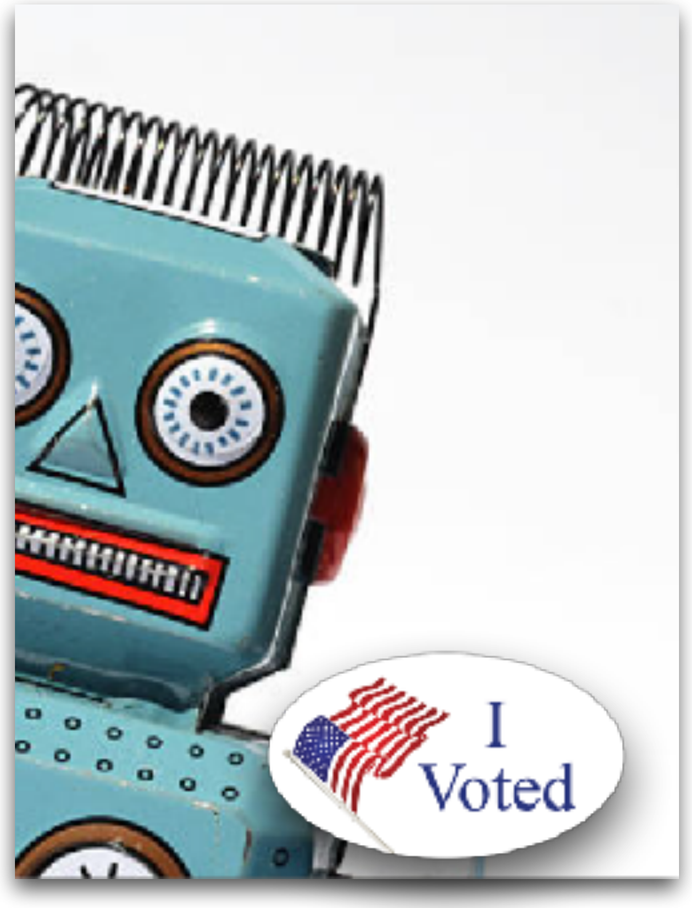
Modeling Democracy

Lecture 10 - **Ranking metrics**



Moon Duchin, Winter 2026 — Mon Feb 9

Let's put a distance on ballots, then on profiles

(this will show us metric structure in elections —
and later tell us whether a synthetic profile is "close" to a real one)

- One reasonable distance: **Kendall tau** (neighbor-swap distance).
  Dates to 1938.

- Another reasonable distance: **Spearman disarray** (sum of rank difference over the objects). Dates to 1906. (Also known as Spearman footrule.)

- Diaconis and Graham showed in 1977 that these are 2-bi-Lipschitz equivalent:

$$d_1 \le d_2 \le 2d_1$$

- (they also derive expectations and variance and suggest normalizations on S_n)

Natural coordinate embeddings

- $\mathfrak{b}(\sigma)(i) = m - \sigma(i)$ — "Borda embedding" in $\mathbb{R}^m$

$$d_B(\sigma, \tau) := \frac{1}{2} \|\mathfrak{b}(\sigma) - \mathfrak{b}(\tau)\|_1$$

Prop: d_H is Kendall tau!
and d_B is (half of) Spearman disarray!

- $\mathfrak{h}(\sigma)(\{i,j\}) = \begin{cases} +1 & \sigma(i) > \sigma(j) \\ 0 & \sigma(i) = \sigma(j) \\ -1 & \sigma(i) < \sigma(j) \end{cases}$ (assuming $i < j$) — "head-to-head embedding" in $\mathbb{R}^{\binom{m}{2}}$

