Improved Orientations of Physical Networks

Iftah Gamzu
TAU

Danny Segev
HAIFA

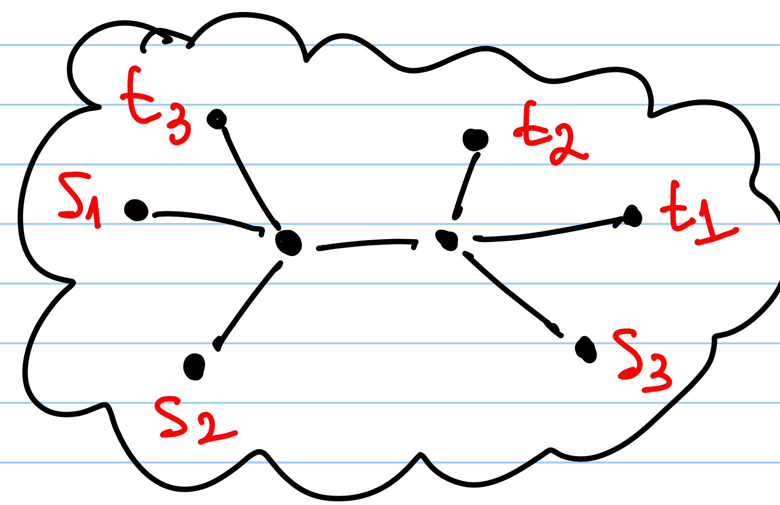
Roded Sharan
TAU

MOTIVATION

- Protein-protein interactions propagate directional signals in the cell
- Current measurement techniques capture only the presence of an interaction but not its direction
- Indirect directionality information can be obtained from knockout experiments that expose cause-effect pairs
- Goal: use pair information to infer interaction directions

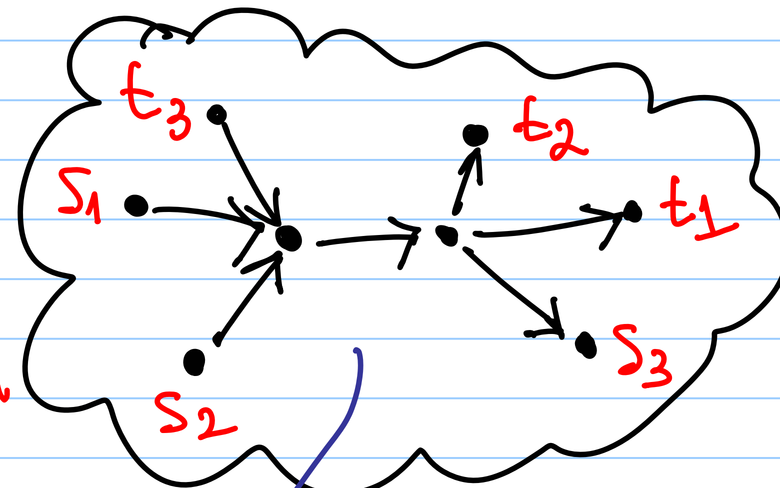
MAXIMUM GRAPH ORIENTATION

- Input: undirected graph $G=(V,E)$ with n vertices
  source-target pairs $(s_1,t_1),\dots,(s_k,t_k)$



MAXIMUM GRAPH ORIENTATION

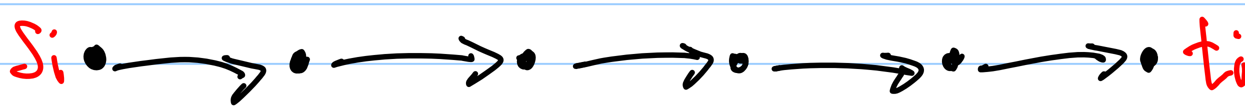
- Input: undirected graph $G=(V,E)$ with n vertices
  source-target pairs $(s_1,t_1),\dots,(s_k,t_k)$
- goal: compute an orientation in which the number of connected pairs is maximized



(s_1,t_1) ✓

(s_2,t_2) ✓

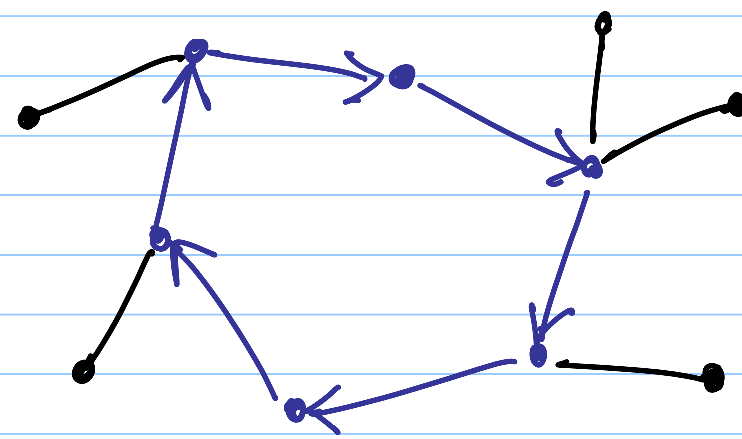
(s_3,t_3) ✗



MAXIMUM GRAPH ORIENTATION

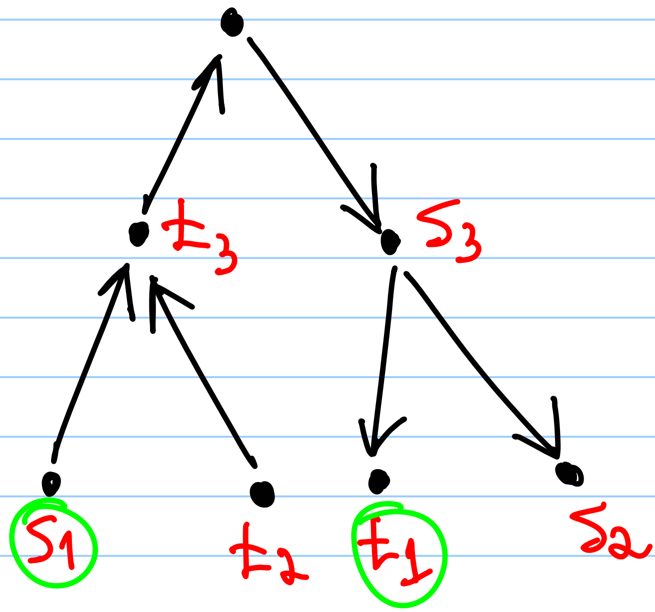
- input: undirected graph $G=(V,E)$ with n vertices
  source-target pairs $(s_1,t_1), \dots, (s_k,t_k)$

