Hyperconvexity and Tight Span Theory for Diversities

David Bryant[a,*], Paul F. Tupper[b]

[a]Dept. of Mathematics and Statistics, University of Otago.
[b]Dept. of Mathematics, Simon Fraser University.

Abstract

The tight span, or injective envelope, is an elegant and useful construction that takes a metric space and returns the smallest hyperconvex space into which it can be embedded. The concept has stimulated a large body of theory and has applications to metric classification and data visualisation. Here we introduce a generalisation of metrics, called diversities, and demonstrate that the rich theory associated to metric tight spans and hyperconvexity extends to a seemingly richer theory of diversity tight spans and hyperconvexity.

Keywords: Tight span; Injective hull; Hyperconvex; Diversity; Metric geometry;

1. Introduction

Hyperconvex metric spaces were defined by Aronszajn and Panitchpakdi in [1] as part of a program to generalise the Hahn-Banach theorem to more general metric spaces (reviewed in [2], and below). Isbell [3] and Dress [4] showed that every metric space could be embedded into a minimum hyperconvex space, called the *tight span* or *injective envelope*.

Our aim is to show that the notion of hyperconvexity, the tight span, and much of the related theory can be extended beyond metrics to a class of multi-way metrics which we call diversities. Specifically, a *diversity* is a pair (X, δ) where X is a set and δ is a function from the finite subsets of X to $\Re \cup \{\infty\}$ satisfying

(D1) $\delta(A) \geq 0$, and $\delta(A) = 0$ if and only if $|A| \leq 1$.
(D2) If $B \neq \emptyset$ then $\delta(A \cup C) \leq \delta(A \cup B) + \delta(B \cup C)$.

We prove below that these axioms imply monotonicity:

(D3) If $A \subseteq B$ then $\delta(A) \leq \delta(B)$.

*Corresponding author
Email addresses: david.bryant@otago.ac.nz (David Bryant), pft3@math.sfu.ca (Paul F. Tupper)

We will show that tight span theory adapts elegantly from metric spaces to diversities. The tight span of a metric (X, d) is formed from the set of point-wise minimal functions $f : X \to \Re \cup \{\infty\}$ such that $f(a_1) + f(a_2) \geq d(a_1, a_2)$ for all $a_1, a_2 \in X$. Letting $\mathcal{P}_{\mathrm{f}}(X)$ denote the finite subsets of X, the tight span of a diversity (X, δ) is formed from the set of point-wise minimal functions $f : \mathcal{P}_{\mathrm{f}}(X) \to \Re \cup \{\infty\}$ such that

$$f(A_1) + f(A_2) + \cdots + f(A_k) \geq \delta(A_1 \cup A_2 \cup \cdots \cup A_k)$$

for all finite collections $\{A_1, A_2, \ldots, A_k\} \subseteq \mathcal{P}_{\mathrm{f}}(X)$. We can embed a metric space in its tight span (with the appropriate metric on the tight span); we can embed a diversity in its tight span (with the appropriate diversity on the tight span). Both constructions have characterisations in terms of injective hulls, and both possess a rich mathematical structure.

The motivation for exploring tight spans of diversities was the success of the metric tight span as a tool for classifying and visualising finite metrics, following the influential paper of Dress [4]. The construction provided the theoretical framework for split decomposition [5] and Neighbor-Net [6], both implemented in the SplitsTree package [7] and widely used for visualising phylogenetic data. By looking at diversities, rather than metrics or distances, our hope is to incorporate more information into the analysis and thereby improve inference [8].

Dress et al. [9] coined the term *T-theory* for the field of discrete mathematics devoted to the combinatorics of the tight span and related constructions. Sturmfels [10] highlighted T-theory as one area where problems from biology have led to substantial new ideas in mathematics. Contributions to T-theory include: profound results on optimal graph realisations of metrics [4, 11, 12]; intriguing connections between the Buneman graph, the tight span and related constructions [9, 12–16]; links with tropical geometry and hyperdeterminants [17, 18]; classification of finite metrics [4, 19]; and properties of the tight span for special classes of metrics [20, 21]. Hirai [22] describes an elegant geometric formulation of the tight span. We believe that there will be diversity analogues for many of these metric space results.

Our use of the term *diversity* comes from the appearance of a special case of our definition in work on phylogenetic and ecological diversity [23–25]. However diversities crop up in a broad range of contexts, for example:

1. *Diameter Diversity.* Let (X, d) be a metric space. For all $A \in \mathcal{P}_{\mathrm{f}}(X)$ let

$$\delta(A) = \max_{a,b \in A} d(a, b) = \mathrm{diam}(A).$$

2. L_1 *diversity* For all finite $A \subseteq \Re^n$ define

$$\delta(A) = \sum_i \max_{a,b} \{a_i - b_i : a, b \in A\}.$$

3. *Phylogenetic Diversity.* Let T be a phylogenetic tree with taxon set X. For each finite $A \subseteq X$, the *phylogenetic diversity* of A is the length $\delta(A)$ of the smallest subtree of T connecting taxa in A.

4. *Length of the Steiner Tree.* Let (X, d) be a metric space. For each finite $A \subseteq X$ let $\delta(A)$ denote the minimum length of a Steiner tree connecting elements in A.

The generalisation of metrics to more than two arguments has a long history. There is an extensive literature on 2-metrics (metrics taking three points as arguments) see, for example, [26]. Generalised metrics defined on n-tuples for arbitrary n go back at least to Menger [27], who took the volume of an n-simplex in Euclidean space as the prototype. Recently various researchers have continued the study of such generalised metrics defined on n-tuples, see [28–30] for examples. However, as of yet, a satisfactory theory of tight spans has not yet been developed for these generalisations. Our definition of diversities also differs from this earlier work in that a diversity function is defined on arbitrary finite subsets of a space, rather that tuples of a fixed length.

The structure of this paper is as follows: In Section 2 we develop the basic theory of tight spans on diversities, defining the appropriate diversity for a tight span and showing that every diversity embeds into its tight span. In Section 3 we characterise diversities that are isomorphic to their tight spans. These are the *hyperconvex* diversities, where the concept of hyperconvexity is a direct extension of metric hyperconvexity. We prove that diversity tight spans, like metric tight spans, are injective, and are formally the injective envelope in the category of diversities. In Section 4 we explore in more detail the direct links between diversity tight spans and metric tight spans. We show when the diversity equals the diameter diversity (as defined above) the metric tight span and the resulting diversity tight span are isomorphic. In Section 5 we study the tight span of a phylogenetic diversity, and prove that taking the tight span of a phylogenetic diversity recovers the underlying tree, in the same way that taking the tight span of an additive metric recovers its underlying tree. This theory is developed for *real trees.* Finally, in Section 6 we examine applications of the theory to the classical Steiner Tree problem, extending results of [31] about the embedding of 'abstract' Steiner trees in tight spans.

2. The tight span of a diversity

We begin by establishing some basic properties of diversities. Recall that $\mathcal{P}_{\mathrm{f}}(X)$ denotes all the finite subsets of the set X, and that a diversity is a function $\delta\colon \mathcal{P}_{\mathrm{f}}(X) \to \Re \cup \{\infty\}$ satisfying axioms (D1) and (D2).

Proposition 2.1. *Let (X, δ) be a diversity.*

1. *If $d\colon X \times X \to \Re \cup \{\infty\}$ is defined as $d(x, y) = \delta(\{x, y\})$ then (X, d) is a metric space. We say that (X, d) is the* induced metric *of (X, δ).*
2. *(D3) holds, that is, for $A, B \in \mathcal{P}_{\mathrm{f}}(X)$, if $A \subseteq B$ then $\delta(A) \leq \delta(B)$.*
3. *For $A, B \in \mathcal{P}_{\mathrm{f}}(X)$ if $A \cap B \neq \emptyset$ then $\delta(A \cup B) \leq \delta(A) + \delta(B)$.*

Proof.

1. That $d(x,y)=0$ if and only if $x=y$ follows from (D1). Symmetry is clear. To obtain the triangle inequality, for any $x,y,z \in X$,

$$d(x,z)=\delta(\{x,z\}) \le \delta(\{x,y\})+\delta(\{y,z\}) = d(x,y)+d(y,z),$$

using (D2).

2. First note for any $a \in A$ and $b \in X$ that by (D2) with C empty

$$\delta(A) \le \delta(A \cup \{b\}) + \delta(\{b\}) = \delta(A \cup \{b\}).$$

The more general result follows by induction.

3. Using (D2)

$$\delta(A \cup B) \le \delta(A \cup (A \cap B)) + \delta(B \cup (A \cap B)) = \delta(A) + \delta(B).$$

□

We now state the diversity analogue for the metric tight span.

Definition 2.2. *Let (X,δ) be a diversity. Let P_X denote the set of all functions $f\colon \mathcal{P}_{\mathrm{f}}(X) \to \Re \cup \{\infty\}$ satisfying $f(\emptyset)=0$ and*

$$\sum_{A \in \mathscr{A}} f(A) \ge \delta\Big(\bigcup_{A \in \mathscr{A}} A\Big) \tag{2.1}$$

for all finite $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$. Write $f \preceq g$ if $f(A) \le g(A)$ for all finite $A \subseteq X$. The tight span of (X,δ) is the set T_X of functions in P_X that are minimal under $\preceq$.

Example. Any diversity δ on $X=\{1,2,3\}$ is determined by the four values

$$d_{12}=\delta(\{1,2\}), \quad d_{23}=\delta(\{2,3\}), \quad d_{13}=\delta(\{1,3\}), \quad d_{123}=\delta(\{1,2,3\}).$$

We write $f_i = f(\{i\})$, $f_{ij} = f(\{i,j\})$ and $f_{123} = f(\{1,2,3\})$ for $i,j \in X$. Condition (2.1) then translates to the following set of inequalities:

$$\begin{aligned} f_i &\ge 0 \\ f_{ij} &\ge d_{ij} \\ f_i + f_j &\ge d_{ij} \\ f_{123} &\ge d_{123} \\ f_i + f_{jk} &\ge d_{123} \\ f_1+f_2+f_3 &\ge d_{123} \end{aligned} \tag{2.2}$$

for distinct $i,j,k \in X$. Note we have omitted inequalities like $f_{ij}+f_{jk} \ge d_{123}$ since these are implied by (2.2) and the triangle inequality (D2). The elements of T_X are the minimal f in P_X. Equivalently, T_X are the set of f that satisfy (2.2) and such that for each nonempty $A \subseteq X$, f_A appears in an inequality in (2.2) that is tight.