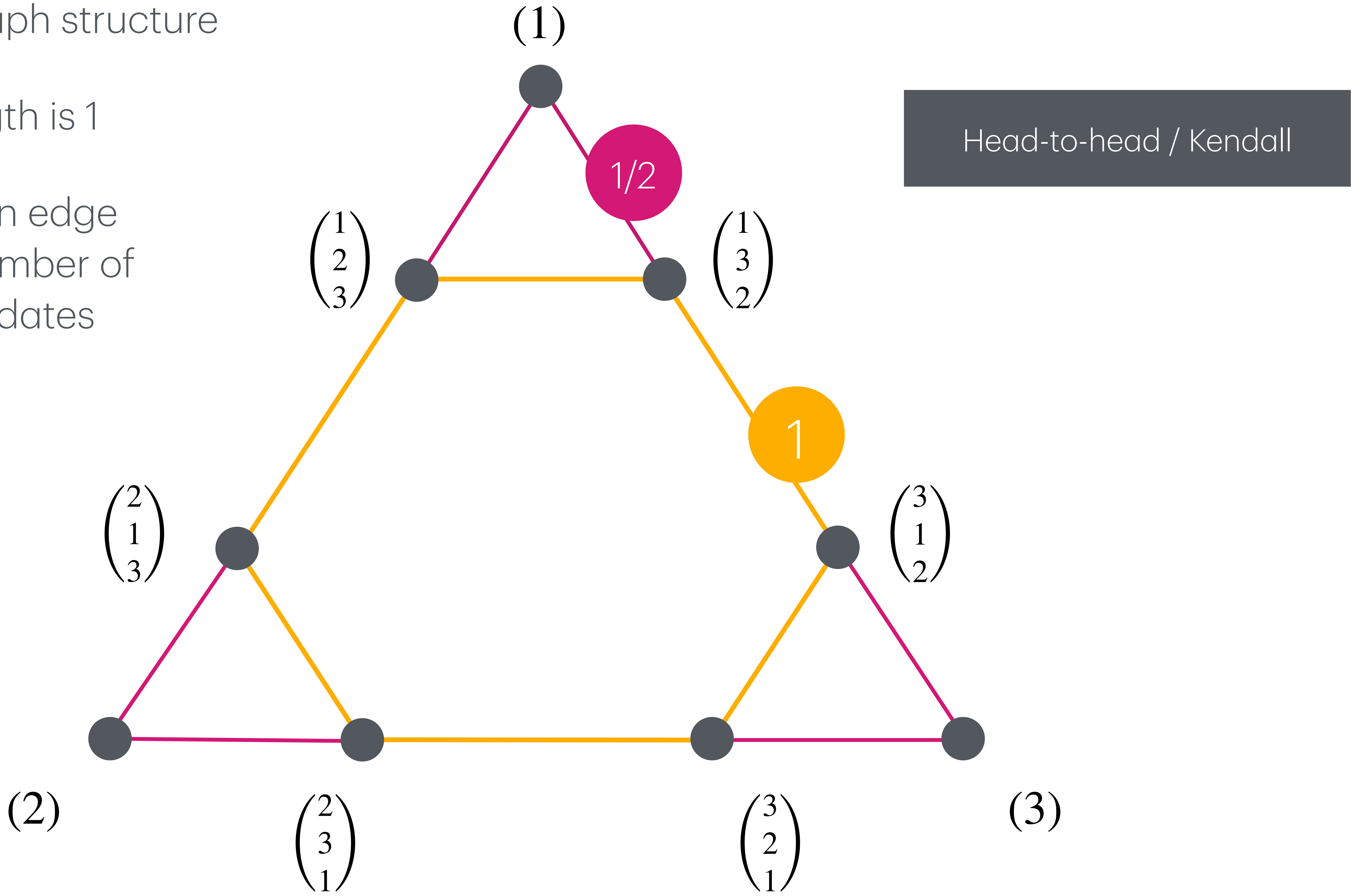
$$d_H(\sigma, \tau) := \frac{1}{2} \|\mathfrak{h}(\sigma) - \mathfrak{h}(\tau)\|_1$$

$$\underset{\text{A B C D}}{(3,\ 1,\ 2,\ 0)} \xleftarrow{\ \mathfrak{b}\ } \begin{pmatrix} A \\ C \\ B \\ D \end{pmatrix} \xrightarrow{\ \mathfrak{h}\ } \underset{\text{AB AC AD BC BD CD}}{(1,\ 1,\ 1,\ -1,\ 1,\ 1)}$$

