Geometric Study on the Split Decomposition of Finite Metrics *

Hiroshi HIRAI
Research Institute for Mathematical Sciences,
Kyoto University, Kyoto 606-8502, Japan
hirai@kurims.kyoto-u.ac.jp

May 2004

Abstract

This paper sheds a new light on the split decomposition theory and T-theory from the viewpoint of convex analysis and polyhedral geometry. By regarding finite metrics as discrete concave function, Bandelt-Dress' split decomposition can be derived as a special case of more general decomposition of polyhedral/discrete concave functions introduced in this paper. It is shown that the combinatorics of splits discussed in connection to the split decomposition corresponds to the geometric properties of a hyperplane arrangement and a point configuration. By our approach, the split decomposition of metrics can be naturally extended for distance functions, which may violate the triangle inequality, using partial split distances.

1 Introduction

Recently, theories of *finite metric spaces* have come to be increasingly important in the area of bioinfomatics and phylogenetics; see [2],[22]. The central problem in phylogenetics is reconstructing phylogenetic trees from given experimental data, e.g., DNA sequences. If the data is given as a distance matrix expressing dissimilarity between species, the problem is to search for a tree metric that "fits" the given distance matrix.

In particular, T-theory [10], developed by A. Dress and coworkers, provides a beautiful mathematical framework for this phylogenetic problem. The central concept of T-theory is the *tight span* of a metric space, which was originally constructed by Isbell in [16] and rediscovered by Dress in [8]. For a finite metric space (V, d), the tight span $T(d)$ is a polyhedral subset of $\mathbf{R}^V$ defined as

$$T(d) = \{p \in \mathbf{R}^V \mid \forall i \in V,\ p(i) = \max_{j \in V}\{d(i,j) - p(j)\}\}. \tag{1.1}$$

The tight span $T(d)$ expresses combinatorial properties of (V, d) in geometric terms. For example, a metric is a tree metric if and only if its tight span is a tree [8]. The *split decomposition*, due to Bandelt and Dress [1], is a phylogenetic tree reconstruction method closely related to the polyhedral structure of the tight span.

The split decomposition theory [1] may be summarized as follows. A *split* of V is a bipartition of V. For each split $S = \{A, B\}$, $A, B \subseteq V$, the *split metric* $\delta_S : V \times V \to \mathbf{R}$

*Major part of this work was done while the author was at Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo.

is defined by

$$\delta_S(i,j) = \begin{cases} 0 & \text{if } i,j \in A \text{ or } i,j \in B, \\ 1 & \text{otherwise} \end{cases} \qquad (i,j \in V)$$

The *isolation index* α_S^d is a nonnegative number that represents a quotient of d by δ_S; see Section 4 for the precise definition. The collection of splits S with $\alpha_S^d > 0$ is denoted by $\mathcal{S}(d)$, is endowed with an interesting property, *weak compatibility*, which is explained in Section 4. Then a metric d can be decomposed as

$$d = \sum_{S \in \mathcal{S}(d)} \alpha_S^d \delta_S + d', \tag{1.2}$$

where d' is a metric, called *split-prime residue*, with $\alpha_{S'}^{d'} = 0$ for each split S'. This decomposition is called the split decomposition of d. Furthermore the split decomposition is *coherent*, i.e., it satisfies

$$P(d) = \sum_{S \in \mathcal{S}(d)} \alpha_S^d P(\delta_S) + P(d'), \tag{1.3}$$

where $P(d)$ denotes the polyhedron $\{p \in \mathbf{R}^V \mid p(i) + p(j) \geq d(i,j) \ (i,j \in V)\}$ associated with metric d and the summation means the Minkowski sum. The tight span $T(d)$ equals the union of bounded faces of $P(d)$ [9]. This indicates that the split decomposition is closely related to the polyhedral structure of the tight span.

One of the main aims of this paper is to derive the split decomposition in a natural way as a special case of a decomposition of (discrete) convex/concave functions that we propose in this paper. Thus, the view point of convex analysis gives a new light on such important concepts of split decomposition and T-theory as tight span, isolation index, weakly compatible splits, and coherent decomposition. We begin with the following observation to indicate the connection to concave functions.

Let $\Lambda \subseteq \mathbf{R}^V$ be a finite set of points defined by

$$\Lambda = \{\chi_i + \chi_j \mid i,j \in V\}, \tag{1.4}$$

where $\chi_i \in \mathbf{R}^V$ is the characteristic (unit) vector of $i \in V$. A metric $d : V \times V \to \mathbf{R}$ is naturally regarded as a function on Λ by the correspondence

$$d(\chi_i + \chi_j) \leftarrow d(i,j) \quad (i,j \in V). \tag{1.5}$$

The function $d : \Lambda \to \mathbf{R}$ has a concave-like structure. Considering the concave closure of d, we obtain a polyhedral concave function corresponding to metric d.

The observation above motivates us to generalize the split decomposition to polyhedral convex functions. Corresponding to split metrics, we define a *split function* $l_H : \mathbf{R}^n \to \mathbf{R}$ associated with a hyperplane $H = \{x \in \mathbf{R}^n \mid \langle a, x\rangle = r\}$ with $\|a\| = 1$ as

$$l_H(x) = |\langle a, x\rangle - r|/2 \quad (x \in \mathbf{R}^n).$$

For a polyhedral convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$, the quotient $c_H(f)$ of f by l_H is defined as the supremum of $t \geq 0$ such that $f - tl_H$ is convex. Let $\mathcal{H}(f)$ be the family of hyperplanes H with $0 < c_H(f) < \infty$.

Our guiding principle in deriving the split decomposition of metrics and its related concepts by means of convex analysis and polyhedral geometry is the following:

A polyhedral convex function f can be decomposed as

$$f = \sum_{H \in \mathcal{H}(f)} c_H(f) l_H + f', \tag{1.6}$$

where f' is a polyhedral convex function with $c_{H'}(f') \in \{0, \infty\}$ for any hyperplane H'. Furthermore this representation is unique.

We call (1.6) the *polyhedral split decomposition*. This is also a refinement of the classical result of Bolker [3] that every polytope has the maximum zonotopic summand. The polyhedral split decomposition is closely related to a polyhedral subdivision induced by polyhedral convex functions. In Section 2, we first discuss polyhedral subdivisions in terms of convex analysis and then derive the polyhedral split decomposition (1.6).

In our approach, metrics are regarded as functions defined on a finite set of points Λ. Therefore we need to discretize the polyhedral split decomposition for discrete functions before we can derive Bandelt-Dress' split decomposition (1.2) from (1.6). Applications of convex analysis to combinatorial and discrete structures are made successfully in the theory of submodular functions by Lovász [18], Frank [11], and Fujishige [12] (see also [13]). Recently Murota [19] developed *discrete convex analysis*, a convex analysis for functions defined on integer lattice points. Hirai and Murota [15] discuss the relationship between tree metrics and discrete convexity. In Section 3, we also follow this line and discuss discrete functions and their convex-extensions. Then we derive the split decomposition of a *discrete convex function* by discretizing the polyhedral split decomposition (1.6).

In Section 4, we regard a metric d as a discrete concave function on Λ by (1.5) and apply the results of Sections 2 and 3. We then obtain the following:

- By discretizing the polyhedral split decomposition (1.6), Bandelt-Dress' split decomposition (1.2) can be derived; see Proposition 4.9 and Theorem 4.10.
- The split decomposition of metrics can be naturally extended for distance functions, which may possibly violate the triangle inequality, using *partial split distances*; see Theorem 4.10, (4.11), and (4.25).
- Weak compatibility of splits can be translated into a geometric property of a hyperplane arrangement and a point configuration; see Theorem 4.15 and Proposition 4.17.

2 Split decomposition of polyhedral convex functions

In this section, we derive the polyhedral split decomposition (1.6), which is the basis for subsequent developments in this paper.

2.1 Basic notation

First we introduce some basic notation. Let $\mathbf{R}$, $\mathbf{R}_+$, $\mathbf{R}_{++}$ be the sets of real numbers, nonnegative real numbers, and positive real numbers, respectively. Let $\mathbf{R}^n$ be the n dimensional Euclid space with the standard inner product $\langle \cdot, \cdot \rangle$. For $x, y \in \mathbf{R}^n$, let $[x, y]$ denotes the closed line segment between x and y. A set $S \subseteq \mathbf{R}^n$ is said to be convex if $[x, y] \subseteq S$ for every $x, y \in S$. For $X \subseteq \mathbf{R}^n$, we denote by $\operatorname{conv} X$, $\operatorname{cone} X$, and $\operatorname{aff} X$, the convex hull, the conical hull, and the affine hull of X, respectively. We refer to an $(n-1)$ dimensional affine subspace of $\mathbf{R}^n$ as a hyperplane. In particular, for $(a, r) \in \mathbf{R}^n \times \mathbf{R}$, we define a hyperplane $H_{a,r}$ by $\{x \in \mathbf{R}^n \mid \langle a, x \rangle = r\}$, closed half spaces $H^-_{a,r}$ and $H^+_{a,r}$

by $\{x \in \mathbf{R}^n \mid \langle a, x\rangle \leq r\}$ and $\{x \in \mathbf{R}^n \mid \langle a, x\rangle \geq r\}$, and open half spaces $H^{--}_{a,r}$ and $H^{++}_{a,r}$ by $\{x \in \mathbf{R}^n \mid \langle a, x\rangle < r\}$ and $\{x \in \mathbf{R}^n \mid \langle a, x\rangle > r\}$. A set $P \subseteq \mathbf{R}^n$ is said to a polyhedron if P is represented as an intersection of finitely many closed half spaces. For $S \subseteq \mathbf{R}^n$, we denote by $\mathrm{int}\, S$ and $\mathrm{ri}\, S$ the sets of interior points and relative interior points of S, respectively.

Second we prepare some basic terms and notation from convex analysis; see [20] as a standard reference. A function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is said to be convex if it satisfies $\lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda) y)$ for all $x, y \in \mathbf{R}^n$, $\lambda \in [0,1]$. For a function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$, we define $\mathrm{dom}\, f = \{x \in \mathbf{R}^n \mid f(x) < +\infty\}$, which is the effective domain of f, and $\mathrm{epi}\, f = \{(x, z) \in \mathbf{R}^n \times \mathbf{R} \mid f(x) \leq z\}$, which is the epigraph of f. The subdifferential of a function f at a point $x \in \mathrm{dom}\, f$ is defined to be the set

$$\partial f(x) = \{p \in \mathbf{R}^n \mid f(y) - f(x) \geq \langle p, y - x\rangle \ (\forall y \in \mathbf{R}^n)\}.$$

The directional derivative of f at $x \in \mathrm{dom}\, f$ in a direction $d \in \mathbf{R}^n$ is defined by

$$f'(x; d) = \lim_{t \searrow 0} \{f(x + td) - f(x)\}/t.$$

The indicator function of a set $S \subseteq \mathbf{R}^n$ is a function $\delta_S : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ defined by

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

The conjugate of a function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$, where $\mathrm{dom}\, f \neq \emptyset$ is assumed, is a function $f^\bullet : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ defined by

$$f^\bullet(p) = \sup_{x \in \mathbf{R}^n} \{\langle p, x\rangle - f(x)\} \quad (p \in \mathbf{R}^n).$$

For a function f and a vector p, $f[-p]$ denotes the function defined by

$$f[-p](x) = f(x) - \langle p, x\rangle \quad (x \in \mathbf{R}^n).$$

The following is fundamental.

Lemma 2.1. *For a convex function f and a vector $p \in \mathbf{R}^n$ with* $\mathrm{argmin}\, f[-p] \neq \emptyset$*, we have:* $x \in \mathrm{argmin}\, f[-p] \iff p \in \partial f(x) \iff x \in \partial f^\bullet(p)$.

A convex function f is said to be polyhedral if its epigraph $\mathrm{epi}\, f$ is a polyhedron. A polyhedral convex function f is represented as

$$f(x) = \max_{i \in I} \{\langle p_i, x\rangle - \beta_i\} + \sum_{j \in J} \delta_{H^-_{a_j, b_j}}(x) \quad (x \in \mathbf{R}^n), \tag{2.1}$$

where $\{(p_i, \beta_i) \mid i \in I\}$ and $\{(a_j, b_j) \mid j \in J\}$ are finite subsets of $\mathbf{R}^n \times \mathbf{R}$. A conjugate function $f^\bullet$ of a polyhedral convex function f is also polyhedral and $f^{\bullet\bullet} = f$ holds. Furthermore if f is positively homogeneous, i.e., $f(\lambda x) = \lambda f(x)$ holds for $\lambda \geq 0$ and $x \in \mathbf{R}^n$, then $f^\bullet = \delta_{\partial f(0)}$ holds and hence $f = (\delta_{\partial f(0)})^\bullet$ is the support function of a polyhedron $\partial f(0)$. We give some fundamental properties of polyhedral convex functions in the following lemmas, where f and g are polyhedral convex functions.

