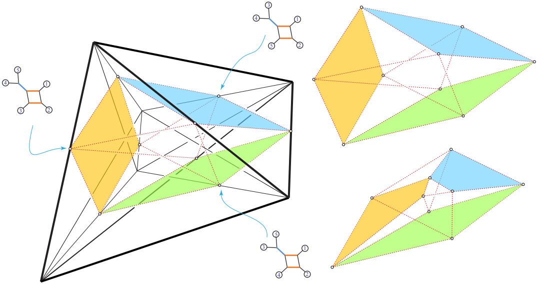
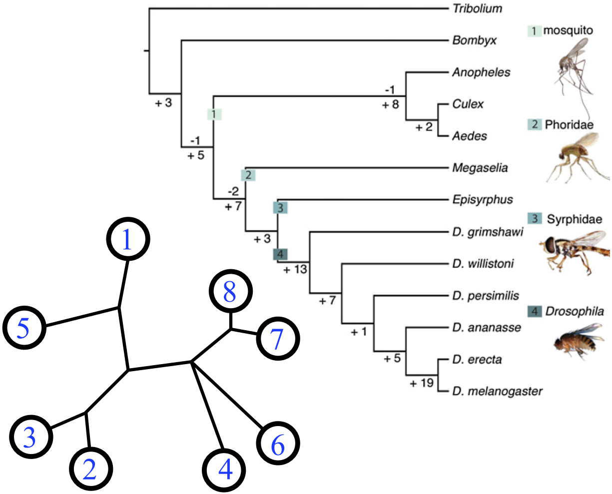
Galois connections for phylogenetic networks and their polytopes.

with S. Devadoss, San Diego; C. Durell and D. Scalzo, Akron.

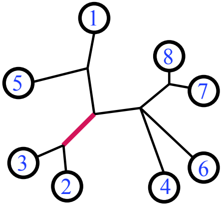


Trees



Episodic radiations in the fly tree of life, [Wiegmann et.al. PNAS 2011]

Splits

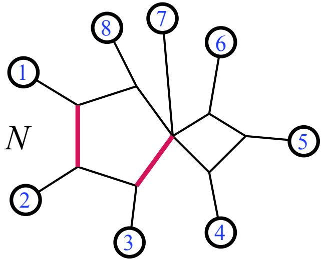


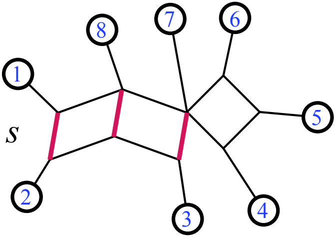
Phylogenetic trees: every edge is a split of $[n]$.

$A | B = \{2, 3\} | \{1, 4, 5, 6, 7, 8\}$

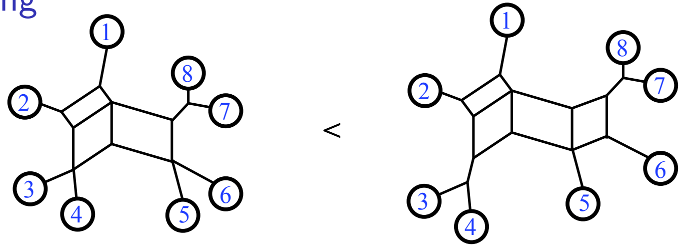
1) Phylogenetic networks: splits are minimal cuts.
These model all the ways heredity can happen.
Here, N is 1-nested: each edge in at most 1 cycle.

2) Split networks: splits are parallel classes of edges.
These model all the ways that genetic distances can be measured.
Here, s is circular: it can be drawn on the plane with leaves on the exterior in circular order c.



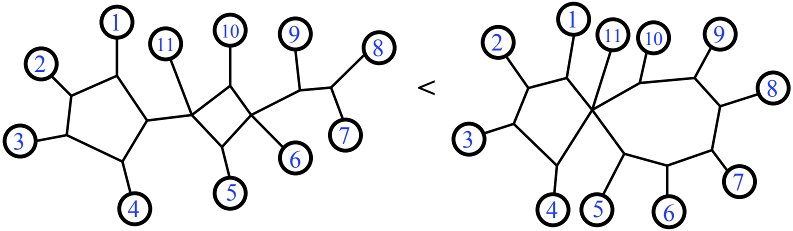


Ordering

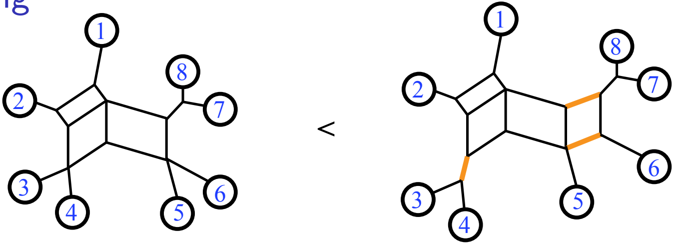


Partial ordering on both sets: $x \leqslant y$ when the set of splits displayed by x is contained in that displayed by y.

