Independent Sets in Sparse Graphs

Approximation Algorithms
via
Ramsey-theoretic ideas

Anupam Gupta
Carnegie Mellon University

work with Nikhil Bansal (TU Eindhoven) and Guru Guruganesh (CMU)

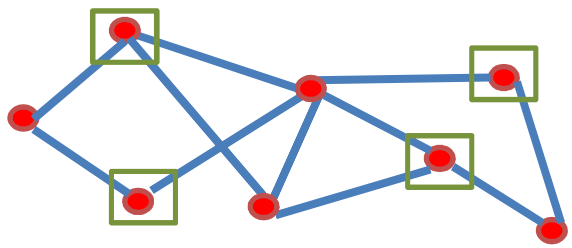
STOC 2015 / arXiv 1504.04767

independent set problem

Given graph G, approximate the largest independent set (size = α(G))

Notoriously hard: $\Omega(n^{0.999\ldots})$ [Hastad 96, …]

Best approx: $n/\log^3 n$ [Feige 04]



independent set problem

Given graph G, approximate the largest independent set (size = α(G))

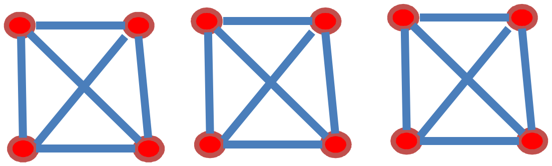
Notoriously hard: $\Omega(n^{0.999\ldots})$ [Hastad 96, …]

Best approx: $n/\log^3 n$ [Feige 04]

Our Focus: max. degree = d (avg. degree d suffices)

(d+1) approximation trivial [Greedy $\geq \frac{n}{d+1}$]

Tight: e.g., n/(d+1) disjoint copies of K_{d+1}



LP/SDP relaxations

IP: $\max \sum_i x_i$ s.t. $x_i + x_j \leq 1$ if $(\mathrm{i,j}) \in E$ $x_i \in \{0,1\}$

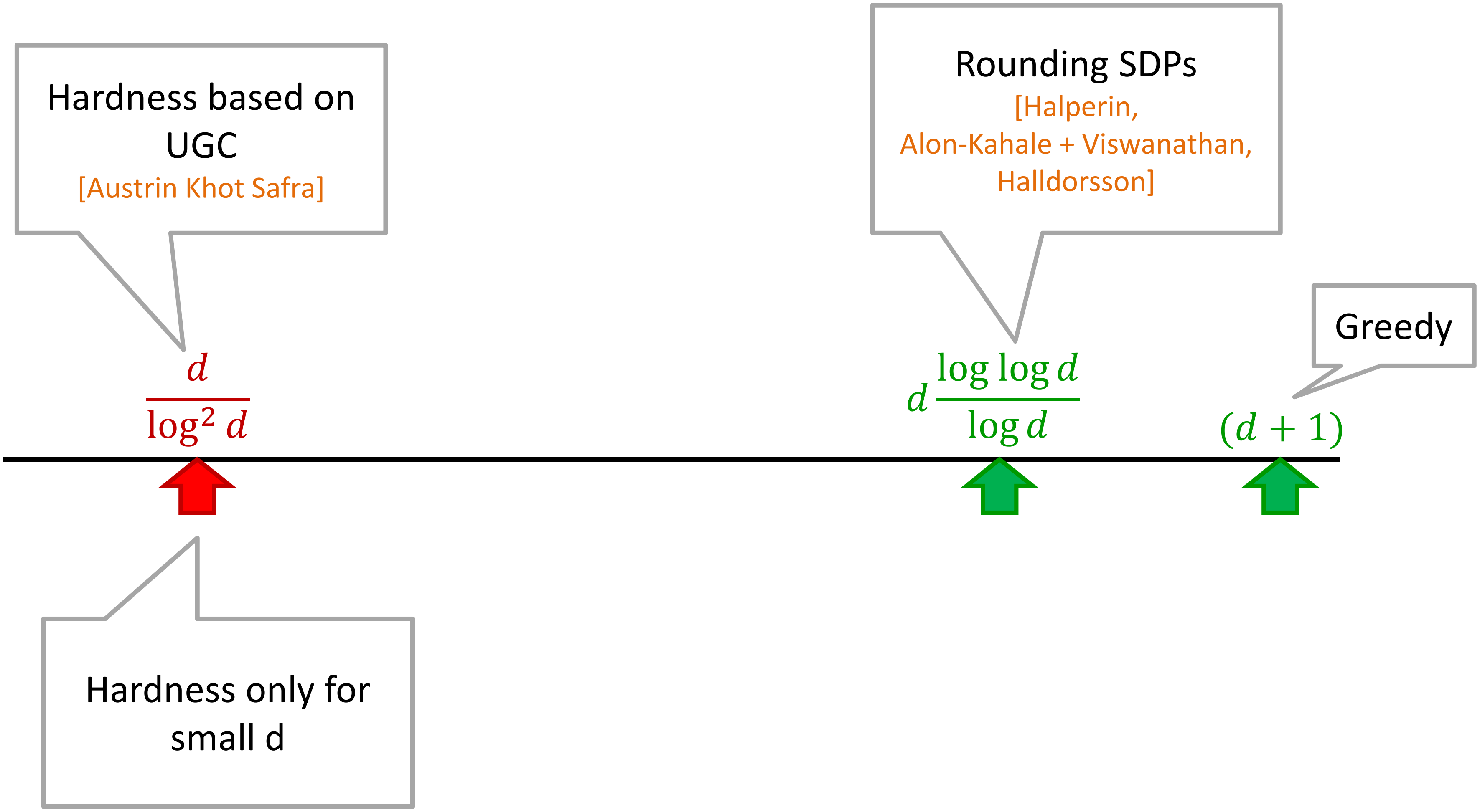
LP relaxation useless:

