

## Exercise 5: Size matters!

### Task 1: As small as possible, please?

Suppose a given graph  $G = (V, E)$  is decomposed into  $f$  forests, such that each  $v \in V$  initially knows the set  $P(v)$  of its (at most)  $f$  parents. Consider the following minimum dominating set (MDS) approximation algorithm.

---

**Algorithm 1** MDS approximation algorithm based on a forest decomposition.

---

- 1:  $H := \left( V, \left\{ \{v, w\} \in \binom{V}{2} \mid P(v) \cap P(w) \neq \emptyset \right\} \right)$
  - 2: compute an MIS  $I$  of  $H$
  - 3:  $D := \bigcup_{v \in I} P(v)$
  - 4: add all  $v \in V \setminus D$  without a neighbor in  $D$  to  $D$
  - 5: return  $D$
- 

- a) Show that Algorithm 1 can be implemented in the synchronous message passing model with running time  $\mathcal{O}(\log n)$  w.h.p.!
- b) Fix an MDS  $M$  of  $G$ . Denote by  $V_C \subseteq V$  the set of nodes that are in  $M$  or have a child in  $M$ . Show that  $|V_C| \leq (f + 1)|M|$ !
- c) Denote by  $V_P \subseteq V$  the set of nodes that have a parent in  $M$ . Show that  $|I \cap V_P| \leq |M|$ !
- d) Prove that after Line 3 of the Algorithm 1, at most  $|M|$  nodes are not covered by  $D$ .
- e) Conclude that Algorithm 1 computes a dominating set that is at most by factor  $\mathcal{O}(f^2)$  larger than the optimum!
- f)\* Show that even if we restrict message size to  $\mathcal{O}(\log n)$  bits, the algorithm can be implemented with running time  $\mathcal{O}(\log n)$  w.h.p.

### Task 2: Lots of wood

Denote by  $A(G)$  the *arboricity* of  $G = (V, E)$ , i.e., the minimum number of forests into which  $E$  can be decomposed. Our goal in this exercise is to decompose  $G$  into  $f \in \mathcal{O}(A)$  forests.

---

**Algorithm 2** Forest decomposition.

---

- 1: **while**  $V \neq \emptyset$  **do**
  - 2:   **for** all  $v \in V$  with  $\delta_v \leq 4A(G)$  in parallel **do**
  - 3:      $v$  assigns its incident edges to different forests  $F_1, \dots, F_{4A(G)}$
  - 4:     delete  $v$  (and its incident edges) from  $G$
  - 5:   **end for**
  - 6: **end while**
  - 7: return the computed forests (each node knows its parent in  $F_1, \dots, F_{4A(G)}$ )
- 

- a) Show that in each iteration of the WHILE loop, at least half of the remaining nodes are deleted! (Hint: Assume that this is false and bound the number of remaining edges from below. Compare the result to the maximum number of edges in  $A(G)$  forests.)

- b) Conclude that the algorithm computes a decomposition of  $G$  into at most  $4A(G)$  forests in  $\mathcal{O}(\log n)$  rounds!
- c) Change the algorithm so that it does not require knowledge of  $A(G)$ , but instead relies on an upper bound  $N \in n^{\mathcal{O}(1)}$  on  $n$ ! You may use up to  $8A(G)$  forests and increase the running time of the algorithm by a factor of  $\mathcal{O}(\log A(G))$ !<sup>1</sup>
- d) Conclude that in graphs of arboricity  $A$ , a factor- $\mathcal{O}(A^2)$  approximation to MDS<sup>2</sup> can be found in  $\mathcal{O}(\log n \log A)$  rounds w.h.p., provided that an upper bound  $N \in n^{\mathcal{O}(1)}$  on  $n$  is known!
- e)\* Can you do it in  $\mathcal{O}(\log n)$  rounds if  $A$  is unknown, but an upper bound  $N \in n^{\mathcal{O}(1)}$  on  $n$  is known?

### Task 3\*: Exponential enhancement

- a) Why is Chernoff's bound called Chernoff's bound?
- b) Show that for independent variables  $X_i, i \in I, \mathbb{E}[\prod_{i \in I} X_i] = \prod[\mathbb{E}[X_i]]$ .
- c) Use Markov's bound to show that for positive  $X_i, i \in I$ , and arbitrary  $t, \delta > 0$ ,

$$P[X \geq (1 + \delta)\mathbb{E}[X]] \leq \frac{\mathbb{E}[\prod_{i \in I} e^{tX_i}]}{e^{t(1+\delta)\mathbb{E}[X]}}.$$

- d) Use a) and b) to infer that if the  $X_i$  are Bernoulli variables, then

$$P[X \geq (1 + \delta)\mathbb{E}[X]] \leq \frac{e^{(e^t - 1)\mathbb{E}[X]}}{e^{t(1+\delta)\mathbb{E}[X]}}.$$

- e) Plug in  $t := \ln(1 + \delta)$ . You obtain the upper tail bound; choosing  $\delta \in (0, 1)$  and  $t = 1 - \delta$  yields the lower tail bound.<sup>3</sup> The bounds derived here are stronger than those in the lecture, but more unwieldy. For most applications, the simpler versions suffice.
- f) Enlarge the knowledge of the exercise group by reporting your findings!

---

<sup>1</sup>Forest decompositions into  $f$  forests are particularly interesting if  $f \geq A(G)$  is small, hence usually  $\log A(G)$  is *very* small!

<sup>2</sup>Read: "a dominating set at most a constant factor larger than an MDS".

<sup>3</sup>Note that one has to introduce a minus sign in the exponents in b) to still be able to apply Markov's inequality.