(Two networks, split or phylogenetic, are considered equivalent when their sets of splits are the same.)

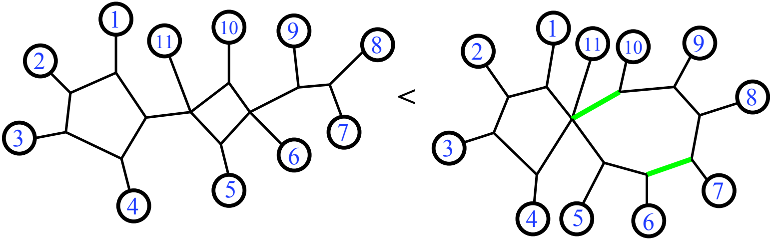


Ordering

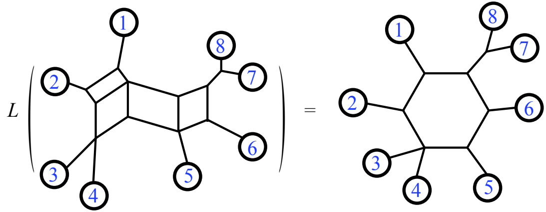


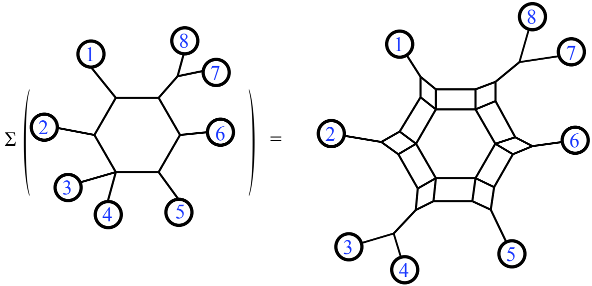
Partial ordering on both sets: $x \leqslant y$ when the set of splits displayed by x is contained in that displayed by y.

(Two networks, split or phylogenetic, are considered equivalent when their sets of splits are the same.)



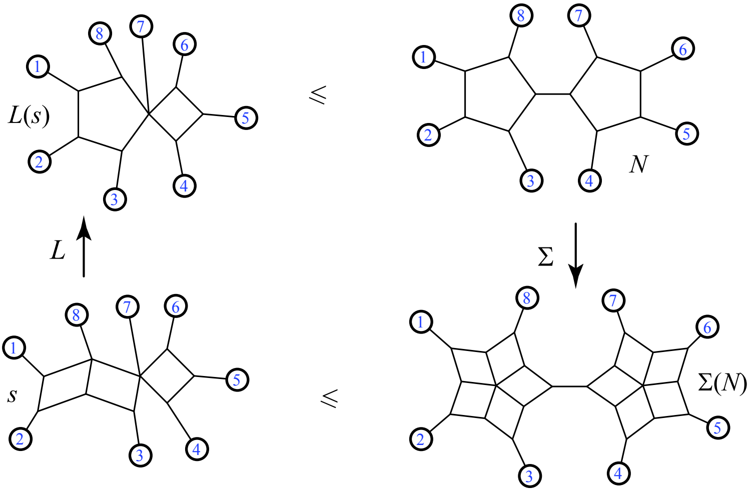
Mapping: [Gambette, Huber, Scholz, 2017]





Both maps are monotone.

Galois connection: Reflection



$$L(s) \leq N \iff s \leq \Sigma(N)$$

L is surjective; Σ is injective.

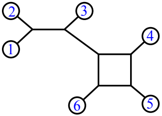
Polytopes

Definition

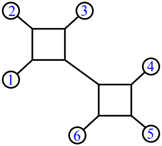
For a binary, level-1 phylogenetic network N, the vector $\mathbf{x}(N)$ is defined to have lexicographically ordered components $x_{ij}(N)$ for each unordered pair of distinct leaves $i, j \in [n]$ as follows:

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

where k is the number of bridges in N and b_{ij} is the number of bridges crossed on any path from i to j.
The convex hull of all the $\mathbf{x}(N)$ with k nontrivial bridges is the level-1 network polytope BME(n, k).



