# Nearly Complete Graphs Decomposable Into Large Induced Matchings

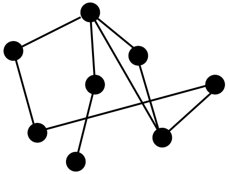
Ankur Moitra, IAS

joint with Noga Alon (IAS, TAU) and Benny Sudakov (UCLA)

May 22, 2012

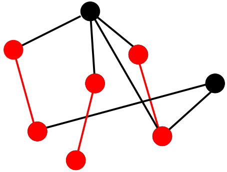
# Rusza-Szemeredi Graphs

**Induced Matching:** an induced subgraph whose edges form a matching



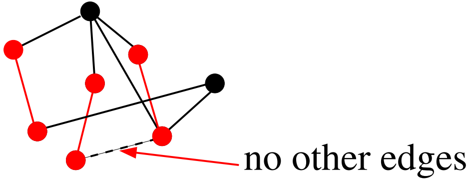
# Rusza-Szemeredi Graphs

**Induced Matching:** an induced subgraph whose edges form a matching



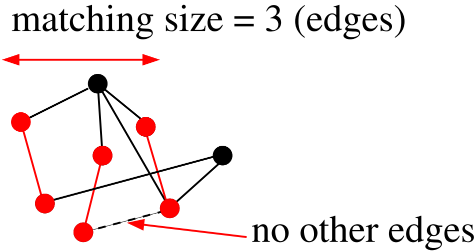
# Rusza-Szemeredi Graphs

**Induced Matching:** an induced subgraph whose edges form a matching



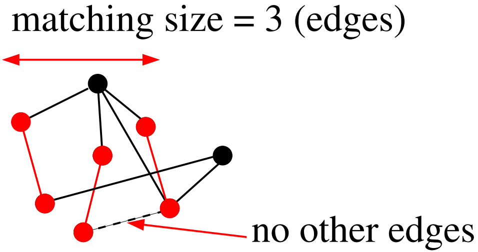
# Rusza-Szemeredi Graphs

**Induced Matching:** an induced subgraph whose edges form a matching



# Rusza-Szemeredi Graphs

**Induced Matching:** an induced subgraph whose edges form a matching

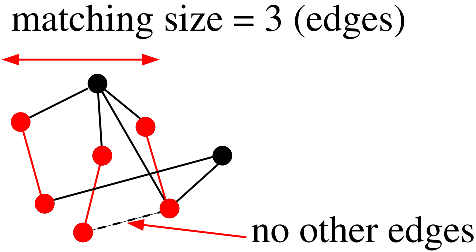



**Definition**

An $(r, t)$-Rusza-Szemeredi graph $G$ is a union of $t$ pairwise disjoint induced matchings, each of size $r$ (hence $rt$ total edges)

# Rusza-Szemeredi Graphs

**Induced Matching:** an induced subgraph whose edges form a matching



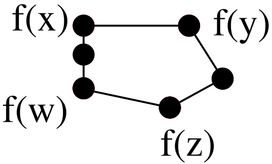

**Definition**

An $(r, t)$-Rusza-Szemeredi graph $G$ is a union of $t$ pairwise disjoint induced matchings, each of size $r$ (hence $rt$ total edges)

**Note:** An induced matching of size $r \Rightarrow$ missing $r(r-1)$ edges

# Applications of Induced Matchings

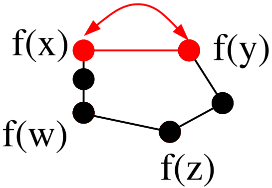
**Graph Test:** [Blum, Luby, Rubinfeld] [Samarodnitsky, Trevisan] [Hastad, Wigderson]



# Applications of Induced Matchings

**Graph Test:** [Blum, Luby, Rubinfeld] [Samarodnitsky, Trevisan] [Hastad, Wigderson]

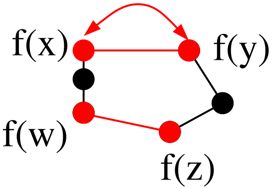
$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$



