A Characterization of Easily Testable Induced Subgraphs (Part II)

Noga Alon and Asaf Shapira
Combinatorics, Probability and Computing **15** (2006) 791–805.

Speaker: Joseph, Chuang-Chieh Lin
Advisor: Professor Maw-Shang Chang

Computation Theory Laboratory
Department of Computer Science and Information Engineering
National Chung Cheng University
Taiwan

May 19, 2009

Outline

1. Introduction
   - Brief introduction to property testing
   - Focus of this talk
2. Two technical skills
   - h-sum-free sets
   - s-blow-up
3. Two main lemmas
4. Proof of the main theorem
5. Go back to the proof of Lemma 1

Outline

Brief introduction to property testing

- Try to answer "yes" or "no" for the following *relaxed* decision problems by observing only a small fraction of the input.
  - Does the input **satisfy a designated property**, or
  - is ϵ-**far from satisfying the property**?

Brief introduction to property testing

- Try to answer "yes" or "no" for the following *relaxed* decision problems by observing only a small fraction of the input.
  - Does the input **satisfy a designated property**, or
  - is **ϵ-far from satisfying the property**?

Brief introduction to property testing (contd.)

- In property testing, we use ϵ-far to say that the input is far from a certain property.
- ϵ: the least fraction of the input needs to be modified.

The model used in this talk (graph property)

- A graph $G(V, E)$ represented by an adjacency-matrix.
  - A query: to see if two vertices u and v are adjacent or not.
- ϵ-far from satisfying $\mathbb{P}$:
  - $\geq \epsilon n^2$ edges should be deleted or added to make G satisfy $\mathbb{P}$.

Focus of this talk

Theorem (Main Theorem)

Let H be a fixed undirected graph that contains at least one triangle. Then there exists a constant $c = c(H) > 0$ such that the query complexity of any one-sided error property tester for induced H-freeness is at least

$$\left(\frac{1}{\epsilon}\right)^{c\log(1/\epsilon)}.$$

Outline

1. Introduction
   - Brief introduction to property testing
   - Focus of this talk
2. Two technical skills
   - h-sum-free sets
   - s-blow-up
3. Two main lemmas
4. Proof of the main theorem
5. Go back to the proof of Lemma 1

h-sum-free sets

- An approach in additive number theory.
  - Invented by Felix A. Behrend (1946)
  - On sets of integers which contain no three terms in arithmetic progression.
- A set $X \subseteq [m] = \{1, \ldots, m\}$ is called h-sum-free if
  - for every pair of positive integers $a, b \leq h$, if $x, y, z \in X$ satisfy the equation $ax + by = (a + b)z$ then $x = y = z$.
- That is, whenever $a, b \leq h$, the only solution to the equation that uses values from X is one of the $|X|$ *trivial solutions*.

h-sum-free sets

- An approach in additive number theory.
  - Invented by Felix A. Behrend (1946)
  - On sets of integers which contain no three terms in arithmetic progression.
- A set $X \subseteq [m] = \{1, \ldots, m\}$ is called h-sum-free if
  - for every pair of positive integers $a, b \leq h$, if $x, y, z \in X$ satisfy the equation $ax + by = (a + b)z$ then $x = y = z$.
- That is, whenever $a, b \leq h$, the only solution to the equation that uses values from X is one of the $|X|$ *trivial solutions*.

h-sum-free sets

- An approach in additive number theory.
  - Invented by Felix A. Behrend (1946)
  - On sets of integers which contain no three terms in arithmetic progression.
- A set $X \subseteq [m] = \{1, \ldots, m\}$ is called h-sum-free if
  - ▷ for every pair of positive integers $a, b \leq h$, if $x, y, z \in X$ satisfy the equation $ax + by = (a + b)z$ then $x = y = z$.
- That is, whenever $a, b \leq h$, the only solution to the equation that uses values from X is one of the $|X|$ *trivial solutions*.

h-sum-free sets (contd.)