A straightforward but tedious analysis of the inequalities (which we omit) gives the following characterisation of T_X. Define the three 'external' vertices

$$\begin{aligned} v^{(1)} &= (0, d_{12}, d_{13}) \\ v^{(2)} &= (d_{12}, 0, d_{23}) \\ v^{(3)} &= (d_{13}, d_{23}, 0) \end{aligned}$$

and the four 'internal' vertices

$$\begin{aligned} u^{(0)} &= (d_{123} - d_{23}, d_{123} - d_{13}, d_{123} - d_{12}). \\ u^{(1)} &= u^{(0)} - (\beta, 0, 0) \\ u^{(2)} &= u^{(0)} - (0, \beta, 0) \\ u^{(3)} &= u^{(0)} - (0, 0, \beta), \end{aligned}$$

where $\beta = \max(2d_{123} - d_{12} - d_{23} - d_{13}, 0)$. Let C be the cell complex formed from the line segments $[u^{(1)}, v^{(1)}]$, $[u^{(2)}, v^{(2)}]$, $[u^{(3)}, v^{(3)}]$ and the tetrahedron with vertices $u^{(1)}, \ldots, u^{(4)}$. A case-by-case analysis gives that $f \in T_X$ if and only if $(f_1, f_2, f_3) \in C$, $f_{23} = \max(d_{23}, d_{123} - f_1)$, $f_{13} = \max(d_{13}, d_{123} - f_2)$, $f_{23} = \max(d_{12}, d_{123} - f_3)$, $f_{123} = d_{123}$. If $\beta \leq 0$ then $u^{(0)}$ to $u^{(3)}$ coincide, and the tight span resembles the metric tight span for the induced metric, albeit sitting in a higher dimensional space (Figure 1a). When $\beta > 0$ the tight span resembles a tetrahedron with three spindles branching off, as in Figure 1b. □

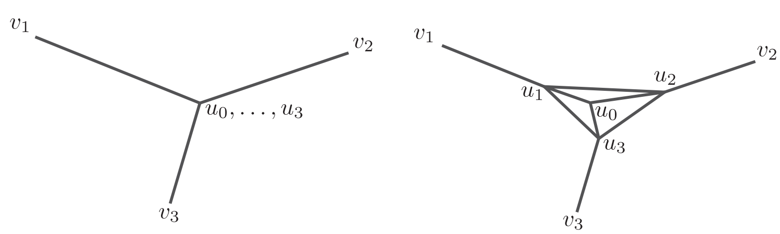

Figure 1: Two examples of the tight span on three points, with different values for $d(\{1, 2, 3\})$. On the left an example where $2d_{123} \leq d_{12} + d_{23} + d_{13}$, and the diversity tight span resembles the tight span of the induced metric. On the right a case with $2d_{123} > d_{12} + d_{23} + d_{13}$.

We now prove a characterisation of the diversity tight span which will be used extensively throughout the remainder of the paper (Theorem 2.3). Note that the diversity function can take value ∞, so expressions like $\delta(A \cup B) - f(B)$ in the statement of Theorem 2.3 can be indeterminate. We follow the convention that suprema and infina are only taken over the elements of a set that are not indeterminate.

Theorem 2.3. *Let $f \colon \mathcal{P}_{\mathrm{f}}(X) \to \Re \cup \{\infty\}$ and suppose $f(\emptyset) = 0$. Then $f \in T_X$ if and only if for all finite $A \subseteq X$,*

$$f(A) = \sup_{\mathscr{B} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta\big(A \cup \bigcup_{B \in \mathscr{B}} B\big) - \sum_{B \in \mathscr{B}} f(B) \colon |\mathscr{B}| < \infty \right\}. \quad (2.3)$$

Proof.
Suppose that $f \in T_X$. As $f(\emptyset) = 0$, the supremum in (2.3) is well defined. For all finite $A \subseteq X$ and all finite $\mathscr{B} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ we have

$$f(A) \geq \delta(A \cup \bigcup_{B \in \mathscr{B}} B) - \sum_{B \in \mathscr{B}} f(B)$$

whenever the right hand side is not indeterminate, giving the required lower bound on $f(A)$. Now suppose that for some finite A_0

$$f(A_0) > \sup_{\mathscr{B} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta\big(A_0 \cup \bigcup_{B \in \mathscr{B}} B\big) - \sum_{B \in \mathscr{B}} f(B) \colon |\mathscr{B}| < \infty \right\}. \qquad (2.4)$$

Define a function $g : \mathcal{P}_{\mathrm{f}}(X) \to \Re_{\geq 0}$ by

$$g(A) = \begin{cases} f(A) & \text{if } A \neq A_0 \\ \sup_{\mathscr{B} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta\big(A_0 \cup \bigcup_{B \in \mathscr{B}} B\big) - \sum_{B \in \mathscr{B}} f(B) \colon |\mathscr{B}| < \infty \right\} & \text{if } A = A_0. \end{cases}$$

Clearly $g \neq f$ and $g \preceq f$. We show that g is in P_X. Let $\mathscr{A}$ be a finite subset of $\mathcal{P}_{\mathrm{f}}(X)$. If $A_0 \notin \mathscr{A}$ then

$$\sum_{A \in \mathscr{A}} g(A) = \sum_{A \in \mathscr{A}} f(A) \geq \delta(\bigcup_{A \in \mathscr{A}} A).$$

If $A_0 \in \mathscr{A}$ then

$$\begin{aligned} \sum_{A \in \mathscr{A}} g(A) &= \sup_{\mathscr{B} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta\big(A_0 \cup \bigcup_{B \in \mathscr{B}} B\big) - \sum_{B \in \mathscr{B}} f(B) \colon |\mathscr{B}| < \infty \right\} + \sum_{B \in \mathscr{A} \setminus \{A_0\}} f(B) \\ &\geq \delta\big(A_0 \cup \bigcup_{B \in \mathscr{A} \setminus \{A_0\}} B\big) \\ &= \delta(\bigcup_{A \in \mathscr{A}} A), \end{aligned}$$

by letting $\mathscr{B} = \mathscr{A} \setminus \{A_0\}$. So $g \in P_X$, $g \neq f$ and $g \preceq f$, contradicting $f \in T_X$. Hence there is no A_0 satisfying (2.4) and that if $f \in T_X$ then (2.3) holds for all finite $A \subseteq X$.

For the converse, suppose that (2.3) holds for all finite $A \subseteq X$. Then $f \in P_X$. Suppose that $g \in P_X$, that $g \preceq f$ and $A \in \mathcal{P}_{\mathrm{f}}(X)$. Then for all finite $\mathscr{B} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ such that

$$\delta\big(A \cup \bigcup_{B \in \mathscr{B}} B\big) - \sum_{B \in \mathscr{B}} f(B)$$

is not indeterminate, we have

$$\delta\big(A \cup \bigcup_{B \in \mathscr{B}} B\big) - \sum_{B \in \mathscr{B}} f(B) \quad \leq \quad \delta\big(A \cup \bigcup_{B \in \mathscr{B}} B\big) - \sum_{B \in \mathscr{B}} g(B) \leq g(A)$$

so that $f(A) \leq g(A)$. Hence f is minimal in P_X. □

The following basic properties of members of T_X will be used subsequently.

Proposition 2.4. *Suppose that $f \in T_X$.*

1. *$f(A) \geq \delta(A)$ for all finite $A \subseteq X$.*
2. *If $A \subseteq B \subseteq X$ and B is finite then $f(A) \leq f(B)$; that is, f is* monotone.
3. *$f(X) = \delta(X)$.*
4. *$f(A \cup C) \leq \delta(A \cup B) + f(B \cup C)$ for all $A, B, C \in \mathcal{P}_{\mathrm{f}}(X)$ with $B \neq \emptyset$.*
5. *$f(A \cup B) \leq f(A) + f(B)$ for all $A, B \in \mathcal{P}_{\mathrm{f}}(X)$; that is, f is* sub-additive.
6. *$f(A) = \sup_B \{\delta(A \cup B) - f(B) : B \in \mathcal{P}_{\mathrm{f}}(X)\}$ for all finite A.*

Proof.
1. Use $\mathscr{A} = \{A\}$ in the definition of P_X.
2. Follows from (2.3) and the monotonicity of δ.
3. From (2.3), $f(X) \leq \delta(X)$ and from part 1, $\delta(X) \leq f(X)$.
4. Let $A, B, C \in \mathcal{P}_{\mathrm{f}}(X)$ with $B \neq \emptyset$. We have

$$f(A \cup C) = \sup_{\mathscr{D} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta(A \cup C \cup \bigcup_{D \in \mathscr{D}} D) - \sum_{D \in \mathscr{D}} f(D) \colon |\mathscr{D}| < \infty \right\}, \quad (2.5)$$

$$f(B \cup C) = \sup_{\mathscr{D} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta(B \cup C \cup \bigcup_{D \in \mathscr{D}} D) - \sum_{D \in \mathscr{D}} f(D) \colon |\mathscr{D}| < \infty \right\}, \quad (2.6)$$

Subtracting (2.6) from (2.5) gives

$$\begin{aligned} f(A \cup C) - f(B \cup C) &\leq \sup_{\mathscr{D} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta\big(A \cup C \cup \bigcup_{D \in \mathscr{D}} D\big) - \delta\big(B \cup C \cup \bigcup_{D \in \mathscr{D}} D\big) \colon |\mathscr{D}| < \infty \right\} \\ &\leq \delta(A \cup B), \end{aligned}$$

by taking $\mathscr{D} = \emptyset$ and using the triangle inequality.
5. Given any $A, B \in \mathcal{P}_{\mathrm{f}}(X)$ and any finite collection $\mathscr{C} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ we have

$$f(A) + f(B) + \sum_{C \in \mathscr{C}} f(C) \geq \delta\big(A \cup B \cup \bigcup_{C \in \mathscr{C}} C\big)$$

so that

$$\begin{aligned} f(A) + f(B) &\geq \sup_{\mathscr{C} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta\big(A \cup B \cup \bigcup_{C \in \mathscr{C}} C\big) - \sum_{C \in \mathscr{C}} f(C) \colon |\mathscr{C}| < \infty \right\} \\ &= f(A \cup B) \end{aligned}$$

by Theorem 2.3.
6. For any finite $\mathscr{B} \subseteq \mathcal{P}_{\mathrm{f}}$, $\sum_{B \in \mathscr{B}} f(A) \geq f(\cup_{B \in \mathscr{B}} B)$. So

$$\sup_{\mathscr{B} \in \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta(A \cup \bigcup_{B \in \mathscr{B}} B) - \sum_{B \in \mathscr{B}} f(B) \colon |\mathscr{B}| < \infty \right\} = \sup_{C \in \mathcal{P}_{\mathrm{f}}(X)} \{\delta(A \cup C) - f(C)\}.$$

$\square$

The distance between any two functions f, g in the metric tight span is given by the l_∞ norm,

$$d_T(f,g) = \sup_{x \in X} |f(x) - g(x)|, \tag{2.7}$$

which Dress [4] shows is equivalent on this set to

$$d_T(f,g) = \sup_{x,y \in X} \{d(x,y) - f(x) - g(y)\}. \tag{2.8}$$

Dress also showed that a metric can be embedded into its tight span using the Kuratowski mapping κ, which takes an element $x \in X$ to the function h_x for which $h_x(y) = d(x,y)$ for all y.

