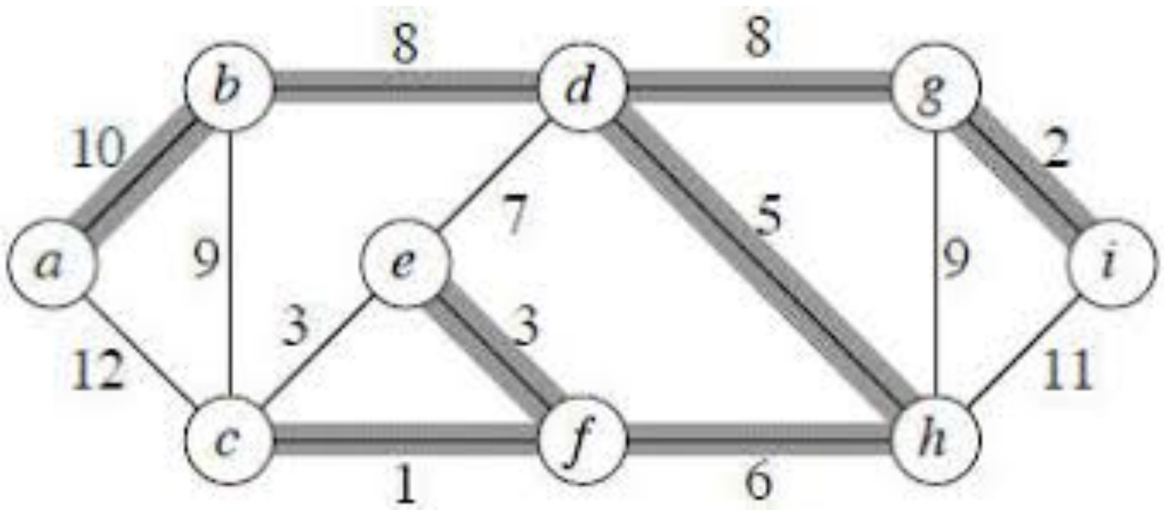
CONGEST Model (cont.)

- Global problems
- Minimum-weight Spanning Tree (MST)
  - Borůvska’s algorithm
  - Matroïd algorithm
  - Sublinear algorithm
- Lower bound for MST

Minimum Spanning Tree (MST)



Input of node u : ID(u), w(e) for every e ∈ E(u)

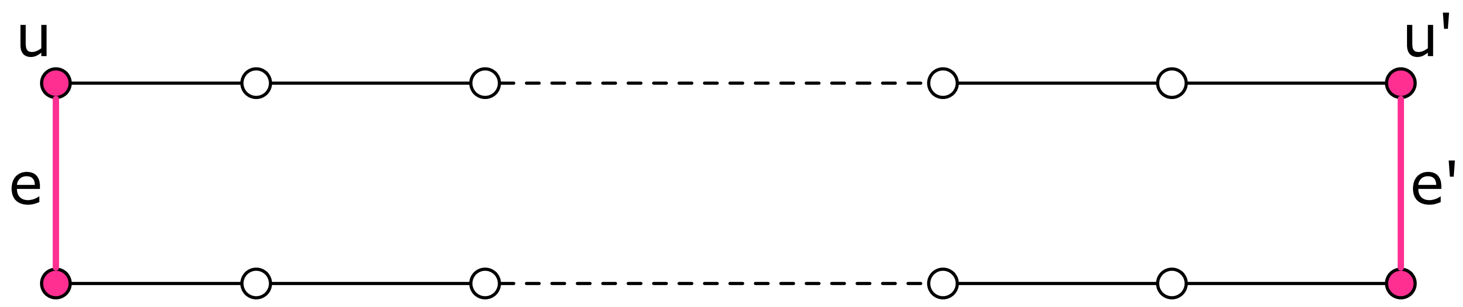
Output of node u : list of edges e ∈ E(u) belonging to MST

Facts about MST

Let G = (V,E) be a connected weighted graph

- Without loss of generality, all weights can be assumed distinct ➠ for every e = {u,v} with ID(u) > ID(v), replace w(e) by (w(e),ID(u),ID(v)).
- For every **cut** (S,V∖S) in G, the edge of **minimum** weight in the cut belongs to the MST.
- For every **cycle** C in G, the edge of **maximum** weight in C does not belong to the MST

MST is a non-local problem



$$I_1 = (1,3) \quad I_2 = (3,2) \quad I_3 = (1,2)$$

Remark MST requires at least D rounds in the cycle.

