Appendix for "Convexity on treespaces"

Bo Lin, Bernd Sturmfels, Xiaoxian Tang, and Ruriko Yoshida

1 Geodesics in Global NPC Orthant Spaces

Definition 1.1 (Path/Path Length/Geodesic)
Suppose $\mathcal{T}$ is a metric space with the metric d. For any $D_1, D_2 \in \mathcal{T}$, if there exists a continuous map $\gamma : [0,1] \to \mathcal{T}$ such that $\gamma(0) = D_1$ and $\gamma(1) = D_2$, then $\gamma([0,1])$ is said to be a path *between D_1 and D_2. For any path $\gamma([0,1])$, the* path length *is defined as the supermum of the set*

$$\{\sum_{i=0}^{k-1} d(\gamma(x_i), \gamma(x_{i+1}))|0 \leq x_0 \leq \cdots \leq x_k \leq 1\}.$$

A path between two points is said to be a geodesic *if the path length is the infimum among all the paths between the two points.*

Definition 1.2 (Geodesic Space)
Suppose $\mathcal{T}$ is a metric space with the metric d. $\mathcal{T}$ is said to be a geodesic space *if for any $D_1, D_2 \in \mathcal{T}$, $d(D_1, D_2)$ is the path length of geodesic between D_1 and D_2.*

Definition 1.3 (Global NPC/CAT(0))
Suppose $\mathcal{T}$ is a geodesic space with the metric d. Consider a triangle T in $\mathcal{T}$ of side lengths a, b, c, and build a comparison triangle T' with the same lengths in Euclidean plane. Consider a chord of length l in T which connects two points on the boundary of T; there is a corresponding comparison chord in T', say of length l'. If for every triangle T in $\mathcal{T}$ and every chord in T we have $l \leq l'$, $\mathcal{T}$ is said to be global nonpositively curved (NPC) *or* CAT(0).

Lemma 1.4 (Lemma 6.2 in [9])
In a global NPC space, the geodesic between any two points is unique.

Notation 1.5
For any two points D_1 and D_2 in a global NPC space, we denote the geodesic between D_1 and D_2 by $G(D_1, D_2)$.

In our context, we assume every convex hull/simplex/simplicial complex is defined in a normed affine space.

Notation 1.6
For any finite set σ, $conv(\sigma)$ denotes the convex hull of σ, particularly, for any two points σ_1 and σ_2, $conv(\{\sigma_1, \sigma_2\})$ is denoted by $[\sigma_1, \sigma_2]$. For any simplex/complex Ω, $ver(\Omega)$ denotes the set of vertices of Ω.

Definition 1.7 (Orthant/Orthant Space)
Suppose Ω is a simplicial complex. For any cell F of Ω, the set

$$\{\sum_{\beta \in ver(F)} l_\beta e_\beta | l_\beta \in \mathbb{R}_{\geq 0}\}$$

is said to be an orthant *w.r.t F, denoted by $\mathcal{O}(F)$. For any simplicial complex Ω, the union $\bigcup_{\text{all cells } F \in \Omega} \mathcal{O}(F)$ is said to be an* orthant space *w.r.t Ω, denoted by $\mathcal{O}(\Omega)$.*

It is pointed out in [9] (Page 38, the third paragraph) that any orthant space is a geodesic space. That means any orthant space is a metric space and the distance between any two points is defined by the path length of geodesic. However, we need extra condition on Ω such that $\mathcal{O}(\Omega)$ is global NPC. See Definition 2.8 and Lemma 2.9.

Definition 1.8 (Flag Complex)
A simplicial complex Ω is said to be a flag complex *if for any $\sigma \subset ver(\Omega)$, if $[\sigma_1, \sigma_2] \in \Omega$ for any $\sigma_1, \sigma_2 \in \sigma$, then $conv(\sigma) \in \Omega$.*

Lemma 1.9 (Proposition 6.14 in [9])
An orthant space $\mathcal{O}(\Omega)$ is global NPC if and only if Ω is a flag complex.

We review the main idea of *Geodesic Treepath Problem* (GTP) algorithm [11] for computing geodesics in BHV space, which can be naturally extended to any global NPC orthant space (see [9, Corollary 6.19]).

Definition 1.10 (Leg/Projection/Projection Norm)
Suppose F is a simplex. For any $D \in \mathcal{O}(F)$, suppose $D = \sum_{\sigma \in ver(F)} l_\sigma e_\sigma$. For each $\sigma \in ver(F)$, l_σ is said to be the leg *of D w.r.t. σ. For any $A \subset ver(F)$, $\sum_{\sigma \in A} l_\sigma e_\sigma$ and $\sqrt{\sum_{\sigma \in A} l_\sigma^2}$ are said to be the* projection *and* projection norm *of D w.r.t. A, noted by D_A and $||D_A||$, respectively.*

Given an n-dimensional flag complex Ω and two points $D^{(\sigma)}, D^{(\tau)} \in \mathcal{O}(\Omega)$, without loss of generality, assume that $D^{(\sigma)} \in \mathcal{O}(F_\sigma)$ and $D^{(\tau)} \in \mathcal{O}(F_\tau)$ where F_σ and F_τ are two cells of Ω with vertex sets $\sigma = \{\sigma_1, \ldots, \sigma_{n+1}\}$ and $\tau = \{\tau_1, \ldots, \tau_{n+1}\}$ ($\sigma \cap \tau = \emptyset$). $G(D^{(\sigma)}, D^{(\tau)})$ is determined by ordered partitions $\mathcal{A} = (A_1, \ldots, A_q)$ and $\mathcal{B} = (B_1, \ldots, B_q)$ of σ and τ, which satisfy three properties:

(P1) for each $i > j$, $conv(A_i \cup B_j) \in \Omega$

(P2) $\frac{||D^{(\sigma)}_{A_1}||}{||D^{(\tau)}_{B_1}||} \leq \frac{||D^{(\sigma)}_{A_2}||}{||D^{(\tau)}_{B_2}||} \leq \cdots \leq \frac{||D^{(\sigma)}_{A_q}||}{||D^{(\tau)}_{B_q}||}$

(P3) for $i = 1, \ldots, q$, there do not exist nontrivial partitions $L_1 \cup L_2$ of A_i and $R_1 \cup R_2$ of B_i such that $conv(L_2 \cup R_1) \in \Omega$ and $\frac{||D^{(\sigma)}_{L_1}||}{||D^{(\tau)}_{R_1}||} < \frac{||D^{(\sigma)}_{L_2}||}{||D^{(\tau)}_{R_2}||}$.