Lemma 2.2. *The subdifferential of f of* (2.1) *is given by*

$$\begin{aligned} \partial f(x) \quad = \quad & \mathrm{conv}\{p_i \mid i \in I, f(x) = \langle p_i, x\rangle - \beta_i\} \\ & + \mathrm{cone}\{a_j \mid j \in J, x \in H_{a_j, b_j}\} \quad (x \in \mathrm{dom}\, f). \end{aligned}$$

Lemma 2.3. *For $x \in \mathrm{dom}\, f$ and $d \in \mathbf{R}^n$, we have*

$$f'(x; d) = \sup\{\langle p, d\rangle \mid p \in \partial f(x)\}.$$

Lemma 2.4. *For $x \in \mathrm{dom}\, f \cap \mathrm{dom}\, g$ and $\alpha, \beta \geq 0$, we have*

$$\partial(\alpha f + \beta g)(x) = \alpha \partial f(x) + \beta \partial g(x).$$

2.2 Polyhedral subdivisions

A polyhedral complex $\mathcal{C}$ is a finite collection of polyhedra such that

(1) if $P \in \mathcal{C}$, all the faces of P are also in $\mathcal{C}$,

(2) the nonempty intersection $P \cap Q$ of two polyhedra $P, Q \in \mathcal{C}$ is a face of P and Q.

The dimension $\dim \mathcal{C}$ is the largest dimension of a polyhedron in $\mathcal{C}$. The underlying set of $\mathcal{C}$ is the point set $|\mathcal{C}| = \bigcup_{P \in \mathcal{C}} P$. A polyhedral subdivision of a polyhedron P is a polyhedral complex $\mathcal{C}$ with $|\mathcal{C}| = P$. A polyhedral subdivision is *pure*, which means that the inclusion maximal elements are of the same dimension.

For a polyhedral convex function f, lower faces of $\operatorname{epi} f$ are bijectively projected on $\operatorname{dom} f$, and determine a polyhedral subdivision of $\operatorname{dom} f$, which is denoted by $\mathcal{T}(f)$. A polyhedral subdivision constructed in this way is said to be *regular* (see [14] [25]). The following is an easy observation.

Lemma 2.5. $\mathcal{T}(f) = \{F \subseteq \mathbf{R}^n \mid \exists p \in \mathbf{R}^n, F = \operatorname{argmin} f[-p]\}$.

Next we investigate the relationship between $\mathcal{T}(f)$ and $\mathcal{T}(f^\bullet)$. For $F \in \mathcal{T}(f)$ and a point $x \in \operatorname{ri} F$, we define $F^\bullet \in \mathcal{T}(f^\bullet)$ as

$$F^\bullet = \partial f(x).$$

In fact, it will be shown that this map $F \mapsto F^\bullet$ is well-defined and establishes a one-to-one correspondence between $\mathcal{T}(f)$ and $\mathcal{T}(f^\bullet)$. The following properties of $\mathcal{T}(f)$ and $\mathcal{T}(f^\bullet)$ are implicitly obtained by Chynoweth [5] and called *generalized Voronoi/Delaunay duality* (see also [6]), which is in fact a polarity between $\operatorname{epi} f^\bullet$ and $\operatorname{epi} f$.

Proposition 2.6. *Suppose that f is a polyhedral convex function and $F, G \in \mathcal{T}(f)$.*

(1) *$F^\bullet$ is determined independently of the choice of $x \in \operatorname{ri} F$.*

(2) *$F^{\bullet\bullet} = F$.*

(3) *$(\operatorname{aff} F - \{x\})^\perp = \operatorname{aff} F^\bullet - \{p\} \quad (x \in F, p \in F^\bullet)$.*

(4) *$F \subseteq G \Leftrightarrow F^\bullet \supseteq G^\bullet$.*

Proof. Since f is represented as (2.1), they can be proved by standard arguments in linear programming (see Appendix). $\square$

For two polyhedral subdivisions $\mathcal{C}_1$ and $\mathcal{C}_2$, the *common refinement* $\mathcal{C}_1 \wedge \mathcal{C}_2$ is defined by $\mathcal{C}_1 \wedge \mathcal{C}_2 = \{F \cap G \mid F \in \mathcal{C}_1, G \in \mathcal{C}_2,\ F \cap G \neq \emptyset\}$. Note that $\mathcal{C}_1 \wedge \mathcal{C}_2$ is a polyhedral subdivision of $|\mathcal{C}_1| \cap |\mathcal{C}_2|$.

Lemma 2.7. *For polyhedral convex functions f and g with $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$, we have*

$$\mathcal{T}(f+g) = \mathcal{T}(f) \wedge \mathcal{T}(g). \tag{2.2}$$

Proof. Both sides of (2.2) are polyhedral subdivisions of $\operatorname{dom} f \cap \operatorname{dom} g$. Hence it is sufficient to show a half inclusion ($\supseteq$). $F \in \mathcal{T}(f)$ and $G \in \mathcal{T}(g)$ with $F \cap G \neq \emptyset$ are represented as $F = \operatorname{argmin} f[-p]$ and $G = \operatorname{argmin} g[-q]$ for some $p, q \in \mathbf{R}^n$ by Lemma 2.5. Then we have $F \cap G = \operatorname{argmin}(f+g)[-p-q]$. Hence $F \cap G \in \mathcal{T}(f+g)$. $\square$

2.3 Split functions

In this paper, we call the support function of a line segment a *split function*.

Definition 2.8. For a hyperplane $H = H_{a,b}$ with $\|a\| = 1$, the split function $l_H : \mathbf{R}^n \to \mathbf{R}$ associated with H is defined as

$$l_H(x) = |\langle a, x\rangle - b|/2 \quad (x \in \mathbf{R}^n). \tag{2.3}$$

By easy applications of Lemma 2.2, the polyhedral subdivision induced by a split function is given as follows.

Proposition 2.9. *Let l_H be the split function associated with a hyperplane $H = H_{a,b}$ with $\|a\| = 1$. The subdifferential of l_H is given by*

$$\partial l_H(x) = \begin{cases} \{a/2\} & \text{if} \quad x \in H^{++} \\ [-a/2, a/2] & \text{if} \quad x \in H \\ \{-a/2\} & \text{if} \quad x \in H^{--} \end{cases}$$

and polyhedral subdivisions $\mathcal{T}(l_H)$ and $\mathcal{T}(l_H^\bullet)$ are given by

$$\begin{aligned} \mathcal{T}(l_H) &= \{H, H^+, H^-\}, \\ \mathcal{T}(l_H^\bullet) &= \{\{a/2\}, \{-a/2\}, [-a/2, a/2]\}. \end{aligned}$$

A finite set of hyperplanes in $\mathbf{R}^n$, say, $\mathcal{H}$ decomposes $\mathbf{R}^n$ into a finite set of polyhedra. We denote this polyhedral subdivision by $\mathcal{A}(\mathcal{H})$; $\mathcal{A}(\mathcal{H})$ is a hyperplane arrangement. A polytope is said to be a *zonotope* if it is represented as the Minkowski sum of a finite number of line segments (see [3],[21]). The conjugacy relationship in convex analysis clarifies the relationship between a hyperplane and a zonotope.

Proposition 2.10. *Let $\mathcal{H} = \{H_{a_1,b_1}, \ldots, H_{a_m,b_m}\}$ be a finite set of hyperplanes with $\|a_j\| = 1$ for $j = 1, \ldots, m$. For $c_1, c_2, \ldots, c_m \in \mathbf{R}^n_{++}$, let $f : \mathbf{R}^n \to \mathbf{R}$ be defined as $f = \sum_{j=1}^m c_j l_{H_{a_j,b_j}}$. Then we have*

(1) *$\mathcal{T}(f) = \mathcal{A}(\mathcal{H})$.*

(2) *$\mathcal{T}(f^\bullet)$ is the projection of lower faces of a zonotope $\sum_{j=1}^m c_j[-(a_j, b_j)/2, (a_j, b_j)/2]$.*

Proof. (1) is immediate from Proposition 2.9 and Lemma 2.7. (2) is obtained from the facts that $\operatorname{epi} l^\bullet_{H_{a,b}} = [-(a, b)/2, (a, b)/2] + \operatorname{cone}(0, 1)$ and that $\operatorname{epi}(f+g)^\bullet = \operatorname{epi} f^\bullet + \operatorname{epi} g^\bullet$ holds for polyhedral convex functions f and g with $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$. □

2.4 Polyhedral split decomposition

Here we derive the split decomposition of polyhedral convex functions, which is our guiding principle in deriving the split decomposition of metrics and its related concepts in terms of convex analysis and polyhedral geometry. Furthermore our derivation leads to an algorithm for the discrete version of polyhedral split decomposition in the next section.

The discussion in Section 2.3 reveals that split functions have a very simple structure. Accordingly, we regard a split function as the most fundamental function in the set of polyhedral convex functions, and consider to decompose a given polyhedral convex function into a sum of split functions and a polyhedral convex function which "contains" no split functions. As preliminaries, we discuss *division* for polyhedral convex functions, and then we derive the polyhedral split decomposition as its special case.

For two polyhedral convex functions $f, g : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$, where $\operatorname{dom} g = \mathbf{R}^n$, we define the quotient $[f : g]$ of f by g as

$$[f : g] = \sup\{t \in \mathbf{R}_+ \mid f - tg \text{ is convex}\}. \tag{2.4}$$

As long as g is not affine over $\operatorname{dom} f$, the supremum is attained by some finite t, and therefore f is decomposed as $f = [f : g]g + r$, where r is a polyhedral convex function with $[r : g] = 0$. The following facts are easy to see.

Lemma 2.11. *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ and $g, h : \mathbf{R}^n \to \mathbf{R}$ be polyhedral convex functions.*

(1) $[f - sg : g] = [f : g] - s \quad (0 \leq s \leq [f : g])$.

(2) $[f + sg : g] \geq s \quad (s \in \mathbf{R}_+)$.

(3) $[f - sg : h] \leq [f : h] \quad (0 \leq s \leq [f : g])$.

(4) $[a_1 f[p_1] : a_2 g[p_2]] = (a_1/a_2)[f : g] \quad (a_1, a_2 \in \mathbf{R}_{++},\ p_1, p_2 \in \mathbf{R}^n)$.

The basic idea for the polyhedral split decomposition (1.6) is dividing a given polyhedral convex function f by the set of split functions successively. For a hyperplane H, we define a nonnegative number $c_H(f)$ as

$$c_H(f) = [f : l_H]. \tag{2.5}$$

We observe the following facts, where H, H_1, and H_2 are hyperplanes:

- $c_H(f) = \infty \Leftrightarrow \operatorname{dom} f \subseteq H^+$ or $\operatorname{dom} f \subseteq H^-$.
- If $0 < c_H(f) < \infty$, then $\{F \in \mathcal{T}(f) \mid F \subseteq H\}$ is a polyhedral subdivision of $H \cap \operatorname{dom} f$.
- If $H_1 \cap \operatorname{int} \operatorname{dom} f \neq \emptyset$ and $H_2 \cap \operatorname{int} \operatorname{dom} f \neq \emptyset$, then $H_1 = H_2 \Leftrightarrow H_1 \cap \operatorname{dom} f = H_2 \cap \operatorname{dom} f$.

By the above observations and the polyhedrality of f, if $\operatorname{int} \operatorname{dom} f \neq \emptyset$, the set of hyperplanes

$$\mathcal{H}(f) = \{H \mid 0 < c_H(f) < \infty\} \tag{2.6}$$

is finite. Hence, we assume the following.

Assumption 2.12. $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is a polyhedral convex function with $\operatorname{int} \operatorname{dom} f \neq \emptyset$.

Put $\mathcal{H}(f) = \{H_1, H_2, \ldots, H_m\}$ and define $f^1, f^2, \ldots, f^m$ by

$$f^k = \begin{cases} f & \text{if} \quad k = 1, \\ f^{k-1} - c_{H_{k-1}}(f^{k-1}) l_{H_{k-1}} & \text{if} \quad k = 2, \ldots, m. \end{cases}$$

Then f is decomposed as

$$f = \sum_{k=1}^{m} c_{H_k}(f^k) l_{H_k} + f^m.$$

By Lemma 2.11 (1),(3), we have $\mathcal{H}(f^{k+1}) \subset \mathcal{H}(f^k)$. Hence $c_H(f^m) \in \{0, \infty\}$ holds for any hyperplane H. In fact, $c_{H_k}(f^k) = c_{H_k}(f)$ holds for $k = 1, \ldots, m$ by the following general fact, which we prove this proposition in Subsection 2.5.

Proposition 2.13. *For $H, H' \in \mathcal{H}(f)$ and $t \in [0, c_H(f)]$, we have*

$$c_{H'}(f - tl_H) = \begin{cases} c_H(f) - t & \text{if} \quad H' = H, \\ c_{H'}(f) & \text{otherwise.} \end{cases} \tag{2.7}$$

The polyhedral split decomposition reads as follows.

Theorem 2.14. *A polyhedral convex function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ with $\operatorname{int} \operatorname{dom} f \neq \emptyset$ can be decomposed as*

$$f = \sum_{H \in \mathcal{H}(f)} c_H(f) l_H + f', \tag{2.8}$$

where $f' : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is a polyhedral convex function with $c_{H'}(f') \in \{0, \infty\}$ for any hyperplane H'. Furthermore this representation is unique.