- goal: compute an orientation
  in which the number of
  connected pairs is maximized



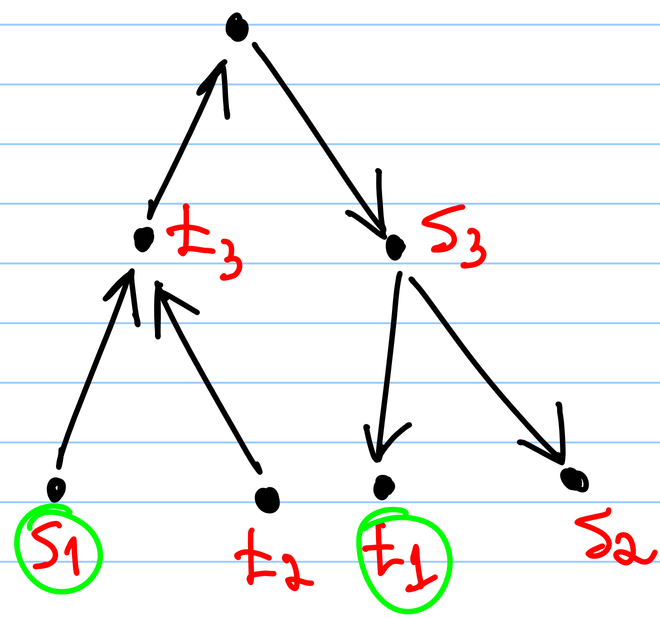
- remark: we may assume that the underlying
  graph is a tree

EXAMPLE

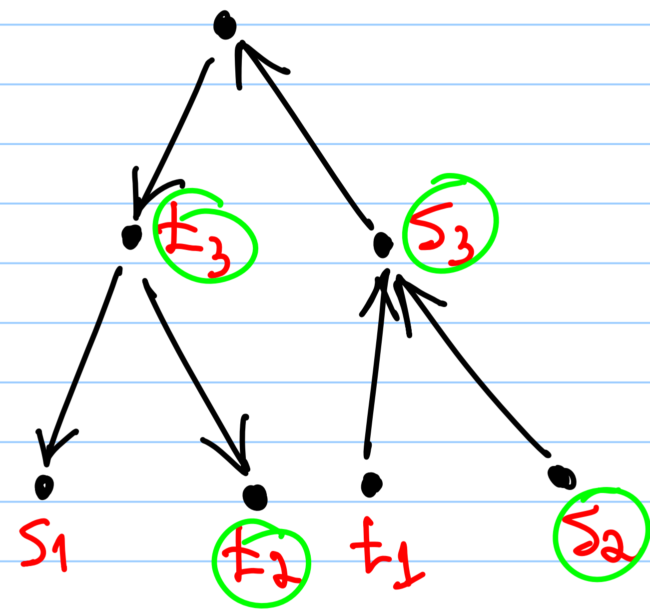


connected : 1

EXAMPLE



connected : 1



connected : 2

PREVIOUS WORK

Medvedovsky, Bafna, Zwick, and Sharan [WABI '08]:

- can be solved to optimality on paths
- star networks $\Longleftrightarrow$ max directed cut

  lower bounds: 0.916, unless P = NP (Hastad [JACM '01])

  0.878, UGC (Khot et al. [SICOMP '07])

  upper bound: 0.874 (Feige and Goemans '95)
- $o(\log n)$-approx for arbitrary trees $\Longleftarrow$

MAIN RESULT

deterministic algorithm with ratio

$$O(\log n / \log\log n)$$

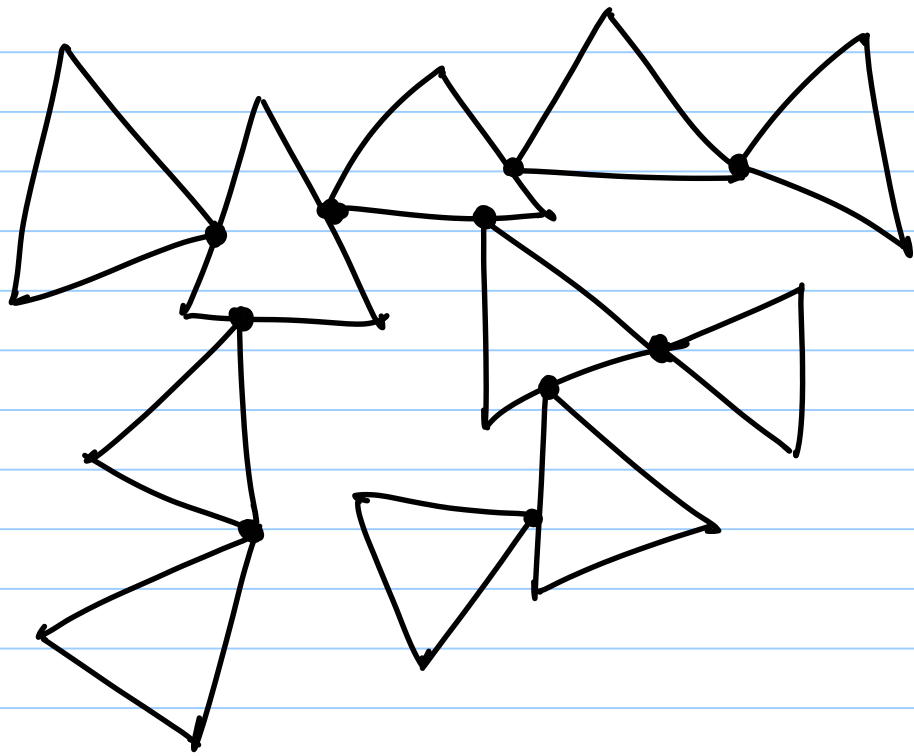
for maximum tree orientation

THE ALGORITHM:

