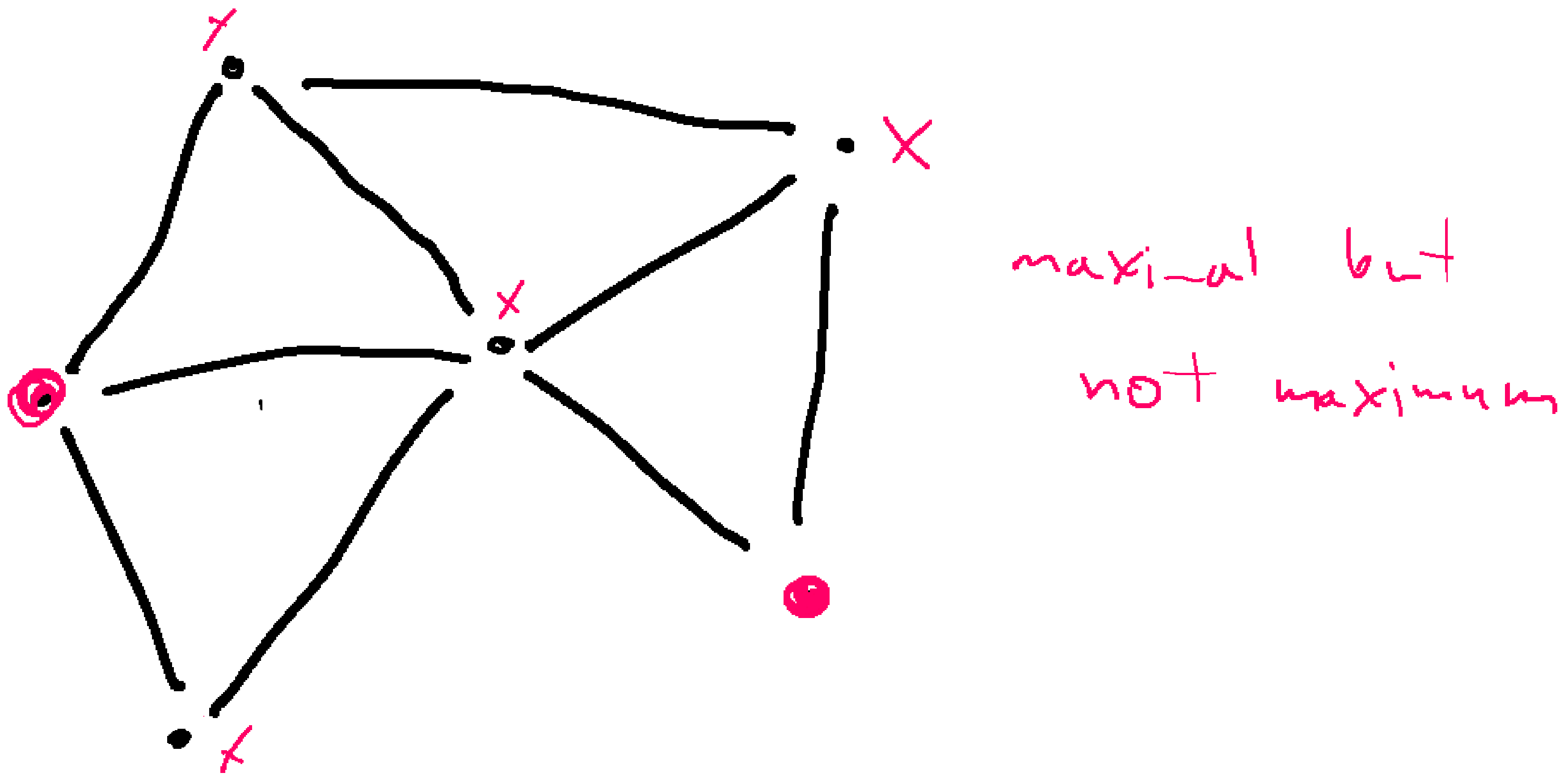
Maximal Independent Set

2021

Independent set in a graph



Maximum independent set

independent set of maximum cardinality

hard problem, even sequential

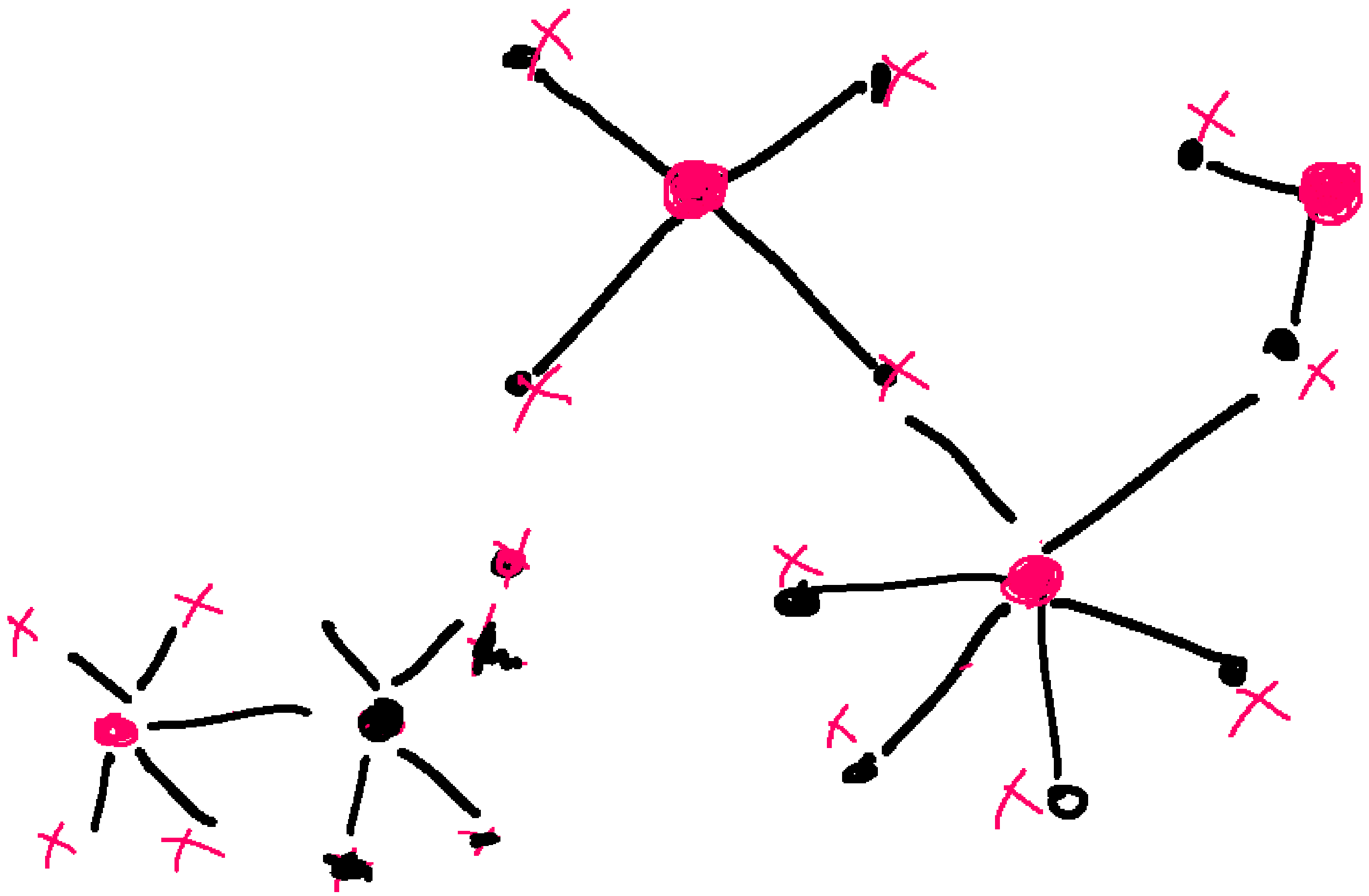
NP complete

Maximal independent set

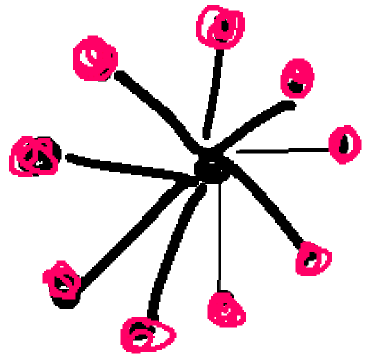
ind set such that it is not a proper subset of another i.s

easy

Greedy sequential algorithm



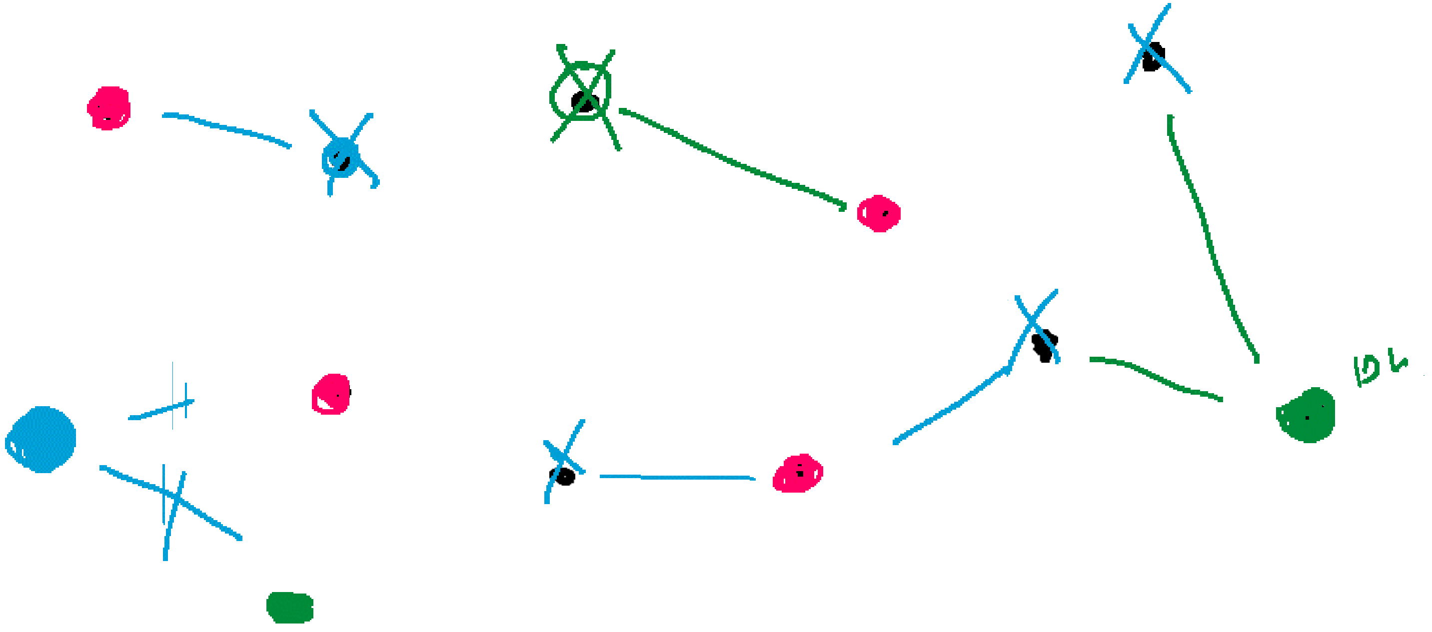
it stops
sequential ?
quality