(4, 2, 1, 0, 1, 2, 1, 0, 1, 2, 0, 2, 4, 0, 4)



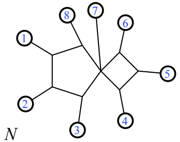
(2, 0, 1, 0, 1, 2, 0, 0, 0, 1, 0, 1, 2, 0, 2)

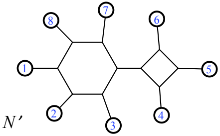
Result: faces of BME(n, k)

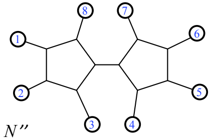
Theorem

Every n leaved 1-nested network N corresponds to a face of each BME(n, k) polytope.

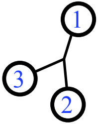
That face has vertices all the binary level-1 k-bridge networks N' such that $N \leq N'$.





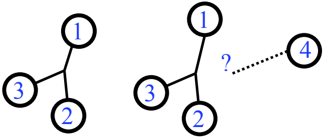


Counting vertices



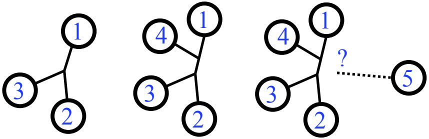
1

Counting vertices



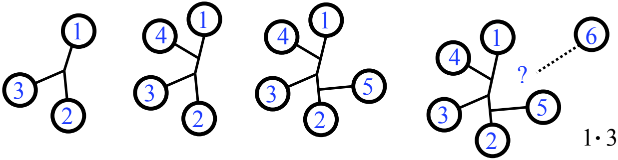
$1 \cdot 3$

Counting vertices



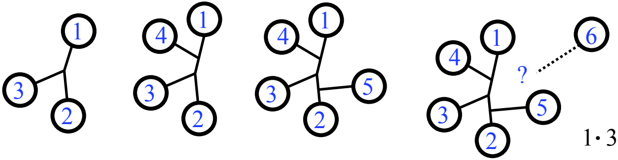
$1 \cdot 3 \cdot 5 \cdot$

Counting vertices

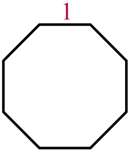


$1 \cdot 3 \cdot 5 \cdot 7 \ldots = (2n-5)!!$

Counting vertices

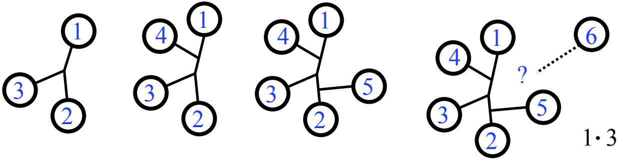


$1 \cdot 3 \cdot 5 \cdot 7 \ldots = (2n\text{-}5)!!$

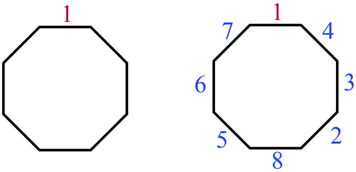


1

Counting vertices

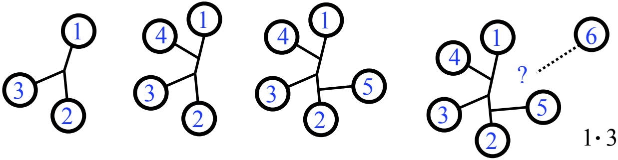


