MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

The Buneman Index via Polyhedral Split Decomposition

Shungo KOICHI

(Communicated by Satoru IWATA)

METR 2006–57 November 2006

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.i.u-tokyo.ac.jp/mi/mi-e.htm

The METR technical reports are published as a means to ensure timely dissemination of scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the authors or by other copyright holders, notwithstanding that they have offered their works here electronically. It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may not be reposted without the explicit permission of the copyright holder.

The Buneman Index via Polyhedral Split Decomposition

Shungo KOICHI*

November, 2006

Abstract

The Buneman index and Bandelt and Dress' isolation index are two well-known tools for constructing a phylogenetic tree from dissimilarity data. A recent paper of Hirai (2006) provides a geometric interpretation of the isolation index by deriving Bandelt and Dress' split decomposition of metrics as a special case of the *polyhedral split decomposition* of polyhedral convex functions in the following way. A finite metric is regarded as a discrete function on a certain type of vector configuration and extended to a polyhedral convex function, which is called its convex extension. Then, the isolation index appears in the polyhedral split decomposition of the convex extension. This paper shows that the same approach works for the Buneman index by taking a different type of vector configuration, namely, Buneman's result is also understood as a polyhedral split decomposition.

By polyhedral split decomposition, a polyhedral convex function can be uniquely represented as a sum of *split functions* and a *residue*. Roughly, a *split-decomposable* function is a discrete function such that the residue of its convex extension is a linear function, and a *split fan* is a simplicial fan consisting of split-decomposable functions. In the case of the Buneman index, the split-decomposable functions coincide with *tree metrics*, and the split fan is essentially identical with the *space of phylogenetic trees*. A geometric property of a split-decomposable function and a vector configuration as its domain is revealed in the paper of Hirai. We newly give a combinatorial characterization for the split-decomposable functions on some type of vector configuration by using the matroid associated with the vector configuration. The combinatorial characterization implies pairwise compatibility of splits arising from a tree metric.

Keywords: metric, the Buneman index, polyhedral split decomposition, the space of phylogenetic trees, vector configuration, hyperplane arrangement, semimodular lattice.

1 Introduction

The problem of reconstructing a tree, called a *phylogenetic tree*, from dissimilarity (or distance) data on biological sequences, e.g., DNA or amino acid sequences is the most fundamental and important issue in phylogeny. By using various alignment methods for biological sequences, we can measure dissimilarities between them. In phylogeny, Buneman's method [8] and Bandelt and Dress' method [1] are well known as tree reconstruction methods. The two methods utilize the *Buneman index* and *isolation index*, respectively, in order to obtain clues to reconstruct a phylogenetic tree. Since the two indices are very similar to each other, they are often compared. The aim of this paper is to reveal the relation between the two methods, especially the two indices.

In order to review the two methods briefly, we classify *dissimilarity maps*, *metrics*, and *tree metrics*. Let X be a set of objects, e.g., sequences or taxa. A nonnegative dissimilarity map is defined as a function $d : X \times X \to \mathbb{R}$ such that $d(i,i) = 0$ for all $i \in X$ and $d(i,j) = d(j,i) \geq 0$

*Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan. E-mail: `Syungo_Koichi@mist.i.u-tokyo.ac.jp`

for all $i, j \in X$. A metric is a nonnegative dissimilarity map satisfying the triangle inequality, i.e., $d(i,j) \le d(i,k) + d(k,j)$ for all $i, j, k \in X$. A tree metric is the path metric of a tree with nonnegative edge weights.

The clues to reconstruct trees are brought as a set of splits. A *split* of X is a partition of X into two non-empty sets. For a metric $d : X \times X \to \mathbb{R}$ and a split $\{A, B\}$, $A, B \subseteq X$, the Buneman index is defined by

$$b^d_{\{A,B\}} = \frac{1}{2} \min_{u,v \in A, x,y \in B} \left\{ \min \left\{ \begin{matrix} d(u,x) + d(v,y), \\ d(u,y) + d(v,x) \end{matrix} \right\} - d(u,v) - d(x,y) \right\},$$

and the isolation index is defined by

$$i^d_{\{A,B\}} = \frac{1}{2} \min_{u,v \in A, x,y \in B} \left\{ \max \left\{ \begin{matrix} d(u,x) + d(v,y), \\ d(u,y) + d(v,x), \\ d(u,v) + d(x,y) \end{matrix} \right\} - d(u,v) - d(x,y) \right\}.$$

By computing the Buneman indices or isolation indices for splits of X, we obtain the set of splits:

$$\Sigma_b(d) = \{\sigma \mid \sigma : \text{a split of } X, b^d_\sigma > 0\}$$

or

$$\Sigma_i(d) = \{\sigma \mid \sigma : \text{a split of } X, i^d_\sigma > 0\}.$$

Many methods to reconstruct a tree or graph from such a set of splits are known [12, 13, 19, 20]. Those methods are based on Splits Equivalence theorem to be described in Section 2. In addition, Bandelt and Dress obtain a decomposition of a metric with the isolation index. The *split metric* $\xi_{\{A,B\}} : X \times X \to \{0, 1\}$ associated with a split $\{A, B\}$ of X is defined by

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

for all $i, j \in X$. By Bandelt and Dress' split decomposition of a metric, a metric d can be decomposed as follows:

$$d = \sum_{\sigma \in \Sigma_i(d)} i^d_\sigma \xi_\sigma + d', \tag{1.1}$$

where d' is a metric with $i^{d'}_{\sigma'} = 0$ for any split σ'. We call d' the *split-prime residue* of d.

In order to extend the results of Bandelt and Dress, Hirai [16, 17] introduces the polyhedral split decomposition of polyhedral convex functions, which is summarized as follows. A function of which the epigraph is a convex polyhedron is called a polyhedral convex function. A polyhedral convex function f on $\mathbb{R}^n$ can be decomposed as

$$f(x) = \sum_{(a,r) \in \mathbb{R}^n \times \mathbb{R}} c^f_{a,r} |\langle a, x \rangle - r| + f'(x) \quad (x \in \mathbb{R}^n),$$

where $c^f_{a,r}$ is the nonnegative number defined by $\sup\{t \ge 0 \mid f(x) - t|\langle a, x \rangle - r| \text{ is convex in } x\}$ and f' is a polyhedral convex function such that $c^{f'}_{a,r} \in \{0, +\infty\}$ for any $(a, r) \in \mathbb{R}^n \times \mathbb{R}$. A function $|\langle a, x \rangle - r|$ of x is called a *split function*. Hirai derives Bandelt and Dress' isolation index geometrically by polyhedral split decomposition.

We briefly explain Hirai's results in [16, 17]. A function is said to be discrete if it is defined on a finite set of points/vectors in $\mathbb{R}^n$. Let $X = \{1, 2, \ldots, n\}$. For $A \subseteq X$, we denote by χ_A the characteristic vector of A defined by $\chi_A(i) = 1$ if $i \in A$ and $\chi_A(i) = 0$ if $i \notin A$. In particular, we write χ_i instead of $\chi_{\{i\}}$ for each $i \in X$.

A nonnegative dissimilarity map d can be regarded as a discrete function d on the point set $\Lambda = \{\chi_i + \chi_j \mid i, j \in X\}$ by the correspondence: $d(\chi_i + \chi_j) \leftarrow d(i, j)$. By using a technique in discrete convex analysis [23], we obtain the convex extension of the discrete function $-d$ on Λ as follows:

$$\overline{(-d)}(x) = \sup_{p \in \mathbb{R}^n} \left\{ \langle p, x \rangle \mid p(i) + p(j) \leq -d(\chi_i + \chi_j) \ (i, j \in X) \right\} \quad (x \in \mathbb{R}^n). \tag{1.2}$$

This $\overline{(-d)}$ is a polyhedral convex function. Thus, by convex extension, we can interpret a dissimilarity map as a geometric object. In the case that d is a metric, split functions appearing in the polyhedral split decomposition of $\overline{(-d)}$ is restricted to those such that $(a, r) = (\chi_A - \chi_B, 0)$ for some split $\{A, B\}$ of X. This type of split function can be regarded as a split metric. Moreover, in this case, $c^f_{a,r}$ corresponds to the isolation index, namely, the polyhedral split decomposition of $\overline{(-d)}$ results in Bandelt and Dress' split decomposition of a metric as shown by (1.1).

In this paper, we derive the Buneman index in the same manner as Hirai, i.e., by polyhedral split decomposition. The only difference between the derivations of the two indices is a discrete function we compose. Our approach is summarized as follows. A metric $d : X \times X \to \mathbb{R}$ can be regarded as a discrete function d on the point set $\Omega = \{\chi_i - \chi_j \mid i, j \in X\}$ by the correspondence: $d(\chi_i - \chi_j) = d(\chi_j - \chi_i) \leftarrow d(i, j)$. The convex extension of d on Ω is as follows:

$$\overline{d}(x) = \sup_{p \in \mathbb{R}^n} \left\{ \langle p, x \rangle \mid p(i) - p(j) \leq d(\chi_i - \chi_j) \ (i, j \in X) \right\} \quad (x \in \mathbb{R}^n). \tag{1.3}$$

This $\overline{d}$ is also a polyhedral convex function. Hence, the polyhedral split decomposition can be applied to $\overline{d}$ with some additional modification of Hirai's decomposition. As a result, split metrics also appear as split functions in the decomposition of $\overline{d}$ and $c^f_{a,r}$ corresponds to the Buneman index. Therefore, we conclude that Buneman's method can be understood as a polyhedral split decomposition of metrics.

We here refer to the dual representation of a polyhedral split decomposition. From (1.2) and (1.3), we realize each of $\overline{(-d)}$ and $\overline{d}$ is the support function of some polyhedron. For the support function of a polyhedron P, the dual operation of the polyhedral split decomposition is to extract line segments from P. As a result, P is decomposed as the Minkowski sum of a *zonotope* Z, which is the Minkowski sum of line segments, and some polyhedron P'; see [17, §2.3]. If P has a vertex, Z can be uniquely defined and it is called the *maximum zonotopic summand* of P. This kind of decomposition of polyhedra is originally due to Bolker [6]. Bandelt and Dress' approach is actually based on a similar perspective, that is, they propose the *coherent* decomposition of the polyhedron supported by $\overline{(-d)}$. Hirai's polyhedral split decomposition can be considered as an extension of Bolker's result to unbounded polyhedra in this context.

In Section 8, we introduce two interesting notions "split-decomposability" and "split fan" suggested by Hirai. A discrete function $g : K \to \mathbb{R}$ is split-decomposable if its convex extension $\overline{g}$ can be decomposed as a sum of split functions and a linear function. The set of all split-decomposable functions on K can be regarded as a simplicial fan of $\mathbb{R}^K$. The fan is called the *split fan* of K. Because each split function corresponds to a hyperplane in $\mathbb{R}^n$, a split-decomposable function makes a hyperplane arrangement. The hyperplane arrangement depends on the vector configuration as the domain of the function. A geometric property of such a hyperplane arrangement and a vector configuration is studied in [16, 17]. In Section 10, we consider a vector configuration Ξ such that Ξ contains the origin 0 and the vector in the opposite direction from the origin for each vector in $\Xi \setminus \{0\}$. By exploring the geometric lattice of the hyperplane arrangement obtained from a split-decomposable function on Ξ and the matroid associated with Ξ, we obtain a combinatorial characterization for the split fan of Ξ. The combinatorial characterization claims that the split fan depends only on the matroid associated with Ξ. In the case of Ω, the split fan $\mathrm{SF}(\Omega)$ coincides with a well-known complex: the space of phylogenetic trees $\mathbf{T}$. Our result designates that $\mathrm{SF}(\Omega)$ is isomorphic to the direct product of a simplex and $\mathbf{T}$. We discuss this result in Remark 9.11.

The present paper is organized as follows. Section 2 introduces notions; X-trees and X-splits for precise arguments about trees and tree metrics. In Section 3, we review dissimilarity maps, metrics, and tree metrics to classify dissimilarity data. Sections 4 and 5 briefly introduce Buneman's method and Bandelt and Dress' method, respectively. From Section 6 to Section 8, we introduce the polyhedral split decomposition of polyhedral convex functions and its extension for discrete functions. Section 6 contains preliminaries about polyhedral convex functions. In Section 7, we discuss the polyhedral split decomposition for more general type of polyhedral convex functions than in Hirai [16, 17]. In Section 8, we obtain the split decomposition of discrete functions from the results in Section 7. In Section 9, the split decomposition is applied to a metric which is regarded as a discrete function on Ω, and the Buneman index is derived geometrically. In Section 10, we rephrase some results on the split decomposition of discrete functions in terms of combinatorics, which is developed on matroids that arise from vector configurations and hyperplane arrangements.

2 X-trees and X-splits

In this section, we introduce X-trees to state precisely the most fundamental theorem about trees; Splits Equivalence theorem, on which Buneman's method and Bandelt and Dress' methods are based to reconstruct a tree.

A tree $T = (V, E)$ is a connected graph with no cycles. A vertex of T of degree one is called a *leaf.*

Definition 2.1 (X-tree). *An X-tree $\mathcal{T}$ is an ordered pair $(T; \phi)$ of a tree T with vertex set V and a map $\phi : X \to V$ such that $v \in \phi(X) = \{\phi(x) | x \in X\}$ holds for each $v \in V$ of degree at most two.*

Two X-trees $\mathcal{T}_1 = (T_1; \phi_1)$ and $\mathcal{T}_2 = (T_2; \phi_2)$, where $T_1 = (V_1, E_1)$ and $T_2 = (V_2, E_2)$, are *isomorphic* if there exists a bijection $\psi : V_1 \to V_2$ which induces a bijection between E_1 and E_2, and satisfies $\phi_2 = \psi \circ \phi_1$, in which case ψ is unique. We write $\mathcal{T}_1 \cong \mathcal{T}_2$ if $\mathcal{T}_1$ is isomorphic to $\mathcal{T}_2$.

Definition 2.2 (X-split). *An X-split is a partition of X into two non-empty sets, i.e., an X-split is a pair $\{A, B\}$ of A and B such that $\emptyset \neq A \subseteq X$, $\emptyset \neq B \subseteq X$, $A \cap B = \emptyset$ and $A \cup B = X$.*

Let $\mathcal{T} = (T; \phi)$ be an X-tree, and let e be an edge of T. Then, $T\backslash e = (V, E \setminus \{e\})$ consists of two connected components. If V_1 and V_2 denote the vertex sets of these two components. then $\{\phi^{-1}(V_1), \phi^{-1}(V_2)\}$ is an X-split. This X-split *corresponds to e in $\mathcal{T}$*. We denote by $\Sigma(\mathcal{T})$ the collection of X-splits that correspond to the edges of $\mathcal{T}$.

Definition 2.3 (compatible). *A pair of X-splits $\{A, B\}$ and $\{C, D\}$ are compatible if at least one of the sets $A \cap C, A \cap D, B \cap C$ and $B \cap D$ is the empty set.*

It is easy to verify that any two X-splits arising from a tree are compatible. The justification for Definition 2.3 is the following central theorem due to Buneman [8].

Theorem 2.4 (Splits Equivalence theorem). *Let Σ be a collection of X-splits. Then, there is an X-tree $\mathcal{T}$ such that $\Sigma = \Sigma(\mathcal{T})$ if and only if the X-splits in Σ are pairwise compatible. Moreover, if such an X-tree exists, then $\mathcal{T}$ is unique up to isomorphism.*

Both Buneman's and Bandelt and Dress' methods are based on Splits Equivalence theorem. They calculate the Buneman index and the isolation index, respectively, in order to mine X-splits from a nonnegative dissimilarity map or metric. Several methods which construct a tree, graph or network from such X-splits are known [12, 13, 19, 20].

3 Dissimilarity maps, metrics and tree metrics

We distinguish a dissimilarity map, metric, and tree metric to clarify the applicable scopes of methods to be described in Sections 4 and 5.

Definition 3.1 (dissimilarity map). *A function $d : X \times X \to \mathbb{R}$ is said to be a dissimilarity map on X, if it satisfies the following two conditions.*

(1) $d(i,i) = 0$ for all $i \in X$, and

(2) $d(i,j) = d(j,i)$ for all $i, j \in X$.

Definition 3.2 (metric). *A function $d : X \times X \to \mathbb{R}$ is said to be a metric on X, if it satisfies the following three conditions.*

(1) $d(i,i) = 0$ for all $i \in X$,

(2) $d(i,j) = d(j,i) \geq 0$ for all $i, j \in X$, and

(3) $d(i,j) \leq d(i,k) + d(k,j)$ for all $i, j, k \in X$.

The inequality in (3) is called the triangle inequality.

Let $T = (V, E)$ be a tree and suppose that $w : E \to \mathbb{R}$ is a map that assigns real-valued weights to the edges of T. This edge-weighting of T induces the following map from $V \times V$ into $\mathbb{R}$. For all $u, v \in V$, let $P(T; u, v)$ denote the unique path in T from u to v. We define the map $d_{(T;w)} : V \times V \to \mathbb{R}$ by setting, for all $u, v \in V$,

$$d_{(T;w)}(u,v) = \begin{cases} \sum_{e \in P(T;u,v)} w(e) & \text{if } u \neq v, \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathcal{T} = (T; \phi)$ is an X-tree, we define the map $d_{(\mathcal{T};w)} : X \times X \to \mathbb{R}$ by setting $d_{(\mathcal{T};w)}(x,y) = d_{(T;w)}(\phi(x), \phi(y))$ for all $x, y \in X$.

