{\mathbf{R}} = (3/2, 3/2, 1/2, 1/2)$ by Theorem 3.2. 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 $S_2 = \{s_3, s_4\}$. Any of m_1, \dots, m_5 is near-uniform on S_1 and on S_2 , but m_5 is not dec-min because it fails to satisfy the tightness condition $\widetilde{m}(C_1) = p(C_1)$. ■

4 Proximity results and continuous relaxation algorithm

In general terms, a proximity result means a statement that the solutions to discrete and continuous versions of an optimization problem are geometrically close to each other. In this section we obtain proximity results for dec-min elements m_Z of an M-convex set $\overset{\dots}{B}$. We shall establish two proximity theorems, which refer to two different continuous problems. The first proximity theorem refers to the (fractional) dec-min element m_R of the base-polyhedron B , and the second to the minimizer of a piecewise-linear function associated with the square-sum $W(x) = \sum[x(s)^2 : s \in S]$. Both types of proximity results will be used as a basis for the continuous relaxation algorithm to be described in Section 4.3.

4.1 Proximity theorem using the fractional dec-min element

Our first proximity theorem reveals the geometric closeness of the dec-min element m_Z of $\overset{\dots}{B}$ to the dec-min element m_R of B . By Theorem 3.2, the dec-min element m_R of B is uniform on each member \hat{S}_i of the principal partition ($m_R(s) = \lambda_i$ for $s \in \hat{S}_i$), whereas the dec-min element m_Z of $\overset{\dots}{B}$ is near-uniform on each member S_j of the canonical partition ($m_Z(s) \in \{\beta_j, \beta_j - 1\}$ for $s \in S_j$) by Theorem 3.4. Combining these results with Theorem 3.9 connecting the principal and canonical partitions, we can obtain the following proximity theorem.

Theorem 4.1. *Let m_R be the dec-min element (minimum norm point) of an integral base-polyhedron B . Then every dec-min element m_Z of the associated M-convex set $\overset{\dots}{B}$ satisfies*

$$\lfloor m_R \rfloor \leq m_Z \leq \lceil m_R \rceil. \quad (4.1)$$

Proof. Fix $s \in S$ and let \hat{S}_i denote the member of the principal partition containing s , and λ_i be the associated critical value. We have $m_R(s) = \lambda_i$ by Theorem 3.2. By Theorem 3.9, $\lceil \lambda_i \rceil$ is an essential value, say, $\lceil \lambda_i \rceil = \beta_j$, where $1 \leq j \leq q$. Since the canonical partition is an aggregation of the principal partition, the corresponding member S_j of the canonical partition contains the element s . We have $m_Z(s) \in \{\beta_j, \beta_j - 1\}$ by Theorem 3.4. Therefore, $m_Z \leq \lceil m_R \rceil$.

Next we apply the above argument to $-B$, which is also an integral base-polyhedron. Since $-m_R$ is the minimum norm point of $-B$ and $-m_Z$ is a dec-min (=inc-max) element for $-\overset{\dots}{B}$, we obtain $-m_Z \leq \lceil -m_R \rceil$, which is equivalent to $m_Z \geq \lfloor m_R \rfloor$. This completes the proof of (4.1). ■

The above theorem states that every dec-min element m_Z of $\overset{\dots}{B}$ is located near the dec-min element m_R of B , satisfying $\lfloor m_R \rfloor \leq m_Z \leq \lceil m_R \rceil$. However, the converse is not true, that is, not every member m of $\overset{\dots}{B}$ satisfying $\lfloor m_R \rfloor \leq m \leq \lceil m_R \rceil$ is a dec-min element of $\overset{\dots}{B}$. This is demonstrated by the following example.

Example 4.1. Recall Example 3.2, where $\overset{\dots}{B}$ 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 $m_5 = (2, 2, 0, 0)$. The first four members, m_1 to m_4 , are the dec-min elements of the M-convex set $\overset{\dots}{B}$, whereas $m_5 = (2, 2, 0, 0)$ is not dec-min. The unique dec-min element of the (integral) base-polyhedron B is $m_R = (3/2, 3/2, 1/2, 1/2)$, for which $\lfloor m_R \rfloor = (1, 1, 0, 0)$ and $\lceil m_R \rceil = (2, 2, 1, 1)$. Every member m of $\overset{\dots}{B}$ satisfies $(1, 1, 0, 0) = \lfloor m_R \rfloor \leq m \leq \lceil m_R \rceil = (2, 2, 1, 1)$. In particular, $m_5 = (2, 2, 0, 0)$ satisfies $\lfloor m_R \rfloor \leq m_5 \leq \lceil m_R \rceil$, but it is not a dec-min element of $\overset{\dots}{B}$. ■

There is another connection between the dec-min elements in the continuous and discrete cases. While Theorem 4.1 above prescribes a region (box) for m_Z in terms of m_R , the following theorem is a statement in the reverse direction, showing that m_R is embraced by the dec-min elements m_Z for the discrete case.

Theorem 4.2. *The unique dec-min element m_R of an integral base-polyhedron B can be represented as a convex combination of the dec-min elements m_Z of the associated M -convex set $\overline{\overline{B}}$.*

Proof. It was shown in [9, Section 5.1] that the dec-min elements of $\overline{\overline{B}}$ lie on the face B^\oplus of B defined by the canonical chain $C_1 \subset C_2 \subset \dots \subset C_q$. This face is the intersection of B with the hyperplanes

$$\{x \in \mathbf{R}^S : \widetilde{x}(C_j) = p(C_j)\} \quad (j = 1, 2, \dots, q).$$

On the other hand, it is known ([13], [14, Section 9.2]) that the minimum norm point m_R of B lies on the face of B defined by the principal chain $\hat{C}_1 \subset \hat{C}_2 \subset \dots \subset \hat{C}_r$, which is the intersection of B with the hyperplanes

$$\{x \in \mathbf{R}^S : \widetilde{x}(\hat{C}_i) = p(\hat{C}_i)\} \quad (i = 1, 2, \dots, r).$$

Since the principal chain is a refinement of the canonical chain (Theorem 3.9), the latter face is a face of B^\oplus . Therefore, m_R belongs to B^\oplus . The point m_R also belongs to

$$T^* := \{x \in \mathbf{R}^S : \beta_j - 1 \leq x(s) \leq \beta_j \text{ for } s \in S_j \ (j = 1, 2, \dots, q)\},$$

since $m_R(s) = \lambda_i$ for $s \in \hat{S}_i$ (Theorem 3.2) and $S_j = \bigcup\{\hat{S}_i : \lceil \lambda_i \rceil = \beta_j\}$ (Theorem 3.9). Therefore, m_R is a member of $B^\bullet := B^\oplus \cap T^*$. Here B^\bullet is an integral base-polyhedron, and Theorem 5.1 of [9] states that the vertices of B^\bullet are precisely the dec-min elements of $\overline{\overline{B}}$. Therefore, m_R can be represented as a convex combination of the dec-min elements of $\overline{\overline{B}}$. ■

Remark 4.1. The convex combination property stated in Theorem 4.2 is no longer true for the intersection of two integral base-polyhedra. See [9, Example 7.1]. ■

4.2 Proximity theorem using a piecewise-linear square-sum minimizer

Our second proximity theorem shows the geometric closeness of the dec-min element m_Z of $\overline{\overline{B}}$ to the minimizer of a piecewise-linear function, to be denoted by $\overline{W}(x)$, arising from the square-sum $W(x) = \sum[x(s)^2 : s \in S]$. Recall from Theorem 2.7 that an element of $\overline{\overline{B}}$ is dec-min if and only if it is a minimizer of $W(x)$ over $\overline{\overline{B}}$.

To define the piecewise-linear function $\overline{W}(x)$, we first consider a piecewise-linear extension of the quadratic function $\varphi(k) = k^2$ in a single integer variable $k \in \mathbf{Z}$. The piecewise-linear extension $\overline{\varphi} : \mathbf{R} \rightarrow \mathbf{R}$ is a function in a real variable whose graph consists of line segments connecting (k, k^2) and $(k+1, (k+1)^2)$ for all $k \in \mathbf{Z}$. That is,

$$\overline{\varphi}(\xi) := (2k+1)|\xi| - k(k+1) \quad \text{with } k = \lfloor \xi \rfloor \quad (\xi \in \mathbf{R}). \quad (4.2)$$

It is noted that $\overline{\varphi}(\xi) = \xi^2$ for integers ξ and $\overline{\varphi}(\xi) > \xi^2$ for non-integral ξ ; for example, $\overline{\varphi}(1/2) = 1/2 > 1/4$. The piecewise-linear function $\overline{W}(x)$ is defined by

$$\overline{W}(x) := \sum[\overline{\varphi}(x(s)) : s \in S] \quad (x \in \mathbf{R}^S). \quad (4.3)$$

We have $\overline{W}(x) = W(x)$ for integral vectors x and $\overline{W}(x) > W(x)$ for non-integral vectors x .

The following fact is implicit in the proof of [14, Theorem 8.3].