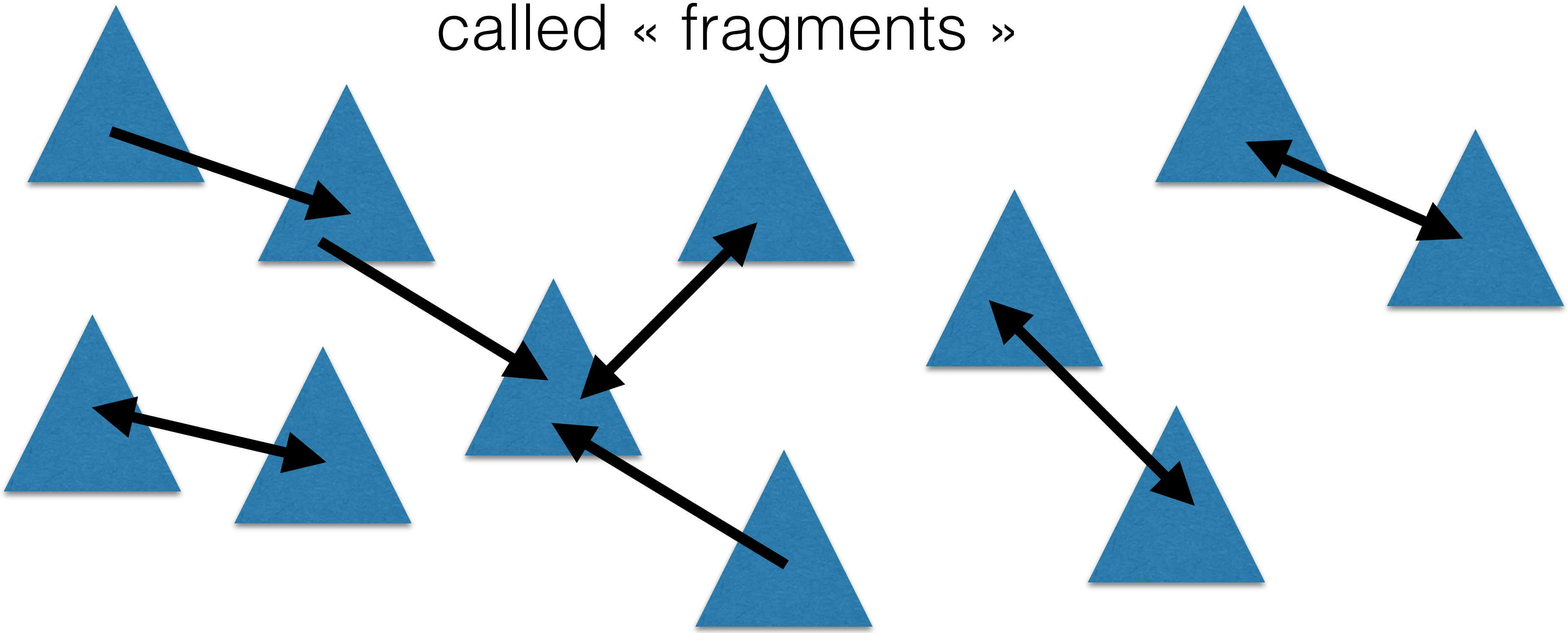
Algorithms with round-complexity O(f(n)+D)
in n-node graphs of diameter D.

Objective: minimizing f(n)

Borůvska's algorithm (1926)

distributed version

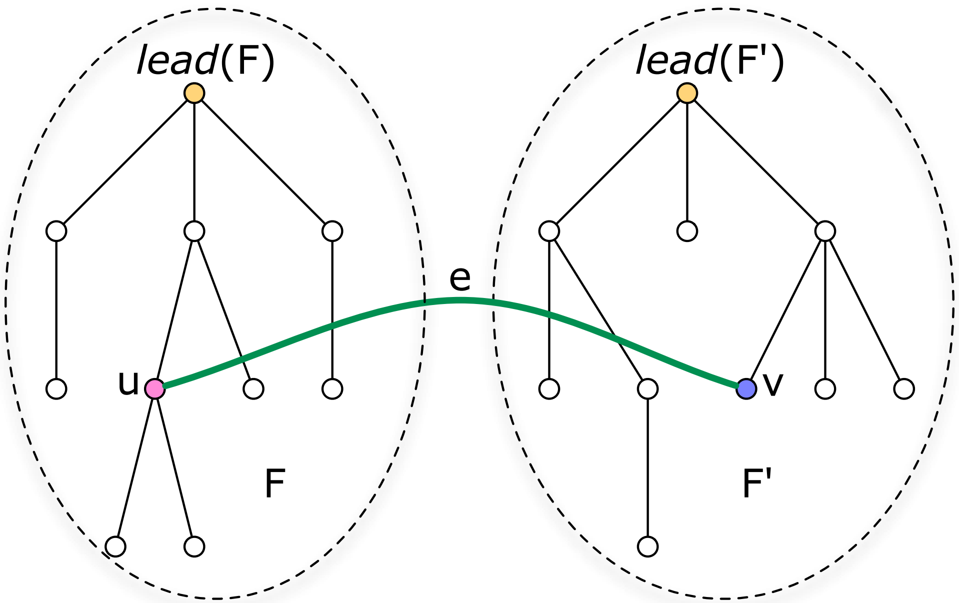
Collection of subtrees
called « fragments »