Definition 3.3 (tree metric). *A function $d : X \times X \to \mathbb{R}$ is said to be a tree metric on X, if there exist an X-tree $\mathcal{T} = (T; \phi)$ and a positive real-valued weighting $w : E(T) \to \mathbb{R}_{++}$ such that, for all $x, y \in X$,*

$$\begin{aligned} d(x,y) &= d_{(\mathcal{T};w)}(x,y) \\ &= d_{(T;w)}(\phi(x), \phi(y)). \end{aligned}$$

We say that $(\mathcal{T}; w)$ is a tree metric representation of d.

We introduce some fundamental theorems about tree metrics.

Theorem 3.4 (cf. [26, Theorem 7.1.8]). *Let d be a tree metric on X. Then, there is a unique tree representation of d up to isomorphism.*

Based on Theorem 3.4, many methods attempt to reconstruct the unique tree representation from a tree metric and succeed in the reconstruction. The next well-known theorem is due independently to Zaretskii [30], Simões-Pereira [27], and Buneman [8, 9]. The theorem connects the various results on tree metrics in several areas, e.g., T-theory [14], discrete convex analysis [23] and tropical geometry [25, 28].

Theorem 3.5 (Tree Metric theorem). *Let d be a nonnegative dissimilarity map on X. Then, d is a tree metric on X if and only if d satisfies the four-point condition:*

$$d(i,j) + d(k,l) \leq \max\{d(i,k) + d(j,l), d(i,l) + d(j,k)\}$$

for every four (not necessarily distinct) elements $i, j, k, l \in X$.

Split metrics in Definition 3.6 are the most fundamental metrics. Split metrics are also known as *cut metrics.*

Definition 3.6 (split metric). *The split metric $\xi_{\{A,B\}} : X \times X \to \{0, 1\}$ associated with an X-split $\{A, B\}$ is defined as*

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

for all $i, j \in X$.

A tree metric can be represented as a sum of split metrics. It is easy to prove the next theorem.

Theorem 3.7. *Let d be a tree metric on X, and let $(\mathcal{T}; w)$ be the tree representation of d. Then d can be represented as*

$$d = \sum_{\sigma \in \Sigma(\mathcal{T})} w(e_\sigma)\xi_\sigma,$$

where e_σ is the edge of $\mathcal{T}$ corresponding to $\sigma \in \Sigma(\mathcal{T})$.

4 Buneman's method

We briefly introduce Buneman's method [8], which utilizes the Buneman index.

Definition 4.1 (Buneman index). *Let $d : X \times X \to \mathbb{R}_+$ be a nonnegative dissimilarity map on X. For an X-split $\{A, B\}$, the Buneman index is defined as follows:*

$$b^d_{\{A,B\}} = \frac{1}{2} \min_{u,v \in A, x,y \in B} \left\{ \min \left\{ \begin{matrix} d(u,x) + d(v,y), \\ d(u,y) + d(v,x) \end{matrix} \right\} - d(u,v) - d(x,y) \right\}.$$

The Buneman index has important property as in Lemmas 4.2 and 4.3.

Lemma 4.2 (Buneman [8]). *Let $d : X \times X \to \mathbb{R}_+$ be a nonnegative dissimilarity map on X, and let $\{A, B\}$ and $\{C, D\}$ be X-splits. If $b^d_{\{A,B\}} > 0$ and $b^d_{\{C,D\}} > 0$, then $\{A, B\}$ and $\{C, D\}$ are compatible.*

Lemma 4.3. *Let $d = d_{(\mathcal{T};w)}$ be a tree metric on X and let σ be an X-split. Then, $b^d_\sigma > 0$ if and only if σ is an X-split induced by $\mathcal{T}$, in which case $b^d_\sigma = w(e)$ for the edge e of $\mathcal{T}$ corresponding to σ.*

For a nonnegative dissimilarity map $d : X \times X \to \mathbb{R}_+$, let

$$\Sigma_b(d) = \{\sigma \mid \sigma : \text{an } X\text{-split}, b^d_\sigma > 0\}.$$

By Lemma 4.2, $\Sigma_b(d)$ is pairwise compatible, and therefore, by Splits Equivalence theorem, there exists a unique X-tree $\mathcal{T}_d$ whose associated set of X-splits is $\Sigma_b(d)$. Let $w : E(\mathcal{T}_d) \to \mathbb{R}_{++}$ be the map defined by setting $w(e) = b^d_{\sigma_e}$ for all $e \in E(\mathcal{T}_d)$, where σ_e is the X-split of $\mathcal{T}_d$ induced by e. Let d_B denote the tree metric $d_{(\mathcal{T}_d;w)}$. Then the following theorem holds.

Theorem 4.4 (Buneman [8]). *Let $d : X \times X \to \mathbb{R}_+$ be a nonnegative dissimilarity map. Then*

(1) $d_B(x, y) \leq d(x, y)$ for all $x, y \in X$, and

(2) $d_B(x, y) = d(x, y)$ for all $x, y \in X$ if and only if d satisfies the four point condition, i.e., d is a tree metric.

Corollary 4.5. *Let d be a tree metric on X. Then d can be decomposed as*

$$d = \sum_{\sigma \in \Sigma_b(d)} b^d_\sigma \xi_\sigma.$$

Furthermore, $\Sigma_b(d)$ is pairwise compatible.

5 Bandelt and Dress' method

We briefly introduce Bandelt and Dress' method [1], which utilizes the isolation index.

Definition 5.1 (isolation index). *Let $d : X \times X \to \mathbb{R}_+$ be a nonnegative dissimilarity map on X. For an X-split $\{A, B\}$, the isolation index is defined as follows:*

$$i^d_{\{A,B\}} = \frac{1}{2} \min_{u,v \in A, x,y \in B} \left\{ \max \left\{ \begin{array}{c} d(u,x) + d(v,y), \\ d(u,y) + d(v,x), \\ d(u,v) + d(x,y) \end{array} \right\} - d(u,v) - d(x,y) \right\}.$$

For a nonnegative dissimilarity map $d : X \times X \to \mathbb{R}_+$, let

$$\Sigma_i(d) = \{\sigma \mid \sigma : \text{an } X\text{-split}, i^d_\sigma > 0\}.$$

In contrast to Buneman's method, the collection of X-splits $\Sigma_i(d)$ is not pairwise compatible. Instead, $\Sigma_i(d)$ is necessarily *weakly compatible*, that is, for any three X-splits $\{A_1, B_1\}$, $\{A_2, B_2\}$ and $\{A_3, B_3\}$ in $\Sigma_i(d)$, there exist no four points $a_0, a_1, a_2, a_3 \in X$ with $\{a_0, a_1, a_2, a_3\} \cap A_i = \{a_0, a_i\}$ for $i = 1, 2, 3$.

The next lemma holds for the isolation index, analogously to Lemma 4.3 for the Buneman index.

Lemma 5.2. *Let $d = d_{(\mathcal{T};w)}$ be a tree metric on X and let σ be an X-split. Then, $i^d_\sigma > 0$ if and only if σ is an X-split induced by $\mathcal{T}$, in which case $i^d_\sigma = w(e)$ for the edge e of $\mathcal{T}$ corresponding to σ.*

Bandelt and Dress provide a decomposition of metrics with split metrics.

Theorem 5.3 (Bandelt and Dress [1]). *Let $d : X \times X \to \mathbb{R}$ be a metric on X. Then d can be decomposed as*

$$d = \sum_{\sigma \in \Sigma_i(d)} i^d_\sigma \xi_\sigma + d',$$

where d' is a metric with $i^{d'}_{\sigma'} = 0$ for any X-split σ'. We call d' the split-prime residue of d.

Since pairwise compatibility implies weak compatibility, the following theorem is obtained.

Theorem 5.4 (Bandelt and Dress [1]). *Let d be a tree metric on X. Then d can be decomposed as*

$$d = \sum_{\sigma \in \Sigma_i(d)} i^d_\sigma \xi_\sigma.$$

Furthermore, $\Sigma_i(d)$ is pairwise compatible.

Since, in general, $\Sigma_i(d)$ is not pairwise compatible, it is impossible to represent d with a tree. For those cases, several methods to construct a network, called a *phylogenetic network*, from $\Sigma_i(d)$ are proposed [12, 13, 19, 20].

6 Polyhedral convex functions

This section is a preliminary to describe the polyhedral split decomposition of polyhedral convex functions in Section 7. Most of notations follow Hirai [16, §2], and proofs of the lemmas and propositions in this section can be found there.

Let $\mathbb{R}^n$ be the n dimensional Euclidean space with the standard inner product $\langle \cdot, \cdot \rangle$. For $x, y \in \mathbb{R}^n$, let $[x, y]$ denote the closed line segment between x and y. We refer to an $(n-1)$ dimensional affine subspace of $\mathbb{R}^n$ as a hyperplane. In particular, for $(a, r) \in \mathbb{R}^n \times \mathbb{R}$, we define a hyperplane $H_{a,r} = \{x \in \mathbb{R}^n \mid \langle a, x \rangle = r\}$, closed half spaces $H^-_{a,r} = \{x \in \mathbb{R}^n \mid \langle a, x \rangle \leq r\}$ and $H^+_{a,r} = \{x \in \mathbb{R}^n \mid \langle a, x \rangle \geq r\}$, and open half spaces $H^{--}_{a,r} = \{x \in \mathbb{R}^n \mid \langle a, x \rangle < r\}$ and $H^{++}_{a,r} = \{x \in \mathbb{R}^n \mid \langle a, x \rangle > r\}$. A set $P \subseteq \mathbb{R}^n$ is said to be a *polyhedron* if P can be represented as an intersection of finitely many closed half spaces. For a set $S \subseteq \mathbb{R}^n$, we denote by $\operatorname{conv} S, \operatorname{cone} S, \operatorname{aff} S$, and $\operatorname{lin} S$ the convex hull, the conical hull, the affine hull, and the linear hull of S, respectively, i.e.,

$$\begin{aligned}
\operatorname{conv} S &= \Big\{ \sum_{t \in T} \lambda_t t \mid T \subseteq S : \text{a finite set}, \lambda \in \mathbb{R}^T_+, \sum_{t \in T} \lambda_t = 1 \Big\}, \\
\operatorname{cone} S &= \Big\{ \sum_{t \in T} \lambda_t t \mid T \subseteq S : \text{a finite set}, \lambda \in \mathbb{R}^T_+ \Big\}, \\
\operatorname{aff} S &= \Big\{ \sum_{t \in T} \lambda_t t \mid T \subseteq S : \text{a finite set}, \lambda \in \mathbb{R}^T, \sum_{t \in T} \lambda_t = 1 \Big\}, \\
\operatorname{lin} S &= \Big\{ \sum_{t \in T} \lambda_t t \mid T \subseteq S : \text{a finite set}, \lambda \in \mathbb{R}^T \Big\}.
\end{aligned}$$

For a set $S \subseteq \mathbb{R}^n$, let $\operatorname{ri} S$ denote the relative interior of S and let $\operatorname{int} S$ denote the interior of S.

For a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, we define $\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$, which is the *effective domain* of f, and $\operatorname{epi} f = \{(x, \beta) \in \mathbb{R}^n \times \mathbb{R} \mid \beta \geq f(x)\}$, which is the *epigraph* of f. The *subdifferential* of a function f at point $x \in \operatorname{dom} f$ is defined to be the set

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

The *directional derivative* of f at point $x \in \operatorname{dom} f$ in a direction $d \in \mathbb{R}^n$ is defined by

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

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

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

The *conjugate* of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ with $\operatorname{dom} f \neq \emptyset$ is the function $f^\bullet : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^\bullet(p) = \sup_{x \in \mathbb{R}^n} \{\langle p, x \rangle - f(x)\} \ (p \in \mathbb{R}^n). \tag{6.1}$$

For a function f and a vector $p \in \mathbb{R}^n$, $f[-p]$ denotes the function defined by

$$f[-p](x) = f(x) - \langle p, x \rangle \quad (x \in \mathbb{R}^n). \tag{6.2}$$

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

$$f(x) = \max_{i \in I} \big\{ \langle p_i, x \rangle - q_i \big\} + \sum_{j \in J} \delta_{H^-_{a_j, b_j}}(x) \quad (x \in \mathbb{R}^n), \tag{6.3}$$

where $\{(p_i, q_i) \mid i \in I\}$ and $\{(a_j, b_j) \mid j \in J\}$ are finite subsets of $\mathbb{R}^n \times \mathbb{R}$. The conjugate function $f^\bullet$ of a polyhedral function f is also polyhedral and $f^{\bullet\bullet} = f$ holds. A function f is said to be *positively homogeneous* if $f(\lambda x) = \lambda f(x)$ holds for $\lambda \geq 0$ and $x \in \mathbb{R}^n$. If f is positively homogeneous, 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.

Lemma 6.1. *The subdifferential of a polyhedral convex function f in (6.3) is given by*

$$\partial f(x) = \operatorname{conv}\{p_i \mid i \in I, f(x) = \langle p_i, x\rangle - q_i\} + \operatorname{cone}\{a_j \mid j \in J, x \in H_{a_j,b_j}\} \quad (x \in \operatorname{dom} f).$$

Lemma 6.2. *For a polyhedral convex function f, the directional derivative $f'(x; d)$ at point $x \in \operatorname{dom} f$ in direction $d \in \mathbb{R}^n$ satisfies*

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

Lemma 6.3. *Let f, g be polyhedral convex functions. For $x \in \operatorname{dom} f \cap \operatorname{dom} g$ and $\alpha, \beta \geq 0$, we have*

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

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 $\mathcal{C}$, and

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

The dimension of $\mathcal{C}$, denoted by $\dim \mathcal{C}$, is the largest dimension of a polyhedron in $\mathcal{C}$. The underlying set of $\mathcal{C}$ is the 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* if its inclusionwise 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*.

Lemma 6.4. *For a polyhedral convex function f, the polyhedral subdivision $\mathcal{T}(f)$ is given by*

$$\mathcal{T}(f) = \{F \subseteq \mathbb{R}^n \mid F = \operatorname{argmin} f[-p] \text{ for some } p \in \mathbb{R}^n\}.$$

The polyhedral subdivisions $\mathcal{T}(f)$ and $\mathcal{T}(f^\bullet)$ are closely related. For $F \in \mathcal{T}(f)$ and a point $x \in \operatorname{ri} F$, we define $F^\bullet$ as

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

By the definition of $\partial f(x)$, we have $F^\bullet \in \mathcal{T}(f^\bullet)$. In fact, this map is well-defined and establishes a one-to-one correspondence between $\mathcal{T}(f)$ and $\mathcal{T}(f^\bullet)$.

Proposition 6.5. *For a polyhedral convex function f and $F, G \in \mathcal{T}(f)$, we have the followings (1) – (4).*

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

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

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

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

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 $|\mathcal{C}_1| \cap |\mathcal{C}_2|$. In particular, for a finite set of hyperplanes $\mathcal{H}$, we define the polyhedral subdivision $\mathcal{A}(\mathcal{H})$ of $\mathbb{R}^n$ as

$$\mathcal{A}(\mathcal{H}) = \bigwedge_{H \in \mathcal{H}} \{H, H^+, H^-\}.$$

Namely, $\mathcal{A}(\mathcal{H})$ is the partition of $\mathbb{R}^n$ by hyperplanes in $\mathcal{H}$.

Lemma 6.6. *For two polyhedral convex functions f, g with $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$, we have*

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

7 Polyhedral split decomposition

We derive the polyhedral split decomposition of a polyhedral convex function f, mostly following the paper of Hirai [16] except for the assumption that the effective domain of f is fully dimensional. The reason why we exclude the assumption is that we attempt, in Section 9, to apply the polyhedral split decomposition to a polyhedral convex function whose effective domain is not fully dimensional.

Definition 7.1 (split function). *For a hyperplane $H = H_{a,b}$ with $\|a\| = 1$, the split function $l_H : \mathbb{R}^n \to \mathbb{R}$ associated with H is defined by*

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

By Lemma 6.1, the polyhedral subdivision induced by a split function is given as follows.

Proposition 7.2. *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*

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

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

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

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

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

For a hyperplane H, we define the nonnegative number $c_H(f)$ as

$$c_H(f) = [f : l_H].$$

Lemma 7.3. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $g, h : \mathbb{R}^n \to \mathbb{R}$ be polyhedral convex functions.*

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

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

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

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

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

(1) $c_H(f) = +\infty$ if and only if $\mathrm{dom}\, f \subseteq H^+$ or $\mathrm{dom}\, f \subseteq H^-$.

(2) If $0 < c_H(f) < +\infty$, then $\{F \in \mathcal{T}(f) \mid F \subseteq H\}$ is a polyhedral subdivision of $H \cap \mathrm{dom}\, f$.

(3) Suppose that $H_1 \cap \mathrm{int}\,\mathrm{dom}\, f \neq \emptyset$ and $H_2 \cap \mathrm{int}\,\mathrm{dom}\, f \neq \emptyset$. Then, $H_1 = H_2$ if and only if $H_1 \cap \mathrm{dom}\, f = H_2 \cap \mathrm{dom}\, f$.

