## LEVEL-1 PHYLOGENETIC NETWORKS AND THEIR BALANCED MINIMUM EVOLUTION POLYTOPES

#### CASSANDRA DURELL AND STEFAN FORCEY

ABSTRACT. Balanced minimum evolution is a distance-based criterion for the reconstruction of phylogenetic trees. Several algorithms exist to find the optimal tree with respect to this criterion. One approach is to minimize a certain linear functional over an appropriate polytope. Here we present polytopes that allow a similar linear programming approach to finding phylogenetic networks. We investigate a two-parameter family of polytopes that arise from phylogenetic networks, and which specialize to the Balanced Minimum Evolution polytopes as well as the Symmetric Travelling Salesman polytopes. We show that the vertices correspond to certain level-1 phylogenetic networks, and that there are facets or faces for every split. We also describe lower bound faces and a family of faces for every dimension.

## 1. INTRODUCTION

Phylogenetics is the theory and practice of using available genetic data to reconstruct the past. Given DNA from a set of taxa (individuals, species, or specific genes) we want to infer the ancestral relationships that existed, or rather were most likely to have existed. Most typically the result of such a study is a biological family tree, or *phylogenetic tree*. Nature, however, is not constrained to tree-like branching alone. Events such as horizontal gene transfer and hybridization result in the rejoining of branches, called reticulation. Biologists refer to this more general situation as a *phylogenetic network*.

Our main results herein are geometric, but they underlie a program of phylogenetic reconstruction using polytopes. A polytope is a well-defined multifaceted convex shape, bounded by linear inequalities. These shapes generalize polygons (2D) and polyhedra (3D) to all dimensions. An *n*-dimensional polytope has (n-1)-dimensional facets, and faces of all smaller dimensions. A polytope can be described by its bounding inequalities, or equivalently by its extreme corner points, or vertices. It is the intersection of the halfspaces given by its facetdefining inequalities, and it is the intersection of all convex sets which contain its vertices (the convex hull). Lower bound and upper bound inequalities give the least and greatest attainable values of the vector components of the vertices. Given a polytope P in  $\mathbb{R}^n$  and any vector  $\mathbf{x} \in \mathbb{R}^n$ , that vector can be used as the coefficients of (parallel, affine) linear inequalities. If (and only if) some vertices of P obey such a linear inequality sharply (by equality), while all other vertices obey it strictly, then (1) that inequality defines a face of P containing those vertices, (2) the vector  $\mathbf{x}$  is normal to that face, and (3) the linear functional defined by  $\mathbf{x}$  (the dot product with  $\mathbf{x}$ ) is either minimized or maximized over P at the vertices of that face.

Mathematically, as described in Steel [2016], a *phylogenetic network* N is a directed, acyclic, simple, connected graph, with n nodes of in-degree one labeled by the taxa  $[n] = \{1, \ldots, n\}$ .

<sup>2000</sup> Mathematics Subject Classification. 90C05, 52B11, 92D15.

Key words and phrases. polytopes, phylogenetics, trees, networks.



FIGURE 1. The directed phylogenetic network on the left is level-2, and has two biconnected components, along with the leaves. In the center is its underlying undirected network, which is 4-nested: for instance the edge labeled (a) is in 4 distinct undirected cycles. On the right is a binary level-5 network, which has a 13-nested underlying network.

(Acyclic means that there are no directed cycles, while simple means there are no multi-edges.) One node of N is the only source, with in-degree 0. This node is the *root*, or common ancestor. (In applications the root is often determine via attachment to a leaf which is an intentionally chosen unrelated *outgroup*.) The degree 1 nodes of N are the the only sinks, (with in-degree 1 and out-degree 0). These are the *leaves*, or extant taxa. All the non-leaf, non-root nodes are required to have total degree  $\geq 3$ . Reticulation nodes are those with in-degree > 1. We focus in the current paper on the *unrooted* or *undirected* phylogenetic networks, which have non-oriented edges, and thus no root. They still have n labeled degree 1 leaves, and all other nodes of degree  $\geq 3$ . In application these underlie directed networks with reticulation, or they can capture the options of an ambiguous ancestry. A *binary* phylogenetic network is required to have all its non-root, non-leaf nodes of degree exactly 3. We also rule out isolated triangles: the biconnected components of our networks N herein never have exactly 3 nodes. (A biconnected component is a subgraph which stays connected even after removal of any one edge.) Examples are in Fig. 1.

Phylogenetic networks vary in complexity, from the phylogenetic trees to the maximum level of connectedness in a complete graph. A commonly used gauge of complexity is *level*. A level-k directed phylogenetic network N has in each biconnected component a maximum of k reticulation nodes. For undirected networks the closest related concept is *nestedness*, as defined by Gambette et al. [2017a]. A k-nested network has a maximum of k cycles containing any one edge. (Cycles here are undirected and have no repeated nodes.) Usually, a level-k directed network will have a j-nested underlying undirected network with  $j \ge k$ . However, since a directed level-1 binary network has an underlying undirected graph with at most 1 cycle containing any given node, undirected 1-nested binary networks are often called level-1 binary as well. Trees, directed or not, are level-0, since they are cycle-free.

Here we show how to organize specified collections of phylogenetic networks by assigning networks to the vertices and faces of polytopes. The value of such an organization is that there are well known algorithms of linear programming (as in the simplex method, and the branch and cut algorithms) which can take advantage of knowing the bounding inequalities in order to find the optimal vertex of a polytope. For us, optimizing means finding the network that best fits the data according to the balanced minimum evolution criterion. In future work we plan to test out these algorithms for networks; in this paper we establish the underlying geometry. Prior work includes the branch and bound algorithms for phylogenetic trees as described in Catanzaro et al. [2012] and Catanzaro et al. [2015], as well as a version using splits in Forcey et al. [2018].

#### 2. Overview

In balanced minimum evolution (BME) methods we try to minimize the total branch length of a candidate phylogenetic tree for a given discrete metric. The BME method can be described as a linear programming problem. The convex hull of solutions for this problem, given n taxa, is the  $\binom{n}{2} - n$  dimensional polytope BME(n). In this paper we present a generalization of the BME polytopes which allows the solutions to be phylogenetic networks rather than trees.

The main results in this paper are theorems about the convex polytopes we define, denoted BME(n,k) for all  $0 \le k \le n-3$ . We prove that their vertices correspond bijectively to binary level-1 phylogenetic networks with n leaves and k non-trivial bridges, in Theorem 5.1, Theorem 5.2, and Corollary 5.3. In Theorem 5.9 we find a formula to count the vertices:

$$\binom{n-3}{k}\frac{(n+k-1)!}{(2k+2)!!}$$

Section 3 contains the basic definitions we need. In Section 4 we define and compare the new polytopes to well-known families: the symmetric travelling salesman polytopes (STSP) and the balanced minimum evolution polytopes. Ours are nested inside the STSP and outside the BME polytopes (after scaling) as shown in Theorem 5.7. In Section 5 we describe a good deal of the facial structure of our new polytopes. Every less-refined phylogenetic network turns out to correspond to multiple faces: one in each polytope with vertices which refine that network. See Theorem 5.5 for details. In Section 6 we show that some of these faces are actually facets, specifically those corresponding to splits as shown in Theorem 6.1. We also describe lower bound faces in Theorem 6.2. Finally in Section 7 we describe further results about the specific case of networks with five leaves and a single bridge. We give a complete classification of the 62 facets in this case, in Theorem 7.1, Theorem 7.2, Theorem 7.3, and Theorem 7.4. In Section 8 we preview the potential of our polytopes (and their relaxations, as mentioned in Theorem 8.1) in linear programming algorithms.

Some of the results in this paper, with abbreviated proofs, were announced in the extended abstract Devadoss et al. [2019], in the proceedings of FPSAC'19. Some of the definitions were also repeated there, for completeness. We thank Satyan Devadoss for collaboration on that extended abstract, and conversations about the work in this paper. We also thank the referees for excellent suggestions on an earlier version.