Proposition 4.3. *The minimum value of \overline{W} over B is equal to the minimum square-sum on the M -convex set $\overline{\overline{B}}$. Moreover, for any minimizer $x_{\mathbf{R}} \in \mathbf{R}^S$ of the function \overline{W} over B , there exists a minimizer $x_{\mathbf{Z}}$ of \overline{W} over B satisfying $x_{\mathbf{Z}} \in \mathbf{Z}^S$ and $\lfloor x_{\mathbf{R}} \rfloor \leq x_{\mathbf{Z}} \leq \lceil x_{\mathbf{R}} \rceil$.*

Proof. (This proof is essentially the same as the proof of [14, Theorem 8.3].) Let $x_{\mathbf{R}}$ be a minimizer of \overline{W} over B , and denote the intersection of B with the box $\{x : \lfloor x_{\mathbf{R}} \rfloor \leq x \leq \lceil x_{\mathbf{R}} \rceil\}$ by B' , which is also an integral base-polyhedron. Since $\|\lceil x_{\mathbf{R}} \rceil - \lfloor x_{\mathbf{R}} \rfloor\|_{\infty} \leq 1$, the function \overline{W} is linear on B' . This implies that $x_{\mathbf{R}}$ can be expressed as a convex combination of some vertices z_1, z_2, \dots, z_k of B' and the function value $\overline{W}(x_{\mathbf{R}})$ is given by the corresponding convex combination of their function values. That is,

$$x_{\mathbf{R}} = \sum_{i=1}^k \alpha_i z_i, \quad \overline{W}(x_{\mathbf{R}}) = \sum_{i=1}^k \alpha_i \overline{W}(z_i),$$

where $\sum_{i=1}^k \alpha_i = 1$ and $\alpha_i > 0$ for all i . Since $x_{\mathbf{R}}$ is a minimizer of \overline{W} , we have $\overline{W}(x_{\mathbf{R}}) \leq \overline{W}(z_i)$ for all i , and hence each z_i is a minimizer of \overline{W} over B , for which $\overline{W}(x_{\mathbf{R}}) = \overline{W}(z_i)$. Moreover, z_i is an integral vector satisfying $\lfloor x_{\mathbf{R}} \rfloor \leq z_i \leq \lceil x_{\mathbf{R}} \rceil$. Therefore, we can take any z_i as $x_{\mathbf{Z}}$. ■

By combining Proposition 4.3 with Theorem 2.7 (characterizing dec-min elements of $\overline{\overline{B}}$ as square-sum minimizers), we obtain the following proximity statement.

Theorem 4.4. *For any minimizer $x_{\mathbf{R}}$ of the function \overline{W} over B , there exists a dec-min element $m_{\mathbf{Z}}$ of the associated M -convex set $\overline{\overline{B}}$ satisfying*

$$\lfloor x_{\mathbf{R}} \rfloor \leq m_{\mathbf{Z}} \leq \lceil x_{\mathbf{R}} \rceil. \quad (4.4)$$

Proof. By Proposition 4.3, there exists an integral vector $x_{\mathbf{Z}}$ that minimizes \overline{W} over B and satisfies $\lfloor x_{\mathbf{R}} \rfloor \leq x_{\mathbf{Z}} \leq \lceil x_{\mathbf{R}} \rceil$. Since \overline{W} coincides with W on $\overline{\overline{B}}$, this vector $x_{\mathbf{Z}}$ is a minimizer of W over $\overline{\overline{B}}$. This implies, by Theorem 2.7, that $x_{\mathbf{Z}}$ is a dec-min element of $\overline{\overline{B}}$ satisfying (4.4). Therefore, we can take this $x_{\mathbf{Z}}$ as $m_{\mathbf{Z}}$ in (4.4). ■

There are substantial differences between the two proximity results given in Theorem 4.4 and in Theorem 4.1. First, the vector $x_{\mathbf{R}}$ in Theorem 4.4, being an arbitrary minimizer of \overline{W} , is not uniquely determined, whereas $m_{\mathbf{R}}$ in Theorem 4.1 denotes the unique dec-min element of B . In particular, $x_{\mathbf{R}}$ is not necessarily dec-min in B (see Example 4.2 below). Second, the box $\lfloor x_{\mathbf{R}} \rfloor \leq x \leq \lceil x_{\mathbf{R}} \rceil$ in (4.4) of Theorem 4.4 may possibly miss some dec-min elements of $\overline{\overline{B}}$ (see Example 4.2), while the box $\lfloor m_{\mathbf{R}} \rfloor \leq x \leq \lceil m_{\mathbf{R}} \rceil$ in (4.1) of Theorem 4.1 contains all dec-min elements of $\overline{\overline{B}}$.

Example 4.2. We continue with the problem treated in Examples 3.2 and 4.1. The M -convex set $\overline{\overline{B}}$ 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 $m_5 = (2, 2, 0, 0)$. We have $W(m_i) = 6$ for $i = 1, 2, 3, 4$ and $W(m_5) = 8$. Hence the minimum of \overline{W} over B is equal to 6. Consider a vector $x_{\mathbf{R}} = (2, 1, 1/3, 2/3)$, which is a minimizer of $\overline{W}(x)$ since

$$\overline{W}(x_{\mathbf{R}}) = \overline{\varphi}(2) + \overline{\varphi}(1) + \overline{\varphi}(1/3) + \overline{\varphi}(2/3) = 4 + 1 + 1/3 + 2/3 = 6.$$

For this vector, we have $\lfloor x_{\mathbf{R}} \rfloor = (2, 1, 0, 0)$ and $\lceil x_{\mathbf{R}} \rceil = (2, 1, 1, 1)$. The box $\lfloor x_{\mathbf{R}} \rfloor \leq x \leq \lceil x_{\mathbf{R}} \rceil$ contains $m_1 = (2, 1, 1, 0)$ and $m_2 = (2, 1, 0, 1)$, but misses the other two dec-min elements,

m_3 and m_4 , of $\overset{\dots}{B}$. Another possible choice of a minimizer of \overline{W} is the minimum norm point $m_{\mathbf{R}} = (3/2, 3/2, 1/2, 1/2)$, which is indeed a minimizer of \overline{W} (see Proposition 4.5 below) with

$$\overline{W}(m_{\mathbf{R}}) = \overline{\varphi}(3/2) + \overline{\varphi}(3/2) + \overline{\varphi}(1/2) + \overline{\varphi}(1/2) = 5/2 + 5/2 + 1/2 + 1/2 = 6.$$

For this vector, we have $\lfloor m_{\mathbf{R}} \rfloor = (1, 1, 0, 0)$ and $\lceil m_{\mathbf{R}} \rceil = (2, 2, 1, 1)$. The box $\lfloor m_{\mathbf{R}} \rfloor \leq x \leq \lceil m_{\mathbf{R}} \rceil$ contains all the four dec-min elements, and additionally, $m_5 = (2, 2, 0, 0)$ (which is not dec-min). ■

We point out here that the minimum norm point $m_{\mathbf{R}}$, which is the unique minimizer of the square-sum $W(x)$ over B , is also a minimizer of the associated piecewise-linear function \overline{W} . This fact is quite natural to expect, but it is a non-trivial fact whose proof relies on the property (Theorem 4.2) that $m_{\mathbf{R}}$ lies in the convex hull of the dec-min elements of $\overset{\dots}{B}$.

Proposition 4.5. *The minimum norm point $m_{\mathbf{R}}$ of B is a minimizer of the piecewise-linear function \overline{W} on B .*

Proof. Theorem 4.2 ensures a convex combination $m_{\mathbf{R}} = \sum_{i=1}^k \alpha_i m_i$, where m_1, m_2, \dots, m_k are dec-min elements of $\overset{\dots}{B}$, $\sum_{i=1}^k \alpha_i = 1$, and $\alpha_i > 0$ for all i . Since $\|m_i - m_j\|_{\infty} \leq 1$, the function \overline{W} is linear on the convex hull of m_1, m_2, \dots, m_k , from which follows that

$$\overline{W}(m_{\mathbf{R}}) = \sum_{i=1}^k \alpha_i \overline{W}(m_i) = \sum_{i=1}^k \alpha_i W(m_i).$$

Here each of m_1, m_2, \dots, m_k is a minimizer of W (by Theorem 2.7) and $\min\{\overline{W}(x) : x \in B\} = \min\{W(x) : z \in \overset{\dots}{B}\}$ by the definition of \overline{W} and the integrality of B . Therefore, $m_{\mathbf{R}}$ is a minimizer of \overline{W} . ■

4.3 Continuous relaxation algorithm

In our continuous relaxation algorithm for computing a dec-min element of $\overset{\dots}{B}$, it is assumed that we are given a real vector x^* such that the box bounded by $\lfloor x^* \rfloor$ and $\lceil x^* \rceil$ contains at least one dec-min element of $\overset{\dots}{B}$. That is, we assume that

$$\lfloor x^* \rfloor \leq m_{\mathbf{Z}} \leq \lceil x^* \rceil \tag{4.5}$$

holds for some dec-min element $m_{\mathbf{Z}}$ of $\overset{\dots}{B}$. By Theorem 4.1 the minimum norm point $m_{\mathbf{R}}$ of B serves as such x^* . Another possibility for x^* is an arbitrary minimizer $x_{\mathbf{R}}$ of the piecewise-linear function $\overline{W}(x)$, as shown in Theorem 4.4. The algorithm of this section relies only on the property (4.5) of the vector x^* , which is regarded as an input of the algorithm. At the end of this section, we indicate some references concerning the computation of $m_{\mathbf{R}}$ and $x_{\mathbf{R}}$ in Remarks 4.3 and 4.4, respectively.

