Deterministic algorithm for MPC
[Czumaj, Davies, Parter '20]

- maximal matching, maximal independent set
  $O(\log \Delta + \log\log n)$ rounds

- connectivity?

Submodular Maximization

Ground set $\mathcal{U}$, $|\mathcal{U}| = n$

$f : 2^{\mathcal{U}} \longrightarrow \mathbb{R}^+$, integer $k \geq 0$

Find subset $X \subseteq \mathcal{U}$, $|X| = k$ st. $f(X)$ maximized

cardinality constraint

f has following properties:

Monotone: $f(A) \leq f(B)$ $\quad A \subseteq B \subseteq \mathcal{U}$

Submodular: $f(A \cup \{e\}) - f(A) \geq f(B \cup \{e\}) - f(B)$
$A \subseteq B \subseteq \mathcal{U}$

$f_A(e) : f(A \cup \{e\}) - f(A)$

diminishing returns: marginal of e w.r.t. A is at least as high as marginal of e w.r.t. supersets B of A

Equivalent: $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$

<u>Example:</u> k-coverage

Given collection of n sets $S_1 .. S_n$

Find k sets st. union of these is maximized

$I \subseteq [n]$ $\quad f(I) = \left| \bigcup_{i \in I} S_i \right|$

NP-hard. Natural greedy algo. gives $1 - \frac{1}{e}$ approx.

(Sequential) Greedy:

$X \leftarrow \phi$

While $|X| < k$

(a) $u \leftarrow \underset{e \in U}{\arg\max} \; \overbrace{f(X \cup \{e\}) - f(X)}^{f_X(e)}$

(b) $X \leftarrow X \cup \{u\}$

Lemma: Greedy is $(1-\frac{1}{e})$ approx

$$f(X) \geq \left(1 - \left(1-\frac{1}{k}\right)^k\right) \cdot OPT$$

Proof: X_i: Greedy set in ith iteration

Claim: $f(X_{i+1}) - f(X_i) \geq \frac{1}{k}(OPT - f(X_i))$

$$(OPT - f(X_i)) - (OPT - f(X_{i+1})) \geq \frac{1}{k}(OPT - f(X_i))$$

$$OPT - f(X_{i+1}) \leq \left(1-\frac{1}{k}\right)(OPT - f(X_i))$$

$$OPT - f(X_k) \leq \left(1-\frac{1}{k}\right)^k \underbrace{(OPT - f(X_0))}_{OPT}$$

$$f(X_k) \geq \left(1 - \left(1-\frac{1}{k}\right)^k\right) \cdot OPT$$

Pf of Claim:

Add k elements of OPT to X_i

$f(\,)$ increases by atleast $OPT - f(X_i)$

At least one elt causes increase of $\frac{OPT - f(X_i)}{k}$

Marginal of this ele. w.r.t. $X_i \geq \frac{OPT - f(X_i)}{k}$

Value of greedy increases by $\geq \frac{OPT - f(X_i)}{k}$

Warm Up: Threshold-Greedy Algorithm

$X \leftarrow \phi$

While $|X| < k$ and $\exists e \in U \setminus X$ s.t. $f_X(e) \geq \tau$

- $X \leftarrow X \cup \{e\}$

<u>Lemma:</u> If $\tau = \frac{OPT}{2k}$, $f(X) \geq \frac{OPT}{2}$

<u>Pf</u>: If $|X| = k$, $f(X) \geq k\tau = k \cdot \frac{OPT}{2k} = \frac{OPT}{2}$

If $|X| < k$, add optimal solⁿ to X
increases solⁿ by $OPT - f(X)$

$\exists$ element with marginal $\frac{OPT - f(X)}{k}$ w.r.t. X

$$\frac{OPT - f(X)}{k} < \frac{OPT}{2k}$$

$$f(X) > \frac{OPT}{2}$$

[Liu, Vondrak '19] $\frac{1}{2} - o(1)$ approx in 2 rounds.

$(1 - \frac{1}{e} - \varepsilon)$ approx in $\sim \frac{1}{\varepsilon}$ rounds.

machines $= \sqrt{\frac{n}{k}} + 1$, memory $O(\sqrt{nk})$

memory: $n^{\alpha} \rightarrow \frac{1}{\alpha}$ overhead in # rounds

Threshold Greedy (S, G, τ)

Input: Set S, partial greedy solⁿ G, $|G| \leq k$, threshold τ

$G' \leftarrow G$

For $e \in S$

if $f_{G'}(e) \geq \tau$ and $|G'| < k$

$G' \leftarrow G' \cup \{e\}$

ThresholdFilter (S, G, τ)