Ω(d) integrality gap (each $x_i = 1/2$)

Q. What do semidefinite programs give?

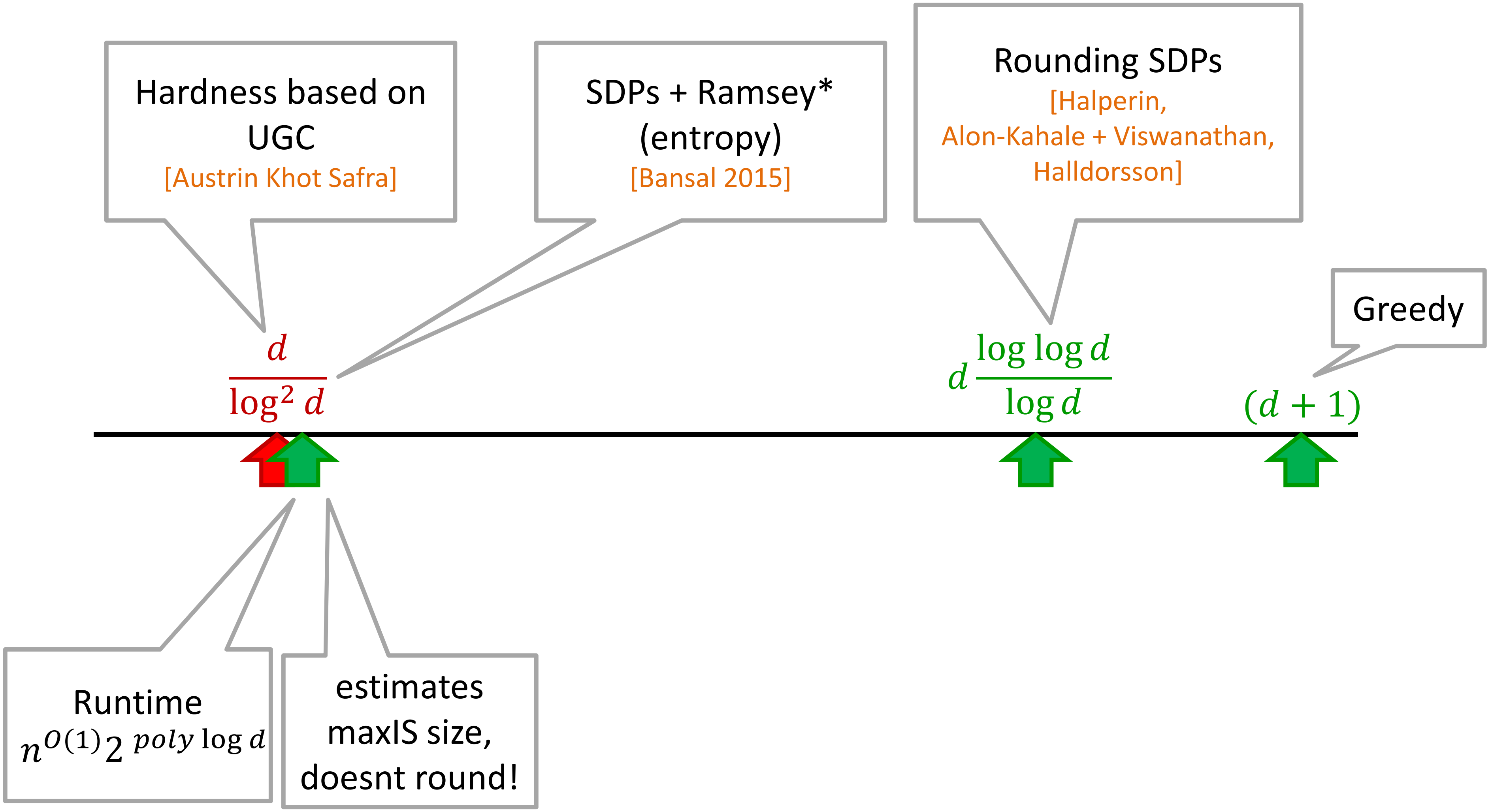
Q. How do "lift/project" hierarchies help?