Using the given vector x^* satisfying (4.5), define

$$\ell := \lfloor x^* \rfloor, \quad u := \lceil x^* \rceil, \tag{4.6}$$

and denote the intersection of $\overset{\dots}{B}$ and box $[\ell, u]$ by $\overset{\dots}{B}_{\ell}^u$, that is,

$$\overset{\dots}{B}_{\ell}^u := \overset{\dots}{B} \cap \{x : \ell \leq x \leq u\}. \tag{4.7}$$

By our assumption, $\overset{\dots}{B}_\ell^u$ contains at least one dec-min element of $\overset{\dots}{B}$. This implies that we can find a dec-min element of $\overset{\dots}{B}$ by computing a dec-min element of $\overset{\dots}{B}_\ell^u$. Thus the dec-min problem on the M-convex set $\overset{\dots}{B}$ is reduced to that on a smaller M-convex set $\overset{\dots}{B}_\ell^u$.

A dec-min element of $\overset{\dots}{B}_\ell^u$ can be computed as follows. Since $\mathbf{0} \leq u - \ell \leq \mathbf{1}$, the set $\overset{\dots}{B}_\ell^u$ is a matroidal M-convex set, and hence it can be represented as

$$\overset{\dots}{B}_\ell^u = \{\ell + \chi_L : L \text{ is a base of } M^\bullet\}$$

for some matroid M^\bullet on S . Define a weight function $\omega : S \rightarrow \mathbf{Z}$ by

$$\omega(s) := u(s)^2 - \ell(s)^2 \quad (s \in S). \quad (4.8)$$

Then the square-sum $W(x)$ of $x = \ell + \chi_L \in \overset{\dots}{B}_\ell^u$ can be expressed as

$$W(x) = \sum_{s \in S} x(s)^2 = \sum_{s \in L} u(s)^2 + \sum_{s \in S-L} \ell(s)^2 = \widetilde{\omega}(L) + \sum_{s \in S} \ell(s)^2.$$

This shows that minimizing $W(x)$ over $\overset{\dots}{B}_\ell^u$ is equivalent to finding a minimum ω -weight base of M^\bullet , whereas, by Theorem 2.7, an element of $\overset{\dots}{B}_\ell^u$ is a minimizer of $W(x)$ over $\overset{\dots}{B}_\ell^u$ if and only if it is dec-min in $\overset{\dots}{B}_\ell^u$. Therefore, a dec-min element of $\overset{\dots}{B}_\ell^u$ can be computed by finding a minimum ω -weight base of matroid M^\bullet . The latter can be done in strongly polynomial time by the greedy algorithm (see, e.g., [7, 35]). In order to apply the greedy algorithm, one needs an evaluation oracle that outputs the rank $r^\bullet(Z)$ of any input subset $Z \subseteq S$ in matroid M^\bullet . But $r^\bullet(Z)$ can be computed from the supermodular function p associated with $\overset{\dots}{B}$ and from the bounding vectors ℓ and u defined in (4.6) with the aid of a submodular function minimization algorithm. Therefore, the above procedure, when x^* is given, finds a dec-min element of $\overset{\dots}{B}$ in strongly polynomial time.

Example 4.3. We illustrate the continuous relaxation algorithm for the problem considered in Examples 3.2, 4.1, and 4.2, where $\overset{\dots}{B}$ 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 $m_5 = (2, 2, 0, 0)$; m_k is dec-min for $k = 1, 2, 3, 4$, while m_5 is not.

Suppose first that the minimum norm point $m_{\mathbf{R}} = (3/2, 3/2, 1/2, 1/2)$ is chosen as x^* . Then we obtain $\ell = \lfloor m_{\mathbf{R}} \rfloor = (1, 1, 0, 0)$ and $u = \lceil m_{\mathbf{R}} \rceil = (2, 2, 1, 1)$, and hence $\overset{\dots}{B}_\ell^u = \overset{\dots}{B}$ and $\omega = (3, 3, 1, 1)$. For $x = m_1$ we have $m_1 - \ell = (2, 1, 1, 0) - (1, 1, 0, 0) = (1, 0, 1, 0) = \chi_{L_1}$ with $L_1 = \{s_1, s_3\}$ and $\widetilde{\omega}(L_1) = 3 + 0 + 1 + 0 = 4$. Similarly, we have $m_k - \ell = \chi_{L_k}$ with $\widetilde{\omega}(L_k) = 4$ for $k = 2, 3, 4$. For $x = m_5$ we have $m_5 - \ell = (2, 2, 0, 0) - (1, 1, 0, 0) = (1, 1, 0, 0) = \chi_{L_5}$ with $L_5 = \{s_1, s_2\}$ and $\widetilde{\omega}(L_5) = 3 + 3 + 0 + 0 = 6$. Therefore, L_k is a minimum ω -weight base for $k = 1, 2, 3, 4$, while L_5 is not. In other words, m_k is dec-min in $\overset{\dots}{B}$ for $k = 1, 2, 3, 4$, while m_5 is not.

As the second choice of x^* , consider a minimizer $x_{\mathbf{R}} = (2, 1, 1/3, 2/3)$ of $\overline{W}(x)$, for which $\ell = \lfloor x_{\mathbf{R}} \rfloor = (2, 1, 0, 0)$, $u = \lceil x_{\mathbf{R}} \rceil = (2, 1, 1, 1)$, and hence $\overset{\dots}{B}_\ell^u = \{m_1, m_2\}$ and $\omega = (0, 0, 1, 1)$. We have $m_1 - \ell = (0, 0, 1, 0) = \chi_{L_1}$ with $L_1 = \{s_3\}$ and $\widetilde{\omega}(L_1) = 1$, while $m_2 - \ell = (0, 0, 0, 1) = \chi_{L_2}$ with $L_2 = \{s_4\}$ and $\widetilde{\omega}(L_2) = 1$. Therefore, we can conclude that both m_1 and m_2 are dec-min in $\overset{\dots}{B}$. The other two dec-min elements of $\overset{\dots}{B}$, m_3 and m_4 , are not captured when $x^* = x_{\mathbf{R}} = (2, 1, 1/3, 2/3)$. \blacksquare

Remark 4.2. The continuous relaxation algorithm using the minimum norm point $m_{\mathbf{R}}$ can cope with the minimum cost dec-min element problem (Section 2.4), since all dec-min elements of $\overset{\dots}{B}$ are captured by $\overset{\dots}{B}_\ell^u$ (Theorem 4.1). In contrast, the continuous relaxation algorithm using a minimizer $x_{\mathbf{R}}$ of $\overline{W}(x)$ cannot be used to solve this problem. ■

Remark 4.3. Several different algorithms are known for computing the minimum norm point $m_{\mathbf{R}}$, with varying theoretical complexity and practical efficiency. Fujishige’s decomposition algorithm [13] (also [14, Section 8.2]) computes $m_{\mathbf{R}}$ in strongly polynomial time. We can also compute $m_{\mathbf{R}}$ by Wolfe’s minimum norm point algorithm [39] tailored to base-polyhedra [14, Section 7.1], which algorithm has drawn a renewed interest as a practically effective subroutine in submodular function minimization (Chakrabarty–Jain–Kothari [2], De Loera–Haddock–Rademacher [3], Fujishige–Isotani [16]). ■

Remark 4.4. 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 [21], Ibaraki–Katoh [23], Katoh–Shioura–Ibaraki [26]). The continuous relaxation approach for discrete variables is considered, e.g., by Hochbaum [20] and Hochbaum–Hong [22]. A paper by Moriguchi–Shioura–Tsuchimura [29] discusses this approach in the more general context of M-convex function minimization. It is known ([22, 29, 36], [26, Theorem 23]) that a convex quadratic function $\sum a_i x_i^2$ in discrete variables can be minimized over an integral base-polyhedron in strongly polynomial time. See Végh [37] for a recent development on the complexity of separable convex function minimization. ■

5 Decomposition algorithms for square-sum minimization on an M-convex set

5.1 General remarks

As a continuation from Section 2.8, we shall present details of the decomposition algorithms for square-sum minimization on an M-convex set. Recall from Theorem 2.7 that minimizing the square-sum $W(x) = \sum [x(s)^2 : s \in S]$ on an M-convex set $\overset{\dots}{B}$ is equivalent to computing a dec-min element of $\overset{\dots}{B}$.

In order to present a clear overview of the existing approaches, we consider two variants of the decomposition algorithms, one based on Groenevelt [18] and the other based on Fujishige [14, Section 8.2] (see Table 1). Although the same name of “decomposition algorithm” is used, Groenevelt’s and Fujishige’s are not exactly the same. To be specific, Groenevelt’s algorithm features a subproblem named “single constraint problem” and decomposes the ground-set into two disjoint subsets, whereas Fujishige’s algorithm uses a variable corresponding to a subgradient (denoted η in [14, Section 8.2]) and decomposes the ground-set into three disjoint subsets. It is noted that Fujishige’s original algorithm [13], targeted to quadratic functions in continuous variables, was based on a simplest special case of “single constraint problem” (without using this terminology) and decomposes the ground-set into two disjoint subsets.

In Section 5.2, we show an adaptation of Groenevelt’s decomposition algorithm to square-sum minimization on an M-convex set, and call it the “Groenevelt-type decomposition algorithm.” In its original form, Groenevelt’s algorithm [18] for Case \mathbf{Z} dealt with an integral

polymatroid (not an integral base-polyhedron). Although the adaptation to an M-convex set is not that difficult, it will be important to have a precise description of the algorithm at hand along with proofs of correctness and strong polynomiality of the algorithm.