Furthermore, $G(D^{(\sigma)}, D^{(\tau)})$ is 1-complex made of $q+1$ line segments living in $q+1$ different orthants determined by $\mathcal{A}$ and $\mathcal{B}$. See Lemma 2.11 below.

Lemma 1.11
$G(D^{(\sigma)}, D^{(\tau)}) = \bigcup_{i=0}^{q} [D_i, D_{i+1}]$ where

- *$D_0 = D^{(\sigma)}$ and $D_{q+1} = D^{(\tau)}$*
- *for any i $(0 \leq i \leq q)$, $[D_i, D_{i+1}] \subset \mathcal{O}(F_i)$*

where again $F_i = conv((\sigma \setminus \cup_{k=1}^{i} A_k) \cup (\cup_{k=1}^{i} B_k))$ for $i = 0, \ldots, q$.

Definition 1.12 (Transition/Codimension/Depth)
$F_0, \ldots, F_q$ stated in Lemma 2.11 are said to be the transitions *w.r.t $G(D^{(\sigma)}, D^{(\tau)})$. For $i = 0, \ldots, q$, the number $\sum_{k=1}^{i}(|A_k| - |B_k|)$ is said to be the* codimension *of F_i. The maximum codimension of the transitions is said to be the* depth *of $G(D^{(\sigma)}, D^{(\tau)})$.*

Based on these facts above, computing $G(D^{(\sigma)}, D^{(\tau)})$ is actually to compute the ordered partitions $\mathcal{A}$ and $\mathcal{B}$ and the "bent" points $D_1, \ldots, D_q$ living on the boundaries of transitions. The sketch of GTP algorithm is given below.

Algorithm 1.13 (Pseudo algorithm on computing geodesics in $\mathcal{O}(\Omega)$)
Input: *$D^{(\sigma)}$ and $D^{(\tau)}$*
Output: *$G(D^{(\sigma)}, D^{(\tau)})$*

1. *Initialize $\mathcal{A}^{(0)} = \{\sigma\}$, $\mathcal{B}^{(0)} = \{\tau\}$.*
2. *Note for each k $(k \geq 0)$,* **(P1)** *and* **(P2)** *always hold for $\mathcal{A}^{(k)}$ and $\mathcal{B}^{(k)}$. Check whether* **(P3)** *holds.*
   (a) *If no, then split some elements A_i and B_i in $\mathcal{A}^{(k)}$ and $\mathcal{B}^{(k)}$ respectively, and re-index the new partitions to get $\mathcal{A}^{(k+1)}$ and $\mathcal{B}^{(k+1)}$. Go to Step 2.*
   (b) *If yes, then we are done.*

Remark 1.14
We have several remarks on GTP algorithm below.

(1) Recall that we have assumed $\sigma \cap \tau = \emptyset$. If this hypothesis is not satisfied then an easy modification of the above construction yields the geodesic as well. See the GTP algorithm with common edges in [11, page 18] for more details.

(2) *If there exists* i $(1 \leq i \leq q-1)$ *such that* $\frac{||D^{(\sigma)}_{A_i}||}{||D^{(\tau)}_{B_i}||} = \frac{||D^{(\sigma)}_{A_{i+1}}||}{||D^{(\tau)}_{B_{i+1}}||}$, $\{A_1, \ldots, A_i \cup A_{i+1}, \ldots, A_q\}$ *and* $\{B_1, \ldots, B_i \cup B_{i+1}, \ldots, B_q\}$ *are also ordered partitions of* σ *and* τ *w.r.t.* $G(D^{(\sigma)}, D^{(\tau)})$. *So every "*$\leq$*" in* **(P2)** *can be replaced with "*$<$*".*

(3) *If there exists* i $(1 \leq i \leq q)$ *such that* $||D^{(\sigma)}_{A_j}|| = 0$ *or* $||D^{(\tau)}_{B_j}|| = 0$, *we consider* A_j *and* B_j *as the common set of* σ *and* τ. *So we can assume each* $\frac{||D^{(\sigma)}_{A_i}||}{||D^{(\tau)}_{B_i}||}$ *in* **(P2)** *is a positive number.*

Example 1.15 (Geodesics in 9-Space)

Experiment 1.16 (Geodesics in BHV Tree Space)
BHV tree space is another typical global NPC orthant space. In BHV tree space, depth is regarded as a quality measure for geodesics: the smaller the depth the better the geodesic. Optimal geodesics have depth 0. Such geodesics are line segments within a single orthant. These occur if and only if the starting point and target point are in the same orthant. Generally, the two given points are not in the same orthant. In this case, the best-case scenario is depth 1, meaning that each transition has codimension 1. On the other extreme are the geodesics of depth $n-2$. These are the *cone paths*: they occur when $\mathcal{A}$ and $\mathcal{B}$ are singletons. Cone paths are bad from a statistical perspective because the give rise to *sticky means*, see e.g. [8] or [9, §5.3]. Fix the number n of taxa between 4 and 20. For each such n, we sampled 1000 random pairs $\left\{D^{(\sigma)}, D^{(\tau)}\right\}$ from tree space $\mathcal{U}_n^{[1]}$, and we computed their geodesic. The depths of these geodesics are integers between 0 and $n-2$. Table 1 shows their distribution.

n\depth	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
4	8.4	58.4	33.2																
5	1.6	26.4	47.4	24.6															
6	0.2	13.2	36.7	31.5	18.4														
7	0	4	25.9	29.9	22.2	18													
8	0	1.1	15	28.9	25	17.1	12.9												
9	0	0.8	8	22.1	25.9	18.3	14.5	10.4											
10	0	0.4	3.3	17.2	22.3	20.6	14.1	13.2	8.9										
11	0	0.2	1.5	10.4	17.6	20.3	16.8	12.8	11.1	9.3									
12	0	0.2	0.1	6	14.1	20.4	13.9	14.6	12.7	10.5	7.5								
13	0	0.2	0.4	4.2	10.1	17.2	15.9	12.5	11	9.8	9.1	9.6							
14	0	0.2	0	2.7	9.3	14.9	15.5	12.2	11.3	10.4	8.7	8	6.8						
15	0	0.1	0	1.4	5.9	12.7	13	13.1	11.3	9.2	8.9	8.5	7.5	8.4					
16	0	0	0	1	5	11.2	11.4	11.3	11.2	9.9	8.1	9.1	7.3	6.7	7.8				
17	0	0	0	0.2	3.4	5.9	10.7	11	11.2	11.5	8.4	7.9	7.9	6.2	8.5	7.2			
18	0	0	0.1	0.4	1.5	6.5	8.7	10.5	10.9	9.7	7.9	7.5	7.1	8.7	7.7	6.5	6.3		
19	0	0	0	0.2	1.6	5	7.2	9.3	9.6	8.5	7.5	8.3	7.4	6.1	9.2	7.4	6.8	5.9	
20	0	0	0	0	0.5	3	6.7	7.6	11.2	9.8	9.4	8.2	5.9	7.5	6.9	6.9	4.5	5.7	6.2

Table 1: The rows are labeled by the number n of taxa and the columns are labeled by the possible depths of a geodesic in tree space. The entries in each row sum to 100%. They are the frequencies of the depths among 1000 geodesics, randomly sampled using Algorithm 2.17.

