

A lower bound for testing Δ -freeness

Last time:

- saw property test for Δ -freeness
 - const time in terms of n
 - dependence on ϵ horrible - tower of 2's
- is this required?

Today:

- answer this question partially (for 1-sided testers)

• When testing H -freeness property,

interesting characterization of bipartiteness

- if H bipartite, $\text{poly}(1/\epsilon)$ is enough
- if H not bipartite no $\text{poly}(1/\epsilon)$ suffices

(We'll actually prove special case of $H = \Delta$ only)

Thm (adj matrix model)

\exists const c st. any 1-sided tester for whether graph G is Δ -free needs $\geq \left(\frac{c}{\epsilon}\right)^{\log 1/\epsilon}$ queries.

Main Tools:

(1) Goldreich-Trevisan Thm:

Adj matrix model

Property P

Tester T with $q(n, \epsilon)$ queries

\Rightarrow Tester T' : "Natural Tester"
 pick $q(n, \epsilon)$ nodes
 query submatrix } $O(q^2)$ queries
 decide

Consequences:

- l.b. for natural tester of $\Omega(q')$
 \Rightarrow l.b. for any tester of $\Omega(\sqrt{q'})$
- note, reduction preserves 1-sidedness,
 so l.b. implication does too.

Main tools (cont.):

(2) Additive Number theory lemma

#theory Lemma $\forall m, \exists X \subset M = \{1, 2, \dots, m\}$

of size $\geq \frac{m}{e^{10\sqrt{\log m}}}$

with no non trivial soln to $X_1 + X_2 = 2X_3$
 \uparrow i.e. $x_1 = x_2 = x_3$ is the trivial soln.

Will use to construct graphs st.

- far from Δ -free
- natural algorithm needs $\Omega\left(\frac{c}{\epsilon}\right)^{\log \frac{1}{\epsilon}}$ queries

examples

Bad X : $\{1, 2, 3\}$
 $\{5, 9, 13\}$

Good X ? $\{1, 2, 4, 5, \cancel{7}, \cancel{8}, \cancel{9}, 10, \dots\}$ \leftarrow how big??
 $\{1, 2, 4, 8, 16, 32, \dots\}$ \leftarrow only size $\log m$

Proof of lemma

• let d be integer (later, set to $e^{10\sqrt{\log m}}$)
 $k \leftarrow \left\lfloor \frac{\log m}{\log d} \right\rfloor - 1$ (so $k \approx \frac{\log m}{10\sqrt{\log m}} \approx \frac{\sqrt{\log m}}{10}$)

Proof of lemma (cont.)

define $X_B = \left\{ \sum_{i=0}^k X_i d^i \mid \begin{array}{l} X_i < \frac{d}{2} \text{ for } 0 \leq i \leq k \\ \sum_{i=0}^k X_i^2 = B \end{array} \right\}$

① $X_i < \frac{d}{2}$ for $0 \leq i \leq k$

② $\sum_{i=0}^k X_i^2 = B$

the two constraints will be used later in a nice way

view each $x \in M$ as represented in base d
 where $x = (X_0 \dots X_k)$
 "digits" of x

Claim $X_B \subseteq M$

Why? largest number in X_B
 $\leq d^{k+1} \leq d^{(\lfloor \frac{\log m}{\log d} \rfloor - 1) + 1} \leq d^{\log_d m} = m^{\log_d d} = m$

What is B ? Pick st. $|X_B|$ maximized

Why the constraints?