Here we establish the analogous results for the diversity tight span. We define the appropriate function κ from a diversity to its tight span, define the diversity δ_T on the tight span itself, and prove that κ is an embedding. We will show that, up to isomorphism, it is the unique way to define a diversity so that (T_X, δ_T) is the injective hull of (X, δ).

Definition 2.5.

1. *Let (Y_1, δ_1) and (Y_2, δ_2) be two diversities. A map $\pi\colon Y_1 \to Y_2$ is an embedding if it is one-to-one (injective) and for all finite $A \subseteq Y_1$ we have $\delta_1(A) = \delta_2(\pi(A))$. In this case, we say that π embeds (Y_1, δ_1) in (Y_2, δ_2).*
2. *Let (X, δ) be a diversity. For each $x \in X$ define the function $h_x : \mathcal{P}_{\mathrm{f}}(X) \to \Re \cup \{\infty\}$ by*
$$h_x(A) = \delta(A \cup \{x\})$$
for all finite $A \subseteq X$. Let κ be the map taking each $x \in X$ to the corresponding function h_x.
3. *Let (X, δ) be a diversity. Let $\delta_T : \mathcal{P}_{\mathrm{f}}(T_X) \to \Re \cup \{\infty\}$ be the function defined by $\delta_T(\emptyset) = 0$ and*
$$\delta_T(F) = \sup_{\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta\left(\bigcup_{A \in \mathscr{A}} A \right) - \sum_{A \in \mathscr{A}} \inf_{f \in F} f(A) : |\mathscr{A}| < \infty \right\} \tag{2.9}$$
for all finite non-empty $F \subseteq T_X$.

Further manipulations give a form for δ_T analogous to (2.8):

$$\delta_T(F) = \sup_{\{A_f\}_{f \in F}} \left\{ \delta\left(\bigcup_{f \in F} A_f \right) - \sum_{f \in F} f(A_f) : A_f \in \mathcal{P}_{\mathrm{f}}(X) \text{ for all } f \in F \right\},$$

for all finite $F \subseteq \mathcal{P}_{\mathrm{f}}(T_X)$. We can also re-express (2.9) in a form closer to (2.7):

Lemma 2.6. *If $f \in F$ then*

$$\delta(F) = \sup_{\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ f\left(\bigcup_{A \in \mathscr{A}} A \right) - \sum_{A \in \mathscr{A}} \inf_{g \in F \setminus \{f\}} g(A) : |\mathscr{A}| < \infty \right\}.$$

Proof.

$$\begin{aligned}\delta(F) &= \sup_{\mathscr{A},\mathscr{B}\subseteq\mathcal{P}_{\mathrm{f}}(X)} \left\{ \delta\left(\bigcup_{A\in\mathscr{A}} A \cup \bigcup_{B\in\mathscr{B}} B \right) - \sum_{A\in\mathscr{A}} \inf_{g\in F\setminus\{f\}} g(A) - \sum_{B\in\mathscr{B}} f(B) : |\mathscr{A}|, |\mathscr{B}| < \infty \right\} \\ &= \sup_{\mathscr{A}\subseteq\mathcal{P}_{\mathrm{f}}(X)} \left\{ f\left(\bigcup_{A\in\mathscr{A}} A \right) - \sum_{A\in\mathscr{A}} \inf_{g\in F\setminus\{f\}} g(A) : |\mathscr{A}| < \infty \right\},\end{aligned}$$

by Theorem 2.3. □

Theorem 2.7. (T_X, δ_T) *is a diversity.*

Proof.
First note that for all $F \subseteq T_X$, when $\mathscr{A} = \{\emptyset\}$,

$$\delta\left(\bigcup_{A\in\mathscr{A}} A \right) - \sum_{A\in\mathscr{A}} \inf_{f\in F} f(A) = 0$$

so that δ_T is non-negative.

If $\emptyset \neq F \subseteq G$ then for all $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ with $|\mathscr{A}| < \infty$ we have

$$\sum_{A\in\mathscr{A}} \inf_{f\in F} f(A) \geq \sum_{A\in\mathscr{A}} \inf_{f\in G} f(A).$$

Hence $\delta_T(F) \leq \delta_T(G)$, showing that δ_T is monotone.

Now suppose $F = \{f\}$. Then $\delta_T(F) \leq \sup\{\delta(A) - f(A) : A \in \mathcal{P}_{\mathrm{f}}(X)\} = 0$ by the subadditivity of f and by part 1 of Proposition 2.4. Conversely, suppose $|F| > 1$. Let $f_1, f_2 \in F$, $f_1 \neq f_2$. By monotonicity and Lemma 2.6 we have

$$\begin{aligned}\delta_T(F) &\geq \delta_T(\{f_1, f_2\}) \\ &= \sup_{\mathscr{A}\subseteq\mathcal{P}_{\mathrm{f}}(X)} \left\{ f_1\left(\bigcup_{A\in\mathscr{A}} A \right) - f_2\left(\bigcup_{A\in\mathscr{A}} A \right) : |\mathscr{A}| < \infty \right\} \\ &= \sup_{A\in\mathcal{P}_{\mathrm{f}}(X)} \{f_1(A) - f_2(A)\} \\ &> 0,\end{aligned}$$

proving that δ_T satisfies diversity property (D1).

For the triangle inequality, suppose F and G are disjoint finite subsets of T_X and that $h \in T_X \setminus (F \cup G)$. Then by Lemma 2.6

$$\delta_T(F \cup \{h\}) = \sup_{\mathscr{A}\subseteq\mathcal{P}_{\mathrm{f}}(X)} \left\{ h\left(\bigcup_{A\in\mathscr{A}} A \right) - \sum_{A\in\mathscr{A}} \inf_{f\in F} f(A) : |\mathscr{A}| < \infty \right\} \tag{2.10}$$

and

$$\delta_T(G \cup \{h\}) = \sup_{\mathscr{B}\subseteq\mathcal{P}_{\mathrm{f}}(X)} \left\{ h\left(\bigcup_{B\in\mathscr{B}} B \right) - \sum_{B\in\mathscr{B}} \inf_{g\in G} g(B) : |\mathscr{B}| < \infty \right\} \tag{2.11}$$

By part 5 of Proposition 2.4 the function h is sub-additive, so

$$h\left(\bigcup_{A\in\mathscr{A}} A\right) + h\left(\bigcup_{B\in\mathscr{B}} B\right) \geq h\left(\bigcup_{C\in\mathscr{A}\cup\mathscr{B}} C\right). \tag{2.12}$$

Combining (2.10)–(2.12) and again applying Lemma 2.6 we have

$$\begin{aligned}\delta_T(F\cup\{h\}) + \delta_T(G\cup\{h\}) &\geq \sup_{\mathscr{C}\subseteq\mathcal{P}_{\mathrm{f}}(X)} \left\{ h\left(\bigcup_{C\in\mathscr{C}} C\right) - \sum_{C\in\mathscr{C}} \inf_{f\in F\cup G} f(C) : |C| < \infty \right\} \\ &= \delta_T(F\cup G\cup\{h\}).\end{aligned}$$

The triangle inequality (D2) now follows by monotonicity.

□

Theorem 2.7 establishes that δ_T is a diversity, but leaves open the question of why this particular diversity. Here we characterise δ_T in terms of a minimality condition and prove that κ is indeed an embedding. In the proof of Theorem 3.7, this result is used to show that δ_T is the unique diversity satisfying the conditions of an injective hull.

Theorem 2.8. 1. *The map κ is an embedding from (X,δ) into (T_X,δ_T).*

2. *For all finite $Y\subseteq X$ and $f\in T_X$,*

$$\delta_T(\kappa(Y)\cup\{f\}) = f(Y).$$

3. *If $\widehat{\delta}$ is any diversity on T_X such that $\widehat{\delta}(\kappa(Y)\cup\{f\}) = f(Y)$ for all finite $Y\subseteq X$ and $f\in T_X$ then*

$$\widehat{\delta}(F) \geq \delta_T(F)$$

for all finite $F\subseteq T_X$.

Proof.
1. Fix $x\in X$. Consider finite $\mathscr{A}\subseteq\mathcal{P}_{\mathrm{f}}(X)$. The triangle inequality for diversities, (D2), gives

$$\sum_{A\in\mathscr{A}} h_x(A) = \sum_{A\in\mathscr{A}} \delta(A\cup\{x\}) \geq \delta\left(\bigcup_{A\in\mathscr{A}} A\right)$$

so that $h_x\in P_X$. There is $g\in T_X$ such that $g\preceq h_x$. Since $h_x(\{x\}) = \delta(\{x\}) = 0$ we have for all finite $A\subseteq X$ that

$$h_x(A) = \delta(A\cup\{x\}) \leq g(A)+g(\{x\}) \leq g(A)+h_x(\{x\}) = g(A) \leq h_x(A).$$

Hence $h_x = g\in T_X$.

To see that κ is one-to-one observe that for $x\neq y$, $h_x(\{x\}) = 0$ but $h_y(\{x\}) = \delta(\{x,y\}) > 0$. So $h_x\neq h_y$ for distinct $x,y\in X$.