For instance, the first row concerns 1000 random geodesics on the BHV surface $\mathcal{U}_4^{[1]}$. Of these geodesics, 8.4% were in a single triangle or quadrangle, 58.4% had depth 1, and 33.2% were cone paths. For $n = 20$, the fraction of cone paths equals 6.2%.

The data in Table 1 depend on the specific probability distribution on $\mathcal{U}_n^{[1]}$ that was used for the sampling. In our experiment the sampling was done using the method described in Algorithm 2.17.

Algorithm 1.17 (Generating a sample of random normalized equidistant trees of n leaves)
Input: *The number n of leaves, and the sample size N.*
Output: *A sample of N random normalized equidistant trees in the tree space $\mathcal{U}_n^{[1]}$.*

1. *Set $\mathcal{S} = \emptyset$.*
2. *For $i = 1, \ldots N$, do*
   - (a) *Generate a tree D_i using the function* `rcoal` *from the* `ape` *package* [13] *in* `R`.
   - (b) *Randomly permute the leaf labels on the metric tree D_i.*
   - (c) *Change the clade nested structure of D_i by randomly applying the nearest neighbor interchange (NNI) operation n times.*
   - (d) *Turn D_i into an equidistant tree using the function* `compute.brtime` *in* `ape`.
   - (e) *Normalize U_i so that the distance from the root to each leaf is* $1/2$.
   - (f) *Add D_i to the output set $\mathcal{S}$.*
3. *Return $\mathcal{S}$.*

We conclude our exposition on global NPC orthant space by mentioning one concrete scenario for how it used in statistical phylogenetics. This is the work of Nye [10] on principal component analysis. Nye works in the usual non-compact BHV space of non-equidistant trees. Our aim here is to just convey the basic geometric ideas.

Nye [10] defines a *line* L in BHV tree space to be an unbounded path such that every bounded subpath is the geodesic between its endpoints. Suppose that L is such a line, and x any tree metric that is not in L. Proposition 2.1 in [10] shows that L contains a unique point y that is closest to x in the BHV metric. We call y the *projection* of x onto the line L. Given x and L, it can be computed as follows. Fixing a base point $L(0)$ on the line, one choses a geodesic parametrization $L(t)$ of the line. This means that t is the distance $d(L(0), L(t))$. Also let r denote the distance from x to $L(0)$. By the triangle inequality, the desired point is $y = L(t^*)$ for some $t^* \in [-r, r]$. The distance $d(x, L(t))$ is a continuous function of t. Our task is to find the minimum t^* of that function on the closed interval $[-r, r]$. This is done easily using numerical methods. The uniqueness of t^* follows from the CAT(0) property.

Suppose we are given a collection $\{x_1, x_2, \ldots, x_N\}$ of tree metrics on n taxa. This is our data set for phylogenetic analysis. Nye's method computes a first principal line (regression line) for these data inside the BHV space. This is done as follows. One first computes the *centroid* x_0 of the N given trees. This can be done using the iterative method in [2, Theorem 4.1]. Now, the desired regression line L is one of lines through x_0. For any such line L, we can compute the projections $y_1, \ldots, y_n$ of the data points $x_1, \ldots, x_n$. The goal is now to find the line L that minimizes a certain objective function. Nye proposes two such functions:

$$f_{\|}(L) := \sum_{i=1}^{N} d(x_0, y_i)^2 \quad \text{or} \quad f_{\perp}(L) := \sum_{i=1}^{N} d(x_i, y_i)^2.$$

This function of L is minimized using an iterative numerical procedure.

While the paper [10] represents a milestone concerning statistical inference in BHV tree space, it left open the problem of computing higher-dimensional principal components. First, what are the geodesic planes? Which of them is the regression plane for $x_1, x_2, \ldots, x_N$? Ideally, a plane in tree space would be a 2-dimensional subcomplex that contains the geodesic triangle formed by any three of its points. Outside a single cone, do such planes even exist? Such questions were raised in [10, §6]. The answer requires a convexity theory in global NPC orthant spaces.

2 Convexity in Global NPC Orthant Spaces

In this section, we always assume that Ω is a flag complex.

Definition 2.1 (Geodesically Convex)
A subset T of $\mathcal{O}(\Omega)$ is said to be geodesically convex *if for any two points $D_1, D_2 \in T$, $G(D_1, D_2) \subset T$.*

Definition 2.2 ("$\leq$" in an orthant)
Suppose F is a cell of Ω and $D_1 = \sum\limits_{\beta \in ver(F)} l_\beta^{(1)} e_\beta, D_2 = \sum\limits_{\beta \in ver(F)} l_\beta^{(2)} e_\beta \in \mathcal{O}(F)$. If for any $\beta \in ver(F)$, $l_\beta^{(1)} \leq l_\beta^{(2)}$, then we say $D_1 \leq D_2$.

Theorem 2.3
Let $T \subset \mathcal{O}(\Omega)$. T is geodesically convex if

(1) for each orthant $\mathcal{O}(F) \subset \mathcal{O}(\Omega)$, the set $T_F \triangleq T \cap \mathcal{O}(F)$ is convex;

(2) each convex set T_F is downward closed, i.e., if $D_1 \in \mathcal{O}(F)$, $D_2 \in T_F$ and $D_1 \leq D_2$, then $D_1 \in T_F$.

Proof. For any $D^{(\sigma)} \in T_{F_\sigma}$ and $D^{(\tau)} \in T_{F_\tau}$, where F_σ and F_τ are two faces of Ω with vertex sets σ and τ, we have to show $G(D^{(\sigma)}, D^{(\tau)}) \subset T$. Assume that $\sigma \cap \tau = \alpha$. Let $\sigma^* = \sigma \backslash \alpha$ and $\tau^* = \tau \backslash \alpha$. Suppose the ordered partitions of σ^* and τ^* *w.r.t.* $G(D^{(\sigma)}, D^{(\tau)})$ are $\mathcal{A} = \{A_1, \ldots, A_q\}$ and $\mathcal{B} = \{B_1, \ldots, B_q\}$. By lemma 2.11, we know

$$G(D^{(\sigma)}, D^{(\tau)}) = [D^{(\sigma)}, D_1] \cup [D_1, D_2] \cup \cdots [D_{q-1}, D_q] \cup [D_q, D^{(\tau)}].$$

By the condition (1), we only need to prove that $D_1, \ldots, D_q \in T$. We prove the conclusion by induction on q.