GENERAL IDEA

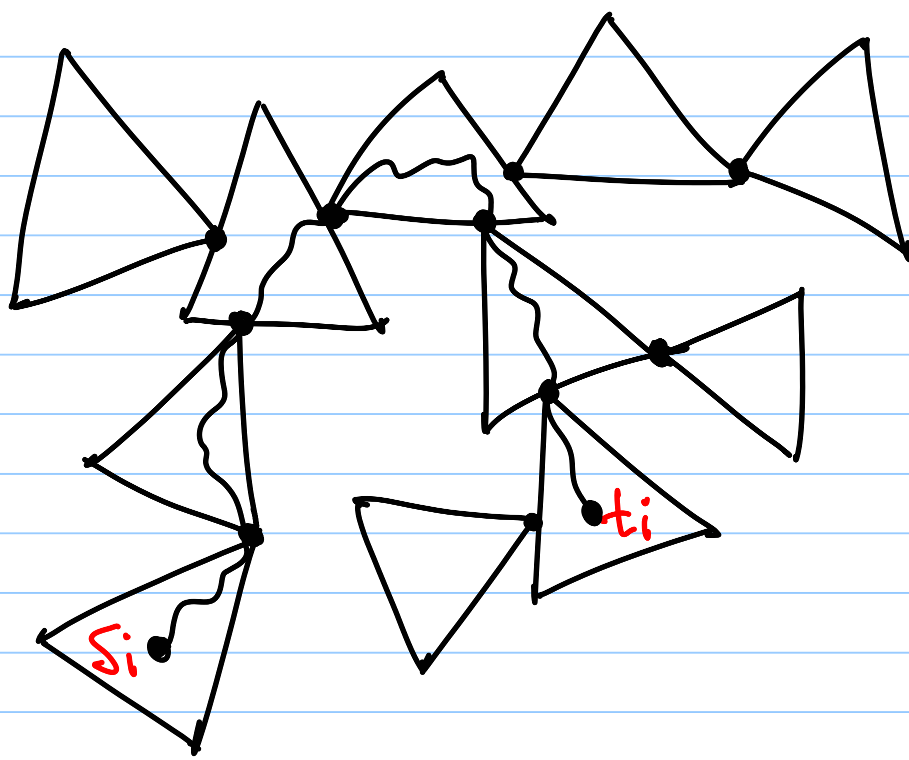
ALMOST BALANCED K-DECOMPOSITIONS

- any tree $T=(V,E)$ can be decomposed into k edge-disjoint subtrees, each containing $O(|E|/k)$ edges



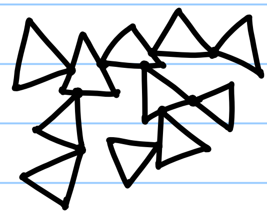
ALMOST BALANCED K-DECOMPOSITIONS

- any tree $T=(V,E)$ can be decomposed into k edge-disjoint subtrees, each containing $O(|E|/k)$ edges