## 3. Basics

A slightly abbreviated version of the definitions in this section is included in Devadoss et al. [2019]. We include them here in order to be self-contained. We begin with the set  $[n] = \{1, 2, ..., n\}$ , in which the integers 1 through n stand for biological taxa. We often also are given a non-negative real-valued pairwise distance function **d**, (also known as a dissimilarity matrix), with outputs (entries) denoted  $d_{ij} = \mathbf{d}(i, j)$  for each pair of taxa  $i, j \in [n]$ . (Note that d can also be described as a discrete metric, or a distance vector.) We try to find an appropriate combinatorial structure to display that data. The structures we directly consider here are split networks and phylogenetic networks. We will restrict to the unrooted versions, and focus on specializations of one or both structures such as: phylogenetic trees, circular split networks, level-1 and 1-nested phylogenetic networks.

The simplest structure we consider is an (unrooted) phylogenetic tree. Mathematically this is a graph with no cycles, and with no nodes of degree 2. The nodes with degree larger than 2 are unlabeled, but the n leaves are labeled bijectively with our n taxa. There is also an option of assigning non-negative lengths to the edges of the tree. The appropriateness of this data structure is clear when the edge-lengths along the path between two leaves i and j sum to the given distance  $d_{ij}$  between those leaves. In fact, if our given pair-wise distances allow such a representation then the weighted tree is unique.

A split system is a more general data structure, (which specializes to a phylogenetic tree). A split of our set is a partition A|B of [n] into two parts. When one part has cardinality 1 we call the split trivial. A split system s on [n] is any collection of splits which contains all the trivial splits. We say a split system s' refines s when  $s' \supset s$ . A split network is a graphical representation of a split system. It is a special connected simple graph. The n taxa are again seen as the labeled leaves (degree 1 vertices). Each split is represented by a set of parallel edges which is a minimal cut of the graph; that is, removing that set of edges separates the graph into two components, whose respective leaves are the parts of the split. Sometimes there are (nontrivial) splits represented by a single edge. We call these edges (nontrivial) bridges. A phylogenetic tree is a special split network. Its splits are all bridges. Upon removing a bridge the leaves of the resulting disconnected components are the parts of the split. The two parts of each split are called clades. We say that the tree displays those splits.



FIGURE 2. A circular split system s, in the center (trivial splits not shown), with its circular split network on the left, and its polygonal representation on the right. This split network has one non-trivial bridge, giving the split  $\{1, 2, 6, 7\}|\{3, 4, 5\}$ . The split network is externally refined, so no bridges can be added.

A *circular* split network is one whose graph may be drawn on the plane without edge crossings, and with its leaves all on the exterior of the diagram. The terminology is due to the fact that any such drawing automatically produces a circular ordering of the leaves. Note that these drawings are not fixed in the plane—twisting around a bridge gives a different

diagram but represents the same split network. As shown in Devadoss and Petti [2017], there is an equivalent *dual polygonal representation* of any circular split network: Given a circular split system with a circular ordering c of the species, consider a regular n-gon, with the edges cyclically labeled according to c. For each split, draw a diagonal partitioning the appropriate edges; see Fig. 2. Note that splits which require multiple parallel edges in the the network picture correspond to diagonals that are *crossing* (they intersect other diagonals in the picture) and that bridges become *noncrossing* diagonals.

**Definition 3.1.** An externally refined split network s is such that there is no split network s' both refining s and possessing more bridges than s.

In the polygonal representation of an externally refined split network there are no noncrossing diagonals that can be added; that is, any additional split added to the network *s* will correspond to a diagonal that intersects existing diagonals. An externally refined phylogenetic tree has non-leaf nodes that are all of degree three; this is usually referred to as a *binary*, or *bifurcated* tree.

The following definitions are from Gambette et al. [2017a] and Gambette et al. [2017b]. Those authors use the term *uprooted* phylogenetic network to refer to what Steel [2016] calls *unrooted* phylogenetic networks. That is due to the overuse of the term: split networks themselves have been called unrooted phylogenetic networks as well, as in Levy and Pachter [2011]. In addition to requiring that the non-leaf nodes have degree larger than 2, Gambette et al. [2017a] require every cycle to be of length at least four. If every edge is part of at most one cycle then the network is called *1-nested*. If every node is part of at most one cycle, the network is called *level-1*. If that is true and all the unlabeled non-leaf nodes also have degree three, then the network is called *binary level-1*. Level-1 networks, as a set, include level-0 networks, which are the phylogenetic trees. The level-1 and 1-nested networks are special versions of galled trees, (which sometimes allow cycles of length three as in Semple and Steel [2006]), and of cactus models (which allow labels for non-leaves as in Brandes and Cornelsen [2009]).

Notice that a phylogenetic tree is both a split network and a level-0 phylogenetic network. In contrast to split networks, phylogenetic networks do not have parallel sets of edges, but sets of edges are still used to represent splits. A minimal cut C of a phylogenetic network is a subset of the edges which, when removed, leaves two connected components. The edge set is minimal in the sense that no more edges are removed than is necessary for the disconnection. The split A|B displayed by such a cut is the pair of sets of leaves of the two connected components. A split A|B is consistent with a phylogenetic network if there is a minimal cut C(A|B) displaying that split. A split system s is consistent with a phylogenetic network if all its splits are consistent. The (maximal) system of all consistent splits for a phylogenetic network N is called  $\Sigma(N)$ . In Fig. 3 we show a binary level-1 network N, and the associated maximal split network  $\Sigma(N)$ . Multiple different phylogenetic networks can map to the same split system under the mapping  $\Sigma$ .

There is an even closer relationship between level-1 (and thus 1-nested) networks and circular split networks. In Gambette et al. [2017a] it is shown that a split network s is circular if and only if there exists an unrooted level-1 network N such that  $s \subset \Sigma(N)$ . For instance the split network s in Fig. 2 has splits a subset of those in  $\Sigma(N)$  seen in Fig. 3.

If s is a circular split system then there is a simple way to associate to s a specific 1-nested phylogenetic network denoted as L(s).



FIGURE 3. A level-1 phylogenetic network N with its associated maximal circular split system  $\Sigma(N)$ , shown both as a network and polygonal representation. Here N is the image L(s) for s the split network in Fig. 2.

**Definition 3.2.** Construct this network L(s) as follows: begin with a split network diagram of s and consider the diagram as a planar drawing of its underlying planar graph, with leaves on the exterior. Then 1) delete all the edges that are not adjacent to the exterior of that graph, and 2) smooth away any resulting degree-2 nodes.

See s in Fig. 2 and L(s) in Fig. 3. We see that L(s) displays all the splits of s, and has the same bridges as s. Explicitly, any bridge of s is displayed by a bridge in the image, and any other split in s by a pair of edges in a cycle of L(s). Refinement is seen easily in the pictures of split networks, by collapsing parallel sets of edges or by removing diagonals from the polygon. The function L preserves refinement in phylogenetic networks. See Fig. 4 for examples. If s is an externally refined split network then L(s) is a binary level-1 phylogenetic network. Note



FIGURE 4. Here r is a split network that refines t, which in turn refines t', which refines t''. Also  $\Sigma(L(t'')) \subset \Sigma(L(t')) \subset \Sigma(L(t)) \subset \Sigma(L(r))$ .

that in Gambette et al. [2017a] the authors define a similar function called N. Their function is equivalent to ours, if the definition in Gambette et al. [2017a] is modified so as not to depend on k-marguerites.

## 4. Polytopes

A circular ordering c of [n] is *consistent* with a circular split system s if a planar network of s may be drawn, such that the leaves lie on the exterior in the order given by c. Similarly a circular ordering c is *consistent* with a level-1 phylogenetic network N if a planar diagram of N may be drawn, such that the leaves lie on the exterior in the order given by c. Thus a circular ordering c is consistent with s if and only if c is consistent with L(s).

We define new families of polytopes by assigning vectors to each externally refined circular split network s, and thus to each binary level-1 phylogenetic network.

**Definition 4.1.** The vector  $\mathbf{x}(s)$  is defined to have lexicographically ordered components  $x_{ij}(s)$  for each unordered pair of distinct leaves  $i, j \in [n]$  as follows:

$$x_{ij}(s) = \begin{cases} 2^{k-b_{ij}} & \text{if there exists } c \text{ consistent with } s; \text{ with } i, j \text{ adjacent in } c, \\ 0 & \text{otherwise.} \end{cases}$$

where k is the number of bridges in s and  $b_{ij}$  is the number of bridges crossed on any path from i to j.