Finally, let $Y \subseteq X$, $Y = \{y_1, \ldots, y_k\}$. Taking $\mathscr{A} = \{\{y_1\}, \ldots, \{y_k\}\}$ in (2.9) gives $\delta_T(\kappa(Y)) \geq \delta(Y)$. By repeatedly using the triangle inequality we have for any finite $\mathscr{A} = \{A_1, A_2, \ldots, A_j\} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ and $z_1, \ldots, z_j \in Y$ that

$$\begin{aligned}
\delta(Y) &\geq \delta(Y \cup A_1) - \delta(\{z_1\} \cup A_1) \\
&\geq \delta(Y \cup A_1 \cup A_2) - \delta(\{z_1\} \cup A_1) - \delta(\{z_2\} \cup A_2) \\
&\geq \delta\left(Y \cup \bigcup_{i=1}^{j} A_i\right) - \sum_{i=1}^{j} \delta(\{z_i\} \cup A_i) \\
&\geq \delta\left(\bigcup_{i=1}^{j} A_i\right) - \sum_{i=1}^{j} h_{z_i}(A_i) \\
&\geq \delta\left(\bigcup_{i=1}^{j} A_i\right) - \sum_{i=1}^{j} \inf_{h \in \kappa(Y)} h(A_i).
\end{aligned}$$

Taking the supremum over all such $\mathscr{A}$ gives $\delta_T(\kappa(Y)) \leq \delta(Y)$. So $\delta_T(\kappa(Y)) = \delta(Y)$ and κ is an embedding.

2. Let $Y \subseteq X$, Y finite, and $f \in T_X$. If $f = h_y$ for $y \in Y$ then, using part 1

$$\delta_T(\kappa(Y) \cup \{f\}) = \delta_T(\kappa(Y)) = \delta(Y) = \delta(Y \cup \{y\}) = f(Y)$$

as required. Otherwise, suppose $f \notin \kappa(Y)$. Let $Y = \{y_1, \ldots, y_k\}$.

$$\delta_T(\kappa(Y) \cup \{f\}) = \sup_{A_i, i=1,\ldots,k, A_f} \left\{ \delta\left(\bigcup_i A_i \cup A_f\right) - \sum_i \delta(\{y_i\} \cup A_i) - f(A_f) \right\}$$

Letting $A_i = \{y_i\}$ for all i shows

$$\delta_T(\kappa(Y) \cup \{f\}) \geq \sup_{A_f} \{\delta(Y \cup A_f) - f(A_f)\} = f(Y),$$

by Proposition 2.4 part 6. On the other hand, following the same reasoning as in part 1 of this proof shows

$$\delta_T(\kappa(Y) \cup \{f\}) \leq \sup_{A_f} \delta(Y \cup A_f) - f(A_f) = f(Y).$$

Therefore $\delta_T(\kappa(Y) \cup \{f\}) = f(Y)$.

3. Suppose that $F = \kappa(Y) \cup G$, where $Y \in \mathcal{P}_{\mathrm{f}}(X)$ and $G \subseteq T_X \setminus \kappa(X)$. For all collections $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ with $|\mathscr{A}| < \infty$ and all collections $\{f_A\}_{A \in \mathscr{A}}$ of elements in F, we have from 1. and 2. that

$$\begin{aligned}
\delta\left(Y \cup \bigcup_{A \in \mathscr{A}} A\right) - \sum_{A \in \mathscr{A}} f_A(A) &= \widehat{\delta}\left(\kappa(Y) \cup \bigcup_{A \in \mathscr{A}} \kappa(A)\right) - \sum_{A \in \mathscr{A}} \widehat{\delta}(\kappa(A) \cup \{f_A\}) \\
&\leq \widehat{\delta}\left(\kappa(Y) \cup \{f_A : A \in \mathscr{A}\}\right) \\
&\leq \widehat{\delta}\left(\kappa(Y) \cup F\right).
\end{aligned}$$

□

3. Hyperconvex diversities and the injective envelope

Aronszajn and Panitchpakdi [1] introduced hyperconvex metric spaces and showed that they are exactly the injective metric spaces.

Definition 3.1. 1. *A metric space (X, d) is said to be* hyperconvex *if for all $r\colon X \to \Re \cup \{\infty\}$ with $r(x) + r(y) \geq d(x, y)$ for all $x, y \in X$ there is a point $z \in X$ such that $d(z, x) \leq r(x)$ for all $x \in X$.*

2. *A metric space (X, d) is* injective *if it satisfies the following property: given any pair of metric spaces (Y_1, d_1), (Y_2, d_2), an embedding $\pi : Y_1 \to Y_2$, and a non-expansive map $\phi : Y_1 \to X$ there is a non-expansive map $\psi : Y_2 \to X$ such that $\phi = \psi \circ \pi$.*

See [2] for a proof for the equivalence of these two concepts, as well as a review of the rich metric structure of hyperconvex spaces. Here we establish diversity analogues for these concepts and show that the equivalence holds in this new setting. We begin by defining diversity analogues of injective and hyperconvex metric spaces.

Definition 3.2. 1. *A map $\phi : Y_1 \to Y_2$ is* non-expansive *if for all $A \subseteq Y_1$ we have $\delta_1(A) \geq \delta_2(\phi(A))$ and it is an* embedding *if it is one-to-one and for all $A \subseteq Y_1$ we have $\delta_1(A) = \delta_2(\pi(A))$.*

2. *A diversity (X, δ) is* injective *if it satisfies the following property: given any pair of diversities (Y_1, δ_1), (Y_2, δ_2), an embedding $\pi : Y_1 \to Y_2$, and a non-expansive map $\phi : Y_1 \to X$ there is a non-expansive map $\psi : Y_2 \to X$ such that $\phi = \psi \circ \pi$.*

3. *A diversity (X, δ) is said to be* hyperconvex *if for all $r\colon \mathcal{P}_{\mathrm{f}}(X) \to \Re \cup \{\infty\}$ such that*
$$\delta\left(\bigcup_{A \in \mathscr{A}} A\right) \leq \sum_{A \in \mathscr{A}} r(A) \tag{3.1}$$
for all finite $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ there is $z \in X$ such that $\delta(\{z\} \cup Y) \leq r(Y)$ for all finite $Y \subseteq X$.

The following theorem establishes the diversity equivalent of Aronszajn and Panitchpakdi's result.

Theorem 3.3. *A diversity (X, δ) is injective if and only if it is hyperconvex.*

Proof.
First suppose that (X, δ) is injective. Consider $r\colon \mathcal{P}_{\mathrm{f}}(X) \to \Re \cup \{\infty\}$ satisfying (3.1) for all finite $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$. Without loss of generality we can assume $r(\emptyset) = 0$ and hence $r \in P_X$. Choose $f \in T_X$ with $f \preceq r$.

Let x^* be a point not in X, let $X^* = X \cup \{x^*\}$ and let $\delta^* : \mathcal{P}_{\mathrm{f}}(X \cup \{x^*\}) \to \Re \cup \{\infty\}$ be the function where for all finite $A \subseteq X$,
$$\begin{aligned} \delta^*(A) &= \delta(A) \\ \delta^*(A \cup \{x^*\}) &= f(A). \end{aligned}$$

From Proposition 2.4 part 2 we have that δ^* is monotonic, and from part 4 and 5 we have that

$$\begin{aligned} \delta^*(A \cup C \cup \{x^*\}) &\leq \delta^*(A \cup \{x^*\}) + \delta^*(C \cup \{x^*\}) \quad (3.2) \\ \delta^*(A \cup B \cup C \cup \{x^*\}) &\leq \delta^*(A \cup B \cup \{x^*\}) + \delta^*(B \cup C). \quad (3.3) \end{aligned}$$

for all finite $A, B, C \subseteq X$ such that $B \neq \emptyset$. These, together monotonicity and the fact that δ^* coincides with δ on $\mathcal{P}_{\mathrm{f}}(X)$ imply the triangle inequality (D2) for (X^*, δ^*).

We now apply the fact that (X, δ) is injective. Let (Y_1, δ_1) be (X, δ); let (Y_2, δ_2) be (X^*, δ^*), let π be the identity embedding from (X, δ) into (X^*, δ^*) and let ϕ be the identity map from (X, δ) to itself. Then there is a non-expansive map $\phi : X^* \to X$ such that $\phi(x) = x$ for all $x \in X$.

Let $\omega = \phi(x^*)$. For all finite $A \subseteq X$ we have

$$\begin{aligned} \delta(A \cup \{\omega\}) &\leq \delta^*(A \cup \{x^*\}) \\ &= f(A) \\ &\leq r(A). \end{aligned}$$

This proves that (X, δ) is hyperconvex.

For the converse, suppose now that (X, δ) is hyperconvex. Let (Y_1, δ_1) and (Y_2, δ_2) be two diversities, let $\pi : Y_1 \to Y_2$ be an embedding and let ϕ be a non-expansive map from Y_1 to X. We will show that there is non-expansive $\psi : Y_2 \to X$ such that $\phi = \psi \circ \pi$.

Let $\mathscr{Y}$ denote the collection of pairs (Y, ψ_Y) such that $\pi(Y_1) \subseteq Y \subseteq Y_2$ and ψ_Y is a non-expansive map from Y to X such that $\phi = \psi_Y \circ \pi$. We want to show that $Y_2 \in \mathscr{Y}$. Suppose this is not the case. We write $(Y, \psi_Y) \trianglelefteq (Z, \psi_Z)$ if $Y \subseteq Z$ and ψ_Z restricted to Y equals ψ_Y. The partially ordered set $(\mathscr{Y}, \trianglelefteq)$ satisfies the conditions of Zorn's lemma, so it contains maximal elements.

Let (Y, ψ_Y) be one such maximal element. Choose $y \in Y_2 \setminus Y$. For each finite $A \subseteq Y$ let $r(A) = \delta_2(A \cup \{y\})$. For any finite collection $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(Y)$ we have

$$\begin{aligned} \delta\left(\bigcup_{A \in \mathscr{A}} \psi_Y(A)\right) &= \delta\left(\psi_Y\left(\bigcup_{A \in \mathscr{A}} A\right)\right) \\ &\leq \delta_2\left(\bigcup_{A \in \mathscr{A}} A\right) \\ &\leq \sum_{A \in \mathscr{A}} \delta_2(A \cup \{y\}) \\ &= \sum_{A \in \mathscr{A}} r(A). \end{aligned}$$

Since (X, δ) is hyperconvex, there is $x \in X$ such that

$$\delta(\psi_Y(A) \cup \{x\}) \leq r(A) = \delta_2(A \cup \{y\})$$

for all finite $A \subseteq Y$. Hence we can extend ψ_Y to $Y \cup \{y\}$ by setting $\psi_Y(y) = x$, giving a non-expansive map from $Y \cup \{y\}$ to X, and contradicting the maximality of Y.