The uniqueness assertion in the theorem can be seen as follows. If f is decomposed as $f = \sum_{H \in \mathcal{H}} \alpha_H l_H + g$ for some finite set of hyperplanes $\mathcal{H}$ meeting $\operatorname{int} \operatorname{dom} f$, positive weight $\alpha \in \mathbf{R}^{\mathcal{H}}_{++}$, and polyhedral convex function g with $c_{H'}(g) \in \{0, \infty\}$ for any hyperplane H', then, by (2.8), we have

$$g = \sum_{H \in \mathcal{H} \cap \mathcal{H}(f)} \{c_H(f) - \alpha_H\} l_H + \sum_{H \in \mathcal{H} \setminus \mathcal{H}(f)} -\alpha_H l_H + \sum_{H \in \mathcal{H}(f) \setminus \mathcal{H}} c_H(f) l_H + f'.$$

By Lemma 2.11 (2), we have $\mathcal{H} \subseteq \mathcal{H}(f)$ and $c_H(f) \geq \alpha_H$ for $H \in \mathcal{H}$. Since $c_{H'}(g) \in \{0, \infty\}$ holds for any hyperplane H', it must be that $\mathcal{H}(f) = \mathcal{H}$, $\alpha_H = c_H(f)$ for $H \in \mathcal{H}(f)$, and $f' = g$ by Lemma 2.11 (2).

Remark 2.15. Applying the polyhedral split decomposition to the support functions of polytopes, we can derive the classical result of Bolker [3] that every polytope has the maximum zonotopic summand.

Remark 2.16. For $H \in \mathcal{H}(f)$, the set of parallel edges

$$\{F^\bullet \mid F \in \mathcal{T}(f),\ \operatorname{aff} F = H\} \tag{2.9}$$

forms a cutset of 1-skeleton of $\mathcal{T}(f^\bullet)$. It is found (cf. the last part of the proof of Proposition 2.20) that $c_H(f)$ is the minimum length of the cutset (2.9). By Lemma 2.4, subtracting a split function l_H from f corresponds to contracting of the cutset (2.9). Hence, in $\mathcal{T}((f - c_H(f) l_H)^\bullet)$, some edges of the cutset (2.9) shrink to the points.

Let $\mathcal{T}_b(f^\bullet)$ be defined by the set of bounded elements of $\mathcal{T}(f^\bullet)$, i.e.,

$$\mathcal{T}_b(f^\bullet) = \{\partial f(x) \mid x \in \operatorname{int} \operatorname{dom} f\}, \tag{2.10}$$

which is a polyhedral complex. Since the union of bounded faces of the polyhedron is contractible, $|\mathcal{T}_b(f^\bullet)|$ is contractible. The structure of $\mathcal{T}_b(f^\bullet)$ expresses the nonlinearity of f over $\operatorname{dom} f$. For example, it is easily observed that $[\dim \mathcal{T}_b(f^\bullet) = 0] \Leftrightarrow [\mathcal{T}_b(f^\bullet)$ is a single point$] \Leftrightarrow [f$ is affine over $\operatorname{dom} f]$. The following is a generalization of this observation and gives a new interpretation to the famous result of Dress [8] that a metric is a tree metric if and only if its tight span is a tree (see Section 4 for further discussion on this issue).

Proposition 2.17. *The following conditions are equivalent.*

(1) $\dim \mathcal{T}_b(f^\bullet) = 1$.

(2) $\mathcal{T}_b(f^\bullet)$ *is a tree.*

(3) *If f is decomposed as* (2.8)*, then f' is affine over* $\operatorname{dom} f$ *and* $H_1 \cap H_2 \cap \operatorname{int}\operatorname{dom} f \neq \emptyset$ *for each pair $H_1, H_2 \in \mathcal{H}(f)$.*

Proof. (1) $\Leftrightarrow$ (2) follows from the contractibility of $|\mathcal{T}_b(f^\bullet)|$. (3) $\Rightarrow$ (1) is immediate from the fact that for $x \in \operatorname{int}\operatorname{dom} f$, $\partial f(x)$ is a line segment or a point by Proposition 2.9. We show (2) $\Rightarrow$ (3). Let $E \in \mathcal{T}_b(f^\bullet)$ be a line segment and $H = \operatorname{aff} E^\bullet$ be a hyperplane. We claim that $E^\bullet = H \cap \operatorname{dom} f$ for each edge $E \in \mathcal{T}_b(f^\bullet)$. Note that all proper faces of $E^\bullet$ are in a boundary of $\operatorname{dom} f$. Obviously $E^\bullet \subseteq H \cap \operatorname{dom} f$. Consider $x \in H \cap \operatorname{int}\operatorname{dom} f$. If $x \notin E^\bullet$, for $y \in E^\bullet$, the line segment $[x, y] \subseteq \operatorname{int}\operatorname{dom} f$ meets proper faces of $E^\bullet$. This is a contradiction. Hence we have $H \cap \operatorname{int}\operatorname{dom} f \subseteq E^\bullet \subseteq H \cap \operatorname{dom} f$. By the closedness of $E^\bullet$, we obtain $E^\bullet = H \cap \operatorname{dom} f$. By Proposition 2.20 (below), we conclude that hyperplanes $\{\operatorname{aff} E^\bullet \mid E \in \mathcal{T}_b(f^\bullet), \dim E = 1\} = \mathcal{H}(f)$ and do not intersect in $\operatorname{int}\operatorname{dom} f$ each other. For an edge $E \in \mathcal{T}_b(f^\bullet)$ and $H = \operatorname{aff} E^\bullet$, E is contracted to a point in $\mathcal{T}_b((f - c_H(f) l_H)^\bullet)$. By repeating the process, $\mathcal{T}_b((f - \sum_{H \in \mathcal{H}(f)} c_H(f) l_H)^\bullet)$ becomes a single point. Hence $f - \sum_{H \in \mathcal{H}(f)} c_H(f) l_H$ is affine over $\operatorname{dom} f$. $\square$

2.5 Proof of Proposition 2.13

To prove Proposition 2.13, we need more detailed study of the quotient $c_H(f)$ and its relation to the structure of $\mathcal{T}(f)$. The quotient $c_H(f)$ of f is represented explicitly.

Proposition 2.18. *Let H be a hyperplane in $\mathbf{R}^n$. Then we have*

$$c_H(f) = \frac{1}{2} \inf \left\{ \frac{f(x) - f(w)}{l_H(x)} + \frac{f(y) - f(w)}{l_H(y)} \;\middle|\; \begin{array}{c} x \in \operatorname{dom} f \cap H^{++}, \\ y \in \operatorname{dom} f \cap H^{--}, \\ \{w\} = [x, y] \cap H \end{array} \right\}. \tag{2.11}$$

Proof. Note that l_H is affine over $\operatorname{dom} f \cap H^{\pm}$. Hence the condition "$f - t l_H$ is convex" is represented as

$$\begin{aligned} &[\lambda l_H(x) + (1-\lambda) l_H(y) - l_H(\lambda x + (1-\lambda) y)] t \\ &\qquad \leq \lambda f(x) + (1-\lambda) f(y) - f(\lambda x + (1-\lambda) y) \\ &\qquad (\forall x \in \operatorname{dom} f \cap H^{++}, \forall y \in \operatorname{dom} f \cap H^{--}, \forall \lambda \in [0,1]). \end{aligned} \tag{2.12}$$

The condition (2.12) holds if and only if it holds for $\lambda = l_H(y)/(l_H(x) + l_H(y))$; see Lemma 2.19 below. For $w = (l_H(y) x + l_H(x) y)/(l_H(x) + l_H(y))$ we have $\{w\} = [x, y] \cap H$. Therefore the condition (2.12) is equivalent to

$$\begin{aligned} &t \leq \frac{1}{2} \left\{ \frac{f(x) - f(w)}{l_H(x)} + \frac{f(y) - f(w)}{l_H(y)} \right\} \\ &\qquad (\forall x \in \operatorname{dom} f \cap H^{++}, \forall y \in \operatorname{dom} f \cap H^{--}), \end{aligned} \tag{2.13}$$

which implies (2.11). $\square$

Lemma 2.19. *Let $g : \mathbf{R} \to \mathbf{R}$ be convex on each of the two intervals $[a, c]$ and $[c, b]$ for $a < c < b \in \mathbf{R}$. Then g satisfies*

$$g(\lambda a + (1-\lambda) b) \leq \lambda g(a) + (1-\lambda) g(b) \tag{2.14}$$

for all $\lambda \in [0,1]$ if and only if g satisfies (2.14) *for $\lambda = (b-c)/(b-a)$.*

Next we characterize the situation $0 < c_H(f) < \infty$ in terms of the structure of $\mathcal{T}(f)$.

Proposition 2.20. *A hyperplane H belongs to $\mathcal{H}(f)$ if and only if H satisfies the following conditions* (1) *and* (2)*.*

(1) $H^{++} \cap \operatorname{dom} f \neq \emptyset$ *and* $H^{--} \cap \operatorname{dom} f \neq \emptyset$*.*

(2) $\{F \in \mathcal{T}(f) \mid F \subseteq H\}$ *is a subdivision of* $H \cap \operatorname{dom} f$*.*

Proof. It suffices to prove "if" part. Let n_H be the unit normal vector of H. Obviously $c_H(f) < \infty$. For $x \in H^{++} \cap \operatorname{dom} f$, $y \in H^{--} \cap \operatorname{dom} f$, and $\{w\} = [x, y] \cap H$, we have

$$\begin{aligned}
&\{f(x) - f(w)\}/2l_H(x) + \{f(y) - f(w)\}/2l_H(y) \\
&= \frac{f(w + \|x - w\|d) - f(w)}{\langle n_H, d\rangle \|x - w\|} + \frac{f(w - \|y - w\|d) - f(w)}{\langle n_H, d\rangle \|y - w\|} \\
&\geq \{f'(w; d) + f'(w; -d)\}/\langle n_H, d\rangle \\
&= \sup\{\langle p - q, d\rangle \mid p, q \in \partial f(w)\}/\langle n_H, d\rangle, \qquad (2.15)
\end{aligned}$$

where $d = (x - y)/\|x - y\|$ and the last equality follows from Lemma 2.3. Let F_w be the unique element of $\mathcal{T}(f)$ satisfying $w \in \operatorname{ri} F_w$. By the condition (2), we have $F_w \subseteq H$. By the pureness of $\{F \in \mathcal{T}(f) \mid F \subseteq H\}$, there exists $G \in \mathcal{T}(f)$ such that $F_w \subseteq G \subseteq H$ and $\dim G = n - 1$. By Proposition 2.6 (4) we have

$$\begin{aligned}
\sup\{\langle p - q, d\rangle \mid p, q \in \partial f(w)\} &= \sup\{\langle p - q, d\rangle \mid p, q \in F_w^\bullet\} \\
&\geq \sup\{\langle p - q, d\rangle \mid p, q \in G^\bullet\}.
\end{aligned}$$

Note that since $G \cap \operatorname{int} \operatorname{dom} f \neq \emptyset$, $G^\bullet$ is bounded. By $\operatorname{aff} G = H$ and Proposition 2.6 (4), $G^\bullet$ is a line segment represented as

$$G^\bullet = p_0 + \alpha_G[-n_H/2, n_H/2]$$

for some $p_0 \in \mathbf{R}^n$ and $\alpha_G > 0$, i.e., α_G is the length of $G^\bullet$. Hence,

$$\sup\{\langle p - q, d\rangle \mid p, q \in G^\bullet\}/\langle n_H, d\rangle = \alpha_G > 0.$$

Since the set $\{F \in \mathcal{T}(f) \mid F \subseteq H,\ \dim F = n - 1\}$ is finite, we obtain $c_H(f) > 0$. □

The following indicates that minimizers of (2.11) can be taken generically.

Lemma 2.21. *Let H be a hyperplane.*

(1) *If $c_H(f) = 0$, then there exists an n-dimensional polyhedron $F \in \mathcal{T}(f)$ such that*

   (1.1) $F \cap H^{++} \neq \emptyset$, $F \cap H^{--} \neq \emptyset$ *and*

   (1.2) *the minimum of* (2.11) *is attained by any $x \in F \cap H^{++}$, $y \in F \cap H^{--}$.*

(2) *If $0 < c_H(f) < \infty$, there exist n-dimensional polyhedra $G_1, G_2 \in \mathcal{T}(f)$ such that*

   (2.1) $G_1 \cup G_2 \in \mathcal{T}(f - c_H(f)l_H)$ *and*

   (2.2) *the minimum of* (2.11) *is attained by any $x \in G_1 \setminus H$, $y \in G_2 \setminus H$.*

Proof. (1) is immediate from Proposition 2.20. We show (2). Put $g = f - c_H(f)l_H$. By $c_H(g) = 0$ and (1), there exists n-dimensional polyhedron G such that $G \cap H^{++} \neq \emptyset$ and $G \cap H^{--} \neq \emptyset$. Put $G_1 = G \cap H^+$ and $G_2 = G \cap H^-$ Then we have $G_1, G_2 \in \mathcal{T}(f)$

by Lemma 2.7 and Proposition 2.9. Therefore we obtain (2.1). Since g is affine over $G = G_1 \cup G_2$, we have

$$\begin{aligned} & \{f(x) - f(w)\}/2l_H(x) + \{f(x) - f(w)\}/2l_H(x) \\ = \ & \{g(x) - g(w) + c_H(f)l_H(x)\}/2l_H(x) + \{g(y) - g(w) + c_H(f)l_H(y)\}/2l_H(y) \\ = \ & c_H(f), \end{aligned}$$

where $x \in G_1 \setminus H$, $y \in G_2 \setminus H$ and $\{w\} = [x, y] \cap H$. This implies (2.2). $\square$

We are now in the position to prove Proposition 2.13.