The formula for  $\mathbf{x}(N)$  works just as well when N is a binary level-1 network. Indeed, we clearly have for any externally refined circular split network that

$$\mathbf{x}(L(s)) = \mathbf{x}(s).$$

The vector is determined entirely by the number and placement of the bridges. Thus two split systems with the same associated binary level-1 network will have the same vector  $\mathbf{x}$ . Therefore we will often use  $\mathbf{x}(L(s))$  interchangeably with  $\mathbf{x}(s)$ , for s any preimage of L(s).

**Definition 4.2.** The convex hull of all the vectors  $\mathbf{x}(s)$  for s any externally refined circular network with n leaves and k nontrivial bridges is the level-1 network polytope BME(n, k). (The set of vectors is the same as the collection  $\mathbf{x}(N)$  for N any binary level-1 network with n leaves and k nontrivial bridges.)

Before proving theorems about these newly discovered level-1 network polytopes, we relate them to well-known examples. We begin with a review of the Balanced Minimal Evolution Polytopes BME(n), and the Symmetric Travelling Salesman polytopes STSP(n). The BME polytopes were first studied in Eickmeyer et al. [2008]. We have found a simple description of the vertices as follows:

**Definition 4.3.** For each given binary phylogenetic tree t with n leaves and k = n - 3 (non-trivial) bridges, the vertex vector  $\mathbf{x}(t)$  has  $\binom{n}{2}$  components

$$x_{ij}(t) = 2^{k - b_{ij}}$$

where  $b_{ij}$  is the number of nontrivial bridges on the tree from leaf i to a different leaf j.

The convex hull of all the (2n-5)!! vertex vectors (for all binary trees t with n leaves), is the polytope BME(n), of dimension  $\binom{n}{2} - n$ .



FIGURE 5. Calculating some vectors  $\mathbf{x}(s)$ .

Note that in Forcey et al. [2016] the formula is given, equivalently, as  $x_{ij}(t) = 2^{n-2-l_{ij}}$  where  $l_{ij}$  is the number of internal nodes along the path from *i* to *j*.

4.1. Facets of BME(n). The (lower-dimensional) *clade faces* of BME (n) were described in Haws et al. [2011]. Recently we have discovered large collections of (maximum dimensional) facets for all n, in Forcey et al. [2016] and Forcey et al. [2017]. In the following list we review our new facets, and show their statistics in Table 1.

- (1) Any split A|B of [n] with m = |A|, p = |B| and both parts  $m, p \ge 3$  corresponds to a facet of BME(n), with vertices all the trees displaying that split.
- (2) A cherry is a clade with two leaves. For each intersecting pair of cherries  $\{a, b\}, \{b, c\}$ , there is a facet of BME(n) whose vertices correspond to trees having either cherry.
- (3) For each pair of leaves  $\{i, j\}$ , the caterpillar trees with that pair fixed at opposite ends constitute the vertices of a facet. These bound BME(n) from below.

4.2. **STSP.** Next we recall the travelling salesman polytopes. We consider symmetric tours (circular orderings of the n taxa). These can be pictured as placing the numbers in order on a circle in which the orientation is not specified—reading around the circle clockwise or counterclockwise gives the same circular ordering.

**Definition 4.4.** For each circular ordering c on the set [n], the incidence vector  $\mathbf{x}(c)$  has  $\binom{n}{2}$  components. The components are

$$x_{ij}(c) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are adjacent in } c \\ 0 & \text{if not.} \end{cases}$$

The Symmetric Traveling Salesman Polytope, denoted as STSP(n), is the convex hull of these  $\frac{(n-1)!}{2}$  vertex vectors. It has dimension  $\binom{n}{2} - n$ .

The two best-known sets of facets of the STSP are the subtour-elimination facets and the lower bound facets. The latter are given by the inequalities  $x_{ij} \geq 0$ . Subtour-elimination facets correspond to any nontrivial split A|B of [n]. The circular orderings which make up the vertices of such a facet are those which contain the elements of A as a contiguous list, and thus B likewise. There can be only two connecting edges between the parts of the splits. Requiring that tours be Hamiltonian means that there must be at least two such connecting edges in any tour, which eliminates the possibility of a subtour through one or the other. Thus for a given split, a facet defining inequality is

$$\sum_{\in A, j \in B} x_{ij} \ge 2$$

We show an alternate inequality in Theorem 13. (Recall that the STSP is of smaller dimension than its ambient space, allowing choices of inequality for all its faces.)

Clearly the three definitions 4.1, 4.3, and 4.4 of  $\mathbf{x}(s)$  agree when their input structures overlap. Circular orderings (seen as unicyclic level-1 networks) have no bridges, so when leaves are adjacent the exponent becomes 0. Trees (seen as networks) allow any two leaves to be adjacent, so the components of  $\mathbf{x}$  are all nonzero in that case. Thus we see that restricting BME(n,k) to the phylogenetic trees, where k = n - 3, recovers the polytopes BME(n). Restricting BME(n,k) to the fully reticulated networks, where k = 0, recovers STSP(n). Next we characterize the vector  $\mathbf{x}$  from a combinatorial viewpoint.

**Theorem 4.5.** For any externally refined circular split network s with k bridges, we have

$$\mathbf{x}(s) = \sum_{\substack{c \text{ consistent}\\ with s}} \mathbf{x}(c)$$

where the sum is over the exactly  $2^k$  circular orderings c consistent with s. Equivalently the component  $x_{ij}(s)$  is the number of circular orderings consistent with that network for which i and j are adjacent.

facets	facet inequalities	number of	number of	
of $BME(n)$		facets	vertices	
		in polytope	in each facet	
Caterpillar	$x_{ab} \ge 1$	$\binom{n}{2}$	(n-2)!	
intersecting-	$x_{ab} + x_{bc} - x_{ac} \le 2^{n-3}$	$\binom{n}{2}(n-2)$	2(2n-7)!!	
cherry				
$(m,3)$ -split, $m \ge 3$	$x_{ab} + x_{bc} + x_{ac} \le 2^{n-2}$	$\binom{n}{3}$	3(2n-9)!!	
(m, p)-split $A B; m, p \ge 3$	$\sum x_{ij} \le (m-1)2^{n-3}$	$2^{n-1} - \binom{n}{2} - n - 1$	(2m-3)!!(2p-3)!!	
	$i,j\in A$			
$T_{1} = -1$ I/		D (D ( ) )	$(\mathbf{D})$	

TABLE 1. Known facets for the BME polytopes, BME(n) = BME(n, n - 3). The third is a special case of the fourth. The inequalities are given for any  $a, b, c, \dots \in [n]$ .



FIGURE 6. A facet (a) in BME(5,0) = STSP(5), a facet (b) in BME(5,1), and a face (c) in BME(5,2) = BME(5). Summing either of the horizontal or vertical pairs of vectors shown in (a) gives a corresponding vector shown in (b). Summing all four vectors in (a) gives the vector shown in (c). Summing all four vectors in (b) gives twice the vector in (c).

*Proof.* We show this equality by considering the sums of respective components. The components of  $\mathbf{x}(c)$  are always 1 or 0. Note that the only way to alter a circular ordering c which is consistent with the externally refined s, to another such consistent c', is to choose a nontrivial bridge and twist the graph around that bridge. That is, we redraw the graph with one side of the bridge reflected vertically; we call this twisting the bridge. The fact that there are exactly  $2^k$  circular orderings contributing over all is seen by independently twisting all k nontrivial bridges–each bridge contributes two options. If c is a circular ordering consistent with s, and further c has i and j adjacent, then the bridges between i and j cannot be twisted without losing this adjacency. However, the other bridges  $(k - b_{ij})$  of them) may be independently

twisted while preserving the adjacency of i and j. Thus each of those latter bridges contributes a factor of 2 to the total count of the consistent circular orderings with i and j adjacent: upon summing we thus achieve the defined value of  $x_{ij} = 2^{k-b_{ij}}$ . Examples are seen in Fig. 6.  $\Box$ 

**Corollary 4.6.** We can infer that for any network s the sum of all the components of  $\mathbf{x}(s)$  obeys  $\sum x_{ij} = n2^k$ .

*Proof.* The total follows from the fact that the sum of components for any tour on [n] is n, the number of edges in the tour.