It follows that $Y_2 \in \mathscr{Y}$, proving that (X, δ) is injective. □

Definition 3.4. *Let (X, δ) be a diversity. For $F \subseteq T_X$ and finite $Y \subseteq X$ let*

$$\Phi_F(Y) = \inf_{\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)} \left\{ \sum_{A \in \mathscr{A}} \inf_{f \in F} f(A) : |\mathscr{A}| < \infty, \bigcup_{A \in \mathscr{A}} A = Y \right\}.$$

Clearly,

$$\delta_T(F) = \sup_{Y \subseteq X} \{\delta(Y) - \Phi_F(Y) \colon |Y| < \infty\}. \tag{3.4}$$

We show that Φ_F also satisfies a sub-additivity type identity.

Lemma 3.5. *For $F, G \subseteq T_X$ and $Y, Z \subseteq \mathcal{P}_{\mathrm{f}}(X)$ we have*

$$\Phi_{F \cup G}(Y \cup Z) \leq \Phi_F(Y) + \Phi_G(Z).$$

Proof.
Given $\epsilon > 0$ there is finite $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ and a collection $\{f_A\}_{A \in \mathscr{A}}$ of elements in T_X such that

$$\Phi_F(Y) \leq \sum_{A \in \mathscr{A}} f_A(A) < \Phi_F(Y) + \epsilon/2.$$

Similarly, there is finite $\mathscr{B} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ and a collection $\{g_B\}_{B \in \mathscr{B}}$ of elements in T_X such that

$$\Phi_G(Z) \leq \sum_{B \in \mathscr{B}} g_B(B) < \Phi_G(Z) + \epsilon/2.$$

Define $\mathscr{C} = \mathscr{A} \cup \mathscr{B}$ and the collection $\{h_C\}_{C \in \mathscr{C}}$ by

$$h_C = \begin{cases} f_C & \text{if } C \in \mathscr{A}; \\ g_C & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} \Phi_F(Y) + \Phi_G(Z) + \epsilon &> \sum_{A \in \mathscr{A}} f_A(A) + \sum_{B \in \mathscr{B}} g_B(B) \\ &\geq \sum_{C \in \mathscr{C}} h_C(C) \\ &\geq \Phi_{F \cup G}(Y \cup Z). \end{aligned}$$

□

Theorem 3.6. *For any diversity (X, δ), (T_X, δ_T) is hyperconvex.*

Proof.
Let $r\colon \mathcal{P}_{\mathrm{f}}(T_X) \to \Re \cup \{\infty\}$ be given such that for all finite $\mathscr{F} \subseteq \mathcal{P}_{\mathrm{f}}(T_X)$

$$\sum_{F \in \mathscr{F}} r(F) \geq \delta_T \left(\bigcup_{F \in \mathscr{F}} F \right).$$

Without loss of generality we can assume $r(\emptyset) = 0$. We need to find $g \in T_X$ so that $\delta_T(G \cup \{g\}) \leq r(G)$ for all $G \subseteq T_X$.

Define ω on $\mathcal{P}_{\mathrm{f}}(X)$ by

$$\omega(A) = \inf_{F \subseteq T_X} \{r(F) + \Phi_F(A) \colon |F| < \infty\}.$$

We have $\omega(\emptyset) = 0$. Suppose that $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$, $|\mathscr{A}| < \infty$ and let $\{F_A : A \in \mathscr{A}\}$ be a collection of finite subsets of T_X indexed by elements of $\mathscr{A}$. From Lemma 3.5 we have

$$\begin{aligned} \delta \left(\bigcup_{A \in \mathscr{A}} A \right) &\leq \delta_T \left(\bigcup_{A \in \mathscr{A}} F_A \right) + \Phi_{\left(\bigcup_{A \in \mathscr{A}} F_A\right)} \left(\bigcup_{A \in \mathscr{A}} A \right) \\ &\leq \sum_{A \in \mathscr{A}} (r(F_A) + \Phi_{F_A}(A)) \end{aligned}$$

so that

$$\delta \left(\bigcup_{A \in \mathscr{A}} A \right) \leq \sum_{A \in \mathscr{A}} \omega(A)$$

and $\omega \in P_X$.

There is $g \in T_X$ such that $g \preceq \omega$. Consider finite $F \subseteq T_X$. Applying Lemma 2.6,

$$\begin{aligned} \delta_T(F \cup \{g\}) &= \sup_{A \in \mathcal{P}_{\mathrm{f}}(X)} \{g(A) - \Phi_F(A)\} \\ &\leq \sup_{A \in \mathcal{P}_{\mathrm{f}}(X)} \{(r(F) + \Phi_F(A)) - \Phi_F(A)\} \\ &= r(F), \end{aligned}$$

as required. □

Isbell [3] and Dress [4] not only proved that the tight span of a metric space is hyperconvex (injective), they showed that the tight span is essentially the minimal hyperconvex metric space into which the metric can be embedded. We prove that the same holds for diversities.

Theorem 3.7. *If there is an embedding from (X, δ) into (Y, δ_Y) and (Y, δ_Y) is injective (hyperconvex) then there is an embedding from (T_X, δ_T) into (Y, δ_Y).*

Proof.
Let π be the embedding of (X, δ) into (Y, δ_Y). Then π is a non-expansive map.

Since (Y, δ_Y) is injective, and κ is an embedding of (X, δ) into (T_X, δ_T), there is a non-expansive map $\phi : T_X \to Y$ such that $\pi = \phi \circ \kappa$. Using Theorem 2.8 part 3 we will show that $\delta_T(F) \leq \delta_Y(\phi(F))$ so that ϕ is an embedding.

Consider $f \in T_X$. Define g on $\mathcal{P}_{\mathrm{f}}(X)$ by $g(A) = \delta_Y(\pi(A) \cup \phi(\{f\}))$ for all finite A. Then for any finite $A \subseteq X$ we have

$$g(A) = \delta_Y(\pi(A) \cup \phi(\{f\})) = \delta_Y(\phi(\kappa(A) \cup \{f\})) \leq \delta_T(\kappa(A) \cup \{f\}) = f(A)$$

for all A. For all finite collections $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ we have

$$\begin{aligned} \sum_{A \in \mathscr{A}} g(A) &= \sum_{A \in \mathscr{A}} \delta_Y(\pi(A) \cup \phi(\{f\})) \\ &\geq \delta_Y\left(\bigcup_{A \in \mathscr{A}} \pi(A)\right) = \delta\left(\bigcup_{A \in \mathscr{A}} A\right) \end{aligned}$$

so that $g \in P_X$ and $g \preceq f$. Hence $g(A) = f(A)$ for all finite $A \subseteq X$. It follows that

$$\delta_Y(\pi(A) \cup \phi(\{f\})) = \delta_T(\kappa(A) \cup \{f\})$$

for all $f \in T_X$ and finite $A \subseteq X$.

Define $\widehat{\delta}$ on T_X by $\widehat{\delta}(F) = \delta_Y(\phi(F))$. Then $\widehat{\delta}$ is a diversity and $\widehat{\delta}(\kappa(Y) \cup \{f\}) = f(Y)$ for all finite $Y \subseteq X$. By Theorem 2.8, $\widehat{\delta}(F) \geq \delta_T(F)$ for all finite F.

Hence $\delta_T(F) = \delta_Y(\phi(F))$ for all finite F and ϕ is an embedding. □

Corollary 3.8. *Let (X, δ) be a diversity. The following are equivalent:*

1. *(X, δ) is hyperconvex;*
2. *(X, δ) is injective;*
3. *There is an isomorphism between (X, δ) and its tight span, (T_X, δ_T).*

Note that the class of all diversities with all non-expansive maps as morphisms forms a category, which we will denote **Dvy**. The definitions of embeddings and injective objects then correspond to the category theory concepts, as reviewed in [32]. Theorem 3.7 establishes that (T_X, δ_T) is the *injective envelope* of (X, δ) in the category **Dvy**.

4. Tight span of the diameter diversity

In this section we prove that tight span theory for metrics is embedded within the tight span theory for diversities. The link between the two is provided by the *diameter diversity* as introduced above.

Definition 4.1. *Given a metric space (X, d) we define the diversity $\delta = \mathrm{diam}_d$ by*

$$\delta(A) = \mathrm{diam}_d(A) = \max\{d(a, a') : a, a' \in A\}$$

for finite $A \subseteq X$, with $\mathrm{diam}_d(\emptyset) = 0$. *We call* diam_d *the* diameter diversity *for d. Note that if we restrict* diam_d *to pairs of elements we recover d as the induced metric.*

In this section we will establish close links between tight spans of metrics and tight spans of their diameter diversities.

$$\begin{array}{ccc} (X,d) & \xrightarrow{\text{tight span}} & (T_X^d, d_T) \\ {\scriptstyle \delta=\mathrm{diam}_d}\big\downarrow & & \big\downarrow{\scriptstyle \delta_T=\mathrm{diam}_{d_T}} \\ (X,\delta) & \xrightarrow{\text{tight span}} & (T_X^\delta, \delta_T) \end{array}$$

Lemma 4.2.

1. *Let (Y,δ) be a diversity with induced metric (Y,d_δ). Let (X,d) be a metric space and let (X,diam_d) be the associated diameter diversity. Then ϕ is a non-expansive map from (Y,δ) to (X,diam_d) if and only if it is a non-expansive map from (Y,d_δ) to (X,d).*
2. *A metric space (X,d) is injective (hyperconvex) if and only if the diameter diversity (X,diam_d) is injective (hyperconvex).*
3. *The tight span (T_X^δ,δ_T) of a diameter diversity is itself a diameter diversity.*

Proof.
1. Suppose that ϕ is a non-expansive map from (Y,δ) to (X,diam_d). For all $y_1,y_2 \in Y$ we have

$$d_\delta(y_1,y_2) = \delta(\{y_1,y_2\}) \geq \mathrm{diam}_d(\{\phi(y_1),\phi(y_2)\}) = d(\phi(y_1),\phi(y_2)),$$

so ϕ is non-expansive from (Y,d_δ) to (X,d). Conversely, suppose ϕ is a non-expansive map from (Y,d_δ) to (X,d). Then for any finite $A \subseteq Y$ we have

$$\begin{array}{rcl} \delta(A) & \geq & \sup\{d_\delta(a_1,a_2) : a_1,a_2 \in A\} \\ & \geq & \sup\{d(\phi(a_1),\phi(a_2)) : a_1,a_2 \in A\} \\ & = & \mathrm{diam}_d(\phi(A)). \end{array}$$

2. Suppose that (X,d) is injective. Let (Y_1,δ_1), (Y_2,δ_2) be two diversities with induced metrics d_1,d_2. Let π be an embedding from (Y_1,δ_1) into (Y_2,δ_2) and let ϕ be a non-expansive map from (Y_1,δ_1) to (X,diam_d). Then π embeds (Y_1,d_1) into (Y_2,d_2), and by part 1., ϕ is a non-expansive map from (Y_1,d_1) to (X,d). Hence there is a non-expansive map ψ from (Y_2,d_2) to (X,d) such that $\phi = \psi \circ \pi$, which by part. 1 is a non-expansive map from $(Y_2,\mathrm{diam}_{d_2})$ to (X,diam_d). Since $\delta_2(A) \geq \mathrm{diam}_{d_2}(A)$ for all A, ψ is non-expansive from (Y_2,δ_2) to (X,diam_d). Hence (X,diam_d) is injective.