- pair (s_i, t_i) separated when s_i and t_i reside in different subtrees

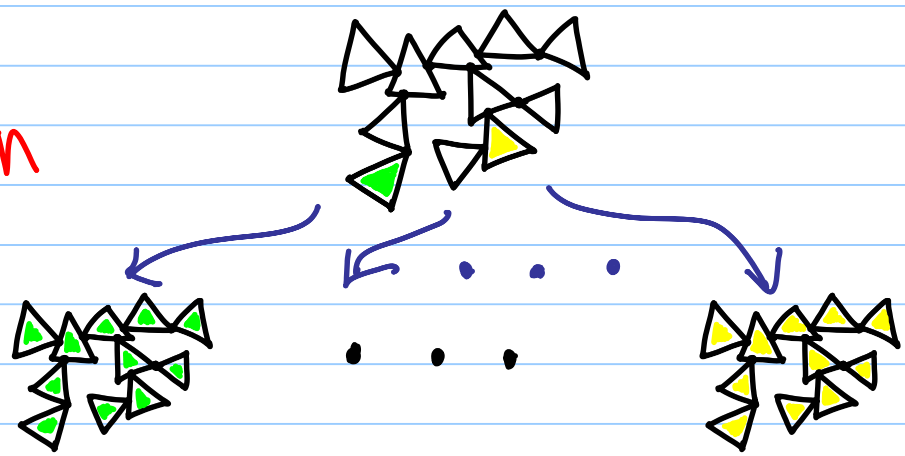
REPEATED K-DECOMPOSITIONS



pick $k = \log n$

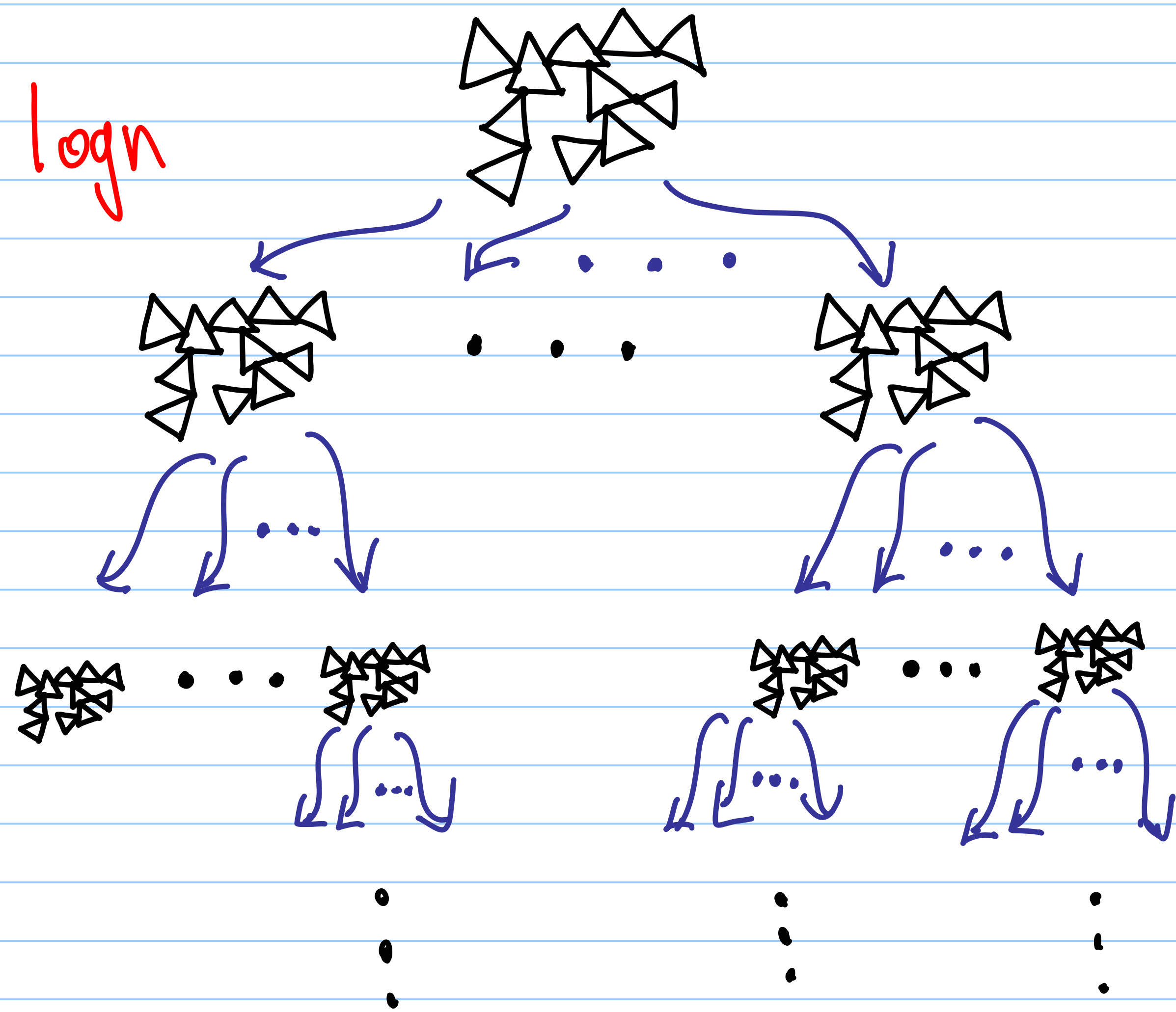
REPEATED K-DECOMPOSITIONS

pick $k = \log n$



REPEATED K-DECOMPOSITIONS

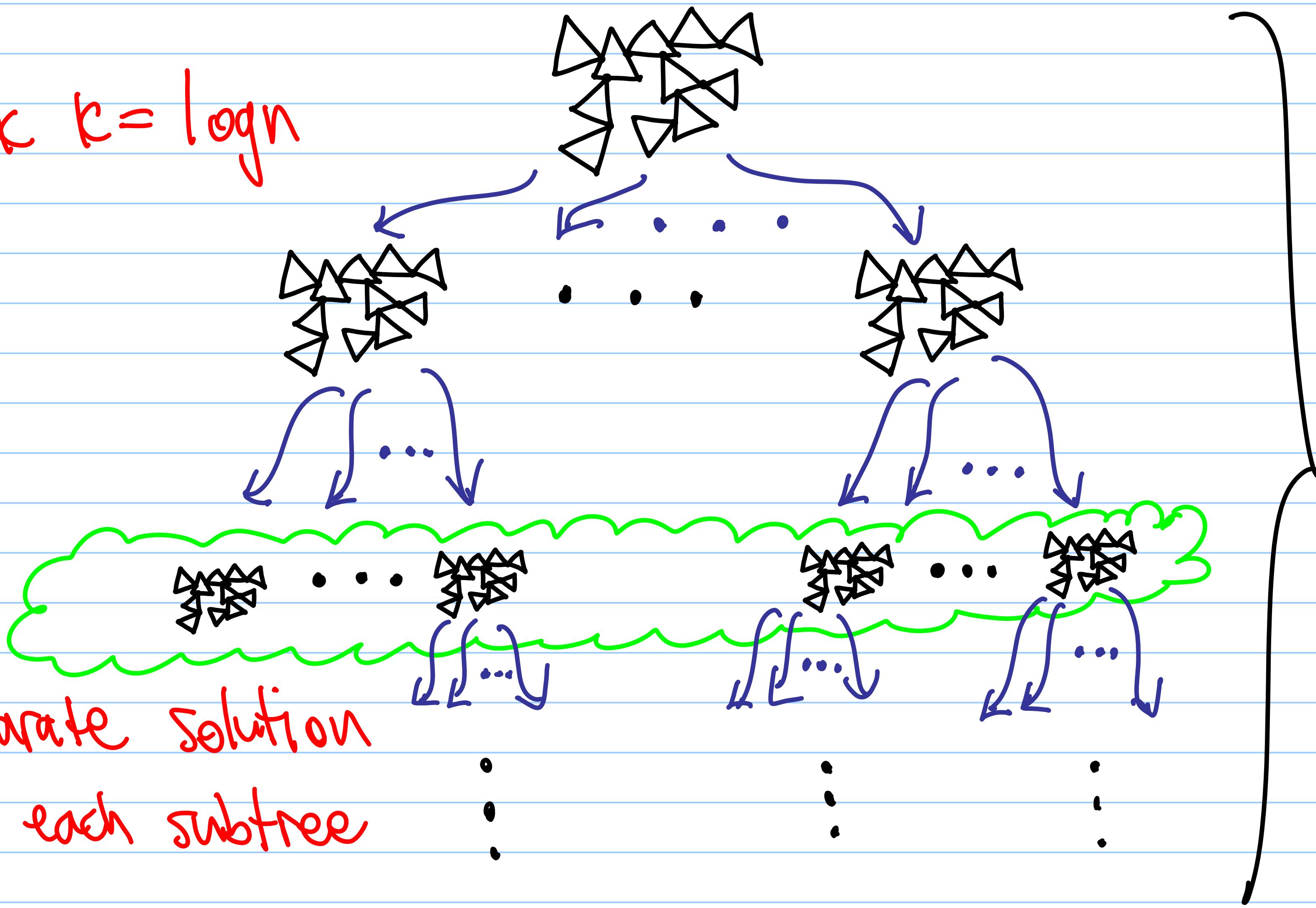
pick $k = \log n$