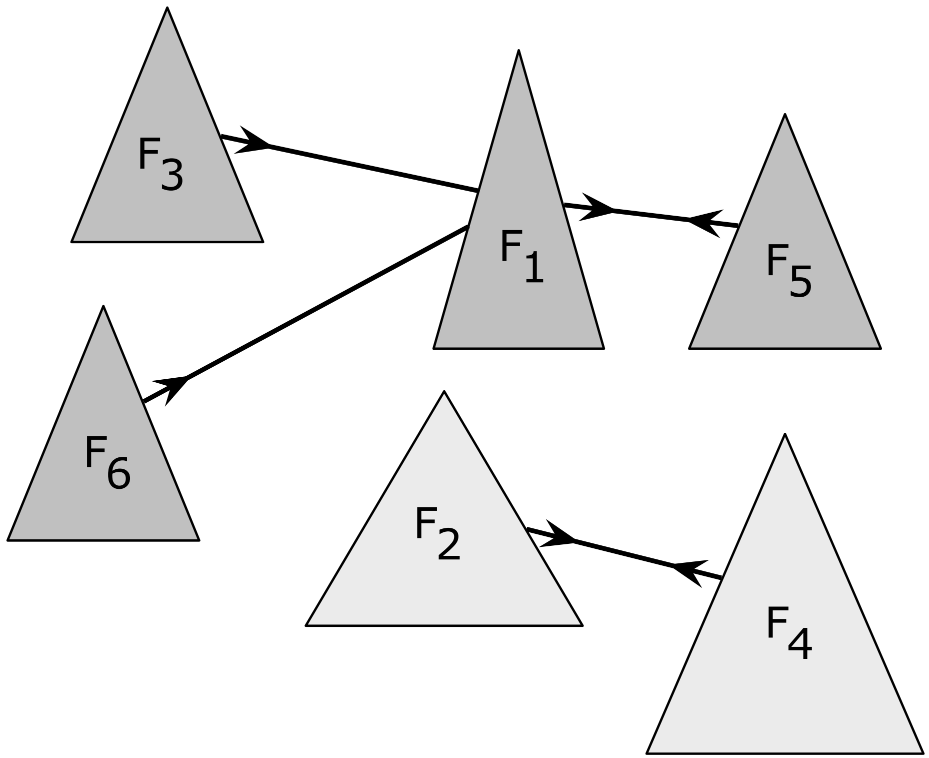


A phase = fragments
are merged

Merges use the edge of minimum
weight going out of each fragment

Fragments & Merging





$e = (ID(u), ID(v), w(e))$

$N(t)$ = #fragments after t phases

$N(0) = n$
$N(t+1) \leq N(t)/2$

$\Rightarrow$ at most $\lceil \log_2 n \rceil$ phases

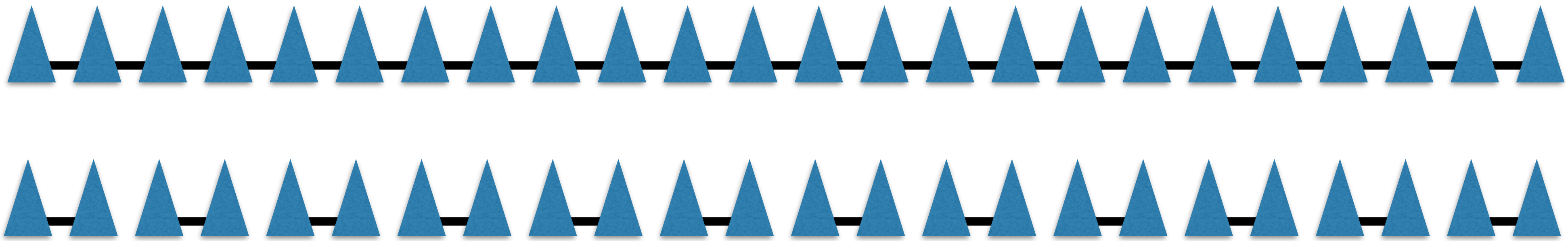
Round complexity

complexity of a phase = O(max$_F$ diam(F))

diam(F) ≤ n-1

Theorem The distributed version of Borůvska's algorithm can be implemented in O(n log n) rounds in the CONGEST model.