Conversely, suppose (X,diam_d) is an injective diversity. Let (Y_1,d_1), (Y_2,d_2) be two metric spaces, let π be an embedding of (Y_1,d_1) into (Y_2,d_2), and let ϕ be a non-expansive map from (Y_1,d_1) to (X,d). Then ϕ is a non-expansive map from $(Y_1,\mathrm{diam}_{d_1})$ to (X,diam_d) and since (X,diam_d) is injective, there is a non-expansive map ψ from $(Y_2,\mathrm{diam}_{d_2})$ to (X,diam_d) such that $\phi = \psi \circ \pi$. Applying part 1. again, we have that ψ is the required non-expansive map from (Y_2,d_2) to (X,d). Hence (X,d) is injective.
3. Since δ is a diameter diversity, for any collection $\{A_f\} \subseteq \mathcal{P}_{\mathrm{f}}(X)$, $f \in F$ with

F finite

$$\delta\left(\bigcup_{f\in F} A_f\right) = \delta(A_{f_1}\cup A_{f_2})$$

for some $f_1, f_2 \in F$. Hence for finite $F \subseteq T_X^\delta$

$$\begin{aligned}
\delta_T(F) &= \sup_{A_f}\left\{\delta\left(\bigcup_{f\in F} A_f\right) - \sum_{f\in F} f(A_f)\right\} \\
&= \max_{f_1,f_2\in F}\ \sup_{A_1,A_2\in\mathcal{P}_f(X)} \{\delta(A_1\cup A_2) - f_1(A_1) - f_2(A_2)\} \\
&= \max_{f_1,f_2\in F} \delta_T(\{f_1,f_2\}).
\end{aligned}$$

□

Theorem 4.3. *Let (X,d) be a metric space with metric tight span (X_T^d, d_T). Let (X,δ) be the associated diameter diversity where $\delta = \mathrm{diam}_d$, and let (T_X,δ_T) be its diversity tight span. Then*

1. *The metric space obtained by restricting δ_T to pairs in T_X is isomorphic to the metric space (T_X^d, d_T).*
2. *The diversity obtained by taking the diameter on the metric space (T_X^d, d_T) is isomorphic to the diversity (T_X,δ_T).*

Proof.
Let d_δ be the induced metric for δ_T. The map κ defined on (X,d) embeds (X,d) in $(\kappa(X), d_\delta)$. By Lemma 4.2 parts 2 and 3, (T_X^δ, d_δ) is injective. The metric tight span is the injective hull, so there is an embedding from (T_X^d, d_T) to (T_X^δ, d_δ).

For the other direction, since (T_X^d, d_T) is injective, so is its associated diversity metric $(T_X^d, \mathrm{diam}_{d_T})$. The map κ defined in (X,δ) embeds (X,δ) into $(T_X^d, \mathrm{diam}_{d_T})$ and since (T_X^δ,δ_T) is the injective hull of (X,δ) there is an embedding from (T_X^δ,δ_T) into $(T_X^d, \mathrm{diam}_{d_T})$. Restricting this embedding to pairs gives an embedding from (T_X^δ, d_δ) into (T_X^d, d_T). This proves 1. and the isomorphism in part 2. follows directly. □

5. Phylogenetic Diversity

A metric space (X,d) is *additive* or *tree-like* if there is a tree with nodes partially labelled by X so that for each x,y the length of the path (including branch-lengths) connecting x and y equals $d(x,y)$. Dress [4] showed that if (X,d) is additive then its tight span corresponds exactly to the smallest tree it can be embedded in. The elements of the tight span correspond not only to the nodes of the original tree, but also the points along the edges. Hence, when

X is finite, the continuous structure arises automatically out of a finite metric. Here we will prove analogous results about phylogenetic diversity.

Following [4] we will work with *real trees* (also called metric-trees or $\Re$-trees), rather than graph-theoretic trees.

Definition 5.1. *See [33, 34]*

1. *Let $(\mathcal{X}, d)$ be a metric space and let x, y be two points at distance $d(x, y) = r$. A* geodesic *joining x, y is a map $c : [0, r] \to \mathcal{X}$ such that $c(0) = x$, $c(r) = y$ and $d(c(s), c(t)) = |t - s|$ for all $s, t \in [0, r]$. The image of c is called a* geodesic segment.
2. *A complete metric space $(\mathcal{X}, d)$ is a* real tree *or $\Re$-tree if*
   (a) *there is a unique geodesic segment $[x, y]$ joining each pair of points $x, y \in \mathcal{X}$.*
   (b) *if $[y, x] \cap [x, z] = \{x\}$ then $[y, x] \cup [x, z] = [y, z]$.*

   Hence if x, y, z are three points in a real tree then

$$[x, y] \subseteq [x, z] \cup [y, z], \tag{5.1}$$

   see [34, Ch. 2, Corr. 1.3].

Phylogenetic diversity, as introduced by [23] and investigated extensively by [24, 25] and others, can be viewed as a generalisation of additive metrics. The phylogenetic diversity of a set of nodes or points in a tree is the length of the smallest subtree connecting them, so that the restriction of a phylogenetic diversity to pairs of points gives an additive metric. A formal definition of phylogenetic diversity on real trees requires a bit more machinery.

For a complete real tree $(\mathcal{X}, d)$, let μ be the one-dimensional Hausdorff measure on it [35]. The important features of μ for our purposes is that it is defined on all Borel sets, it is monotone, and it is additive on disjoint sets. Furthermore, for any points $a, b \in \mathcal{X}$, $\mu([a, b]) = d(a, b)$, and naturally $\mu(\{a\}) = 0$. See [36] for a related measure on real trees.

Definition 5.2.

1. *The convex hull of a set $A \subseteq \mathcal{X}$ is*

$$\mathrm{conv}(A) = \bigcup_{a,b \in A} [a, b]$$

   and we say that A is convex if $A = \mathrm{conv}(A)$.
2. *Let $(\mathcal{X}, d)$ be a complete real tree. The* tree-diversity *$(\mathcal{X}, \delta_t)$ for $(\mathcal{X}, d)$ is defined by*

$$\delta_t(A) := \mu(\mathrm{conv}(A))$$

   for all finite $A \subseteq \mathcal{X}$.

Note that since A is finite, $\mathrm{conv}(A)$ is closed and hence $\mu(\mathrm{conv}(A))$ is defined.

First we prove that this phylogenetic diversity satisfies the diversity axioms (D1) and (D2).

Theorem 5.3. *Let $(\mathcal{X}, d)$ be a complete real tree. Then $(\mathcal{X}, \delta_t)$ is a diversity.*

Proof.
Since μ is a measure, δ_t is non-negative and also monotonic. If $|A| \leq 1$ then $\mathrm{conv}(A) = A$ and so $\delta_t(A) = \mu(A) = 0$. If $|A| > 1$ then select distinct $a, b \in A$. Since $\mathrm{conv}([a,b]) = [a,b]$ and $\mu([a,b]) = d(a,b)$ we have $\delta_t(A) \geq \delta_t(\{a,b\}) = d(a,b) > 0$. This proves (D1).

Let $A, B, C \in \mathcal{P}_{\mathrm{f}}(\mathcal{X})$ and suppose that $B \neq \emptyset$. From (5.1) we have

$$[a,c] \subseteq [a,b] \cup [b,c] \tag{5.2}$$

for all $a \in A$, $b \in B$ and $c \in C$. Hence

$$\mathrm{conv}(A \cup C) \subseteq \mathrm{conv}(A \cup B) \cup \mathrm{conv}(B \cup C)$$

and

$$\begin{aligned} \delta_t(A \cup C) &= \mu(\mathrm{conv}(A \cup C)) \\ &\leq \mu(\mathrm{conv}(A \cup B)) + \mu(\mathrm{conv}(B \cup C)) \\ &= \delta_t(A \cup B) + \delta_t(B \cup C), \end{aligned}$$

giving us the triangle equality (D2).

□

The remainder of this section characterises the tight span of a phylogenetic diversity in terms of real trees.

Definition 5.4.
1. *A diversity (X, δ) is a* phylogenetic diversity *or* tree-like *if it can be embedded in the tree diversity $(\mathcal{X}, \delta_t)$ for some complete real tree $(\mathcal{X}, d)$.*
2. *The* convex closure *of A is the topological closure of the convex hull of A, denoted* $\mathrm{cl}(A)$.

Consider a phylogenetic diversity (X, δ) with induced metric (X, d). Let (T_X^δ, δ_T) denote the tight span of (X, δ) and let (T_X^d, d_T) denote the metric tight span of (X, d), which is itself a complete real tree. We denote by $(\bar{X}, \bar{\delta})$ the tree diversity on (T_X^d, d_T). Our main result in this section, Theorem 5.8, shows that (T_X^δ, δ_T) is isomorphic to $(\bar{X}, \bar{\delta})$.

Our first lemma shows that (X, δ) can be embedded in $(\bar{X}, \bar{\delta})$.

Lemma 5.5. *Any phylogenetic diversity (X, δ) with induced metric (X, d) can be embedded in $(\bar{X}, \bar{\delta})$, the tree diversity of the metric tight span of (X, d). Moreover, $\bar{X}$ is the convex closure of the image of X under the embedding.*

Proof.
First note that if a real tree $(\mathcal{X}, d_X)$ is embedded into a real tree $(\mathcal{Y}, d_Y)$, then the corresponding tree diversity $(\mathcal{X}, \delta_X)$ is embedded into the corresponding tree diversity $(\mathcal{Y}, \delta_Y)$. This fact follows from the definition of tree diversities and the properties of μ, the Hausdorff outer measure.