By the above observations and the polyhedrality of f, if $\dim \mathrm{dom}\, f = n$, the set of hyperplanes

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

is finite. The basic idea for the polyhedral split decomposition is to subtract split functions associated with hyperplanes in $\mathcal{H}(f)$ from a given polyhedral convex function successively. This idea is based on the following proposition [17, Lemma 2.5], which leads us to Polyhedral Split Decomposition theorem.

Proposition 7.4. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a polyhedral convex function with $\dim \mathrm{dom}\, f = n$. Then, for $H, H' \in \mathcal{H}(f)$ and $t \in [0, c_H(f)]$, we have*

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

One of the important results in Hirai [16, 17] is the following theorem.

Theorem 7.5 (Polyhedral Split Decomposition theorem [17, Theorem 2.2]). *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a polyhedral convex function with $\dim \mathrm{dom}\, f = n$. Then f can be decomposed as*

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

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

If the effective domain of a polyhedral convex function f is not fully dimensional, Proposition 7.4 and thus Polyhedral Split Decomposition theorem cannot be applied to f because $\mathcal{H}(f)$ may be a infinite set of hyperplanes. This fact follows from the next proposition [16, Proposition 2.20].

Proposition 7.6. *A hyperplane H belongs to $\mathcal{H}(f)$ if and only if H satisfies the following conditions (1) and (2).*

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

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

Although $\mathcal{H}(f)$ may be infinite, we notice from Proposition 7.6 that $\mathcal{H}(f)$ is determined by $\mathcal{T}(f)$ rather than f. Moreover, the next proposition holds for any polyhedral convex function. Compare this with Proposition 7.4.

Proposition 7.7. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a polyhedral convex function. Then, for $H, H' \in \mathcal{H}(f)$ and $t \in [0, c_H(f)]$, we have*

$$H = H' \Rightarrow c_{H'}(f - tl_H) = c_{H'}(f) - t,$$
$$H \cap \operatorname{dom} f \neq H' \cap \operatorname{dom} f \Rightarrow c_{H'}(f - tl_H) = c_{H'}(f).$$

Our idea for the polyhedral split decomposition of f is basically the same as Hirai. Our additional idea is to restrict $\mathcal{H}(f)$ to a set of hyperplanes such that there are no hyperplanes having the same intersection with $\operatorname{dom} f$ in the set. Technically speaking, we define the equivalence relation $\sim$ by letting $H \sim H'$ if $H \cap \operatorname{dom} f = H' \cap \operatorname{dom} f$. Since a collection of representatives from the equivalence classes has the desirable property, we decompose f by using the representatives. In general, we can select representatives of $\mathcal{H}(f)/\sim$ arbitrarily. However, in the case that f is a metric as in Section 9, we choose representatives having an interesting property to be described in Remark 9.10. We denote the chosen representatives of $\mathcal{H}(f)/\sim$ by $\mathcal{H}^\diamond(f)$. The representatives $\mathcal{H}^\diamond(f)$ is finite because $\mathcal{T}(f)$ is finite. Since $H \neq H'$ implies $H \cap \operatorname{dom} f \neq H' \cap \operatorname{dom} f$ for $H, H' \in \mathcal{H}^\diamond(f)$, we can decompose f uniquely with the hyperplanes in $\mathcal{H}^\diamond(f)$ by Proposition 7.7. The main result in this section is the following.

Theorem 7.8. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a polyhedral convex function. Then f can be decomposed as*

$$f = \sum_{H \in \mathcal{H}^\diamond(f)} c_H(f) l_H + f', \tag{7.1}$$

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

The rest of this section is devoted to proving Propositions 7.6 and 7.7, In particular, we prove Proposition 7.6 without the assumption on the full dimensionality of effective domains. The quotient $c_H(f)$ of f by l_H is written explicitly as in the next proposition [16, Proposition 2.18], [17, Lemma 2.7].

Proposition 7.9. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a polyhedral convex function, and let H be a hyperplane in $\mathbb{R}^n$. Then we have*

$$c_H(f) = \frac{1}{2} \inf \left\{ \frac{f(x) - f(w)}{l_H(x)} + \frac{f(y) - f(w)}{l_H(y)} \;\middle|\; \begin{array}{c} x \in \operatorname{dom} f \cap H^{++}, \\ y \in \operatorname{dom} f \cap H^{--}, \\ \{w\} = [x, y] \cap H \end{array} \right\}. \tag{7.2}$$

Proposition 7.6 is proved by using the explicit expression (7.2). Proposition 7.6 and Lemma 7.10 characterize the relation between the hyperplanes $\mathcal{H}(f)$ and subdivision $\mathcal{T}(f)$.

The proof of Proposition 7.6. The only-if part is easily observed. Indeed, if there exists $F \in \mathcal{T}(f)$ such that both $F \cap H^{++}$ and $F \cap H^{--}$ are nonempty, then $f - tl_H$ is not convex on F for any

$t > 0$. We show the 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

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

where $d = (x - y)/\|x - y\|$ and the last equality follows from Lemma 6.2. Let F_w be the unique minimal element of $\mathcal{T}(f)$ satisfying $w \in \operatorname{ri} F_w$. By the condition (2) in Proposition 7.6, 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 $\dim G = \dim \operatorname{dom} f - 1$. By Proposition 6.5 (4), we have

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

Since $x, y \in \operatorname{dom} f$ and $\operatorname{aff} G = H \cap \operatorname{aff} \operatorname{dom} f$, $G^\bullet$ is bounded in the directions d and $-d$. By Proposition 6.5 (2), we have

$$\begin{aligned}
\operatorname{aff} G^\bullet &= p_0 + (\operatorname{aff} G - \{z\})^\perp \\
&= p_0 + (H \cap \operatorname{aff} \operatorname{dom} f - \{z\})^\perp
\end{aligned}$$

for some $z \in G$ and $p_0 \in \mathbb{R}^n$. Note that αn_H and $-\beta n_H$ belong to $(H \cap \operatorname{aff} \operatorname{dom} f - \{z\})^\perp$ for some $\alpha, \beta > 0$. Therefore,

$$\begin{aligned}
\frac{\sup\{\langle p, d\rangle \mid p \in G^\bullet\} + \sup\{\langle q, -d\rangle \mid q \in G^\bullet\}}{\langle n_H, d\rangle} &\geq \frac{\langle \alpha n_H + p_0, d\rangle + \langle -\beta n_H + p_0, -d\rangle}{\langle n_H, d\rangle} \\
&= \alpha + \beta > 0.
\end{aligned}$$

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

Lemma 7.10. *Let H be a hyperplane, and let k be the dimension of* $\operatorname{dom} f$*.*

(1) If $c_H(f) = 0$, there exists a k-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 (7.2) is attained by any $x \in F \cap H^{++}, y \in F \cap H^{--}$.

(2) If $0 < c_H(f) < +\infty$, there exist k-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 (7.2) is attained by any $x \in G_1 \setminus H, y \in G_2 \setminus H$.

Proof. (1) is immediate from Proposition 7.6. We show (2). Put $g = f - c_H(f)l_H$. By $c_H(g) = 0$ and (1), there exists a k-dimensional polyhedron G such that $G \cap H^{++} \neq \emptyset$ and $G \cap H^{--} \neq \emptyset$. We define $G_1 = G \cap H^+$ and $G_2 = G \cap H^-$. Then we have $G_1, G_2 \in \mathcal{T}(f)$ by Lemma 6.6 and Proposition 7.2. Therefore, we obtain (2.1). For $x \in G_1 \setminus H, y \in G_2 \setminus H$ and $\{w\} = [x, y] \cap H$, we have $\|w-x\|/\|y-w\| = l_H(x)/l_H(y)$ and $l_H(w) = 0$. Since g is affine over $G = G_1 \cup G_2$, we have $\{g(w)-g(x)\}/\|w-x\| = \{g(y)-g(w)\}/\|y-w\|$. Hence, it follows that

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

This implies (2.2) □

We are now ready to prove Proposition 7.7.

The proof of Proposition 7.7. The case $H' = H$ is immediate from Lemma 7.3 (1). Hence we consider the case that $\operatorname{dom} f \cap H \neq \operatorname{dom} f \cap H'$. By Lemma 7.3 (3), it is sufficient to show $c_{H'}(f - tl_H) \geq c_{H'}(f)$. By the assumption that $\operatorname{dom} f \cap H \neq \operatorname{dom} f \cap H'$ and Lemma 7.10, we may assume that the minimum of (7.9) for H' is attained by some $x, y \in H^+ \cap \operatorname{dom} f$, or $x, y \in H^- \cap \operatorname{dom} f$, which implies that l_H is affine over $[x, y]$. Because $\|w-x\|/\|y-w\| = l_{H'}(x)/l_{H'}(y)$ and $\{l_H(w)-l_H(x)\}/\|w-x\| = \{l_H(y)-l_H(w)\}/\|y-w\|$ for $\{w\} = [x, y] \cap H'$, we have

$$\begin{aligned}
c_{H'}(f-tl_H) &= \frac{1}{2}\left\{\frac{f(x)-tl_H(x)-f(w)+tl_H(w)}{l_{H'}(x)} + \frac{f(y)-tl_H(y)-f(w)+tl_H(w)}{l_{H'}(y)}\right\} \\
&= \frac{1}{2}\left\{\frac{f(x)-f(w)}{l_{H'}(x)} + \frac{f(y)-f(w)}{l_{H'}(y)} - t\left\{\frac{l_H(x)-l_H(w)}{l_{H'}(x)} + \frac{l_H(y)-l_H(w)}{l_{H'}(y)}\right\}\right\} \\
&= \frac{1}{2}\left\{\frac{f(x)-f(w)}{l_{H'}(x)} + \frac{f(y)-f(w)}{l_{H'}(y)}\right\} \\
&\geq c_{H'}(f).
\end{aligned}$$

□

In the case of $\dim \operatorname{dom} f = n$, if $H \cap \operatorname{dom} f \neq H' \cap \operatorname{dom} f$, then $H \neq H'$. Hence, a hyperplane with $0 < c_H(f) < +\infty$ is uniquely determined from the subdivision $\mathcal{T}(f)$ by Proposition 7.6, and thus Proposition 7.4 holds instead of Proposition 7.7. As a result, we obtain Polyhedral Split Decomposition theorem. On the other hand, if $\operatorname{dom} f$ is not full-dimensional, there exist infinitely many hyperplanes having the same intersection with $\operatorname{dom} f$. Moreover, $c_{H'}(f - tl_H)$ may not be equal to $c_{H'}(f)$ for $H, H' \in \mathcal{H}(f)$ and $t \in [0, c_H(f)]$ despite that $H \neq H'$. Thus, Polyhedral Split Decomposition theorem needs some modification, and its modified version is Theorem 7.8.

We conclude this section by a remark, which is used in Section 8.

Remark 7.11. By Proposition 7.2, we have $\mathcal{T}(\alpha l_H) = \{H, H^+, H^-\}$ for $\alpha \in \mathbb{R}_{++}$. Hence, by Lemma 6.6, corresponding to the decomposition (7.1), $\mathcal{T}(f)$ is decomposed as

$$\mathcal{T}(f) = \mathcal{A}(\mathcal{H}(f)) \wedge \mathcal{T}(f'). \tag{7.3}$$

8 Split decomposition of discrete functions

In this section, we describe the split decomposition of discrete functions defined on a finite set K of points of $\mathbb{R}^n$. The split decomposition of a discrete function is summarized as the polyhedral split decomposition of the convex extension of the discrete function. Although Hirai [16, 17] assumes Assumption 8.3, which guarantees the full-dimensionality of effective domains, some of his results do not require Assumption 8.3. Then we rearrange or restate the results and introduce the split decomposition of discrete functions.

Let K be a finite set of points in $\mathbb{R}^n$. For a function $f : K \to \mathbb{R}$, the *homogeneous convex extension* of f is defined by

$$\overline{f}(x) = \inf\Big\{\sum_{y\in K} \lambda_y f(y) \mid \sum_{y\in K} \lambda_y y = x, \lambda_y \geq 0\ (y \in K)\Big\} + \delta_{\mathrm{cone}\,K}(x) \quad (x \in \mathbb{R}^n). \tag{8.1}$$

By definition, $\overline{f}$ is a positively homogeneous polyhedral convex function with $\mathrm{dom}\, f = \mathrm{cone}\, K$. By linear programming duality, $\overline{f}$ is also expressed as

$$\overline{f}(x) = \sup\big\{\langle p, x\rangle \mid p \in \mathbb{R}^n, \langle p, y\rangle \leq f(y)\ (y \in K)\big\} \quad (x \in \mathbb{R}^n). \tag{8.2}$$

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

$$Q(f) = \{p \in \mathbb{R}^n \mid \langle p, y\rangle \leq f(y)\ (y \in K)\},$$

and $Q(f) = \partial\overline{f}(0)$. The polyhedral subdivision $\mathcal{T}(\overline{f})$ of $\mathrm{cone}\, K$ is the intersection of the normal fan of $Q(f)$ with $\mathrm{cone}\, K$.

For a function $g : \mathbb{R}^n \to \mathbb{R}$, we denote the restriction of g to K by g^K. A function $f : K \to \mathbb{R}$ is said to be *convex-extensible* if it satisfies $\overline{f}^K = f$. The set of convex-extensible functions is recognized as a fundamental class in discrete convex analysis [23].

We give a fundamental property of discrete functions and their homogeneous convex extensions.

Lemma 8.1. *Let $f : K \to \mathbb{R}$ be a convex-extensible discrete function. $F \in \mathcal{T}(\overline{f})$ is represented as* $\mathrm{cone}\{y \mid y \in K, \langle p, y\rangle = f(y)\}$ *for some $p \in Q(f)$. Furthermore $\overline{f}(x) = \overline{f^{F\cap K}}(x)$ for $x \in F$.*

The next proposition [17, Theorem 3.2] leads us to Discrete Split Decomposition theorem.

Proposition 8.2. *For $f : K \to \mathbb{R}$, $H \in \mathcal{H}(\overline{f})$, and $t \in [0, c_H(\overline{f})]$, we have*

$$\overline{f} = tl_H + \overline{f - tl_H^K}.$$

For a discrete function $f : K \to \mathbb{R}$, the next assumption guarantees the full dimensionality of $\mathrm{dom}\,\overline{f} = \mathrm{cone}\, K$.

Assumption 8.3. *$K \subseteq \mathbb{R}^n$ is a finite set such that* $\mathrm{aff}\, K = U$ *for some hyperplane U not containing the origin of $\mathbb{R}^n$.*

Under Assumption 8.3, the next theorem follows from Polyhedral Split Decomposition theorem and Proposition 8.2.

Theorem 8.4 (Discrete Split Decomposition theorem [17, Theorem 3.2]). *A discrete function $f : K \to \mathbb{R}$ satisfying Assumption 8.3 can be decomposed as*

$$f = \sum_{H\in\mathcal{H}(\overline{f})} c_H(\overline{f}) l_H^K + f',$$

where $f' : K \to \mathbb{R}$ satisfies $c_{H'}(\overline{f'}) \in \{0, +\infty\}$ for any linear hyperplane H'. Furthermore, we have

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

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

Before we give Discrete Split Decomposition theorem without Assumption 8.3, we observe the relation between K and $\mathcal{H}(\overline{f})$. Let f be a convex-extensible discrete function on K. Note that since $\mathcal{T}(\overline{f})$ is the intersection of the normal fan of $Q(f)$ with $\operatorname{cone} K$, each hyperplane $H \in \mathcal{H}(\overline{f})$ is linear, i.e., $H = H_{a,0}$ for some $a \in \mathbb{R}^n$. From the regularity of the subdivision $\mathcal{T}(\overline{f})$ induced by $\overline{f}$, we notice possible hyperplanes appearing in $\mathcal{H}(\overline{f})$ is limited by the point set K. Motivated by this observation, we make the next definition.

Definition 8.5 (K-admissible). *A set of linear hyperplanes $\mathcal{H}$ is K-admissible if $\mathcal{H}$ satisfies*

(A1) $H \cap \operatorname{ri} \operatorname{cone} K \neq \emptyset$ *for each* $H \in \mathcal{H}$, *and*

(A2) $\operatorname{cone}(F \cap K) = F \cap \operatorname{cone} K$ *for each* $F \in \mathcal{A}(\mathcal{H})$.

Note that K-admissibility is determined solely by K. Under Assumption 8.3, K-admissibility can be restated as in [17, §3.2].

