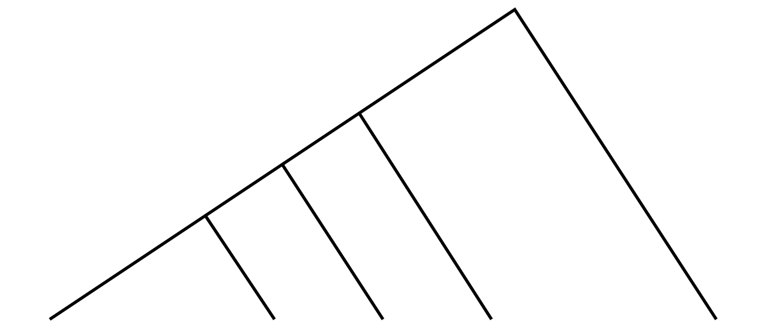
Aalto University
School of Science

PHYLOGENETIC SEMIGROUPS

KAIE KUBJAS, KAIE.KUBJAS@AALTO.FI

HISTORY

- The Jukes-Cantor binary model is the simplest group-based model with the underlying group $\mathbb{Z}_2$.
- The phylogenetic semigroup on a trivalent graph was defined by Buczyńska as a generalization of the affine semigroup of the Jukes-Cantor binary model on a trivalent tree (2012).
- Buczyńska, Buczyński, Michałek and Kubjas further generalized the definition of the phylogenetic semigroup to arbitrary graphs (2013).



DEFINITION

- Let G be a graph with edge set E and inner vertex set I.
- Define lattices

$$L_G = \{\omega \in \mathbb{Z}^E : \sum_{v \in e} \omega_e \in 2\mathbb{Z} \text{ for every } v \in I\}$$

and

$$L_G^{gr} = L_G \oplus \mathbb{Z}$$

together with the degree map

$$\deg : L_G^{gr} = L_G \oplus \mathbb{Z} \to \mathbb{Z}$$

given by the projection on the last summand.

- The lattice polytope P_T associated with the Jukes-Cantor binary model on T is

$$\text{conv}\{\omega \in L_T : \omega_e \in \{0,1\} \text{ for every } e \in E\}.$$

- If T is a claw tree, then the inequality description of P_T is

$$\{\omega \in [0,1]^E : \sum_{e \in F} \omega_e - \sum_{e \in E \setminus F} \omega_e \leq |F| - 1$$
$$\text{for all } F \subseteq E \text{ of odd cardinality}\}.$$

- For a general tree T, we get the inequality description of P_T by taking the union of inequalities for each claw tree around an internal vertex of T.
- Similarly, we define a polytope P_G for any graph. In general, it is not a lattice polytope anymore.
- The *phylogenetic semigroup* $\tau(G)$ on G is

$$\tau(G) = \text{cone}(P_G \times \{1\}) \cap L_G^{gr}.$$

PHYLOGENETIC NETWORKS

- Can phylogenetic semigroups be obtained as discrete Fourier transforms of generalizations of group-based models to phylogenetic networks?
- What is the relationship between phylogenetic semigroups and directed graphical models of phylogenetic networks (Strimmer & Moulton 2000)?
- What is the relationship between phylogenetic semigroups and extensions of tree models to splits networks (Bryant 2005)?
- Can minimal generators of phylogenetic semigroups be interpreted as splits of phylogenetic networks?

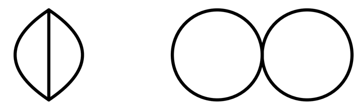
MINIMAL GENERATORS

- The phylogenetic semigroups on trees are generated by degree one labelings (Buczyńska & Wiśniewski 2007, Donten-Bury & Michałek 2012).
- Buczyńska proved that any minimal generator of the phylogenetic semigroup on a trivalent graph with first Betti number 1 has degree at most two, and explicitly described the minimal generating sets (2012).
- The maximal degree of the minimal generating set of the phylogenetic semigroup on a graph with first Betti number g is at most $g+1$ (Buczyńska, Buczyński, Michałek, Kubjas 2013).