In Section 5.3, we develop another decomposition algorithm for square-sum minimization on an M-convex set on the basis of the framework of Fujishige [14, Section 8.2] for separable convex minimization on a base-polyhedron. We refer to the resulting algorithm as the “Fujishige-type decomposition algorithm.” As suggested in [14, Section 8.3], the framework of [14, Section 8.2] for Case **R** can be adapted to Case **Z** with the use of the piecewise-linear extension $\overline{W}(x)$ in (4.3). Our contribution consists in devising a concrete computational procedure for a key subroutine assumed in the general framework, giving self-contained rigorous proofs of correctness and strong polynomiality of the algorithm, and revealing the relation between the computed decomposition and the canonical chain as well as the certifying chain in Theorem 2.3.

In relation to the structural and algorithmic results in [9, 10], the two decomposition algorithms, Groenevelt-type and Fujishige-type, have the following characteristics.

- Unlike the algorithm in [10], the two decomposition algorithms do not rely on the Newton–Dinkelbach algorithm.
- The two decomposition algorithms decompose the ground-set, but the resulting decomposition may or may not coincide with the canonical partition, whereas the algorithm in [10] iteratively construct the canonical partition. It should be noted, however, that if a single dec-min element of an M-convex set is available, the canonical partition can be computed quite easily [10, Algorithm 2.3].
- The Groenevelt-type algorithm is simpler than the Fujishige-type, both in steps of the algorithm and the proof of correctness, and is rather independent of the structural results found in [9]. In contrast, the decomposition computed in the Fujishige-type algorithm is consistent with a characterization of dec-minimality given in [9], as pointed out in Remark 5.3.

5.2 Groenevelt-type decomposition algorithm

In this section, we show an adaptation of Groenevelt’s decomposition algorithm to square-sum minimization on an M-convex set, which we call “Groenevelt-type decomposition algorithm” in this paper.

Let $B = B'(p)$ be an integral base-polyhedron on a ground-set S described by an integer-valued supermodular function p , and $\overline{B} = B \cap \mathbf{Z}^S$ be the associated M-convex set. For any subset S_+ of S , the restriction of B to S_+ means the base-polyhedron $B_+ := B'(p_+)$ described by the supermodular function p_+ defined by $p_+(X) = p(X)$ for $X \subseteq S_+$. For any subset S_- of S , the contraction of B to S_- is the base-polyhedron $B_- := B'(p_-)$ described by the supermodular function p_- defined by $p_-(X) = p(X \cup (S - S_-)) - p(S - S_-)$ for $X \subseteq S_-$. We also define the supermodular polyhedron

$$Q = \{y \in \mathbf{R}^S : y(X) \geq p(X) (\forall X \subseteq S)\}, \quad (5.1)$$

which can also be defined as $Q = \{y \in \mathbf{R}^S : y \geq x \text{ for some } x \in B\}$.

Groenevelt's framework employs an auxiliary subproblem, called "single constraint problem." In our case of square-sum minimization, this subproblem is given by

$$\text{Minimize } \sum_{s \in S} x(s)^2 \quad \text{subject to } x \in \mathbf{Z}^S, x(S) = p(S), \quad (5.2)$$

which (fortunately) admits an explicit solution. Let

$$a := \lfloor p(S)/|S| \rfloor, \quad k := p(S) - a|S|, \quad (5.3)$$

where $0 \leq k \leq |S| - 1$. Then a vector x is a solution to the problem (5.2) if and only if $x(s) \in \{a, a + 1\}$ for all $s \in S$ and precisely k of $x(s)$ are equal to $a + 1$; that is,

$$x = a\chi_S + \chi_U \quad (5.4)$$

for any $U \subset S$ with $|U| = k$. For example, $x = (\overbrace{a + 1, \dots, a + 1}^k, a, \dots, a)$.

With above preparations we can describe the algorithm to minimize the square-sum $W(x)$ on an M-convex set.

Groenevelt-type decomposition algorithm for square-sum on an M-convex set $\overset{\dots}{B}$

- 1: Let $a := \lfloor p(S)/|S| \rfloor$, $k := p(S) - a|S|$, and $x := a\chi_S + \chi_U$ for an arbitrary U with $|U| = k$.
- 2: If x belongs to $\overset{\dots}{B}$, then let $z^* := x$ and stop.
- 3: Find a minimal vector y satisfying $y \geq x$ and $y \in \overset{\dots}{Q}$.
- 4: Let S_+ be the largest subset $X \subseteq S$ satisfying $y(X) = p(X)$, and let $S_- := S \setminus S_+$.
- 5: Apply this algorithm recursively to the restriction $\overset{\dots}{B}_+$ to S_+ . Let $z_+^* \in \mathbf{Z}^{S_+}$ be the output.
- 6: Apply this algorithm recursively to the contraction $\overset{\dots}{B}_-$ to S_- . Let $z_-^* \in \mathbf{Z}^{S_-}$ be the output.
- 7: Define $z^*(s) := z_+^*(s)$ for $s \in S_+$ and $z^*(s) := z_-^*(s)$ for $s \in S_-$, and stop.

Basic properties of the above algorithm are given below. The second property (2) implies that the recursive calls in Step 5 and Step 6 make sense.

Proposition 5.1.

- (1) When the algorithm terminates in Step 2, the output z^* is a square-sum minimizer.
- (2) In Step 4, we have $\emptyset \neq S_+ \neq S$ and $\emptyset \neq S_- \neq S$.
- (3) z^* defined in Step 7 is a member of $\overset{\dots}{B}$.
- (4) $y(s) = x(s)$ for all $s \in S - S_+$.

Proof. (1) z^* is near-uniform at the termination in Step 2, and hence it is decreasingly minimal, or equivalently, a square-sum minimizer.

(2) We have $y \geq x$, $y \neq x$, $x \notin \overset{\dots}{B}$, and $y \in \overset{\dots}{Q}$ in Step 4. Since y is not equal to x , there exists an element $s \in S$ with $y(s) > x(s)$. Then the minimality of y implies that such s must belong to S_+ , implying $S_+ \neq \emptyset$. Because of such s , we have $\widetilde{y}(S) > \widetilde{x}(S) = p(S)$, which shows $S_+ \neq S$.

(3) This is immediate from the fact that $z^*|_{S_+} = z_+^* \in \overset{\dots}{B}_+$ and $z^*|_{S_-} = z_-^* \in \overset{\dots}{B}_-$, where $\overset{\dots}{B}_+$ is the restriction of $\overset{\dots}{B}$ to S_+ and $\overset{\dots}{B}_-$ is the contraction of $\overset{\dots}{B}$ to S_- .

(4) By definition, y is a minimal element of $\overset{\dots}{Q}$ satisfying $y \geq x$, and S_+ is the largest y -tight set with respect to p . Suppose, indirectly, that $y(s) > x(s)$ for some $s \in S - S_+$. Define $\hat{y} := y - \chi_s$. This \hat{y} belongs to $\overset{\dots}{Q}$, since $y(X) > p(X)$ for any X containing s . Also we have $\hat{y} \geq x$, a contradiction to the minimality of y . \blacksquare

The correctness of the algorithm is established in the following proposition.

Proposition 5.2. *The output z^* of the Groenevelt-type decomposition algorithm is a minimizer of square-sum $W(x)$ over $\overset{\dots}{B}$, or equivalently, z^* is a dec-min element of $\overset{\dots}{B}$.*

Proof. By induction on the size of the ground-set S , we prove that the output of the algorithm is a square-sum minimizer. By Proposition 5.1(2), we have $|S_+| < |S|$ and $|S_-| < |S|$. First we show that the output z^* of the algorithm satisfies

$$z^*(s) \geq a \quad (s \in S_+), \quad z^*(s) \leq a + 1 \quad (s \in S_-). \quad (5.5)$$

To prove the first inequality $z^*(s) \geq a$ ($s \in S_+$) in (5.5), let z_+^* denote the subvector of z^* on S_+ (i.e., $z_+^* = z^*|_{S_+}$), which is the outcome of the recursive call to the restriction $\overset{\dots}{B}_+$ to S_+ . By the induction hypothesis, z_+^* minimizes the square-sum $\sum[x(s)^2 : s \in S_+]$ over $\overset{\dots}{B}_+$. This implies, by Theorems 2.2 and 2.7, that z_+^* is an inc-max element of $\overset{\dots}{B}_+$. On the other hand, $y|_{S_+}$ is a member of $\overset{\dots}{B}_+$ (since S_+ is y -tight with respect to p) satisfying $(y|_{S_+})(s) \geq x(s) \geq a$ for all $s \in S_+$. It then follows that each component of z_+^* is bounded from below by a . Therefore, $z^*(s) \geq a$ for $s \in S_+$.

The second inequality $z^*(s) \leq a + 1$ ($s \in S_-$) in (5.5) can be proved as follows. By the induction hypothesis, the vector $z_-^* := z^*|_{S_-}$ minimizes the square-sum over the contraction $\overset{\dots}{B}_-$ to S_- , and hence z_-^* is a dec-min element of $\overset{\dots}{B}_-$ by Theorem 2.7, which further implies that z_-^* is a dec-min element of $\overset{\dots}{Q}_- = \{v \in \mathbf{Z}^{S_-} : v \geq u \text{ for some } u \in \overset{\dots}{B}_-\}$. On the other hand, $y|_{S_-}$ is a member of $\overset{\dots}{Q}_-$ satisfying $(y|_{S_-})(s) \leq a + 1$ for all $s \in S_-$, since $y(s) = x(s) \leq a + 1$ for all $s \in S_-$ by Proposition 5.1(4) and (5.4). Therefore $z^*(s) \leq a + 1$ for all $s \in S_-$. Thus (5.5) is proved.