Lemma 8.6. *For $f : K \to \mathbb{R}$, the set of hyperplanes $\mathcal{H}(\overline{f})$ is K-admissible.*

Proof. (A1) is clearly satisfied. We show (A2). The inclusion ($\subseteq$) is obvious. We show ($\supseteq$). By (7.3) and Lemma 8.1, we have

$$\begin{aligned}
F \cap \operatorname{cone} K &= \bigcup \{G \mid G \in \mathcal{T}(\overline{f}), G \subseteq F\} \\
&= \bigcup \{\operatorname{cone}(G \cap K) \mid G \in \mathcal{T}(\overline{f}), G \subseteq F\} \\
&\subseteq \operatorname{cone}(K \cap \bigcup \{G \mid G \in \mathcal{T}(\overline{f}), G \subseteq F\}) \\
&= \operatorname{cone}(F \cap K)
\end{aligned}$$

□

Note that if a set of linear hyperplanes $\mathcal{H}$ is K-admissible, then any subset of $\mathcal{H}$ is also K-admissible. So we define the set of linear hyperplanes $\mathcal{H}_K$ as

$$\mathcal{H}_K = \{H \mid H : \text{a linear hyperplane}, \{H\} \text{ is } K\text{-admissible}\}.$$

By Lemma 8.6, $\mathcal{H}(f) \subseteq \mathcal{H}_K$ holds for any $f : K \to \mathbb{R}$. Therefore, from algorithmic viewpoint, we restrict $\mathcal{H}_K$ to representatives, denoted by $\mathcal{H}_K^\diamond$, of $\mathcal{H}_K/\sim$ rather than $\mathcal{H}(\overline{f})$, and we determine $\mathcal{H}^\diamond(\overline{f}) \subseteq \mathcal{H}_K^\diamond$, that is, $\mathcal{H}^\diamond(\overline{f}) = \mathcal{H}(\overline{f}) \cap \mathcal{H}_K^\diamond$. Without Assumption 8.3, a discrete function f can be decomposed uniquely by using the split functions associated with hyperplanes in $\mathcal{H}^\diamond(\overline{f})$. Thus, we obtain the following theorem.

Theorem 8.7. *A discrete function $f : K \to \mathbb{R}$ can be decomposed as*

$$f = \sum_{H \in \mathcal{H}^\diamond(\overline{f})} c_H(\overline{f}) l_H^K + f',$$

where $f' : K \to \mathbb{R}$ satisfies $c_{H'}(\overline{f'}) \in \{0, +\infty\}$ for any linear hyperplane H'. Furthermore, we have

$$\overline{f} = \sum_{H \in \mathcal{H}^\diamond(\overline{f})} c_H(\overline{f}) l_H + \overline{f'}.$$

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

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

Theorem 8.8 (Hirai [17, Theorem 3.4]). *For a discrete function $f : K \to \mathbb{R}$ and a hyperplane $H \in \mathcal{H}_K$, let $\tilde{c}_H(f)$ be defined by*

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

Then we have

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

Here, we introduce two interesting notions "split-decomposability" and "split fan". A function $f \in \mathbb{R}^K$ is said to be *split-decomposable* if $f - \sum_{H \in \mathcal{H}^\diamond(f)} c_H(\overline{f}) l_H^K$ is (the restriction of) a linear function. The *split fan* of K is the fan consisting of all split-decomposable functions on K. Split-decomposable functions are closely related to the totally split-decomposable metrics defined by Bandelt and Dress [1] and tree metrics in the case of $K = \Omega$ as mentioned in Section 1. We explain the relation between split-decomposable functions on K and K-admissible sets of hyperplanes.

We begin by showing a fundamental lemma about the homogeneous convex extensions of discrete functions.

Lemma 8.9. *Let $f : K \to \mathbb{R}$ be a discrete function. Then we have*

$$\overline{cf + (\langle q, \cdot \rangle)^K} = c\overline{f} + \langle q, \cdot \rangle + \delta_{\operatorname{cone} K} \quad (c \in \mathbb{R}_+, q \in \mathbb{R}^n).$$

Proof. If $c = 0$, then it is immediate from the definition (8.1). If $c > 0$, by (8.2), we have

$$\begin{aligned} \overline{cf + (\langle q, \cdot \rangle)^K}(x) &= \sup\bigl\{\langle p, x\rangle \mid p \in \mathbb{R}^n, \langle p, y\rangle \le cf(y) + \langle q, y\rangle \ (y \in K)\bigr\} \\ &= \sup\bigl\{c\langle (p-q)/c, x\rangle + \langle q, x\rangle \mid p \in \mathbb{R}^n, \langle (p-q)/c, y\rangle \le f(y) \ (y \in K)\bigr\} \\ &= c\sup\bigl\{\langle p', x\rangle \mid p' \in \mathbb{R}^n, \langle p', y\rangle \le f(y) \ (y \in K)\bigr\} + \langle q, x\rangle \\ &= (c\overline{f} + \langle q, \cdot\rangle)(x). \end{aligned}$$

In the third equality, we define $p' = (p-q)/c$. $\square$

Note that the quotient of $\overline{cf + (\langle q, \cdot \rangle)^K}$ by a split function depends only on $c\overline{f}$. Hence, the discrete split decomposition of $cf + (\langle q, \cdot \rangle)^K$ is determined by $c\overline{f}$. The next proposition can be proved in the same way as [16, Proposition 3.10], [17, Proposition 3.5].

Proposition 8.10. *For $\mathcal{H} \subseteq \mathcal{H}_K^\diamond$ and $\alpha \in \mathbb{R}_{++}^{\mathcal{H}}$, let $f = \sum_{H \in \mathcal{H}} \alpha_H l_H^K$. Then the following conditions (a), (b) and (c) are equivalent.*

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

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

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

By Lemma 8.9 and Proposition 8.10, every split-decomposable function is constructed from a K-admissible set of hyperplanes, i.e., the sum of a positive combination of (the restrictions of) the split functions associated with the hyperplanes and (the restriction of) a linear function. Thus, split-decomposable functions are also determined by K through the K-admissible sets of hyperplanes since the K-admissibility depends on K. Moreover, we obtain the next proposition, which can be verified in the same way as [16, Proposition 3.12], [17, Proposition 3.6].

Proposition 8.11. *Let $\mathcal{H} \subseteq \mathcal{H}_K^\diamond$ be a K-admissible set of hyperplanes. Then the set of vectors $\{l_H^K \mid H \in \mathcal{H}\} \cup \{e_i^K \mid 1 \leq i \leq n\}$ is linearly independent in $\mathbb{R}^K$, where $e_i : \mathbb{R}^n \to \mathbb{R}$ is the ith coordinate function defined by $e_i(x) = x_i$ for $x \in \mathbb{R}^n$.*

By Proposition 8.11, the split fan of K can be naturally regraded as a simplicial fan of $\mathbb{R}^K$ and it is obviously isomorphic to the set of K-admissible sets of hyperplanes in $\mathcal{H}_K^\diamond$ as an abstract simplicial complex. We give a combinatorial characterization for split fans in Section 10.

9 The Buneman index

In this section, we derive the Buneman index by discrete split decomposition.

9.1 Hirai's split decomposition of nonnegative dissimilarity maps

We briefly introduce Hirai's split decomposition of nonnegative dissimilarity maps.

Let $X = \{1, 2, \ldots, n\}$. A nonnegative dissimilarity map is naturally regarded as a discrete function defined on the set $\Lambda = \{\chi_i + \chi_j \mid i, j \in X\}$ by the correspondence:

$$d(\chi_i + \chi_j) \leftarrow d(i, j) \quad (i, j \in X).$$

Lemma 9.1. *A discrete function $f : \Lambda \to \mathbb{R}$ with $f(2\chi_i) = 0$ for all $i \in X$ is convex-extensible if and only if it satisfies $f(\chi_i + \chi_j) \leq 0$ for all $i, j \in X$.*

Hence it is natural to regard a nonnegative dissimilarity map $d : \Lambda \to \mathbb{R}$ as a discrete concave function on Λ. Since $\operatorname{aff} \Lambda = \{x \in \mathbb{R}^n \mid \sum_{i \in X} x(i) = 2\}$, we can apply Discrete Split Decomposition theorem to $-d$. The convex extension of $-d$ is as follows:

$$\begin{aligned}
\overline{(-d)}(x) &= \inf\Big\{ \sum_{i,j \in X} \lambda_{ij}(-d)(\chi_i + \chi_j) \mid \sum_{i,j \in X} \lambda_{ij}(\chi_i + \chi_j) = x, \lambda_{ij} \geq 0 \ (i, j \in X)\Big\} + \delta_{\operatorname{cone}\Lambda}(x) \\
&= \sup\big\{\langle p, x\rangle \mid p \in \mathbb{R}^n, \langle p, \chi_i + \chi_j\rangle \leq -d(\chi_i + \chi_j) \ (i, j \in X)\big\} \quad (x \in \mathbb{R}^n).
\end{aligned} \tag{9.1}$$

From (9.1), $\overline{(-d)}$ is the support function of the polyhedron

$$Q(-d) = \big\{p \in \mathbb{R}^n \mid \langle p, \chi_i + \chi_j\rangle \leq -d(\chi_i + \chi_j) \ (i, j \in X)\big\}.$$

By discrete split decomposition, Hirai extended the results of Bandelt and Dress' split decomposition of nonnegative dissimilarity maps by using a partial X-split, which is a pair $\{A, B\}$ such that $\emptyset \neq A, B \subseteq X, A \cap B = \emptyset, A \cup B \subseteq X$.

9.2 Deriving the Buneman index

We consider the finite set

$$\Omega = \{\chi_i - \chi_j \mid i, j \in X\}.$$

A metric γ is regarded as a discrete function defined on the set Ω by the correspondence:

$$\gamma(\chi_i - \chi_j) = \gamma(\chi_j - \chi_i) \leftarrow \gamma(i, j) \quad (i, j \in X).$$

Lemma 9.2. *A discrete function $f : \Omega \to \mathbb{R}$ with $f(0) = 0$ is convex-extensible if and only if f satisfies $f(\chi_i - \chi_j) \leq f(\chi_i - \chi_k) + f(\chi_k - \chi_j)$ for all $i, j, k \in X$.*

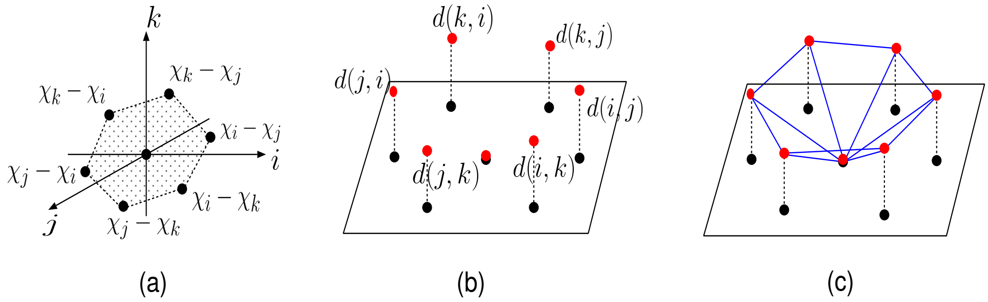

Figure 1: The homogeneous convex extension of a metric γ on $X = \{i, j, k\}$.

Proof. Suppose that f is convex-extensible. Since $(\chi_i - \chi_j) = (\chi_i - \chi_k) + (\chi_k - \chi_j)$ for $i, j, k \in X$, the convex-extensibility indicates $f(\chi_i - \chi_j) \leq f(\chi_i - \chi_k) + f(\chi_k - \chi_j)$.

Next suppose that f satisfies, for all $i, j, k \in X$, $f(\chi_i - \chi_j) \leq f(\chi_i - \chi_k) + f(\chi_k - \chi_j)$, which we call the triangle inequality for convenience. For a point $\chi_u - \chi_v \in \Omega$, we consider the following linear program:

$$\begin{aligned} \text{minimize} \quad & \sum_{i,j \in X} \lambda_{ij} f(\chi_i - \chi_j) \\ \text{subject to} \quad & \sum_{i,j \in X} \lambda_{ij}(\chi_i - \chi_j) = \chi_u - \chi_v, \\ & \lambda_{ij} \geq 0 \quad (i, j \in X). \end{aligned}$$

We interpret this linear program as a problem on the complete graph K_n each edge of which is two-way directed, that is, has an edge of the opposite direction. For the first constraint, an arbitrary representation of $\chi_u - \chi_v$ as a nonnegative combination of other points in Ω is regarded as a union of directed paths from u to v with nonnegative weights. The objective value for the representation is equal to the nonnegative weighted sum of the lengths of the directed paths. Hence, $\overline{f}(\chi_u - \chi_v)$ is equal to the minimum of such objective values. The triangle inequalities imply that objective values decrease by taking a shortcut along the directed paths. Therefore, the triangle inequalities suffice the convex-extensibility of f. $\square$

By Lemma 9.2, metrics are convex-extensible on the set Ω. The homogeneous convex extension of γ is defined by

$$\begin{aligned} \overline{\gamma}(x) &= \inf\{\sum_{i,j \in X} \lambda_{ij}\gamma(\chi_i - \chi_j) \mid \sum_{i,j \in X} \lambda_{ij}(\chi_i - \chi_j) = x,\ \lambda_{ij} \geq 0\ (i, j \in X)\} + \delta_{\operatorname{cone}\Omega}(x) \\ &= \sup\{\langle p, x\rangle \mid p \in \mathbb{R}^n, \langle p, \chi_i - \chi_j\rangle \leq \gamma(\chi_i - \chi_j)\ (i, j \in X)\} \quad (x \in \mathbb{R}^n). \end{aligned} \tag{9.2}$$

The effective domain of $\overline{\gamma}$ is $\operatorname{aff}\Omega = \{x \in \mathbb{R}^n \mid \sum_{i \in X} x(i) = 0\}(= \operatorname{cone}\Omega = \operatorname{lin}\Omega)$. From (9.2), $\overline{\gamma}$ is the support function of the polyhedron

$$Q(\gamma) = \{p \in \mathbb{R}^n \mid \langle p, \chi_i - \chi_j\rangle \leq \gamma(\chi_i - \chi_j)\ (i, j \in X)\}.$$

Figure 1 (c) illustrates the homogeneous convex extension of a metric γ on $X = \{i, j, k\}$. Since X on a linear space as in Figure 1 (a), we can project $\{(\chi_i - \chi_j, d(i, j)) \mid i, j \in X\}$ to 3-dimensional space as shown in Figure 1 (b). Although $\overline{\gamma}$ is a polyhedral convex function, its effective domain $\operatorname{dom}\overline{\gamma}$ is not fully dimensional. Then we define representatives $\mathcal{H}_\Omega^\diamond$ as mentioned in Section 8. We begin by revealing $\mathcal{H}_\Omega := \{H \mid H : \text{a linear hyperplane}, \{H\} : \Omega\text{-admissible}\}$. We denote $H_{(\chi_A - \chi_B)/2\sqrt{|X|},0}$

by $H_{\{A,B\}}$ and all-one vector by $\mathbf{1}$. For $x \in \mathbb{R}^n$, we define $\mathrm{supp}^+ x = \{i \mid x(i) > 0, i \in X\}$ and $\mathrm{supp}^- x = \{i \mid x(i) < 0, i \in X\}$. Hereafter, coefficients for scaling vectors to unit ones are omitted for simplicity.

Proposition 9.3. $\mathcal{H}_\Omega = \{H_{\alpha(\chi_A - \chi_B)+\beta\mathbf{1},0} \mid \{A,B\} : \text{an } X\text{-split},\ \alpha,\beta \in \mathbb{R}\}$.

Proof. Because $\langle \beta\mathbf{1}, x\rangle = 0$ for any $x \in \mathrm{aff}\,\Omega$, the term $\beta\mathbf{1}$ of a coefficient vector can be neglected. Hence, it suffices for showing ($\supseteq$) that each $H_{\{A,B\}}$ satisfies the conditions (A1) and (A2) in Definition 8.5. (A1) is clearly satisfied since $0 \in \mathrm{ri}\,\mathrm{cone}\,\Omega$ and $0 \in H_{\{A,B\}}$. We show (A2). Obviously, $\mathrm{cone}\,\Omega \cap H_{\{A,B\}} \supseteq \mathrm{cone}\,(\Omega \cap H_{\{A,B\}})$. To show ($\subseteq$), we take an arbitrary $x \in \mathrm{cone}\,\Omega \cap H_{\{A,B\}}$, which can be represented as

$$x = \sum_{\chi_i - \chi_j \in H^{++}_{\{A,B\}}} \lambda_{ij}(\chi_i - \chi_j) + \sum_{\chi_k - \chi_l \in H_{\{A,B\}}} \lambda_{kl}(\chi_k - \chi_l) + \sum_{\chi_u - \chi_v \in H^{--}_{\{A,B\}}} \lambda_{uv}(\chi_u - \chi_v), \tag{9.3}$$

where $\lambda_{st} \geq 0$ for all $(s,t) \in X \times X$. If there is a term $\lambda_{uv}(\chi_u - \chi_v)$ with $\chi_u - \chi_v \in H^{--}_{\{A,B\}}$ and $\lambda_{uv} > 0$, there necessarily exists a term $\lambda_{ij}(\chi_i - \chi_j)$ with $\chi_i - \chi_j \in H^{++}_{\{A,B\}}$ and $\lambda_{ij} > 0$ because $x \in H_{\{A,B\}}$. By the definition of the set Ω, there exist $\chi_i - \chi_v$ and $\chi_u - \chi_j$ in $\mathrm{cone}\,\Omega$. Moreover, we have $\chi_i - \chi_v, \chi_u - \chi_j \in H_{\{A,B\}}$ since $i, v \in A, j, u \in B$.