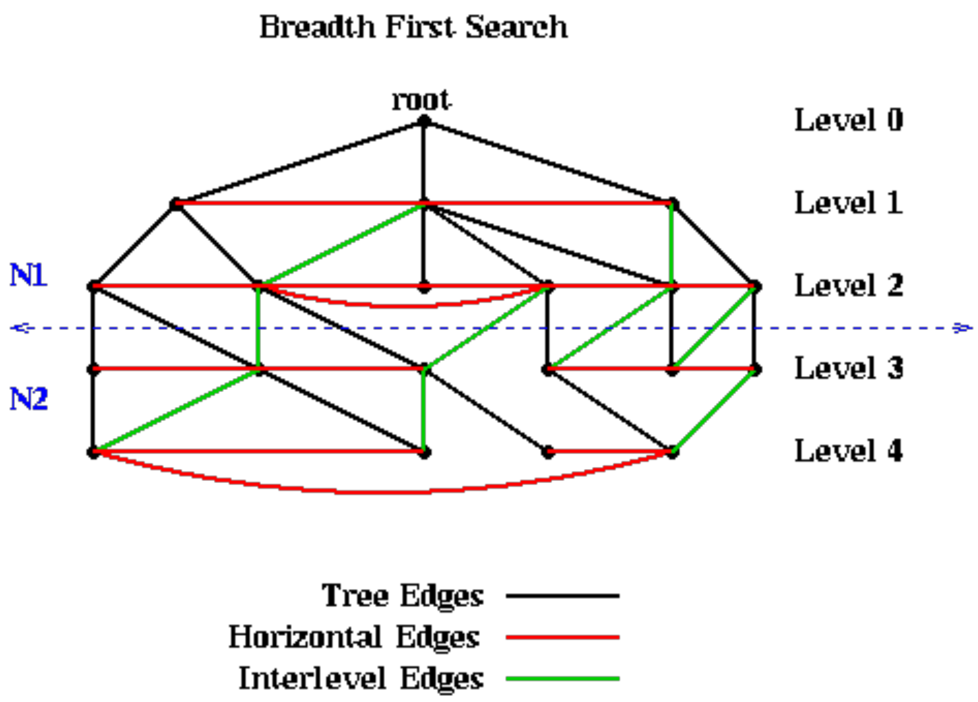
The bound is tight:



Another MST algorithm

Based on a **Breadth-First Search (BFS)** tree

Lemma BFS construction requires $O(D)$ rounds in the CONGEST model



Algorithm of node u

$\mathtt{id}_{min} \leftarrow ID(u)$

repeat

send $\mathtt{id}_{min}$ to neighbors, and receive IDs from neighbors

if $\exists$ $\mathtt{id} \in$ {IDs sent by neighbors} : $\mathtt{id} < \mathtt{id}_{min}$ **then**

$\mathtt{id}_{min} \leftarrow \mathtt{id}$

parent(u) ← ID(v) where v is the neighbor which sent $\mathtt{id}$

Matroid Algorithm (1)

Algorithm for a node u

$K \leftarrow E(u)$ edges incident to node u
wait until having received an edge from each child
repeat
know $K \leftarrow K \cup \{\text{received edges}\}$
up $U \leftarrow \{\text{edges previously sent to parent}(u)\}$
remove $R \leftarrow \{e \in K \setminus U : U \cup \{e\} \text{ contains a cycle}\}$
candidate $C \leftarrow K \setminus (U \cup R)$
if $C \neq \varnothing$ **then**
send $e \in C$ with minimum weight to parent
receive edges from children
else terminate

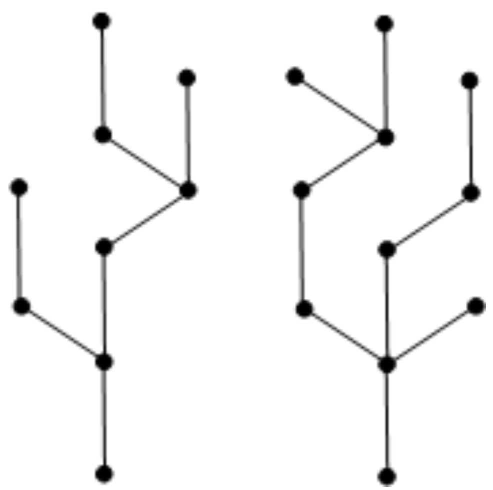
Proof of correctness

Theorem The Matroid algorithm performs in $O(n + D)$ rounds in the CONGEST model, and enables the root of the tree to construct a MST.

Lemma 0 Let A and B be acyclic subsets of edges. If $|A|>|B|$ then there exists $e \in A \setminus B$ such that $B \cup \{e\}$ is acyclic.

This is a matroid axiom

Proof B is a forest $\{T_1,\ldots,T_k\}$. Let $n_i = |V(T_i)|$. We have $|E(T_i)|=n_i-1$.



For every i, there are at most n_i-1 edges of A connecting nodes in T_i.