Closely related is the following *twisting* lemma, useful in the next Section:

**Lemma 4.7.** Let s have at least one bridge b. Then there exist two ways to add a single split to s, to achieve by those additions two split networks s' and s'' each with one less bridge than s, and such that:

$$\mathbf{x}(s) = \mathbf{x}(s') + \mathbf{x}(s'').$$

*Proof.* Notice that in the polygonal picture of a split network s with k+1 bridges, for any given bridge b there is always a way to add a split, (as a new diagonal), which crosses that bridge b but no other existing bridge. This is true since (even if the maximum number of bridges is present), the bridge b can be seen as one of the two diagonals of a quadrilateral which is nested inside the polygon. See Fig. 7. The other diagonal of that quadrilateral is, of course, missing—so it is always available to become the new split which crosses b. That new diagonal crosses no other bridge since a quadrilateral can have only two diagonals. We use this new diagonal to



FIGURE 7. Example of the three split networks described in Lemma 4.7.

reduce the number of bridges by one. We create s' and s'' as the two possible results of adding that new split which crosses b, with and without a twist around b. From Theorem 4.5 we have that  $\mathbf{x}(s')$  is the sum of the vectors  $\mathbf{x}(c)$  for all the circular orderings consistent with s', and likewise for s''. The key here is that the circular orderings consistent with s are partitioned into those consistent with s' and those consistent with s''.

Thus the vector  $\mathbf{x}(s)$  is the sum of the two vectors  $\mathbf{x}(s')$  and  $\mathbf{x}(s'')$ .

#### CASSANDRA DURELL AND STEFAN FORCEY

#### 5. Metrics and faces

For our proofs about the structure of BME(n, k), we need weighted networks. A weighting of a split system s is a function  $w : s \to \mathbb{R}_{\geq 0}$ . In practice each split is assigned a positive weight, since when splits are assigned weight = 0 this system can be equated to the system minus those splits. Given such a weighted split system we can derive a metric  $\mathbf{d}_s$  on [n], where

$$\mathbf{d}_s(i,j) = \sum_{i \in A, j \in B} w(A|B)$$

where the sum is over all splits of s with i in one part and j in the other. The metric is often referred to as the distance vector  $\mathbf{d}_s$ . We can also derive a weighting on the edges of the 1-nested network L(s), extending our function L to weighted networks. Here the weight function is from the edges of L(s) to positive real numbers, and given by

$$w_s(e) = \sum_{e \in C(A|B)} w(A|B)$$

where the sum is over the splits of s which are represented by a minimal cut containing e. Clearly the sum of weights on a shortest path in L(s) from i to j equals  $\mathbf{d}_s(i, j)$ . See Fig. 8 for an example.



FIGURE 8. A weighted split network and its associated weighted level-1 network. Here the trivial splits (edges adjacent to leaves) are all given weight 1 for simplicity. The right-most picture is the same weighted network, with a tour shown by arrows. The tour length is twice the total weight, 72.6.

We define the total weight of the network to be the sum of all the weights:

$$W(s) = \sum_{A|B \in s} w(A|B).$$

Now we can still calculate  $\mathbf{x}(s)$  for a weighted network, the vector does not depend on the weights. Instead we are interested in the dot product:

**Theorem 5.1.** For an externally refined weighted split network s, and thus for the level-1 binary L(s), the dot product of our network vector with the distance vector gives a multiple of the sum of the weights:

$$\mathbf{x}(s) \cdot \mathbf{d}_s = 2^{k+1} W(s).$$

*Proof.* Since s is circular, given a network diagram for s, the distance  $\mathbf{d}_s(i, j)$  for adjacent i and j can be found by adding the weights on edges between i and j on the exterior of the

diagram. (That is also the same as the weights on the edges of L(s) between them). For a circular ordering c that is consistent with the split system s we see that  $\mathbf{x}(c) \cdot \mathbf{d}_s$  is equal to summing the distances between adjacent pairs of taxa, and that this sum includes the weight of each split exactly twice: it totals to 2W(s). See Fig. 8 for an example. Since there are  $2^k$  such circular orderings c (whose vectors sum to  $\mathbf{x}(s)$ , as shown in Theorem 4.5) we have that

$$\mathbf{x}(s) \cdot \mathbf{d}_s = \sum_c \mathbf{x}(c) \cdot \mathbf{d}_s = 2(2^k)W(s) = 2^{k+1}W(s).$$

**Theorem 5.2.** If s' is any other externally refined split network on [n] with k bridges, with binary level-1 network  $L(s') \neq L(s)$ , then  $\mathbf{x}(s') \cdot \mathbf{d}_s > \mathbf{x}(s) \cdot \mathbf{d}_s$ .

*Proof.* As just shown in the proof of Theorem 5.1, for any circular ordering c consistent with s, we have that  $\mathbf{x}(c) \cdot \mathbf{d}_s = 2W(s)$ , which is the minimal length of a tour of [n] given the distances in  $\mathbf{d}_s$ . Indeed it is a minimum length tour for the given metric  $\mathbf{d}_s$ , since visiting any of the taxa out of this order would involve retracing some part of the path between them. This is most clearly seen by considering the level-1 network L(s) with weighted edges, where the distance between each pair is the minimum length path. See Fig. 8. The circular ordering minimizing a tour length using those minimum length paths must be consistent with the network, or else some portion of some path between leaves will be traversed twice, increasing the length of the tour. Thus since the dot product  $\mathbf{x}(c) \cdot \mathbf{d}_s$  is minimized for each consistent c, and  $\mathbf{x}(s) \cdot \mathbf{d}_s$  is the sum of those products as seen in Theorem 5.1, then the latter sum is minimized for the network L(s), and forming s' by exchanging any c for a non-consistent alternative c' would increase that term in the sum.)

The previous two results do have a geometric interpretation, which is:

**Corollary 5.3.** The vertices of BME(n,k) are the vectors  $\mathbf{x}(s)$  corresponding to the distinct binary level-1 networks L(s). That is for each externally refined circular network s, with n leaves and k bridges, we get a vertex of the polytope (but it is determined only by L(s).)

Remark: Levy and Pachter [2011], generalizing the work of Semple and Steel [2004], define a coefficient  $\eta_s$  which takes values the components of our vector  $\mathbf{x}(s)$ . For an arbitrary distance vector  $\mathbf{d}$ , Levy and Pachter call the dot product  $\mathbf{x}(s) \cdot \mathbf{d}$  the *length* of  $\mathbf{d}$  with respect to s. They point out that neighbor-net is a greedy algorithm for minimizing this quantity. Our results show how to minimize this length via linear programming. We also see as a consequence that that length is minimized precisely by a binary level-1 network, (or several if the number of bridges is larger than k.)

The question is raised: if the vertices of BME(n, k) correspond to binary level-1 networks, but minimize a length that is a function of the weighted split network, then what role is left for an arbitrarily weighted 1-nested network? In fact, any weighted 1-nested network N has the following property:

**Theorem 5.4.** If  $\mathbf{d}_N$  is the metric on the leaves of N defined by  $\mathbf{d}_N(i, j)$  equal to the least sum of weights along a path between leaves i and j, then there is a unique circular weighted split system  $s = S_w(N)$  which has the same associated metric. That is,  $\mathbf{d}_N = \mathbf{d}_s$ . *Proof.* First we show that  $\mathbf{d}_N$  obeys the Kalmanson condition: there exists a circular ordering of [n] such that for all  $1 \leq i < j < k < l \leq n$  in that ordering,

$$\max\{\mathbf{d}_N(i,j) + \mathbf{d}_N(k,l), \mathbf{d}_N(j,k) + \mathbf{d}_N(i,l)\} \le \mathbf{d}_N(i,k) + \mathbf{d}_N(j,l).$$

The circular ordering that meets our specifications is just any choice of one of the circular orderings consistent with N. The two paths involved on the right hand side of the condition intersect each other while crossing. Then since the leaves are on the exterior, the four paths involved on the left hand side of the condition are each bounded above in length by a path made by following first one intersecting path and then the other, (switching at their shared portion). Two paths in a sum on the left hand side of the condition can at most use exactly all of both the crossing paths, so that the inequality is guaranteed. For example, in the following network N we choose to look at the four taxa 1,2,8,7 in that order. The intersecting paths go from 1 to 8 and 2 to 7, with lengths of 12 and 11 respectively. The graph edges used by the crossing paths are highlighted. Notice that the shortest path from 1 to 2, length 9, is bounded above by the path from 1 to 2 using highlighted edges. The other three paths, from 2 to 8, from 8 to 7, and from 7 to 1, all actually use highlighted edges borrowed from the crossing paths. In this case the inequality becomes max $\{19, 17\} \leq 23$ .