Suppose $q = 1$. By Definitions 2.10 and 3.2 and the condition (2), $D_\alpha^{(\sigma)}, D_\alpha^{(\tau)} \in T$. By the step (c) of GTP algorithm with common edges in [9, page 18], we see $D_1 \in [D_\alpha^{(\sigma)}, D_\alpha^{(\tau)}]$. By the condition (1), we have $D_1 \in T$.

Suppose the conclusion holds for $q = m$ ($m \geq 1$). Then we want to prove the conclusion for $q = m + 1$. By Remark 2.14 (2–3), assume $0 < \frac{||D_{A_1}^{(\sigma)}||}{||D_{B_1}^{(\tau)}||} < \frac{||D_{A_2}^{(\sigma)}||}{||D_{B_2}^{(\tau)}||}$. First, we show

$D_1 \in T$. More specifically, we show there exists $D \in T_{F_\sigma} \cap T_{F_1}$ such that $D_1 \leq D$ where $F_1 = conv(B_1 \cup A_2 \cup \cdots A_q \cup \alpha)$. In fact, we calculate D_1 according to the GTP algorithm [9]. Note $D_1 \in F_\sigma \cap F_1 = conv(A_2 \cup \cdots A_q \cup \alpha)$. So we only need to calculate $D_{1_{A_j}}$ $(2 \leq j \leq q)$ and D_{1_α}:

(i). for each j $(2 \leq j \leq q)$, by (**P2**),

$$D_{1_{A_j}} = \frac{D^{(\sigma)}_{A_j}}{||D^{(\sigma)}_{A_j}||}\left(\frac{||D^{(\tau)}_{B_1}||||D^{(\sigma)}_{A_j}|| - ||D^{(\sigma)}_{A_1}||||D^{(\tau)}_{B_j}||}{||D^{(\sigma)}_{A_1}|| + ||D^{(\tau)}_{B_1}||}\right);$$

(ii).

$$D_{1_\alpha} = \frac{||D^{(\tau)}_{B_1}||}{||D^{(\sigma)}_{A_1}|| + ||D^{(\tau)}_{B_1}||}D^{(\sigma)}_\alpha + \frac{||D^{(\sigma)}_{A_1}||}{||D^{(\sigma)}_{A_1}|| + ||D^{(\tau)}_{B_1}||}D^{(\tau)}_\alpha$$

Note also we have the facts below.

(iii). Since $D^{(\sigma)} \in T_{F_\sigma}$, we have $D^{(\sigma)}_{\sigma\backslash A_1} \in T_{F_\sigma}$ by the condition (2).

(vi). Since $D^{(\tau)} \in T_{F_\tau}$, we have $D^{(\tau)}_\alpha \in T_{F_\tau}$ by the condition (2). Note $\alpha = \sigma \cap \tau$. So $D^{(\tau)}_\alpha \in \mathcal{O}(F_\sigma)$ and Hence $D^{(\tau)}_\alpha \in T_{F_\sigma}$.

Let $t = \frac{||D^{(\tau)}_{B_1}||}{||D^{(\sigma)}_{A_1}||+||D^{(\tau)}_{B_1}||}$. By the facts (i–iv),

$$\begin{aligned}
D_1 &= \Sigma^q_{j=2} D_{1_{A_j}} + D_{1_\alpha} \\
&= \Sigma^q_{j=2} D^{(\sigma)}_{A_j}\left(t - (1-t)\frac{||D^{(\tau)}_{B_j}||}{||D^{(\sigma)}_{A_j}||}\right) + tD^{(\sigma)}_\alpha + (1-t)D^{(\tau)}_\alpha \\
&\leq \Sigma^q_{j=2} tD^{(\sigma)}_{A_j} + tD^{(\sigma)}_\alpha + (1-t)D^{(\tau)}_\alpha \\
&= t\left(\Sigma^q_{j=2} D^{(\sigma)}_{A_j} + D^{(\sigma)}_\alpha\right) + (1-t)D^{(\tau)}_\alpha \\
&= tD^{(\sigma)}_{\sigma\backslash A_1} + (1-t)D^{(\tau)}_\alpha \\
&\in T_{F_\sigma} \cap T_{F_1}
\end{aligned}$$

The second "=" follows from (i–ii), the "$\leq$" follows from the assumption $||D^{(\sigma)}_{A_1}|| > 0$ and the "$\in$" follows from (iii–vi) and the condition (1). By Lemma 2.11, $G(D_1, D^{(\tau)}) = [D_1, D_2] \cup [D_2, D_3] \cup \cdots \cup [D_q, D^{(\tau)}]$. By the induction hypothesis, we know $D_2, \ldots, D_q \in T$. □

Definition 2.4 (Geodesic Convex Hull)
Let $S = \{D_1, D_2, \ldots, D_k\}$ be a finite set of points in $\mathcal{O}(\Omega)$. The geodesic convex hull $conv(S)$ *of S is the smallest geodesically convex set that contains S. If $k = 3$ then we call it a* geodesic triangle.

Our main result in this section offers a decomposition of an arbitrary geodesic polytope into a finite union of convex polytopes (in the usual sense) that fit together along faces.

Definition 2.5
Let $S \subset \mathcal{O}(\Omega)$. The set $g(S)$ is defined as

$$\bigcup_{D_1, D_2 \in S} G(D_1, D_2).$$

For positive integer $n \geq 2$, we define $g^n(S)$ recursively as $g^n(S) = g(g^{n-1}(S))$.

Lemma 2.6
Let S be a set of points in $\mathcal{O}(\Omega)$. Then

$$conv(S) = \bigcup_{n=1}^{\infty} g^n(S).$$

In other words, the geodesic convex hull of S is the set of all points in $\mathcal{O}(\Omega)$ that can be generated in finite steps from S by taking geodesic paths.