# Applications of Induced Matchings

**Graph Test:** [Blum, Luby, Rubinfeld] [Samarodnitsky, Trevisan] [Hastad, Wigderson]

$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$

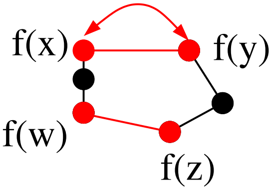



parallel linearity tests

# Applications of Induced Matchings

**Graph Test:** [Blum, Luby, Rubinfeld] [Samarodnitsky, Trevisan] [Hastad, Wigderson]

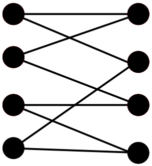
$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$




parallel linearity tests

**Broadcast Network:**

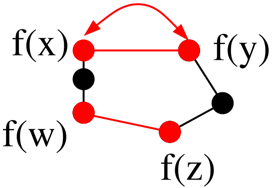
[Birk, Linial, Meshulam]



# Applications of Induced Matchings

**Graph Test:** [Blum, Luby, Rubinfeld]
[Samarodnitsky, Trevisan] [Hastad, Wigderson]

$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$

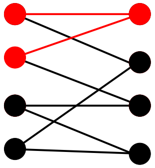



parallel linearity tests

**Broadcast Network:**

[Birk, Linial, Meshulam]

Collision!

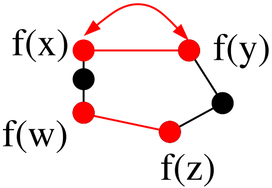


# Applications of Induced Matchings

## Graph Test: [Blum, Luby, Rubinfeld]

[Samarodnitsky, Trevisan] [Hastad, Wigderson]

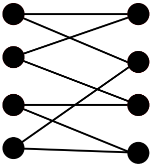
$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$




parallel linearity tests

## Broadcast Network:

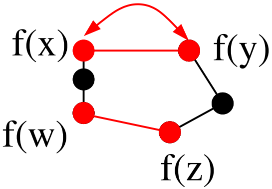
[Birk, Linial, Meshulam]



# Applications of Induced Matchings

**Graph Test:** [Blum, Luby, Rubinfeld] [Samarodnitsky, Trevisan] [Hastad, Wigderson]

$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$

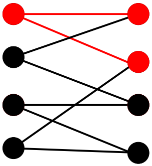



parallel linearity tests

**Broadcast Network:**

[Birk, Linial, Meshulam]

Collision!

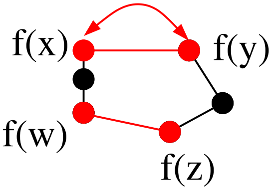


# Applications of Induced Matchings

## Graph Test: [Blum, Luby, Rubinfeld]

[Samarodnitsky, Trevisan] [Hastad, Wigderson]

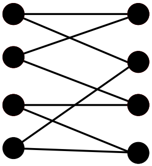
$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$




parallel linearity tests

## Broadcast Network:

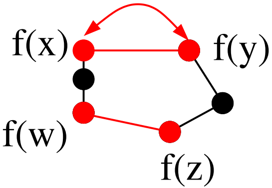
[Birk, Linial, Meshulam]



# Applications of Induced Matchings

**Graph Test:** [Blum, Luby, Rubinfeld]
[Samarodnitsky, Trevisan] [Hastad, Wigderson]

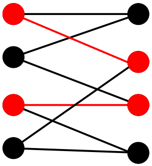
$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$




parallel linearity tests

**Broadcast Network:**

[Birk, Linial, Meshulam]



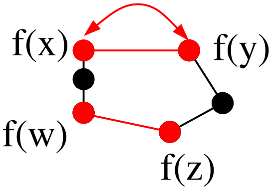
collisionless broadcast

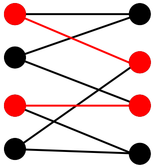
# Applications of Induced Matchings

**Graph Test:** [Blum, Luby, Rubinfeld] [Samarodnitsky, Trevisan] [Hastad, Wigderson]

**Broadcast Network:** [Birk, Linial, Meshulam]

$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$



parallel linearity tests

collisionless broadcast

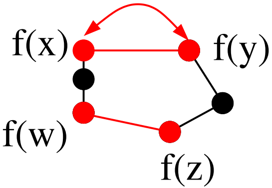
**Property Testing:** triangle–free, monotonicity, ...