$O\left(\frac{\log n}{\log\log n}\right)$ levels

REPEATED K-DECOMPOSITIONS

pick $k = \log n$



separate solution
for each subtree

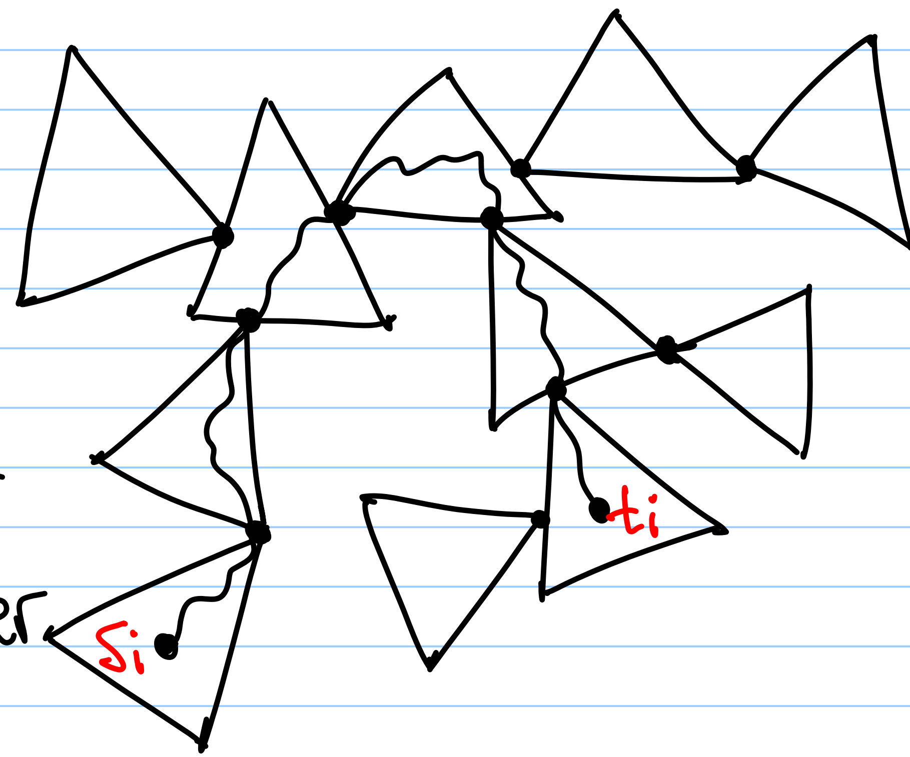
$O\left(\frac{\log n}{\log\log n}\right)$ levels

NEED O(1)-APPROX FOR...

- Input: tree $T=(V,E)$, decomposed into $T_1, \dots, T_k$
  source-target pairs $(s_1,t_1), \dots, (s_k,t_k)$