Proof of Proposition 2.13. The case $H' = H$ is immediate from Lemma 2.11 (1). Hence we consider the case $H' \neq H$. It is sufficient to show $c_{H'}(f - tl_H) \geq c_{H'}(f)$ by Lemma 2.11 (3). By Lemma 2.21, we may assume that the minimum of (2.11) is attained by some $x, y \in H^+$ or $x, y \in H^-$. This implies

$$c_{H'}(f - tl_H) = \frac{1}{2}\left\{\frac{f(x) - f(w)}{l_{H'}(x)} + \frac{f(y) - f(w)}{l_{H'}(y)}\right\} \geq c_{H'}(f),$$

where $\{w\} = [x, y] \cap H$. $\square$

3 Split decomposition of discrete functions

In this section, we describe the split decomposition of functions defined on a finite set X of points of $\mathbf{R}^n$. The basic idea is to apply the polyhedral split decomposition to the convex-extension of a given discrete function.

3.1 Discrete functions and convex-extension

For a function $f : X \to \mathbf{R}$, the *convex-extension* of f is defined by

$$\inf\{\textstyle\sum_{y \in X} \lambda_y f(y) \mid \sum_{y \in X} \lambda_y (y, 1) = (x, 1),\ \lambda_y \geq 0\ (y \in X)\} \quad (x \in \mathbf{R}^n). \tag{3.1}$$

For subsequent discussions, however, it is more convenient to employ the *homogeneous convex extension* of f defined by

$$\overline{f}(x) = \inf\{\textstyle\sum_{y \in X} \lambda_y f(y) \mid \sum_{y \in X} \lambda_y y = x,\ \lambda_y \geq 0\ (y \in X)\} \quad (x \in \mathbf{R}^n), \tag{3.2}$$

where we assume:

Assumption 3.1. X is included in some hyperplane K with $0 \notin K$.

It is noted that the subsequent results can also be adapted to the case of convex-extension (3.1). By the definition, $\overline{f}$ is a positively homogeneous polyhedral convex function with $\operatorname{dom} \overline{f} = \operatorname{cone} X$. By linear programming duality, $\overline{f}$ is also represented as

$$\overline{f}(x) = \sup\{\langle p, x\rangle \mid p \in \mathbf{R}^n, \langle p, y\rangle \leq f(y)\ (y \in X)\} \quad (x \in \mathbf{R}^n). \tag{3.3}$$

Hence $\overline{f}$ is the support function of a polyhedron

$$Q(f) = \{p \in \mathbf{R}^n \mid \langle p, y\rangle \leq f(y),\ (y \in X)\} (= \partial \overline{f}(0)). \tag{3.4}$$

Note that $\mathcal{T}(\overline{f})$ forms a *fan*, i.e., each element of $\mathcal{T}(\overline{f})$ is a cone. Let $\mathcal{T}^X(\overline{f})$ denote the subdivision of $\operatorname{conv} X$ which is defined by $\{F \cap K \mid F \in \mathcal{T}^X(\overline{f})\}$. For a function

$g : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$, we denote the restriction of g to X by g^X. f is said to be *convex-extensible* if it satisfies $\overline{f}^X = f$. The set of convex-extensible functions is recognized as a fundamental class of the *discrete convex functions* (see [19]).

We give some fundamental properties of discrete functions and homogeneous convex-extensions.

Lemma 3.2. *Let $f, g : X \to \mathbf{R}$ be a function.*

(1) $\overline{cf[q]} = c\overline{f}[q] \quad (c \in \mathbf{R}_+, q \in \mathbf{R}^n)$.

(2) *$F \in \mathcal{T}(\overline{f})$ is represented as* $\mathrm{cone}\{y \mid y \in X,\ \langle p, y\rangle = f(y)\}$ *for some $p \in Q(f)$. Furthermore $\overline{f}(x) = \overline{f^{F\cap X}}(x)$ for $x \in F$. In particular, $\overline{f}(y) = f(y)$ if $y \in X$ is a vertex of $\mathcal{T}^X(\overline{f})$.*

3.2 Split decomposition for discrete functions

The split decomposition for discrete functions is derived here. As shown in the next section, this turns out to be a generalization of Bandelt-Dress' split decomposition of metrics. Furthermore some interesting properties of weakly compatible splits are also generalized. Note that since $\mathcal{T}(\overline{f})$ for $f : X \to \mathbf{R}$ is a fan, each hyperplane $H \in \mathcal{T}(\overline{f})$ is linear, i.e., $H = H_{a,0}$ for some $a \in \mathbf{R}^n$.

Corresponding to Assumption 2.12, we assume full-dimensionality of the domain of the homogeneous convex extension.

Assumption 3.3. $X \subseteq \mathbf{R}^n$ is a finite set with $\mathrm{aff}\, X = K$.

The following proposition indicates that the polyhedral split decomposition of convex extensions can be carried out on the discrete side.

Proposition 3.4. *For $f : X \to \mathbf{R}$, $H \in \mathcal{H}(\overline{f})$, and $t \in [0, c_H(\overline{f})]$, we have*

$$\overline{f} = tl_H + \overline{f - tl_H^X}. \tag{3.5}$$

Proof. Suppose that $H = H_{a,0}$ for $a \in \mathbf{R}^n$ with $\|a\| = 1$. (3.5) is equivalent to

$$Q(f) = t[a/2, -a/2] + Q(f - tl_H^X).$$

We show this. The inclusion $(\supseteq)$ follows from

$$\langle p + sa/2, y\rangle \le f(y) - t|\langle a, y\rangle|/2 + s\langle a, y\rangle/2 \le f(y) \quad (y \in X)$$

for $s \in [-t, t]$ and $p \in Q(f - tl_H^X)$. We next show $(\subseteq)$. By $\overline{f} = tl_H + (\overline{f} - tl_H)$ and Lemma 2.4, we have

$$Q(f) = \partial f(0) = t[a/2, -a/2] + \partial(\overline{f} - tl_H)(0).$$

By the definition of $Q(\cdot)$ and $\overline{f}^X \le f$, we have

$$\partial(\overline{f} - tl_H)(0) \subseteq Q((\overline{f} - tl_H)^X) \subseteq Q(f - tl_H^X).$$

Hence, we obtain $(\subseteq)$. ∎

With this proposition, we obtain the split decomposition of discrete functions.

Theorem 3.5. *A discrete function $f : X \to \mathbf{R}$ can be decomposed as*

$$f = \sum_{H \in \mathcal{H}(\overline{f})} c_H(\overline{f}) l_H^X + \gamma, \tag{3.6}$$

where $\gamma : X \to \mathbf{R}$ satisfies $c_{H'}(\overline{\gamma}) \in \{0, \infty\}$ for any hyperplane H'. Furthermore we have

$$\overline{f} = \sum_{H \in \mathcal{H}(\overline{f})} c_H(\overline{f}) l_H + \overline{\gamma}. \tag{3.7}$$

If, in addition, f is convex-extensible, then γ is also convex-extensible.

Proof. The first part of the theorem follows from Proposition 3.4 and Theorem 2.14. If f is convex-extensible, then $\overline{f}^X = f$. By restricting both side of (3.7) to X, we have $\gamma = \overline{\gamma}^X$. Hence γ is convex-extensible. □

Next we consider an algorithmic aspect of the decomposition (3.6). We begin with the relationship between (linear) hyperplanes $\mathcal{H}(\overline{f})$ and points X.

Lemma 3.6. *For $f : X \to \mathbf{R}$, any hyperplane $H \in \mathcal{H}(\overline{f})$ satisfies the following conditions* (0), (1), (2) *and* (3).

(0) $H \cap \operatorname{ri} \operatorname{conv} X \neq \emptyset$.

(1) $\operatorname{conv}(X \cap H) = H \cap \operatorname{conv} X$.

(2) $\operatorname{conv}(X \cap H^+) = H^+ \cap \operatorname{conv} X$.

(3) $\operatorname{conv}(X \cap H^-) = H^- \cap \operatorname{conv} X$.

Proof. The condition (0) is obviously satisfied. In the conditions (1), (2), and (3), the inclusion ($\subseteq$) always holds. Hence it suffices to show ($\supseteq$). Let $x \in H^+ \cap \operatorname{conv} X$. Since $H \in \mathcal{H}(\overline{f})$ and Proposition 2.20, there exists $F \in \mathcal{T}(\overline{f})$ such that $F \subseteq H^+$ and $x \in F$. By Lemma 3.2 (2), x can be represented as the convex combination of $F \cap X$. This implies ($\supseteq$) and hence we have (2). The conditions (1) and (3) are also obtained in a similar way. □

Motivated by Lemma 3.6 above, we define a set of hyperplanes $\mathcal{H}_X$ that is determined solely by X, independently of f.

Definition 3.7. Let $\mathcal{H}_X$ denote the set of all linear hyperplanes satisfying the conditions (0), (1), (2), and (3) of Lemma 3.6.

Note that $\mathcal{H}(\overline{f}) \subseteq \mathcal{H}_X$ holds for any $f : X \to \mathbf{R}$ by Lemma 3.6. Also note that $\mathcal{H}_X$ is a finite set, since $H \in \mathcal{H}_X$ is represented as a linear hull of some $Y \subseteq X$.

The next theorem implies that the discrete split decomposition (3.6) can be carried out without explicit construction of convex-extensions; the quotient $c_H(\overline{f})$ can be calculated on the discrete side.

Theorem 3.8. *For a function $f : X \to \mathbf{R}$ and a hyperplane $H \in \mathcal{H}_X$, let $\tilde{c}_H(f)$ be defined by*

$$\frac{1}{2} \min \left\{ \frac{f(x) - \overline{f^{X \cap H}}(w)}{l_H(x)} + \frac{f(y) - \overline{f^{X \cap H}}(w)}{l_H(y)} \;\middle|\; \begin{array}{c} x \in X \cap H^{++} \\ y \in X \cap H^{--} \\ \{w\} = H \cap [x, y] \end{array} \right\}. \tag{3.8}$$

Then we have

$$c_H(\overline{f}) = \max(0, \tilde{c}_H(f)). \tag{3.9}$$

Proof. By Lemma 2.21 and Lemma 3.2 (2), $c_H(\overline{f})$ can be represented as

$$c_H(\overline{f}) = \frac{1}{2} \min \left\{ \frac{f(y) - \overline{f}(w)}{l_H(y)} + \frac{f(z) - \overline{f}(w)}{l_H(z)} \;\middle|\; \begin{array}{c} y \in X \cap H^{++} \\ z \in X \cap H^{--} \\ \{w\} = H \cap [x, y] \end{array} \right\}.$$

Since $\overline{f^{X \cap H}}(w) \geq \overline{f}(w)$ holds for $w \in H \cap \operatorname{cone} X$, we have $c_H(\overline{f}) \geq \tilde{c}_H(f)$. If $c_H(\overline{f}) > 0$ holds, then for $w \in H \cap \operatorname{cone} X$, there exists $F \in \mathcal{T}(\overline{f})$ such that $w \in F$ and $F \subseteq H$ by Proposition 2.20. Therefore, by Lemma 3.2 (2), we have $\overline{f}(w) = \overline{f^{X \cap F}}(w) = \overline{f^{X \cap H}}(w)$ and $c_H(\overline{f}) = \tilde{c}_H(f)$. □

The theorem yields an algorithm for the split decomposition of $f : X \to \mathbf{R}$ as follows:

1. Determine $\mathcal{H}_X$ from the points X.
2. Calculate $c_H(\overline{f})$ for $H \in \mathcal{H}_X$ by formulas (3.8) and (3.9).
3. Decompose f into the form of (3.6).

In Section 4, we derive Bandelt-Dress' split decomposition from this recipe. We continue to study the sets of hyperplanes $\mathcal{H}_X$ and $\mathcal{H}(\overline{f})$. Similarly to the proof of Lemma 3.6, we obtain the following fact.

Lemma 3.9. *For $f : X \to \mathbf{R}$, the set of hyperplanes $\mathcal{H}(\overline{f})$ satisfies the following property:*

$$\forall F \in \mathcal{A}(\mathcal{H}(\overline{f})),\ \operatorname{conv}(X \cap F) = F \cap \operatorname{conv} X. \tag{3.10}$$

A set of hyperplanes $\mathcal{H} \subseteq \mathcal{H}_X$ is said to be X-*admissible* if it satisfies the property (3.10); $\mathcal{H}(\overline{f})$ is X-admissible by Lemma 3.10. As will be shown in Section 4, Theorem 4.15, X-admissibility of hyperplanes corresponds to weak compatibility of splits. By the definition, any singleton $\{H\} \subseteq \mathcal{H}_X$ is X-admissible. In addition, any subset of X-admissible hyperplanes is also X-admissible. The next proposition corresponds to [1, Corollary 10].

Proposition 3.10. *For $\mathcal{H} \subseteq \mathcal{H}_X$ and $\alpha \in \mathbf{R}^{\mathcal{H}}_{++}$, let $f = \sum_{H \in \mathcal{H}} \alpha_H l^X_H$. Then the following conditions (a), (b) and (c) are equivalent;*

(*a*) $\overline{f} = \sum_{H \in \mathcal{H}} \alpha_H l_H + \delta_{\operatorname{cone} X}$

(*b*) $\mathcal{H} = \mathcal{H}(\overline{f})$ *and* $\alpha_H = c_H(\overline{f})$ *for* $H \in \mathcal{H}$.

(*c*) $\mathcal{H}$ *is X-admissible.*

Proof. (a) ⇔ (b) follows from Theorems 2.14 and 3.5. (b) ⇒ (c) follows from Lemma 3.9.