Suppose (X, δ) is embedded in the tree diversity $(\mathcal{X}, \delta_t)$, where the underlying real tree $(\mathcal{X}, d_t)$ is complete and hence hyperconvex. Then (X, d) is embedded in $(\mathcal{X}, d_t)$. By Theorem 3.7, (T_X^d, d_T) can be embedded in $(\mathcal{X}, d_t)$. By the comments above, $(\bar{X}, \bar{\delta})$ is embeddable in $(\mathcal{X}, \delta_t)$. This is sufficient to imply that (X, δ) is embedded in $(\bar{X}, \bar{\delta})$.

The last statement follows from the convex closure of the image of X being the minimal complete real tree containing the image of X, and $(\bar{X}, \bar{d})$ being isomorphic to (T_X^d, d_T). □

In what follows, for a phylogenetic diversity (X, δ), we call a tree diversity $(\mathcal{X}, \delta_t)$ a *minimal embedding* if $X \subseteq \mathcal{X}$, $\mathrm{cl}(X) = \mathcal{X}$, and $\delta(A) = \delta_t(A)$ for all $A \subseteq X$. The previous theorem shows that a minimal embedding exists for all phylogenetic trees, and all minimal embeddings of (X, δ) are isomorphic to $(\bar{X}, \bar{\delta})$, the tree diversity of (T_X^d, d_T).

Lemma 5.6. *Let (X, δ) be a phylogenetic diversity with tight span (T_X, δ_T). Let $(\mathcal{X}, \delta_t)$ be a minimal embedding of (X, δ). For all $f \in T_X$ there is $v \in \mathcal{X}$ such that $f(A) = \delta_t(A \cup \{v\})$ for all finite $A \subseteq X$.*

Proof.
First note that for any finite but non-empty $A \subseteq X$ the function

$$\phi : \mathcal{X} \to \Re : x \mapsto \delta_t(A \cup \{x\})$$

is continuous. Hence if $r \geq \delta_t(A)$ the ball

$$B(A, r) := \phi^{-1}(A) = \{x \in \mathcal{X} : \delta_t(A \cup \{x\}) \leq r\}$$

is closed. Furthermore, for $x_1, x_2 \in B(A, r)$ and $a \in A$ then $[a, x_1] \subseteq \mathrm{cl}(A \cup \{x_1\}) \subseteq B(A, r)$ and $[a, x_2] \subseteq \mathrm{cl}(A \cup \{x_2\}) \subseteq B(A, r)$. Again by (5.1) we have

$$[x_1, x_2] \subseteq [a, x_1] \cup [a, x_2] \subseteq B(A, r)$$

so that $B(A, r)$ is both closed and convex.

Let Γ be the collection

$$\Gamma = \{B(A, f(A)) : A \in \mathcal{P}_{\mathrm{f}}(X),\ A \neq \emptyset,\ f(A) < \infty\}.$$

We will show that Γ has non-empty intersection.

Firstly consider a pair $A_i, A_j \in \Gamma$. We show that there is v such that $\delta_t(A_i \cup \{v\}) \leq f(A_i)$ and $\delta_t(A_j \cup \{v\}) \leq f(A_j)$. This clearly holds if there is $v \in \mathrm{cl}(A_i) \cap \mathrm{cl}(A_j)$. Suppose then that $\mathrm{cl}(A_i)$ and $\mathrm{cl}(A_j)$ are disjoint. By [34, Ch. 2, Lemma 1.9] there exists $a_i \in \mathrm{cl}(A_i)$ and $a_j \in \mathrm{cl}(A_j)$ such that $[a_i, a_j] \cap \mathrm{cl}(A_i) = \{a_i\}$ and $[a_i, a_j] \cap \mathrm{cl}(A_j) = \{a_j\}$ and for all $x \in A_i$ and $y \in A_j$ we have $[a_i, a_j] \subseteq [x, y]$. Then

$$\begin{aligned} f(A_i) + f(A_j) &\geq \delta_t(A_i \cup A_j) \\ &\geq \mu(\mathrm{cl}(A_i)) + \mu([a_i, a_j]) + \mu(\mathrm{cl}(A_j)) \\ &= \delta(A_i) + d(a_i, a_j) + \delta(A_j). \end{aligned}$$

Hence there is $v \in [a_i, a_j]$ such that $d(a_i, v) \leq f(A_i) - \delta(A_i)$ and $d(a_j, v) \leq f(A_j) - \delta(A_j)$, so that

$$\begin{aligned}\delta_t(A_i \cup \{v\}) &= \delta_t(A_i \cup \{a_i\}) + \delta_t(\{a_i, v\}) \\ &= \delta(A_i) + d(a_i, v) \\ &\leq f(A_i),\end{aligned}$$

and likewise $\delta_t(A_j \cup \{v\}) \leq f(A_j)$.

We have established that Γ satisfies the pairwise intersection property. The closed, convex sets of a real tree satisfy the Helly property [33], so every finite subcollection of Γ has non-empty intersection. Let Γ^* be the collection of all intersections of finite subcollections of Γ. Then Γ^* is closed under pairwise intersections and so for all $A, B \in \Gamma^*$ there is a $C \in \Gamma^*$ such that $C \subseteq A \cap B$. Hence Γ^* satisfies the properties of Proposition 3.1 in [33] and has non-empty intersection.

Let v be contained in the intersection of Γ and let $f_v(A) = \delta_t(A \cup \{v\})$ for all finite $A \subseteq X$. Then $f_v \preceq f$. For any finite collection $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$ we have

$$\sum_{A \in \mathscr{A}} f_v(A) = \sum_{A \in \mathscr{A}} \delta_t(A \cup \{v\}) \geq \delta\left(\sum_{A \in \mathscr{A}} A\right)$$

so $f_v \in P_X$ and by the minimality of f, $f_v = f$. □

Lemma 5.7. *Let (X, δ) be a phylogenetic diversity with tight span (T_X, δ_T). Let $(\mathcal{X}, \delta_t)$ be a minimal embedding of (X, δ). For all $v \in \mathcal{X}$ if we define $f(A) = \delta_t(A \cup \{v\})$ for all finite $A \subseteq X$, then $f \in T_X$.*

Proof.
For all finite $\mathscr{A} \subseteq \mathcal{P}_{\mathrm{f}}(X)$

$$\sum_{A \in \mathscr{A}} f(A) = \sum_{A \in \mathscr{A}} \delta(\{v\} \cup A) \geq \bigcup_{A \in \mathscr{A}} \delta(A),$$

and so $f \in P_X$. Choose $g \in T_X$ with $g \preceq f$. By Lemma 5.6, there is a $w \in \mathcal{X}$ such that $g(A) = \delta(\{w\} \cup A)$ for all finite $A \subseteq X$. Now

$$\delta(\{v, w\}) = g(\{v\}) \leq f(\{v\}) = \delta(\{v\}) = 0,$$

so $v = w$. □

Theorem 5.8. *Let (X, δ) be a phylogenetic diversity with induced metric (X, d). Let (T^δ_X, δ_T) be the tight span of (X, δ) and let (T^d_X, d_T) be the metric tight span of (X, d). Then (T^δ_X, δ_T) is isomorphic to the tree diversity of (T^d_X, d_T).*

Proof.
Let $(\mathcal{X}, \delta_t)$ be a minimal embedding of (X, δ). We will show that (T^δ_X, δ_T) is isomorphic to $(\mathcal{X}, \delta_t)$.

Let γ be the map taking $v \in \mathcal{X}$ to the function f_v, where $f_v(A) = \delta_t(A \cup \{v\})$ for all finite $A \subseteq X$. Lemma 5.6 shows that γ maps $\mathcal{X}$ to T_X. Lemma 5.7 shows that γ is surjective—all of T_X is in the image of γ. It remains to show γ preserves the diversity. This will then imply that γ is a one-to-one correspondence.

We need to show that for finite $V \subseteq \mathcal{X}$, $\delta_T(\gamma(V)) = \delta_t(V)$. We have

$$\begin{aligned}
\delta_T(\gamma(V)) &= \sup_{A_v \in \mathcal{P}_{\mathrm{f}}(X), v \in V} \left\{ \delta\left(\bigcup_{v \in V} A_v \right) - \sum_{v \in V} f_v(A_v) \right\} \\
&= \sup_{A_v \in \mathcal{P}_{\mathrm{f}}(X), v \in V} \left\{ \delta_t\left(\bigcup_{v \in V} A_v \right) - \sum_{v \in V} \delta_t(A_v \cup \{v\}) \right\} \\
&\leq \delta_t(V).
\end{aligned}$$

Note that equality can be attained by letting $A_v = \{v\}$ for all v, and thus δ_T and δ_t are identical. □

6. Tight span and the Steiner tree problem

Let X be a finite set of points in a metric space (M, d). The *(Metric) Steiner Tree Problem* is to find the shortest network that connects them. Clearly this network will always be a tree. More formally

METRIC STEINER PROBLEM.
Input: Subset X of a metric space (M, d).
Problem: Find a (graph theoretic) tree T for which $X \subseteq V(T) \subseteq M$ and

$$\sum_{\{u,v\} \in E(T)} d(u, v)$$

is minimised.

We let $L_d(X)$ denote the length of a Steiner tree for X.

Dress and Krüger [31] examined an 'abstract' metric Steiner problem where one effectively drops the constraint that $V(T) \subseteq M$. This abstract Steiner tree was actually one of the first distance based criteria proposed for the inference of phylogenetic trees [37, 38], though it is now not widely used. Suppose that T is a tree with edge weights $w : E(T) \to \Re_{\geq 0}$. Given $u, v \in V(T)$ we let $d_w(u, v)$ denote the sum of edge weights along the path from u to v.

ABSTRACT STEINER PROBLEM.
Input: Finite metric space (X, d).
Problem: Find a (graph theoretic) tree T and edge weighting $w : E(T) \to \Re$ such that $X \subseteq V(T)$, $d_w(x, y) \geq d(x, y)$ for all $x, y \in X$ and

$$\sum_{e \in E(T)} w(e)$$

is minimised.

Suppose that T is a solution to the Metric Steiner Problem for $X \subseteq M$. Define the weight function $w : E(T) \rightarrow \Re$ by $w(\{u,v\}) = d(u,v)$. Then, by the triangle inequality, $d_w(x,y) \geq d(x,y)$ for all $x,y \in X$. It follows then that the length of the minimum abstract Steiner tree for $(X, d|_X)$ is a lower bound for the metric Steiner problem. Dress and Krüger showed that the lower bound becomes tight when (M,d) equals the tight span of X.