MIS via graph coloring

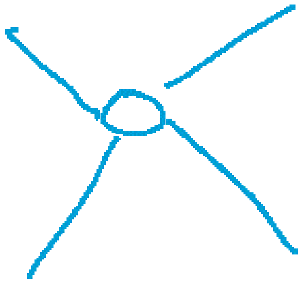
Corollary 7.5. *Given a coloring algorithm that runs in time T and needs C colors, we can construct a MIS in time $T + C$.*

nodes of color A are independent



time

step1: activate if red
and check neighbors



return

step2: activate if green
and check neighbors

return

local

Algorithm 7.6 Fast MIS

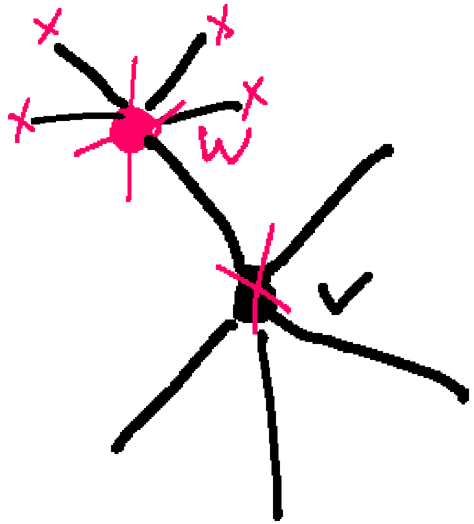
The algorithm operates in synchronous rounds, grouped into phases.