progress timeline



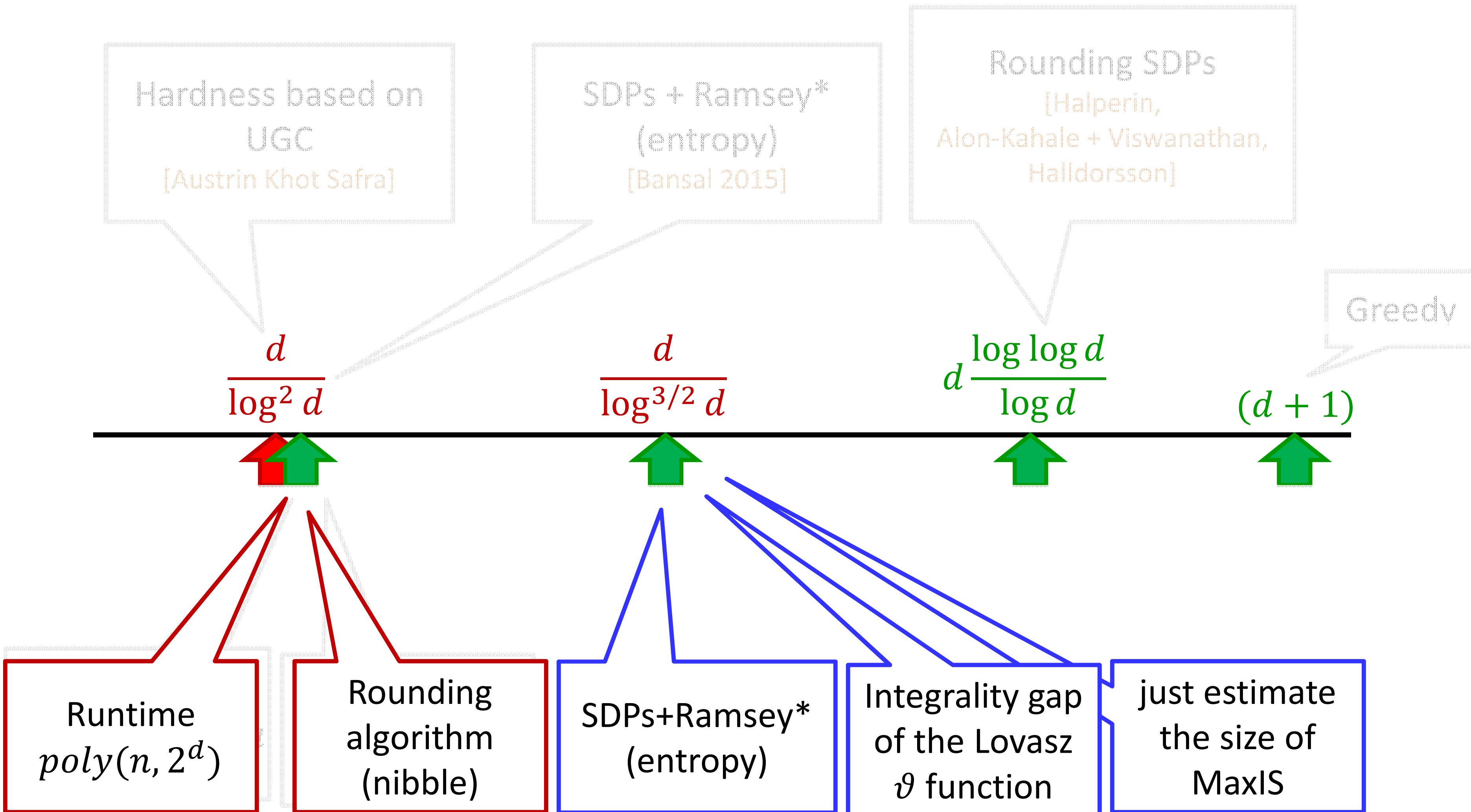
*Ignoring poly(log log d) factors

progress timeline



*Ignoring poly(log log d) factors

our results



*Ignoring poly(log log d) factors

our results (in words)

Theorem 1:

A randomized algorithm for MaxIS on degree-d graphs with (almost-tight) approximation factor

$$\tilde{O}\left(\frac{d}{\log^2 d}\right)$$

and run time $poly(n, 2^d)$.

Theorem 2:

The integrality gap of the Lovasz ϑ-function is at most

$$\tilde{O}\left(\frac{d}{\log^{3/2} d}\right)$$

$\Rightarrow$ estimate MaxIS size to within this factor in poly-time.

our chief weapon is…

+

why  +  ?

$\alpha(G)/n$

$1/k$ — 1

here, even greedy is (d/k) approx.

but need to certify we are here

if ϑ small, α small too
$\Rightarrow$ certificate!

need to do something smart

if ϑ large, $\bar{\chi}$ large $\Rightarrow$ clique covers of G are large

"$\Rightarrow$" small cliques in G

$\Rightarrow$ must have large ind.sets!

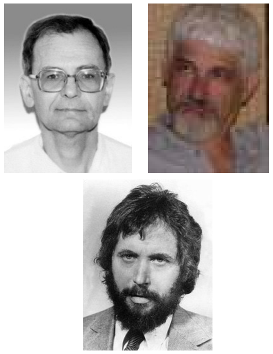
computable by SDPs

$$\alpha(G) \leq \vartheta(G) \leq \bar{\chi}(G)$$

Frank P. Ramsey (1903-30)

Ramsey results for sparse graphs

Thm: If K_3-free, $\alpha(G) \geq \left(\frac{n}{d}\right) \log d$ [Ajtai Komlos Szemeredi 80]



Celebrated result; pioneered "Rödl" nibble method

Several proofs known.

Tight.

Thm: If K_r-free, $\alpha(G) \geq \left(\frac{n}{d}\right) \log d \cdot \frac{1}{\mathrm{r} \log \log d}$ [Shearer 95]

Beautiful application of entropy method (non-algorithmic)

Remove the $\log \log d$?

What is the right dependence on r?