Input: Set S, partial greedy sol^n G, threshold τ

Output: $S' \subseteq S$ s.t. $f_G(e) \geq \tau$ for all $e \in S'$

Algorithm:

round 1

$E[|S|] = 4\sqrt{kn}$

$S \leftarrow$ sample each $e \in V$ w. prob. $p = 4\sqrt{\frac{k}{n}}$ $\quad E[|V_i|] = \sqrt{nk}$

partition V randomly into $V_1 \dots V_m$ to m machines (one per machine)

Send S to each machine and central machine C

On each machine M_i (in parallel) do

$\tau = \frac{OPT}{2k}$

$G_0 \leftarrow$ ThresholdGreedy (S, ϕ, τ)

If $|G_0| < k$ then

$R_i \leftarrow$ ThresholdFilter (V_i, G_0, τ)

else $R_i \leftarrow \phi$

send R_i to central machine C

round 2 (only on C)

Compute G_0 from S as in in round 1

$G \leftarrow$ ThresholdGreedy $(\cup R_i, G_0, \tau)$

Lemma: Approx ratio $\geq \frac{1}{2}$

Lemma: W. prob. $1 - e^{-\Omega(k)}$, #elements sent to central machine is $\leq \sqrt{nk}$

Proof: $E[|S|] = 4\sqrt{nk}$

Chernoff: $\Pr[|S| < 3\sqrt{nk}] \leq e^{-\Omega(\sqrt{nk})} \leq e^{-\Omega(k)}$

Assume that $|S| \geq 3\sqrt{nk}$

elements sent to C in round 2 $= |\underset{\uparrow}{\cup} R_i|$

eles. w. marginal $\geq \frac{OPT}{2k}$ w.r.t. G_0

If $|\cup R_i| \leq \sqrt{nk}$, then we're done!

Assume $|\cup R_i| > \sqrt{nk}$

Break S into 3k blocks of size $\sqrt{n/k}$

Imagine choosing S block by block

Before each block, at least $\sqrt{nk}$ elements of marginal value $\geq \frac{OPT}{2k}$

$Pr(\text{not sampling high margin eles}] \leq \left(1-\sqrt{\frac{k}{n}}\right)^{\sqrt{\frac{n}{k}}} < \frac{1}{2}$

High margin ele sampled w. prob $\geq \frac{1}{2}$ in each block

$\Rightarrow$ G_0 increases w. prob. $\geq \frac{1}{2}$ in each block.

3k blocks $\Rightarrow$ expected size of $G_0 \geq \frac{3k}{2}$

$Pr[|G_0| \geq k] \geq 1 - e^{-\Omega(k)}$

ie. we pick k eles in G_0

Fixing knowledge of OPT:

Guess within factor of $(1+\varepsilon)$. Increase memory by $\frac{\log k}{\varepsilon}$

2 classes of inputs:

① (Dense) At least $\sqrt{nk}$ eles. of value $\frac{OPT}{2k}$

v: max value of single ele in sample S

likely to be in $\left[\frac{OPT}{2k}, OPT\right]$

Use guesses $\tau_j = v(1+\varepsilon)^j$

at least one is within factor $(1+\varepsilon)$ of OPT

② (Sparse) $< \sqrt{nk}$ eles of value $\frac{OPT}{2K}$

Send all large elements to central machine and run threshold greedy on this machine

$\left(\frac{1}{2} - \varepsilon\right)$ approx in 2 rounds.

$\left(1 - \frac{1}{e} - \varepsilon\right)$ approx in $\sim \frac{1}{\varepsilon}$ rounds

Repeat algorithm t times w. different thresholds:

$\alpha_\ell = \left(1 - \frac{1}{t+1}\right)^\ell \cdot \frac{OPT}{K}$ $\qquad \ell = 1 .. t$

Build on greedy solⁿ from previous iterations

Lemma: Approx ratio $\geq 1 - \left(1 - \frac{1}{t+1}\right)^t$

Proof outline: By induction: value of first $\frac{\ell}{t} K$ elements picked is at least $\left(1 - \left(1 - \frac{1}{t+1}\right)^\ell\right) \cdot OPT$

Suppose true for $\ell - 1$

2 cases:

① Either all $\frac{1}{t} K$ eles picked above α_ℓ

First $\frac{\ell-1}{t} K$ eles: $\geq \left(1 - \left(1 - \frac{1}{t+1}\right)^{\ell-1}\right) \cdot OPT$

$\frac{1}{t} \cdot K$ eles $\geq \frac{1}{t} \cdot K \left(1 - \frac{1}{t+1}\right)^\ell \cdot \frac{OPT}{K}$

② Not all $\frac{\ell}{t} \cdot K$ eles picked are above α_ℓ

$\rightarrow$ no eles of high marginal value left