A single phase is as follows:

1) Each node v marks itself with probability $\frac{1}{2d(v)}$, where $d(v)$ is the current degree of v.

2) If no higher degree neighbor of v is also marked, node v joins the MIS. If a higher degree neighbor of v is marked, node v unmarks itself again. (If the neighbors have the same degree, ties are broken arbitrarily, e.g., by identifier).

3) Delete all nodes that joined the MIS and their neighbors, as they cannot join the MIS anymore.



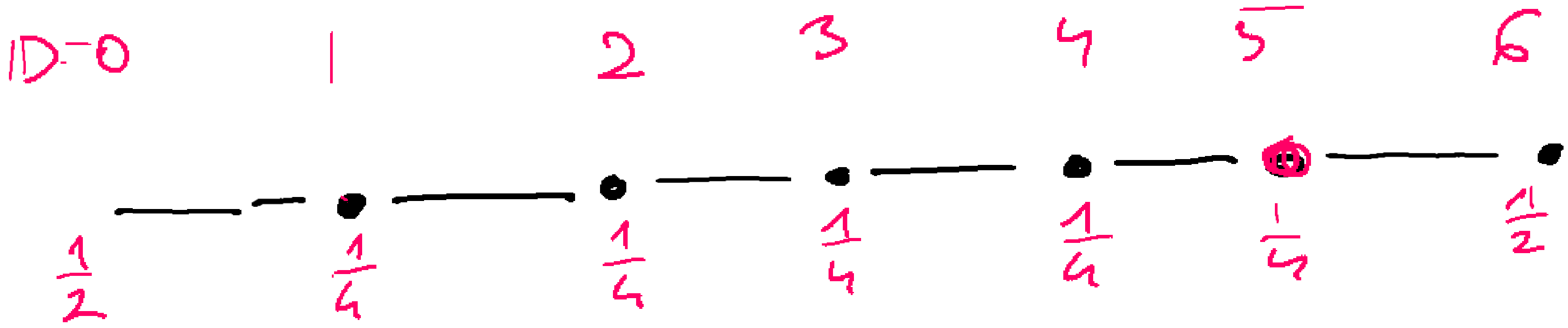
deg(v) = 5

marked with pbb $\frac{1}{2 \cdot 5}$

Correctness

- always no neighbors added
- elements: neighbors of MIS members are removed

⇝ like greedy



goal: to show that runtime is low with high probability

randomized algorithm

Lemma 7.7 (Joining MIS). *A node v joins the MIS in Step 2 with probability $p \geq \frac{1}{4d(v)}$.*

M = marked nodes

$$
\begin{aligned}
P[v \notin \text{MIS} | v \in M] &= P[\text{there is a node } w \in H(v), w \in M | v \in M] \\
&= P[\text{there is a node } w \in H(v), w \in M] \\
&\leq \sum_{w \in H(v)} P[w \in M] = \sum_{w \in H(v)} \frac{1}{2d(w)} \\
&\leq \sum_{w \in H(v)} \frac{1}{2d(v)} \leq \frac{d(v)}{2d(v)} = \frac{1}{2}.
\end{aligned}
$$

$$
P[v \in \text{MIS}] = P[v \in \text{MIS} | v \in M] \cdot P[v \in M] \geq \frac{1}{2} \cdot \frac{1}{2d(v)}.
$$

Lemma 7.8 (Good Nodes). *A node v is called* good *if*

$$\sum_{w \in N(v)} \frac{1}{2d(w)} \geq \frac{1}{6},$$

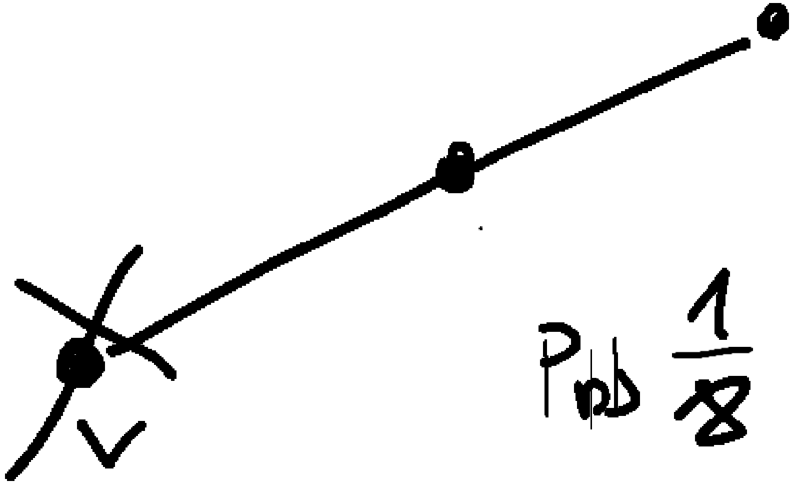
where $N(v)$ is the set of neighbors of v. Otherwise we call v a bad *node. A good node will be removed in Step 3 with probability* [illegible].