It is well known that for any Kalmanson metric  $\mathbf{d}_N$  there exists a unique weighted split system s whose weighting gives that metric:  $\mathbf{d}_N = \mathbf{d}_s$ . See Chapter 6 in Steel [2016]. To actually calculate this split system, the algorithm neighbor-net can be used; since it is guaranteed to return the unique answer for any Kalmanson metric, as shown in Bryant et al. [2007].

Now we show that for any 1-nested network, we get faces of our polytopes. In fact we get multiple faces from each network: one in each of the polytopes BME(n, k) for which that network has more bridges than k. Precisely:

**Theorem 5.5.** Every n leaved 1-nested network N with m bridges corresponds to a face  $F_k(N)$  of each BME(n,k) polytope for  $0 \le k \le m$ . That face has vertices all the binary level-1 k-bridge networks N' whose splits refine those of N, that is such that  $\Sigma(N) \subset \Sigma(N')$ .

*Proof.* Without loss of generality we choose a split network s which has the exterior form of N, that is L(s) = N. Then for any N' as described in the theorem, we then have  $s \subseteq \Sigma(N')$ . Let  $s' = \Sigma(N')$ ; since N' is binary, s' is externally refined and L(s') = N'.

Next let s be weighted by assigning the value of 1 to each split. (Then W(s) is the total number of splits in s). Let  $\mathbf{d}_s$  be the distance vector derived from that weighting, so that the i, j component of  $\mathbf{d}_s$  is the number of splits between those leaves on s. We claim that the

PHYLOGENETIC NETWORKS AND THEIR BME POLYTOPES

 $(\mathbf{1})$ 

(2)

N

(2)

S

4

3



(4)

2

3

*s*"

FIGURE 9. Demonstration of Theorem 5.5. Split networks on the bottom row have all splits weight 1, except for the dashed splits which have weight 0.

(2)

dot product  $\mathbf{x}(s') \cdot \mathbf{d}_s$  is minimized simultaneously at each of the externally refined k-bridge networks s' which refine s. In fact we have that the following inequality holds:

$$\mathbf{x}(s') \cdot \mathbf{d}_s \ge 2^{k+1} W(s)$$

for all k-bridge externally refined networks s', and is an equality precisely when s' refines s. Therefore

$$\mathbf{x}(N') \cdot \mathbf{d}_s \ge 2^{k+1} W(s)$$

for all k-bridge binary level-1 networks N', and is an equality precisely when N' refines N.

The reason for our claim is that when s' refines s, we have that  $\mathbf{d}_s$  is equivalent to a distance vector  $\mathbf{d}_{s'}$  derived from a weighting of s', where the splits are given weight = 1 if they are also in s, and weight = 0 if not. Also then W(s) = W(s'). Thus in that case,  $\mathbf{x}(s') \cdot \mathbf{d}_s = \mathbf{x}(s') \cdot \mathbf{d}_{s'} = 2^{k+1}W(s') = 2^{k+1}W(s)$ . The second equality in that string is by Theorem 5.1, which holds even for weights equal to zero. Notice that this could also be described as summing the tours on s for consistent circular orders. For example, see Fig. 9.

Now, for any binary level-1 R not refining N, meaning that N = L(s) displays a split A|B not displayed by R, we see that R will be consistent with a circular order c that is not consistent with L(s). That is because when A|B is displayed then every consistent circular order has both parts A and B contiguous. If A|B is not displayed then there will be a consistent circular order c in which the set of leaves A is not found contiguously.

Similar to the argument for Theorem 5.2, R having a circular order c not consistent with N = L(s) implies that  $\mathbf{x}(s') \cdot \mathbf{d}_s > 2^{k+1}W(s)$ . That is because the dot product is equivalent to summing all the tour lengths for tours of s using circular orders consistent with R, and for c this results in a tour that uses some splits of s more than twice. For example, see below.

Examples: In Fig. 9, for n = 8 and k=1, we show two networks N', N'' which are both vertices of the face determined by network N. The relevant vectors are:

$$\begin{aligned} \mathbf{d}_s &= \mathbf{d}_{s'} = \ \mathbf{d}_{s''} = (3, 5, 5, 6, 5, 4, 3, 4, 6, 7, 6, 5, 4, 4, 5, 4, 3, 4, 3, 4, 3, 4, 3, 4, 5, 3, 4, 3) \\ \mathbf{x}(N') &= (2, 0, 0, 0, 0, 0, 2, 2, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 2, 0, 1, 0, 2, 0, 0, 1, 0, 2) \\ \mathbf{x}(N'') &= (2, 0, 0, 0, 0, 0, 2, 2, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 2, 0, 0, 1, 2, 0, 0, 2, 0, 1) \end{aligned}$$

And thus:  $\mathbf{x}(N') \cdot \mathbf{d}_s = \mathbf{x}(N'') \cdot \mathbf{d}_s = 52 = 4(13) = 2^2 W(s)$ . In contrast for R as follows:



with  $\mathbf{x}(R) = (2, 0, 1, 0, 0, 0, 1, 2, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 2, 0, 0, 0, 2, 0, 0, 2, 0, 2)$ , we see that N displays the split  $\{1, 2, 8\}|\{3, 4, 5, 6, 7\}$  but that R does not. Also, R is consistent with the circular order c = (8, 7, 6, 5, 4, 1, 2, 3, 8), but N and thus s are not. The tour that c describes on s uses the split  $\{1, 2, 8\}|\{3, 4, 5, 6, 7\}$  four times (so adds the value of that split twice more than in a minimal tour.) Indeed  $\mathbf{x}(R) \cdot \mathbf{d}_s = 54$ .

For some smaller examples, in Fig. 10 there are three faces shown. The first two (a) and (b) are in BME(5,0), and the third (c) is in BME(5,1). They are pictured in context in Fig. 6. Here we include the vector  $\mathbf{d}_s$  for each of the three. In (a) we have W(s) = 7 and the vertices of BME(5,0) obey  $\mathbf{x}(s') \cdot \mathbf{d}_s \ge 14$ . In (b) we have W(s) = 8 and the vertices of BME(5,0) obey  $\mathbf{x}(s') \cdot \mathbf{d}_s \ge 14$ . In (b) we have W(s) = 8 and the vertices of BME(5,0) obey  $\mathbf{x}(s') \cdot \mathbf{d}_s \ge 16$ . In (c) we have W(s) = 7 and the vertices of BME(5,0) obey  $\mathbf{x}(s') \cdot \mathbf{d}_s \ge 28$ . In the figure we show vertices that obey the inequality sharply. To see the strict inequality take dot products with any other vertex vector from the respective polytope.

**Remark 5.6.** For a 1-nested network s with n leaves and k bridges, embedded in each polytope BME(n, j) for  $k > j \ge 0$  is a collection of faces corresponding to networks which refine our given network s. Those faces link up (by sharing subfaces) to make an interesting complex, as shown by the shaded subfaces in Fig. 6. The topology of these complexes is an interesting open question.

Two polytopes are called *nested* when one lies inside the other, and the smaller polytope has all its vertices on the surface of the larger. It turns out that all the level-1 network polytopes are, up to scaling, nested sequentially inside each other. Even more, they are all at the same time nested inside the Symmetric Travelling Salesman polytope. This is pictured in Fig. 11. Precisely:

**Theorem 5.7.** We can scale the network polytopes so that the polytope BME(n,k) is nested inside BME(n, k - 1) for  $0 < k \le n - 3$ . Furthermore, we can simultaneously scale all the BME(n,k) polytopes so that they are all nested inside of STSP(n) (with vertices at facial barycenters) and each BME(n,k) is nested inside BME(n,j) for j < k.

*Proof.* Consider  $0 < j < k \leq n-3$ . We show that the vertices of the scaled polytope  $(2^{n-3-k})BME(n,k)$  lie on faces of the scaled polytope  $(2^{n-3-j})BME(n,j)$  which in turn lie on the faces of the scaled polytope  $(2^{n-3})STSP(n)$ .