$$1 \cdot 3 \cdot 5 \cdot 7 \ldots = (2n-5)!!$$

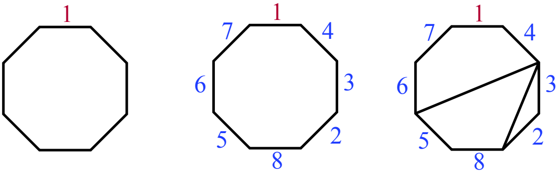


$$1 \cdot \frac{(n-1)!}{2}$$

Counting vertices

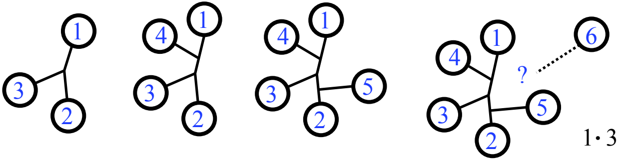


$$1 \cdot 3 \cdot 5 \cdot 7 \ldots = (2n-5)!!$$

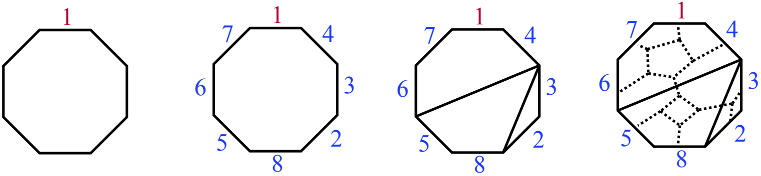


$$1 \cdot \frac{(n-1)!}{2}\ T(n,k)$$

Counting vertices

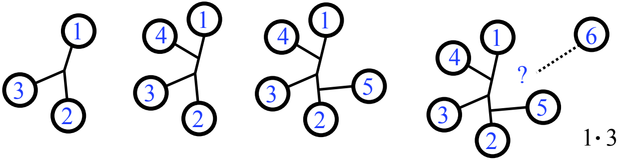


$$1 \cdot 3 \cdot 5 \cdot 7 \ldots = (2n-5)!!$$

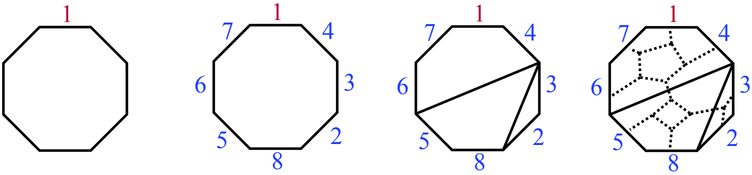


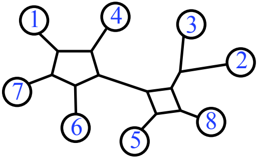
$$1 \cdot \frac{(n-1)!}{2}\; T(n, k)$$

Counting vertices



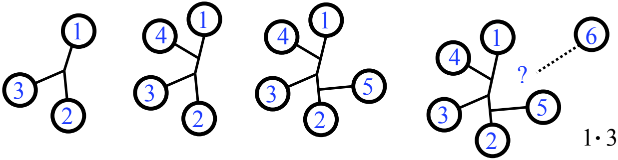
$$1 \cdot 3 \cdot 5 \cdot 7 \ldots = (2n\text{-}5)!!$$



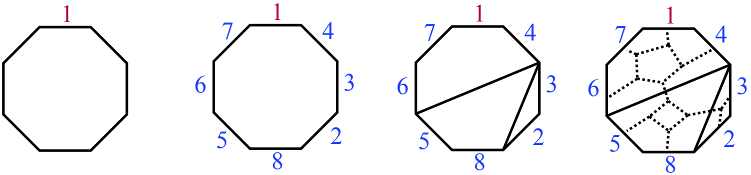


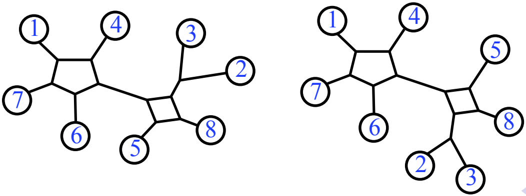
$$1 \cdot \frac{(n\text{-}1)!}{2}\; T(n,\, k\,)$$

Counting vertices



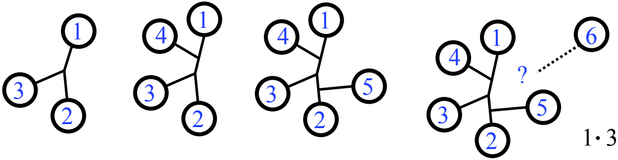
$$1 \cdot 3 \cdot 5 \cdot 7 \ldots = (2n-5)!!$$



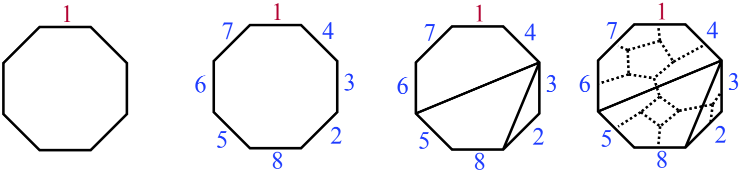


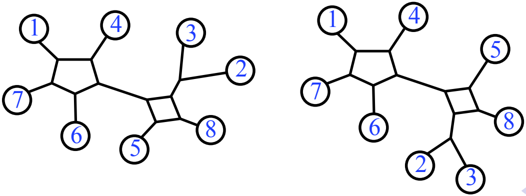
$$1 \cdot \frac{(n-1)!}{2}\; T(n,k)\; \frac{1}{2^k}$$