# Applications of Induced Matchings

**Additive Combinatorics:** arithmetic progressions

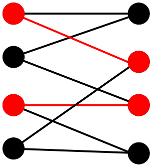
**Graph Test:** [Blum, Luby, Rubinfeld]
[Samarodnitsky, Trevisan] [Hastad, Wigderson]

$$f(x+y) \stackrel{?}{=} f(x) + f(y)$$




parallel linearity tests

**Broadcast Network:**
[Birk, Linial, Meshulam]



collisionless broadcast

**Property Testing:** triangle–free, monotonicity, ...

# Known Constructions

# Known Constructions

- [Rusza, Szemeredi, 1978] $r = \frac{N}{e^{\sqrt{\log N}}}$, $t = \frac{N}{3}$
  (using a construction of Behrend)

# Known Constructions

- [Rusza, Szemeredi, 1978] $r = \frac{N}{e^{\sqrt{\log N}}}$, $t = \frac{N}{3}$ (using a construction of Behrend)
- [Frankl, Furedi, 1987], for any $r = O(1)$, $rt = (1 - o(1))\binom{N}{2}$

# Known Constructions

- [Rusza, Szemeredi, 1978] $r = \frac{N}{e^{\sqrt{\log N}}}$, $t = \frac{N}{3}$ (using a construction of Behrend)
- [Frankl, Furedi, 1987], for any $r = O(1)$, $rt = (1 - o(1))\binom{N}{2}$
- [Fischer et al, 2002] $r = \frac{N}{3} - o(N)$, $t = N^{1/\log\log N}$

# Known Constructions

- [Rusza, Szemeredi, 1978] $r = \frac{N}{e^{\sqrt{\log N}}}, t = \frac{N}{3}$ (using a construction of Behrend)
- [Frankl, Furedi, 1987], for any $r = O(1)$, $rt = (1 - o(1))\binom{N}{2}$
- [Fischer et al, 2002] $r = \frac{N}{3} - o(N), t = N^{1/\log\log N}$
- [Birk, Linial, Meshulam, 1993] and [Meshulam, 2011] $r = (\log N)^{\log\log N/(\log\log\log N)^2}$, $rt = \frac{N^2}{24}$

# Known Constructions

- [Rusza, Szemeredi, 1978] $r = \frac{N}{e^{\sqrt{\log N}}}, t = \frac{N}{3}$
  (using a construction of Behrend)
- [Frankl, Furedi, 1987], for any $r = O(1)$, $rt = (1 - o(1))\binom{N}{2}$
- [Fischer et al, 2002] $r = \frac{N}{3} - o(N), t = N^{1/\log\log N}$
- [Birk, Linial, Meshulam, 1993] and [Meshulam, 2011]
  $r = (\log N)^{\log\log N/(\log\log\log N)^2}$, $rt = \frac{N^2}{24}$

**Question**

*Can Rusza-Szemeredi graphs be dense (for $r = N^{\Omega(1)}$)?*

# Our Results

There are very dense graphs with nearly linear induced matchings!

# Our Results

There are very dense graphs with nearly linear induced matchings!

**Theorem**

*There are $(r, t)$-Rusza-Szemeredi graphs with:*

- $rt = \binom{N}{2} - N^{2-\delta(\epsilon)}$, *yet* $r = N^{1-\epsilon}$ *for any* $\epsilon > 0 \quad (\delta(\epsilon) \sim \epsilon^4)$

# Our Results

There are very dense graphs with nearly linear induced matchings!

**Theorem**

*There are $(r, t)$-Rusza-Szemeredi graphs with:*

- $rt = \binom{N}{2} - N^{2-\delta(\epsilon)}$, *yet* $r = N^{1-\epsilon}$ *for any* $\epsilon > 0 \quad (\delta(\epsilon) \sim \epsilon^4)$
- $rt = \binom{N}{2} - N^{3/2+\epsilon}$, *yet* $r = N^{\delta(\epsilon)}$ *for any* $\epsilon > 0 \quad (\delta(\epsilon) \sim \epsilon^2)$

*and if* $rt = \binom{N}{2} - N^{3/2}$, *then* $r = O(1)$.