Proof. By Definition 3.4, if a set $T \subset conv(S)$, then $g(T) \subset conv(S)$. Since $S \subset conv(S)$, by induction on n we can prove that $g^n(S) \subset conv(S)$ for all positive integers n. Then $\bigcup_{n=1}^{\infty} g^n(S) \subset conv(S)$. On the other hand, for any two points $D_1, D_2 \in \bigcup_{n=1}^{\infty} g^n(S)$, there exist positive integers n_1, n_2 such that $D_1 \in g^{n_1}(S), D_2 \in g^{n_2}(S)$. Then the geodesic path $G(D_1, D_2) \subset g^{\max(n_1,n_2)+1}(S)$, so $G(D_1, D_2) \subset \bigcup_{n=1}^{\infty} g^n(S)$. Furthermore $\bigcup_{n=1}^{\infty} g^n(S)$ contains S, so it is the smallest set that satisfying the conditions in Definition 3.4, hence it equals to $conv(S)$. □

Definition 2.7 (Geodesics between two sets of points)
Suppose in a global NPC orthant space A is a set of points within one orthant, and B is a set of points within another orthant. Define

$$Geo(A, B) = \bigcup_{a \in A, b \in B} G(a, b).$$

The following proposition shows that locally the ”Geo” operation is not commutative with taking (ordinary) convex hull.

Proposition 2.8
Suppose A, B are the same as in the above definition. If F is a cell of a global NPC orthant space, then in general

$$conv(Geo(A, B) \cap F) \neq Geo(conv(A), conv(B)) \cap F. \tag{1}$$

The following example is an example to demonstrate (1).

Example 2.9 (Geodesics in 5-Space)
Consider a flag complex with 5 *vertices* $1,2,3,4,5$, *where the maximal dimensional cliques are* $\{1,2,3\}$, $\{2,3,4\}$ *and* $\{3,4,5\}$. *Then the corresponding global NPC orthant space has* 3 *orthant with dimension* 3. *We use five coordinates to denote each point.*

Now let $A = \{(1,1,1,0,0)\}$, $B = \{(0,0,1,0,1),(0,0,0,1,0)\}$ *and* F *is the quadrant of axis* $2,3$ *(*$F = \{(0,b,c,0,0)|b,c \geq 0\}$*). Then*

$$Geo(A,B) \cap F = \{(0,0,1,0,0),(0,\frac{1}{2},\frac{1}{2},0,0)\}$$

so $conv(Geo(A,B) \cap F)$ *is the line segment connecting the two points* $(0,0,1,0,0),(0,\frac{1}{2},\frac{1}{2},0,0)$.

On the other hand, $conv(A) = A$ *and* $conv(B) = \{(0,0,t,1-t,t)|0 \leq t \leq 1\}$. *In particular* $(0,0,\frac{1}{2},\frac{1}{2},\frac{1}{2}) \in conv(B)$. *And note that* $G((1,1,1,0,0),(0,0,\frac{1}{2},\frac{1}{2},\frac{1}{2}))$ *contains the point* $(0,0,\frac{2}{3},0,0)$, *which belongs to* F. *Then*

$$(0,0,\frac{2}{3},0,0) \in Geo(conv(A),conv(B)) \cap F.$$

But $(0,0,\frac{2}{3},0,0) \notin conv(Geo(A,B) \cap F)$, *so this example proves the above proposition.*

Lemma 2.10
Suppose $G(D^{(\sigma)},D^{(\tau)}) = [D^{(\sigma)},D_1] \cup G(D_1,D^{(\tau)})$ *and* $[D^{(\sigma)},D_1] \subset \mathcal{O}(F)$ *where* D_1 *is defined as that in Lemma 2.11. For any* $D \in \mathcal{O}(F)$, *if* D *is on the line determined by* $D^{(\sigma)}$ *and* D_1, *then* $G(D,D^{(\tau)}) = [D,D_1] \cup G(D_1,D^{(\tau)})$.

Proof. Assume $D = tD^{\sigma} + (1-t)D_1$. Remark that $t \geq 0$ since $D \in \mathcal{O}(F)$. But we don't require $1-t \geq 0$ since D may not be in the interval $[D^{(\sigma)},D_1]$. Assume that $\sigma \cap \tau = \alpha$. Let $\sigma^* = \sigma \backslash \alpha$ and $\tau^* = \tau \backslash \alpha$. Suppose the ordered partitions of σ^* and τ^* *w.r.t.* $G(D^{(\sigma)},D^{(\tau)})$ are $\mathcal{A} = \{A_1,\ldots,A_q\}$ and $\mathcal{B} = \{B_1,\ldots,B_q\}$. By lemma 2.11, we know

$$G(D^{(\sigma)},D^{(\tau)}) = [D^{(\sigma)},D_1] \cup [D_1,D_2] \cup \cdots [D_{q-1},D_q] \cup [D_q,D^{(\tau)}].$$

We prove the conclusion by induction on q. If $q = 1$, then according to GTP algorithm, $D_{1_{\sigma^*}} = 0$ and $D_{1_\alpha} = \lambda D_\alpha^{(\sigma)} + (1-\lambda)D_\alpha^{(\tau)}$ where $\lambda = \frac{||D_{\tau^*}^{(\tau)}||}{||D_{\sigma^*}^{(\sigma)}||+||D_{\tau^*}^{(\tau)}||}$. So

$$D = tD^{(\sigma)} + (1-t)(\lambda D_\alpha^{(\sigma)} + (1-\lambda)D_\alpha^{(\tau)}).$$

Now it is easy to check by GTP alogirthm that $G(D,D^{(\tau)}) = [D,D_1] \cup [D_1,D^{(\tau)}]$. Suppose the conclusion holds for $q = m$ ($m \geq 1$). Now assume $q = m+1$. Note

$$G(D^{(\sigma)},D^{(\tau)}) = [D^{(\sigma)},D_1] \cup [D_1,D_2] \cup \cdots [D_m,D_{m+1}] \cup [D_{m+1},D^{(\tau)}].$$

By the induction hypothesis, we have $G(D,D_{m+1}) = [D,D_1] \cup G(D_1,D_{m+1})$. Therefore,

$$G(D,D^{(\tau)}) = G(D,D_{m+1}) \cup [D_{m+1},D^{(\tau)}] = [D,D_1] \cup G(D_1,D^{(\tau)}).$$