We modify the representation (9.3) as follows. If $\lambda_{ij} \geq \lambda_{uv}$,

$$\lambda_{ij}(\chi_i - \chi_j) + \lambda_{uv}(\chi_u - \chi_v) = \lambda_{uv}(\chi_i - \chi_v) + \lambda_{uv}(\chi_u - \chi_j) + (\lambda_{ij} - \lambda_{uv})(\chi_i - \chi_j)$$

and, if $\lambda_{ij} < \lambda_{uv}$,

$$\lambda_{ij}(\chi_i - \chi_j) + \lambda_{uv}(\chi_u - \chi_v) = \lambda_{ij}(\chi_i - \chi_v) + \lambda_{ij}(\chi_u - \chi_j) + (\lambda_{uv} - \lambda_{ij})(\chi_u - \chi_v).$$

Either modification gives another representation of x, and the sum of the coefficients in the first term of RHS of (9.3); $\sum_{\chi_i - \chi_j \in H^{++}_{\{A,B\}}} \lambda_{ij}$ and the sum of the coefficients in the third term of RHS of (9.3); $\sum_{\chi_u - \chi_v \in H^{--}_{\{A,B\}}} \lambda_{uv}$ are smaller than before, respectively. In the case where there is a term $\lambda_{ij}(\chi_i - \chi_j)$ with $\chi_i - \chi_j \in H^{++}_{\{A,B\}}$ and $\lambda_{ij} > 0$, we can the same modifications as above. We can repeat these modifications unless $\sum_{\chi_i - \chi_j \in H^{++}_{\{A,B\}}} \lambda_{ij} = 0$ and $\sum_{\chi_u - \chi_v \in H^{--}_{\{A,B\}}} \lambda_{uv} = 0$, and the two values decrease monotonically in the repetition. Moreover, Ω is a finite set. Therefore, the repetition terminates and the first and third terms of RHS of (9.3) is not needed to represent x, which means x is represented as a nonnegative combination of points in $\Omega \cap H_{\{A,B\}}$, namely, $x \in \mathrm{cone}\,(\Omega \cap H_{\{A,B\}})$.

Next we show ($\subseteq$). Let $H_{a,0} \in \mathcal{H}_\Omega$. By Ω-admissibility, $H_{a,0}$ satisfies that $\mathrm{cone}\,\Omega \cap H_{a,0} = \mathrm{cone}\,(\Omega \cap H_{a,0})$. Since $\dim(\mathrm{cone}\,\Omega \cap H_{a,0}) = n-2$, $\Omega \cap H_{a,0}$ contains $n-2$ linearly independent vectors such as $\chi_i - \chi_j$ for some $i, j \in X$. Note that (i) if $\chi_i - \chi_j$ is contained in $\Omega \cap H_{a,0}$, then $\chi_j - \chi_i$ is also contained in $\Omega \cap H_{a,0}$, and (ii) if $\chi_i - \chi_j$ and $\chi_j - \chi_k$ are contained in $\Omega \cap H_{a,0}$, then $\chi_i - \chi_k$ is also contained in $\Omega \cap H_{a,0}$. Hence, by row and column permutations, we may assume, without loss of generality, that a coefficient vector a satisfies

$$\begin{bmatrix} I & -c_{n-1} & -c_n \end{bmatrix} a = 0, \tag{9.4}$$

where I is $(n-2) \times (n-2)$ unit matrix and two column vectors c_{n-1}, c_n are 0-1 vectors such that $\mathrm{supp}^+ c_{n-1} \cap \mathrm{supp}^+ c_n = \emptyset$ and $\mathrm{supp}^+ c_{n-1} \cup \mathrm{supp}^+ c_n = X$. In the case that $\mathrm{supp}^+ c_{n-1} = X$ or $\mathrm{supp}^+ c_n = X$, it follows from the equation (9.4) that $a = \beta\mathbf{1}$ for some $\beta \in \mathbb{R}$. In the case that $\mathrm{supp}^+ c_{n-1} \neq X$ and $\mathrm{supp}^+ c_n \neq X$, we define $A := \mathrm{supp}^+ c_{n-1}$ and $B := \mathrm{supp}^+ c_n$. Then, $\{A,B\}$ is obviously an X-split. Moreover, because of the equation (9.4), vector a can be represented as $a = \xi\chi_A + \eta\chi_B = \frac{\xi-\eta}{2}(\chi_A - \chi_B) + \frac{\xi+\eta}{2}\mathbf{1}$ for some $\xi, \eta \in \mathbb{R}$. □

It is obvious that, for all $\alpha, \beta \in \mathbb{R}$,

$$H_{\{A,B\}} \cap \operatorname{dom} \overline{\gamma} = H_{\alpha(\chi_A - \chi_B) + \beta \mathbf{1}, 0} \cap \operatorname{dom} \overline{\gamma}.$$

Then we define the hyperplanes:

$$\mathcal{H}_\Omega^\diamond := \{H_{\chi_A - \chi_B, 0} \mid \{A, B\} : \text{an } X\text{-split}\}.$$

Proposition 9.4. *For each $H \in \mathcal{H}_\Omega$, there exists a hyperplane $H^\diamond$ such that*

$$H \cap \operatorname{dom} \overline{\gamma} = H^\diamond \cap \operatorname{dom} \overline{\gamma}$$

in $\mathcal{H}_\Omega^\diamond$. Moreover, for all $H, H' \in \mathcal{H}_\Omega^\diamond$, $H \cap \operatorname{dom} \overline{\gamma} = H' \cap \operatorname{dom} \overline{\gamma}$ if and only if $H = H'$.

By Proposition 9.4, $\mathcal{H}_\Omega^\diamond$ constitutes representatives of $\mathcal{H}_\Omega/\sim$. Hence, Theorem 8.7 can be applied to γ and γ is decomposed uniquely with hyperplanes in $\mathcal{H}(\overline{\gamma}) \cap \mathcal{H}_\Omega^\diamond$. Moreover, $\mathcal{H}(\overline{\gamma}) \cap \mathcal{H}_\Omega^\diamond$ provides an interesting result described in Remark 9.10.

Before applying Theorem 8.7 to γ, we show the main result in this paper. For $H \in \mathcal{H}_\Omega^\diamond$, $c_H(\overline{\gamma})$ is represented by Buneman index.

Theorem 9.5. *Let $\gamma : X \times X \to \mathbb{R}$ be a metric, and let $H_{\{A,B\}}$ be the hyperplane associated with an X-split $\{A, B\}$. Then we have*

$$c_{H_{\{A,B\}}}(\overline{\gamma}) = \sqrt{|X|} \max\{0, b^\gamma_{\{A,B\}}\},$$

where $b^\gamma_{\{A,B\}}$ is the Buneman index for the X-split $\{A, B\}$.

Proof. We apply Theorem 8.8 to $\overline{\gamma}$. $\tilde{c}_{H_{\{A,B\}}}(\gamma)$ is equal to the minimum of

$$\frac{\gamma(\chi_i - \chi_k) - \overline{\gamma^{\Omega \cap H_{\{A,B\}}}}(w)}{2 l_{H_{\{A,B\}}}(\chi_i - \chi_k)} + \frac{\gamma(\chi_l - \chi_j) - \overline{\gamma^{\Omega \cap H_{\{A,B\}}}}(w)}{2 l_{H_{\{A,B\}}}(\chi_l - \chi_j)}$$

where $i, j \in A, k, l \in B$, and $\{w\} = H_{\{A,B\}} \cap [\chi_i - \chi_k, \chi_l - \chi_j]$. Hence, we have

$$\begin{aligned}\tilde{c}_{H_{\{A,B\}}}(\gamma) = \frac{\sqrt{|X|}}{2} \min_{i,j \in A, k,l \in B} \Bigg\{ &\gamma(\chi_i - \chi_k) + \gamma(\chi_l - \chi_j) - 2\overline{\gamma^{\Omega \cap H_{\{A,B\}}}}\Big(\frac{\chi_i - \chi_k + \chi_l - \chi_j}{2}\Big), \\ &\gamma(\chi_i - \chi_l) + \gamma(\chi_k - \chi_j) - 2\overline{\gamma^{\Omega \cap H_{\{A,B\}}}}\Big(\frac{\chi_i - \chi_l + \chi_k - \chi_j}{2}\Big)\Bigg\}.\end{aligned}$$

Since γ satisfies the triangle inequality, we obtain

$$\overline{\gamma^{\Omega \cap H_{\{A,B\}}}}\Big(\frac{\chi_i - \chi_k + \chi_l - \chi_j}{2}\Big) = \frac{1}{2}(\gamma(\chi_i - \chi_j) + \gamma(\chi_l - \chi_k))$$

and

$$\overline{\gamma^{\Omega \cap H_{\{A,B\}}}}\Big(\frac{\chi_i - \chi_l + \chi_k - \chi_j}{2}\Big) = \frac{1}{2}(\gamma(\chi_i - \chi_j) + \gamma(\chi_k - \chi_l)).$$

Thus, we have

$$\begin{aligned}\tilde{c}_{H_{\{A,B\}}}(\gamma) &= \frac{\sqrt{|X|}}{2} \min_{i,j \in A, k,l \in B} \Big\{\gamma(i,k) + \gamma(l,j) - \gamma(i,j) - \gamma(l,k), \gamma(i,l) + \gamma(k,j) - \gamma(i,j) - \gamma(k,l)\Big\} \\ &= \frac{\sqrt{|X|}}{2} \min_{i,j \in A, k,l \in B} \left\{ \min \left\{ \begin{matrix} \gamma(i,k) + \gamma(j,l), \\ \gamma(i,l) + \gamma(j,k) \end{matrix} \right\} - \gamma(i,j) - \gamma(k,l) \right\} \\ &= \sqrt{|X|}\, b^\gamma_{\{A,B\}}.\end{aligned}$$

Therefore, we have $c_{H_{\{A,B\}}}(\overline{\gamma}) = \max\{0, \tilde{c}_{H_{\{A,B\}}}(\gamma)\} = \sqrt{|X|} \max\{0, b^\gamma_{\{A,B\}}\}$. $\square$

As a result of the discrete split decomposition of metrics on Ω, the next theorem is obtained.

Theorem 9.6. *Let $\gamma : X \times X \to \mathbb{R}$ be a metric. Then γ can be decomposed as*

$$\gamma = \sum_{\sigma \in \Sigma_b(\gamma)} b_\sigma^\gamma \xi_\sigma + \gamma', \tag{9.5}$$

where $\gamma' : X \times X \to \mathbb{R}$ is a metric with $b_{\sigma'}^{\gamma'} \leq 0$ for any X-split σ'.

Proof. By applying Theorem 8.7 with $\mathcal{H}_\Omega^\diamond$ to γ, we obtain

$$\gamma = \sum_{H \in \mathcal{H}(\overline{\gamma}) \cap \mathcal{H}_\Omega^\diamond} c_H(\overline{\gamma}) l_H^\Omega + \gamma'.$$

From Theorem 9.5, it is immediate that $\Sigma_b(\gamma) = \{\{A, B\} \mid \{A, B\} : \text{an } X\text{-split}, H_{\{A,B\}} \in \mathcal{H}(\overline{\gamma}) \cap \mathcal{H}_\Omega^\diamond\}$ and it follows that

$$\gamma = \sum_{\sigma \in \Sigma_b(\gamma)} \sqrt{|X|}\, b_\sigma^\gamma l_{H_\sigma}^\Omega + \gamma'.$$

It is clear that $\sqrt{|X|}\, l_{H_\sigma}^\Omega$ is the split metric ξ_σ on Ω. As a result, we obtain (9.5). In addition, by Theorem 8.7, γ' is convex-extensible on Ω. Hence, γ' is a metric on X by Lemma 9.2. $\square$

By the property of the Buneman index as in Lemma 4.2 and Proposition 8.10, we obtain the following propositions.

Proposition 9.7. *A metric γ is a tree metric if and only if $\overline{\gamma}$ is decomposed as*

$$\overline{\gamma} = \sum_{\sigma \in \Sigma_b(\gamma)} \sqrt{|X|}\, b_\sigma^\gamma l_{H_\sigma} + \delta_{\mathrm{cone}\,\Omega}.$$

Proposition 9.8. *Let $\mathcal{H}$ be a subset of $\mathcal{H}_\Omega^\diamond$, and let $\Sigma = \{\{A, B\} \mid \{A, B\} : \text{an } X\text{-split}, H_{\{A,B\}} \in \mathcal{H}\}$. Then, $\mathcal{H}$ is Ω-admissible if and only if Σ is pairwise compatible.*

Figure 2 illustrates the polyhedral split decomposition of a metric on X with $|X| = 3$. It is known that every 3-point metric can be represented by split metrics, i.e., $\gamma' = 0$ in the decomposition (9.5).

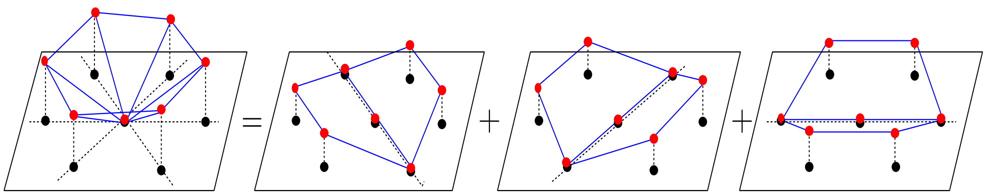

Figure 2: The polyhedral split decomposition of a metric on $X = \{i, j, k\}$.

Remark 9.9. Our approach can be applied to an asymmetric distance which may take negative values. We denote by $\gamma(i, j)$ the distance from i to j for all $i, j \in X$. We do not necessarily assume that $\gamma(i, j) \geq 0$ and $\gamma(i, j) = \gamma(j, i)$ for all distinct $i, j \in X$. Note, however, that $\gamma(i, i) = 0$ for all $i \in X$. This γ can be regarded as a discrete function on the set Ω by the correspondence:

$$\gamma(\chi_i - \chi_j) \leftarrow \gamma(i, j) \quad (i, j \in X).$$

Since γ is assumed to be convex-extensible on Ω, by Lemma 9.2, γ should satisfy $\gamma(\chi_i - \chi_j) \leq \gamma(\chi_i - \chi_k) + \gamma(\chi_k - \chi_j)$ for all $i, j, k \in X$, which implies γ satisfies the "directional" triangle inequality, i.e., $\gamma(i,j) \leq \gamma(i,k) + \gamma(k,j)$ for all $i, j, k \in X$. This γ is said to be an *asymmetric distance*.

For an asymmetric distance γ, the homogeneous convex extension of γ is represented as the same as in (9.2). Then, we use $\mathcal{H}(\overline{\gamma}) \cap \mathcal{H}^{\diamond}_{\Omega}$ to decompose $\overline{\gamma}$. The quotient $c_{H_{\{A,B\}}}(\overline{\gamma})$ for $H_{\{A,B\}} \in \mathcal{H}(\overline{\gamma}) \cap \mathcal{H}^{\diamond}_{\Omega}$ can be expressed as $c_{H_{\{A,B\}}}(\overline{\gamma}) = \max\{0, \tilde{c}_{H_{\{A,B\}}}(\overline{\gamma})\}$ where

$$\begin{aligned}\tilde{c}_{H_{\{A,B\}}}(\gamma) =& \frac{\sqrt{|X|}}{2} \min_{i,j \in A, k,l \in B} \Big\{ \gamma(\chi_i - \chi_k) + \gamma(\chi_l - \chi_j) - \gamma(\chi_i - \chi_j) - \gamma(\chi_l - \chi_k), \\ & \qquad \gamma(\chi_i - \chi_l) + \gamma(\chi_k - \chi_j) - \gamma(\chi_i - \chi_j) + \gamma(\chi_k - \chi_l) \Big\} \\ =& \frac{\sqrt{|X|}}{2} \min_{i,j \in A, k,l \in B} \Big\{ \gamma(i,k) + \gamma(l,j) - \gamma(i,j) - \gamma(l,k), \gamma(i,l) + \gamma(k,j) - \gamma(i,j) - \gamma(k,l) \Big\}.\end{aligned}$$

Remark 9.10. The tight span of a metric space is the central concept in T-theory [14]. For a metric d, the polyhedron $P(d) \subseteq \mathbb{R}^n$ is defined by

$$P(d) = \big\{ p \in \mathbb{R}^n \mid \langle p, \chi_i + \chi_j \rangle \geq d(\chi_i + \chi_j) \ (i, j \in X) \big\}.$$

By the definition (9.1), $\overline{(-d)}$ is the support function of $-P(d) = Q(-d)$. The *tight span* of metric d is the subset of $P(d)$ defined as

$$T(d) = \{ p \in \mathbb{R}^n \mid \forall i \in X, p(i) = \sup_{j \in X} \{ d(i,j) - p(j) \} \}. \tag{9.6}$$

By the definition (9.6), $T(d)$ is the set of all minimal elements in $P(d)$ relative to the order $p \leq q \iff p(i) \leq q(i)$ for every $i \in X$. The tight span was originally constructed by Isbell in [21] and rediscovered by Dress in [14]; see also [15] and [18]. It is known that $T(d)$ coincides with the union of all bounded faces of $P(d)$ [11, Lemma 1].

The tight span $T(d)$ expresses combinatorial properties of a finite metric space (X, d) in geometric terms. For example, a metric is a tree metric if and only if its tight span is a tree [10]. In this remark, we describe that, if d is a tree metric, we obtain essentially the same set as the tight span $T(d)$ by our decomposition of d as in Theorem 9.6.

By Bandelt and Dress' coherent decomposition [1] or Hirai's split decomposition of a metric d [17, Remark 4.10], $P(d)$ is decomposed as

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

where $Z(d)$ is given by

$$Z(d) = \sum_{\{A,B\} \in \Sigma_i(d)} i^d_{\{A,B\}} ([\chi_A - \chi_B, \chi_B - \chi_A]/2 + \chi_X/2).$$

If the split-prime residue d' is zero, we have $P(d) = Z(d) + \mathbb{R}^n_+$. In this case, the tight span $T(d)$ is the union of the faces of $Z(d)$ whose normal cone contains negative vectors.