Counting vertices: note the cases $k = 0$, $k = n - 3$.



$$1\cdot 3\cdot 5\cdot 7\ldots = (2n-5)!!$$





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

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

Weighting

A *weighting* of a split network s is a function of splits, $w : s \to \mathbb{R}_{\geq 0}$.
For a weighted split network s we define the distance vector $\mathbf{d}_s$ 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.

A weighting for a phylogenetic network N is a function from edges to positive reals.
We define $\mathbf{d}_N$ where

$$\mathbf{d}_N(i,j) = \min\{\sum_{e \in P} w(e) \mid P \text{ path } i \to j\}$$

is the minimum sum of the weights of edges over all paths i to j.

Weighting

Theorem

For any weighted planar phylogenetic network N, the distance vector $\mathbf{d}_N$ obeys the Kalmanson condition: there exists a circular ordering c of $[n]$ such that for all $1 \leq i < j < k < l \leq n$ in that ordering,

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

Theorem

Theorem: For any Kalmanson vector $\mathbf{d}$, there is a unique weighted split network $s = NN(\mathbf{d})$ such that $\mathbf{d}_s = \mathbf{d}$.

Weighting

For any Kalmanson $\mathbf{d}$ for $[n]$, we can restrict the poset of phylogenetic networks (with weighting) to those that have $\mathbf{d}_N = \mathbf{d}$.

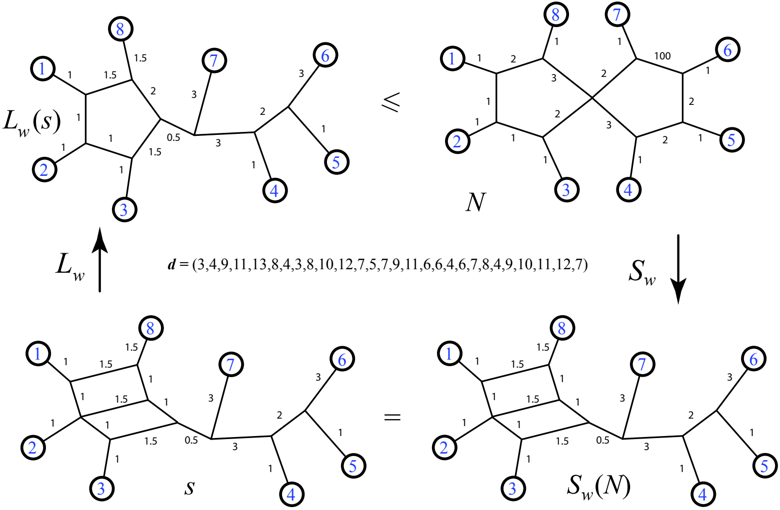
For a weighted split network s define $L_w(s)$ to be $L(s)$ with weight function

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

where the sum is of weights in s of splits whose cuts the edge is a member of. For a weighted phylogenetic network N define

$$S_w(N) = NN(d_N)$$

Galois connection: Coreflection



$$L_w(s) \leq N \iff s \leq S_w(N)$$

L_w is 1-1; S_w is onto.

Result

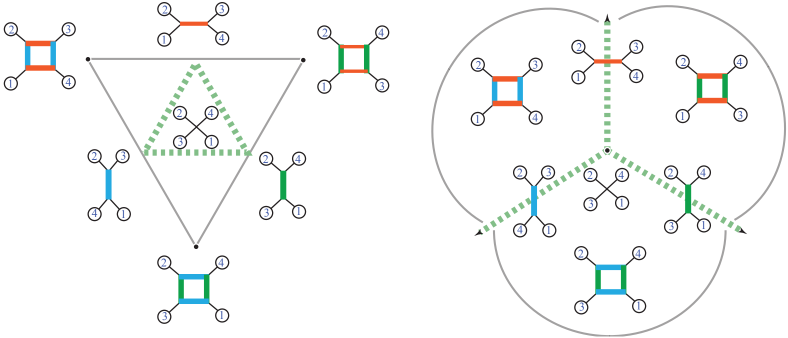
By comparing the two connections and the face theorem, we see that:

Theorem

Given any weighted phylogenetic network N, the product $\mathbf{x}(\hat{N}) \cdot \mathbf{d}_N$ is minimized simultaneously for the binary networks $\hat{N}$ with k bridges such that $\overline{S_w(N)} \leq \Sigma(\hat{N})$.