➡ There is an edge in A whose extremities do not belong to a same tree T_i. ❑

A node u is said **active** at phase t if it has not terminated at phase $t - 1$.

Let h(u) = height of u = length of longest path to a leaf of the subtree T_u rooted at u.

Lemma 1 For every active child v of a node u, the set C of candidates for u at time t contains at least one edge sent by v to u before time t. ➪ **no premature termination**

Proof Induction on h(u). Lemma holds for h(u)=0.

Assume lemma hold for all nodes at height $\leq$ k.

Let u with h(u)=k+1, and v active child of u. Note h(v) $\leq$ k.

Let E_u and E_v be edges sent by u to p(u), and by v to u=p(v) before phase t.

Since h(v)<h(u) we have $|E_v| > |E_u|$.

By Lemma 0, $\exists\ e \in E_v \setminus E_u$ such that $E_u \cup \{e\}$ is acyclic ➡ $e \in C$. ❑

Lemma 2

(a) If u sends e to p(u) at phase t then

1. all edges received by u at phase t-1 from its active children were of weight ≥ w(e), and
2. all edges to be received by u at phases ≥ t will be of weight ≥ w(e).

(b) The weights of the edges sent by u to its parent are ↗

➜ it is legitimate to remove edges creating cycles with previously sent edges.

Proof True for height 0. Assume holds for height k.

(a.1) Let u with h(u) = k+1.

Let e' be edge sent by child v at phase t-1.

Let e'' ∈ C whose existence follows from Lemma 1.

By induction, property (b) implies w(e'') ≤ w(e').

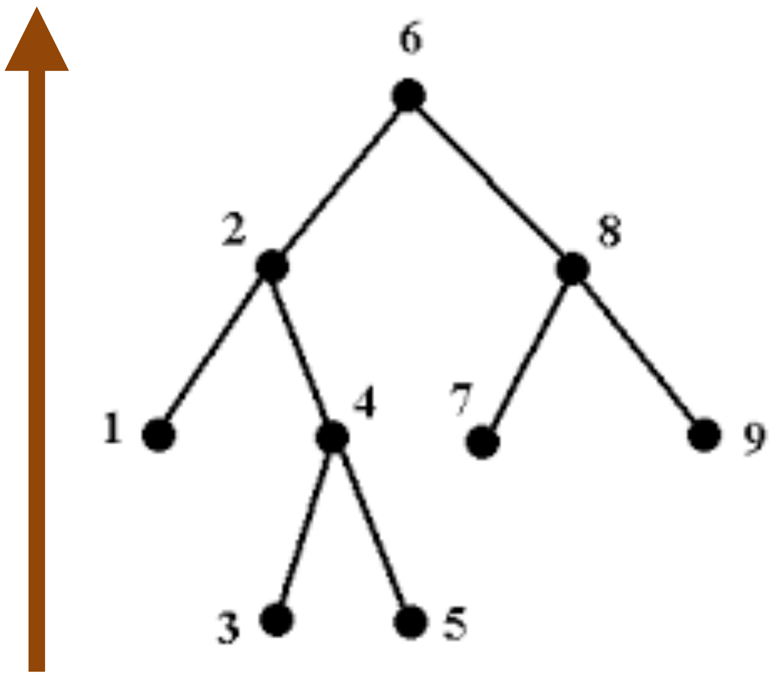
By the choice of the edge in C, we have w(e) ≤ w(e'').

➥ w(e') ≥ w(e).

(a.2) follows from (a.1) and by induction from (b).