By our decomposition of a metric d, $Q(d)$ is decomposed as

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

where $Z'(d)$ is given by

$$Z'(d) = \sum_{\{A,B\} \in \Sigma_b(d)} b^d_{\{A,B\}} ([\chi_A - \chi_B, \chi_B - \chi_A]/2).$$

If the split-prime residue d'' is zero, i.e., d is a tree metric, we have $Q(d) = Z'(d) + \{a\mathbf{1} \mid a \in \mathbb{R}\}$. Moreover, in this case, we have $\Sigma_b(d) = \Sigma_i(d)$ and $b^d_{\{A,B\}} = i^d_{\{A,B\}}$ for each $\{A, B\} \in \Sigma_b(d)$. Therefore, $Z'(d)$ is a translation of $Z(d)$, and, analogously to $Z(d)$, the set of faces of $Z'(d)$ whose normal cone contains negative vectors is a tree as the graph consisting of the 1-skeletons of the set. In fact, it is known that if d is a tree metric, the finite metric space (X, d) can be isometrically embedded into $(T(d), \|\cdot\|_\infty)$ [10].

Figure 3 illustrates $P(d)$ and $Q(d)$ for a 3-point metric d. Since every 3-point metric can be represented only by split metrics, $Z'(d)$ is a translation of $Z(d)$.

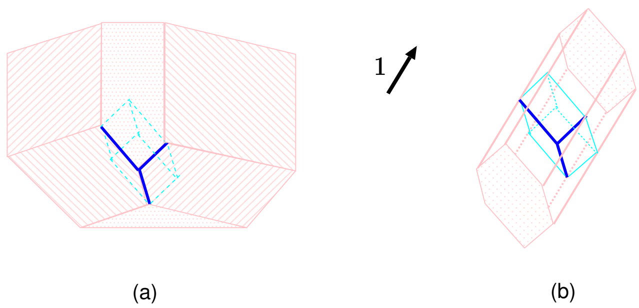

Figure 3: (a) $P(d)$ (pink) and $Z(d)$ (blue and cyan) for a metric d on $X = \{i, j, k\}$ and (b) $Q(d)$ (pink) and $Z'(d)$ (blue and cyan).

Remark 9.11. We reveal the relation between the split fan of Ω and the space of phylogenetic trees. Recall that each cone of the split fan consists of split-decomposable functions on Ω. Obviously, the split fan is a simplicial fan isomorphic to the set of Ω-admissible sets of hyperplanes in $\mathcal{H}^\diamond_\Omega$ as an abstract simplicial complex. More basically, the split fan is isomorphic to the set of pairwise compatible sets of X-splits.

The set of pairwise compatible sets of X-splits was studied by Billera, Holmes, and Vogtmann in [4]. We begin by reviewing their study. An X-split $\{A, B\}$ with $\min\{|A|, |B|\} = 1$ is called a *trivial* X-split. Every trivial X-split is compatible for any X-split. Then, we exclude the trivial X-splits and consider a simplicial complex $\mathbf{T}$ as follows. The vertex set of $\mathbf{T}$ consists of all X-splits $\{A, B\}$ such that A and B have cardinality at least two. We first define a graph whose vertex set corresponds to non-trivial X-splits. Two vertices $\{A, B\}$ and $\{C, D\}$ of the graph are connected by an edge if $\{A, B\}$ and $\{C, D\}$ are compatible. The *space of phylogenetic trees* $\mathbf{T}$ is now defined as the flag complex associated with the graph, that is, each face of $\mathbf{T}$ corresponds to a clique of the graph. Thus, for every face F of $\mathbf{T}$, any pair $\{\{A, B\}, \{C, D\}\} \subseteq F$ is compatible.

Clearly, the set of trivial X-splits is isomorphic to $(n-1)$-simplex, i.e., a simplex with n vertices. The set of pairwise compatible sets of X-splits is isomorphic to the direct product of $(n-1)$-simplex and $\mathbf{T}$, and thus so is the split fan of Ω.

10 Combinatorics of split fans

In this section, we discuss the combinatorics of split fans and give a combinatorial characterization of a K-admissible set of hyperplanes for some particular K, though Definition 8.5 describes a geometric characterization of a K-admissible set of hyperplanes. By the combinatorial characterization, we obtain Proposition 9.8 without Lemma 4.2 and Proposition 8.10. Moreover, the characterization proposes K-admissibility as a new concept for matroids that arise from *vector configurations* and *hyperplane arrangements*. This concept is closely related to *adjoints* of a matroid, which is described in Remark 10.19.

Consider a matroid M_Ξ associated with a set of vectors $\Xi = \{\xi_1, \xi_2, \dots, \xi_k\} \subseteq \mathbb{R}^n$ such that Ξ contains the origin 0 and the vector in the opposite direction from the origin for each vector in $\Xi \setminus \{0\}$. (Without loss of generality, we may assume that $\Xi = -\Xi = \{-\xi_1, -\xi_2, \dots, -\xi_k\}$.) Note that Ω in Section 9 is such a set of vectors. The ground set of M_Ξ is Ξ and the independent sets of M_Ξ consist of all linearly independent subsets of Ξ. Let r be the rank function of M_Ξ. For $U \subseteq \Xi$, the rank $r(U)$ of U is defined as the number of linearly independent vectors in U. In the matroid sense, a hyperplane of M_Ξ is a maximal subset of Ξ having rank $r(\Xi) - 1$.

In this section, we say that a hyperplane H is Ξ-admissible if $\{H\}$ is Ξ-admissible, and so "a set of Ξ-admissible hyperplanes" means that each hyperplane in the set is Ξ-admissible. Another matroid emerges from a hyperplane arrangement $\mathcal{H}$ consisting of Ξ-admissible hyperplanes. In particular, we are interested in the intersection poset $\mathcal{L}(\mathcal{H})$ of the hyperplane arrangement, which will be a geometric lattice.

The combinatorial characterization is based on the fact that each hyperplane of M_Ξ can be identified with a Ξ-admissible hyperplane H, which is described in Proposition 10.13. Given Ξ-admissible hyperplanes $\mathcal{H}$, we can consider the intersection poset $\mathcal{L}(\mathbb{H})$ of hyperplanes $\mathbb{H}$ of M_Ξ corresponding to $\mathcal{H}$ analogously to the intersection poset $\mathcal{L}(\mathcal{H})$. Note that each element of $\mathcal{L}(\mathbb{H})$ is a subset of vectors in Ξ. Then, the characterization is as follows.

Theorem 10.1. *Let $\mathcal{H} \subseteq \mathcal{H}_\Xi$ be a set of Ξ-admissible hyperplanes, and let $\mathbb{H}$ be the set of hyperplanes of M_Ξ corresponding to $\mathcal{H}$. Then, $\mathcal{H}$ is Ξ-admissible if and only if $\mathcal{L}(\mathbb{H})$ is a geometric lattice and the height of U is equal to $r(\Xi) - r(U)$ for each U in $\mathcal{L}(\mathbb{H})$.*

We briefly apply Theorem 10.1 to the set Ω. It is easy to see that M_Ω arises from the two-way directed complete graph $K_n = (X, E)$. A flat of a matroid is an intersection of hyperplanes of the matroid. It is known that each flat of M_Ω can be identified with a partition of X. In particular, each hyperplane of M_Ω corresponds to a bipartition of X, i.e., an X-split. Let H_1 and H_2 be hyperplanes of M_Ω. If the set of Ω-admissible hyperplanes corresponding to H_1 and H_2 is Ω-admissible, the partition $H_1 \cap H_2$ must be composed of three blocks by Theorem 10.1. This implies that X-splits H_1 and H_2 are compatible. Although the pairwise compatibility implies Ω-admissibility, we show it in Section 10.4 since it needs preparation. As a result, we can obtain Proposition 9.8 as a corollary of Theorem 10.1. This is much simpler than the arguments for obtaining Proposition 9.8 in Section 9.

We review the lattice of flats of a matroid in Section 10.1 and hyperplane arrangements in Section 10.2.

10.1 Lattice of flats

Let M be a matroid having ground set E and rank function $r : 2^E \to \mathbb{Z}_+$. Let cl be the function 2^E into 2^E defined by

$$\mathrm{cl}(U) = \{u \in E \mid r(U \cup \{u\}) = r(U)\} \quad (U \subseteq E).$$

This function is called the *closure operator* of M and $\mathrm{cl}(U)$ is called the *closure* of U in M. If $U = \mathrm{cl}(U)$, then U is called a *closed set* or *flat* of M. In particular, a flat U having $r(U) = r(E) - 1$ is called a *hyperplane*. We denote the set of all hyperplanes of M by $\mathbb{H}_M$.

The flats of a matroid has a special structure. For a matroid M, let $\mathcal{L}_M$ denote the set of flats of M ordered by inclusion.

Proposition 10.2 (cf. [24]). *The partially ordered set $\mathcal{L}_M$ forms a lattice with meet "$\wedge$" and join "$\vee$" operations given by*

$$U \wedge V = U \cap V \quad \text{and} \quad U \vee V = \mathrm{cl}(U \cup V)$$

for all flats U and V of M.

In fact, $\mathcal{L}_M$ is a rather special type of lattice. To characterize these matroid lattices, we shall require some more terminology. Let $(P, \leq)$ be a finite partially ordered set. If $u < v$ in P but there is no element w of P such that $u < w < v$, then we say that v *covers* u in P. A *chain* in P from u_0 to u_n is a subset $\{u_0, u_1, \dots, u_n\}$ of P such that $u_0 < u_1 < \cdots < u_n$. The *length* of such a chain is n. The chain is *maximal* if u_i covers u_{i-1} for all $i \in \{1, 2, \dots, n\}$. If for every pair $\{u, v\}$ of elements of P with $u < v$, all maximal chains from u to v have the same length, then P is said to satisfy the *Jordan-Dedekind chain condition*.

If the poset P has an element v such that $v \leq u$ for all u in P, then v is called the *zero* of P and denoted by $\mathbf{0}$. Clearly, the zero of P is unique if it exists. Similarly, if P has an element w such that $w \geq u$ for all u in P, then w is called the *one* of P and denoted by $\mathbf{1}$. The one of P is unique if it exists.

Now suppose that P is a partially ordered set having the zero. An element u is called an *atom* of P if u covers $\mathbf{0}$. The *height* $h(v)$ of an element v of P is the maximum length of a chain from $\mathbf{0}$ to v. Thus, in particular, the atoms of P are precisely the elements of height one.

If a poset P is a finite lattice, P has the zero and the one. In particular, for a matroid M, the zero of $\mathcal{L}_M$ is $\mathrm{cl}(\emptyset)$, while the one is the ground set of M. A finite lattice $\mathcal{L}$ is called *semimodular* if it satisfies the Jordan-Dedekind chain condition and, for every pair u and v of elements of $\mathcal{L}$,

$$h(u) + h(v) \geq h(u \vee v) + h(u \wedge v).$$

A finite lattice $\mathcal{L}$ is called *atomic* if every element is a join of atoms. A *geometric lattice* is a finite atomic semimodular lattice.

The following theorem motivates the lattice-theoretical approach to matroids.

Theorem 10.3 (Birkhoff [3]). *A lattice $\mathcal{L}$ is geometric if and only if it is the lattice of flats of a matroid.*

By the next proposition, for the lattice $\mathcal{L}_M$ of flats of a matroid M, the height $h(U)$ of a flat U corresponds with $r(U)$, where r is the rank function of M.

Proposition 10.4 (cf. [24]). *If U and V be flats of M and $U \subseteq V$, then every maximal chain of flats from U to V has length $r(V) - r(U)$.*

10.2 Hyperplane arrangements

In this paper, a *finite hyperplane arrangement* $\mathcal{A}$ is a finite set of affine hyperplanes in $\mathbb{R}^n$. A hyperplane arrangement $\mathcal{A}$ is *central* if $\bigcap_{H \in \mathcal{A}} H \neq \emptyset$. Equivalently, $\mathcal{A}$ is a translation of a *linear hyperplane arrangement.*

Definition 10.5 (intersection poset). *Let $\mathcal{A}$ be a hyperplane arrangement in $\mathbb{R}^n$, and $\mathcal{L}(\mathcal{A})$ be the set of all nonempty intersections of hyperplanes in $\mathcal{A}$, including $\mathbb{R}^n$ itself as the intersection over the empty set. Define $x \leq y$ in $\mathcal{L}(\mathcal{A})$ if $x \supseteq y$ (as subsets of $\mathbb{R}^n$). In other word, $\mathcal{L}(\mathcal{A})$ is partially ordered by reverse inclusion. We call $\mathcal{L}(\mathcal{A})$ the intersection poset of $\mathcal{A}$.*

Note that $\mathbb{R}^n \in \mathcal{L}(\mathcal{A})$ satisfies $x \geq \mathbb{R}^n$ for all $x \in \mathcal{L}(\mathcal{A})$. Hence, $\mathbb{R}^n$ is the zero of the intersection poset $\mathcal{L}(\mathcal{A})$.

A *graded poset* is defined as a poset P with a function $g : P \to \mathbb{Z}$ from P to the chain of all integers such that

(1) $x > y$ implies $g(x) > g(y)$, and

(2) if x covers y, then $g(x) = g(y) + 1$.

Any graded poset satisfies the Jordan-Dedekind chain condition.

Lemma 10.6 (cf. [29]). *Let $\mathcal{A}$ be a hyperplane arrangement in $\mathbb{R}^n$. Then the intersection poset $\mathcal{L}(\mathcal{A})$ is graded by height equal to codimension, namely, for each $x \in \mathcal{L}(\mathcal{A})$,*

$$h(x) = \operatorname{codim} x = n - \dim x.$$

We give some more terminology about lattices. A *meet-semilattice* is a poset P for which any two elements have a meet. Dually, a *join-semilattice* is a poset for which any two elements have a join.

Lemma 10.7 (cf. [29]). *A finite meet-semilattice L with a unique maximal element $\mathbf{1}$ is a lattice. Dually, a finite join-semilattice L with a unique minimal element $\mathbf{0}$ is a lattice.*

If $\bigcap_{H \in \mathcal{A}} = \emptyset$, then we adjoin $\emptyset$ to $\mathcal{L}(\mathcal{A})$ as the one, so that the augmented intersection poset $\mathcal{L}'(\mathcal{A})$ is obviously a join-semilattice by definition. Since $\mathcal{L}'(\mathcal{A})$ has the zero, $\mathcal{L}'(\mathcal{A})$ is a lattice by Lemma 10.7. Hence, the next proposition claims that $\mathcal{L}(\mathcal{A})$ is a meet-semilattice.

Proposition 10.8 (cf. [29]). *Let $\mathcal{A}$ be a hyperplane arrangement in $\mathbb{R}^n$. Then $\mathcal{L}(\mathcal{A})$ is a meet-semilattice. In particular, every interval $[x, y]$ of $\mathcal{L}(\mathcal{A})$ is a lattice. Moreover, $\mathcal{L}(\mathcal{A})$ is a lattice if and only if $\mathcal{A}$ is central.*

In fact, the following lemma and theorem are well-known.

Lemma 10.9. *Let $\mathcal{A}$ be a central hyperplane arrangement in $\mathbb{R}^n$. Then $\mathcal{L}(\mathcal{A})$ is a semimodular lattice.*

By definition, $\mathcal{L}(\mathcal{A})$ is clearly atomic, from which the next theorem follows.

Theorem 10.10 (cf. [29]). *Let $\mathcal{A}$ be a central hyperplane arrangement in $\mathbb{R}^n$. Then $\mathcal{L}(\mathcal{A})$ is a geometric lattice.*

We close this subsection by introducing an isomorphism of posets. A function $\theta : P \to Q$ from a poset P to a poset Q is called *isotone* if it satisfies

$$x \leq y \Rightarrow \theta(x) \leq \theta(y). \tag{10.1}$$

An isotone function which has an isotone two-sided inverse is called an *isomorphism*. In other word, an isomorphism between two poset P and Q is a bijection which satisfies (10.1) and also

$$\theta(x) \leq \theta(y) \Rightarrow x \leq y.$$

Two posets P and Q are called *isomorphic*, in symbols, $P \cong Q$, if and only if there exists an isomorphism between them.

For a central hyperplane arrangement $\mathcal{A}$, its intersection poset is isomorphic to the lattice of flats of the matroid associated with the normal vectors to the hyperplanes in $\mathcal{A}$.

10.3 Combinatorial characterization of a K-admissible set of hyperplanes

In this subsection, we consider a matroid M_Ξ associated with a set of vectors $\Xi = \{\xi_1, \xi_2, \dots, \xi_k\} \subseteq \mathbb{R}^n$ such that Ξ contains the origin and $\Xi = -\Xi$. For $U \subseteq \Xi$, the rank $r(U)$ of U is defined as the number of linearly independent vectors in U. The next lemma is immediate from the definition of r.

Lemma 10.11. *Let U be a subset of Ξ. Then we have $r(U) = \dim \operatorname{lin} U$.*

We are interested in a relation between flats of M_Ξ and intersections of Ξ-admissible hyperplanes. To distinguish linear spaces in $\mathbb{R}^n$ and flats of M_Ξ, we use small letters for the former and capital letters for the latter.