Ramsey results to remember…

Thm:  If K_r-free, $\alpha(G) \geq \tilde{\Omega}\left(\left(\frac{n}{d}\right) \cdot \frac{\log d}{r}\right)$

Conjecture: $\alpha(G) \geq \Omega\left(\left(\frac{n}{d}\right) \cdot \frac{\log d}{\log r}\right)$

Laci Lovasz (c. 1982)

the Lovasz ϑ-function

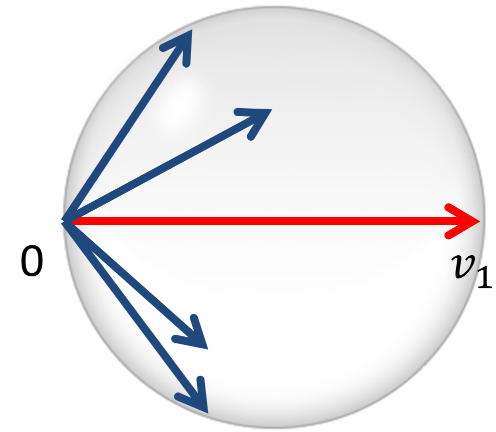
SDP: vector v_i for vertex i.

$$\vartheta(G) = \quad \text{Max } \Sigma_i \; v_i \cdot v_i$$

$$v_i \cdot v_j = 0 \qquad \text{if } (i,j) \in E$$

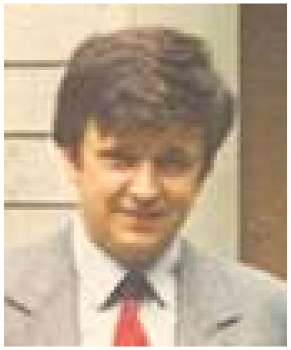
$$v_i \cdot v_1 = v_i \cdot v_i$$

$$v_1 \cdot v_1 = 1$$



Intended solution: $v_i = v_1$ if i chosen, 0 otherwise.

Define: $x_i \coloneqq v_i \cdot v_i$



Facts:

1. $\alpha(G) \leq \vartheta(G) \leq \bar{\chi}(G)$
2. If C is a clique in G, then $x(C) \leq 1$.

Halperin gives $\approx \frac{d}{\log d}$ approx

Use greedy on "low" vertices

Pr[vertex in IndSet] $\geq \frac{1}{d}$

$\Rightarrow \left(d\ \frac{8 \log \log d}{\log d}\right)$-approx on these

Round the "high" vertices using Karger-Motwani-Sudan rounding

Get a $\frac{d}{\log^2 d}$ -approx on these!

$x_i = \frac{8 \log \log d}{\log d}$

$x_i = 1$

Overall $\tilde{O}\left(\frac{d}{\log d}\right)$ approximation

Same factor as Ramsey!

Can we combine Halperin + Ramsey?

how to do better?

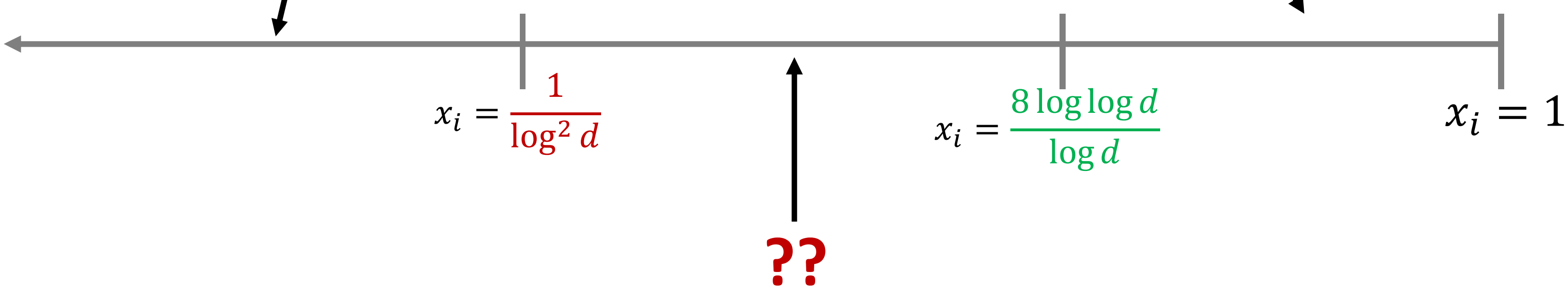
Use greedy on "low" vertices

Pr[vertex in IndSet] $\geq \frac{1}{d}$

$\left(\frac{d}{\log^2 d}\right)$-approx on these

Round the "high" vertices using Karger-Motwani-Sudan rounding

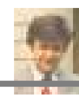
Get a $\frac{d}{\log^2 d}$ -approx on these!



how to do better?