# Our Results

There are very dense graphs with nearly linear induced matchings!

**Theorem**

*There are $(r, t)$-Rusza-Szemeredi graphs with:*

- $rt = \binom{N}{2} - N^{2-\delta(\epsilon)}$, *yet* $r = N^{1-\epsilon}$ *for any* $\epsilon > 0$ $\quad (\delta(\epsilon) \sim \epsilon^4)$
- $rt = \binom{N}{2} - N^{3/2+\epsilon}$, *yet* $r = N^{\delta(\epsilon)}$ *for any* $\epsilon > 0$ $\quad (\delta(\epsilon) \sim \epsilon^2)$

*and if* $rt = \binom{N}{2} - N^{3/2}$, *then* $r = O(1)$.

**Corollary:** $K_N = G_1 + G_2 \ldots G_{f(\epsilon)}$, where each $G_i$ can be covered by $N^{1+\epsilon}$ induced matchings

# Our Results

There are very dense graphs with nearly linear induced matchings!

**Theorem**

*There are $(r, t)$-Rusza-Szemeredi graphs with:*

- $rt = \binom{N}{2} - N^{2-\delta(\epsilon)}$, *yet* $r = N^{1-\epsilon}$ *for any* $\epsilon > 0$ $\quad (\delta(\epsilon) \sim \epsilon^4)$
- $rt = \binom{N}{2} - N^{3/2+\epsilon}$, *yet* $r = N^{\delta(\epsilon)}$ *for any* $\epsilon > 0$ $\quad (\delta(\epsilon) \sim \epsilon^2)$

*and if* $rt = \binom{N}{2} - N^{3/2}$, *then* $r = O(1)$.

**Corollary:** $K_N = G_1 + G_2 \ldots G_{f(\epsilon)}$, where each $G_i$ can be covered by $N^{1+\epsilon}$ induced matchings

We give applications to linearity testing, routing in broadcast networks, and disprove conjectures of Meshulam and Vempala

# Our Constructions

# Our Constructions

- Our first construction is **geometric**, and uses only basic volume arguments

  (inspired by a construction of Fox and Loh)

# Our Constructions

- Our first construction is **geometric**, and uses only basic volume arguments

  (inspired by a construction of Fox and Loh)

- Our second construction is based on **error correcting codes**

# A Geometric Construction

# A Geometric Construction

### The Graph

*Let* $C$ *be a large constant,* $V = \{1, 2, ..., C\}^d$ **(nodes)** $(|V| = N)$

# A Geometric Construction

### The Graph

*Let* $C$ *be a large constant,* $V = \{1, 2, \ldots, C\}^d$ **(nodes)** $(|V| = N)$

*Let* $\mu = E_{x,y \in V}[\|x - y\|_2^2]$,

# A Geometric Construction

### The Graph

*Let $C$ be a large constant,* $V = \{1, 2, ..., C\}^d$ **(nodes)** $(|V| = N)$

*Let* $\mu = E_{x,y \in V}[\|x - y\|_2^2]$, $E = \left\{(x, y) \middle| \left| \|x - y\|_2^2 - \mu \right| \leq d \right\}$ **(edges)**

# A Geometric Construction

### The Graph

*Let $C$ be a large constant, $V = \{1, 2, ..., C\}^d$* **(nodes)** $(|V| = N)$

*Let $\mu = E_{x,y\in V}[\|x-y\|_2^2]$, $E = \left\{(x,y) \middle| \left|\|x-y\|_2^2 - \mu\right| \le d\right\}$* **(edges)**

**Claim** (easy): $\binom{N}{2} - |E| \le \binom{N}{2} e^{-d/2C^4} = N^{2-\delta}$ via Hoeffding Bound

# A Geometric Construction

**The Graph**

*Let $C$ be a large constant,* $V = \{1, 2, ..., C\}^d$ **(nodes)** $(|V| = N)$

*Let* $\mu = E_{x,y \in V}[\|x - y\|_2^2]$, $E = \left\{(x, y) \middle| \left| \|x - y\|_2^2 - \mu \right| \leq d \right\}$ **(edges)**