(c) ⇒ (a). Let $\mathcal{H} = \{H_{a_i,0} \mid i \in I\}$ with $\|a_i\| = 1$ $(i \in I)$. For $x \in \operatorname{cone} X$, there exists $F \in \mathcal{A}(\mathcal{H})$ such that $x \in F$. By the condition (c), x can be represented as a conical combination

$$x = \sum_{y \in F \cap X} \lambda_y y. \tag{3.11}$$

Then the coefficient λ is feasible to the linear program corresponding (3.2). On the other hand, p defined as

$$p = \sum_{i \in I,\ F \subseteq H_i^+} (\alpha_{H_i}/2) a_i - \sum_{i \in I,\ F \subseteq H_i^-} (\alpha_{H_i}/2) a_i$$

is also feasible to the dual program corresponding (3.3). It is easy to check that λ and p satisfy the complementary slackness condition. This implies (a). □

In particular, from the equivalence between (b) and (c) of this proposition, we see that the decomposition into a sum of X-admissible split functions is uniquely determined.

Remark 3.11. In the decomposition (3.6), the property (3.7) is equivalent to

$$Q(f) = \sum_{H \in \mathcal{H}(\overline{f})} c_H(\overline{f})[n_H/2, -n_H/2] + Q(\gamma), \tag{3.12}$$

where n_H is the unit normal vector of $H \in \mathcal{H}(\overline{f})$. This implies that the decomposition (3.6) extracts a zonotope from (unbounded) polyhedron $Q(f)$. (3.12) is also rewritten as

$$Q(f) = \sum_{H \in \mathcal{H}(\overline{f})} c_H(\overline{f}) Q(l_H^X) + Q(\gamma). \tag{3.13}$$

This corresponds to the coherent decomposition of metrics.

The number of X-admissible hyperplanes is bounded by $|X| - n$. This fact corresponds to [1, Corollary 4].

Proposition 3.12. *Let $\mathcal{H} \subseteq \mathcal{H}_X$ be an X-admissible set of hyperplanes. Then the set of vectors $\{l_H^X \mid H \in \mathcal{H}\} \cup \{\chi_i^X \mid 1 \leq i \leq n\}$ is linearly independent in $\mathbf{R}^X$. Therefore $|\mathcal{H}| \leq |X| - n$.*

Proof. Suppose that it is linearly dependent. Then there exists nonzero $(\alpha, p) \in \mathbf{R}^{\mathcal{H}} \times \mathbf{R}^n$ such that

$$\sum_{H \in \mathcal{H},\ \alpha_H > 0} \alpha_H l_H^X + \langle p, \cdot \rangle = \sum_{H \in \mathcal{H},\ \alpha_H < 0} -\alpha_H l_H^X. \tag{3.14}$$

By Proposition 3.10 and Lemma 3.2 (1), the convex extensions of both sides of (3.14) lead to a contradiction to the uniqueness of the polyhedral split decomposition (Theorem 2.14). □

If an X-admissible $\mathcal{H}$ has the maximal cardinality $|X| - n$, the cone

$$\{\sum_{H \in \mathcal{H}} \alpha_H l_H^X + g^X \mid \alpha \in \mathbf{R}_+^{\mathcal{H}}, g : \text{linear function on } \mathbf{R}^n\}, \tag{3.15}$$

as a subset of $\mathbf{R}^X$, has interior points. Therefore, for sufficiently generic $f : X \to \mathbf{R}$ from the cone (3.15), $\mathcal{T}(\overline{f})$ forms a simplicial fan (see [14, Chapter 7]).

Corollary 3.13. *Let $\mathcal{H} \subseteq \mathcal{H}_X$ be an X-admissible set of hyperplanes with maximal cardinality $|X| - n$ and let $f = \sum_{H \in \mathcal{H}} l_H^X$. Then $\mathcal{T}^X(\overline{f})$ is a triangulation. Furthermore, the set of vertices of $\mathcal{T}^X(\overline{f})$ coincides with X.*

Proof. The latter part is immediate from the fact that if $y \in X$ is not a vertex of $\mathcal{T}^X(\overline{f})$, then $\mathcal{H}$ is also $(X \setminus \{y\})$-admissible. □

4 Metrics as discrete concave functions

By regarding metrics as discrete concave functions and applying the results of Sections 2 and 3, we derive in this section the tight span, Bandelt-Dress' split decomposition of metrics and some other important concepts of T-theory.

First we briefly review T-theory and the split decomposition of metrics. Let V be a finite set. A function $d : V \times V \to \mathbf{R}$ is said to be a *metric* if it satisfies $d(i, i) = 0$,

$d(i,j) = d(j,i) \geq 0$, and $d(i,j) \geq d(i,k)+d(j,k)$ for $i,j,k \in V$. A polyhedron $P(d) \subseteq \mathbf{R}^V$ associated with metric d is defined as

$$P(d) = \{p \in \mathbf{R}^V \mid p(i) + p(j) \geq d(i,j)\}. \tag{4.1}$$

The tight span of metric d is a subset of $P(d)$ defined as

$$T(d) = \{p \in \mathbf{R}^V \mid \forall i \in V,\ p(i) = \sup_{j \in V}(d(i,j) - p(j))\}. \tag{4.2}$$

It is known that $T(d)$ coincides with the union of all bounded faces of $P(d)$ [9, Lemma 1].

A *split* $\{A,B\}$ is a bipartition of V, which means that $A \cap B = \emptyset$, $A \cup B = V$, $A, B \neq \emptyset$. A *split metric* $\delta_{\{A,B\}}$ associated with a split $\{A,B\}$ is defined as

$$\delta_{\{A,B\}}(i,j) = \begin{cases} 0 & \text{if } i,j \in A \text{ or } i,j \in B, \\ 1 & \text{otherwise} \end{cases} \qquad (i,j \in V).$$

For a metric d and a split $\{A,B\}$, the *isolation index* $\alpha^d_{\{A,B\}}$ is a nonnegative number defined as

$$\alpha^d_{\{A,B\}} = \frac{1}{2} \min_{i,j \in A,\ k,l \in B} \left\{ \max \left\{ \begin{array}{c} d(i,k) + d(j,l) \\ d(i,l) + d(j,k) \\ d(i,j) + d(k,l) \end{array} \right\} - d(i,j) - d(k,l) \right\}. \tag{4.3}$$

Let $\mathcal{S}(d)$ be the collection of splits defined as

$$\mathcal{S}(d) = \{S : \text{split on } V \mid \alpha^d_S > 0\}. \tag{4.4}$$

The split decomposition theorem is as follows:

Theorem 4.1 (Bandelt-Dress [1]). *A metric d can be decomposed as*

$$d = \sum_{S \in \mathcal{S}(d)} \alpha^d_S \delta_S + d', \tag{4.5}$$

where d' is a metric with $\alpha^{d'}_{S'} = 0$ for each split S'. Furthermore, the decomposition is coherent, i.e.,

$$P(d) = \sum_{S \in \mathcal{S}(d)} \alpha^d_S P(\delta_S) + P(d') \tag{4.6}$$

holds.

The collection of splits $\mathcal{S}(d)$ is necessarily *weakly compatible*, that is, for any three splits $S_1 = \{A_1, B_1\}$, $S_2 = \{A_2, B_2\}$, and $S_3 = \{A_3, B_3\}$ in $\mathcal{S}(d)$, there exist no four points $a, a_1, a_2, a_3 \in V$ with $\{a, a_1, a_2, a_3\} \cap A_i = \{a, a_i\}$ for $i = 1,2,3$.

As already mentioned in the introduction, we regard a metric as a function defined on a finite set of points. In the following we deal with a more general object, *distance function*, rather than a metric. A distance function on V is a function $d : V \times V \to \mathbf{R}$ such that $d(i,i) = 0$ and $d(i,j) = d(j,i) \geq 0$ for $i,j \in V$, where the triangle inequality is not imposed. For each $A \subseteq V$, we denote by χ_A the characteristic vector of A defined as: $\chi_A(i) = 1$ if $i \in A$ and 0 otherwise. In particular we write simply χ_i instead of $\chi_{\{i\}}$ for each $i \in V$. A distance function d is naturally regarded as a function defined on the set $\Lambda = \{\chi_i + \chi_j \mid i,j \in V\}$ by the correspondence

$$d(\chi_i + \chi_j) \leftarrow d(i,j) \quad (i,j \in V). \tag{4.7}$$

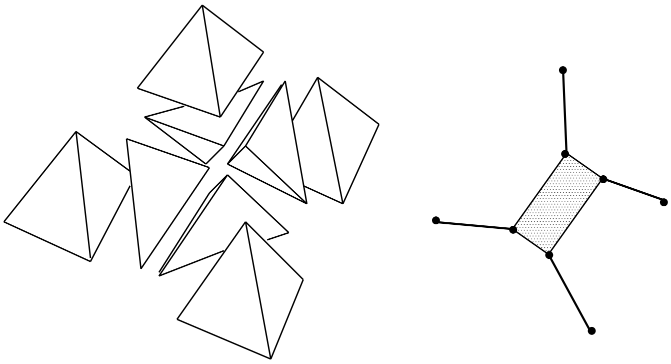

Figure 1: $\mathcal{T}^\Lambda(\overline{(-d)})$ (left) and $T(d)$ (right) of a generic 4-point metric d

Remark 4.2. Any symmetric function $f : V \times V \to \mathbf{R}$ with $f(i,j) = f(j,i)$ for $i,j \in V$ can be regarded as $f : \Lambda \to \mathbf{R}$ by the correspondence (4.7). Similarly, $P(f)$ and $T(f)$ are also definable. The following arguments can be adapted to any symmetric function on V; in fact, [1] discusses the split decomposition (4.5) for symmetric functions.

The following is easily observed.

Lemma 4.3. *A function $f : \Lambda \to \mathbf{R}$ with $f(2\chi_i) = 0$ for $i \in V$ is convex-extensible if and only if it satisfies $f(\chi_i + \chi_j) \leq 0$ for $i, j \in V$.*

Hence it is natural to regard any distance function $d : V \times V \to \mathbf{R}$ as a discrete concave function on Λ. Since aff Λ coincides with $K = \{x \in \mathbf{R}^V \mid \sum_{i \in V} x(i) = 2\}$, we can apply the results of the previous section in a straightforward manner.

The homogeneous convex extension of $-d$ is given by

$$\begin{aligned}
&\overline{(-d)}(x) \\
&= \inf\{\sum_{i,j \in V} \lambda_{ij}(-d(i,j)) \mid \sum_{i,j \in V} \lambda_{ij}(\chi_i + \chi_j) = x,\ \lambda_{ij} \geq 0 \quad (i,j \in V)\} \\
&= \sup\{\langle p, x\rangle \mid p(i) + p(j) \leq -d(i,j)\} \\
&= \sup\{\langle p, x\rangle \mid -p \in P(d)\} \quad (x \in \mathbf{R}^V),
\end{aligned} \tag{4.8}$$

where for $i, j \in V$, ij denotes an unordered pair, which means that ij and ji are not distinguished from each other. Hence $\overline{-d}$ is the support function of $-P(d)$.

Proposition 4.4. *$\mathcal{T}(\overline{-d}^\bullet)$ is a face lattice of $-P(d)$.*

Using the notation (3.4) and (2.10), we have $Q(-d) = -P(d)$ and $|\mathcal{T}_b(\overline{-d}^\bullet)| = -T(d)$. This implies that the tight span $T(d)$ has a dual structure of $\mathcal{T}^\Lambda(\overline{-d})$ (recall Proposition 2.6 and see Figure 1). This duality relation is also suggested by Sturmfels and Yu [24] in the connection of the triangulation of the *second hypersimplex* $\mathrm{conv}\{\chi_i + \chi_j \mid i, j \in V,\ i \neq j\}$.

Remark 4.5. An additive decomposition of a distance function $d = d_1 + d_2$ is said to be *coherent* if $P(d) = P(d_1) + P(d_2)$ holds. Note that this condition is equivalent to $\overline{(-d)} = \overline{(-d_1)} + \overline{(-d_2)}$. By Lemma 2.7, we have $\mathcal{T}^\Lambda(\overline{(-d)}) = \mathcal{T}^\Lambda(\overline{(-d_1)}) \wedge \mathcal{T}^\Lambda(\overline{(-d_2)})$. This implies that the coherent decomposition corresponds to a representation of $\mathcal{T}^\Lambda(\overline{(-d)})$ as a common refinement of coarser subdivisions induced by distance functions. In particular, it is observed from the subdivision of $\mathrm{conv}\{2\chi_i, 2\chi_j, 2\chi_k\}$ induced by $\mathcal{T}^\Lambda(\overline{(-d)})$ that d satisfies $d(i,j) \leq d(i,k) + d(j,k)$ if and only if $\mathrm{conv}\{2\chi_k, \chi_i + \chi_j\} \notin \mathcal{T}^\Lambda(\overline{(-d)})$ (see Figure 2). This indicates that in any coherent decomposition $d = d_1 + d_2$, if d is a metric, then both d_1 and d_2 are necessarily metrics.