Finally we show that z^* satisfies the condition

$$z^*(t) \geq z^*(s) + 2 \implies z^* + \chi_s - \chi_t \notin \overset{\dots}{B} \quad (5.6)$$

for all (s, t) with $s, t \in S_+ \cup S_-$. Suppose that $z^*(t) \geq z^*(s) + 2$. By (5.5) we may assume $(s, t) \notin S_+ \times S_-$. The condition (5.6) holds when $(s, t) \in S_+ \times S_+$, since the vector $z_+^* = z^*|_{S_+}$ is a dec-min element of $\overset{\dots}{B}_+$. Similarly, (5.6) holds when $(s, t) \in S_- \times S_-$, since the vector $z_-^* = z^*|_{S_-}$ is a dec-min element of $\overset{\dots}{B}_-$. It remains to consider the case of $(s, t) \in S_- \times S_+$. Since $z_+^* \in \overset{\dots}{B}_+$, the set S_+ is z^* -tight with respect to p . It then follows that $z^* + \chi_s - \chi_t \notin \overset{\dots}{B}$ if $s \in S_-$ and $t \in S_+$. Hence, the condition (5.6) holds for $(s, t) \in S_- \times S_+$. By Theorems 2.3 and 2.7, this completes the proof of Proposition 5.2. \blacksquare

For clarity we make an explicit statement about strong polynomiality of the algorithm.

Proposition 5.3. *The Groenevelt-type decomposition algorithm computes a square-sum minimizer (i.e., dec-min element) of an M -convex set in strongly polynomial time.*

Proof. The number of recursive calls is bounded by $|S|$, and each step can be done in strongly polynomial time (using a submodular function minimization subroutine). \blacksquare

The algorithm is illustrated for a simple example.

Example 5.1. Let $S = \{s_1, s_2\}$ and $\overset{\dots}{B} = \{(3, 2), (4, 1), (5, 0)\}$, which has a unique dec-min element $m_Z = (3, 2)$. The defining supermodular function p is given by

$$p(\emptyset) = 0, \quad p(\{s_1\}) = 3, \quad p(\{s_2\}) = 0, \quad p(\{s_1, s_2\}) = 5.$$

In Step 1 of the Groenevelt-type decomposition algorithm, we obtain $a = \lfloor p(S)/|S| \rfloor = \lfloor 5/2 \rfloor = 2$ and $k = p(S) - a|S| = 1$. We have two choices for x , namely, $x^{(1)} = (2, 3)$ and $x^{(2)} = (3, 2)$.

The first vector $x^{(1)} = (2, 3)$ does not belong to $\overset{\dots}{B}$, and the vector y in Step 3 is given (uniquely) by $y = (3, 3)$, for which we obtain $S_+ = \{s_1\}$ and $S_- = \{s_2\}$ in Step 4. In Step 5, the restriction $\overset{\dots}{B}_+$ consists of a single number (one-dimensional vector) 3, that is, $\overset{\dots}{B}_+ = \{3\}$, for which $z_+^* = 3$. In Step 6, the contraction is given by $\overset{\dots}{B}_- = \{2\}$, for which $z_-^* = 2$. In Step 7, we obtain $z^* = (z_+^*, z_-^*) = (3, 2)$, which is the dec-min element of $\overset{\dots}{B}$.

The second vector $x^{(2)} = (3, 2)$ is already in the given M-convex set $\overset{\dots}{B}$. Therefore, the algorithm terminates at Step 2 with $z^* = x^{(2)} = (3, 2)$. Note that no recursive calls are involved, which means that the dec-min element is computed without decomposing the ground-set. ■

The Groenevelt-type decomposition algorithm may not find the canonical partition. The resulting decomposition can be coarser or finer than the canonical partition, which is shown in the following examples.

Example 5.2. For the problem of Example 5.1, we have

$$p(X) - \beta|X| = \begin{cases} 0 & (X = \emptyset), \\ 3 - \beta & (X = \{s_1\}), \\ -\beta & (X = \{s_2\}), \\ 5 - 2\beta & (X = \{s_1, s_2\}). \end{cases}$$

By Proposition 3.8, there are two essential values $\beta_1 = 3$ and $\beta_2 = 2$, and the canonical partition is a bipartition $\{S_1, S_2\}$ given by

$$\begin{aligned} S_1 &= L(\beta_1 - 1) - L(\beta_1) = L(2) - L(3) = \{s_1\} - \emptyset = \{s_1\}, \\ S_2 &= L(\beta_2 - 1) - L(\beta_2) = L(1) - L(2) = \{s_1, s_2\} - \{s_1\} = \{s_2\}. \end{aligned}$$

The first choice $x^{(1)} = (2, 3)$ in Example 5.1 results in the decomposition with $S_+ = \{s_1\}$ and $S_- = \{s_2\}$, which coincides with the canonical partition, whereas the second choice $x^{(2)} = (3, 2)$ does not decompose S at all, that is, results in a decomposition coarser than the canonical partition. ■

Example 5.3. Let $S = \{s_1, s_2, s_3, s_4\}$ and $\overset{\dots}{B} = \{(1, 0, 1, 0), (1, 0, 0, 1), (0, 1, 0, 1), (0, 1, 1, 0)\}$, in which every element is dec-min. The defining supermodular function p is given by

$$\begin{aligned} p(\emptyset) &= 0, & p(S) &= 2, & p(\{s_i\}) &= 0, & p(S - \{s_i\}) &= 1 \quad (i = 1, \dots, 4), \\ p(\{s_1, s_2\}) &= p(\{s_3, s_4\}) &= 1, & p(X) &= 0 & \text{for other } X \text{ with } |X| = 2. \end{aligned}$$

We can verify, as in Example 5.2, that the canonical partition is a trivial partition consisting of S itself. The Groenevelt-type decomposition algorithm may result in a finer decomposition. In Step 1, we get $a = \lfloor p(S)/|S| \rfloor = \lfloor 2/4 \rfloor = 0$ and $k = p(S) - a|S| = 2$. Suppose we have chosen $x = (0, 0, 1, 1)$, which does not belong to $\overset{\dots}{B}$. For this x , we may take $y = (0, 1, 1, 1) \in \overset{\dots}{Q}$ in Step 3, for which $S_+ = \{s_1, s_2\}$ and $S_- = \{s_3, s_4\}$. We have $\overset{\dots}{B}_+ = \{(1, 0), (0, 1)\}$ and may take $z_+^* = (1, 0)$, whereas we have $\overset{\dots}{B}_- = \{(1, 0), (0, 1)\}$ and may take $z_-^* = (1, 0)$. Then we obtain $z^* = (z_+^*, z_-^*) = (1, 0, 1, 0)$. The resulting decomposition of S is a bipartition $\{\{s_1, s_2\}, \{s_3, s_4\}\}$, which is finer than the canonical partition. It is noteworthy that the first member $S_+ = \{s_1, s_2\}$ is not even a z^* -top set. ■

5.3 Fujishige-type decomposition algorithm

In this section, we develop another decomposition algorithm for square-sum minimization on an M-convex set using the framework of Fujishige [14, Section 8.2] for separable convex minimization on a base-polyhedron. The proposed algorithm is based on the natural idea to apply the framework of [14, Section 8.2] to the piecewise-linear extension $\overline{W}(x)$ in (4.3). To ensure strong polynomiality, however, we need to devise a non-trivial gadget to cope with complications arising from integrality. The relation of the proposed algorithm to the framework of [14, Section 8.2] is explained in Remark 5.4 at the end of this section.

By definition, an element z of an M-convex set $\overline{\overline{B}}$ admits no 1-tightening step if it satisfies the condition:

$$z(t) \geq z(s) + 2 \implies z + \chi_s - \chi_t \notin \overline{\overline{B}}. \quad (5.7)$$

This condition, Condition (B) in Theorem 2.3, is necessary and sufficient for $z \in \overline{\overline{B}}$ to be a dec-min element of $\overline{\overline{B}}$. It is worth noting that this condition coincides with the local optimality condition [31, Theorem 6.26] for M-convex function minimization applied to the function $W(x)$ over $\overline{\overline{B}}$.

We consider a relaxation of the condition (5.7) on z , which, for any given integer a , requires that

$$z(s) \leq a, \quad z(t) \geq a + 1, \quad z(t) \geq z(s) + 2 \implies z + \chi_s - \chi_t \notin \overline{\overline{B}}. \quad (5.8)$$

Obviously, the condition (5.7) is stronger than (5.8) for any fixed a .

To consider algorithmic aspects of (5.8), it is convenient to relate (5.8) to convex minimization. Define functions

$$g_a(k) := \max(a - k, 0, k - a - 1) \quad (k \in \mathbf{Z}), \quad (5.9)$$

$$\begin{aligned} G_a(x) &:= \sum_{s \in S} g_a(x(s)) \\ &= \sum_{s \in S} \max(a - x(s), 0, x(s) - a - 1) \quad (x \in \mathbf{Z}^S). \end{aligned} \quad (5.10)$$

Then, (5.8) can be recognized as a local optimality condition for the minimization of $G_a(x)$ over $\overline{\overline{B}}$.

