0368.416701 Sublinear Time Algorithms	December 29, 2014
Lecture 10	
Lecturer: Ronitt Rubinfeld	*Scribe: Daniel Zaidman*

0.1 Today's lecture

Today's lecture is about property testing in dense graphs: a lower bound in epsilon.

1 Theorem (adjacent matrix model)

$\exists$ const c s.t. any 1-sided* tester for whether graph is $\triangle$ - free needs $\geq (\frac{c}{\epsilon})^{clog(\frac{c}{\epsilon})}$ queries.

* Always pass if $\triangle$ - free. You have to <u>see</u> a $\triangle$ to fail.

Tools:

1.1 (1) Goldreich-Trevisan Theorem

Adj matrix model. Property P. Tester T using q(n,ϵ)
$\Rightarrow \exists$ 1-sided error tester T' "Natural Tester" with 1-sided error.

1. Pick q(n, ϵ) nodes randomly.
2. Query submetrix.
3. Decide.

All are O(q^2) queries.
Lower bound (l.b.) for 1-sided natural tester of $\Omega(q')$
$\Rightarrow$ 1-sided l.b. for any tester of $\Omega(\sqrt{q'})$
Note - reduction preserves 1-sidedness. So l.b. consequence holds.
$\Omega(q')$ queries 1-sided, natural $\Rightarrow$
$\Omega(\sqrt{q'})$ queries 1-sided tester

1.2 (2) Additive # theory lemma

$\forall m, \exists x \subset M = \{1, 2, ..., m\}$ s.t. $|x| \geq \frac{m}{e^{10\sqrt{logm}}}$
with no nontrivial solution to $x_1 + x_2 = 2x_3$
(trivial $x_1 = x_2 = x_3$)
Bad x: $\{1, 2, 3\}, \{5, 9, 13\}$
Good x: $\{1, 2, 4, 5, 10, ...\}, \{1, 2, 3, 8, 16, 32, ...\}$ Both too small

<u>Proof:</u>
Let d be an integer $\approx e^{10\sqrt{logm}}$
k $\leftarrow \lfloor \frac{\log m}{\log d} \rfloor - 1$ (so $k \approx \frac{\log m}{10\sqrt{\log m}} \approx \frac{\sqrt{\log m}}{10}$)
Define $X_B = \left\{ \sum_{i=0}^{k} x_i d^i | x_i <^* \frac{d}{2} \quad for \quad 0 \leq i \leq k; \sum_{i=0}^{k} x_i^2 = B \right\}$
$\forall B \in [0, ?]$

* Summing elements of X_B doesn't generate a carry.

Claim: $X_B \leq M$
Why? Largest # in $X_B \leq d^{k+1} \leq d^{\log_d m} = m$
What is B? Pick B to maximize $|X_B|$
$B \leq (k+1)(\frac{d}{2})^2$
$|\bigcup_B X_B| \leq \sum\limits_B |X_B| = \frac{d}{2}^{k+1}$
$\exists B$ s.t. $|X_B| \geq \frac{(\frac{d}{2})^{k+1}}{kd^2}$
With settings $\Rightarrow |X_B| \geq \frac{m}{e^{10 \log m}}$
Claim: X_B is "sum - free— if $\forall x, y, z \in X_B$ if $x + y = 2z \Rightarrow x = y = z$

Proof: $x, y, z \in X_B$
$x + y = 2z \Leftrightarrow \sum\limits_{i=0}^{k} x_i d^i + \sum\limits_{i=0}^{k} y_i d^i = 2 \sum\limits_{i=0}^{k} z_i d^i$
$\Leftrightarrow$ since no carries
$x_0 + y_0 = 2z_0$
$x_1 + y_1 = 2z_1$
.
.
.
$x_k + y_k = 2z_k$