In $\overline{S_w(N)}$ the overline indicates forgetting the weighting.

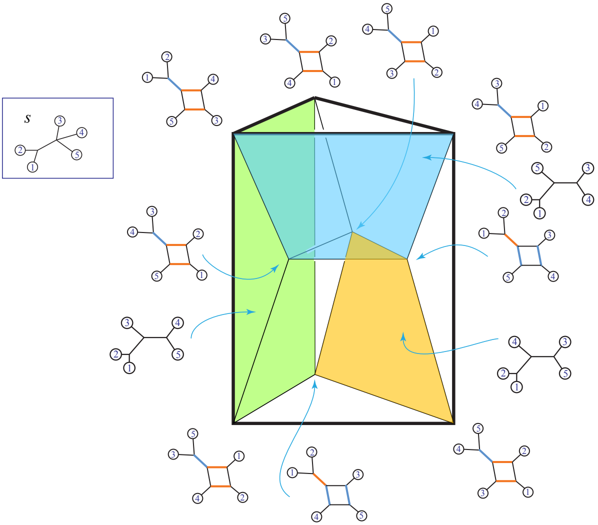
Polytope Pictures: Duality



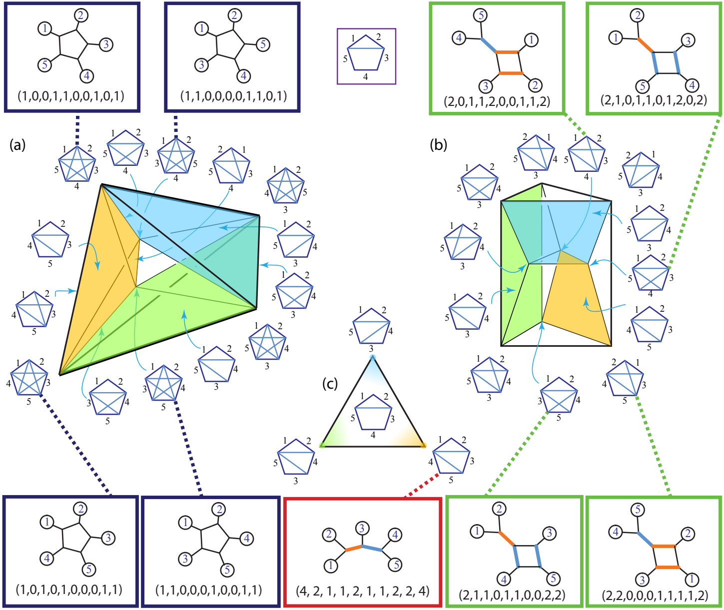
BME(4,0) on left, with BME(4,1) included.

CSN(4) on the right, with BHV(4) included.