Proposition 5.4. $z \in \overline{\overline{B}}$ satisfies (5.8) if and only if z is a minimizer of G_a over $\overline{\overline{B}}$.

Proof. Since G_a is a separable convex function, its restriction to the M-convex set $\overline{\overline{B}}$ is an M-convex function (see [31, (6.31)]). By the optimality criterion [31, Theorem 6.26] for M-convex functions, z is a minimizer of G_a over $\overline{\overline{B}}$ if and only if, for any $s, t \in S$, we have the local optimality that

$$z + \chi_s - \chi_t \in \overline{\overline{B}} \implies G_a(z + \chi_s - \chi_t) \geq G_a(z),$$

which is equivalent to

$$G_a(z + \chi_s - \chi_t) < G_a(z) \implies z + \chi_s - \chi_t \notin \overline{\overline{B}}.$$

Here we have

$$\begin{aligned} G_a(z + \chi_s - \chi_t) &< G_a(z) \\ \iff (g_a(z(s) + 1) - g_a(z(s))) + (g_a(z(t) - 1) - g_a(z(t))) &< 0 \\ \iff [z(s) \leq a, z(t) \geq a + 1, z(t) \geq z(s) + 2]. \end{aligned}$$

Hence follows the claim of Proposition 5.4. ■

The following proposition shows that an integral vector $z \in \overset{\dots}{B}$ satisfying (5.8) can be computed in strongly polynomial time.

Proposition 5.5. *For any given $a \in \mathbf{Z}$, the function G_a can be minimized over $\overset{\dots}{B}$ in strongly polynomial time. Equivalently, a member of $\overset{\dots}{B}$ satisfying (5.8), for any given $a \in \mathbf{Z}$, can be found in strongly polynomial time.*

Proof. We express a vector $x \in \mathbf{Z}^S$ as $x = x^1 + x^2 + x^3$, where x^1 , x^2 , and x^3 are integer vectors and

$$x^1(s) \leq a, \quad 0 \leq x^2(s) \leq 1, \quad 0 \leq x^3(s) \quad (s \in S).$$

Consider three disjoint copies S^1, S^2, S^3 of the ground-set S , and regard x^i as a vector on S^i for $i = 1, 2, 3$. The set defined by

$$\begin{aligned} \overset{\dots}{B}' = & \{(x^1, x^2, x^3) \in \mathbf{Z}^{S^1 \cup S^2 \cup S^3} \mid \\ & x^1 + x^2 + x^3 \in \overset{\dots}{B}, \quad x^1 \leq a\mathbf{1}, \quad \mathbf{0} \leq x^2 \leq \mathbf{1}, \quad \mathbf{0} \leq x^3\} \end{aligned}$$

is an M-convex set (which can be proved directly or by using [33, Proposition 3.3]). Furthermore, the minimization of $G_a(x)$ over $\overset{\dots}{B}$ is reduced to the minimization of $\sum_{s \in S^1} (a - x^1(s)) + \sum_{s \in S^3} x^3(s)$ over $\overset{\dots}{B}'$. The latter problem is a linear optimization over a base-polyhedron, and the required subroutines for $\overset{\dots}{B}'$ can be realized from those of the given M-convex set $\overset{\dots}{B}$ with the aid of a submodular function minimization algorithm. For technical details, the reader is referred to [7] (Section 14.3 (in particular, Theorem 14.3.39) and Section 14.5). ■

We are now ready to present the Fujishige-type decomposition algorithm for computing a dec-min element of $\overset{\dots}{B}$. We assume that B is given by an integer-valued supermodular function p as $B = B'(p)$ in (2.3).

Fujishige-type decomposition algorithm for square-sum on an M-convex set $\overset{\dots}{B}$

- 1: Set $a := \lfloor p(S)/|S| \rfloor$.
- 2: Find $z \in \overset{\dots}{B}$ that satisfies (5.8) (by the method in the proof of Proposition 5.5).
- 3: Define S_+, S_-, S_0 by

$$S_+ := \bigcup_{t \in S: z(t) \geq a+2} \{s \in S \mid z + \chi_s - \chi_t \in \overset{\dots}{B}\}, \quad (5.11)$$

$$S_- := \bigcup_{s \in S: z(s) \leq a-1} \{t \in S \mid z + \chi_s - \chi_t \in \overset{\dots}{B}\}, \quad (5.12)$$

$$S_0 := S - (S_+ \cup S_-), \quad (5.13)$$

and let $z^*(s) := z(s)$ for each $s \in S_0$.

- 4: If $S_+ \neq \emptyset$, then apply the present algorithm recursively to the restriction to S_+ .
If $S_- \neq \emptyset$, then apply the present algorithm recursively to the contraction to S_- .
- 5: If both S_+ and S_- are empty, then stop. ■

When the algorithm terminates in Step 5, the component values $z^*(s)$ have already been defined for all $s \in S$. The computed vector z^* is a dec-min element of $\overset{\dots}{B}$, which we show in Proposition 5.7. Step 3 of the algorithm defines the component values $z^*(s)$ for $s \in S_0$ (the subset S_0 may be an empty), whereas the component values on S_+ and S_- are determined within the recursive calls in Step 4. The recursive call to the restriction to S_+ means applying

the above algorithm to $B_+ = B'(p_+)$ on S_+ , where p_+ is the supermodular function on S_+ defined by $p_+(X) = p(X)$ for $X \subseteq S_+$. The recursive call to the contraction to S_- means applying the above algorithm to $B_- = B'(p_-)$ on S_- , where p_- is the supermodular function on S_- defined by $p_-(X) = p(X \cup (S - S_-)) - p(S - S_-)$ for $X \subseteq S_-$.

The following proposition implies the termination of the algorithm with at most $|S|$ recursive calls.

Proposition 5.6. $S_+ \neq S$ and $S_- \neq S$.

Proof. (This is an adaptation of the argument of [14, Section 8.2] to the discrete case.) To show $S_+ \neq S$ by contradiction, assume $S_+ = S$. By (5.11), this implies that for every $s \in S$, there exists $t \in S$ such that $z(t) \geq a + 2$ and $z + \chi_s - \chi_t \in \overset{\dots}{\bar{B}}$. Combining this with (5.8) we have $z(s) \geq a + 1$ for all $s \in S$, which implies $p(S) = \sum[z(s) : s \in S] \geq (a + 1)|S|$. We cannot have equality here, because, by (5.11), our assumption $S_+ = S$ implies $z(t_0) \geq a + 2$ for some $t_0 \in S$. Therefore, $p(S) > (a + 1)|S|$, which is a contradiction to $a = \lfloor p(S)/|S| \rfloor$.

To show $S_- \neq S$ by contradiction, assume $S_- = S$. By (5.12), this implies that for every $t \in S$, there exists $s \in S$ such that $z(s) \leq a - 1$ and $z + \chi_s - \chi_t \in \overset{\dots}{\bar{B}}$. Combining this with (5.8) we have $z(t) \leq a$ for all $t \in S$, which implies $p(S) = \sum[z(t) : t \in S] \leq a|S|$. We cannot have equality here, because, by (5.12), our assumption $S_- = S$ implies $z(s_0) \leq a - 1$ for some $s_0 \in S$. Therefore, $p(S) < a|S|$, which is a contradiction to $a = \lfloor p(S)/|S| \rfloor$. ■

The correctness of the algorithm is established in the following proposition.

Proposition 5.7. The output z^* of the Fujishige-type decomposition algorithm is a minimizer of square-sum $W(x)$ over $\overset{\dots}{\bar{B}}$, or equivalently, z^* is a dec-min element of $\overset{\dots}{\bar{B}}$.

Proof. (This is an adaptation of the argument in [14, Section 8.2] to the discrete case.) We prove that the output of the algorithm is a square-sum minimizer by induction on the size $|S|$ of the ground-set S . We have $|S_+| < |S|$ and $|S_-| < |S|$ by Proposition 5.6. By the definition in Step 3 and the property (5.8) of z , we have

$$\{s : z(s) \geq a + 2\} \subseteq S_+ \subseteq \{s : z(s) \geq a + 1\}, \quad (5.14)$$

$$\{s : z(s) \leq a - 1\} \subseteq S_- \subseteq \{s : z(s) \leq a\}, \quad (5.15)$$

which imply that S_+ and S_- are disjoint and

$$S_0 \subseteq \{s : a \leq z(s) \leq a + 1\}. \quad (5.16)$$

First we show that the output z^* of the algorithm satisfies

$$z^*(s) \begin{cases} \geq a + 1 & (s \in S_+), \\ \leq a & (s \in S_-), \\ \in \{a, a + 1\} & (s \in S_0). \end{cases} \quad (5.17)$$

To prove $z^*(s) \geq a + 1$ for $s \in S_+$, let z_+^* denote the subvector of z^* on S_+ , which is the outcome of the recursive call to the restriction to S_+ . By the induction hypothesis, z_+^* minimizes the square-sum $\sum[x(s)^2 : s \in S_+]$ over the M-convex set $\overset{\dots}{\bar{B}}_+$, where $B_+ = B'(p_+)$ is the restriction to S_+ . Then Theorems 2.3 and 2.7 show that z_+^* is an inc-max element of $\overset{\dots}{\bar{B}}_+$, whereas the subvector of z on S_+ is a member of $\overset{\dots}{\bar{B}}_+$ satisfying $z(s) \geq a + 1$ for all $s \in S_+$. Therefore, $z^*(s) \geq a + 1$ for $s \in S_+$. The second case, $z^*(s) \leq a$ for $s \in S_-$, can be proved symmetrically.