joining MIS

good nodes — fraction $\frac{1}{36}$ should go

bad nodes - 2

Case 1 : There is a neighbor of degree 2 or 1



Prob $\frac{1}{8}$ join IS

$$\frac{1}{8} > \frac{1}{36}$$

Case 2 : all neighbors have degree >2

For any neighbor w of v we have $\frac{1}{2d(w)} \leq \frac{1}{6}$. Since $\displaystyle\sum_{w \in N(v)} \frac{1}{2d(w)} \geq \frac{1}{6}$ there is a subset of neighbors $S \subseteq N(v)$ such that $\displaystyle\frac{1}{6} \leq \sum_{w \in S} \frac{1}{2d(w)} \leq \frac{1}{3}$

$$\frac{1}{2\,d(w)} \leq \frac{1}{6}$$

$$\frac{1}{6}$$

$$\frac{1}{3} < \frac{1}{2d(w_1)} + \frac{1}{2\,d(w_2)} + \cdots \geq \frac{1}{6}$$

S - fixed

R – event that v removed

$u_1, u_2, u_3, \ldots$

$$\begin{aligned} P[R] &\geq P[\text{there is a node } u \in S, u \in \text{MIS}] \\ &\geq \sum_{u \in S} P[u \in \text{MIS}] - \sum_{u,w \in S; u \neq w} P[u \in \text{MIS} \textbf{ and } w \in \text{MIS}]. \end{aligned}$$

MIS MIS

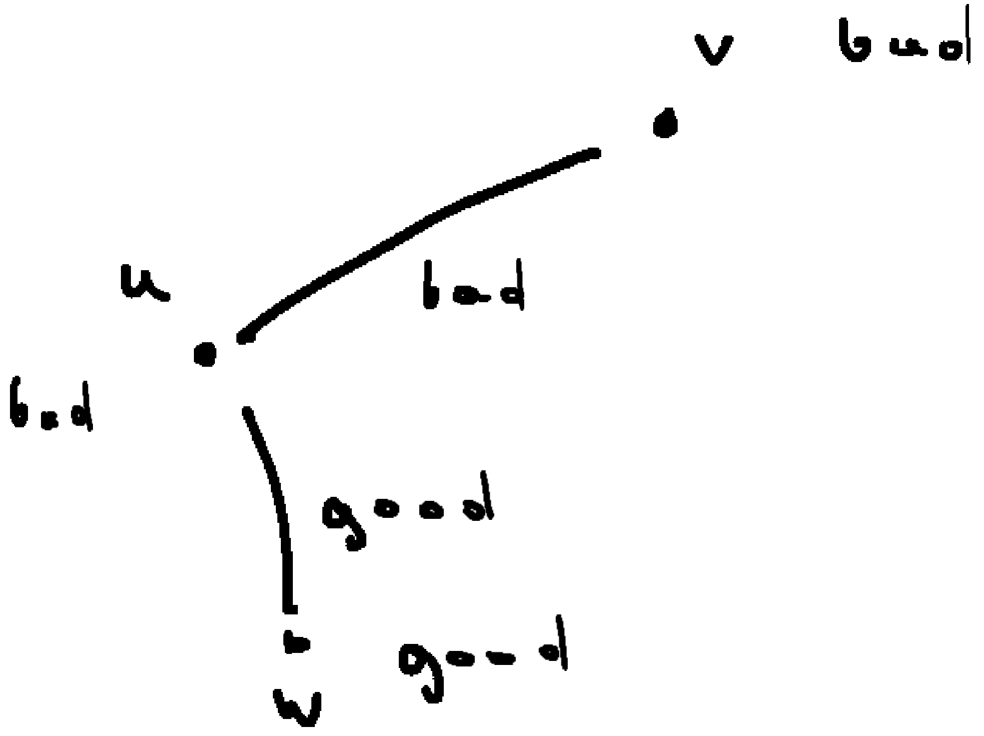
$$
\begin{aligned}
P[R] &\geq \sum_{u \in S} P[u \in \mathrm{MIS}] - \sum_{u,w \in S; u \neq w} P[u \in M \text{ and } w \in M] \\
&\geq \sum_{u \in S} P[u \in \mathrm{MIS}] - \sum_{u \in S} \sum_{w \in S} P[u \in M] \cdot P[w \in M] \\
&\geq \sum_{u \in S} \frac{1}{4d(u)} - \sum_{u \in S} \sum_{w \in S} \frac{1}{2d(u)} \frac{1}{2d(w)} \\
&\geq \sum_{u \in S} \frac{1}{2d(u)} \left(\frac{1}{2} - \sum_{w \in S} \frac{1}{2d(w)} \right) \geq \frac{1}{6} \left(\frac{1}{2} - \frac{1}{3} \right) = \frac{1}{36}.
\end{aligned}
$$