① X_i 's $< \frac{d}{2} \Rightarrow$ summing pairs of elements in X_B doesn't generate a carry in any location!

we'll see why this is useful soon

(2) will use \checkmark (along with (1)) to show that X_B is "sum-free"

Claim X_B is "sum free" i.e. $\nexists x, y, z \in X_B$ s.t.
 $x + y = 2z$

Pf of claim

for $x, y, z \in X_B$

$$x + y = 2z \iff \sum_{i=0}^k x_i d^i + \sum_{i=0}^k y_i d^i = 2 \sum_{i=0}^k z_i d^i$$

$$\iff \begin{aligned} x_0 + y_0 &= 2z_0 \\ x_1 + y_1 &= 2z_1 \\ &\vdots \\ x_k + y_k &= 2z_k \end{aligned}$$

} since no carries

Note $\forall i \quad x_i + y_i = 2z_i \implies \forall i \quad x_i^2 + y_i^2 \geq 2z_i^2$
with equality only if $x_i = y_i = z_i$

why? $f(a) = a^2$ is convex

Jensen's $\frac{\sum_{i=1}^n f(a_i)}{n} \geq f\left(\frac{\sum_{i=1}^n a_i}{n}\right)$ with equality only if a_i 's are all =

$$\implies \frac{x_i^2 + y_i^2}{2} \geq \left(\frac{2z_i}{2}\right)^2 = z_i^2 \quad \text{+ equal only if}$$

$$x_i = y_i = z_i$$

▣ (proof of note)

finishing proof of claim:

if x, y, z s.t. $\text{not}(x=y=z)$

then $\exists i$ s.t. $\text{not}(x_i=y_i=z_i)$

then note $\Rightarrow x_i^2 + y_i^2 > 2z_i^2$

+ for all other j , $x_j^2 + y_j^2 \geq 2z_j^2$

but then:

$$\underbrace{\sum x_i^2}_{=B} + \underbrace{\sum y_i^2}_{=B} > \sum 2z_i^2 = 2 \underbrace{\sum z_i^2}_{=B}$$

$$\underbrace{\hspace{10em}}_{2B}$$

\Leftarrow

but how do we know that X_B is big?

- $B \leq (k+1) \left(\frac{d}{a}\right)^2 < kd^2$
 \uparrow
 bound on digits of B

- $| \cup_B X_B | \geq \left(\frac{d}{a}\right)^{k+1} > \left(\frac{d}{a}\right)^k$
 \parallel
 $\sum_B |X_B|$