- Example 1: $h = 1$, $m = 8$,
  - The only equation is $x + y = 2z$,
  - $X = \{1, 2, 4, 8\}$ is h-sum-free (i.e., no three terms in arithmetic progression).
- Example 2: $h = 2$, $m = 8$,
  - The possible equations are $x + y = 2z$, $x + 2y = 3z$, $2x + y = 3z$, and $2x + 2y = 4z$.
  - $X = \{1, 2, 4, 8\}$ is NOT h-sum-free.
  - $X' = \{1, 2, 8\}$ is h-sum-free.

h-sum-free sets (contd.)

- Example 1: $h = 1$, $m = 8$,
  - The only equation is $x + y = 2z$,
  - $X = \{1, 2, 4, 8\}$ is h-sum-free (i.e., no three terms in arithmetic progression).
- Example 2: $h = 2$, $m = 8$,
  - The possible equations are $x + y = 2z$, $x + 2y = 3z$, $2x + y = 3z$, and $2x + 2y = 4z$.
  - $X = \{1, 2, 4, 8\}$ is NOT h-sum-free.
  - $X' = \{1, 2, 8\}$ is h-sum-free.

h-sum-free sets (contd.)

- Example 1: $h = 1$, $m = 8$,
  - The only equation is $x + y = 2z$,
  - $X = \{1, 2, 4, 8\}$ is h-sum-free (i.e., no three terms in arithmetic progression).
- Example 2: $h = 2$, $m = 8$,
  - The possible equations are $x + y = 2z$, $x + 2y = 3z$, $2x + y = 3z$, and $2x + 2y = 4z$.
  - $X = \{1, 2, 4, 8\}$ is NOT h-sum-free.
  - $X' = \{1, 2, 8\}$ is h-sum-free.

h-sum-free sets (contd.)

- Example 1: $h = 1$, $m = 8$,
  - The only equation is $x + y = 2z$,
  - $X = \{1, 2, 4, 8\}$ is h-sum-free (i.e., no three terms in arithmetic progression).
- Example 2: $h = 2$, $m = 8$,
  - The possible equations are $x + y = 2z$, $x + 2y = 3z$, $2x + y = 3z$, and $2x + 2y = 4z$.
  - $X = \{1, 2, 4, 8\}$ is NOT h-sum-free.
  - $X' = \{1, 2, 8\}$ is h-sum-free.

h-sum-free sets (contd.)

- Example 1: $h = 1$, $m = 8$,
  - The only equation is $x + y = 2z$,
  - $X = \{1, 2, 4, 8\}$ is h-sum-free (i.e., no three terms in arithmetic progression).
- Example 2: $h = 2$, $m = 8$,
  - The possible equations are $x + y = 2z$, $x + 2y = 3z$, $2x + y = 3z$, and $2x + 2y = 4z$.
  - $X = \{1, 2, 4, 8\}$ is NOT h-sum-free.
  - $X' = \{1, 2, 8\}$ is h-sum-free.

Lemma 1

For every positive integer m, there exists an h-sum-free subset $X \subset [m] = \{1, 2, \dots, m\}$ *of size at least*

$$|X| \geq \frac{m}{e^{10\sqrt{\log h \log m}}}.$$

Outline

s-blow-up

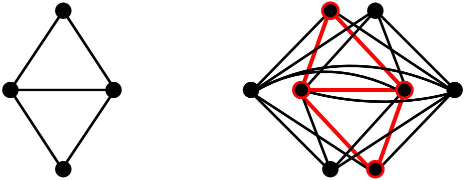
For convenience, we start the discussion with digraphs (the results for undirected graphs will be obtained as a special case).

- An s-blow-up of a digraph $H = (V(H), E(H))$ on h vertices:
  - $v_i \in V(H) \xrightarrow{\text{replaced by}}$ an independent set I_i of size s;
  - $(v_i, v_j) \in E(H) \xrightarrow{\text{replaced by}}$ a complete bipartite directed subgraph (I_i, I_j) with edges directed from I_i to I_j.

s-blow-up (contd.)