PHYLOGENETIC NETWORKS AND THEIR BME POLYTOPES



FIGURE 10. Three subfaces, shown with their representative networks, normal vectors, and vertices. All three are from facets shown in Fig. 6. Edge (b) is a subface of tetrahedron (a), from BME(5,0). Quadrilateral (c) is from BME(5,1). Notice that the normal vectors are used more than once.

First, by Lemma 4.7, any vertex  $\mathbf{x}(s)$  of BME(n, k) is the sum of two vertices of BME(n, k - 1), found by adding a single split to s in two ways. Therefore if BME(n, k - 1) is first scaled by 2, the sum of those two vertices will also be scaled by 2. Thus the (unscaled) vertex  $\mathbf{x}(s)$  of BME(n, k) is the midpoint of those two scaled vertices: after adding, divide by 2. By convexity, this vertex is thus on a face of (2)BME(n, k - 1).

Secondly, any vertex  $\mathbf{x}(s)$  of BME(n, k) is on the surface of  $(2^k)$ STSP(n). To see this, recall from Theorem 4.5 that  $\mathbf{x}(s)$  is the sum of all  $\mathbf{x}(c)$  for c a circular ordering consistent with s. Note that the consistent circular orderings are precisely the vertices of the face of STSP(n)corresponding to the binary level-1 network s. There are  $2^k$  of them. Thus by first multiplying each by  $2^k$  and then dividing their sum by  $2^k$  we see that the vertex  $\mathbf{x}(s)$  is at the barycenter of the corresponding face of  $(2^k)$ STSP(n). Together these facts show that while each BME(n, k) is nested in a scaled version of BME(n, k-1), all are simultaneously nested in a scaled version of STSP(n). By scaling STSP(n) by  $2^{n-3}$  and each BME(n, k) by  $2^{n-3-k}$  we can see them all nested simultaneously and sequentially. For example see Fig. 11.



FIGURE 11. The facet of STSP(5) pictured is the same one as in Fig. 6(a), but scaled by a factor of 4. The corners of the shaded quadrilaterals are located at barycenters of faces (edges) of STSP(5). On the right we show just the convex hull of those barycenters, and then a Schlegel diagram obtained by rotating it. The vertices of this convex hull are those of the split facet of BME(5,1) shown in Fig. 6(b) but scaled by 2.

**Theorem 5.8.** The dimension of BME(n,k) is  $\binom{n}{2} - n$ . The dimension reducing equalities are as follows: For each leaf j = 1, ..., n the vertices  $\mathbf{x}(s)$  satisfy

$$\sum_{\in [n] - \{j\}} x_{ij} = 2^{k+1}$$

where k is the number of bridges (non-crossing diagonals) in the diagram.

*Proof.* The dimension results from the nesting property, Theorem 5.7. Since the STSP(n) and BME(n) polytopes are both of dimension  $\binom{n}{2} - n$ , the BME(n,k) polytopes nested between them must also have that same dimension.

The equalities generalize the *Kraft* equalities for phylogenetic trees (k = n - 3), and the *degree* equations for the STSP(n) for all n, (k = 0.) In fact we can use the latter as base cases for the proof by induction on the number of bridges k. Assuming that the equality holds for all s with n leaves and k bridges, we show that it holds for s with n leaves and k+1 bridges. Recall

from Lemma 4.7 that there are networks s' and s'' with k bridges and with  $\mathbf{x}(s) = \mathbf{x}(s') + \mathbf{x}(s'')$ . By our inductive assumption, both smaller vectors obey the required formula, both with the sum of components  $= 2^{k+1}$ . Thus after adding them together the resulting sum is  $= 2^{k+2}$ .  $\Box$ 

5.1. Counting binary level-1 networks. The associahedra K(n) are a sequence of polytopes, one in each dimension. Their faces correspond to sets of non-crossing diagonals in the *n*-sided polygon. In this case the polygon is fixed in the plane, with no rotations or flips allowed. However, we can still use the number of associahedron faces of each dimension to enumerate the vertices of the BME(n, k) polytopes.

**Theorem 5.9.** The number of vertices of BME(n, k), and thus the number of binary level-1 networks with n leaves and k non-trivial bridges, with  $0 \le k \le n-3$  is:

$$v(n,k) = T(n,k) \frac{(n-1)!}{2^{k+1}}$$

where T(n,k) gives the components of the face vector of the associahedron K(n).

Here, as seen in entry A033282 of Sloane [2018],

$$T(n,k) = \left(\frac{1}{k+1}\right) \binom{n-3}{k} \binom{n+k-1}{k}$$

which allows the simpler count:

$$v(n,k) = \binom{n-3}{k} \frac{(n+k-1)!}{(2k+2)!!}.$$

Table 2 shows the number of vertices of BME(n, k), (the number of binary level-1 networks with n leaves and k bridges) for small values of n, k. Note that the cases k = 0 and k = n - 3 count the circular orderings and phylogenetic trees of length n, respectively.

Proof. Vertices in the polytope BME(n, k) correspond to the binary level-1 networks with n leaves and k nontrivial bridges. The bridges in a level-1 network L = L(s) are the same as the bridges in any preimage split network s. We construct a binary level-1 network N as follows. We start with the n-sided polygon, label one side as 1, and then label the remaining sides in any order. The number of circular orderings is (n-1)!/2. Then, independently, we choose non-crossing diagonals for the k bridges, counted by T(n, k). Crucially, the side labeled 1 can be though of as the rooted edge so that each of the subdivisions of the polygon counted by T(n, k) is actually a distinct choice. Since we are only counting up to twists around each bridge , we need to divide by  $2^k$ . That completes the counting: next we construct the final binary level-1 network N by making a graph cycle out of each region in the subdivided polygon (except the triangular regions, which become tree-like degree-3 nodes), and attaching bridges and leaves according to the labeled polygon edges and diagonals. Equivalently, we can see the polygonal picture of  $\Sigma(N)$  by adding crossing diagonals to each region of our subdivided polygon.

Fig. 12 shows the process of constructing a binary level-1 network using a cyclic ordering and a face of the associahedron. In Levy and Pachter [2011] the authors point out that a vertex of the associahedron, together with a cyclic ordering, corresponds to a phylogenetic tree. We see that the correct extension of that correspondence is to the binary level-1 phylogenetic networks.



FIGURE 12. Starting with a cyclic ordering of the edges of the octagon, and adding bridges (to get a face of the associahedron) to construct a binary level-1 network.

Also note that Semple and Steel use a generating function of three variables to derive a more general formula for counting galled trees in Semple and Steel [2006]. Thus their formula implies ours in the case that no cycles of length three are allowed. It is suggestive that their proof method uses Lagrange inversion of the generating function. As shown in Aguiar and Ardila [2017] Lagrange inversion of a series uses the face numbers of the associahedra. Here that fact instead allows us to enumerate the binary level-1 networks with n taxa directly, by summing our formula from k = 0 to k = n - 3. The row sums of Table 2 are 1, 6, 57, 750, 12645, 260190... which is sequence A032119 of Sloane [2018]. However it does not appear that that sequence has been used to count level-1 networks; rather it is described as counting rooted planar trees with labeled leaves, with equivalence under rotating subtrees left to right at any given branch node. The bijection from binary level-1 networks to these planar trees is straightforward, by replacing our cycles with nodes of higher degree.

ſ	n =	k =	0	1	2	3	4	5	6
ſ	3		1						
	4		3	3					
	5		12	30	15				
	6		60	270	315	105			
	7		360	2520	5040	3780	945		
	8		2520	25200	75600	94500	51975	10395	
	9		20160	272160	1134000	2079000	1871100	810810	135135

TABLE 2. Numbers of vertices for the BME(n, k) polytopes, that is, numbers of binary level-1 phylogenetic networks with n leaves and k bridges.

## 6. Facets

We have found that many of the known facets of BME(n) have analogues in BME(n, k).

**Theorem 6.1.** Any split A|B of [n] with |A| > 1, |B| > 1 corresponds to a face of BME(n, k), for all n, k with  $k \le n - 3$ . The vertices of that face are the binary level-1 networks which

display the split A|B. Furthermore, if |A| > 2, |B| > 2, the face is shown to be a facet of the polytope. The face inequality is:

$$\sum_{i,j \in A} x_{ij} \le (|A| - 1)2^k.$$

*Proof.* First we show that the collection of vertices corresponding to networks displaying a split A|B of [n] obey our linear equality, and that all other vertices obey a corresponding inequality.