- $\exists B$ s.t. $|X_B| \geq \frac{\left(\frac{d}{a}\right)^k}{kd^2}$

use settings of d, k , get

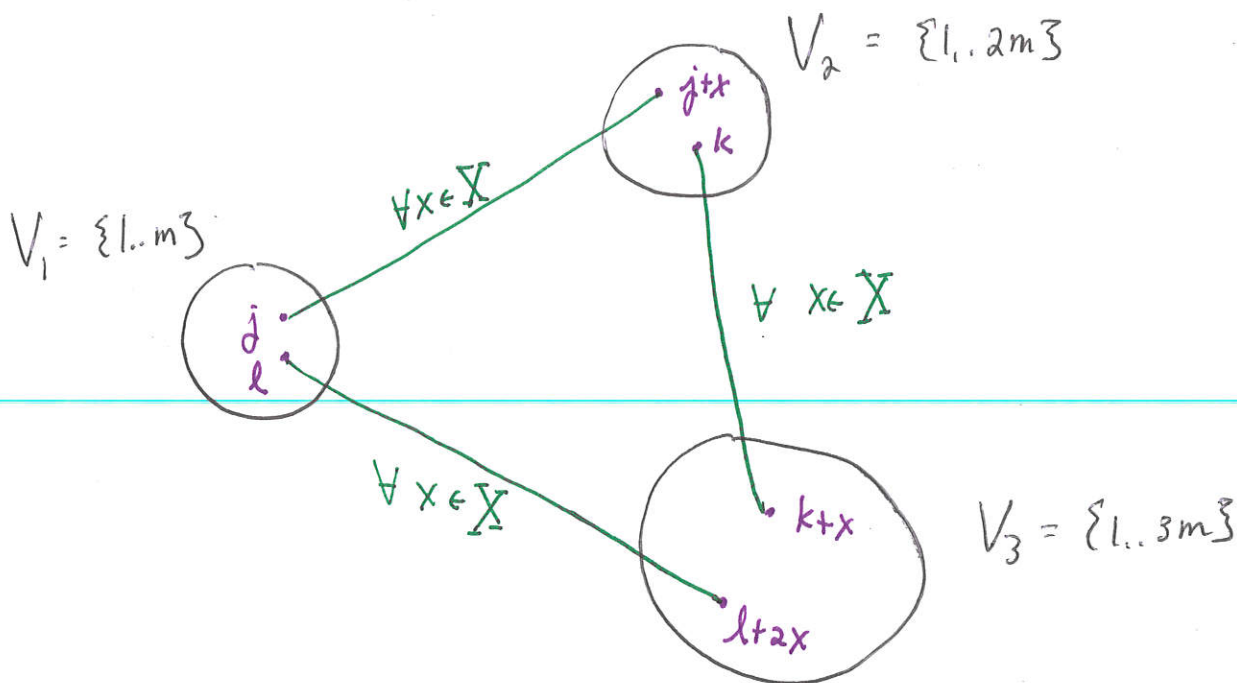
$$|X_B| \geq \frac{m}{e^{10 \log m}}$$

Not enough! need another idea, but won't do it here

Proof of Thm (prop testing bound)

given sum-free $X \subseteq \{1..m\}$

construct a graph:



• will abuse notation:

node should be (i, j)
 $i \in \{1, 2, 3\}$ $j \in \{1..cm\}$

will drop i if easy to see from context

• #nodes = $6m$

• #edges = $\Theta(m \cdot |X|) = \Theta(n^2 / e^{10\sqrt{\log n}})$ ← not exactly dense

• # cycles :

intended Δ 's : $j, j+x, j+2x$

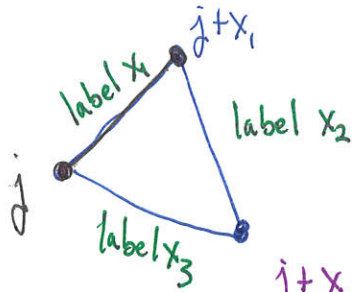
intended Δ 's is $m/|X| = \Theta\left(\frac{n^2}{e^{10\sqrt{\log n}}}\right)$

nonintended Δ 's :

• no edges internal to V_1, V_2 or V_3

\therefore any Δ has

$$\begin{aligned} u &\in V_1 \\ v &\in V_2 \\ w &\in V_3 \end{aligned}$$



$$\left. \begin{aligned} j+x_1+x_2 \\ = j+2x_3 \end{aligned} \right\} \Rightarrow x_1+x_2 = 2x_3$$

$$\Rightarrow x_1 = x_2 = x_3 \quad \text{since } X \text{ is sum-free}$$

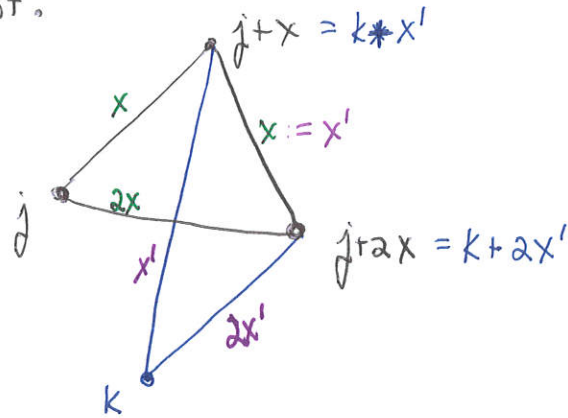
but these are intended!