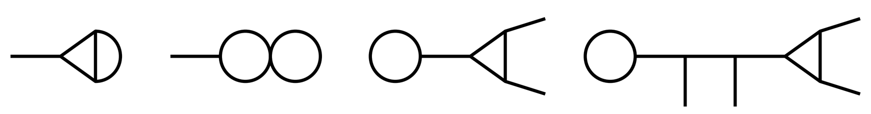
MAXIMAL DEGREES FOR $g = 2$

Theorem 1 *Let G be a graph with first Betti number 2. The maximal degree of a minimal generator of $\tau(G)$ is*

- *one if and only if G does not contain any cycle legs that are not cycle edges;*



- *two if and only if*
  - *the cycles of G live in different connected components, or*
  - *G contains at least one cycle leg that is not a cycle edge, all cycles of G live in the same connected component, and they are not separated by an inner vertex;*
- *three if and only if the minimal cycles of G live in the same connected component and are more than one edge apart from each other;*

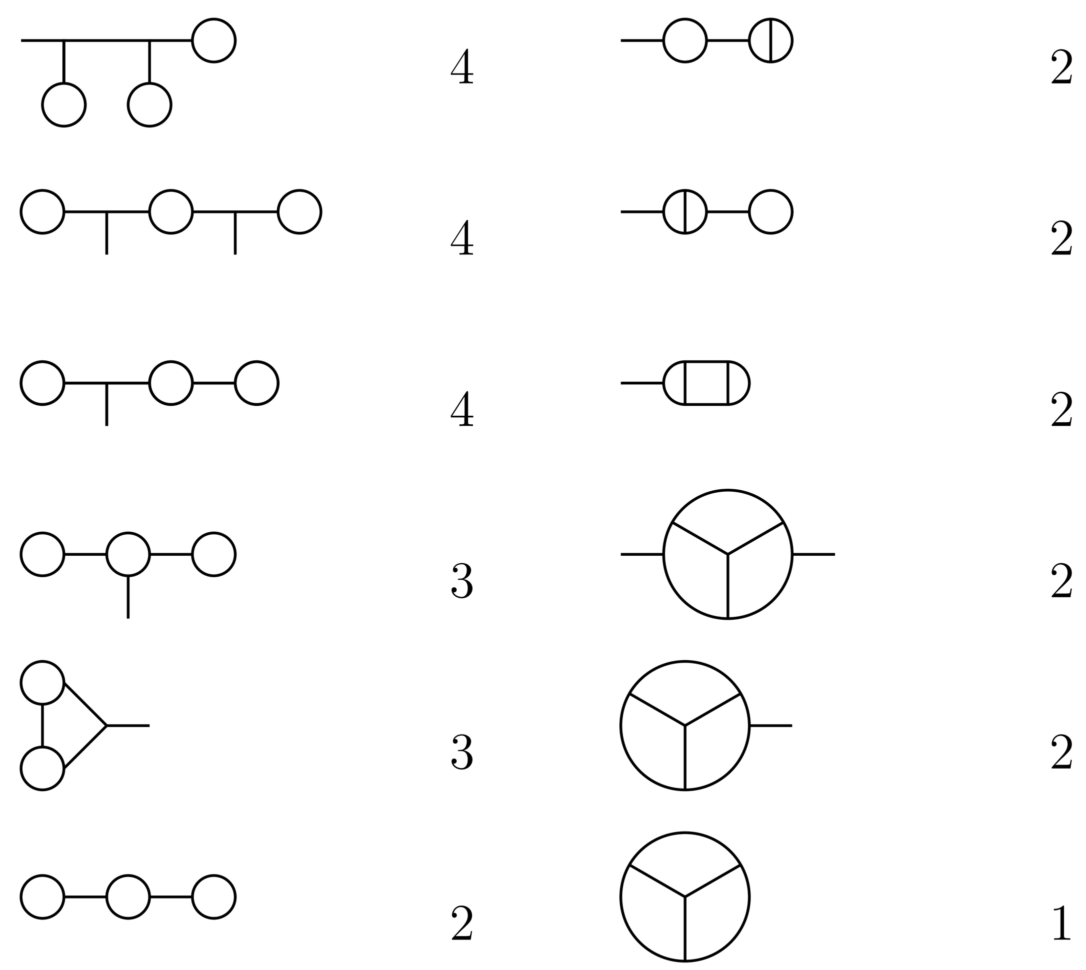


DEGREE 3 GENERATORS

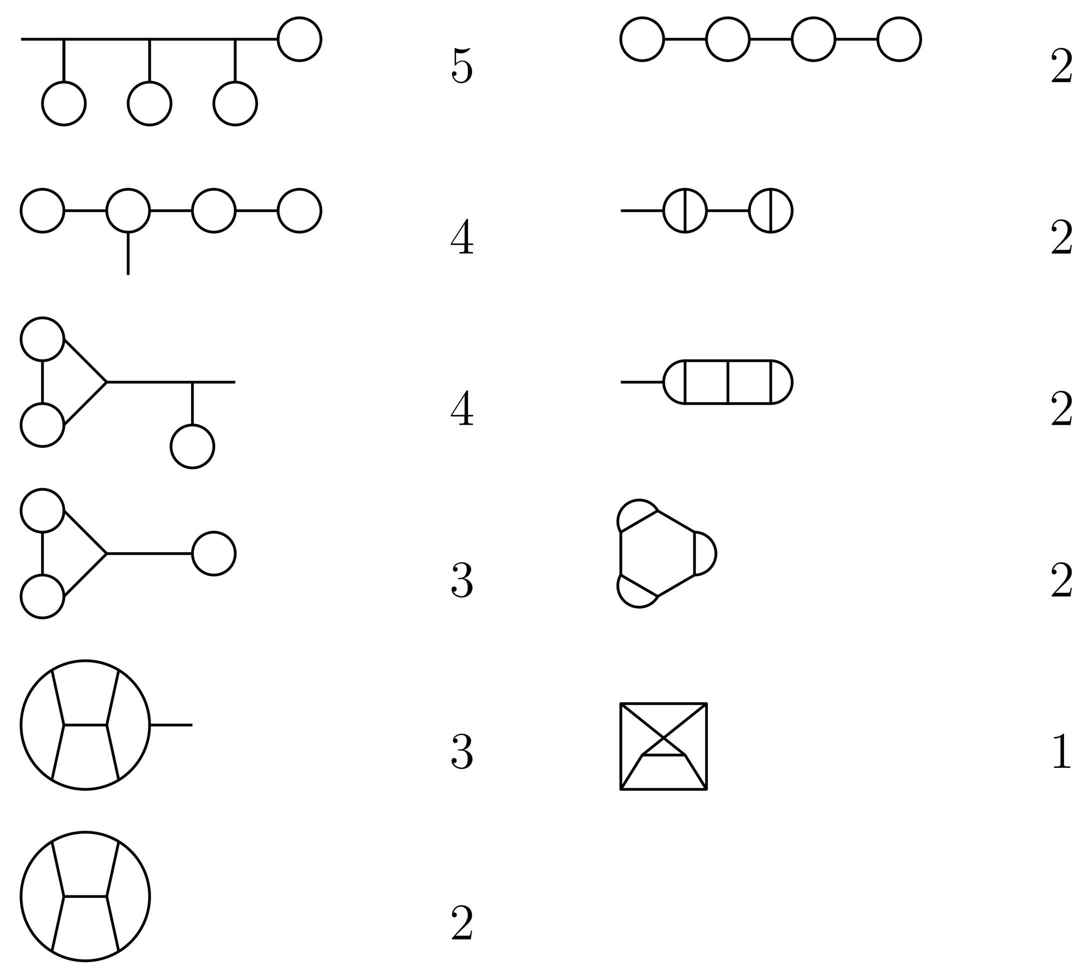
Theorem 2 *Let G be a trivalent graph with first Betti number 2 and the cycles of G more than one edge apart from each other. Then $\omega \in \tau(G)$ is a degree three minimal generator of $\tau(G)$ if and only if it satisfies the following three conditions:*