- 3-blow-up of an edge.

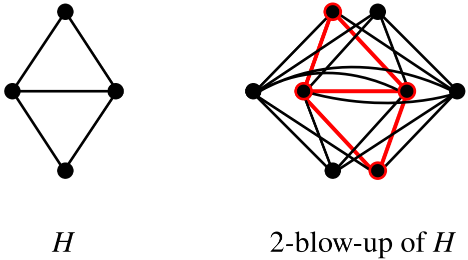
s-blow-up (contd.)



H 2-blow-up of H

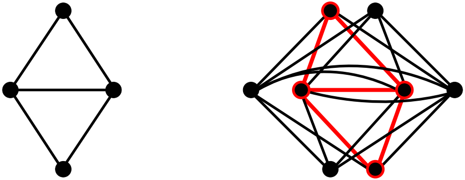
- Taking an s-blow-up of H $\Rightarrow$ getting a digraph on sh vertices that contains s^h induced copies of H.
- Each of these copies is called a special copy of H.

s-blow-up (contd.)



- Taking an s-blow-up of H $\Rightarrow$ getting a digraph on sh vertices that contains s^h induced copies of H.
- Each of these copies is called a special copy of H.

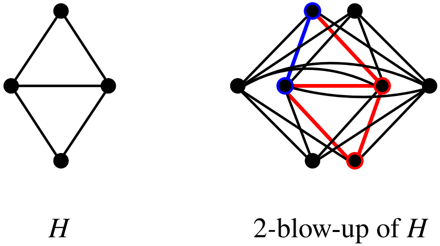
s-blow-up (contd.)



H 2-blow-up of H

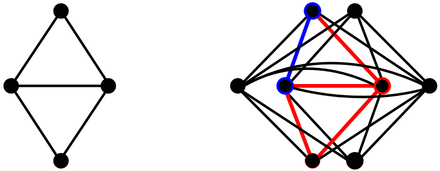
- Taking an s-blow-up of H $\Rightarrow$ getting a digraph on sh vertices that contains s^h induced copies of H.
- Each of these copies is called a special copy of H.

s-blow-up (contd.)



- Each pair of vertices in the blow-up is contained in $\leq s^{h-2}$ special copies of H.
- $\therefore$ adding or removing an edge from the blow-up can destroy $\leq s^{h-2}$ special copies of H.
- One must add or remove $\geq s^h/s^{h-2} = s^2$ edges from the blow-up to destroy all its special copies of H.

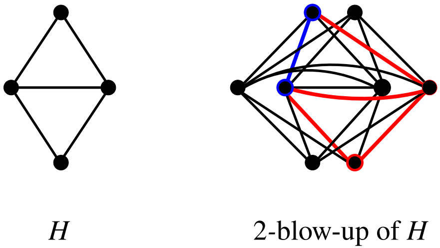
s-blow-up (contd.)



H 2-blow-up of H

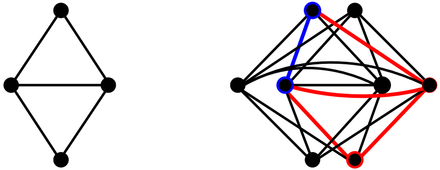
- Each pair of vertices in the blow-up is contained in $\leq s^{h-2}$ special copies of H.
- $\therefore$ adding or removing an edge from the blow-up can destroy $\leq s^{h-2}$ special copies of H.
- One must add or remove $\geq s^h/s^{h-2} = s^2$ edges from the blow-up to destroy all its special copies of H.

s-blow-up (contd.)



- Each pair of vertices in the blow-up is contained in $\leq s^{h-2}$ special copies of H.
- $\therefore$ adding or removing an edge from the blow-up can destroy $\leq s^{h-2}$ special copies of H.
- One must add or remove $\geq s^h / s^{h-2} = s^2$ edges from the blow-up to destroy all its special copies of H.

s-blow-up (contd.)