**Claim** (easy): $\binom{N}{2} - |E| \leq \binom{N}{2} e^{-d/2C^4} = N^{2-\delta}$ via Hoeffding Bound

**Plan**

*1. Cover $G$ with $N$ induced subgraphs each of max degree $\Delta$*

# A Geometric Construction

## The Graph

*Let $C$ be a large constant, $V = \{1, 2, ..., C\}^d$* **(nodes)** $(|V| = N)$

*Let $\mu = E_{x,y \in V}[\|x - y\|_2^2]$, $E = \left\{(x, y) \middle| \left| \|x - y\|_2^2 - \mu \right| \leq d\right\}$* **(edges)**

**Claim** (easy): $\binom{N}{2} - |E| \leq \binom{N}{2} e^{-d/2C^4} = N^{2-\delta}$ via Hoeffding Bound

## Plan

1. *Cover $G$ with $N$ induced subgraphs each of max degree $\Delta$*
2. *Decompose each induced subgraph into $O(\Delta^2)$ induced matchings*

# Goal: Cover $G$ by $N$ Induced Subgraphs

What induced subgraphs should we use?

# Goal: Cover $G$ by $N$ Induced Subgraphs

What induced subgraphs should we use?

**Definition**

$G_z \equiv$ induced subgraph on $V_z = \left\{ x \middle| \left| \|x - z\|_2^2 - \mu/4 \right| \leq 3/4d \right\}$

# Goal: Cover $G$ by $N$ Induced Subgraphs

What induced subgraphs should we use?

**Definition**

$G_z \equiv$ induced subgraph on $V_z = \left\{ x \middle| \left| \|x - z\|_2^2 - \mu/4 \right| \leq 3/4d \right\}$

Do $\cup_z G_z$ cover the edges of $G$?

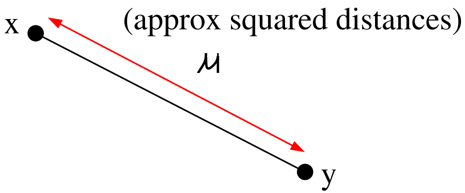
# Goal: Cover $G$ by $N$ Induced Subgraphs

What induced subgraphs should we use?

**Definition**

$G_z \equiv$ induced subgraph on $V_z = \left\{ x \middle| \left| \|x-z\|_2^2 - \mu/4 \right| \leq 3/4d \right\}$

Do $\cup_z G_z$ cover the edges of $G$?



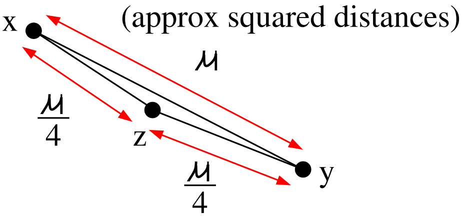
# Goal: Cover $G$ by $N$ Induced Subgraphs

What induced subgraphs should we use?

**Definition**

$G_z \equiv$ induced subgraph on $V_z = \left\{ x \middle| \left| \|x - z\|_2^2 - \mu/4 \right| \leq 3/4d \right\}$

Do $\cup_z G_z$ cover the edges of $G$?



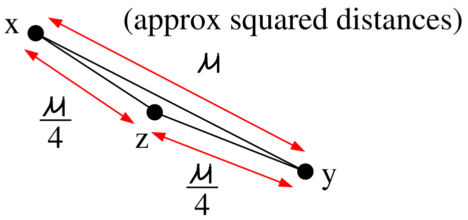
# Goal: Cover $G$ by $N$ Induced Subgraphs

What induced subgraphs should we use?