□

Example 2.11 (A Geodesic Triangle in 6-Space)
Consider a flag complex Ω *with* 6 *vertices* $1,2,3,4,5$ *and* 6*, where the maximal dimensional cliques are* $\{1,2,3\}$*,* $\{2,3,4\}$*,* $\{3,4,5\}$ *and* $\{4,5,6\}$*. Then the corresponding global NPC orthant space* $\mathcal{O}(\Omega)$ *has* 4 *orthants with dimension* 3*. We use six coordinates to denote each point. Note the four orthants can be embedded in* $\mathbb{R}^3$ *by setting* $l_4 = -l_1$*,* $l_5 = -l_2$*,* $l_6 = -l_3$ *for any* $(l_1, l_2, l_3, l_4, l_5, l_6) \in \mathcal{O}(\Omega)$*. Consider the polyhedral complex made of four polytopes* $OABH$*,* $OFHBD$*,* $OFED$ *and* $OFEC$ *in the four different orthants respectively, which fit together along three faces* OBH*,* OFD *and* OFE*, where*

$$O = (0,0,0,0,0,0), A = (4,6,6,0,0,0), B = (0,5,8,0,0,0), C = (0,0,0,1,2,3)$$

$$H = (0, \frac{14}{19}, \frac{14}{19}, 0, 0, 0), D = (0, 0, \frac{1}{7}, \frac{5}{7}, 0, 0), F = (0, 0, 0, \frac{14}{25}, 0, 0), E = (0, 0, 0, \frac{8}{11}, \frac{1}{11}, 0).$$

Denote the polyhedral complex by $\mathcal{G}$ *and denote polytopes* $OABH$*,* $OFHBD$*,* $OFED$ *and* $OFEC$ *by* $\mathcal{G}_1$*,* $\mathcal{G}_2$*,* $\mathcal{G}_3$ *and* $\mathcal{G}_4$ *respectively. We draw* $\mathcal{G}$ *in* $\mathbb{R}^3$*, see Figure 1. The main goal of this example is to show* $\mathcal{G} = conv(\{A, B, C\})$*.*

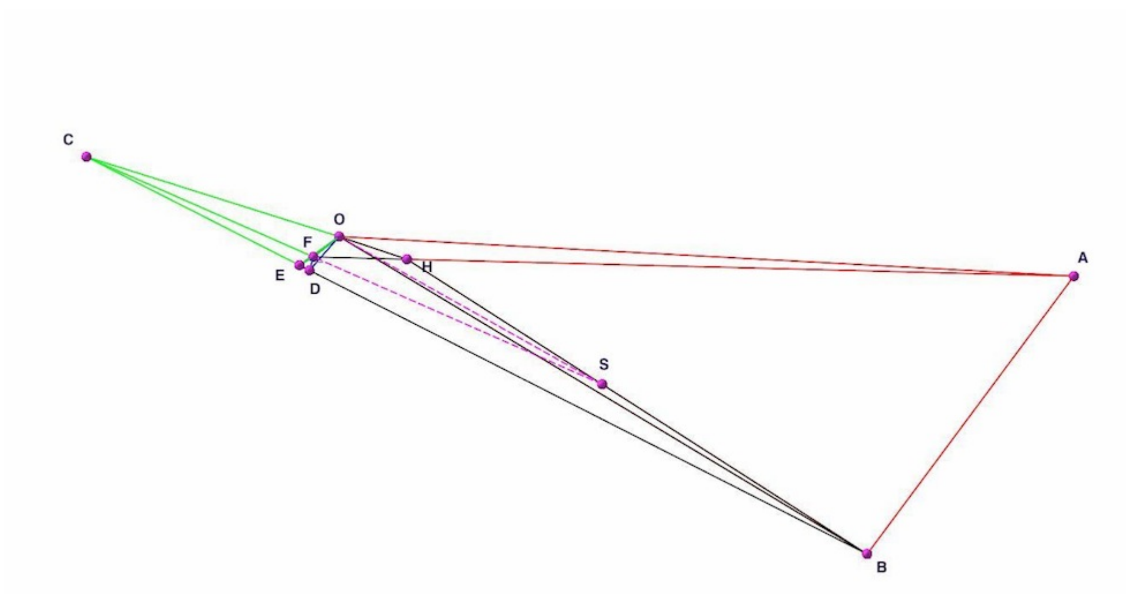

Figure 1: Geodesic Triangle $conv(\{A, B, C\})$

First, we show $\mathcal{G} \subset conv(\{A, B, C\})$*, we only need to show the eight points* O, A, B, C, H, D, F *and* E *(the vertices of the polytopes* $\mathcal{G}_i$ $(i = 1, 2, 3, 4)$*) are contained in* $conv(\{A, B, C\})$*. It is easy to check by GTP algorithm that* $G(A, C) = [A, O] \cup [O, C]$ *and* $G(B, C) = [B, D] \cup [D, E] \cup [E, C]$*. So by Definition 3.4,* $O, A, B, C, D, E \in conv(\{A, B, C\})$*. It is easy to check by GTP algorithm again that* $G(A, E) = [A, X] \cup [X, Y] \cup [Y, E]$ *where* $X = (0, \frac{11}{13}, \frac{12}{13}, 0, 0, 0)$ *and* $Y = (0, 0, \frac{6}{67}, \frac{44}{67}, 0, 0)$*. By Definition 3.4,* $X \in conv(\{A, B, C\})$ *and hence* $G(B, X) = [B, X] \subset conv(\{A, B, C\})$*. Now it is easy to check that* $S = \frac{9}{22}B + (1 - \frac{9}{22})X$ *and thus* $S \in conv(\{A, B, C\})$*. By GTP algorithm, we get* $G(S, C) = [S, F] \cup [F, C]$ *and* *therefore $F \in conv(\{A,B,C\})$. Finally, by GTP algorithm, we get $G(A,F)=[A,H]\cup[H,F]$ and hence $H \in conve(\{A,B,C\})$.*

Now in order to prove $conv(\{A,B,C\})\subset\mathcal{G}$, we prove $\mathcal{G}$ is geodesically convex. We only need to show for any $D^{(\sigma)}\in\mathcal{G}_i$ and $D^{(\sigma)}\in\mathcal{G}_j$ $(i\neq j)$, $G(D^{(\sigma)},D^{(\tau)})\subset\mathcal{G}$. If $i=j+1$ or $j=i+1$, it is easy to check that $\mathcal{G}_i\cup\mathcal{G}_j$ is convex in $\mathbb{R}^3$ and hence the conclusion holds. Now we discuss the three non-trivial cases below.