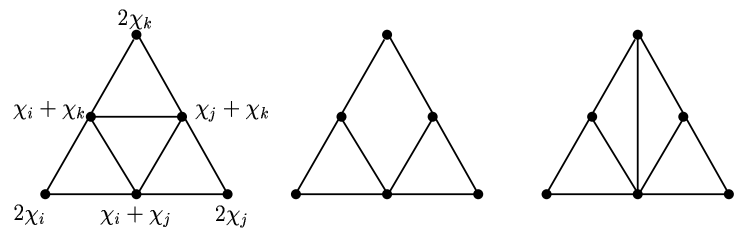

Figure 2: The role of triangle inequality: $d(i,j) < d(i,k) + d(j,k)$ (left), $d(i,j) = d(i,k) + d(j,k)$ (center) and $d(i,j) > d(i,k) + d(j,k)$ (right)

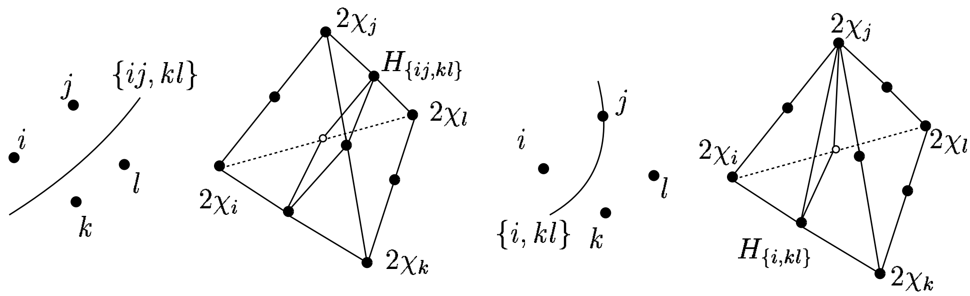

Figure 3: Examples of the split-hyperplane correspondence

Remark 4.6. Koolen, Moulton and Tönges [17] introduced the *coherency index* as a direct generalization of isolation index. For two metrics d' and d, the coherency index is a nonnegative value $\alpha_{d'}^{d}$ which has the property that $d = \alpha d' + (d - \alpha d')$ is a coherent decomposition if and only if $0 \leq \alpha \leq \alpha_{d'}^{d}$ [17, Theorem 4.1]. Hence we have $0 \leq \alpha \leq \alpha_{d'}^{d}$ $\Leftrightarrow P(d) = \alpha P(d') + P(d - \alpha d') \Leftrightarrow \overline{(-d)} = \alpha\overline{(-d')} + \overline{(d - \alpha d')} \Leftrightarrow 0 \leq \alpha \leq [\overline{(-d)} : \overline{(-d')}]$. This shows that coherency index is essentially equivalent to the quotient $[\cdot : \cdot]$ considered in this paper.

Next we give a derivation of the split decomposition of metrics (Theorem 4.1) using the result of the present paper. For this, we begin by establishing a relationship between splits and hyperplanes. For $A, B \subseteq V$ with $A, B \neq \emptyset$ and $A \cap B = \emptyset$, we call unordered pair $\{A, B\}$ a *partial split*. A hyperplane $H_{\{A,B\}}$ associated with partial split $\{A, B\}$ is defined as

$$H_{\{A,B\}} = \{x \in \mathbf{R}^V \mid x(A) = x(B)\}, \tag{4.9}$$

where we denote $\sum_{i \in A} x(i)$ by $x(A)$ for $x \in \mathbf{R}^V$ and $A \subseteq V$.

Lemma 4.7. *Let $\{A, B\}$ be a partial split on V and $C = V \setminus (A \cup B)$. Then we have*

$$\left\{ \begin{array}{lcl} \Lambda \cap H_{\{A,B\}} & = & \{\chi_i + \chi_j \mid i \in A, j \in B \text{ or } i, j \in C\}, \\ \Lambda \cap H_{\{A,B\}}^{+} & = & \{\chi_i + \chi_j \mid i, j \in A \cup C\}, \\ \Lambda \cap H_{\{A,B\}}^{-} & = & \{\chi_i + \chi_j \mid i, j \in B \cup C\}, \\ \Lambda \cap H_{\{A,B\}}^{++} & = & \{\chi_i + \chi_j \mid i \in A, j \in A \cup C\}, \\ \Lambda \cap H_{\{A,B\}}^{--} & = & \{\chi_i + \chi_j \mid i \in B, j \in A \cup C\}, \end{array} \right. \tag{4.10}$$

where we regard $H_{\{A,B\}}$ as $H_{\chi_A - \chi_B, 0}$.

Figure 3 illustrates small examples of the correspondence between partial splits and hyperplanes. In this figure, a partial split and an element of V are represented by a line and a point, respectively. The line corresponding to a partial split $\{A, B\}$ separates points of A and B and contains points of $V \setminus A \cup B$. The family $\mathcal{H}_\Lambda$ of linear hyperplanes, as defined in Definition 3.7, can be identified in terms of partial splits as follows.

Proposition 4.8. $\mathcal{H}_\Lambda = \{H_{\{A,B\}} \mid \{A, B\}$ *is a partial split on* $V\}$.

Proof. First we show that $H_{\{A,B\}}$ satisfies the conditions (0), (1), (2), and (3) of Lemma 3.6. Similarly to the previous lemma, we regard $H_{\{A,B\}}$ as $H_{\chi_A - \chi_B, 0}$. By $\Lambda \cap H^{\pm\pm}_{\{A,B\}} \neq \emptyset$, we have (0). In the condition (1), it suffices to show $H^{+}_{\{A,B\}} \cap \operatorname{conv} \Lambda \subseteq \operatorname{conv}(\Lambda \cap H^{+}_{\{A,B\}})$. Consider $x \in H^{+}_{\{A,B\}} \cap \operatorname{conv} \Lambda$. By definition, we have $x(A) \geq x(B)$ and $x \in \mathbf{R}^V_+$. If $x(B) = 0$, then we have $x(i) = 0$ for $i \in B$. Hence x can be represented as $x = \sum_{i \in V \setminus B} (x(i)/2)(2\chi_i)$ and therefore we have $x \in \operatorname{conv}(\Lambda \cap H^{+}_{\{A,B\}})$. If $x(B) \neq 0$, then there exists $i \in B$ with $x(i) > 0$. By $x(A) \geq x(B)$, there exists $j \in A$ with $x(j) > 0$. Let x' be defined by $x - \min(x(j), x(i))(\chi_i + \chi_j)$. Then the cardinality of nonzero elements of x' is smaller than that of x. Repeat this process for x'. After finitely many steps, we obtain the desired convex combination of x. In the similar way, we obtain (2) and (3). This implies ($\supseteq$). Conversely, let $H_{a,0} \in \mathcal{H}_\Lambda$. Define $A = \{i \in V \mid 2\chi_i \in H^{++}_{a,0}\}$ and $B = \{i \in V \mid 2\chi_i \in H^{--}_{a,0}\}$. Then $\{A, B\}$ is a partial split on V. We show $H_{a,0} = H_{\{A,B\}}$. By the definition of $\{A, B\}$, we have $a(i) = 0$ for $i \in V \setminus (A \cup B)$. For any $2\chi_i \in H^{++}_{a,0}$ and $2\chi_i \in H^{--}_{a,0}$, by Λ-admissibility of $\{H_{a,0}\}$, it is necessary that $\chi_i + \chi_j \in H_{a,0}$. Hence we have $a(i) = -a(j)$ for $i \in A$, $j \in B$. This implies $H_{a,0} = H_{\{A,B\}}$. □

For a partial split $\{A, B\}$, we define a *partial split distance* $\zeta_{\{A,B\}}$ associated with $\{A, B\}$ as

$$\zeta_{\{A,B\}}(i, j) = \begin{cases} 1 & \text{if } i \in A, j \in B \text{ or } i \in B, j \in A \\ 0 & \text{otherwise} \end{cases} \qquad (i, j \in V). \tag{4.11}$$

If $A \cup B \neq V$, then $\zeta_{\{A,B\}}$ is not a metric. If $A \cup B = V$, i.e., $\{A, B\}$ is a split, then $\zeta_{\{A,B\}}$ coincides with the split metric $\delta_{\{A,B\}}$.

The following shows that a partial split distance $\zeta_{\{A,B\}}$ associated with partial split $\{A, B\}$ is represented as split function $l_{H_{\{A,B\}}}$ associated with hyperplane $H_{\{A,B\}}$.

Proposition 4.9. *For a partial split* $\{A, B\}$*, the partial split distance* $\zeta_{\{A,B\}} : \Lambda \to \mathbf{R}$ *is represented as*

$$\zeta_{\{A,B\}}(x) = -|x(A) - x(B)|/2 + x(A \cup B)/2 \quad (x \in \Lambda). \tag{4.12}$$

Moreover, we have

$$\overline{(-\zeta_{\{A,B\}})}(x) = |x(A) - x(B)|/2 - x(A \cup B)/2 + \delta_{\mathbf{R}^V_+} \quad (x \in \mathbf{R}^V). \tag{4.13}$$

Proof. The equation (4.12) is obtained by direct calculations. The equation (4.13) follows from Λ-admissibility of $\{H_{\{A,B\}}\}$ (Proposition 4.8), Proposition 3.10, and Lemma 3.2 (1). □

The quotient $c_{H_{\{A,B\}}}(-d)$ can be explicitly calculated as follows.

Theorem 4.10. *For a distance function* d *and a partial split* $\{A, B\}$*, we have*

$$c_{H_{\{A,B\}}}(\overline{(-d)}) = \max\{0, \min\{\rho^d_{A,B}, \sigma^d_{A,B}, \sigma^d_{B,A}, \tau^d_{A,B}\}\}\sqrt{|A \cup B|}, \tag{4.14}$$

where

$$\rho^d_{A,B} = \frac{1}{2}\min\left\{\max\left\{\begin{array}{c} d(i,k)+d(j,l) \\ d(i,l)+d(j,k)\end{array}\right\} - d(i,j) - d(k,l) \;\middle|\; \begin{array}{c} i,j\in A \\ k,l\in B\end{array}\right\}, \quad (4.15)$$

$$\sigma^d_{A,B} = \frac{1}{2}\min\left\{d(i,k)+d(i,l)-d(k,l)-2d(i,j) \;\middle|\; \begin{array}{c} i\in A,\ k,l\in B \\ j\in V\setminus(A\cup B)\end{array}\right\}, \quad (4.16)$$

$$\tau^d_{A,B} = \min\left\{d(i,k)+d(j,l)-d(i,j)-d(k,l) \;\middle|\; \begin{array}{c} i\in A,\ k\in B \\ j,l\in V\setminus(A\cup B)\end{array}\right\} \quad (4.17)$$

and we define $\sigma^d_{A,B}=\infty$ *and* $\tau^d_{A,B}=\infty$ *if* $A\cup B = V$.

Proof. We apply the formulas (3.8) and (3.9) in Theorem 3.8. Let $C = V\setminus(A\cup B)$. Then $\tilde{c}_{H_S}(-d)$ is the minimum of

$$\frac{-d(i,j) - \overline{(-d)^{\Lambda\cap H_{\{A,B\}}}}(w)}{2l_{H_{\{A,B\}}}(\chi_i+\chi_j)} + \frac{-d(k,l) - \overline{(-d)^{\Lambda\cap H_{\{A,B\}}}}(w)}{2l_{H_{\{A,B\}}}(\chi_k+\chi_l)} \quad (4.18)$$

over $i\in A$, $j\in A\cup C$, $k\in B$, and $l\in B\cup C$, where $\{w\} = H_{\{A,B\}}\cap[\chi_i+\chi_j,\chi_k+\chi_l]$.

(Case 1: $i,j\in A$, $k,l\in B$). (4.18) is given by

$$\frac{\sqrt{|A\cup B|}}{2}\left\{-d(i,j)-d(k,l)-2\overline{(-d)^{\Lambda\cap H_S}}((\chi_i+\chi_j+\chi_k+\chi_l)/2)\right\}.$$

Furthermore $\overline{(-d)^{\Lambda\cap H_{\{A,B\}}}}((\chi_i+\chi_j+\chi_k+\chi_l)/2)$ is given by the optimal value of linear program

$$\begin{array}{ll} \text{min.} & \displaystyle\sum_{uv\in\{ik,jl,jk,jl\}} \lambda_{uv}(-d(u,v)) \\ \text{s.t.} & \displaystyle\sum_{uv\in\{ik,jl,jk,jl\}} \lambda_{uv}(\chi_u+\chi_v) = (\chi_i+\chi_j+\chi_k+\chi_l)/2, \\ & \lambda_{uv}\ge 0 \quad (uv\in\{ik,jl,jk,jl\}). \end{array}$$

By direct calculations, the optimal value of the problem is

$$\min\{-d(i,k)-d(j,l), -d(i,l)-d(j,k)\}/2.$$

Hence we have

$$(4.18) = \frac{\sqrt{|A\cup B|}}{2}\left\{-d(i,j)-d(k,l)+\max\left\{\begin{array}{c} d(i,k)+d(j,l) \\ d(i,l)+d(j,k)\end{array}\right\}\right\}. \quad (4.19)$$

(Case 2: $i\in A, j\in C, k,l\in B$). (4.18) is given by

$$\frac{\sqrt{|A\cup B|}}{2}\left\{-2d(i,j)-d(k,l)-3\overline{(-d)^{\Lambda\cap H_{\{A,B\}}}}((2\chi_i+2\chi_j+\chi_k+\chi_l)/3)\right\}.$$

In a similar way of (Case 1), we have

$$\overline{(-d)^{\Lambda\cap H_{\{A,B\}}}}((2\chi_i+2\chi_j+\chi_k+\chi_l)/3) = -(d(i,k)+d(i,l))/3.$$

Hence we obtain

$$(4.18) = \frac{\sqrt{|A\cup B|}}{2}\left\{-2d(i,j)-d(k,l)+d(i,k)+d(i,l)\right\}. \quad (4.20)$$

(Case 3: $i, j \in A, k \in B, l \in C$). It follows from (Case 2) by interchanging A and B.