Useful observation: these are compatible with graph structure

Neighbor swap length is 1

Extension/truncation edge length is half the number of unmentioned candidates

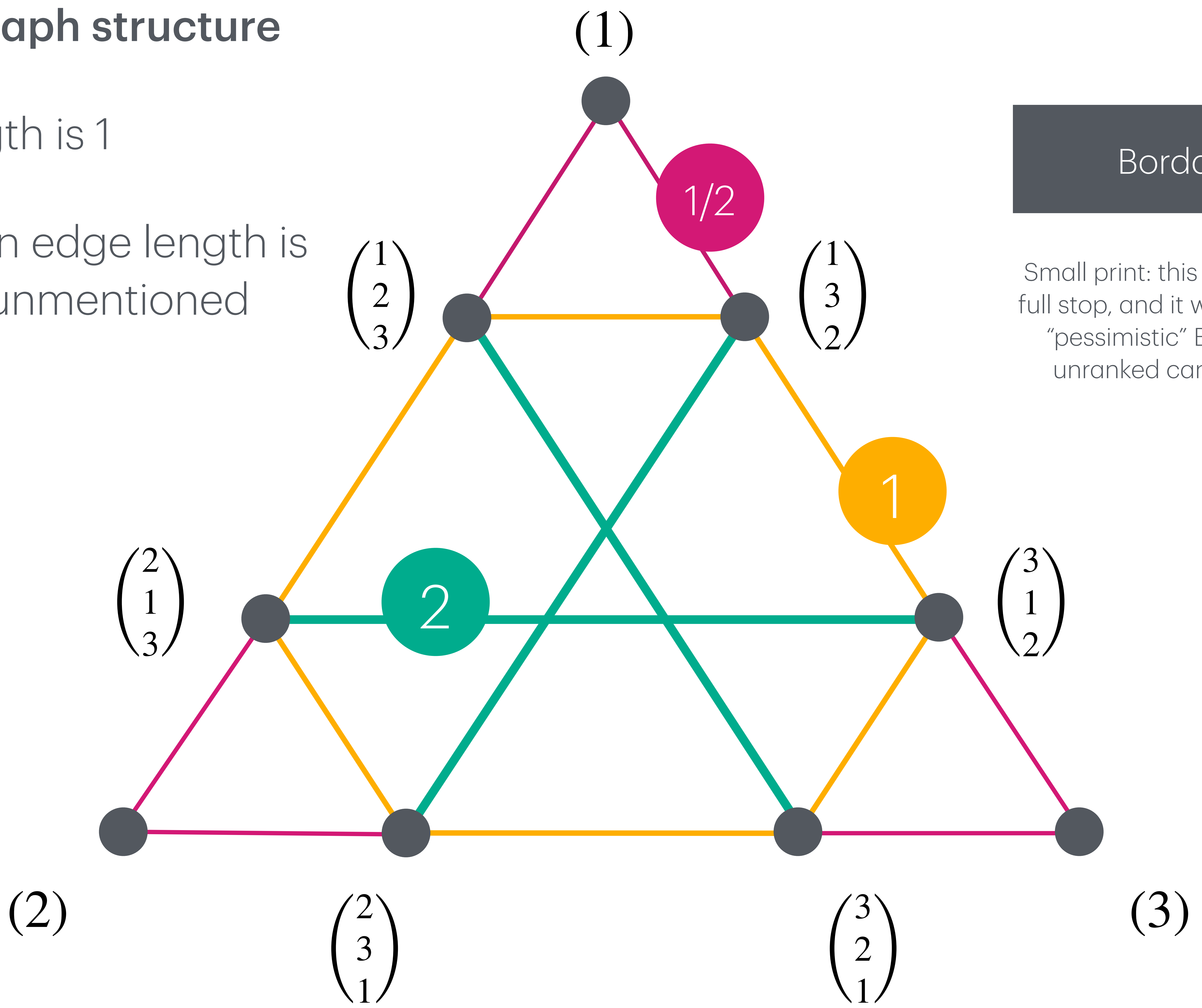


Useful observation: these are
compatible with graph structure

Neighbor swap length is 1

Extension/truncation edge length is half the number of unmentioned candidates