- *(I) For $i=1$ and $j=3$, it is easy to check that $\mathcal{G}_1\cup\mathcal{G}_2\cup\mathcal{G}_3$ is convex in $\mathbb{R}^3$. So the geodesic between $D^{(\sigma)}$ and $D^{(\tau)}$ is a straight line in $\mathbb{R}^3$ and the straight line is contained in $\mathcal{G}_1\cup\mathcal{G}_2\cup\mathcal{G}_3$.*

- *(II) For $i=2$ and $j=4$, it is easy to check that F is strictly contained in the Euclidean triangle CBH. Compute the intersect point of line FC and interval BH, we get $(0,\frac{28}{11},\frac{42}{11},0,0,0)$. Denote this point by S. It is seen that if $D^{(\sigma)}$ belongs to the polytope $OSBDF$, then $G(D^{(\sigma)},D^{(\tau)})$ is contained in the union of $OSBDF$ and $\mathcal{G}_3\cup\mathcal{G}_4$ since this union is the polytope $OSBC$ in $\mathbb{R}^3$. Now we only need to show if $D^{(\sigma)}$ belongs to the polytope $OSHF$, then the conclusion holds. Assume*

  $$D^{(\sigma)}=\lambda_1F+\lambda_2H+\lambda_3S+\lambda_4O=(0,\frac{14}{19}\lambda_2+\frac{28}{11}\lambda_3,\frac{14}{19}\lambda_2+\frac{42}{11}\lambda_3,\frac{14}{25}\lambda_1,0,0)$$

  $$D^{(\tau)}=\beta_1F+\beta_2E+\beta_3C+\beta_4O=(0,0,0,\frac{14}{25}\beta_1+\frac{8}{11}\beta_2+\beta_3,\frac{1}{11}\beta_2+2\beta_3,3\beta_3)$$

  where

  $$\Sigma_i\lambda_i=1 \tag{2}$$

  $$\Sigma_i\beta_i=1 \tag{3}$$

  $$\lambda_i,\beta_i\geq 0 \tag{4}$$

  Note $\sigma=\{2,3,4\}$ and $\tau=\{4,5,6\}$ $(\sigma\cap\tau=\{4\})$. According to GTP algorithm, we have two cases below.

- *(i) Suppose the ordered partitions w.r.t $G(D^{(\sigma)},D^{(\tau)})$ are $\{\{2,3\}\}$ and $\{\{5,6\}\}$. According to the condition (**P2**) in GTP algorithm, we have*

  $$\frac{\frac{14}{19}\lambda_2+\frac{28}{11}\lambda_3}{\frac{1}{11}\beta_2+2\beta_3}>\frac{\frac{14}{19}\lambda_2+\frac{42}{11}\lambda_3}{3\beta_3} \tag{5}$$

  Without loss of generality, we assume that β_2 and β_3 are not 0 at the same time (otherwise, $D^{(\sigma)}$ and $D^{(\tau)}$ are in the same orthant). Then (5) is equivalent to

  $$3\beta_3(\frac{14}{19}\lambda_2+\frac{28}{11}\lambda_3)-(\frac{14}{19}\lambda_2+\frac{42}{11}\lambda_3)(\frac{1}{11}\beta_2+2\beta_3)>0 \tag{6}$$

  According to Lemma 2.11, $G(D^{(\sigma)},D^{(\tau)})=[D^{(\sigma)},D_1]\cup[D_1,D^{(\tau)}]$. According to GTP algorithm, we compute

  $$D_1=(0,0,0,\lambda\frac{14}{25}\lambda_1+(1-\lambda)(\frac{14}{25}\beta_1+\frac{8}{11}\beta_2+\beta_3),0,0)$$

where $\lambda = \frac{\frac{1}{11}\sqrt{\beta_2^2+44\beta_2\beta_3+1573\beta_3^2}}{\frac{1}{11}\sqrt{\beta_2^2+44\beta_2\beta_3+1573\beta_3^2}+\frac{14}{209}\sqrt{242\alpha_2^2+2090\alpha_2\alpha_3+4693\alpha_3^2}}$. *In order to show* $D_1 \in \mathcal{G}$, *we only need to show* $D_1 \in [O, F]$, *i.e.,*

$$\lambda\frac{14}{25}\lambda_1 + (1-\lambda)(\frac{14}{25}\beta_1 + \frac{8}{11}\beta_2 + \beta_3) \leq \frac{14}{25} \tag{7}$$

We substitute $\alpha_1 = 1 - \alpha_2 - \alpha_3 - \alpha 4$ *and* $\beta_1 = 1 - \beta_2 - \beta_3 - \beta_4$ *into (6) and find (6) is equivalent to*

$$\frac{-2926\sqrt{\beta_2^2+44\beta_2\beta_3+1573\beta_3^2}(\lambda_2+\lambda_3+\lambda_4)+14\sqrt{242\lambda_2^2+2090\lambda_2\lambda_3+4693\lambda_3^2}(46\beta_2+121\beta_3-154\beta_4)}{5225\sqrt{\beta_2^2+44\beta_2\beta_3+1573\beta_3^2}+3850\sqrt{242\lambda_2^2+2090\lambda_2\lambda_3+4693\lambda_3^2}} \leq 0 \tag{8}$$

Under the assumption that β_2 *and* β_3 *are not* 0 *at the same time, the denominator in (7) is strictly positive. (7) is equivalent to*

$$-2926\sqrt{\beta_2^2+44\beta_2\beta_3+1573\beta_3^2}(\lambda_2+\lambda_3+\lambda_4)+14\sqrt{242\lambda_2^2+2090\lambda_2\lambda_3+4693\lambda_3^2}(46\beta_2+121\beta_3-154\beta_4) \leq 0 \tag{9}$$

If $46\beta_2 + 121\beta_3 - 154\beta_4 \leq 0$, *then (8) holds naturally. If* $46\beta_2 + 121\beta_3 - 154\beta_4 > 0$, *then (8) is equivalent to*

$$2926^2(\beta_2^2+44\beta_2\beta_3+1573\beta_3^2)(\lambda_2+\lambda_3+\lambda_4)^2 - 14^2(242\lambda_2^2+2090\lambda_2\lambda_3+4693\lambda_3^2)(46\beta_2+121\beta_3-154\beta_4)^2 \geq 0 \tag{10}$$

Let

$$f_1 = 3\beta_3(\frac{14}{19}\lambda_2 + \frac{28}{11}\lambda_3) - (\frac{14}{19}\lambda_2 + \frac{42}{11}\lambda_3)(\frac{1}{11}\beta_2 + 2\beta_3)$$
$$f_2 = 46\beta_2 + 121\beta_3 - 154\beta_4$$
$$f_3 = 1 - \lambda_2 - \lambda_3 - \lambda_4$$
$$f_4 = 1 - \beta_2 - \beta_3 - \beta_4$$
$$f_5 = 2926^2(\beta_2^2+44\beta_2\beta_3+1573\beta_3^2)(\lambda_2+\lambda_3+\lambda_4)^2 - 14^2(242\lambda_2^2+2090\lambda_2\lambda_3+4693\lambda_3^2)(46\beta_2+121\beta_3-154\beta_4)^2$$