From now on, imagine all vertices have x_i in this range

But: $x(C) \leq 1$ for all cliques!



$x_i = \frac{1}{\log^2 d}$ $x_i = \frac{8 \log\log d}{\log d}$ $x_i = 1$

So: largest clique has size $r \coloneqq \log^2 d$

Find a large ind.set in a K_r-free graph, where $r = \log^2 d$.

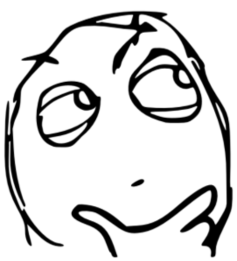
can we use Shearer?!?

Find a large ind.set in a K_r-free graph, where $r = \log^2 d$.

Thm:  If K_r-free, $\alpha(G) \geq \widetilde{\Omega}\left(\left(\frac{n}{d}\right) \cdot \frac{\log d}{r}\right)$

Conjecture: $\alpha(G) \geq \Omega\left(\left(\frac{n}{d}\right) \cdot \frac{\log d}{\log r}\right)$

For us, $r = \log^2 d$, so Shearer is worse than greedy n/d !!



If only conjecture were true…

Theorem 3: $\alpha(G) \geq \widetilde{\Omega}\left(\left(\frac{n}{d}\right) \cdot \sqrt{\frac{\log d}{\log r}}\right)$

putting it in the picture

$x_i = \frac{1}{\log^2 d}$ $\qquad$ $x_i = \frac{8 \log\log d}{\log d}$ $\qquad$ $x_i = 1$

greedy here

$\left(\frac{d}{\log^2 d}\right)$-approx

our results here

$\tilde{O}\left(\frac{d}{\log^{3/2} d}\right)$-approx

hyperplane rounding

$\frac{d}{\log^2 d}$ -approx

Theorem 2:
The integrality gap of the Lovasz ϑ-function is at most

$$\tilde{O}\left(\frac{d}{\log^{3/2} d}\right)$$

$\Rightarrow$ estimate MaxIS size to within this factor in poly-time.

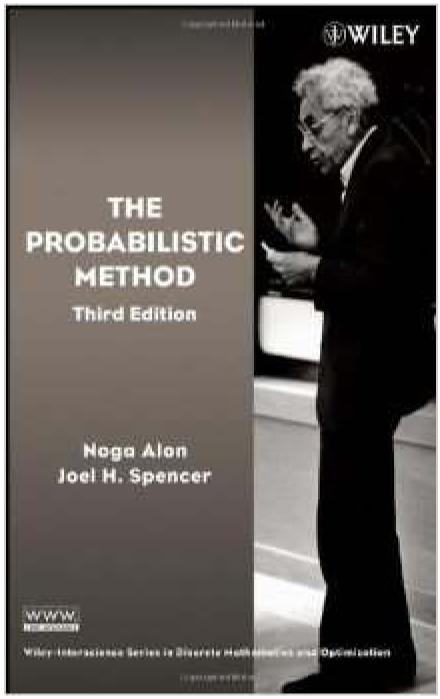
so now to prove…

Theorem 3: if G is K_r-free, then $\alpha(G) \geq \widetilde{\Omega}\left(\left(\frac{n}{d}\right) \cdot \sqrt{\frac{\log d}{\log r}}\right)$

Let's first prove

Theorem: if G is K_3-free, then $\alpha(G) \geq \Omega\left(\frac{n}{d} \log d\right)$

via Shearer's entropy approach



Assume d-regular graph G.