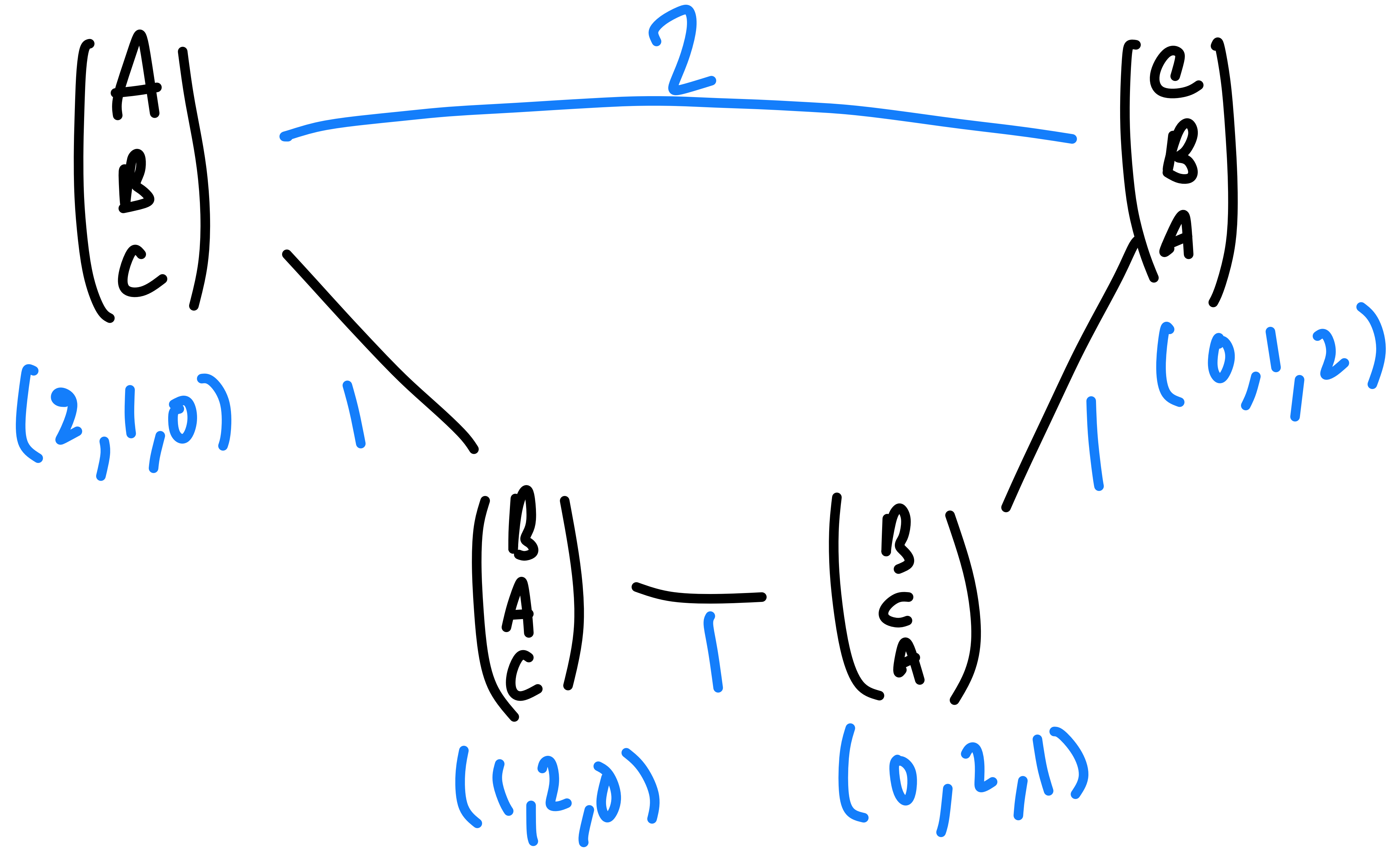
Shortcut edges: arbitrary swaps, length is rank difference



Borda / Spearman

Small print: this works for complete ballots full stop, and it works for partial ballots with "pessimistic" Borda embedding, where unranked candidates get zero points.

Let's see about those shortcuts.