Now we only need prove that if $f_1 > 0$, $f_2 > 0$, $f_3 \geq 0$, $f_4 \geq 0$ *and* $\lambda_i, \beta_i \geq 0$ $(i = 2, 3, 4)$, *then* $f_5 \geq 0$. *We show how to prove by cylindrical algebraic decomposition (CAD) tool if* $f_1 > 0$, $f_2 > 0$, $f_3 > 0$, $f_4 > 0$ *and* $\lambda_i, \beta_i > 0$ $(i = 2, 3, 4)$, *then* $f_5 \geq 0$. *For the case* $f_3 = 0$ *or* $f_4 = 0$ *or some* $\lambda_i = 0$ *or some* $\beta_i = 0$, *we can similarly do it. Let* $f = \Pi_{i=2}^4 \lambda_i \beta_i \Pi_{i=1}^5 f_i$. *By the theory of real algebraic geometry, we know that* $f \neq 0$ *is a finite union of open connected sets in* $\mathbb{R}^6$ *and the signs of* f_i $(i = 1, 2, 3, 4, 5)$ *and* λ_i, β_i $(i = 2, 3, 4)$ *don't change over each open connected set. By CAD tool, we can theoretically compute at least one rational point with all positive coordinates, namely* sample point, *in each open connected set and then check the signs of* f_i $(i = 1, 2, 3, 4, 5)$ *at each of the sample point. However, practically, we can not compute the sample points for* $f \neq 0$ *in reasonable time due to the high complexity. Remark that by Lemma 3.10, we only need to check the cases that* $D^{(\sigma)}$ *belongs to some face of* $OFHS$ *and* $D^{(\tau)}$ *belongs to some face of* $OFEC$. *That means when we compute sample points of* $f \neq 0$, *we can compute* 9 *smaller cases. In each case, we assume that* λ_k *(*k *is one number among* $2, 3, 4$*) and* β_t *(*t *is one number among* $2, 3, 4$*) are* 0, *which saves computational time. For instance, we substitute* $\lambda_2 = 0$ *and* $\beta_3 = 0$ *into* f *and then we obtain* 465 *sample points. By checking each sample point, we confirm that if* $f_1 > 0$, $f_2 > 0$, $f_3 > 0$, $f_4 > 0$ *and* $\lambda_i, \beta_i > 0$ $(i = 2, 3, 4)$, *then* $f_5 \geq 0$.

(ii) Suppose the ordered partitions w.r.t $G(D^{(\sigma)}, D^{(\tau)})$ are $\{\{2\},\{3\}\}$ and $\{\{5\},\{6\}\}$. The proof is similarly as that in (i).

(III) For $i = 1$ and $j = 4$, we have four cases below.

(i) If the geodesic path is cone path, then the conclusion holds since the only break point is O.

(ii) Suppose the ordered partitions w.r.t $G(D^{(\sigma)}, D^{(\tau)})$ are $\{\{1\},\{2,3\}\}$ and $\{\{4\},\{5,6\}\}$. Similarly as the proof in (II), we can check that the first break point $D_1 \in G(D^{(\sigma)}, D^{(\tau)})$ belongs to the face OBH and hence $[D^{(\sigma)}, D_1] \in \mathcal{G}_1$. By the conclusion (II), we see that $G(D_1, D^{(\tau)}) \in \mathcal{G}_2 \cup \mathcal{G}_3 \cup \mathcal{G}_4$. Therefore, $G(D^{(\sigma)}, D^{(\tau)}) \in \mathcal{G}$.

(iii) Suppose the ordered partitions are $\{\{1\},\{2\},\{3\}\}$ and $\{\{4\},\{5\},\{6\}\}$. The proof is similar as that in (ii).

(iv) Suppose the ordered partitions are $\{\{1,2\},\{3\}\}$ and $\{\{4,5\},\{6\}\}$. The proof is similar as that in (ii).

Theorem 2.12
Every geodesic polytope in global NPC orthant space contains the structure of a finite polyhedral complex. That complex can have the full dimension $n-2$ even for triangles ($k=3$).

Proof. GOAL. THE PROOF GOES HERE.
BOTH STATEMENTS ARE TERRIFIC NEW RESULTS,
ASSUMING THEY ARE TRUE. □

3 Tropical Convexity

4 Generalization to Tropical Linear Spaces

References

[1] F. Ardila and C. Klivans: *The Bergman complex of a matroid and phylogenetic trees*, Journal of Combinatorial Theory Ser. B **96** (2006) 38–49.

[2] L. Billera, S. Holmes and K. Vogtman: *Geometry of the space of phylogenetic trees*, Advances in Applied Mathematics **27** (2001) 733–767.

[3] E. Feichtner: *Complexes of trees and nested set complexes*, Pacific Journal of Mathematics **227** (2006) 271–286.

[4] E. Feichtner and B. Sturmfels: *Matroid polytopes, nested sets and Bergman fans*, Portugaliae Mathematica **62** (2005) 437–468.

[5] A. Gavruskin and A. Drummond: *The space of ultrametric phylogenetic trees*, `arXiv:1410.3544`.

[6] E. Gawrilow and M. Joswig: *polymake: a framework for analyzing convex polytopes*, in Polytopescombinatorics and computation (Oberwolfach, 1997), 43–73, DMV Sem., 29, Birkhäuser, Basel, 2000.

[7] D. Maclagan and B. Sturmfels: *Introduction to Tropical Geometry*, Graduate Studies in Mathematics, 161, American Mathematical Society, Providence, RI, 2015.

[8] E. Miller: Fruit flies and moduli: Interactions between biology and mathematics, *Notices of the AMS*, to appear.

[9] E. Miller, M. Owen and S. Provan: *Polyhedral computational geometry for averaging metric phylogenetic trees*, `arXiv:1211.7046`.

[10] T. Nye: *Principal components analysis in the space of phylogenetic trees*, Annals of Statistics **39** (2011) 2716–2739.

[11] M. Owen and S. Provan: *A fast algorithm for computing geodesic distances in tree space*, IEEE/ACM Trans. Computational Biology and Bioinformatics **8** (2011) 2–13.

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

[13] E. Paradis, J. Claude and K. Strimmer: *APE: analyses of phylogenetics and evolution in R language*, Bioinformatics **20** (2004) 289–290.

[14] C. Semple and M. Steel: *Phylogenetics*, Oxford Lecture Series in Mathematics and its Applications, 24, Oxford University Press, 2003.

[15] M. Steel and D. Penny: *Distribution of tree comparison metrics–Some new results*, Systematic Biology **42**(2) (1993) 126–141.

[16] D. Speyer and B. Sturmfels: *Tropical mathematics*, Mathematics Magazine **82** (2009) 163–173.