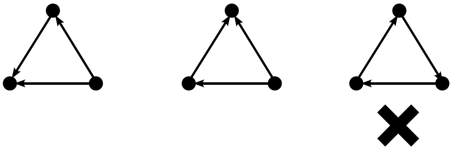
H 2-blow-up of H

- Each pair of vertices in the blow-up is contained in $\leq s^{h-2}$ special copies of H.
- $\therefore$ adding or removing an edge from the blow-up can destroy $\leq s^{h-2}$ special copies of H.
- One must add or remove $\geq s^h/s^{h-2} = s^2$ edges from the blow-up to destroy all its special copies of H.

Outline

Assumptions

- We start the discussion with **digraphs**.
- A triangle in a digraph is like:



The first main lemma

We have seen the following lemma:

Lemma 1

For every positive integer m, there exists an h-sum-free subset $X \subset [m] = \{1, 2, \ldots, m\}$ of size at least

$$|X| \geq \frac{m}{e^{10\sqrt{\log h \log m}}}.$$

The second main lemma

Lemma 2

For every fixed digraph H on h vertices, that contains at least one triangle, there is a constant $c = c(H) > 0$, such that for every positive $\epsilon < \epsilon_0(H)$ and every integer $n > n_0(\epsilon)$, there is an n-vertex digraph G such that

- *G is ϵ-far from being induced H-free;*
- *yet G contains $\leq \epsilon^{c \log(1/\epsilon)} n^h$ induced copies of H.*

Proof of Lemma 2

- Given a small $\epsilon > 0$, and let m be the largest integer satisfying

$$\frac{1}{h^4 e^{10\sqrt{\log m \log h}}} \geq \epsilon.$$

- It is easy to check that this m satisfies

$$m \geq \left(\frac{1}{\epsilon}\right)^{c \log(1/\epsilon)},$$

for an appropriate $c = c(H) > 0$.

Proof of Lemma 2 (contd.)

- Let $X \subseteq \{1, 2, \ldots, m\}$ be the set as in Lemma 1.
- Call the vertices of H $v_1, v_2, \ldots, v_h$.
- Let $V_1, V_2, \ldots, V_h$ be pairwise disjoint sets of vertices, where
  - $|V_i| = im$ and the vertices in V_i are denoted by $1, 2, \ldots, im$.
  - With a slight abuse of notation, we think of the sets V_i as being pairwise disjoint.

Proof of Lemma 2 (contd.)

- We now construct a graph F whose vertex set is $V_1 \cup V_2 \cup \ldots \cup V_h$.
- For each j, $1 \leq j \leq m$, for each $x \in X$ and for each directed edge (v_p, v_q) of H:

$$j + (p-1)x \in V_p \quad \longrightarrow \quad j + (q-1)x \in V_q.$$

  - That is, for each $1 \leq j \leq m$ and $x \in X$, the graph F contains a **copy** of H spanned by the vertices $j,\ j+x,\ j+2x,\ \ldots,\ j+(h-1)x$.
  - $$t = j + (p-1)x \quad \longrightarrow \quad j + (q-1)x$$

    i.e.,

    $$t \quad \longrightarrow \quad t + (q-p)x.$$
  - $m|X|$ copies of H.

Proof of Lemma 2 (contd.)

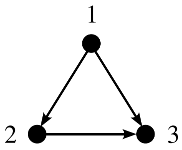
- Each of these $m|X|$ copies of H corresponds to an arithmetic progression whose first element is j $(1 \leq j \leq m)$ and whose difference is x $(x \in X)$.
- F contains $m|X|$ copies of H such that **each pair of copies have at most one common vertex**.
- Since each edge of F belongs to one of these copies, these $m|X|$ copies of H in F are in particular induced.
- We call these copies essential copies of H.