The equality follows from similar logic to that in the proofs of Theorem 4.5 and Corollary 4.6. If a network displays a split A|B then so must every circular ordering of [n] consistent with that network. There are  $2^k$  such circular orderings. If a circular ordering c displays the split A|Bthen it must have the leaves of A contiguous in that circular ordering. Thus the components  $x_{ij}(c)$  with i, j both corresponding to leaves from A will contain exactly |A| - 1 entries that are equal to 1, for the edges connecting those leaves. The equality follows: for s displaying the split A|B we have  $\sum_{i,j\in A} x(s)_{ij} = (|A| - 1)2^k$ .

The strict inequality holds for any binary level-1 network which does not display the split A|B. The same reasoning as above holds, but this time there will be at least one circular ordering c which does not display the split. In that circular ordering, the components corresponding to the leaves from |A| have fewer 1's than the maximum |A| - 1, so the sum will be less.

Next we show that the face just described, denoted  $F_A(n,k)$ , is of codimension 1 when both parts of the split are larger than 2. The polytopes are all of dimension  $\binom{n}{2} - n$ , so we show the facets are of dimension 1 less. We use the fact that for any polytope, its scaling by m is of the same dimension. Equivalently, taking sets of m vectors all from the same facet of a polytope, the vector sums of m such vectors will all lie in an affine space of the same dimension as that facet. (Thus any subset of those sums of m vectors each will have a convex hull of smaller or equal dimension than the original facet.)

We have that a given split, with both parts larger than 2, corresponds to a subtourelimination facet of STSP(n) = BME(n, 0) and also to a split-facet of BME(n, n-3) = BME(n). Both are dimension  $\binom{n}{2} - n - 1$ . From Lemma 4.7, we know that each vertex  $\mathbf{x}(s)$  of the proposed split-facet  $F_A(n, k)$  is the vector sum  $\mathbf{x}(s') + \mathbf{x}(s'')$ . The two summands are both vertices of the proposed split-facet  $F_A(n, k-1)$ , since s' and s'' both display the split A|B.

Therefore the facet  $F_A(n, k)$  cannot be of greater dimension than the facet  $F_A(n, k-1)$ . Since the dimension cannot increase at any step between k = 0 and k = n - 3, and it has the same value for k = 0 and for k = n - 3, then it must remain constant for each k at  $\binom{n}{2} - n - 1$ .  $\Box$ 

For the case of splits with one part of size two, we know that these do correspond to a facet when k = 0, the STSP(n), but not when k = n - 3, in BME(n). It is an open question for which other n, k the splits of size two correspond to facets of BME(n, k). We conjecture this for all k < n - 3, but we can only report the positive result for n = 5.

Fig. 6(a) shows a subtour elimination facet of STSP(5) = BME(5,0), corresponding to the split  $s = \{\{1,2\}, \{3,4,5\}\}$ . In this case it is combinatorially equivalent to the 4D Birkhoff polytope. Split networks label subfaces of this facet. Fig. 6(b) shows the corresponding split-facet of BME(5,1), in which the vertices are nine networks with a single bridge that each refine s. Fig. 13 shows the same split-facet, with alternate labels. Fig. 6(c) shows the 2D face of BME(5) corresponding to the same split. Summing either of the horizontal or vertical pairs of

vectors shown in (a) gives the four vectors shown in (b). Summing all four vectors in (a) gives the vector shown in (c).

Next we look at the existence of lower bound faces, and conjecture that they are in fact facets as well.

**Theorem 6.2.** For each pair  $i, j \in [n]$  we get a face of BME(n, k) for all n, k with  $k \leq n-3$ . For k = n-3 these are the caterpillar facets, described above in Table 1. For  $k \leq n-4$  these faces contain the vertices associated to networks with no consistent circular orderings such that i, j are adjacent. The face inequality for each of these latter is  $x_{ij} \geq 0$ .

*Proof.* For  $k \leq n-4$  the equality  $x_{ij}(s) = 0$  holds by definition for networks s with no consistent circular orderings such that i, j are adjacent. For any network s which possesses a consistent circular ordering with i, j adjacent, we see the component  $x_{ij}(s) > 0$ . Recall from the introduction that any such inequality defines a face of the polytope.

We would like to know which of these lower bound faces are facets. They are facets for the case of k = 0, the Symmetric Travelling Salesman polytopes, and for k = n - 3, the Balanced Minimum Evolution polytopes. For the case of k = n - 3 however the caterpillar facets have vertices that obey  $x_{ij} \ge 1$ . In fact none of their vector components are zero, so they are not the sum of a pair of lower bound face vertices from a given lower bound face. However the lower the lower bound face is a facet for n = 5, k = 1, (pictured in the next section, Fig. 14), and we conjecture this is true for all lower bound faces.

## 7. BME(5,1)

We have investigated more fully the case of BME(5,1), by using polymake, Gawrilow and Joswig [2000], to find all the facets and then observing the patterns they obey. There are 62 facets altogether, of four different types. The (5,1) networks are especially simple: each has the same underlying graph, with a cherry clade attached to a length-4 cycle with one central leaf across from the cherry.

# **Theorem 7.1.** In BME(5,1) there are $\binom{5}{2} = 10$ split facets, each of which has nine vertices.

*Proof.* These are predicted to be faces by Theorem 5.8, but it is a surprise that they are indeed facets since each nontrivial split of [5] has a part of size 2. We check that they are facets by inspection in polymake. Let the split be  $\{a, b\}|\{c, d, e\}$ . Each split facet in BME(5, 1) has nine vertices: These are the networks formed by 3 ways to put  $\{a, b\}$  on the cherry with a choice of the other 3 on the central leaf, 3 ways to put a on the central leaf, and 3 ways to put b on the central leaf.

From polymake, we find that these lower bound facets are each a product of two triangles. For an example see the split facet for the split  $\{1,2\}|\{3,4,5\}$  pictured in Fig. 13, where the inequality is  $x_{1,2} \leq 2$ . Compare to Fig. 6 where the same facet is shown labeled by polygons.

**Theorem 7.2.** In BME(5,1) there are 10 lower bound facets, one for each component of  $\mathbf{x}_{ij}(s)$ . Each has nine vertices.

*Proof.* These are predicted to be faces by Theorem 6.2, but here again we see they are indeed facets via polymake. Each lower bound facet in BME(5, 1) has nine vertices: three ways to put *i* on the central leaf and *j* on the cherry with one of the other three taxa, three ways to



FIGURE 13. A split facet in BME(5,1). These vertices obey  $x_{1,2} = 2$ 

put j on the central leaf, and three ways to put i, j non-adjacent on the 4-cycle with one of the other three on the central leaf.

From polymake, we find that these lower bound facets are each a product of two triangles. For an example see the lower bound facet for the inequality  $x_{1,2} \ge 0$  pictured in Fig. 14.

**Theorem 7.3.** In BME(5,1) there are 30 facets which we call the excluded node facets. They have 8 vertices each, which obey (sharply) the facet inequality:

 $x_{ab} + x_{cd} - x_{ac} - x_{bd} \le 3$ 

where a, b, c, d, are four taxa in cyclic order.

*Proof.* These facets correspond to choosing 4 of the 5 taxa, excluding one taxon. Then the four are given a cyclic order, and then that cyclic order is split into two contiguous pairs. These



FIGURE 14. A lower bound facet in BME(5,1). These vertices obey  $x_{1,2} = 0$ .

choices are independent, giving 5(3)2 = 30 facets. Each facet has 8 networks as its vertices, found by choosing to place the excluded taxon on either the cherry or the 4-cycle (but not on the central leaf); followed by placing the chosen four taxa on the remaining leaves in their cyclic order, but not allowing either contiguous pair to be separated by the excluded taxon. By

inspection, each excluded node facet obeys the following inequality, where the circular ordering is a, b; c, d with the first and second pairs contiguous:  $x_{ab} + x_{cd} - x_{ac} - x_{bd} \leq 3$ .

Each excluded node facet is a 4D prism: the interval crossed with a tetrahedron. For an example see the excluded node facet for the cycle 1, 2; 5, 3 pictured in Fig. 15.