triangle-free graph

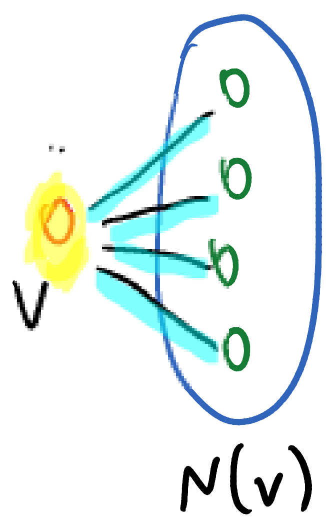
Theorem: if G is K_3-free, then $\alpha(G) \geq \Omega\left(\frac{n}{d}\log d\right)$

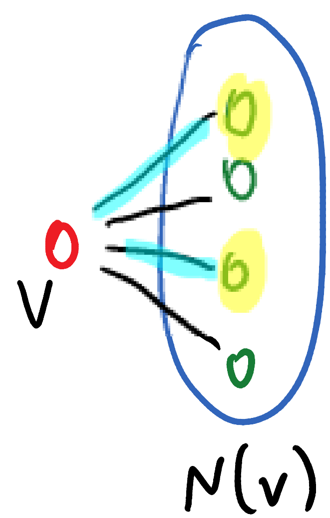
Pick a random independent set S from G. (show $\mathbb{E}[\#S]$ is large.)

$$\Phi_v \;\coloneqq\; 1(v \in S) + \frac{|N(v) \cap S|}{d}$$

so $\sum_v \Phi_v \leq 2\,|S|$

Suffices to show: for all vertices, $\mathbb{E}[\,\Phi_v\,] \geq \frac{\log d}{d}$





triangle-free graph

Theorem: if G is K_3-free, then $\alpha(G) \geq \Omega\left(\frac{n}{d}\log d\right)$

Pick a random independent set S from G.

$$\Phi_v := 1(v \in S) + \frac{|N(v) \cap S|}{d}$$

Suffices to show: for all vertices, $\mathbb{E}[\ \Phi_v\ ] \geq \frac{\log d}{d}$

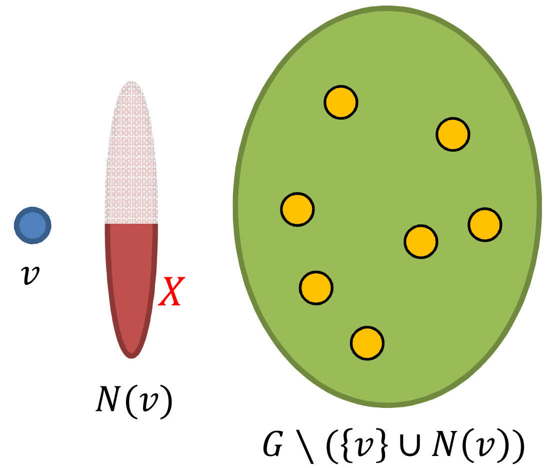
Condition on $S \setminus (\{v\} \cup N(v))$

Let X be the still-available vertices in $N(v)$

Want to pick a random ind.set from $X \cup \{v\}$

But X has no edges! $\Leftarrow$ G is triangle free!

1. Number of independent sets is $2^x + 1$
2. Average size of independent set in X is $x/2$



triangle-free graph

Theorem: if G is K_3-free, then $\alpha(G) \geq \Omega\left(\frac{n}{d}\log d\right)$

Pick a random independent set S from G.

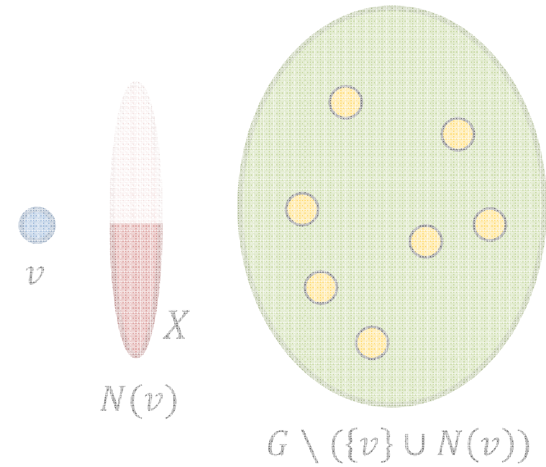
$$\Phi_v \coloneqq 1(v \in S) + \frac{|N(v) \cap S|}{d}$$