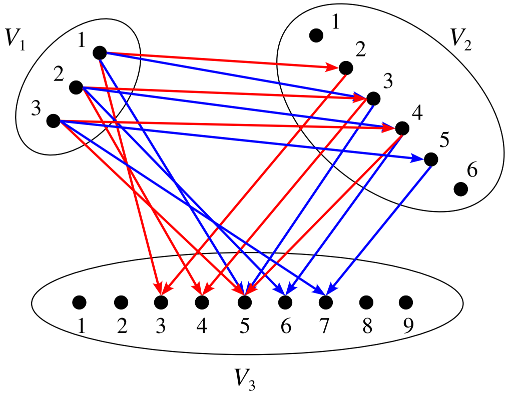
Proof of Lemma 2 (contd.)

- Define
$$s = \left\lfloor \frac{n}{|V(F)|} \right\rfloor = \left\lfloor \frac{2n}{h(h+1)m} \right\rfloor .$$
- Let G be the s-blow-up of F
  - Add some isolated vertices, if needed, to make sure the number of vertices is precisely n.
- After s-blow-up of F, we will derive **special copies** of the essential copies of H.

An illustration of F

Assume that $h = 3$, $m = 3$, so we have an h-sum-free set $X = \{1, 2\}$.





$$\text{Use } X \text{ and } H \xrightarrow{\text{construct } F} \text{essential copies of } H$$

$$\text{essential copies} \xrightarrow{s\text{-blow-up (construct } G)} \text{special copies of } H$$

The following two claims complete the proof of this lemma.

Claim 1

The digraph G is ϵ-far from being induced H-free.

Claim 2

The digraph G contains at most $\epsilon^{c\log(1/\epsilon)}n^h$ induced copies of H.

Proof of Claim 1

Claim 1

The digraph G is ϵ-far from being induced H-free.

Proof.

- The main idea of the proof:
  - Show that adding or removing an edge from G can destroy special copies that belong to at most one of the blow-ups of the essential copies of H in F.
    - (Recall) Two essential copies of H in F share at most one common vertex in F.
    - Their corresponding blow-ups in G, say Y_1 and Y_2, share at most one common **independent set**.
    - Hence a special copy of H in Y_1 and a special copy of H in Y_2 share at most one common vertex.

Proof of Claim 1

Claim 1

The digraph G is ϵ-far from being induced H-free.

Proof.

- The main idea of the proof:
  - Show that adding or removing an edge from G can destroy special copies that belong to at most one of the blow-ups of the essential copies of H in F.
    - (Recall) Two essential copies of H in F share at most one common vertex in F.
    - Their corresponding blow-ups in G, say Y_1 and Y_2, share at most one common **independent set**.
    - Hence a special copy of H in Y_1 and a special copy of H in Y_2 share at most one common vertex.

Proof of Claim 1

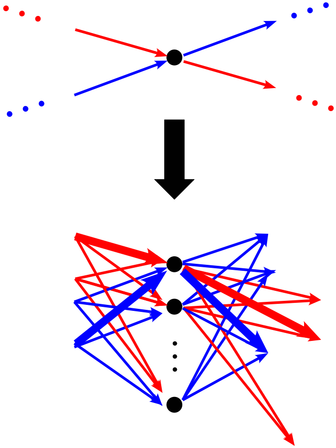
Claim 1

The digraph G is ϵ-far from being induced H-free.

Proof.

- The main idea of the proof:
  - Show that adding or removing an edge from G can destroy special copies that belong to at most one of the blow-ups of the essential copies of H in F.
    - (Recall) Two essential copies of H in F share at most one common vertex in F.
    - Their corresponding blow-ups in G, say Y_1 and Y_2, share at most one common **independent set**.
    - Hence a special copy of H in Y_1 and a special copy of H in Y_2 share at most one common vertex.

Proof of Claim 1 (contd.)



Proof of Claim 1 (contd.)

Proof. (contd.)