□

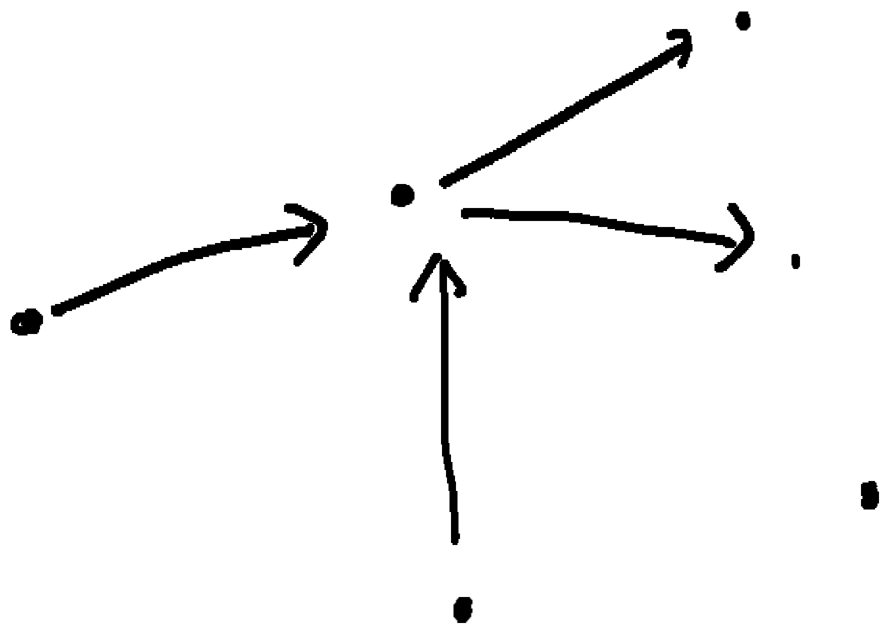
Good nodes die out

We cannot prove that there is always a certain fraction of good nodes

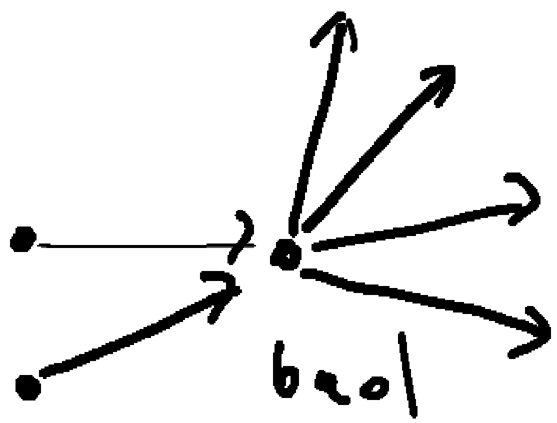
Lemma 7.9 (Good Edges). *An edge $e = (u, v)$ is called* bad *if both u and v are bad; else the edge is called* good. *The following holds: At any time at least half of the edges are good.*



Auxiliary graph -- directed edges towards higher degree

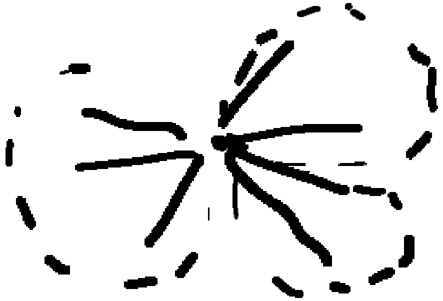


Lemma 7.10. *A bad node has outdegree (number of edges pointing away from bad node) at least twice its indegree (number of edges pointing towards bad node).*



By contradiction, let more than 1/3 neighbors have not smaller degree

Set S of such nodes

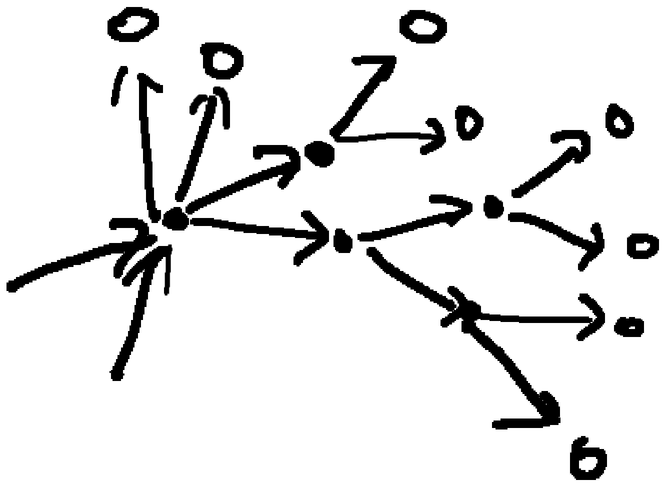


$\frac{1}{3}$ neigh. have deg. $\leq$ deg(v)

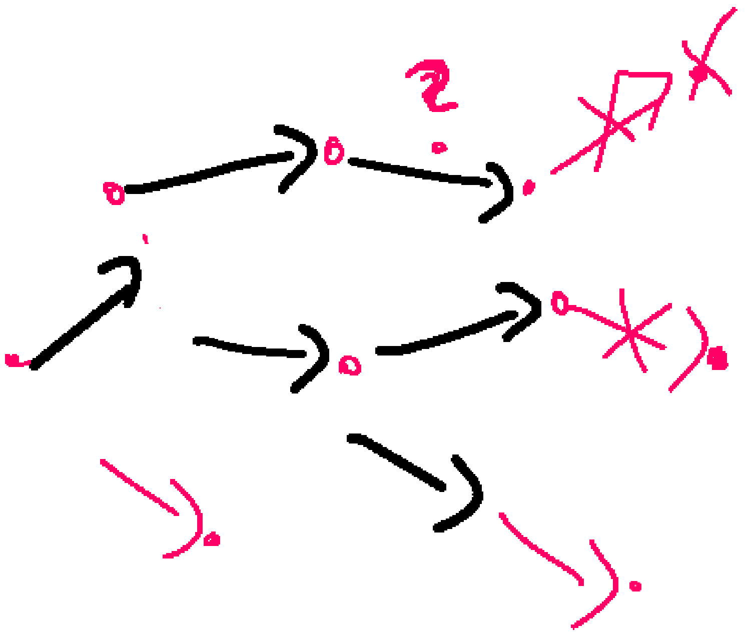
$$\sum_{w\in N(v)} \frac{1}{2d(w)} \geq \sum_{w\in S} \frac{1}{2d(w)} \geq \sum_{w\in S} \frac{1}{2d(v)} \geq \frac{d(v)}{3}\frac{1}{2d(v)} = \frac{1}{6}$$

S [illegible] good!?

Corollary: number of edges pointing to bad nodes is at most half



Theorem 7.11 (Analysis of Algorithm 7.6). *Algorithm 7.6 terminates in expected time $\mathcal{O}(\log n)$.*



$\frac{1}{36}$

50% ≥ good edges

fraction dies out

start $\frac{n^2}{2} \leq$ edges

$\log_c \frac{n^2}{2} = \frac{?}{?}(2\log n - 1)$

Algorithm 7.12 Fast MIS 2

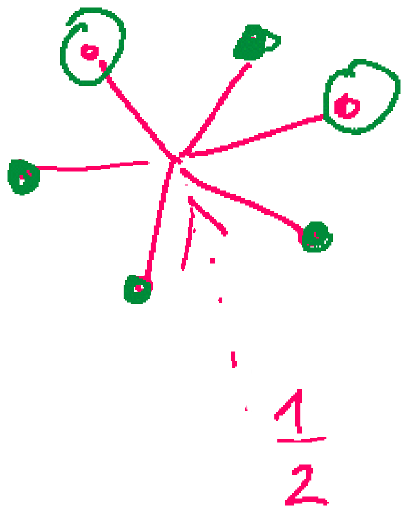
The algorithm operates in synchronous rounds, grouped into phases.