**Definition**

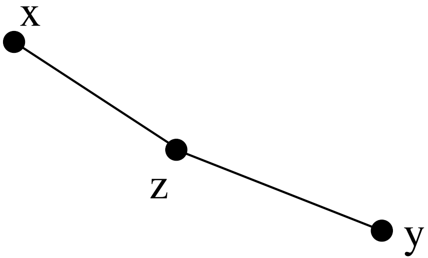
$G_z \equiv$ induced subgraph on $V_z = \left\{ x \middle| \left| \|x - z\|_2^2 - \mu/4 \right| \leq 3/4d \right\}$

Do $\cup_z G_z$ cover the edges of $G$? Yes, choose $z \approx$ the midpoint of $x, y$



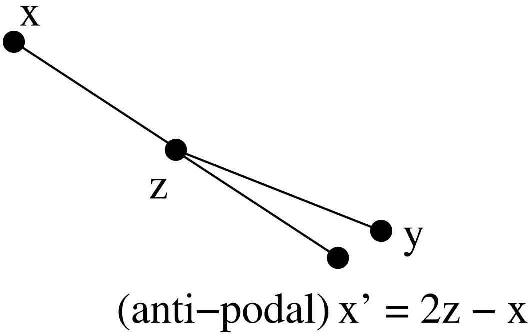
# Goal: Bound the Max Degree

**Claim:** $(x, y) \in E_z$ implies $y$ is close to $x'$



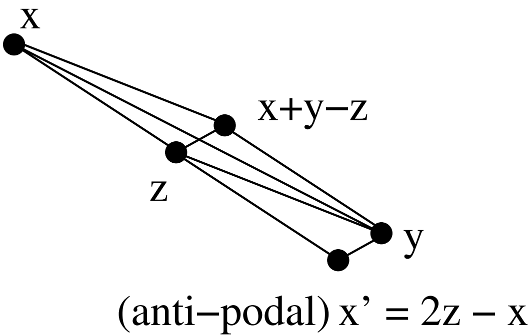
# Goal: Bound the Max Degree

**Claim:** $(x, y) \in E_z$ implies $y$ is close to $x'$



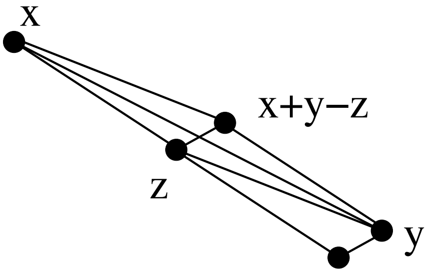
# Goal: Bound the Max Degree

**Claim:** $(x, y) \in E_z$ implies $y$ is close to $x'$



# Goal: Bound the Max Degree

**Claim:** $(x, y) \in E_z$ implies $y$ is close to $x'$




(anti–podal) x’ = 2z − x

**Parallelogram Law:**

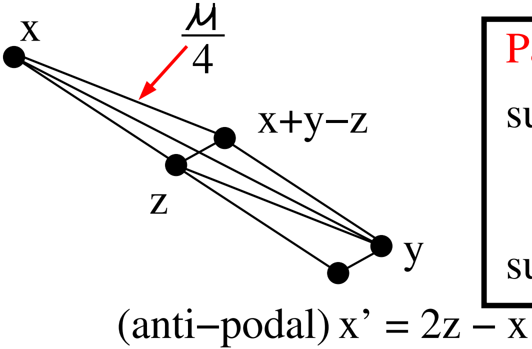
sum of squared sides

=

sum of squared diags

# Goal: Bound the Max Degree

**Claim:** $(x, y) \in E_z$ implies $y$ is close to $x'$




Parallelogram Law:

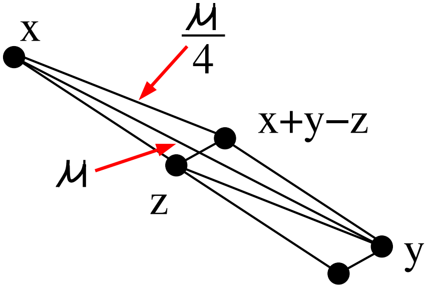
sum of squared sides

=

sum of squared diags

# Goal: Bound the Max Degree

**Claim:** $(x, y) \in E_z$ implies $y$ is close to $x'$




(anti–podal) $x' = 2z - x$

**Parallelogram Law:**

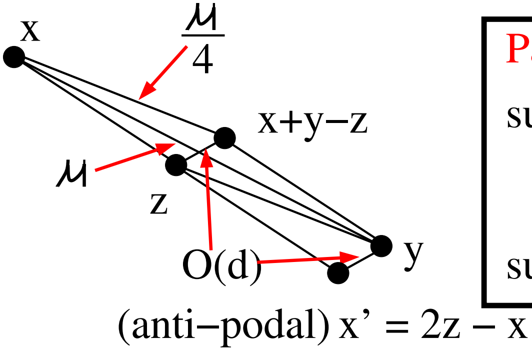
sum of squared sides

=

sum of squared diags

# Goal: Bound the Max Degree

**Claim:** $(x, y) \in E_z$ implies $y$ is close to $x'$




Parallelogram Law:

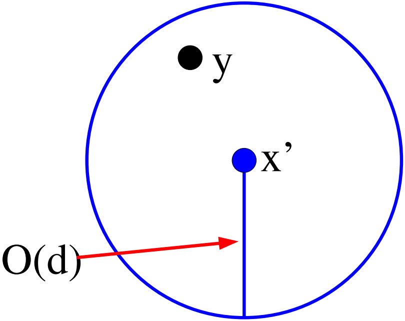
sum of squared sides

=

sum of squared diags

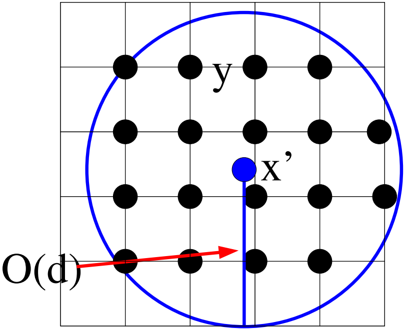
# Goal: Bound the Max Degree

Degree of $x$ (in $G_z$) is bounded by number of $y$ close to $x'$:



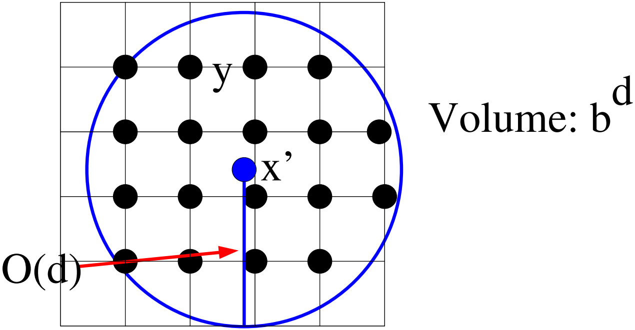
# Goal: Bound the Max Degree

Degree of $x$ (in $G_z$) is bounded by number of $y$ close to $x'$:



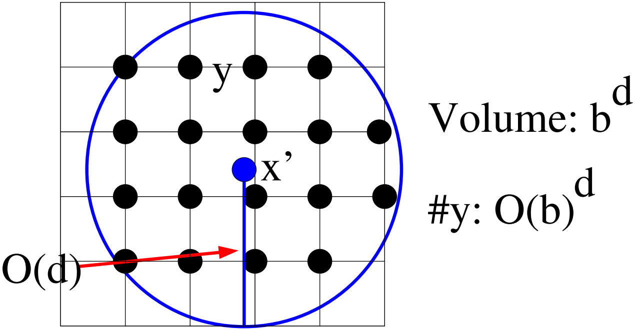
# Goal: Bound the Max Degree

Degree of $x$ (in $G_z$) is bounded by number of $y$ close to $x'$:

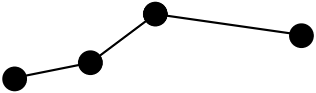


# Goal: Bound the Max Degree

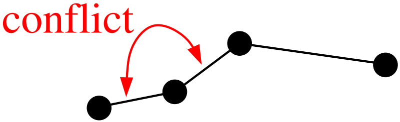
Degree of $x$ (in $G_z$) is bounded by number of $y$ close to $x'$:



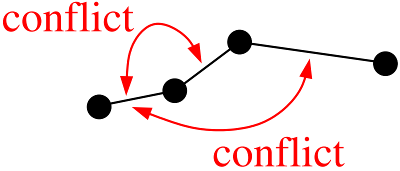
# Partitioning $G_z$ into Induced Matchings