Namely, by the induction hypothesis, the subvector z_-^* of z^* on S_- minimizes the square-sum over the M -convex set $\overset{\dots}{B}_-$ defined by the contraction $B_- = B'(p_-)$ to S_- , and hence z_-^* is a dec-min element of $\overset{\dots}{B}_-$ by Theorem 2.7. This implies $z^*(s) \leq a$ for $s \in S_-$, since the subvector of z on S_- is a member of $\overset{\dots}{B}_-$ satisfying $z(s) \leq a$ for all $s \in S_-$. The third case, $z^*(s) \in \{a, a+1\}$ for $s \in S_0$, is obvious from (5.16) and the definition $z^*(s) = z(s)$ for $s \in S_0$ in Step 3.

We will show that z^* satisfies the condition (5.7) for all (s, t) with $s, t \in S_+ \cup S_0 \cup S_-$. Suppose that $z^*(t) \geq z^*(s) + 2$. By (5.17) we may assume $(s, t) \notin (S_+ \cup S_0) \times (S_- \cup S_0)$.

The condition (5.7) holds when $(s, t) \in S_+ \times S_+$, since the subvector z_+^* is an inc-max (and hence dec-min) element of $\overset{\dots}{B}_+$, as already mentioned. Similarly, (5.7) holds when $(s, t) \in S_- \times S_-$, since the subvector z_-^* is a dec-min element of $\overset{\dots}{B}_-$.

It remains to consider the other three cases: $(s, t) \in S_- \times S_+$, $(s, t) \in S_0 \times S_+$, and $(s, t) \in S_- \times S_0$. Since z_+^* is a base of the restriction B_+ , we have $\widetilde{z}^*(S_+) = p(S_+)$, that is, S_+ is z^* -tight with respect to p . It then follows that $z^* + \chi_s - \chi_t \notin \overset{\dots}{B}$ if $s \in S - S_+$ and $t \in S_+$. Hence, the condition (5.7) holds for $(s, t) \in (S - S_+) \times S_+ = (S_- \cup S_0) \times S_+$. Similarly, since z_-^* is a base of the contraction B_- , we have $\widetilde{z}^*(S_-) = p_-(S_-) = p(S) - p(S - S_-) = \widetilde{z}^*(S) - p(S - S_-)$, which shows that $S - S_-$ is z^* -tight with respect to p . Hence $z^* + \chi_s - \chi_t \notin \overset{\dots}{B}$ if $s \in S_-$ and $t \in S - S_-$. Hence, the condition (5.7) holds for $(s, t) \in S_- \times (S - S_-) = S_- \times (S_+ \cup S_0)$. This completes the proof of Proposition 5.7. ■

The Fujishige-type decomposition algorithm can be executed in strongly polynomial time. The key fact here is that we can carry out Step 2, which is characteristic of the discrete case, in polynomial time.

Proposition 5.8. *The Fujishige-type decomposition algorithm computes a square-sum minimizer (i.e., dec-min element) of an M -convex set in strongly polynomial time.*

Proof. By Proposition 5.6, the number of recursive calls is bounded by $|S|$. In each call of the algorithm, the vector z in Step 2 can be found in strongly polynomial time by Proposition 5.5. The subsets S_+ and S_- in Step 3 can be determined in strongly polynomial time by a standard method using submodular function minimization [14, 35]. ■

Remark 5.1. The subsets S_+ and S_- constructed in the Fujishige-type decomposition algorithm have crucial properties that each of S_+ and $S - S_-$ is a z^* -top and z^* -tight set (with respect to p), where the former property is obvious from (5.17) and the latter property is shown in the proof of Proposition 5.7. In addition, z^* is near-uniform on their difference $S_0 = (S - S_-) - S_+$, where S_0 may possibly be empty. It follows from these properties that an integral vector $z = (z_+, z_0, z_-)$ is a dec-min element of $\overset{\dots}{B}$ if and only if z_+ is a dec-min element of $\overset{\dots}{B}_+$, z_- is a dec-min element of $\overset{\dots}{B}_-$, and z_0 is a near-uniform element of $\overset{\dots}{B}_0$, where $B_0 = B'(p_0)$ with $p_0(X) = p(X \cup S_+) - p(S_+)$ for $X \subseteq S_0$. This fact justifies the recursive calls in Step 4. ■

Remark 5.2. Condition (C) in Theorem 2.3 refers to a chain $C_1 \subset C_2 \subset \dots \subset C_\ell$ to characterize a dec-min element. The subsets S_+ and S_- constructed in the algorithm correspond to this chain as follows. First note that $S_+ \subseteq S - S_-$, where equality may occur. As mentioned in Remark 5.1, both S_+ and $S - S_-$ are z^* -top and z^* -tight sets, and moreover, z^* is near-uniform on their difference $S_0 = (S - S_-) - S_+$. Through the recursive calls to S_+ and S_- , the algorithm constructs, in effect, the chain $C_1 \subset C_2 \subset \dots \subset C_\ell$. More precisely, if S_+ and $S - S_-$ are distinct, they are consecutive members of the chain; otherwise, $S_+ (= S - S_-)$ is a member of the chain. ■

Remark 5.3. The Fujishige-type decomposition algorithm may not find the canonical chain, in spite of the fact explained in Remark 5.2. In Example 5.1, for instance, we have $S = \{s_1, s_2\}$ and $a = \lfloor p(S)/|S| \rfloor = \lfloor 5/2 \rfloor = 2$. The element of $\overset{\dots}{B}$ satisfying (5.8) in Step 2 is given (uniquely) by $z = (3, 2)$, for which $S_+ = \emptyset$, $S_- = \emptyset$, and $S_0 = S$ in Step 3. Thus the ground-set S is not decomposed at all, which corresponds to a trivial chain consisting of a single member S . In contrast, as we have seen in Example 5.2, the canonical partition is a bipartition $\{S_1, S_2\}$ with $S_1 = \{s_1\}$ and $S_2 = \{s_2\}$, which corresponds to the canonical chain: $\{s_1\} \subset \{s_1, s_2\}$. ■

Remark 5.4. We explain here how our Fujishige-type decomposition algorithm is derived from the framework in [14, Section 8.2] with additional integrality considerations. We apply the framework to the piecewise-linear function $\overline{W}(x) = \sum [\overline{\varphi}(x(s)) : s \in S]$ defined in (4.3), where $\overline{\varphi} : \mathbf{R} \rightarrow \mathbf{R}$ is the piecewise-linear extension (4.2) of $\varphi(k) = k^2$ ($k \in \mathbf{Z}$).

The notation of [14, Section 8.2] is as follows. The ground-set is E , and a base-polyhedron B is described by a submodular function f . For $x \in B$, the smallest x -tight set (with respect to f) containing $e \in E$ is denoted as $\text{dep}(x, e)$. The objective function to be minimized on B is a separable convex function $\sum_{e \in E} w_e(x(e))$, where each $w_e : \mathbf{R} \rightarrow \mathbf{R}$ is a real-valued convex function on \mathbf{R} . The left and right derivatives of w_e at $\xi \in \mathbf{R}$ are denoted, respectively, by $w_e^-(\xi)$ and $w_e^+(\xi)$. For any $\eta \in \mathbf{R}$, $[i_e^-(\eta), i_e^+(\eta)]$ denotes the set (interval) of minimizers ξ of function $w_e(\xi) - \eta\xi$. In our problem to minimize \overline{W} , we have

$$w_e(\xi) = \overline{\varphi}(\xi), \quad i_e^-(\eta) = \left\lfloor \frac{\eta - 1}{2} \right\rfloor, \quad i_e^+(\eta) = \left\lfloor \frac{\eta + 1}{2} \right\rfloor \quad (5.18)$$

for all $e \in E$.

Step 1 in [14, page 258] says:

$$\text{Choose } \eta \in \mathbf{R} \text{ such that } \sum_{e \in E} i_e^-(\eta) \leq f(E) \leq \sum_{e \in E} i_e^+(\eta).$$

It follows from (5.18) that the choice of $\eta = 2a + 1$ with $a = \lfloor f(E)/|E| \rfloor$ satisfies this condition; then $i_e^-(\eta) = a$ and $i_e^+(\eta) = a + 1$. This explains our Step 1 to set $a := \lfloor p(S)/|S| \rfloor$.

Step 2 in [14, page 258] is a tricky step where the substantial condition (5.8) is hidden. This step requires to find a base $x \in B$ such that, for each $s, t \in E$,

1. if $w_s^+(x(s)) < \eta$ and $w_t^-(x(t)) > \eta$, then we have $t \notin \text{dep}(x, s)$,
2. if $w_s^+(x(s)) < \eta$ and $w_t^-(x(t)) = \eta$, and $t \in \text{dep}(x, s)$, then for any $\alpha > 0$ we have $w_t^-(x(t) - \alpha) < \eta$, i.e., $x(t) = i_t^-(\eta)$,
3. if $w_s^+(x(s)) = \eta$ and $w_t^-(x(t)) > \eta$, and $t \in \text{dep}(x, s)$, then for any $\alpha > 0$ we have $w_s^+(x(s) + \alpha) > \eta$, i.e., $x(s) = i_s^+(\eta)$.