(Case 4: $i \in A, k \in B, j, l \in C$). (4.18) is given by

$$\sqrt{|A \cup B|}\left\{-d(i,j) - d(k,l) - 2\overline{(-d)^{\Lambda \cap H_{\{A,B\}}}}((\chi_i + \chi_j + \chi_k + \chi_l)/2)\right\}.$$

In a similar way of (Case 1), we have

$$\overline{(-d)^{\Lambda \cap H_{\{A,B\}}}}((\chi_i + \chi_j + \chi_k + \chi_l)/2) = -(d(i,k) + d(j,l))/2.$$

Hence, we obtain

$$(4.18) = \sqrt{|A \cup B|}\left\{-d(i,j) - d(k,l) + d(i,k) + d(j,l)\right\}. \tag{4.21}$$

Combining (4.19), (4.20), and (4.21), we obtain the desired formula (4.14). □

In particular, if a partial split $\{A, B\}$ forms a split, then we have

$$c_{H_{\{A,B\}}}(\overline{-d}) = \alpha^d_{\{A,B\}}\sqrt{|V|} \tag{4.22}$$

(see the definitions of $\alpha^d_{\{A,B\}}$ in (4.3) and $\rho^d_{A,B}$ in (4.15)). Accordingly, the isolation index can be extended for a partial split $\{A, B\}$ by defining

$$\alpha^d_{\{A,B\}} = \max\{0, \min\{\rho^d_{A,B}, \sigma^d_{A,B}, \sigma^d_{B,A}, \tau^d_{A,B}\}\}. \tag{4.23}$$

Similarly, $\mathcal{S}(d)$ can be extended for partial splits by defining

$$\mathcal{S}(d) = \{S : \text{partial split on } V \mid \alpha^d_S > 0\}. \tag{4.24}$$

As a consequence of above arguments, we obtain an extension of Bandelt-Dress' split decomposition (Theorem 4.1).

Theorem 4.11. *Every distance function $d : V \times V \to \mathbf{R}$ can be coherently decomposed as*

$$d = \sum_{S \in \mathcal{S}(d)} \alpha^d_S \zeta_S + d', \tag{4.25}$$

where d' is a distance function with $\alpha^{d'}_{S'} = 0$ for any partial split S'.

Proof. Decompose $-d$ into the form (3.6) in Theorem 3.5 and apply Propositions 4.8 and 4.9, Theorem 4.10 and Lemma 4.3. Then we obtain (4.25). The coherency of the decomposition (4.25) follows from $Q(-d) = -P(d)$ and Remark 3.11. □

The cardinality of $\mathcal{S}(d)$ is bounded by $|V|(|V|-1)/2$ by Proposition 3.12. Therefore the decomposition (4.25) is also obtained in polynomial time by an algorithm similar to that for the split decomposition of a metric [1]. It is easily observed that if d satisfies the triangle inequality, $\tau^d_{\{A,B\}} \leq 0$ holds for any proper partial split $\{A, B\}$ with $A \cup B \neq V$. This implies that $\mathcal{S}(d)$ consists of splits. Hence if d is a metric, the decomposition (4.25) coincides with the split decomposition of metrics (4.5).

Remark 4.12. By the decomposition (4.25), $P(d)$ is decomposed as

$$P(d) = Z(d) + P(d'),$$

where $Z(d)$ is a zonotope defined as

$$Z(d) = \sum_{\{A,B\} \in \mathcal{S}(d)} \alpha^d_{\{A,B\}}([\chi_A - \chi_B, \chi_B - \chi_A]/2 + \chi_{A \cup B}/2)$$

(see Remark 3.11). If $d' = 0$. we have $P(d) = Z(d) + \mathbf{R}^V_+$. In this case, tight span $T(d)$ is the union of the faces of $Z(d)$ whose normal cone contains negative vectors.

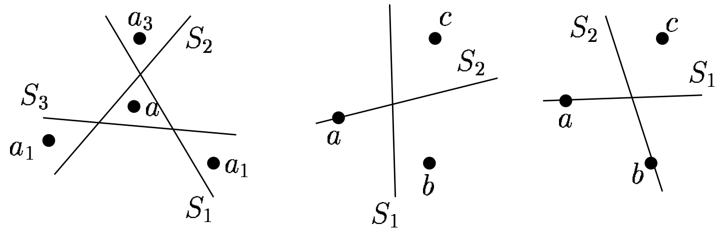

Figure 4: Forbidden partial splits:(C1)(left), (C2)(center) and (C3)(right)

Remark 4.13. It is well known that every 4-point metric is totally split decomposable, i.e., a split-prime residue d' of (4.5) vanishes. It is easily seen that every 3-point distance function is also totally split decomposable (in our sense), i.e., $d' = 0$ in the decomposition (4.25). However, 4-point distance function d on $\{i, j, k, l\}$ defined as

$$d = \begin{array}{|c|cccc|} \hline & i & j & k & l \\ \hline i & 0 & 1 & 1 & 0 \\ j & 1 & 0 & 0 & 1 \\ k & 1 & 0 & 0 & 0 \\ l & 0 & 1 & 0 & 0 \\ \hline \end{array} \tag{4.26}$$

is (partial) split prime, i.e., $\alpha^d_{\{A,B\}} = 0$ for any partial split $\{A, B\}$. This demonstrates that not every 4-point distance function is totally split decomposable.

Next we characterize Λ-admissibility of hyperplanes in terms of combinatorial properties of the corresponding partial splits. First we observe that $\mathcal{S}(d)$ does not contain the following types of partial splits (see Figure 4):

(C1) three partial splits $\{A_1, B_1\}$, $\{A_2, B_2\}$, $\{A_3, B_3\}$ and four points $a, a_1, a_2, a_3 \in V$ such that $\{a, a_1, a_2, a_3\} \cap A_i = \{a, a_i\}$ for $i = 1, 2, 3$ (the violation of weak compatibility).

(C2) two partial splits $\{A_1, B_1\}$, $\{A_2, B_2\}$ and three points $a, b, c \in V$ such that $a \in A_1$, $b, c \in B_1$, $b \in A_2$, $c \in B_2$, and $a \in V \setminus (A_2 \cup B_2)$.

(C3) two partial splits $\{A_1, B_1\}$, $\{A_2, B_2\}$ and three points $a, b, c \in V$ such that $b \in A_1$, $c \in B_1$, $a \in V \setminus (A_1 \cup B_1)$, $a \in A_2$, $b \in B_2$, and $c \in V \setminus (A_2 \cup B_2)$.

For (C1), observe that not all three $\rho^d_{A_1,B_1}$, $\rho^d_{A_2,B_2}$ and $\rho^d_{A_3,B_3}$ are positive. Similarly, for (C2), observe that both $\rho^d_{A_1,B_1}$ and $\tau^d_{A_2,B_2}$ are not positive. To see (C3), observe that both $\tau^d_{A_1,B_1}$ and $\tau^d_{A_2,B_2}$ are not positive. A collection of partial splits free from (C1), (C2) and (C3) is an extension of weakly compatible collection of splits. The following is an extension of [1, Theorem 3].

Theorem 4.14. *Let $\mathcal{S}$ be a collection of partial splits on V which does not contain partial splits of the types (C1), (C2) and (C3). For $\lambda \in \mathbf{R}^{\mathcal{S}}_{++}$, let a distance function $d : V \times V \to \mathbf{R}$ be defined as $d = \sum_{S \in \mathcal{S}} \lambda_S \zeta_S$. Then we have $\mathcal{S} = \mathcal{S}(d)$ and $\lambda_S = \alpha^d_S$ for each $S \in \mathcal{S}$.*

Proof. We adapt the proof of [1, Theorem 3]. For a partial split $\{A, B\}$, let $\tilde{\alpha}^d_{\{A,B\}}$ be defined as

$$\tilde{\alpha}^d_{\{A,B\}} = \min\{\rho^d_{A,B}, \sigma^d_{A,B}, \sigma^d_{B,A}, \tau^d_{A,B}\}.$$

It suffices to show

$$\tilde{\alpha}^d_S \geq \lambda_S \quad (S \in \mathcal{S}). \tag{4.27}$$

By the formula of α^d_S for a partial split $S = \{A, B\}$, there exists $Y \subseteq V$ with $A \cap Y \neq \emptyset$, $B \cap Y \neq \emptyset$, and $|Y| \leq 4$ such that

$$\tilde{\alpha}^d_{\{A,B\}} = \tilde{\alpha}^{d_Y}_{\{A\cap Y, B\cap Y\}},$$

where $d_Y : Y \times Y \to \mathbf{R}$ denotes the restriction of d to Y. Hence it is sufficient to show (4.27) for the case $|V| \leq 4$.

In the case of $|V| = 2$, (4.27) is obvious. In the case of $|V| = 3$, let $V = \{i, j, k\}$. For the simplicity of notation, we denote a partial split $\{\{i, j\}, \{k\}\}$ by $\{ij, k\}$. It suffices to check (4.27) for $S = \{ij, k\}$ and $S = \{i, j\}$. For $S = \{ij, k\}$, $\mathcal{S}$ does not contain $\{i, j\}$ by (C2). Hence we have $\tilde{\alpha}^d_{\{ij,k\}} = \rho^d_{ij,k} = \min\{(d(i, k) + d(j, k) - d(i, j))/2, d(i, k), d(j, k)\} \geq \lambda_{\{ij,k\}}$. For $S = \{i, j\}$, $\mathcal{S}$ does not contain $\{j, k\}$, $\{i, k\}$ and $\{ij, k\}$ from (C2) and (C3). We have $\tau^d_{i,j} = d(i, j) - d(i, k) - d(j, k) = \lambda_{\{i,j\}}$, $\rho^d_{i,j} = d(i, j) \geq \lambda_{\{i,j\}}$, $\sigma^d_{i,j} \geq \tau^d_{i,j}$ and $\sigma^d_{j,i} \geq \tau^d_{i,j}$. Hence we obtain $\tilde{\alpha}^d_{\{i,j\}} \geq \lambda_{\{i,j\}}$. Therefore (4.27) holds for $|V| = 3$.

In the case of $|V| = 4$, let $V = \{i, j, k, l\}$. We may assume that the minimum of $\tilde{\alpha}^d_S$ is attained by all different four points (otherwise it can be reduced to the case $|V| \leq 3$). It suffices to check (4.27) for $S = \{ij, kl\}$, $S = \{ij, k\}$ and $S = \{i, j\}$. For $S = \{ij, kl\}$, we have

$$\tilde{\alpha}^d_{\{ij,kl\}} = \rho^d_{ij,kl} = \{\max\{d(i, k) + d(j, l), d(i, l) + d(j, k)\} - d(i, j) - d(k, l)\}/2. \tag{4.28}$$

From the condition (C1), $\mathcal{S}$ does not contain both $\{ik, jl\}$ and $\{il, jk\}$ simultaneously. Suppose that $\{ik, jl\} \in \mathcal{S}$ and $\{il, jk\} \notin \mathcal{S}$. By the conditions (C2) and (C3), the possible partial splits in $\mathcal{S}$ that contribute to the term $d(i, j) + d(k, l)$ are only five partial splits $\{i, jkl\}$, $\{j, ikl\}$, $\{k, ijl\}$, $\{l, ijk\}$, and $\{ik, jl\}$. First four partial splits contribute equally to $d(i, k) + d(j, l)$, $d(i, l) + d(j, k)$ and $d(i, j) + d(k, l)$. Hence we have $\tilde{\alpha}^d_{\{ij,kl\}} \geq \max\{\lambda_{\{ij,kl\}}, \lambda_{\{ij,kl\}} + \lambda_{\{ik,jl\}}\} - \lambda_{\{ik,jl\}} = \lambda_{\{ij,kl\}}$. For $S = \{ij, k\}$, we have

$$\tilde{\alpha}^d_{\{ij,k\}} = \sigma^d_{k,ij} = \{d(i, k) + d(j, k) - d(i, j) - 2d(k, l)\}/2. \tag{4.29}$$

From the conditions (C2) and (C3), the possible partial splits in $\mathcal{S}$ that contribute to the term $d(i, j) + 2d(k, l)$ are four partial splits $\{i, jkl\}$, $\{j, ikl\}$, $\{k, ijl\}$, and $\{ik, j\}$. However it is easily examined that these partial split distances cancel out in (4.29). Therefore we obtain $\tilde{\alpha}^d_{\{ij,k\}} \geq \lambda_{\{ij,k\}}$. For $S = \{i, j\}$, we may assume

$$\tilde{\alpha}^d_{\{i,j\}} = \tau^d_{i,j} = d(i, j) + d(k, l) - d(i, k) - d(j, l). \tag{4.30}$$

From the conditions (C2) and (C3), the possible partial splits in $\mathcal{S}$ that contribute to the term $d(i, k) + d(j, l)$ are five partial splits $\{i, jkl\}$, $\{j, ikl\}$, $\{i, jk\}$, $\{j, il\}$ and $\{il, jk\}$. But these partial split distances cancel out in (4.30). Therefore we obtain $\tilde{\alpha}^d_{\{i,j\}} \geq \lambda_{\{i,j\}}$. Hence we conclude that (4.27) holds for $|V| = 4$. $\square$

By Theorem 4.14 and Proposition 3.10, weak compatibility of a collection of splits is also characterized by Λ-admissibility of the corresponding collection of hyperplanes.