- To destroy all the special copies of *one* s-blow-up of H, one needs to add or delete $\geq s^2$ edges from the blow-up.
- Since G contains $m|X|$ blow-ups of essential copies of H which are all induced in F, we conclude that one has to add or delete

$$\geq s^2 m|X| = \frac{4n^2 m|X|}{h^2(h+1)^2 m^2} \geq \frac{|X|n^2}{h^4 m} \geq \frac{n^2}{h^4 e^{10\sqrt{\log m \log h}}} \geq \epsilon n^2$$

edges to make G induced H-free. □

Proof of Claim 2

Claim 2

The digraph G contains at most $\epsilon^{c\log(1/\epsilon)}n^h$ induced copies of H.

Proof.

- Our goal is to show that G contains $\leq \epsilon^{c\log(1/\epsilon)}n^3$ triangles.
  - $\because$ H contains ≥ 1 triangle and each triangle belongs to $\leq \binom{n}{h-3} \leq n^{h-3}$ copies of H.
- Let $\mathrm{BP}(V_i)$ denote the blow-up of the *im* vertices that belonged to V_i in F.
- We denote by I_v the independent set of vertices in G which replace the vertex v in F ($\therefore \mathrm{BP}(V_i) = \bigcup_{v\in V_i} I_v$).
- Consider a partition of $V(G)$ into h subsets $U_1, \ldots, U_h$, where $\mathrm{BP}(V_i) \subseteq U_i$.

A remark

- Note that if we show that:

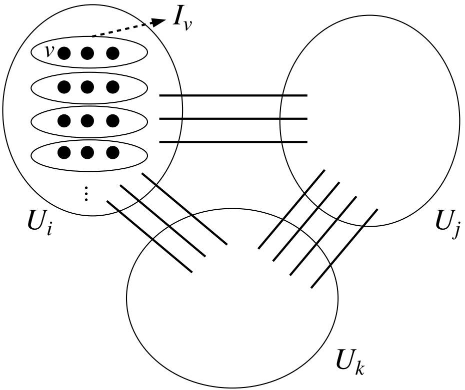
  the induced subgraphs of G on any three of the subsets $U_1, \ldots, U_h$ contains $\leq \epsilon^{c' \log(1/\epsilon)} n^3$ triangles,

  then the total number of triangles in G is $\leq \binom{h}{3} \epsilon^{c' \log(1/\epsilon)} n^3$,

  which is still $\leq \epsilon^{c \log(1/\epsilon)} n^3$, when a small enough $c = c(H)$ is chosen.

Proof. (contd.)

- Fix any three subsets U_i, U_j, U_k such that $1 \leq i < j < k \leq h$.
- A triangle spanned by U_i, U_j, U_k must have exactly one vertex in each of them.



Proof. (contd.)

- If U_i, U_j, U_k span a triangle with vertices belonging to $I_x \subseteq U_i$, $I_y \subseteq U_j$, and $I_z \subseteq U_k$, then the three vertices $x \in V_i$, $y \in V_j$, $z \in V_k$ in F must also span a triangle.
- Conversely, if $x \in V_i$, $y \in V_j$, $z \in V_k$ span a triangle in F, then for every choice of three vertices $u \in I_x \subseteq U_i$, $v \in I_y \subseteq U_j$, $w \in I_z \subseteq U_k$, the vertices u, v, w span a triangle in G.
- Therefore,

$$\#\{\text{triangles spanned by } U_i, U_j, U_k\}$$
$$= s^3 \cdot \#\{\text{triangles spanned by } V_i, V_j, V_k\}.$$

Proof. (contd.)

- Assume that v_i, v_j, v_k span a triangle in H in the following discussion.
  - If not, then by the definition of F, V_i, V_j, V_k do not span any triangle, and similarly U_i, U_j, U_K in G.
- Then by the definition of F, for any triangle spanned by V_i, V_j, V_k, there are $x, y \in X$ and $1 \leq t \leq im$ such that the three vertices of this triangle are

$$t \in V_i, \quad t + (j-i)x \in V_j, \quad t + (j-i)x + (k-j)y \in V_k.$$

  - t connects to $t + (j-i)x$ and $t + (j-i)x$ connects to $t + (j-i)x + (k-j)y$.

Proof. (contd.)