\therefore no nonintended Δ 's

• # disjoint cycles:

intended Δ 's are disjoint

suppose not:



since $x=x'$, $k=j$ $\rightarrow \leftarrow$

• distance to Δ -free:

must remove ≥ 1 edge from each Δ

\Downarrow

"Absolute" distance

$$\begin{aligned} \text{from } \Delta\text{-free} &= \Theta(\#\Delta\text{'s}) \\ &= \Theta\left(\frac{n^2}{e^{10\sqrt{\log n}}}\right) \\ &= \Theta(m/|X|) \end{aligned}$$

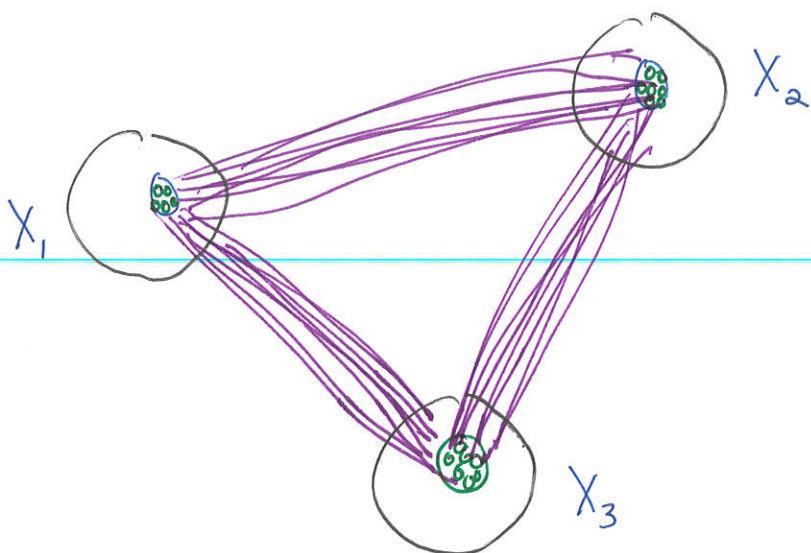
Problem need $\Omega(\epsilon n^2)$ distance

Idea for fix

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 $G^{(s)}$



Note: Δ in $G \Rightarrow s^3 \Delta$'s in $G^{(s)}$

nodes in $G^{(s)} \sim m \cdot s$ (actually $6ms$)

edges " " $\sim m/x \cdot s^2$

triangles " " $\sim m/x \cdot s^3$

Lemma dist of $G^{(s)}$ from Δ -free

\geq #edge disjoint Δ 's

$\geq m/x \cdot s^2$

Proof show each triangle in $G \Rightarrow s^2$ disjoint Δ 's in $G^{(s)}$

Given ϵ , pick m to be largest int
 satisfying

$$\epsilon \leq \frac{1}{e^{10\sqrt{\log m}}}$$

this m satisfies

$$m \geq \left(\frac{c}{\epsilon}\right)^{c \log c/\epsilon}$$

Pick $S = \lfloor \frac{n}{6m} \rfloor \approx n \left(\frac{\epsilon}{c}\right)^{c \log c/\epsilon}$

\Rightarrow #edges \sim distance $\sim \epsilon n^2$

#triangles $\sim \left(\frac{\epsilon}{c}\right)^{c \log c/\epsilon} n^3$



$$m|X| \cdot S^3 = \frac{m^2}{e^{10\sqrt{\log m}}} S^3$$

$$= \frac{1}{\epsilon} \left(\left(\frac{c}{\epsilon}\right)^{c \log c/\epsilon}\right)^2 \cdot \left(\frac{\epsilon}{c}\right)^{c \log c/\epsilon} n^3$$

(since $\approx \frac{m|X|S^2}{m^2 S^2} \leftarrow$ size of adj matrix
 $= \frac{|X|}{m} \geq \frac{1}{e^{10\sqrt{\log m}}} \geq \epsilon$
 \uparrow $|X| = \frac{m}{e^{10\sqrt{\log m}}}$)

Finally if take sample of size q

$$E[\#\Delta\text{'s in sample}] < \binom{q}{3} \left(\frac{\epsilon}{c}\right)^{c \log c/\epsilon}$$

$$<< 1 \quad \text{unless } q > \left(\frac{c}{\epsilon}\right)^{c \log c/\epsilon}$$

by Markov's $\neq \Rightarrow \Pr[\text{see } \Delta] << 1$

But since 1-sided error,
 must find Δ in order to fail \blacksquare