(b) follows from (a.2) by the choice of the edge in C. ❑

Complexity

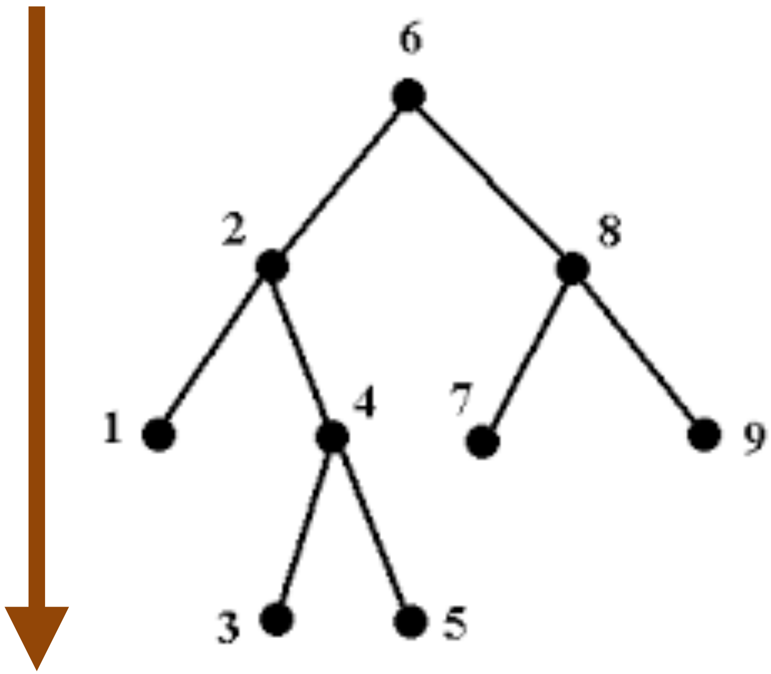


In n-node graphs, any set of n edges includes a cycle

➥ every node sends ≤ n-1 edges

➥ #rounds ≤ D + n - 1

Broadcasting the MST from the root to all nodes



Pipelining the edges of $T = \{e_1, e_2, \ldots, e_{n-1}\}$ down the BFS tree

➥ #rounds $\leq D + n - 1$

Wrap Up

- **Borůvska:** O(n log n) rounds — this is because fragments can have arbitrarily large diameter
- **Matroid:** O(D+n) rounds — this is because too many edges are gathered at a single node.
- **Combining Borůvska and Matroid:**
  - control the diameter of the fragment, and stops when fragments have too large diameter
  - carry on with matroid for computing the (few) edges connecting the fragments already computed by Borůvska

Tool

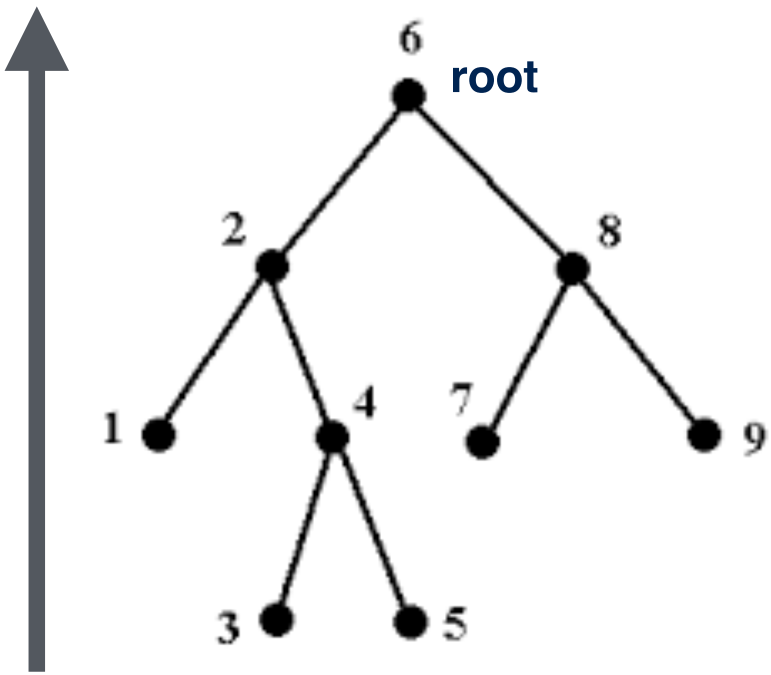
- $D \subseteq V$ is a dominating set if every $u \notin D$ has a neighbor in D.

- Remarks:
  - Every maximal independent set is a dominating set.
  - Every tree has a dominating set of size $\leq n/2$

- **Objective:** Distributed computing of a dominating set of size $\leq n/2$ in consistently oriented trees.

MIS in Rooted Trees

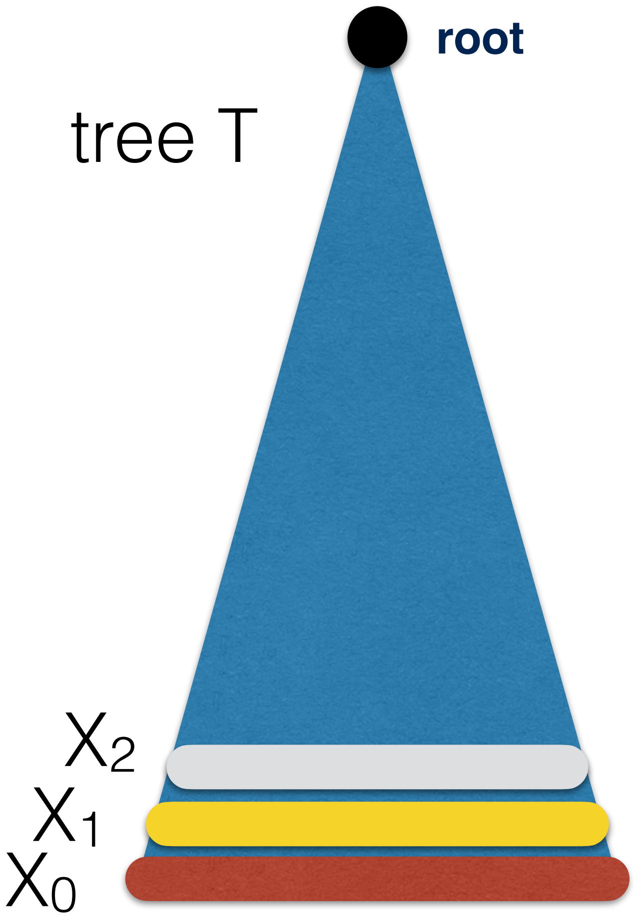
Every node has pointer to its parent



- Perform Cole and Vishkin algorithm with parent
- When colors are on 3 bits, every node pushes down its color
- Performs 5 rounds to get all colors in {1,2,3}.