Theorem 6.1. *[31] Let (X,d) be a finite metric space. For every solution (T,w) to the abstract Steiner tree problem there is a map $\phi : V(T) \rightarrow T_X$ such that $\phi(x) = \kappa(x)$ for all $x \in X$ and $w(\{u,v\}) = d_T(\phi(u), \phi(v))$ for all $\{u,v\} \in E(T)$.*

Hence the length of the minimal Steiner tree for $\kappa(X)$ in (T_X, d_T) equals the length of the minimal abstract Steiner tree for (X,d) and the minimal abstract Steiner trees can be embedded within the tight span. A direct corollary is that if d is tree-like then the abstract Steiner tree equals the tree corresponding to d.

Here we show that, using diversities, we can obtain a tighter bound on the metric Steiner problem than that given by the abstract Steiner problem. Given a tree T with edge weights w and $A \subseteq V(T)$ we let $\delta_w(A)$ be the sum of edge weights in the smallest subtree of T connecting A. Hence $(X, \delta_w|_X)$ is a phylogenetic diversity.

Diversity Steiner Problem.
Input: Finite diversity (X, δ).
Problem: Find a (graph theoretic) tree T and edge weighting $w : E(T) \rightarrow \Re$ such that $X \subseteq V(T)$, $\delta_w(Y) \geq \delta(Y)$ for all $Y \subseteq X$, and

$$\sum_{e \in E(T)} w(e)$$

is minimised.

Proposition 6.2. *Given $k \geq 2$ define the diversity $\delta^{(k)}$ by*

$$\delta^{(k)}(A) = \max\{L_d(B) : |B| \leq k,\ B \subseteq A\}.$$

If T is a solution to the Diversity Steiner problem for $\delta^{(k)}$ then the length of T is a lower bound for the metric Steiner problem.

Proof.
Let T' be a solution to the metric Steiner problem and let δ_w be the associated phylogenetic diversity. Then for all B such that $|B| \leq k$ we have that $\delta_w(B)$, the length of T' restricted to B, is bounded below by $L_d(B) = \delta^{(k)}(B)$. It follows that $\delta^{(k)}(A) \leq \delta_w(A)$ for all $A \subseteq X$, so that T' is a potential solution for the diversity Steiner problem. □

As k increases, the bounds returned by the diversity Steiner tree applied to $\delta^{(k)}$ will tighten, until eventually the diversity Steiner tree will coincide with the metric Steiner tree. Furthermore, we have a direct extension of Theorem 6.1, stating that these diversity Steiner trees will all be contained in the diversity state span.

Theorem 6.3. *Let (X, δ) be a finite diversity with induced metric space (X, d). For every solution (T, w) to the diversity Steiner tree problem for (X, δ) there is a map $\phi : V(T) \to T_X$ such that $\phi(x) = \kappa(x)$ for all $x \in X$ and $w(\{u, v\}) = \delta_T(\{\phi(u), \phi(v)\})$ for all $\{u, v\} \in E(T)$.*

Proof.
Let δ_w be the diversity on $V(T)$ given by (T, w), as defined above. Since (T, w) solves the diversity Steiner problem, $\delta_w(A) \geq \delta(A)$ for all $A \subseteq X$. Let κ denote the canonical embedding from X to T_X. Then κ is a non-expansive map from $(X, \delta_w|_X)$ to (T_X, δ_T).

The tight span (T_X, δ_T) is injective. Hence there is a non-expansive map ϕ from $(V(T), \delta_w)$ to (T_X, δ_T) such that $\phi(x) = \kappa(x)$ for all $x \in X$. For each u, v let $w'(\{u, v\}) = \delta_T(\{\phi(u), \phi(v)\})$. Then

$$w(\{u, v\}) = \delta_w(\{u, v\}) \geq \delta_T(\{\phi(u), \phi(v)\}) = w'(\{u, v\})$$

for all $u, v \in V$.

Consider $A \subseteq X$, and let E_A be the set of edges in the smallest subtree of T containing A. By the triangle inequality,

$$\delta_{w'}(A) = \sum_{e \in E_A} w'(e) \geq \delta(X).$$

Hence (T, w') is a candidate for the diversity Steiner problem, but since (T, w) is already minimum, $\sum_{e \in E(T)} w(e) \leq \sum_{e \in E(T)} w'(e)$. It follows that $w(e) = w'(e)$ for all $e \in E(T)$. □

References

[1] N. Aronszajn, P. Panitchpakdi, Extension of uniformly continuous transformations and hyperconvex metric spaces, Pacific J. Math. 6 (1956) 405–439.

[2] R. Espínola, M. A. Khamsi, Introduction to hyperconvex spaces, in: Handbook of metric fixed point theory, Kluwer Acad. Publ., Dordrecht, 2001, pp. 391–435.

[3] J. R. Isbell, Six theorems about injective metric spaces, Comment. Math. Helv. 39 (1964) 65–76.

[4] 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, Adv. in Math. 53 (3) (1984) 321–402.

[5] H.-J. Bandelt, A. W. M. Dress, A canonical decomposition theory for metrics on a finite set, Adv. Math. 92 (1) (1992) 47–105.

[6] D. Bryant, V. Moulton, NeighborNet: An agglomerative algorithm for the construction of planar phylogenetic networks, Molecular Biology And Evolution 21 (2004) 255–265.

[7] D. Huson, D. Bryant, Application of phylogenetic networks in evolutionary studies, Mol. Biol. Evol. 23 (2006) 254–267.

[8] L. Pachter, D. Speyer, Reconstructing trees from subtree weights, Appl. Math. Lett. 17 (6) (2004) 615–621.

[9] A. Dress, V. Moulton, W. Terhalle, T-theory: an overview, European J. Combin. 17 (2-3) (1996) 161–175, discrete metric spaces (Bielefeld, 1994).

[10] Sturmfels, Can biology lead to new theorems?, Annual report of the Clay mathematics institute.

[11] A. Dress, K. T. Huber, A. Lesser, V. Moulton, Hereditarily optimal realizations of consistent metrics, Ann. Comb. 10 (1) (2006) 63–76.

[12] K. T. Huber, J. Koolen, V. Moulton, A. Spillner, Characterizing cell-decomposable metrics, Electron. J. Combin. 15 (1) (2008) Note 7, 9.

[13] P. Buneman, A note on the metric properties of trees, J. Combinatorial Theory Ser. B 17 (1974) 48–50.

[14] A. Dress, K. Huber, V. Moulton, Some variations on a theme by Buneman, Ann. Comb. 1 (4) (1997) 339–352.

[15] A. Dress, M. Hendy, K. Huber, V. Moulton, On the number of vertices and edges of the Buneman graph, Ann. Comb. 1 (4) (1997) 329–337.

[16] K. T. Huber, J. H. Koolen, V. Moulton, On the structure of the tight-span of a totally split-decomposable metric, European J. Combin. 27 (3) (2006) 461–479.

[17] M. Develin, B. Sturmfels, Tropical convexity, Doc. Math. 9 (2004) 1–27 (electronic).

[18] P. Huggins, B. Sturmfels, J. Yu, D. S. Yuster, The hyperdeterminant and triangulations of the 4-cube, Math. Comp. 77 (263) (2008) 1653–1679.

[19] B. Sturmfels, J. Yu, Classification of six-point metrics, Electron. J. Combin. 11 (1) (2004) Research Paper 44, 16 pp. (electronic).

[20] K. T. Huber, J. H. Koolen, V. Moulton, The tight span of an antipodal metric space. I. Combinatorial properties, Discrete Math. 303 (1-3) (2005) 65–79.

[21] D. Eppstein, Manhattan orbifolds, Topology Appl. 157 (2) (2010) 494–507.

[22] H. Hirai, A geometric study of the split decomposition, Discrete Comput. Geom. 36 (2) (2006) 331–361.

[23] D. Faith, Conservation evaluation and phylogenetic diversity, Biological Conservation 61 (1992) 1–10.

[24] M. A. Steel, Phylogenetic diversity and the greedy algorithm, Systematic Biology 54 (527–529).

[25] B. Minh, S. Klaere, A. von Haeseler, Taxon selection under split diversity, Systematic Biology 58 (6) (2009) 586–594.

[26] S. Gähler, Untersuchungen über verallgemeinerte m-metrische Räume, I, II, III, Math. Nachr. 40 (1969), 165–189; ibid. 40 (1969), 229–264; ibid. 41 (1969) 23–36.

[27] K. Menger, Untersuchungen über allgemeine Metrik, Math. Ann. 100 (1) (1928) 75–163.

[28] M.-M. Deza, I. G. Rosenberg, Small cones of m-hemimetrics, Discrete Math. 291 (1-3) (2005) 81–97.

[29] V. Chepoi, B. Fichet, A note on three-way dissimilarities and their relationship with two-way dissimilarities, in: Selected contributions in data analysis and classification, Stud. Classification Data Anal. Knowledge Organ., Springer, Berlin, 2007, pp. 465–475.

[30] M. J. Warrens, n-way metrics, J. Classification 27 (2010) In press.

[31] A. Dress, M. Krüger, Parsimonious phylogenetic trees in metric spaces and simulated annealing, Adv. in Appl. Math. 8 (1) (1987) 8–37.

[32] J. Adámek, H. Herrlich, G. E. Strecker, Abstract and concrete categories, Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, 1990, the joy of cats, A Wiley-Interscience Publication.

[33] R. Espínola, W. A. Kirk, Fixed point theorems in $\mathbb{R}$-trees with applications to graph theory, Topology Appl. 153 (7) (2006) 1046–1055.

[34] I. Chiswell, Introduction to Λ-trees, World Scientific Publishing Co. Inc., River Edge, NJ, 2001.

[35] G. Edgar, Measure, topology, and fractal geometry, 2nd Edition, Undergraduate Texts in Mathematics, Springer, New York, 2008.

[36] S. N. Evans, A. Winter, Subtree prune and regraft: a reversible real tree-valued Markov process, Ann. Probab. 34 (3) (2006) 918–961.

[37] W. Beyer, M. Stein, T. Smith, S. Ulam, A molecular sequence metric and evolutionary trees, Mathematical Biosciences 19 (1974) 9–25.

[38] M. Waterman, T. Smith, M. Singh, W. Beyer, Additive evolutionary trees, Journal of Theoretical Biology 64 (1977) 199–213.