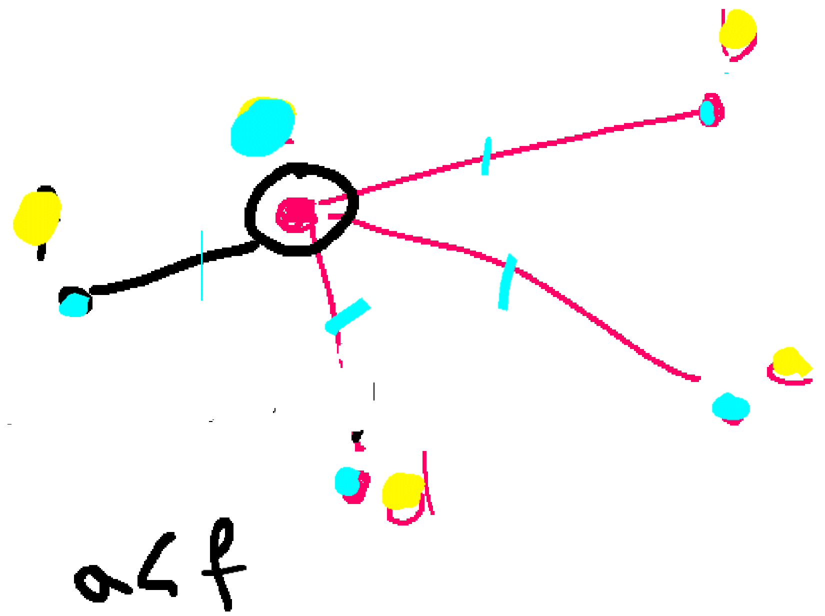
A single phase is as follows:

1) Each node v chooses a random value $r(v) \in [0,1]$ and sends it to its neighbors.

2) If $r(v) < r(w)$ for all neighbors $w \in N(v)$, node v enters the MIS and informs its neighbors.

3) If v or a neighbor of v entered the MIS, v terminates (v and all edges adjacent to v are removed from the graph), otherwise v enters the next phase.





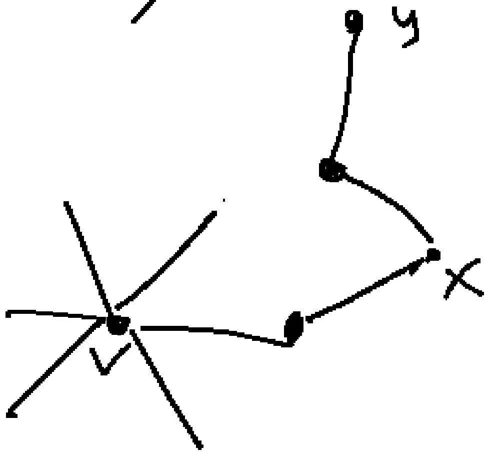
Lemma 7.14 (Edge Removal). *In a single phase, we remove at least half of the edges in expectation.*

$[v \to w]$

$$\#E \leq \frac{n(n-1)}{2} \approx n^2$$

$\approx \log(n^2)$ steps to reach $O(1)$

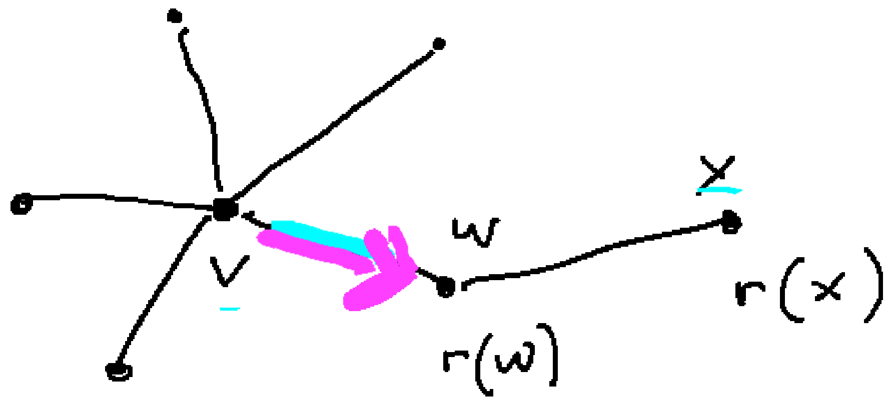
$= 2 \lg n$



$r(x) < r(v)$

$r(y) < r(x)$

Suppose that a node v joins the MIS in this phase, i.e., $r(v) < r(w)$ for all neighbors $w \in N(v)$. If in addition we have $r(v) < r(x)$ for all neighbors x of a neighbor w of v, we call this event $(v \to w)$. The probability of event $(v \to w)$ is at least $1/(d(v) + d(w))$, since $d(v) + d(w)$ is the maximum number of nodes