- *ω restricted to any cycle with its cycle legs does not decompose as a sum of degree one labelings,*
- *ω restricted to an edge on the shortest path between two cycles has value one or two, and*
- *ω restricted to exactly one edge incident to an edge on the shortest path between two cycles that is not a cycle edge or an edge on the shortest path has value one or two, and has value zero or three on all other such edges.*

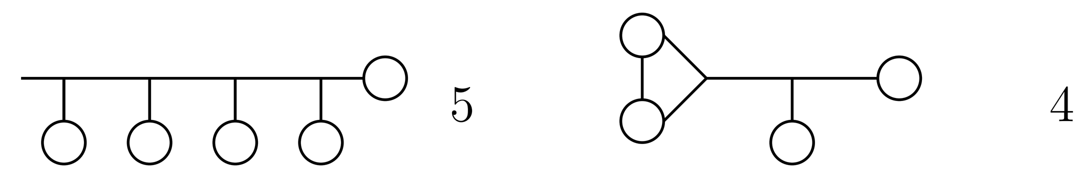
MAXIMAL DEGREES FOR $g = 3$



MAXIMAL DEGREES FOR $g = 4$



MAXIMAL DEGREES FOR $g = 5$



SOME RECENT PUBLICATIONS

- W. Buczyńska, *Phylogenetic toric varieties on graphs*, Algebraic Combin., 2012, 35(3), 421–460
- W. Buczyńska, J. Buczyński, K. Kubjas, M. Michałek, *Degrees of generators of phylogenetic semigroups on graphs*, Cent. Eur. J. Math., 2013, 11(9), 1577–1592
- K. Kubjas, *Low degree minimal generators of phylogenetic semigroups*, to appear in Eur. J. Math.