In our case, B is an integral base-polyhedron, $x \in \overset{\dots}{B}$, and $w_s(\xi) = \overline{\varphi}(\xi)$, for which $w_s^+(\xi) = 2\xi + 1$ and $w_s^-(\xi) = 2\xi - 1$ for integer ξ . Under the integrality requirement, the first condition is replaced by

$$x(s) < a, \quad x(t) > a + 1 \implies x + \chi_s - \chi_t \notin \overset{\dots}{B}. \quad (5.19)$$

A literal translation of the second condition results in the condition

$$x(s) < a, \quad x(t) = a + 1, \quad x + \chi_s - \chi_t \in \overset{\dots}{B} \implies x(t) = a,$$

which should be interpreted as

$$x(s) < a, \quad x(t) = a + 1 \implies x + \chi_s - \chi_t \notin \overset{\dots}{B}. \quad (5.20)$$

Similarly, the third condition is replaced by

$$x(s) = a, \quad x(t) > a + 1 \implies x + \chi_s - \chi_t \notin \overset{\dots}{B}. \quad (5.21)$$

The combination of (5.19)–(5.21) is equivalent to

$$x(s) \leq a, \quad x(t) \geq a + 1, \quad x(t) \geq x(s) + 2 \implies x + \chi_s - \chi_t \notin \overset{\dots}{B},$$

which coincides with the condition (5.8) in our Step 2.

The remaining steps of our algorithm is a straightforward translation of the corresponding steps in [14, page 258] with obvious integrality requirements. ■

Acknowledgement We thank Satoru Iwata and Akiyoshi Shioura for discussion about algorithms, and Arie Tamir for indicating references. We also thank the anonymous referee for helpful comments. Insightful questions posed by Tamás Király led to Theorem 4.2. The research was partially supported by the National Research, Development and Innovation Fund of Hungary (FK_18) – No. NKFI-128673, and by JSPS KAKENHI Grant Number JP20K11697.

References

- [1] Arnold, B.C., Sarabia, J.M.: Majorization and the Lorenz Order with Applications in Applied Mathematics and Economics. Springer International Publishing, Cham (2018), (1st edn., 1987)
- [2] Chakrabarty, D., Jain, P., Kothari, P.: Provable submodular minimization via Fujishige-Wolfe’s algorithm. *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, 802–809 (2014)
- [3] De Loera, J.A., Haddock, J., Rademacher, L.: The minimum Euclidean-norm point in a convex polytope: Wolfe’s combinatorial algorithm is exponential. *SIAM Journal on Computing* **49**, 138–169 (2020)
- [4] Dutta, B.: The egalitarian solution and reduced game properties in convex games. *International Journal of Game Theory* **19**, 153–169 (1990)
- [5] Dutta, B., Ray, D.: A concept of egalitarianism under participation constraints. *Econometrica* **57**, 615–635 (1989)
- [6] Edmonds, J.: Submodular functions, matroids and certain polyhedra. In: Guy, R., Hanani, H., Sauer, N., Schönheim, J. (eds.) *Combinatorial Structures and Their Applications*, pp. 69–87. Gordon and Breach, New York (1970)
- [7] Frank, A.: *Connections in Combinatorial Optimization*. Oxford University Press, Oxford (2011)
- [8] Frank, A., Murota, K.: Discrete decreasing minimization, Part II: Views from discrete convex analysis. arXiv: 1808.08477 (August 2018)

- [9] Frank, A., Murota, K.: Decreasing minimization on M-convex sets: Background and structures. *Mathematical Programming*, published online (October 27, 2021)
<https://doi.org/10.1007/s10107-021-01722-2>
- [10] Frank, A., Murota, K.: Decreasing minimization on M-convex sets: Algorithms and applications. *Mathematical Programming*, published online (October 15, 2021)
<https://doi.org/10.1007/s10107-021-01711-5>
- [11] Frank, A., Murota, K.: Fair integral network flows. Submitted for publication; arXiv: 1907.02673v4 (January 2022)
- [12] Frank, A., Murota, K.: Fair integral submodular flows. Submitted for publication; arXiv: 2012.07325 (December 2020)
- [13] Fujishige, S.: Lexicographically optimal base of a polymatroid with respect to a weight vector. *Mathematics of Operations Research* **5**, 186–196 (1980)
- [14] Fujishige, S.: *Submodular Functions and Optimization*, 1st edn. North-Holland, Amsterdam (1991); 2nd edn. Elsevier, Amsterdam (2005)
- [15] Fujishige, S.: Theory of principal partitions revisited. In: Cook, W., Lovász, L., Vygen, J. (eds.) *Research Trends in Combinatorial Optimization*, pp. 127–162. Springer, Berlin (2009)
- [16] Fujishige, S., Isotani, S.: A submodular function minimization algorithm based on the minimum-norm base. *Pacific Journal of Optimization* **7**, 3–15 (2011)
- [17] Goemans, M.X., Gupta, S., Jaillet, P.: Discrete Newton’s algorithm for parametric submodular function minimization. In: Eisenbrand, F., Koenemann, J. (eds.) *Integer Programming and Combinatorial Optimization. Lecture Notes in Computer Science*, vol. 10328, pp. 212–227 (2017)
- [18] Groenevelt, H.: Two algorithms for maximizing a separable concave function over a polymatroid feasible region. *European Journal of Operational Research* **54**, 227–236 (1991); The technical report version appeared as Working Paper Series No. QM 8532, Graduate School of Management, University of Rochester (1985)
- [19] Harvey, N.J.A., Ladner, R.E., Lovász, L., Tamir, T.: Semi-matchings for bipartite graphs and load balancing. *Journal of Algorithms* **59**, 53–78 (2006)
- [20] Hochbaum, D.S.: Lower and upper bounds for the allocation problem and other nonlinear optimization problems. *Mathematics of Operations Research* **19**, 390–409 (1994)
- [21] Hochbaum, D.S.: Complexity and algorithms for nonlinear optimization problems. *Annals of Operations Research* **153**, 257–296 (2007)
- [22] Hochbaum, D.S., Hong, S.-P.: About strongly polynomial time algorithms for quadratic optimization over submodular constraints. *Mathematical Programming* **69**, 269–309 (1995)
- [23] Ibaraki, T., Katoh, N.: *Resource Allocation Problems: Algorithmic Approaches*. MIT Press, Boston (1988)
- [24] Iri, M. : A review of recent work in Japan on principal partitions of matroids and their applications. *Annals of the New York Academy of Sciences* **319**, 306–319 (1979)
- [25] Iwata, S., Murota, K., Shigeno, M.: A fast parametric submodular intersection algorithm for strong map sequences. *Mathematics of Operations Research* **22**, 803–813 (1997)

- [26] Katoh, N., Shioura, A., Ibaraki, T.: Resource allocation problems. In: Pardalos, P.M., Du, D.-Z., Graham, R.L. (eds.) *Handbook of Combinatorial Optimization*, 2nd ed., Vol. 5, pp. 2897-2988, Springer, Berlin (2013)
- [27] Marshall, A.W., Olkin, I., Arnold, B.C.: *Inequalities: Theory of Majorization and Its Applications*, 2nd edn. Springer, New York (2011)
- [28] Megiddo, N.: Optimal flows in networks with multiple sources and sinks. *Mathematical Programming* **7**, 97–107 (1974)
- [29] Moriguchi, S., Shioura, A., Tsuchimura, N.: M-convex function minimization by continuous relaxation approach—Proximity theorem and algorithm. *SIAM Journal on Optimization* **21**, 633–668 (2011)
- [30] Murota, K.: Discrete convex analysis. *Mathematical Programming* **83**, 313–371 (1998)
- [31] Murota, K.: *Discrete Convex Analysis*. Society for Industrial and Applied Mathematics, Philadelphia (2003)
- [32] Murota, K.: Recent developments in discrete convex analysis. In: Cook, W., Lovász, L., Vygen, J. (eds.) *Research Trends in Combinatorial Optimization*, Chapter 11, pp. 219–260. Springer, Berlin (2009)
- [33] Murota, K.: On basic operations related to network induction of discrete convex functions. *Optimization Methods and Software* **36**, 519–559 (2021)
- [34] Radzik, T.: Fractional combinatorial optimization. In: Pardalos, P.M., Du, D.-Z., Graham, R.L. (eds.) *Handbook of Combinatorial Optimization*, 2nd edn., pp. 1311–1355. Springer Science+Business Media, New York (2013)
- [35] Schrijver, A.: *Combinatorial Optimization—Polyhedra and Efficiency*. Springer, Heidelberg (2003)
- [36] Tamir, A.: A note on the polynomial solvability of the resource allocation problem over polymatroids. Technical Report, School of Mathematical Sciences, Tel Aviv University (March, 1993)
- [37] Végh, L.A.: A strongly polynomial algorithm for a class of minimum-cost flow problems with separable convex objectives. *SIAM Journal on Computing* **45**, 1729–1761 (2016)
- [38] Veinott, Jr., A.F.: Least d -majorized network flows with inventory and statistical applications. *Management Science* **17**, 547–567 (1971)
- [39] Wolfe, P.: Finding the nearest point in a polytope. *Mathematical Programming* **11**, 128–149 (1976)