Advanced Distributed Algorithms **Spring 2019**

Lecture 4: May 2

Lecturer: Merav Parter

Randomized Maximal Independent Set (MIS)

Luby's Round: Every vertex v picks a number r_v sampled u.a.r in $[0, 1]$, and joins the MIS if it is the local minima in its neighborhood, i.e., if $r_v < r_u$ for every $u \in N(v)$. MIS nodes and their neighbors are removed from the graph.

The analysis for this algorithm is based on showing that in each round, the number of *active edges*[1] is reduced by a factor of 2 in expectation. Thus, after $O(\log n)$ rounds, one can show that w.h.p. there are no active edges and the entire graph is decided.

In this class, we will show an improved algorithm by Ghaffari [Gha16] that solves MIS w.h.p within $O(\log \Delta)+2^{O(\sqrt{\log\log n})}$ rounds. We will in fact show a simplified variant with round complexity of $O(\log \Delta)+2^{O(\sqrt{\log \Delta+\log\log n})}$. This algorithm follows the two-stage structure: a pre-shattering randomized algorithm of $O(\log \Delta)$ rounds, followed by a deterministic algorithm that solves the remaining undecided subgraph.

Local Complexity vs. Global Complexity. The local complexity of a distributed algorithm is the number of rounds until a fixed vertex v terminates with *good* probability. Throughout a good probability is polynomially small in Δ, i.e., $1 - 1/\Delta^c$ for some constant c. The randomized algorithm with a small local complexity can complemented by a deterministic algorithm that solves the shattered graph.

The randomized phase of Ghaffari's algorithm has $O(\log \Delta)$ rounds with the guarantee that each vertex v remains undecided[2] after this phase with probability of $1/\Delta^c$. This holds even if the coins of all nodes not in $N_2^+(v) = \{v\} \cup \{u \mid \texttt{dist}(u, v) \leq 2\}$ are determined in an adversarial manner.

The intuition that underlies this algorithm is that there are two good scenarios for a given vertex v. Either v has a small number of neighbors that are still undecided. In such a case, it has a small number of competitors and thus has a good chance of joining MIS. Alternatively, v has many neighbors, most of them of low-degree. In this case, there is a good chance that at least one of v's neighbors will join the MIS, and v will be removed. Ghaffari's randomized algorithm is based upon creating a dynamic that guarantees that for a vertex v that is not yet decided after $\Theta(\log \Delta)$ rounds, w.p. 1 its spends a constant fraction of its life time in the above mentioned good scenarios. The (pre-shattering) algorithm is given below.

Ghaffari's MIS Algorithm:
Set $p_1(v) = 1/2$.

$$p_{t+1}(v) = \begin{cases} 1/2 \cdot p_t(v), & \text{if } d_t(v) \geq 2 \\ \min\{2p_t(v), 1/2\}, & \text{if } d_t(v) < 2, \end{cases}$$

where $d_t(v) = \sum_{u \in N(v)} p_t(u)$ is the *effective degree* of node v in round t. In each round t, the node v gets marked with probability $p_t(v)$ and if none of its neighbors is marked, v joins the MIS and gets removed along with its neighbors.

Lemma 4.1 *After $T = \beta(\log \Delta + \log 1/\epsilon)$ rounds, the probability that a vertex v is undecided is at most ϵ.*

[1] Edges whose both endpoints are still undecided (not in MIS and their neighbors are not in the MIS).
[2] By undecided we mean that the vertex is not yet in the MIS and that none of its neighbors is yet in the MIS.

We can in fact replace Δ with $deg(v)$ in the above expression, i.e., vertex v is decided with probability at least $1-\epsilon$ after $\beta(\log \deg(v) + \log 1/\epsilon)$.

To prove this lemma, we take the following strategy. First, we define two types of rounds that are good for a given vertex v, by good we mean that v has a constant chance, say of $1/100$, of being removed in these rounds. Then, we show that after T rounds, either v is decided or that it had experienced at least $T/13$ good rounds throughout its lifetime. Then, due to independencies between rounds, we can conclude that the probability that v has survived that many rounds is $(1/100)^{T/13} \leq \epsilon/\Delta \leq \epsilon$ for a sufficiently large β.

A vertex v is low-deg in round t if $d_t(v) < 2$, and otherwise it is high-deg. Let $L_t(v) = \{u \in N(v) \mid d_t(u) < 2\}$ be the low-deg neighbors of v in round t.