- additional structure:
  all pairs are separated

- goal: find orientation that
  maximizes the number
  of connected pairs



CONSTANT-FACTOR APPROXIMATION

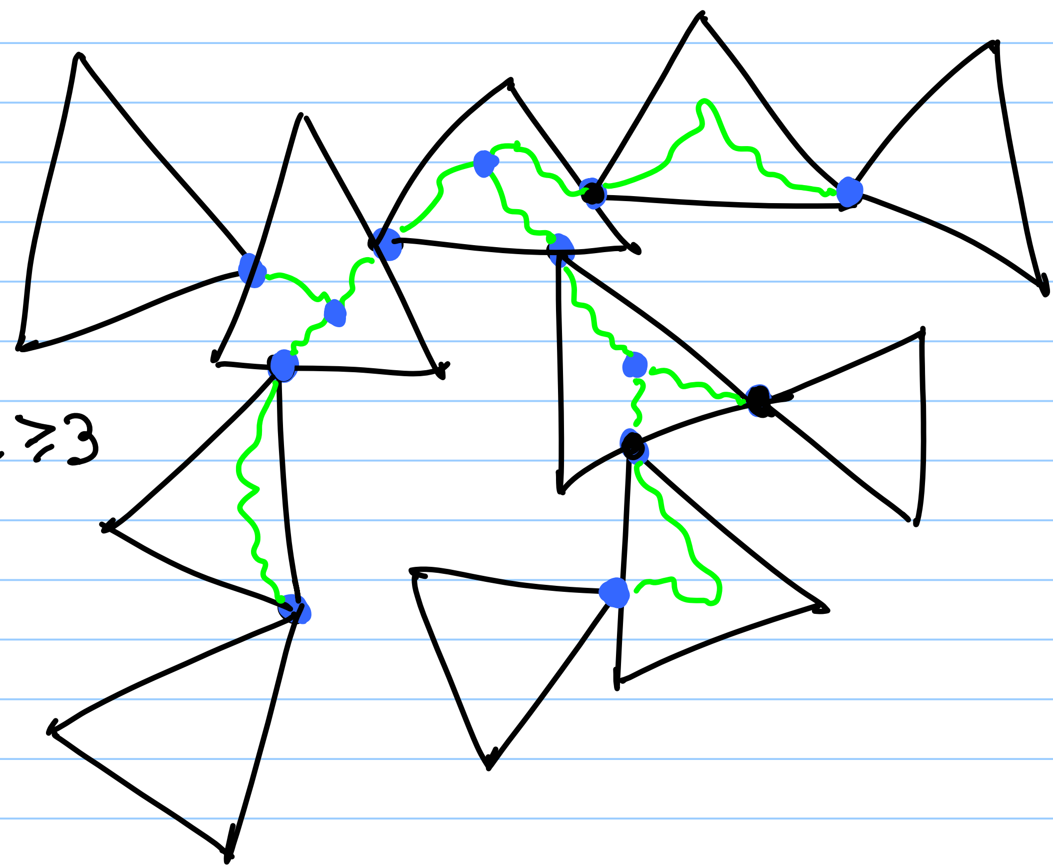
FOR

K-DECOMPOSITIONS

NOTATION AND TERMINOLOGY

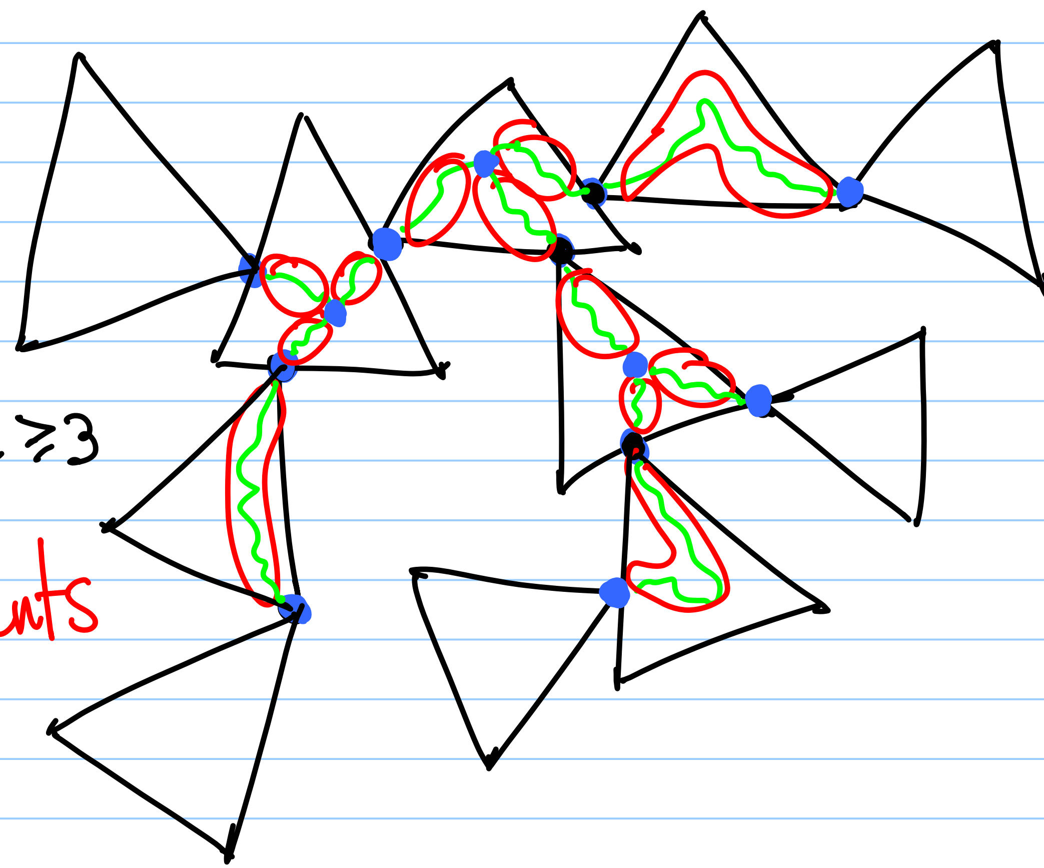
- skeleton $\mathcal{S}$ = minimal subtree that spans all intersection vertices

- special vertices = intersection vertices + vertices in $\mathcal{S}$ with degree ≥ 3



NOTATION AND TERMINOLOGY

- skeleton $\mathcal{S}$ = minimal subtree that spans all intersection vertices
- special vertices = intersection vertices + vertices in $\mathcal{S}$ with degree ≥ 3
- $\Sigma(\mathcal{S})$ = set of segments



SEGMENT GUESSING

In optimal orientation, each segment $\sigma \in \Sigma(\mathcal{S})$ can be

(1) right-oriented: $\square \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \square$

(2) left-oriented: $\square \longleftarrow \bullet \longleftarrow \bullet \longleftarrow \bullet \longleftarrow \square$

(3) mixed: $\square \longleftarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longleftarrow \square$

SEGMENT GUESSING

- assume that we know the optimal orientation of all skeleton segments $\sigma \in \Sigma(\xi')$
- number of guesses is:

$$3^{|\Sigma(\xi')|} = 3^{O(k)}$$

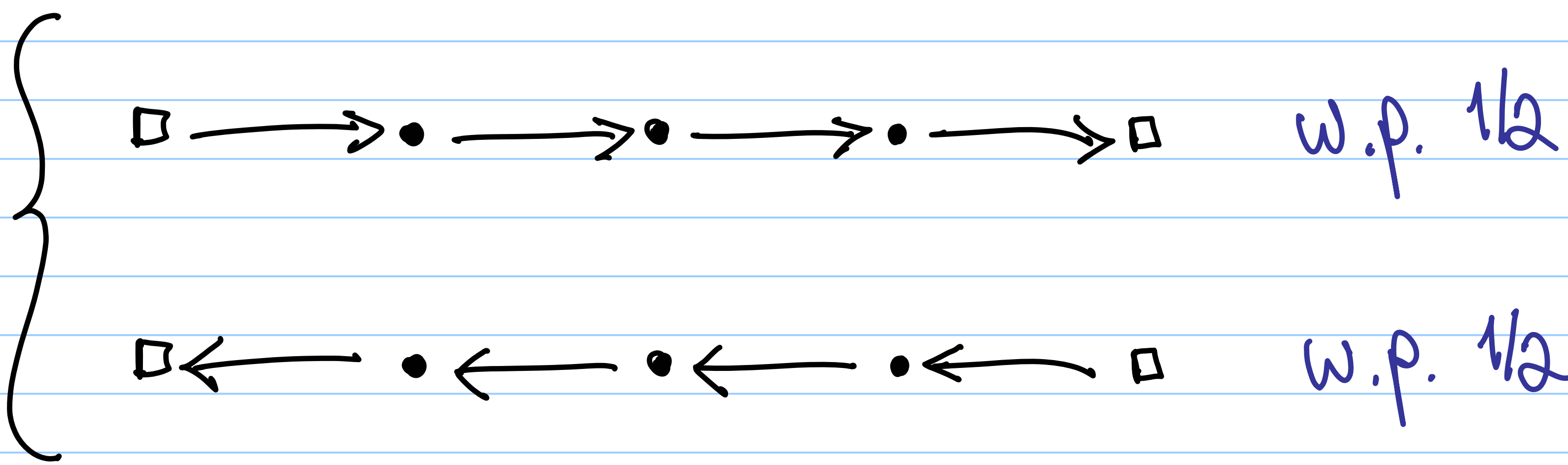
$$= 3^{O(\log n)}$$

$$= n^{O(1)}$$

THE RANDOMIZED ORIENTATION

(1) for skeleton segments:

- if right or left oriented, set accordingly
- if mixed, pick direction at random

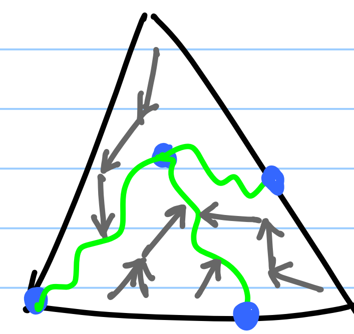


THE RANDOMIZED ORIENTATION

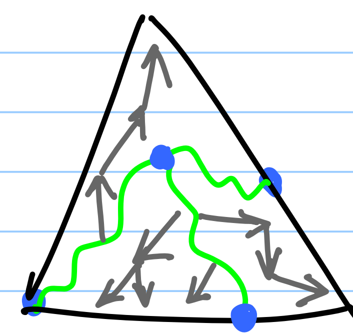
(2) for each of the decomposition trees:

- randomly pick whether SENDER or RECEIVER

- in RECEIVERS: all non-skeleton edges are directed toward the skeleton

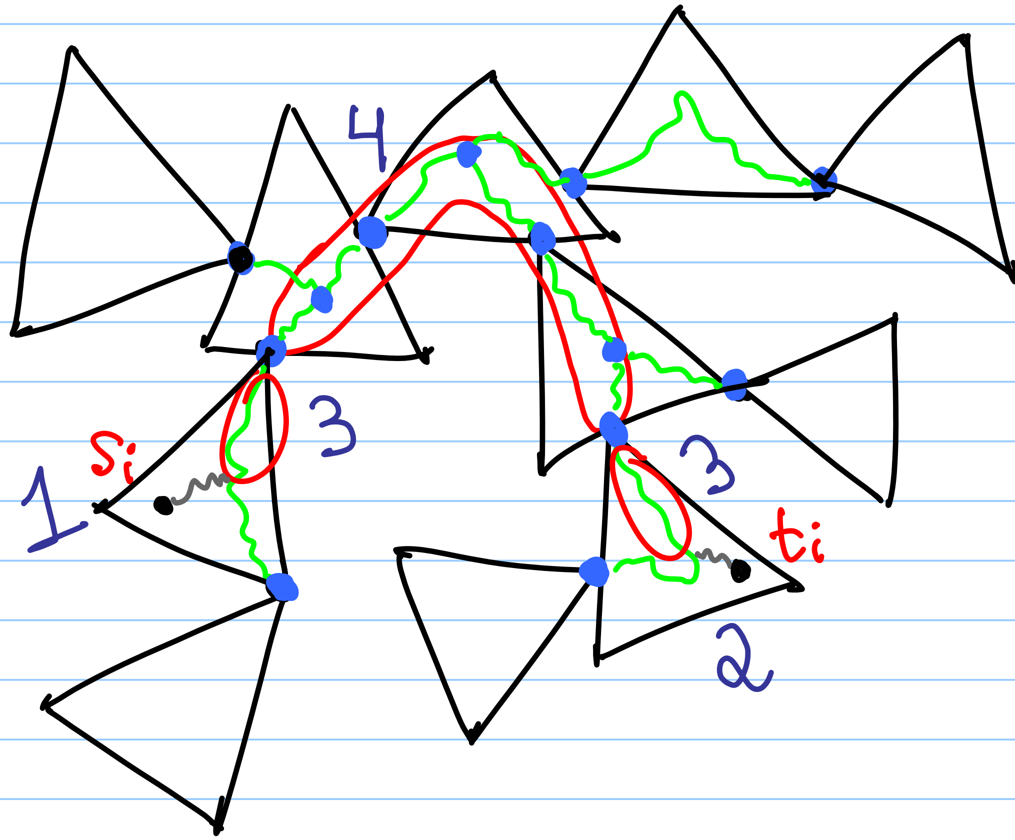


- in SENDERS: all non-skeleton edges are directed away from the skeleton



ANALYSIS

Claim: each pair connected by OPT is connected by the random orientation with probability $\geq 1/16$