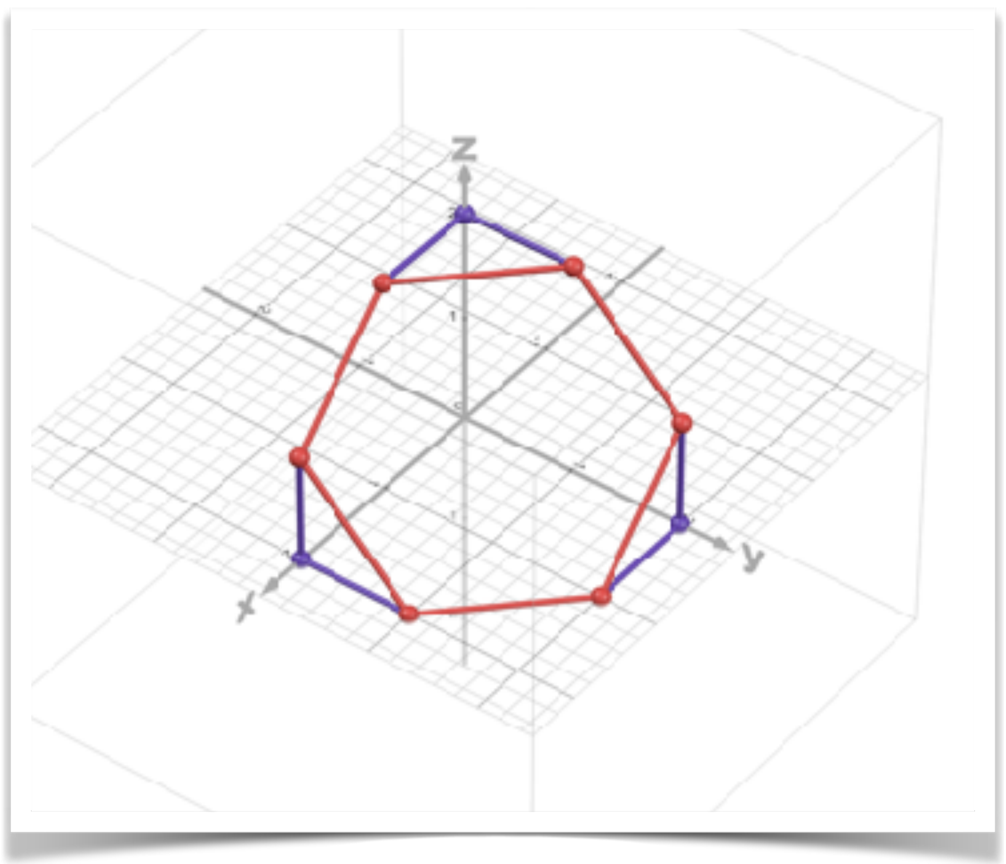


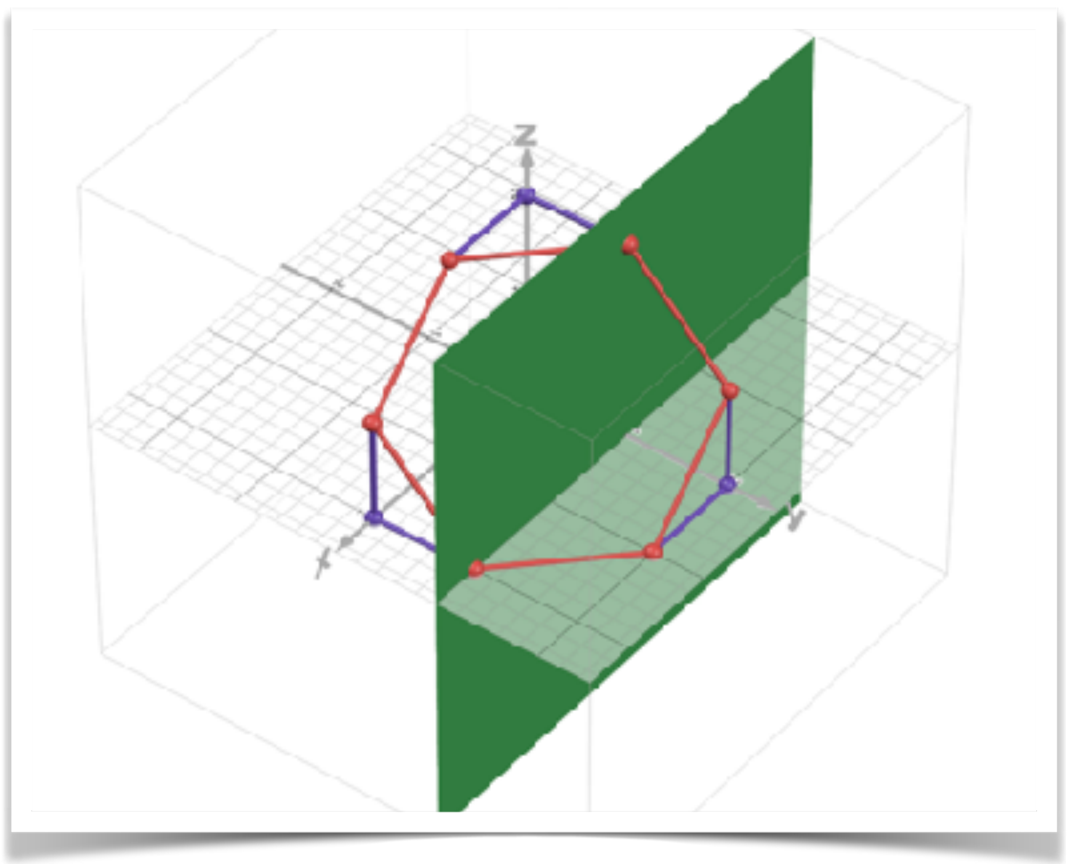
note that the empty ballot,