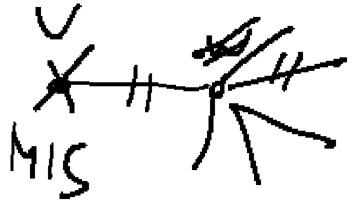


1. $r(x) > r(v)$
2. $r(x) < r(v)$

1) $r(v) < r(w)$ ∀ neighbors w

2) ∀ neigh. x of w $r(v) < r(x)$

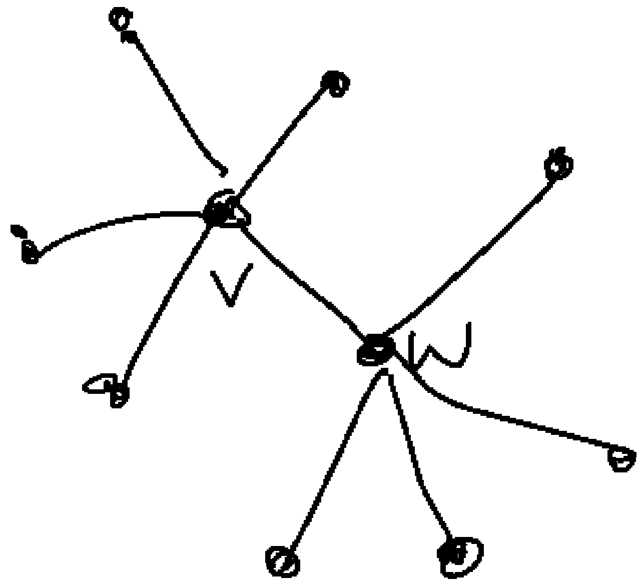
$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{\{v,w\}\in E} \mathbb{E}[X_{(v\to w)}] + \mathbb{E}[X_{(w\to v)}] \\
&= \sum_{\{v,w\}\in E} P\left[\text{Event } (v \to w)\right] \cdot d(w) + P\left[\text{Event } (w \to v)\right] \cdot d(v) \\
&\geq \sum_{\{v,w\}\in E} \left(\frac{d(w)}{d(v)+d(w)} + \frac{d(v)}{d(w)+d(v)} \right) \\
&= \sum_{\{v,w\}\in E} 1 = |E|.
\end{aligned}
$$



$X =$ # such edges removed

$$
X_{(v\to w)} = \begin{cases} 0 & \text{if not } (v \to w) \\ \#\text{ edges removed because of } v \text{ joining MIS} \end{cases}
$$

$$
X = \sum_{(v,w)\in E} X_{(v\to w)}
$$



$[v \to w]$

$r(v)$ is the smallest one

$d(v) + d(w)$

$$\frac{1}{d(v) + d(w)}$$

Stochastic process

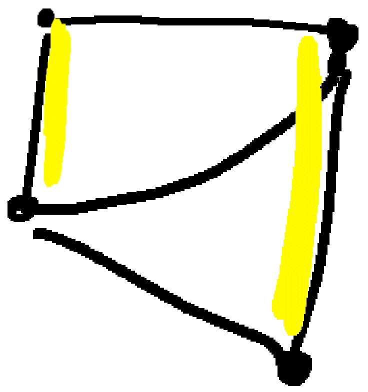
n^2 tokens

kill $\geq \frac{1}{2}$ of ~~tokens~~ at expectation

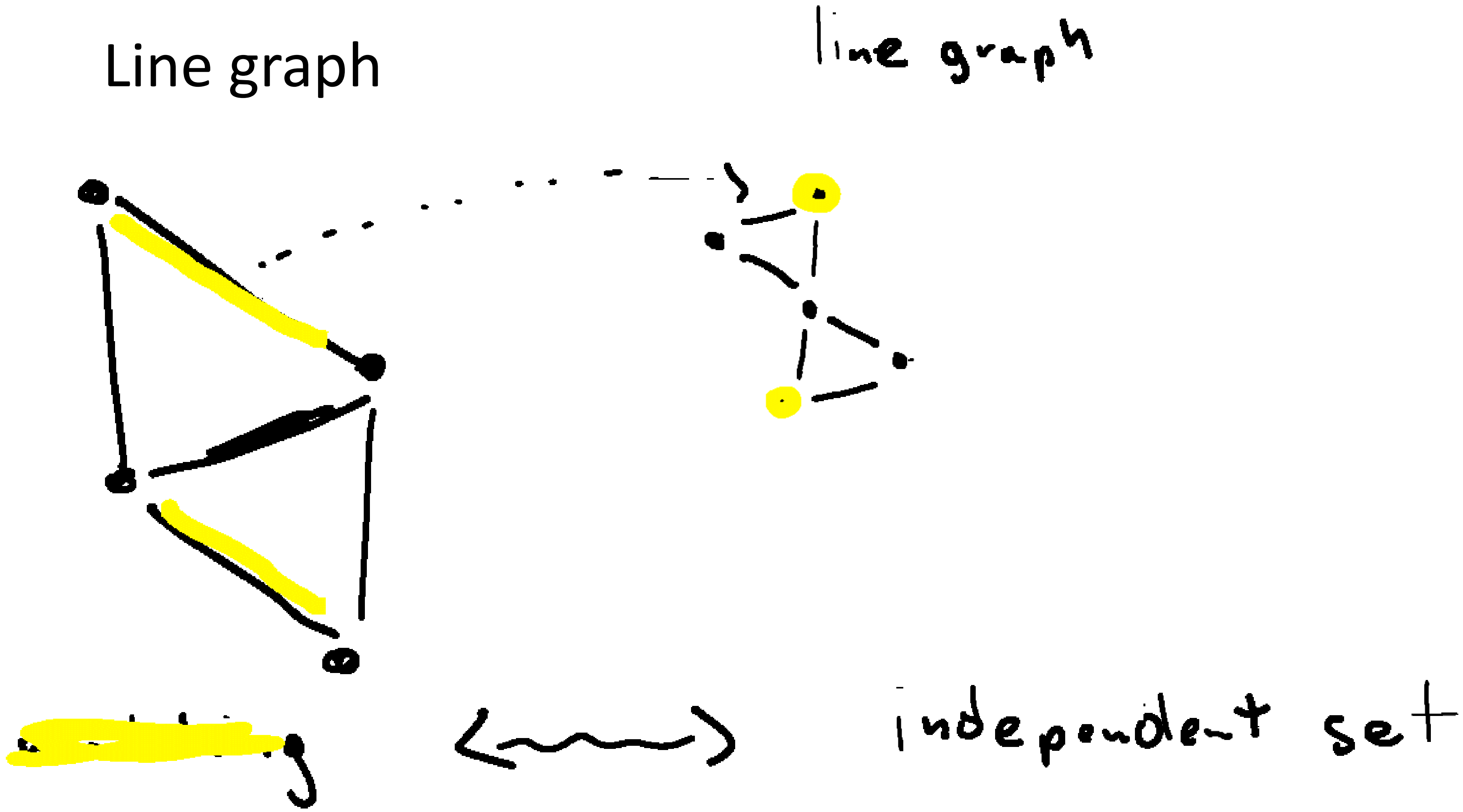
<u>Theorem</u>

$c \cdot \log n^2$ suffices with high prob

Matching in a graph



Line graph



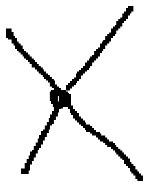
Independent set in line graph ---- matching in the graph