- Assume that v_i, v_j, v_k span a triangle in H in the following discussion.
  - If not, then by the definition of F, V_i, V_j, V_k do not span any triangle, and similarly U_i, U_j, U_K in G.
- Then by the definition of F, for any triangle spanned by V_i, V_j, V_k, there are $x, y \in X$ and $1 \leq t \leq im$ such that the three vertices of this triangle are

$$t \in V_i, \quad t + (j-i)x \in V_j, \quad t + (j-i)x + (k-j)y \in V_k.$$

  - t connects to $t + (j-i)x$ and $t + (j-i)x$ connects to $t + (j-i)x + (k-j)y$.

Proof. (contd.)

- Assume that v_i, v_j, v_k span a triangle in H in the following discussion.
  - If not, then by the definition of F, V_i, V_j, V_k do not span any triangle, and similarly U_i, U_j, U_K in G.
- Then by the definition of F, for any triangle spanned by V_i, V_j, V_k, there are $x, y \in X$ and $1 \leq t \leq im$ such that the three vertices of this triangle are

$$t \in V_i, \quad t + (j - i)x \in V_j, \quad t + (j - i)x + (k - j)y \in V_k.$$

  - t connects to $t + (j - i)x$ and $t + (j - i)x$ connects to $t + (j - i)x + (k - j)y$.

Proof. (contd.)

- Assume that v_i, v_j, v_k span a triangle in H in the following discussion.
  - If not, then by the definition of F, V_i, V_j, V_k do not span any triangle, and similarly U_i, U_j, U_K in G.
- Then by the definition of F, for any triangle spanned by V_i, V_j, V_k, there are $x, y \in X$ and $1 \leq t \leq im$ such that the three vertices of this triangle are

$$t \in V_i, \quad t + (j - i)x \in V_j, \quad t + (j - i)x + (k - j)y \in V_k.$$

  - t connects to $t + (j - i)x$ and $t + (j - i)x$ connects to $t + (j - i)x + (k - j)y$.

Proof. (contd.)

- As this is a triangle, there must also be an edge connecting t to $t + (j - i)x + (k - j)y$.
- Hence there exists $z \in X$ such that

$$t + (k - i)z = t + (j - i)x + (k - j)y.$$

- Thus we have $(j - i)x + (k - j)y = (k - i)z$.
- Since X is h-sum-free, we have $x = y = z$.

Proof. (contd.)

- Therefore, V_i, V_j, V_k span precisely $m|X|$ triangles, which are spanned by the vertices

$$t + (i-1)x \in V_i, \quad t + (j-1)x \in V_j, \quad t + (k-1)x \in V_k.,$$

  for every possible choice $t \in \{1, \ldots, m\}$ and $x \in X$.

- We conclude that U_i, U_j, U_k span

$$m|X|s^3 < m^2(n/m)^3 \le n^3/m \le \frac{n^3}{(1/\epsilon)^{c \log(1/\epsilon)}} = \epsilon^{c \log(1/\epsilon)} n^3$$

  triangles.

Outline

The main theorem can be proved by the previous lemmas

Main Theorem

Let H be a fixed undirected graph that contains at least one triangle. Then there exists a constant $c = c(H) > 0$ such that the query complexity of any one-sided error property tester for induced H-freeness is at least

$$\left(\frac{1}{\epsilon}\right)^{c\log(1/\epsilon)}.$$

- Here we left the details of the proof as an exercise.
    - *Hint: use Lemma 2, and apply two probabilistic strategies: union bound and Markov's inequality.*

Outline

1. Introduction
   - Brief introduction to property testing
   - Focus of this talk
2. Two technical skills
   - h-sum-free sets
   - s-blow-up
3. Two main lemmas
4. Proof of the main theorem
5. Go back to the proof of Lemma 1

Recall Lemma 1

Lemma 1

For every positive integer m, there exists an h-sum-free subset $X \subset [m] = \{1, 2, \ldots, m\}$ of size at least

$$|X| \geq \frac{m}{e^{10\sqrt{\log h \log m}}}.$$

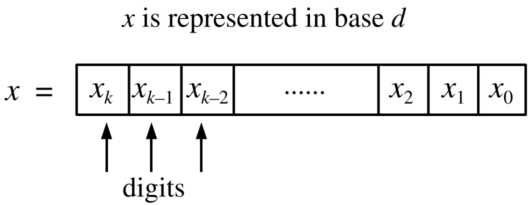
Proof of Lemma 1

Proof.