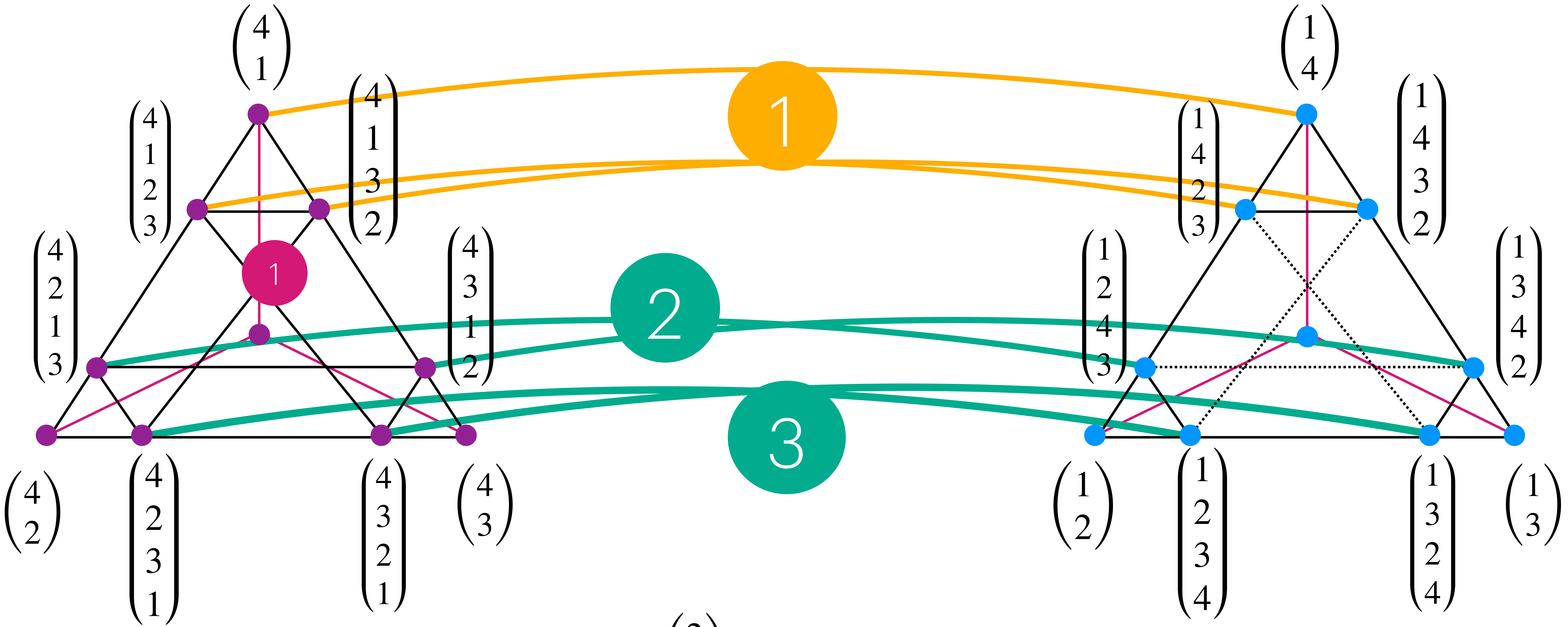
i.e., $(\{A,B,C\})$

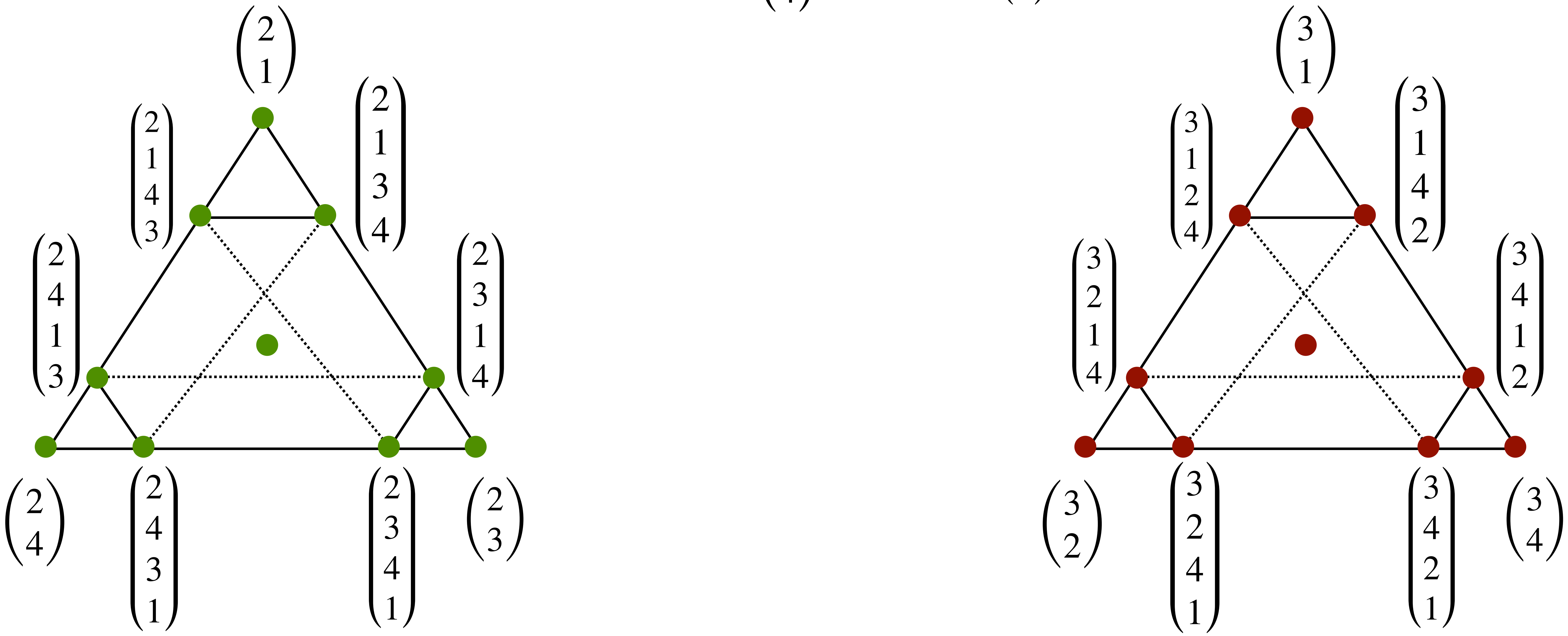
$(1,1,1)$ $(\mathfrak{b}_{\text{avg}})$

is the only point between them that is a ballot.