Complexity : O(log*n) rounds

Computing small dominating sets in rooted trees

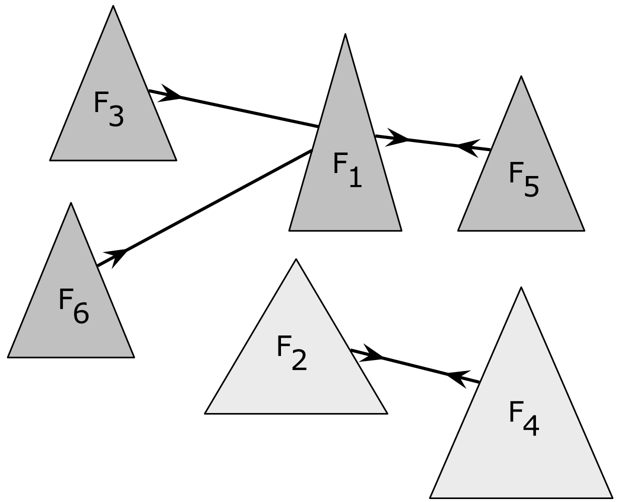


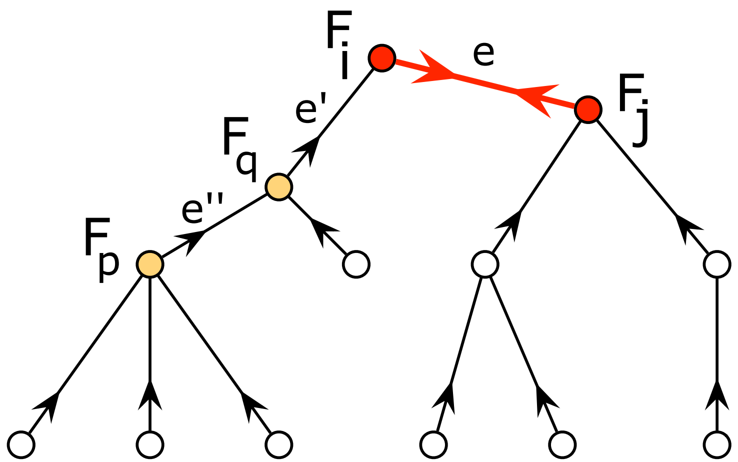
- X_d = {nodes at distance d from a leaf}
- $Y = V(T) \setminus (X_0 \cup X_1 \cup X_2)$
- Let J be MIS in Y (comput. in $O(\log^* n)$ rounds)
- Let $D = J \cup X_1$

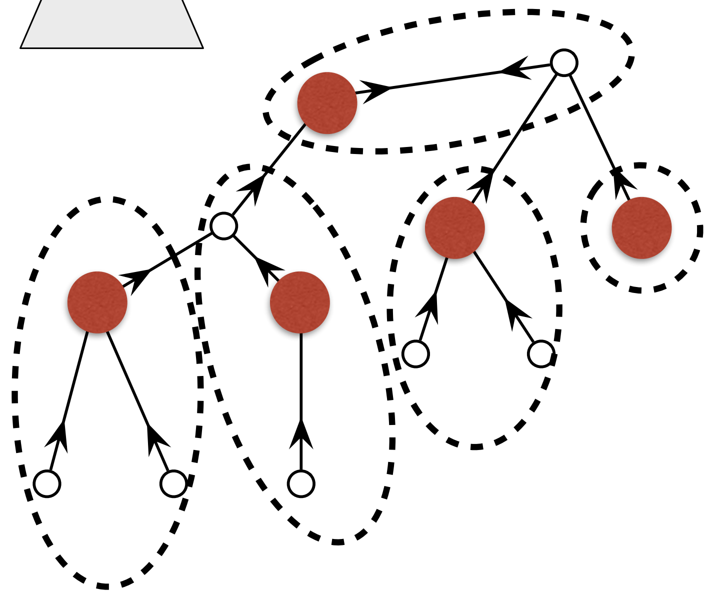
- D is a dominating set
- $|X_1| \leq |X_0| \Rightarrow |X_1| \leq \frac{1}{2} |X_0 \cup X_1|$
- $|J| \leq |(Y \cup X_2) \setminus J| \Rightarrow |J| \leq \frac{1}{2} |Y \cup X_2|$