Definition 4.2 (Golden-In) *A round t is* golden-in *for a vertex v if $p_t(v) = 1/2$ and $d_t(v) < 2$.*

Definition 4.3 (Golden-Out) *A round t is* golden-out *for a vertex v if $d_t(v) \geq 1$ and at least $d_t(v)/10$ is due to low-deg nodes. That is, $\sum_{u \in L_t(v)} d_t(u) \geq d_t(v)/10$.*

Claim 4.4 (Golden-In is good) *If round t is golden-in for a vertex v, then v joins MIS in round t w.p. at least $1/32$.*

Proof: For v to join the MIS, it has to be marked (which happens w.p. $p_t(v)$) and that none of its neighbors is marked. Thus,

$$\begin{aligned}
\Pr[v \text{ joins the MIS}] &\geq 1/2 \cdot \prod_{u \in N(v)} (1 - p_t(u)) \geq 1/2 \cdot 1/4^{\sum_{u \in N(v)} p_t(u)} \\
&= 1/2 \cdot 1/4^{d_t(v)} \geq 1/32 \ ,
\end{aligned}$$

where the third inequality follows by using the inequality of $(1-x) \geq 1/4^x$ for every $x \in [0, 1/2]$. ■

Claim 4.5 (Golden-Out is good) *If round t is golden-in for a vertex v, then w.p at least $1/100$, v has a least one low-deg neighbor that joins the MIS in round t, and thus v is removed.*

Proof: We first bound the probability that at least one of the low-deg neighbors of v is marked. Then we will show that a low-deg vertex has a good chance of joining the MIS given that it is marked.

$$\begin{aligned}
\Pr[v \text{ has a low-deg neighbor that gets marked}] &\geq 1 - \prod_{u \in L_t(v)} (1 - p_t(u)) \geq 1 - e^{-\sum_{u \in L_t(v)} p_t(u)} \\
&\geq 1 - e^{-d_t(v)/10} \geq 1 - e^{-1/10} \ .
\end{aligned}$$

In addition, for every low-deg vertex v it holds that:

$$\Pr[\text{A low-deg vertex } u \text{ joins MIS} \mid u \text{ is marked}] \geq \prod_{w \in N(u)} (1 - p_t(w)) \geq 1/4^{\sum_{w \in N(u)} p_t(w)} = 4^{-d_t(u)} \geq 1/16 \ .$$

Overall, we get that v is removed w.p. at least $(1 - e^{-1/10}) \cdot 1/16 \geq 1/100$. ■

We next show that a vertex that remains undecided had many golden rounds throughout the execution.

Lemma 4.6 *After T rounds, either v is removed or that it had at least $T/13$ golden rounds.*

Proof: Let g_{in} be the number of golden-in rounds for v, i.e., rounds in which $p_t(v) = 1/2$ and $d_t(v) < 2$. Let g_{out} be the number of golden-out rounds for v, i.e., rounds in which $d_t(v) \geq 1$ and at least $d_t(v)/10$ is due to low-deg vertices (with $d_t(u) < 2$). Finally, let h denote the number of rounds in which $d_t(v) \geq 2$.

Step 1: Small $g_{in} \rightarrow$ Large h. Assume that $g_{in} < T/13$ (otherwise we are done). Observe that there are three types of rounds: rounds in which we increase the $p_t(v)$ by a factor of 2, reduce it by a factor of 2 or

do not change its value. In the latter case it must be that $p_t(v) = 1/2$. Also observe that since the value of $p_t(v)$ is at most $1/2$, the number of rounds in which we increase it by a factor of 2 is at most the number of rounds in which we reduce it by that factor. Finally, observe that whenever $d_t(v) \geq 2$, we reduce $p_t(v)$ be a factor of 2. Combining these observations we get that (# rounds in which $p_t(v) = 1/2) \geq T - 2h$.
Thus, $g_{in} \geq T - 3h$ as in golden-in rounds, $p_t(v) = 1/2$ and $d_t(v) < 2$. As $g_{in} < T/13$, we get that $h \geq 4/13 \cdot T$.

Step 2: Large $h \rightarrow$ Large g_{out}. Note that a golden-out round should satisfy two properties: (1) $d_t(v) \geq 1$ and (2) at least $d_t(v)/10$ is due to low-deg vertices. All the rounds of h satisfy the first property (with a slack!), but they do not necessarily satisfy the second property. We will show that if there are many rounds in which (1) holds then there should be many rounds in which the two properties hold, and thus g_{out} is large. We call a round *bad* if it satisfies (1) but does not satisfy (2). Note that for every bad round t it holds that:

$$d_{t+1}(v) \leq 1/2 \cdot 9/10 \cdot d_t(v) + 2 \cdot 1/10 \cdot d_t(v) \leq 2/3 d_t(v) \ .$$

To see this observe that since the bad round t does not satisfy (2), the total degree due to high-deg nodes is at least $9/10 \cdot d_t(v)$. Those high-deg nodes u have $d_t(u) \geq 2$, and therefore for each such u, $p_{t+1}(u) = p_t(u)/2$. The marking probabilities of the remaining low-deg nodes can be at most doubled, and thus the total contribution of the low-deg nodes in at most $1/5 \cdot d_t(v)$.