- Let d and r be integers (to be chosen later) and define:

$$S_r = \left\{ \sum_{i=0}^{k} x_i d^i \mid x_i < \frac{d}{2h} \text{ for } 0 \le i \le k \text{ and } \sum_{i=0}^{k} x_i^2 = r \right\},$$

where $k = \lfloor \log m / \log d \rfloor - 1 = \lfloor \log_d m \rfloor - 1$.



Proof. (contd.)

- We claim that S_r is h-sum-free for every d and r.
- Assume that there are $x, y, z \in S_r$ that satisfy the equation $ax + by = (a + b)z$, where $a, b \leq h$ are positive integers and

$$x = \sum_{i=0}^{k} x_i d^i, \quad y = \sum_{i=0}^{k} y_i d^i, \quad z = \sum_{i=0}^{k} z_i d^i.$$

- By definition, $x_i, y_i, z_i < d/(2h)$, and $a, b \leq h$, there is no carry in the base-d addition of the numbers in S_r.
  - That is, $ax_i + by_i = (a + b)z_i$ (i.e., z_i is a weighted average of x_i and y_i).

Proof. (contd.)

- **Fact:** $f(z) = z^2$ is a convex function, so by Jensen's inequality we have
$$ax_i^2 + by_i^2 \geq (a+b)z_i^2,$$
and the inequality is *strict* unless $x_i = y_i = z_i$.
- However, if for some i the inequality is strict, we have
$$a\sum_{i=0}^{k} x_i^2 + b\sum_{i=0}^{k} y_i^2 > (a+b)\sum_{i=0}^{k} z_i^2,$$
which is impossible since by definition
$$\sum_{i=0}^{k} x_i^2 = \sum_{i=0}^{k} y_i^2 = \sum_{i=0}^{k} z_i^2 = r.$$
- Thus $x_i = y_i = z_i$ for all i and S_r is h-sum-free.

Proof of Lemma 1 (contd.)

Proof. (contd.)

- Next we complete the proof by showing that, for some r, the set S_r has size at least $m/e^{10\sqrt{\log h \log m}}$.
- The integer r in the definition of S_r satisfies $r = \sum_{i=0}^{k} x_i^2 \leq (k+1)(d/2h)^2 < kd^2$.
- The union of the sets S_r has size $(d/2h)^{k+1} > (d/2h)^k$.
- It follows that for some r, the set S_r satisfies $|S_r| \geq (d/2h)^k/kd^2$.

Proof of Lemma 1 (contd.)

Proof. (contd.)

- Setting $d = e^{\sqrt{\log m \log h}}$

$$\therefore k = \left\lfloor \frac{\log m}{\log d} \right\rfloor = \left\lfloor \frac{\log m}{\sqrt{\log m \log h}} \right\rfloor \approx \sqrt{\frac{\log m}{\log h}}.$$

-

$$|S_r| \geq \frac{d^k}{(2h)^k k d^2} = \frac{e^{\sqrt{\log m \log h} \cdot \sqrt{\log m / \log h}}}{(2h)^k k d^2}$$

$$= \frac{m}{(2h)^{\sqrt{\log m / \log h}} \cdot \sqrt{\log m / \log h} \cdot e^{2\sqrt{\log m \log h}}}$$

$$= \frac{m}{e^{(\log 2h)\sqrt{\log m / \log h}} \cdot \sqrt{\log m / \log h} \cdot e^{2\sqrt{\log m \log h}}}$$

$$> \frac{m}{e^{10\sqrt{\log m \log h}}}.$$

which is as required.

Thank you!