Note: $\forall i \quad x_i + y_i = 2z_i \Rightarrow x_i^2 + y_i^2 \geq 2z_i^2$
with equality only if $x_i = y_i = z_i$
why fact? $f(a) = a^2$ is convex.
Jensen's inequality for convex f: $\frac{\sum\limits_{i=1}^{n} f(a_i)}{n} \geq f\left(\frac{\sum\limits_{i=1}^{n} a_i}{n}\right)$
with "=" $\Leftrightarrow$ all a_i's are equal.
$\Rightarrow \frac{x_i^2 + y_i^2}{2} \geq (\frac{2z_i}{2})^2 = z_i^2$ with "=" $\Leftrightarrow x_i = y_i = z_i$

Suppose $!(x = y = z) \Rightarrow \exists i \quad s.t. \quad x_i^2 + y_i^2 > 2z_i^2$
$\forall j \neq i \quad x_j^2 + y_j^2 \geq 2z_j^2$
So $\sum\limits_{l=0}^{k} x_l^2 + \sum\limits_{l=0}^{k} y_l^2 > 2 \sum\limits_{l=0}^{k} z_l^2$
So $B + B > 2B$ contradiction
So $x = y = z$

1.3 Proof of property terting l.b.

Given sum-free $x \subset \{1, 2, ..., m\}$ construct a graph:
$V_1 = \{1, 2, ..., m\}, V_2 = \{1, 2, ..., 2m\}, V_3 = \{1, 2, ..., 3m\}$
Nodes: (i,j), $i \in \{1, 2, 3\}, j \in \{1, 2, ..., 3m\}$
$\forall node(1, j), \forall x \in X$ there is an edge to $(2, j + x)$, and an edge to $(2, j + 2x)$.
$\forall node(2, j), \forall x \in X$ there is an edge to $(3, j + x)$.

nodes = 6m
nodes = $\Theta(m|x|) = \Theta(\frac{n^2}{e^{10\sqrt{logn}}})$

Reminder:

- if $\triangle$ - free $\Rightarrow$ Pass
- if ϵ - far (remove $> \epsilon n^2$ edges) $\Rightarrow$ Fail w.h.p.

triangles =

- "intended $\triangle$'s": k, j+x, j+2x
- "non intended $\triangle$'s": none of these.

Distance to $\triangle$ - free $\geq m|x|$ edges.
S-blow up $G \rightarrow G^{(s)}$
Vertex in $G \rightarrow$ size s independent set in $G^{(s)}$
Edge in $G \rightarrow$ complete bipartite graph in $G^{(s)}$

$\triangle$ in $G \Rightarrow s^3 in G^{(s)}$
$G^{(s)}$:

- nodes: ms
- edges: $m|x|s^2$
- $\triangle$'s: $m|x|s^3$

1.4 <u>Lemma</u>: # edge disjoint $\triangle's \geq m|x|s^2$

Given ϵ, pick m largest integer, s.t. $\epsilon \leq \frac{1}{10^{10\sqrt{\log m}}}$
$m \geq (\frac{c}{\epsilon})^{clog(\frac{c}{\epsilon})}$
Pick $s = \lfloor \frac{n}{6m} \rfloor \sim n(\frac{c}{\epsilon})^{clog(\frac{c}{\epsilon})}$

- # edges distance to $\triangle$-free ϵn^2
- # $\triangle's \sim (\frac{c'}{\epsilon})^{c'log(\frac{c'}{\epsilon})}$

$Pr[pick \quad \triangle] = (\frac{c''}{\epsilon})^{c''log(\frac{1}{\epsilon})}$
Sample of size q
How many $\triangle$'s in sample?

$E[\#\triangle's] = \binom{d}{3}(\frac{c''}{\epsilon})^{c''log(\frac{1}{\epsilon})} << 1$
Unless $q > (\frac{c''}{\epsilon})^{c''log(\frac{1}{\epsilon})}$
By Markov's inequality, Pr[see $\triangle$] << 1
Since 1-sided error, we'll pass even though ϵ-far.