This implies that every bad round cuts $d_t(v)$ by a constant factor. Observe that $d_t(v)$ is at most Δ. These reductions can be somewhat compensated by the interleaving good rounds (golden-out rounds), but their effect is quite limited: every golden-out round can at most double the effective degree, and can therefore cancel the reduction of at most *two* bad rounds as $(2/3)^2 \cdot 2 < 1$. Overall, by ignoring the g_{out} rounds and their cancellations, we get that $h - g_{out} - 2g_{out} \leq \log_{3/2} \Delta$, and thus $g_{out} \geq (h - \log_{3/2} \Delta)/3$. Taking $h \geq 400(\log \Delta + \log 1/\epsilon)$, we get that $g_{out} > 100(\log \Delta + \log 1/\epsilon)$. ■

By using Claims 4.4, 4.5, Lemma 4.6 and taking $\epsilon = 1/\Delta^c$, we get that each vertex is decided after $O(\log \Delta)$ rounds w.p. at least $1 - 1/\Delta^c$.

The Post-Shattering Algorithm. After applying the first phase of Ghaffari's algorithm, the remaining undecided subgraph is solved deterministically using the same approach as we saw last class from [BEPSv3]. Recall that an MIS can be solved deterministically in $2^{O(\sqrt{\log n})}$ rounds by applying network decomposition (see first class). We will show that after applying the first phase each connected component is small, i.e., has $\widetilde{O}(\Delta^4)$ vertices, which leads to the total round complexity of $O(\log \Delta) + 2^{O(\sqrt{\log \Delta + \log \log n})}$.

Observation 4.7 *Fix a set S of 5-independent vertices (i.e., $\mathtt{dist}(u, v, G) \geq 5$ for every $u, v \in S$). The probability that all vertices in S are undecided after the first phase is at most $1/\Delta^{c \cdot |S|}$.*

Proof: For a vertex v the probability of joining the MIS in the given round, only depends on its immediate neighbors. However, the probability that v is removed depends also on the probability that one of v's neighbors joins the MIS. Therefore the decisions of v depend only on its 2-hop neighbors. Since the 2-hop neighborhoods of vertices at distance at least 5 are vertex disjoint, and since each vertex is undecided w.p. $1/\Delta^c$, we get that all vertices in S are undecided is bounded by $1/\Delta^{c|S|}$. ■

Lemma 4.8 *W.h.p. after the first phase, the size of each connected component in the remaining undecided subgraph is at most $\log_\Delta n \cdot \Delta^4$.*

Proof: The proof is very similar to the one we showed in the previous class for the coloring problem. The main distinction is that here we have independency between vertices at distance at least 5, and in the coloring algorithm, we had independency between vertices at distance at least 3.

Fix a subset T of $\log_\Delta n$ vertices that is (1) 5-independent and (2) connected in G^5. By Obs. 4.7, the probability that all vertices in T are undecided at the end of the first phase is at most $1/n^c$. We next bound the number of such sets T. There are $4^{\log_\Delta n}$ distinct unlabeled rooted trees with $\log_\Delta n$ vertices. In addition, there are $n \cdot \Delta^{5 \log_\Delta n}$ to embed each such tree in T. Thus in total there are at most $4^{\log_\Delta n} \cdot n \cdot \Delta^{5 \log_\Delta n} \leq n^6$

such sets. By taking c to be large enough, we can applying the union bound and conclude that w.h.p. there are no sets T that satisfy (i) and (ii) and remains undecided after the first phase.

Assume now towards contraction that the undecided subgraph contains a large component C of size at least $\log_\Delta n \cdot \Delta^4$. By applying a greedy procedure over C, we can compute a set T of size $|C|/\Delta^4$ that satisfies (i) and (ii), thus leading to a contradiction. ■

By employing the deterministic MIS algorithm (via network decomposition), we get:

Corollary 4.9 *The remaining undecided subgraph can be solved in* $2^{O(\sqrt{\log \Delta + \log \log n})}$ *rounds.*

References

[BEPSv3] Leonid Barenboim, Michael Elkin, Seth Pettie, and Johannes Schneider. The locality of distributed symmetry breaking. In *Foundations of Computer Science (FOCS) 2012*, pages 321–330. IEEE, 2012, also coRR abs/1202.1983v3.

[Gha16] Mohsen Ghaffari. An improved distributed algorithm for maximal independent set. In *SODA*, 2016.