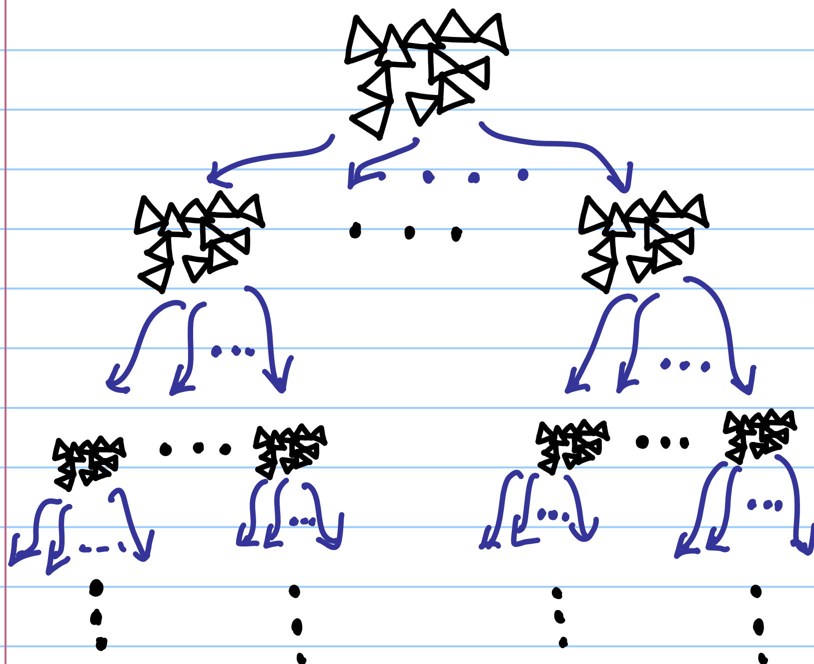
(1) SENDER w.p. 1/2

(2) RECEIVER w.p. 1/2

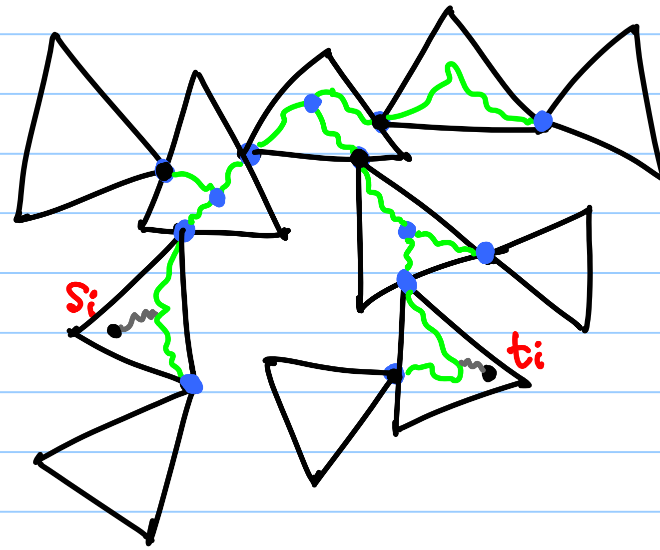
(3) as in OPT w.p. $\geq 1/2$

(4) as in OPT w.p. 1

QUICK RECAP



$\Theta\left(\frac{\log n}{\log\log n}\right)$

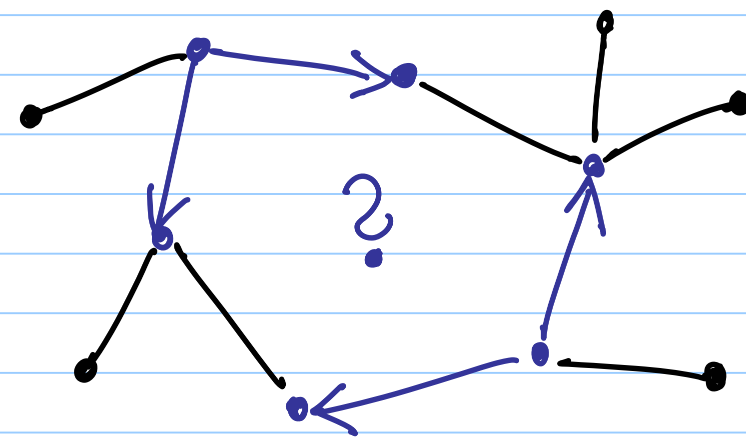


$O(1)$

can be derandomized

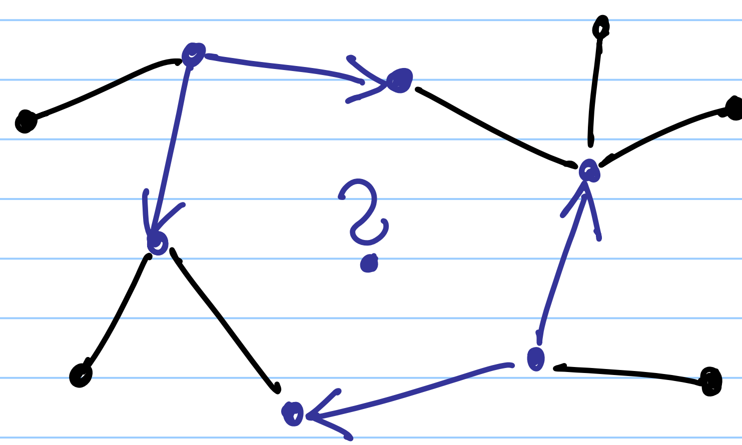
OPEN QUESTIONS

- improved approximation factor ?
- predetermined edge directions (mixed graphs) ?
  - reduction to trees does not work
  - $O(\log^M n)$-approx, where $M \approx$ number of PDI segments in the underlying pathways



OPEN QUESTIONS

- improved approximation factor ?
- predetermined edge directions (mixed graphs) ?
  - reduction to trees does not work
  - $O(\log^M n)$-approx, where
    $M \approx$ number of PDI segments
    in the underlying pathways



THANK YOU