For simplification, we assume:

Assumption 10.12. $\mathrm{lin}\,\Xi = \mathrm{aff}\,\Xi$ *is fully dimensional, i.e.,* $\dim \mathrm{aff}\,\Xi = n$.

This assumption makes some equalities simple, for example, $\mathrm{lin}\,(h \cap \Xi) = h$ for a Ξ-admissible hyperplane h. Without this assumption, we have $\mathrm{lin}\,(h \cap \Xi) = h \cap \mathrm{lin}\,\Xi$. In fact, the statements in this subsection hold for the set Ω with slight modification.

Proposition 10.13.

(1) For a hyperplane H of matroid M_Ξ, $\mathrm{lin}\,H$ is Ξ-admissible.

(2) For a Ξ-admissible hyperplane h in $\mathbb{R}^n$, $h \cap \Xi$ is a hyperplane of M_Ξ.

Proof. (1) Let H be a hyperplane of M_Ξ. Obviously, $\mathrm{lin}\,H$ satisfies the condition (A1) in Definition 8.5. Moreover, we have

$$\mathrm{cone}\,(\mathrm{lin}\,H \cap \Xi) = \mathrm{lin}\,H, \text{ and } \mathrm{cone}\,\Xi \cap \mathrm{lin}\,H = \mathrm{lin}\,H.$$

Hence, $\mathrm{lin}\,H$ is Ξ-admissible.

(2) Let h be a Ξ-admissible hyperplane. By Ξ-admissibility, we have $\mathrm{cone}\,(h \cap \Xi) = \mathrm{cone}\,\Xi \cap h$. Because $\mathrm{lin}\,(\mathrm{cone}\,(h \cap \Xi)) = \mathrm{lin}\,(h \cap \Xi)$ and $\mathrm{lin}\,(\mathrm{cone}\,\Xi \cap h) = h$, it follows that $\mathrm{lin}\,(h \cap \Xi) = h$. Since $\dim h = n - 1$, we have $r(h \cap \Xi) = \dim \mathrm{lin}\,(h \cap \Xi) = n - 1$. In addition, $h \cap \Xi$ is obviously maximal. Thereby, $h \cap \Xi$ is a hyperplane of M_Ξ. □

By Proposition 10.13, there is a one-to-one correspondence between Ξ-admissible hyperplanes and hyperplanes of M_Ξ. And then, we consider a poset of flats induced by a set of Ξ-admissible hyperplanes.

Definition 10.14. *Let $\mathbb{H} \subseteq \mathbb{H}_{M_\Xi}$ be a set of hyperplanes in M_Ξ. We define $\mathcal{L}(\mathbb{H})$ as the set of all intersections of hyperplanes in $\mathbb{H}$ and ground set Ξ. Define $U \leq V$ in $\mathcal{L}(\mathbb{H})$ if $U \supseteq V$. We call $\mathcal{L}(\mathbb{H})$ the intersection poset of $\mathbb{H}$.*

Note that $\mathcal{L}(\mathbb{H})$ is a poset of flats since the intersection of flats is also a flat. In fact, $\mathcal{L}(\mathbb{H})$ is a lattice. Note that $\mathcal{L}(\mathbb{H})$ is distinct from the lattice of flats of a matroid described in Section 10.1.

Lemma 10.15. *The intersection poset $\mathcal{L}(\mathbb{H})$ is a lattice.*

Proof. Since every flat in $\mathcal{L}(\mathbb{H})$ contains the origin, we have $\bigcap_{H \in \mathcal{L}(\mathbb{H})} H \neq \emptyset$. Hence, any two elements U, V in $\mathcal{L}(\mathbb{H})$ have a join $U \vee V = U \cap V$, that is, $\mathcal{L}(\mathbb{H})$ is a join-semilattice. Because $\mathcal{L}(\mathbb{H})$ has $\mathbf{0}$, which is the ground set Ξ, $\mathcal{L}(\mathbb{H})$ is a lattice by Lemma 10.7. □

For a matroid M_Ξ and its rank function r, we define a function r^* by setting $r^*(U) = r(\Xi) - r(U) = n - r(U)$ for each $U \subseteq \Xi$. Some authors call r^* the *corank* function of M_Ξ. The next lemma is immediate from Lemma 10.11.

Lemma 10.16. *Let U be a subset of Ξ. Then we have $r^*(U) = \mathrm{codim}\,\mathrm{lin}\,U$.*

The next proposition allows us to consider only linear spaces in the intersection poset of Ξ-admissible hyperplanes.

Proposition 10.17. *Let $\mathcal{H} \subseteq \mathcal{H}_\Xi$ be a set of Ξ-admissible hyperplanes. Then, $\mathcal{H}$ is Ξ-admissible if and only if each $F \in \mathcal{A}(\mathcal{H})$ such that $F \cup -F$ is a linear subspace in $\mathbb{R}^n$ satisfies the condition (A2) in Definition 8.5.*

Proof. The only-if part is obvious. We show the if part. For $G \in \mathcal{A}(\mathcal{H})$, let $\mathcal{A}'$ be the set of all faces $F \in \mathcal{A}(\mathcal{H})$ such that $F \subseteq G$ and $F \cup -F$ is a linear subspace in $\mathbb{R}^n$. Because $G \in \mathcal{A}(\mathcal{H})$ is a cone, any element $x \in G$ can be represented as a nonnegative combination of elements belonging to some face in $\mathcal{A}'$. Since every $F \in \mathcal{A}'$ satisfies the condition (A2), any element $x \in F$ can be expressed as a nonnegative combination of elements in $\Xi \cap F \subseteq \Xi \cap G$. Therefore, any element of $x \in G \cap \operatorname{cone}\Xi (= G)$ can be represented as a nonnegative combination of $\Xi \cap G$, which results in $\operatorname{cone}\Xi \cap G = \operatorname{cone}(G \cap \Xi)$. $\square$

Let $\mathcal{H} \subseteq \mathcal{H}_\Xi$ be a set of Ξ-admissible hyperplanes. If there exists $F \in \mathcal{A}(\mathcal{H})$ such that $F \cup -F$ is a linear subspace in $\mathbb{R}^n$ and F does not satisfy the condition (A2), there exists $u \in \mathcal{L}(\mathcal{H})$ such that $\operatorname{cone}(\Xi \cap u) \subsetneq \operatorname{cone}\Xi \cap u$ since $\Xi = -\Xi$.

We are prepared to prove Theorem 10.1, which is restated in the following because $\mathcal{L}(\mathbb{H})$ is obviously atomic.

Theorem 10.1. *Let $\mathcal{H} \subseteq \mathcal{H}_\Xi$ be a set of Ξ-admissible hyperplanes, and let $\mathbb{H} \subseteq \mathbb{H}_{M_\Xi}$ be the set of hyperplanes corresponding to $\mathcal{H}$. Then, $\mathcal{H}$ is Ξ-admissible if and only if $\mathcal{L}(\mathbb{H})$ is a semimodular lattice graded by r^*.*

Theorem 10.1 lets us conclude that a Ξ-admissible set of hyperplanes depends only on the matroid M_Ξ. If $\mathcal{H}$ is a Ξ-admissible set of hyperplanes, the corresponding lattice $\mathcal{L}(\mathbb{H})$ is graded by r^*. Hence, the opposite (or, order dual) lattice of $\mathcal{L}(\mathbb{H})$ can be naturally embedded into the lattice of flats of M_Ξ.

The proof of Theorem 10.1. First, we show the only-if part. Since $\mathcal{H}$ is central, $\mathcal{L}(\mathcal{H})$ is a semimodular lattice by Lemma 10.9. Hence, it suffices for showing semimodularity that $\mathcal{L}(\mathbb{H})$ is isomorphic to $\mathcal{L}(\mathcal{H})$. Let U be a flat in $\mathcal{L}(\mathbb{H})$. By Definition 10.14 and Proposition 10.13, U can be represented as follows:

$$U = \bigcap_{h \in \mathcal{H}'} (\Xi \cap h) = \Xi \cap (\bigcap_{h \in \mathcal{H}'} h)$$

for some $\mathcal{H}' \subseteq \mathcal{H}$. Note that $\mathcal{H}'$ is not necessarily unique while $\bigcap_{h \in \mathcal{H}'} h$ is unique. It is clear that $\bigcap_{h \in \mathcal{H}'} h$ is in $\mathcal{L}(\mathcal{H})$. The map $S : U \mapsto \bigcap_{h \in \mathcal{H}'} h$ is obviously injective map from $\mathcal{L}(\mathbb{H})$ to $\mathcal{L}(\mathcal{H})$.

Let u be a linear space in $\mathcal{L}(\mathcal{H})$. By Definition 10.5, u can be represented as follows:

$$u = \bigcap_{h \in \mathcal{H}'} h$$

for some $\mathcal{H}' \subseteq \mathcal{H}$. By Ξ-admissibility of $\mathcal{H}$,

$$\operatorname{cone}\Xi \cap u = \operatorname{cone}(\Xi \cap u) = \operatorname{cone}(\Xi \cap (\bigcap_{h \in \mathcal{H}'} h)) = \operatorname{cone}(\bigcap_{h \in \mathcal{H}'} (\Xi \cap h)).$$

Because $\operatorname{lin}(\operatorname{cone}\Xi \cap u) = u$ and $\operatorname{lin}(\operatorname{cone}(\bigcap_{h \in \mathcal{H}'} (\Xi \cap h))) = \operatorname{lin}(\bigcap_{h \in \mathcal{H}'} (\Xi \cap h))$, we have

$$u = \operatorname{lin}(\bigcap_{h \in \mathcal{H}'} (\Xi \cap h)).$$

Since $\Xi \cap h$ is a hyperplane of M_Ξ for each $h \in \mathcal{H}'$, it follows that $\bigcap_{h \in \mathcal{H}'} (\Xi \cap h)$ is a flat and in $\mathcal{L}(\mathbb{H})$. The map $T : u \mapsto \bigcap_{h \in \mathcal{H}'} (\Xi \cap h)$ is obviously injective map from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathbb{H})$.

Therefore, the map S is a bijective map from $\mathcal{L}(\mathbb{H})$ to $\mathcal{L}(\mathcal{H})$, and the map T is a bijective map from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathbb{H})$. It is obvious that S and T are isomorphisms, that is, $\mathcal{L}(\mathbb{H})$ is isomorphic to $\mathcal{L}(\mathcal{H})$.

We show that $\mathcal{L}(\mathbb{H})$ is graded by the corank of M_Ξ. Let U be a flat of M_Ξ, and let $u = S(U) \in \mathcal{L}(\mathcal{H})$. We denote the height of $\mathcal{L}(\mathbb{H})$ and that of $\mathcal{L}(\mathcal{H})$ by k and k', respectively. By Lemma 10.6, we have $k'(u) = \operatorname{codim} u$. From the definition of the map S, it follows that $U = \Xi \cap u$. Since $\operatorname{lin} U = \operatorname{lin}(\Xi \cap u) = \operatorname{lin}(\operatorname{cone}(\Xi \cap u)) = \operatorname{lin}(\operatorname{cone}\Xi \cap u) = u$ by Ξ-admissibility, we have $\operatorname{lin} U = u$. Thus, we have $r^*(U) = \operatorname{codim}(\operatorname{lin} U) = \operatorname{codim} u$ by Lemma 10.16. Since $\mathcal{L}(\mathbb{H})$ is isomorphic to $\mathcal{L}(\mathcal{H})$, we have $k(U) = k'(u) = \operatorname{codim} u = r^*(U)$.

Next we show the if part. Suppose that $\mathcal{H}$ is not Ξ-admissible. By Proposition 10.17, there exists $u \in \mathcal{L}(\mathcal{H})$ such that

$$\operatorname{cone}(\Xi \cap u) \subsetneq \operatorname{cone}\Xi \cap u.$$

Since $\operatorname{lin}(\operatorname{cone}(\Xi \cap u)) = \operatorname{lin}(\Xi \cap u)$ and $\operatorname{lin}(\operatorname{cone}\Xi \cap u) = u$, we have

$$\operatorname{codim}\operatorname{lin}(\Xi \cap u) > \operatorname{codim} u.$$

Note that $\Xi \cap u$ is a flat of M_Ξ. By Lemma 10.6 and Lemma 10.16,

$$k(\Xi \cap u) = r^*(\Xi \cap u) = \operatorname{codim}\operatorname{lin}(\Xi \cap u) > \operatorname{codim} u = k'(u). \tag{10.2}$$

Suppose that $u = \bigvee_{h\in\mathcal{H}'} h = \bigcap_{h\in\mathcal{H}'} h$ for some $\mathcal{H}' \subseteq \mathcal{H}$ and that u is one of the lowest elements which do not satisfy the condition (A2) in Definition 8.5. Note that $\mathcal{H}'$ has cardinality at least two since each $h \in \mathcal{H}$ is Ξ-admissible. Without loss of generality, we may assume that $k'(u) > k'(\bigvee_{h\in\mathcal{H}'\setminus\{h'\}} h)$ for any $h' \in \mathcal{H}'$. We choose a hyperplane $h_0 \in \mathcal{H}'$. Since $h_0 \wedge (\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} h) = \mathbb{R}^n$, it follows from the semimodularity of $\mathcal{L}(\mathcal{H})$ that

$$k'(\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} h) + k'(h_0) \geq k'(\mathbb{R}^n) + k'(u). \tag{10.3}$$

Because $k'(u) > k'(\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} h)$, the inequality (10.3) gives $k'(u) = k'(\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} h) + 1$. Moreover, by the inequality (10.2), we have $k(\Xi\cap u) > k'(\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} h) + 1$. Since u is one of the lowest elements which violate the condition (A2) in Definition 8.5, we know that

$$k(\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} (\Xi \cap h)) = k'(\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} h).$$

From the semimodularity of $\mathcal{L}(\mathbb{H})$, it follows that

$$k(\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} (\Xi \cap h)) + k(\Xi \cap h_0) \geq k(\Xi) + k(\Xi \cap u).$$

Therefore, we have

$$k'(\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} h) + 1 \geq k(\Xi \cap u) > k'(\bigvee_{h\in\mathcal{H}'\setminus\{h_0\}} h) + 1,$$

which is a contradiction. □

Remark 10.18. Similar arguments can be made for a vector set, say $\check{\Xi}$, satisfying Assumption 8.3, that is, an acyclic point configuration in terms of oriented matroids. In this case, a hyperplanes of $M_{\check{\Xi}}$ such that its linear hull contains a facet of $\operatorname{conv}\check{\Xi}$ does not correspond to a $\check{\Xi}$-admissible hyperplane. Moreover, the intersection poset $\mathcal{L}(\mathbb{H})$ of hyperplanes $\mathbb{H}$ of $M_{\check{\Xi}}$ is not necessarily a lattice. However, $\widehat{\mathcal{L}}(\mathbb{H}) = \mathcal{L}(\mathbb{H}) \cup \{\mathbf{1}\}$, i.e, the poset $\mathcal{L}(\mathbb{H})$ with the maximal element $\mathbf{1}$ adjoined is a lattice. The lattice $\widehat{\mathcal{L}}(\mathbb{H})$ is isomorphic to the intersection poset $\mathcal{L}(\mathcal{H})$ of the corresponding $\check{\Xi}$-admissible hyperplanes $\mathcal{H}$ and graded by r^* if $\mathcal{H}$ is $\check{\Xi}$-admissible.

Remark 10.19. We here introduce adjoints, which are "duals" of a matroid in the lattice-theoretical sense. For a lattice L, we denote the opposite lattice by L^{op}. An *adjoint* of a matroid M is defined as a matroid M^{ad} of the same rank, such that there is an order preserving embedding

$$L^{\mathrm{op}} \hookrightarrow L^{\mathrm{ad}}$$

which identifies the coatoms of L (hyperplanes of M) with the atoms of L^{ad}. Here we think of M and M^{ad} as simple matroids with geometric lattices L and L^{ad}, respectively [5]. In general, adjoints are not unique and they may also fail to exist, which begets the problem whether an adjoint exists or not and to find it if it exists. However, a geometric lattice L such that its elements correspond to the set of linear subspaces spanned by a finite set of points has an adjoint, i.e., a matroid arising from the finite set of points. If a set of hyperplanes $\mathcal{H}$ is Ξ-admissible, $\mathcal{L}(\mathcal{H})$ is such a geometric lattice. This intimates that $\mathcal{L}(\mathbb{H})$ can be embedded into the lattice $\mathcal{L}_{M_\Xi}$ of flats of M_Ξ because $\mathcal{L}_{M_\Xi}$ includes the lattice of an adjoint of the matroid having $\mathcal{L}(\mathcal{H})$. Note that, however, our aim is to find a Ξ-admissible set of hyperplanes instead of adjoints. Roughly speaking, we have known an adjoint M^{ad} and we are seeking M.