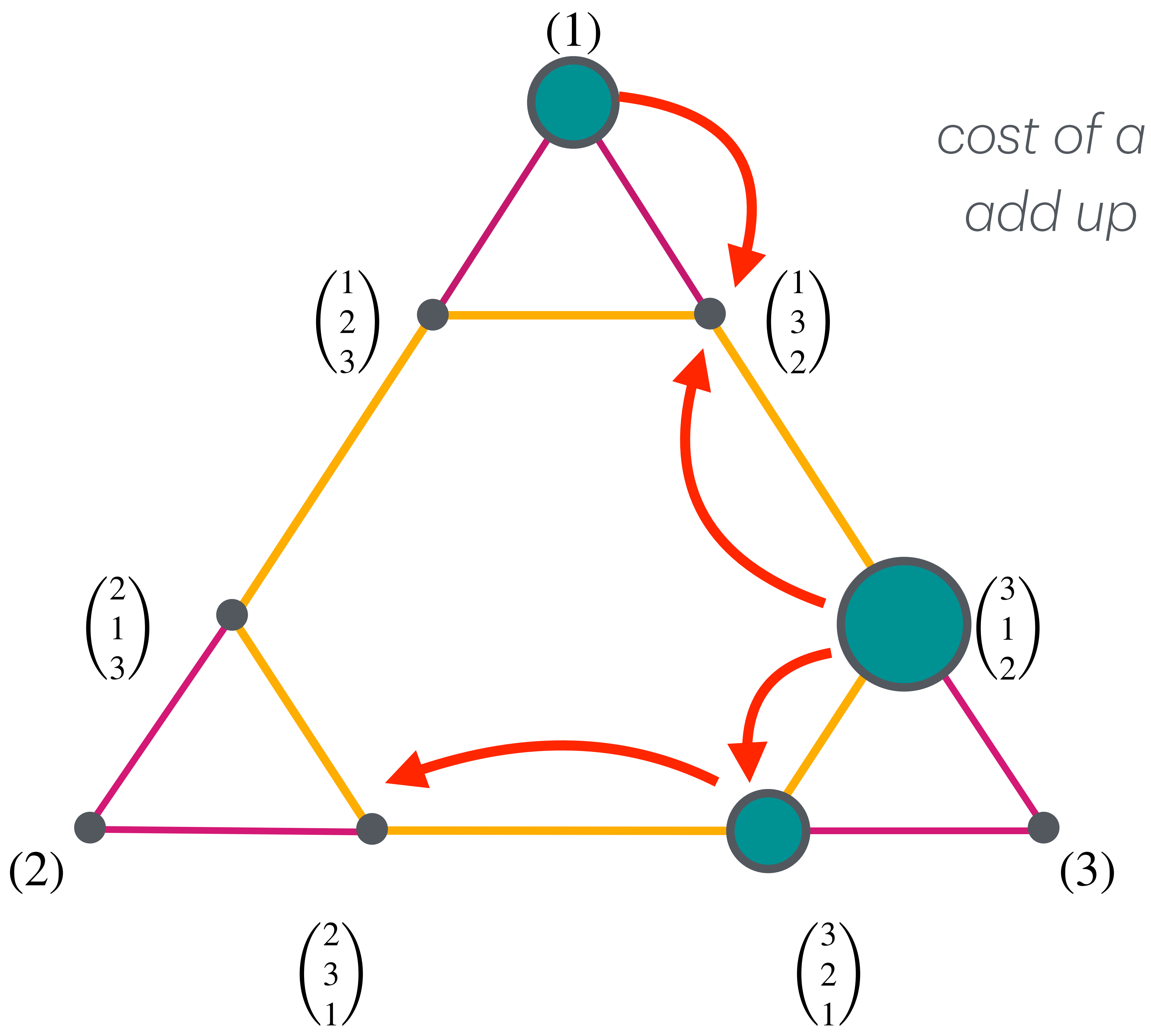


A representative part of the 4-candidate ballot graph, with shortcut edges shown in **teal**

Distance between profiles

- Now that we have a graph, we can realize a profile by node weights on the graph—a distribution on the graph

- Can use “**earth mover distance**” (optimal transport) to get a distance on profiles



cost of a “transportation plan”:
add up mass times distance