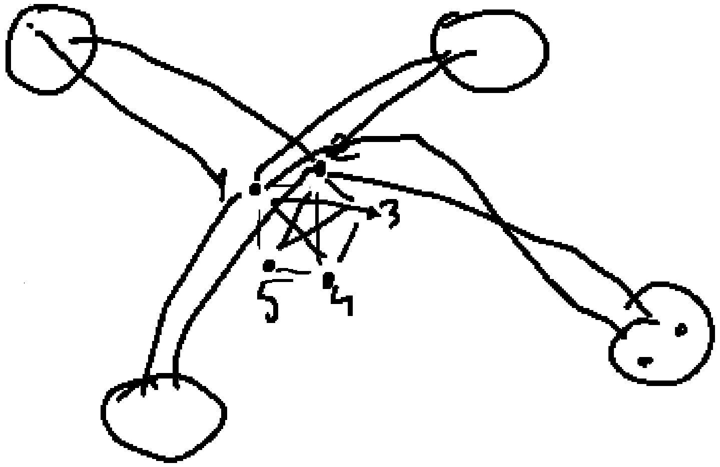
algo for MIS (~~~)

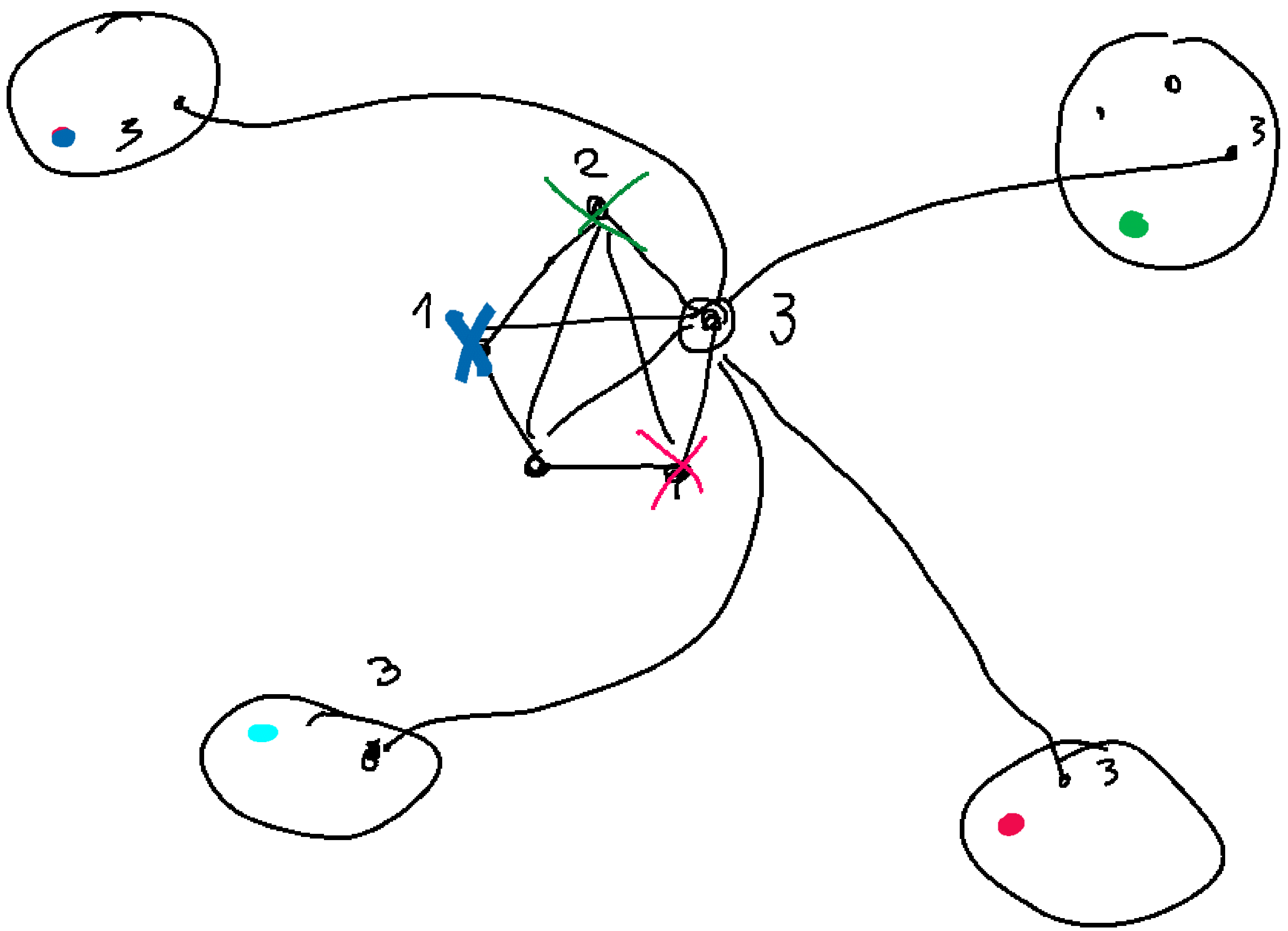
algo for M matching

Algorithm 7.20 General Graph Coloring

```
Given a graph G = (V, E) we virtually build a graph G' = (V', E') as follows:
Every node v ∈ V clones itself d(v)+1 times (v_0, ..., v_{d(v)} ∈ V'), d(v) being the degree of v in G.
The edge set E' of G' is as follows:
First all clones are in a clique: (v_i, v_j) ∈ E', for all v ∈ V and all 0 ≤ i < j ≤ d(v)
Second all i^th clones of neighbors in the original graph G are connected: (u_i, v_i) ∈ E', for all (u, v) ∈ E and all 0 ≤ i ≤ min(d(u), d(v)).
Now we simply run (simulate) the fast MIS Algorithm 7.12 on G'.
If node v_i is in the MIS in G', then node v gets color i.
```







2
1
3
3
3
3
3