FIGURE 15. An excluded node facet in BME(5,1). These vertices obey  $x_{1,2} + x_{5,3} - x_{1,5} - x_{2,3} = 3$ .

**Theorem 7.4.** In BME(5,1) there are 12 facets which we call the cyclic order facets. They have 5 vertices each, which obey (sharply) the facet inequality:

$$x_{ab} + x_{bc} + x_{cd} + x_{df} + x_{af} \le 8$$

where (a, b, c, d, f) is a cyclic order on the taxa.

*Proof.* By inspection, there is a cyclic order facet for each of the 12 circular orderings. Each network (vertex) in the facet represents the same circular ordering when reading the leaves around the network in a circle (up to twists around the bridge). There are five such networks for a given circular ordering, distinct by the choice of which taxa to place on the central leaf of the 4-cycle. By inspection, the vectors of the trees in these facets adhere to the equality  $x_{ab} + x_{bc} + x_{cd} + x_{df} + x_{af} = 8$ .

For an example see the cyclic order facet for the cycle (a, b, c, d, f) pictured in Fig. 16.



FIGURE 16. A generic cyclic order facet in BME(5,1). These vertices obey  $x_{ab} + x_{bc} + x_{cd} + x_{df} + x_{af} = 8$ .

## 8. Applications and further investigation.

As mentioned above, we conjecture that certain inequalities of BME(n, k) such as the lower bounds in Theorem 6.2 and the upper bounds from a split with one part of size two in Theorem 6.1 are actually facet inequalities (except for the latter in the case of k = n - 3). However we already have the following:

**Theorem 8.1.** The split inequalities and lower bound inequalities together define a relaxation of the polytope BME(n, k) for  $n \ge 3$  and  $0 \le k \le n - 3$ .

*Proof.* This is straightforward, since by restricting to some subset of the inequalities we will be considering a convex shape that contains BME(n,k). The result is bounded above and below via the inequalities  $0 \le x_{ij} \le 2^k$ . (The lower bound is  $1 \le x_{ij}$  for k = n - 3.)

For example, the relaxation of BME(5,1) given by the lower bound and split inequalities is just the 10-dimensional hypercube. The relaxation of a polytope can be used as a starting point for algorithms that attempt to find the optimal solution among the vertices of the original polytope. Our relaxation will allow future implementations of two such algorithms: branch-and-bound, and branch-and-cut. The first uses the fact that all the coordinates of the actual solutions (vertices of BME(n, k), representing networks) are powers of 2. That allows an initial approximate solution on the relaxation to be forced closer to the correct solution by looking for coordinates in the approximate solution that are not powers of 2, and branching into subproblems by subdividing the polytope with hyperplanes perpendicular to that coordinate axis. The branch-and-bound algorithm with split facets is implemented for BME(n) = BME(n, n - 3) in Forcey et al. [2018]. The second algorithm, branch-and-cut, becomes important when for larger values of n the number of split facets grows exponentially. Rather than use all those inequalities immediately, they can be added in little-by-little when they are seen to cut off some coordinate in the approximate solution that is not a power of 2.

#### References

- M. Aguiar and F. Ardila. Hopf monoids and generalized permutahedra. arXiv:1709.07504 [math.CO], pages 1–113, 2017.
- Ulrik Brandes and Sabine Cornelsen. Phylogenetic graph models beyond trees. *Discrete Appl. Math.*, 157(10):2361–2369, 2009. ISSN 0166-218X. doi: 10.1016/j.dam.2008.06.031. URL https://doi.org/10.1016/j.dam.2008.06.031.
- David Bryant, Vincent Moulton, and Andreas Spillner. Consistency of the neighbor-net algorithm. *Algorithms for Molecular Biology*, 2(1):8, Jun 2007. ISSN 1748-7188. doi: 10.1186/1748-7188-2-8. URL https://doi.org/10.1186/1748-7188-2-8.
- Daniele Catanzaro, Martine Labbé, Raffaele Pesenti, and Juan-José Salazar-González. The balanced minimum evolution problem. *INFORMS J. Comput.*, 24(2):276–294, 2012. ISSN 1091-9856. doi: 10.1287/ijoc.1110.0455. URL http://dx.doi.org/10.1287/ijoc.1110. 0455.
- Daniele Catanzaro, Roberto Aringhieri, Marco Di Summa, and Raffaele Pesenti. A branchprice-and-cut algorithm for the minimum evolution problem. European Journal of Operational Research, 244(3):753 – 765, 2015. ISSN 0377-2217. doi: http://dx.doi.org/ 10.1016/j.ejor.2015.02.019. URL http://www.sciencedirect.com/science/article/pii/ S0377221715001204.

- S. Devadoss and S. Petti. A space of phylogenetic networks. SIAM Journal on Applied Algebra and Geometry, 1:683–705, 2017.
- Satyan Devadoss, Cassandra Durell, and Stefan Forcey. Split network polytopes and network spaces. SLC 82B, Proceedings of FPSAC 31, to appear, 2019.
- K. Eickmeyer, P. Huggins, L. Pachter, and R. Yoshida. On the optimality of the neighborjoining algorithm. *Algorithms for Molecular Biology*, 3(5), 2008. URL http://www.almob. org/content/3/1/5.
- S. Forcey, L. Keefe, and W. Sands. Facets of the balanced minimal evolution polytope. *Journal* of Mathematical Biology, 73(2):447–468, 2016.
- S. Forcey, L. Keefe, and W. Sands. Split-facets for balanced minimal evolution polytopes and the permutoassociahedron. *Bulletin of Mathematical Biology*, 79(5):975–994, May 2017. doi: 10.1007/s11538-017-0264-7.
- S. Forcey, G. Hamerlinck, and W. Sands. Optimization problems in phylogenetics: Polytopes, programming and interpretation. In R. Robeva and M. Macauley, editors, *Algebraic and Combinatorial Computational Biology*. Academic Press, 2018.
- P. Gambette, K. T. Huber, and G. E. Scholz. Uprooted phylogenetic networks. Bulletin of Mathematical Biology, 79(9):2022–2048, Sep 2017a. ISSN 1522-9602. doi: 10.1007/ s11538-017-0318-x. URL https://doi.org/10.1007/s11538-017-0318-x.
- Philippe Gambette, Katharina Huber, and Steven Kelk. On the challenge of reconstructing level-1 phylogenetic networks from triplets and clusters. *Journal of Mathematical Biology*, 74(7):1729–1751, 2017b. doi: 10.1007/s00285-016-1068-3. URL https://hal-upec-upem. archives-ouvertes.fr/hal-01391430.
- Ewgenij Gawrilow and Michael Joswig. polymake: a framework for analyzing convex polytopes. In Gil Kalai and Günter M. Ziegler, editors, *Polytopes — Combinatorics and Computation*, pages 43–74. Birkhäuser, 2000.
- D. Haws, T. Hodge, and R. Yoshida. Optimality of the neighbor joining algorithm and faces of the balanced minimum evolution polytope. *Bull. Math. Biol.*, 73(11):2627–2648, 2011. ISSN 0092-8240. doi: 10.1007/s11538-011-9640-x. URL http://dx.doi.org/10.1007/ s11538-011-9640-x.
- D. Levy and Lior Pachter. The neighbor-net algorithm. *Advances in Applied Mathematics*, 47: 240–258, 2011.
- C. Semple and M. Steel. Unicyclic networks: compatibility and enumeration. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 3(1):84–91, Jan 2006. ISSN 1545-5963. doi: 10.1109/TCBB.2006.14.
- Charles Semple and Mike Steel. Cyclic permutations and evolutionary trees. *Adv. in Appl. Math.*, 32(4):669–680, 2004. ISSN 0196-8858. doi: 10.1016/S0196-8858(03)00098-8. URL http://dx.doi.org/10.1016/S0196-8858(03)00098-8.
- N. J. A. Sloane. *The On-Line Encyclopedia of Integer Sequences*. 2018. published electronically at www.oeis.org.
- Mike Steel. Phylogeny—discrete and random processes in evolution, volume 89 of CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2016. ISBN 978-1-611974-47-8. doi: 10.1137/1. 9781611974485.ch1. URL https://doi.org/10.1137/1.9781611974485.ch1.

(S. Forcey) DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF AKRON, AKRON, OH 44325-4002 *E-mail address:* sf34@uakron.edu *URL:* http://www.math.uakron.edu/~sf34/