10.4 Application to the set Ω

We apply Theorem 10.1 to the set Ω. The matroid M_Ω arises from the two-way directed complete graph $K_n = (X, E)$. Moreover, every flat of M_Ω can be identified as a partition of X as follows. If F is a flat of M_Ω, we denote by π_F the partition of X in which i and j are in the same block if and only if the vector $\chi_i - \chi_j$ is in F. This correspondence $F \mapsto \pi_F$ determines a map from the set of flats of M_Ω into the set of partitions of X. Moreover, this map is a bijection. Indeed, the map is an isomorphism from $\mathcal{L}_{M_\Omega}$ to the set of partitions of X, where, for partitions U and V, we define $U \leq V$ if U is a refinement of V, that is, every block of U is contained in a block of V. We call the set of partitions the *partition lattice* of X. It is easy to see that hyperplanes of M_Ω correspond to bipartitions of X, i.e., X-splits. Hence, Proposition 9.3 follows from Proposition 10.13.

In order to obtain Proposition 9.8, we show the following lemma.

Lemma 10.20. *Let $\mathbb{H}$ be a set of hyperplanes of M_Ω, and let Σ be the set of X-splits corresponding to $\mathbb{H}$. Then, $\mathcal{L}(\mathbb{H})$ is a semimodular lattice graded by r^* if and only if Σ is pairwise compatible.*

Proof. We first show the only-if part. Let $\{H_1, H_2\} \subseteq \mathbb{H}$. For H_1 and H_2, the semimodularity of $\mathcal{L}(\{H_1, H_2\})$ graded by r^* implies that $r^*(H_1 \vee H_2) = r^*(H_1 \cap H_2) = 2$. In terms of "partition", this condition means that the partition corresponding to $H_1 \cap H_2$ is composed of three blocks, i.e., the partitions corresponding to H_1 and H_2 are compatible. For this reason, Σ must be pairwise compatible.

Next we show the if part. We regard $\mathcal{L}(\mathbb{H})$ as the lattice of partitions of X. By definition, $r^*(U)$ is the number of blocks in U minus one for each $U \in \mathcal{L}(\mathbb{H})$. We begin by exhibiting that the height of U is equal to $r^*(U)$ for each $U \in \mathcal{L}(\mathbb{H})$. This follows from showing that $r^*(V) = r^*(U) + 1$ in the case that V covers U. Since V covers U, there exists an X-split $H \in \mathbb{H}$ such that $U \vee H = V$, where $U \vee H$ is the common refinement of U and H. Moreover, by pairwise compatibility of $\mathbb{H}$, the number of blocks in $U \vee H$ is one more than that of U, that is, $r^*(V) = r^*(U) + 1$.

We then show the semimodularity of $\mathcal{L}(\mathbb{H})$. Let $U = \bigvee_{H \in \mathbb{H}'} H \in \mathcal{L}(\mathbb{H})$ for a subset $\mathbb{H}' \subseteq \mathbb{H}$. Then, by pairwise compatibility of $\mathbb{H}$, we have $r^*(U) \leq |\mathbb{H}'|$. In particular, equality holds for some $\mathbb{H}'' \subseteq \mathbb{H}$, i.e., $U = \bigvee_{H \in \mathbb{H}''} H$ and $r^*(U) = |\mathbb{H}''|$.

Let $V = \bigvee_{H' \in \mathbb{H}'} H'$ for some $\mathbb{H}' \subseteq \mathbb{H}$ with $r^*(V) = |\mathbb{H}'|$, and let $W = \bigvee_{H'' \in \mathbb{H}''} H''$ for some $\mathbb{H}'' \subseteq \mathbb{H}$ with $r^*(W) = |\mathbb{H}''|$. We choose $\mathbb{H}'$ and $\mathbb{H}''$ such that $|\mathbb{H}' \cap \mathbb{H}''|$ is maximal for those subsets of $\mathbb{H}$. Since $V \vee W = \bigvee_{H \in \mathbb{H}' \cup \mathbb{H}''} H$, we have

$$\begin{aligned} r^*(V \vee W) &\leq |\mathbb{H}' \cup \mathbb{H}''| \\ &= |\mathbb{H}'| + |\mathbb{H}''| - |\mathbb{H}' \cap \mathbb{H}''| \\ &= r^*(V) + r^*(W) - |\mathbb{H}' \cap \mathbb{H}''|. \end{aligned}$$

Then, we show that $V \wedge W = \bigvee_{H \in \mathbb{H}' \cap \mathbb{H}''} H$. Indeed, it follows from this that $r^*(V \wedge W) \leq |\mathbb{H}' \cap \mathbb{H}''|$, and the semimodularity of $\mathcal{L}(\mathbb{H})$ is immediate. Because $V \wedge W = \bigvee_{Y \leq V, Y \leq W} Y$ and $H \leq V$ and $H \leq W$ for each $H \in \mathbb{H}' \cap \mathbb{H}''$, we have $V \wedge W \geq \bigvee_{H \in \mathbb{H}' \cap \mathbb{H}''} H$. Suppose that the inequality is strict, that is, $V \wedge W > \bigvee_{H \in \mathbb{H}' \cap \mathbb{H}''} H$, which means that $V \wedge W$ is a finer partition of X than $\bigvee_{H \in \mathbb{H}' \cap \mathbb{H}''} H$. Hence, there exists an X-split $H_0 \in \mathbb{H} \setminus (\mathbb{H}' \cup \mathbb{H}'')$ such that $V \wedge W \geq (\bigvee_{H \in \mathbb{H}' \cap \mathbb{H}''} H) \vee H_0$. Since $H_0 \leq V$ and $H_0 \leq W$, there exist $\mathbb{H}' \cup \mathbb{H}'' \cup \{H_0\} \subseteq \hat{\mathbb{H}} \subseteq \mathbb{H}$ such that $V = \bigvee_{H \in \hat{\mathbb{H}}} H$ and $r^*(V) = |\hat{\mathbb{H}}|$ and $\mathbb{H}' \cup \mathbb{H}'' \cup \{H_0\} \subseteq \check{\mathbb{H}} \subseteq \mathbb{H}$ such that $W = \bigvee_{H \in \check{\mathbb{H}}} H$ and $r^*(W) = |\check{\mathbb{H}}|$. This contradicts the maximality of $|\mathbb{H}' \cap \mathbb{H}''|$. □

As a result, we can obtain Proposition 9.8 as a corollary of Theorem 10.1. This is simpler than the arguments for obtaining Proposition 9.8 in Section 9.

Figure 4 shows the set Ω with $X = \{i, j, k, l\}$. In Figure 4 (a), we denote by st the vector $\chi_s - \chi_t$ for each $s, t \in X$. Figure 4 (b) illustrates one of the hyperplanes of M_Ω by the red points and (c) shows (the intersection of) the corresponding hyperplane (with $\operatorname{conv}\Omega$) in $\mathbb{R}^n$ by the blue region.

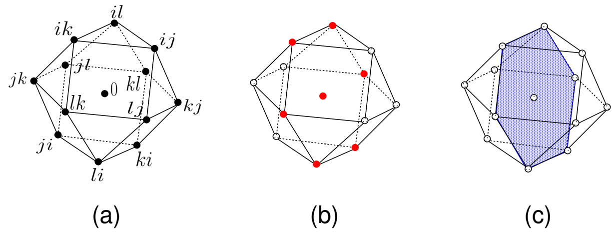

Figure 4: (a) $\operatorname{conv}\Omega$, (b) a hyperplane of M_Ω as shown by the red points, and (c) the hyperplane corresponding to the hyperplane in (b) as shown by the blue region.

Figure 5 illustrates the intersection poset of a hyperplane arrangement, and Figure 6 shows the lattice of flats generated by joining the hyperplanes of M_Ω corresponding to the hyperplanes in Figure 5. Because the lattice in Figure 6 is isomorphic to the lattice in Figure 5 and is graded by the corank of M_Ω, the set of hyperplanes in Figure 5 is Ω-admissible by Theorem 10.1.

Figure 7 illustrates (a) a lattice of flats and (b) the intersection poset of the corresponding hyperplanes. Although the lattice (a) is isomorphic to the intersection poset (b), the lattice (a) is not grade by the corank of M_Ω. Therefore, the set of hyperplanes in (b) is not Ω-admissible by Theorem 10.1. Indeed, the maximal element of the lattice (b) does not satisfy the condition (A2) in Definition 8.5.

11 Conclusion

We have shown that Buneman's method can be understood as the polyhedral split decomposition of the convex extension of a metric which is regarded as a discrete function on $\Omega = \{\chi_i - \chi_j \mid i, j \in X\}$. In [26], Semple and Steel say that Theorem 4.4 can be interpreted as the continuity of Buneman's method and the property is important and remarkable. We have provided Theorem 9.6 as a comprehensible explanation about the property in geometric terms. Moreover, we have given

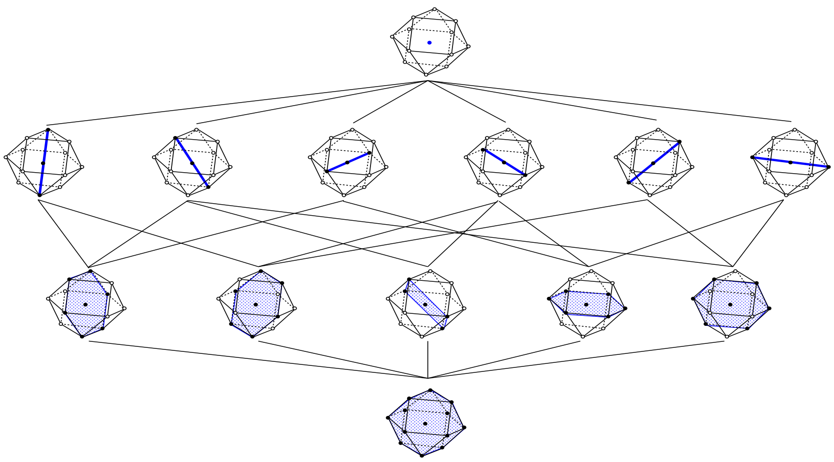

Figure 5: An intersection poset $\mathcal{L}(\mathcal{H})$.

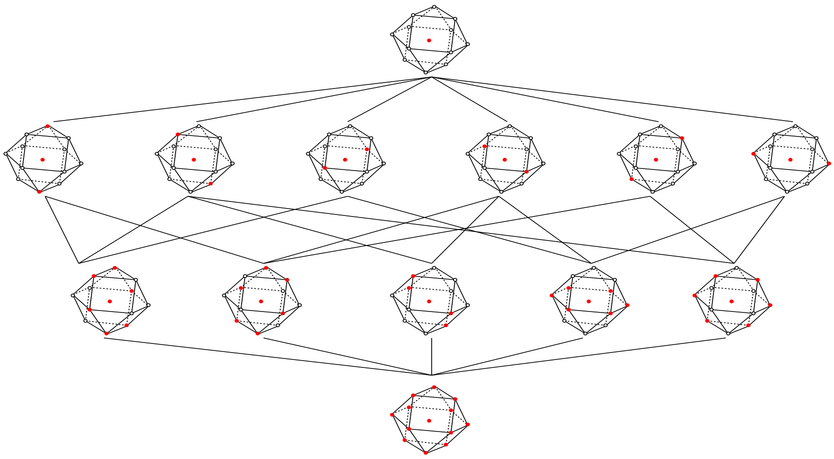

Figure 6: The lattice $\mathcal{L}(\mathbb{H})$ generated by joining the hyperplanes of M_Ω corresponding to the the hyperplanes in Figure 5.

a combinatorial characterization for split fans by exploring the geometric lattice of the hyperplane arrangement obtained from a split-decomposable function and the matroid associated with the vector configuration as the domain of the function. The combinatorial characterization claims that the split fan of the vector configuration depends only on the matroid. In the case of Ω, the split fan $\mathrm{SF}(\Omega)$ coincides with a well-known complex: the space of phylogenetic trees $\mathbf{T}$. Our result designates that $\mathrm{SF}(\Omega)$ is isomorphic to the direct product of a simplex and $\mathbf{T}$.

Several modifications for Buneman's method are proposed [2, 7, 22]. It would be interesting to provide geometric interpretation to them.

Acknowledgments

The author is very grateful to Hiroshi Hirai for his valuable discussions and useful comments. The author is indebted to Satoru Iwata for his guidance in this paper. The author also thanks Kazuo Mutora and Kazuhisa Makino for their helpful suggestions. The present work is supported by the 21st Century COE Program on Information Science and Technology Strategic Core.

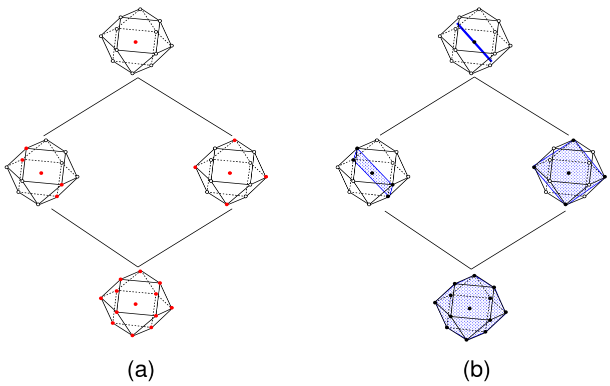

Figure 7: (a) a lattice of flats and (b) the intersection poset of the corresponding hyperplanes.

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, pp. 47–105.

[2] V. Berry and D. Bryant: Faster Reliable Phylogenetic Analysis. *Proceedings of the third annual international conference on Computational molecular biology*, 1999, pp. 59–68.

[3] G. Birkhoff: Abstract Linear Dependence and Lattices. *American Journal of Mathematics*, vol. 57 (1935), no. 4, pp. 800–804.

[4] L. J. Billera, S. P. Holmes, and K. Vogtmann: Geometry of the Space of Phylogenetic Trees. *Advances in Applied Mathematics*, 27 (2001), pp. 733–767.

[5] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. Ziegler: *Oriented Matroids*. Encyclopedia of Mathematics and its Applications, vol. 46, Cambridge University Press, Cambridge, 1993.

[6] E. D. Bolker: A Class of Convex Bodies. *Transactions of the American Mathematical Society*, vol. 145 (1969), pp. 323–345.

[7] D. Bryant and V. Berry: A Structured Family of Clustering and Tree Construction Methods. *Advances in Applied Mathematics*, 27 (2001), pp. 705–732.

[8] P. Buneman: The Recovery of Trees from Measures of Dissimilarity. *Mathematics in the Archaeological and Historical Sciences*, F. R. Hodson, D. G. Kendall, and P. Tautu eds., Edinburgh University Press, Edinburgh, pp. 387–395, 1971.

[9] P. Buneman: A Note on the Metric Property of Trees. *Journal of Combinatorial Theory (B)*, 17 (1974), pp. 48–50.

[10] 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, pp. 321–402.

[11] A. W. M. Dress: Towards a Classification of Transitive Group Actions on Finite Metric Spaces. *Advances in Mathematics*, 74 (1989), no. 2, pp. 163–189.

[12] A. W. M. Dress and D. H. Huson: Constructing Splits Graphs. *IEEE Transactions on Computational Biology and Bioinformatics*, vol. 1 (2004), no. 3, pp. 109–115.

[13] A. Dress, D. Huson, and V. Moulton: Analyzing and Visualizing Sequence and Distance Data Using SplitsTree. *Discrete Applied Mathematics*, 71 (1996), pp. 95–109.

[14] A. Dress, V. Moulton, and W. Terhalle: T-theory: An overview. *European Journal of Combinatorics*, 17 (1996), pp. 161–175.

[15] A. Dress and W. Terhalle: The Tree of Life and Other Affine Building. *Documenta Mathematica*, Extra Vol. ICM III (1998), pp. 565–574 (electronic).

[16] H. Hirai: Geometric Study on the Split Decomposition of Finite Metrics. *RIMS Preprint 1459*, Kyoto University, 2004.

[17] H. Hirai: A Geometric Study of the Split Decomposition. *Discrete and Computational Geometry*, vol. 36 (2006), no. 2, pp. 331–361.

[18] H. Hirai: Characterization of the Distance between Subtrees of a Tree by the Associated Tight Span. *Annals of Combinatorics*, 10 (2006), pp. 111–128.

[19] D. H. Huson: SplitsTree: Analyzing and Visualizing Evolutionary Data. *Bioinformatics*, vol. 14 (1998), no. 1, pp. 68–73.

[20] D. H. Huson and D. Bryant: Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution*, 23(2), pp. 254–267, 2006.

[21] J. R. Isbell: Six Theorems about Injective Metric Spaces. *Commentarii Mathematici Helvetici*, 39 (1964), pp. 65–76.

[22] V. Moulton and M. Steel: Retractions of Finite Distance Functions onto Tree Metrics. *Discrete Applied Mathematics*, 91 (1999), pp. 215–233.

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

[24] J. G. Oxley: *Matroid Theory*. Oxford University Press, Oxford, 1992.

[25] L. Pachter and B. Sturmfels (Eds.): *Algebraic Statistics for Computational Biology*. Cambridge University Press, Cambridge, 2005.

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

[27] J. M. S. Simões-Pereira: A Note on the Tree Realizability of a Distance Matrix. *Journal of Combinatorial Theory*, 6 (1969), pp. 303–310.

[28] D. Speyer and B. Sturmfels: The Tropical Grassmannian. *Advances in Geometry*, 4 (2004), pp. 389–411.

[29] R. P. Stanley: An Introduction to Hyperplane Arrangements, Fall 2004. MIT OpenCourceWare 18.315, Massachusetts Institute of Technology.

[30] K. A. Zaretskii: Constructing Trees from the Set of Distances between Pendant Vertices. *Uspehi Matematiceskih Nauak*, 20 (1965), pp. 90–92.