Theorem 4.15. *For a collection of partial splits $\mathcal{S}$, the following conditions* (a) *and* (b) *are equivalent.*

(a) *$\mathcal{S}$ does not contain partial splits of the types (C1),(C2), and (C3).*

(*b*) $\{H_S \mid S \in \mathcal{S}\}$ *is* Λ*-admissible.*

In particular, if $\mathcal{S}$ *consists of splits, then* $\mathcal{S}$ *is weakly compatible if and only if* $\{H_S \mid S \in \mathcal{S}\}$ *is* Λ*-admissible.*

Remark 4.16. The fundamental fact [1, Corollary 4] that the number of weakly compatible splits is bounded by $|V|(|V|-1)/2$ also follows from Proposition 3.12. It is shown that this bound is attained by the maximum *circular split system*, which is obtained from a convex $|V|$-gon [1, Theorem 5]. By Corollary 3.13, the sum of maximum circular split metrics yields a triangulation of $\mathrm{conv}\,\Lambda$. We point out that this construction of a triangulation of $\mathrm{conv}\,\Lambda$ is essentially equivalent to the construction of the triangulation of the second hypersimplex $\mathrm{conv}\{\chi_i + \chi_j \mid i, j \in V, i \neq j\}$ due to de Loera, Sturmfels, and Thomas [7] (see also [23, Chapter 9]).

A collection of splits $\mathcal{S}$ is said to be *compatible* if for any pair of splits $\{A, B\}, \{C, D\} \in \mathcal{S}$, at least one of four sets $A \cap C$, $A \cap D$, $B \cap C$ and $B \cap D$ is empty (see [4],[2],[22]). Compatibility of a collection of splits can also be captured as a geometric property of the corresponding collection of hyperplanes.

Proposition 4.17. *For a collection of splits* $\mathcal{S}$*, the following conditions* (*a*) *and* (*b*) *are equivalent.*

(*a*) $\mathcal{S}$ *is compatible.*

(*b*) $H_1 \cap H_2 \cap \mathrm{ri\,conv}\,\Lambda = \emptyset$ *holds for each pair* $H_1, H_2 \in \{H_S \mid S \in \mathcal{S}\}$.

Proof. For two splits $\{A, B\}, \{C, D\} \in \mathcal{S}$, consider two hyperplanes $H_{\{A,B\}}$ and $H_{\{C,D\}}$. The nonemptiness of $H_{\{A,B\}} \cap H_{\{C,D\}} \cap \mathrm{ri\,conv}\,\Lambda$ is equivalent to the existence of a solution to linear equality, inequality system

$$\left\{ \begin{array}{rcl} x(A) - x(B) & = & 0, \\ x(C) - x(D) & = & 0, \\ x(i) & > & 0 \quad (i \in V). \end{array} \right. \tag{4.31}$$

$(a) \Rightarrow (b)$. By compatibility of splits $\mathcal{S}$, we may assume $A \subset C$ and $D \subset B$. By subtracting the second of (4.31) from the first. we have $x(B \cap C) = 0$ and hence (4.31) is empty. $(b) \Rightarrow (a)$. Suppose that $\{A, B\}$ and $\{C, D\}$ are incompatible, i.e., all the four sets $A \cap C$, $A \cap D$, $B \cap C$ and $B \cap D$ are nonempty. Then $x \in \mathbf{R}^V$ defined by

$$x(i) = \left\{ \begin{array}{lll} 1/|A \cap C| & \text{if} & i \in A \cap C \\ 1/|A \cap D| & \text{if} & i \in A \cap D \\ 1/|B \cap C| & \text{if} & i \in B \cap C \\ 1/|B \cap D| & \text{if} & i \in B \cap D \end{array} \right. \tag{4.32}$$

is a solution to (4.31). □

A metric d is a *tree metric* if it is represented as the path metric of some weighted tree. It is well known that d is a tree metric if and only if it is represented as

$$d = \sum_{S \in \mathcal{S}} \alpha_S \delta_S \tag{4.33}$$

for some compatible collection of splits $\mathcal{S}$ and a positive weight $\alpha \in \mathbf{R}^{\mathcal{S}}_{++}$ (see [4],[2],[22]). From Propositions 2.17 and 4.17, one of the central theorems in T-theory can be derived.

Theorem 4.18 (Dress [8]). *A metric* d *is a tree metric if and only if* $T(d)$ *is a tree.*

Acknowledgment

The author thanks Kazuo Murota for helpful comments.

A Proof of Proposition 2.6

We may assume that f is represented as (2.1). $F = \operatorname{argmin} f[-p] \in \mathcal{T}(f)$ is given by the projection of the lower face of $\operatorname{epi} f$ that is represented as the set of optimal solutions of the linear program

$$\begin{aligned} \mathrm{LP}(p): \text{Maximize} \quad & \langle p, y\rangle - z \\ \text{subject to} \quad & \langle a_j, y\rangle \le b_j \quad (j \in J), \\ & \langle p_i, y\rangle - z \le \beta_i \quad (i \in I). \end{aligned}$$

Then there exist $I^* \subseteq I$ and $J^* \subseteq J$ such that the set of optimal solutions of $\mathrm{LP}(p)$ is given by the solution of the linear inequality

$$\begin{cases} \langle a_j, y\rangle = b_j & (j \in J^*) \\ \langle a_j, y\rangle \le b_j & (j \in J \backslash J^*) \\ \langle p_i, y\rangle - z = \beta_i & (i \in I^*) \\ \langle p_i, y\rangle - z \le \beta_i & (i \in I \backslash I^*) \end{cases} \tag{A.1}$$

and its relative interior is given by (A.1) with $\le$ replaced by $<$. Then $x \in \operatorname{ri} F$ if and only if $(x, f(x))$ satisfies (A.1) with strict inequalities. By Lemma 2.2 we have

$$\forall x \in \operatorname{ri} F,\ \partial f(x) = \operatorname{conv}\{p_i \mid i \in I^*\} + \operatorname{cone}\{a_j \mid j \in J^*\} = F^{\bullet}. \tag{A.2}$$

Hence $F^{\bullet}$ is well-defined. This establishes (1).

Next we show (2). $F^{\bullet\bullet} = \partial f^{\bullet}(p^*) = \operatorname{argmin} f[-p^*]$ with $p^* \in \operatorname{ri} F^{\bullet}$. Therefore $F^{\bullet\bullet}$ is also represented by the optimal solutions of $\mathrm{LP}(p^*)$. By (A.2), we have

$$p^* = \sum_{i \in I^*} \lambda_i p_i + \sum_{j \in J^*} \mu_j a_j,\ \sum_{i \in I^*} \lambda_i = 1,\ \lambda_i > 0\ (i \in I^*),\ \mu_j > 0\ (j \in J^*).$$

Hence (λ, μ) is a feasible solution to the dual program of $\mathrm{LP}(p^*)$,

$$\begin{aligned} \mathrm{DLP}(p^*): \text{Maximize} \quad & \sum_{i \in I} \lambda_i \beta_i + \sum_{j \in J} \mu_j b_j \\ \text{subject to} \quad & p^* = \sum_{i \in I} \lambda_i p_i + \sum_{j \in J} \mu_j a_j, \\ & \lambda_i \ge 0\ (i \in I), \quad \mu_j \ge 0\ (j \in J). \end{aligned}$$

Let (x, z) be a solution to the linear inequalities (A.1) with strict inequalities. Then (x, z) and (λ, μ) satisfy the strict complementary slackness condition of $\mathrm{LP}(p^*)$ and $\mathrm{DLP}(p^*)$. Therefore, the set of the optimal solutions of $\mathrm{LP}(p^*)$ coincides with the set of the solutions of (A.1). This means $F = F^{\bullet\bullet}$ and we obtain (2).

(3). By (A.1) and (A.2), the vector subspaces parallel to affine spaces $\operatorname{aff} F$ and $\operatorname{aff} F^{\bullet}$ are given by

$$\begin{aligned} \operatorname{aff} F - \{x\} &= \{y \mid \langle a_j, y\rangle = 0\ (j \in J^*),\ \langle p_i, y\rangle - z = 0\ (i \in I^*)\}, \\ \operatorname{aff} F^{\bullet} - \{q\} &= \left\{ p \;\middle|\; p = \sum_{i \in I^*} \lambda_i p_i + \sum_{j \in J^*} \mu_j a_j,\ \sum_{i \in I^*} \lambda_i = 0 \right\}. \end{aligned}$$

By the well-known relation $(\mathrm{Ker}\, A)^{\perp} = \mathrm{Im}\, A^{\top}$ for a matrix A, we obtain (3).

(4). By the above discussions we may assume that $F = \mathrm{argmin}\, f[-p]$ and $G = \mathrm{argmin}\, f[-q] \in \mathcal{T}(f)$ are represented respectively by

$$\begin{cases} \langle a_j, y\rangle = b_j & (j \in J_p^*) \\ \langle a_j, y\rangle \leq b_j & (j \in J\backslash J_p^*) \\ \langle p_i, y\rangle - z = \beta_i & (i \in I_p^*) \\ \langle p_i, y\rangle - z \leq \beta_i & (i \in I\backslash I_p^*) \end{cases}, \quad \begin{cases} \langle a_j, y\rangle = b_j & (j \in J_q^*) \\ \langle a_j, y\rangle \leq b_j & (j \in J\backslash J_q^*) \\ \langle p_i, y\rangle - z = \beta_i & (i \in I_q^*) \\ \langle p_i, y\rangle - z \leq \beta_i & (i \in I\backslash I_q^*). \end{cases}$$

Hence $F \subseteq G \Leftrightarrow J_q^* \subseteq J_p^*$ and $I_q^* \subseteq I_p^* \Rightarrow G^\bullet \subseteq F^\bullet$.

References

[1] H.-J. Bandelt and A. W. M. Dress: A canonical decomposition theory for metrics on a finite set, *Advances in Mathematics* **92** (1992), no. 1, 47–105.

[2] J.-P. Barthélémy and A. Guénoche: *Trees and Proximity Representations*, Translated from the French by Gregor Lawden, Wiley, Chichester, 1991.

[3] E. D. Bolker: A class of convex bodies, *Transactions of the American Mathematical Society* **145** (1969), 323–345

[4] P. Buneman: The recovery of trees from measures of dissimilarity, *Mathematics in the Archaeological and Historical Sciences* (F. R. Hodson, D. G. Kendall, P. Tautu, eds.) Edinburgh University Press, Edinburgh, 1971, 387–395.

[5] S. Chynoweth: Generalised Voronoi/Delaunay duality, with applications, *Zeitschrift für Angewandte Mathematik und Mechanik* **76** (1996), no. 3, 339–342.

[6] S. Chynoweth and M. J. Sewell: Mesh duality and Legendre duality, *Proceedings of the Royal Society London Series A* **428** (1990), no. 1875, 351–377.

[7] J. A. de Loera, B. Sturmfels and R. R. Thomas: Gröbner bases and triangulations of the second hypersimplex, *Combinatorica* **15** (1995), no. 3, 409–424.

[8] A. W. M. Dress: Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: a note on combinatorial properties of metric spaces, *Advances in Mathematics* **53** (1984), no. 3, 321–402.

[9] A. W. M. Dress: Towards a classification of transitive group actions on finite metric spaces, *Advances in Mathematics* **74** (1989), no. 2, 163–189.

[10] A. Dress, V. Moulton and W. Terhalle: T-theory: an overview, *European Journal of Combinatorics* **17** (1996), no. 2-3, 161–175.

[11] A. Frank: An algorithm for submodular functions on graphs, *Annals of Discrete Mathematics* **16** (1982), 97–120.

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

[13] S. Fujishige: *Submodular Functions and Optimization*, Annals of Discrete Mathematics, **47**, North-Holland, Amsterdam, 1991.

[14] I. M. Gel'fand, M. M. Kapranov and A. V. Zelevinsky: *Discriminants, Resultants, and Multidimensional Determinants*, Birkhäuser, Boston, MA, 1994.

[15] H. Hirai and K. Murota: M-convex functions and tree metrics, METR 2003-23, University of Tokyo, July 2003.

[16] J. R. Isbell: Six theorems about injective metric spaces, *Commentarii Mathematici Helvetici* **39** (1964), 65–76.

[17] J. Koolen, V. Moulton and U. Tönges: The coherency index, *Discrete Mathematics* **192** (1998), no. 1-3, 205–222.

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

[19] K. Murota: *Discrete Convex Analysis*, SIAM, Philadelphia, PA, 2003.

[20] R. T. Rockafellar: *Convex Analysis*, Princeton University Press, Princeton, N.J., 1970.

[21] R. Schneider: *Convex Bodies: the Brunn-Minkowski theory*, Cambridge University Press, Cambridge, 1993.

[22] C. Semple and M. Steel: *Phylogenetics*, Oxford University Press, Oxford, 2003.

[23] B. Sturmfels: *Gröbner Bases and Convex Polytopes*, American Mathematical Society, Providence, RI, 1996.

[24] B. Sturmfels and J. Yu: Classification of six-point metrics, `arXiv:math.MG/0403147`.

[25] G. M. Ziegler: *Lectures on Polytopes*, Springer-Verlag, New York, 1995.