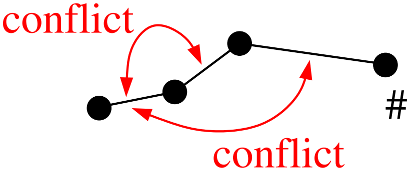
# Partitioning $G_z$ into Induced Matchings



# Partitioning $G_z$ into Induced Matchings

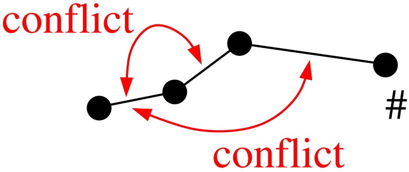


# Partitioning $G_z$ into Induced Matchings




# conflicts: $O(\triangle^2)$

# Partitioning $G_z$ into Induced Matchings

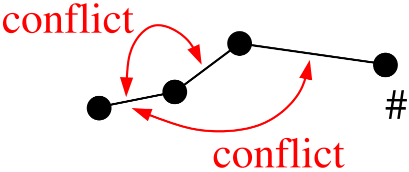



# conflicts: $O(\triangle^2)$

Intialize: empty matchings

$M_1 \quad M_2 \quad \cdots \quad M_{O(\triangle^2)}$

# Partitioning $G_z$ into Induced Matchings




# conflicts: $O(\triangle^2)$

Intialize: empty matchings

$$M_1 \qquad M_2 \qquad \cdots \qquad M_{O(\triangle^2)}$$

Place each edge in first matching w/o conflicts

# Our Constructions

- Our first construction is **geometric**, and uses only basic volume arguments

  (inspired by a construction of Fox and Loh)

- Our second construction is based on **error correcting codes**

# Our Constructions

- Our first construction is **geometric**, and uses only basic volume arguments

  (inspired by a construction of Fox and Loh)

- Our second construction is based on **error correcting codes**

- We use an entropy argument to show that if $rt = \binom{N}{2} - N^{3/2}$, then $r = O(1)$

# Our Constructions

- Our first construction is **geometric**, and uses only basic volume arguments
  (inspired by a construction of Fox and Loh)
- Our second construction is based on **error correcting codes**
- We use an entropy argument to show that if $rt = \binom{N}{2} - N^{3/2}$, then $r = O(1)$
- **Known:** The triangle removal lemma is **equivalent** to no $(r, t)$-RS graphs in certain ranges (e.g. no dense graphs with $t = N \log^* N$)

# Implications and Open Questions

We apply our constructions to:

# Implications and Open Questions

We apply our constructions to:

- extend the analysis of Hastad and Wigderson of the Graph Test (modestly better parameters)

# Implications and Open Questions

We apply our constructions to:

- extend the analysis of Hastad and Wigderson of the Graph Test (modestly better parameters)
- give a broadcast protocol for a problem of Birk, Linial and Meshulam that runs in $N^{1+\epsilon}$ rounds (previous best was $N^2/(\log N)^{\log \log N}$)

# Implications and Open Questions

We apply our constructions to:

- extend the analysis of Hastad and Wigderson of the Graph Test (modestly better parameters)
- give a broadcast protocol for a problem of Birk, Linial and Meshulam that runs in $N^{1+\epsilon}$ rounds (previous best was $N^2/(\log N)^{\log\log N}$)

Are there RS graphs when $r = \Omega(N)$, $t = N^{\Omega(1)}$?
(applications to monotonicity testing)

# Implications and Open Questions

We apply our constructions to:

- extend the analysis of Hastad and Wigderson of the Graph Test (modestly better parameters)
- give a broadcast protocol for a problem of Birk, Linial and Meshulam that runs in $N^{1+\epsilon}$ rounds (previous best was $N^2/(\log N)^{\log\log N}$)

Are there RS graphs when $r = \Omega(N)$, $t = N^{\Omega(1)}$?
(applications to monotonicity testing)

Can our graphs be used to give an integrality gap for Directed Steiner Tree that is polynomial in the number of nodes?
(best so far is poly-logarithmic)

Questions?

Thanks!