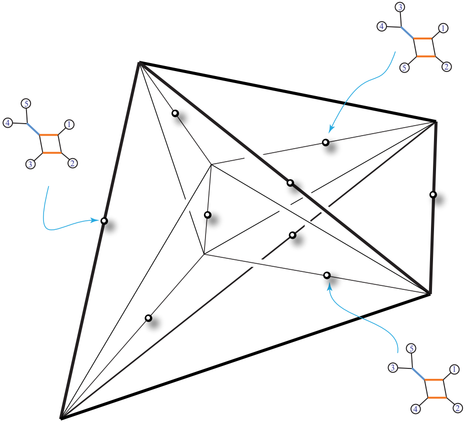
Facets



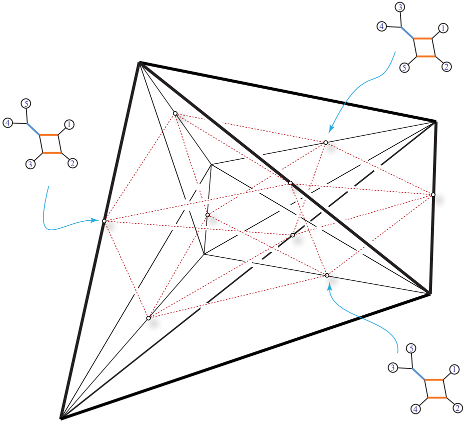
Facets


(1,0,0,1,1,0,0,1,0,1)
(1,1,0,0,0,0,1,1,0,1)
(2,0,1,1,2,0,0,1,1,2)
(2,1,0,1,1,0,1,2,0,2)
(a)
(b)
(c)
(1,0,1,0,1,0,0,0,1,1)
(1,1,0,0,0,1,0,0,1,1)
(4, 2, 1, 1, 2, 1, 1, 2, 2, 4)
(2,1,1,0,1,1,0,0,2,2)
(2,2,0,0,0,1,1,1,1,2)

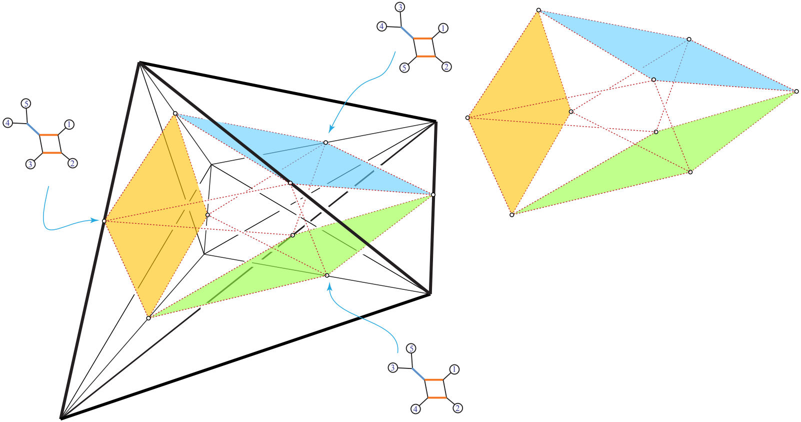
Nesting



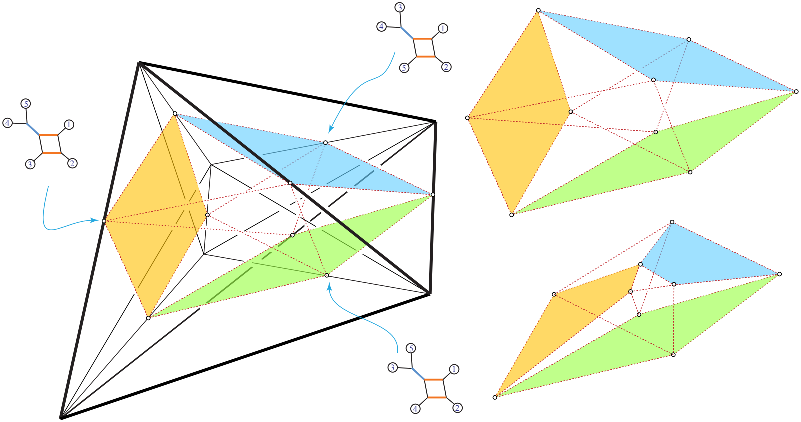
Nesting



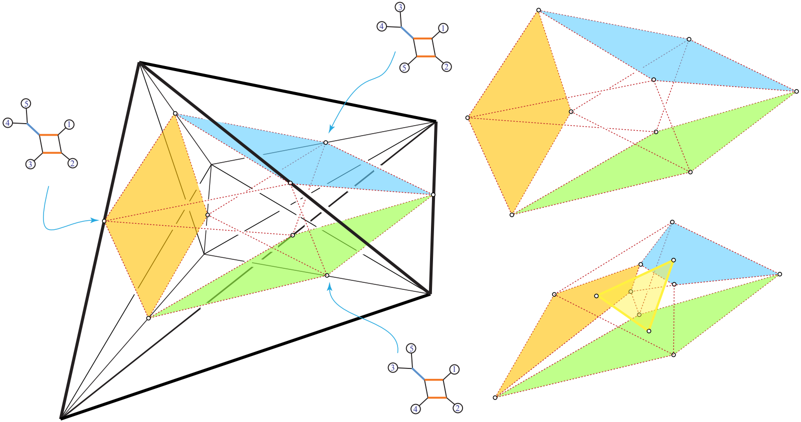
Nesting



Nesting



Nesting



Thanks so much!

Thanks so much!

Questions...

Thanks so much!

Questions...

What are the antipodes of the incidence Hopf algebras of these posets?