➥ $|D| \leq n/2$

Bounding the diameter of fragments







Fast MST algorithm

Two stages:

1. Few phases of Borůvska
2. Completed by Matroid

$N(t) \leq N(t-1)/2$

$\Longrightarrow N(t) \leq n/2^t$

$\mathrm{diam}(t) \leq 3\ \mathrm{diam}(t-1) + 2$

$\Longrightarrow \mathrm{diam}(t) \leq 3^t - 1$

N(t) = #frags after t phases

diam(t) = max diameter frags

Phase t costs

$O(\mathrm{diam}(t) \log^* n)$ rounds

τ phases Borůvska costs

$\tilde{O}(3^\tau)$ rounds

Matroid completes in

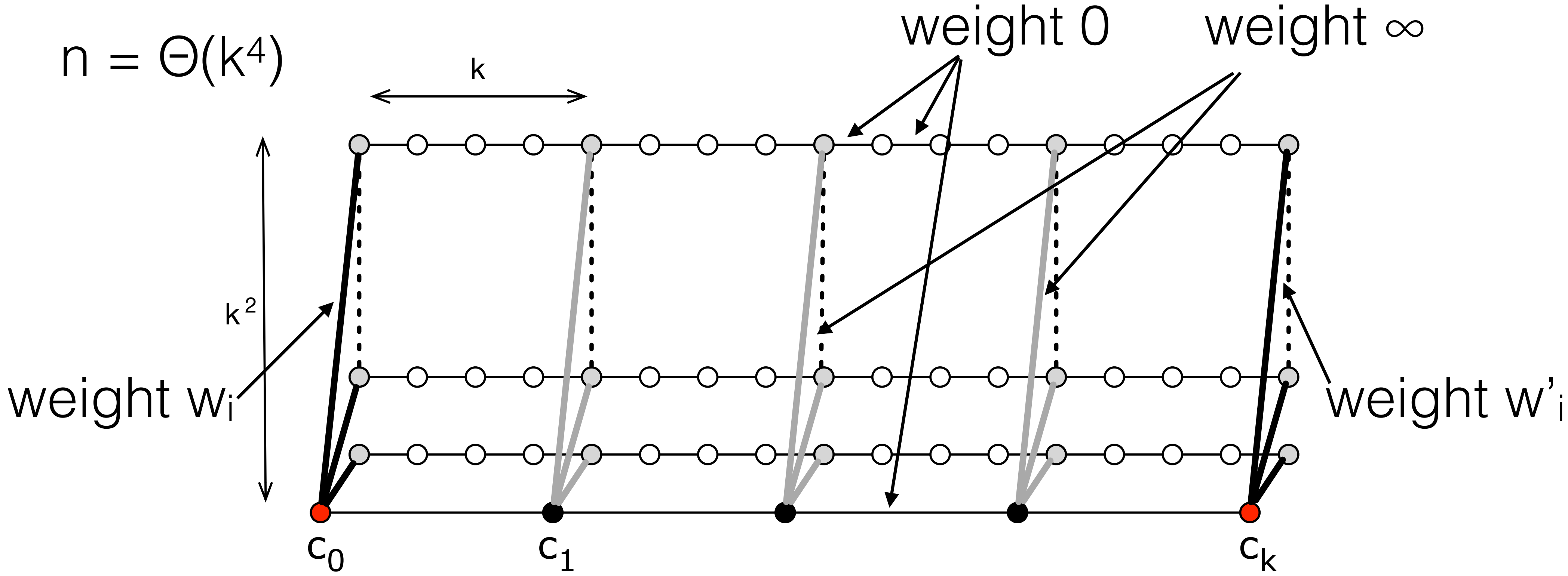
$O(D+N(\tau))$ rounds

$3^\tau = n/2^\tau \Longrightarrow \#\text{rounds} = \tilde{O}(D + n^{0.6131})$

Theorem MST construction can be achieved in $\tilde{O}(D + \sqrt{n})$ rounds in the CONGEST model.

$\Omega(\sqrt{n})$ lower bound for MST

$n = \Theta(k^4)$



$w_i = 2$ $w'_i = 2x_i + 1$ with $x_i \in \{0,1\}$

Lemma Transmitting k^2 bits from c_k to c_1 takes $\Omega(k^2)$ rounds

Proof (simplified: no recombination)

- $\exists$ i, x_i uses $\leq k/2$ of highway ⇨ $\Omega(k \cdot k/2)$ rounds
- $\forall$ i, x_i uses $> k/2$ of highway ⇨ $\Omega((k^2 \cdot k/2)/(k \log n))$ rounds ☐

End Lecture 6