Suffices to show: for all vertices, $\mathbb{E}[\ \Phi_v\ ] \geq \frac{\log d}{d}$

1. Number of independent sets is $2^x + 1$
2. Average size of independent set in X is $x/2$

$$\mathbb{E}[\Phi_v] = \frac{1}{2^x+1} \cdot 1 + \frac{2^x}{2^x+1} \cdot \frac{\mathbb{E}[size\ of\ ind.set\ in\ X]}{d}$$

$$\approx \frac{1}{2^x} + \frac{x/2}{d} \qquad \geq \Omega\left(\frac{\log d}{d}\right)$$



extending to Shearer

Theorem: if G is K_3-free, then $\alpha(G) \geq \Omega\left(\frac{n}{d}\log d\right)$

Pick a random independent set S from G.

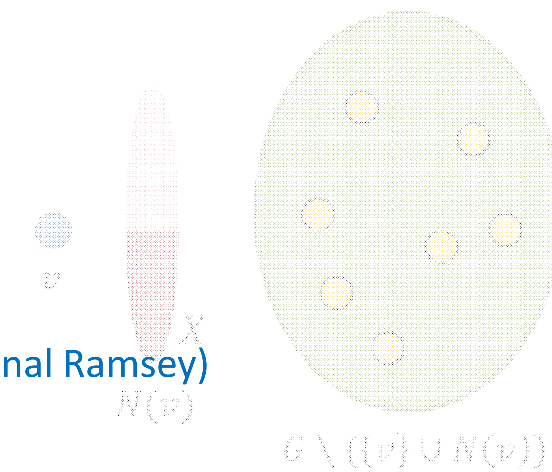
$$\Phi_v \;\coloneqq\; 1(v \in S) + \frac{|N(v) \cap S|}{d}$$

Suffices to show: for all vertices, $\mathbb{E}[\,\Phi_v\,] \geq \frac{\log d}{d}$

1. Number of independent sets is $2^x + 1$
2. Average size of independent set in X is $x/2$

For K_r-free:

1. Number of ind.sets is $2^{\alpha(X)}$ and $\alpha(X) \geq x^{1/r}$ (off-diagonal Ramsey)
2. If number of ind.sets is $2^{\varepsilon x}$ then average size is $\frac{\varepsilon\, x}{\log 1/\varepsilon}$



$\Rightarrow$ **Shearer's Theorem:** if G is K_r-free, then $\alpha(G) \geq \Omega\left(\frac{n}{d} \cdot \frac{\log d}{r \log\log d}\right)$

for our theorem: a puzzle

suppose K_3-free graph G, so $\alpha(G) \geq \frac{n}{d}\log d$ $=: A$

puzzle: how many ind.sets in G?

at least $2^{\alpha(G)} \approx 2^{A}$

at most $\binom{n}{\alpha(G)} \approx d^{A}$

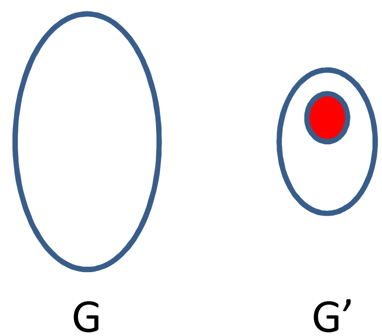
what's the truth?

---

thought experiment: sample each node w.p. $p := \frac{1}{\sqrt{d}}$

New graph G' has $n' = pn$, $d' = pd$, so $\frac{n}{d} \approx \frac{n'}{d'}$

So $\alpha(G') \geq \left(n' \frac{\log d'}{d'}\right) = \frac{n}{d} \cdot \frac{1}{2} \log d$ $=: A/2$



But any ind.set of size $A/2$ in G survives only w.p. $\left(\frac{1}{\sqrt{d}}\right)^{A/2}$

$\Rightarrow$ G has $\approx \sqrt{d}^{A/2}$ independent sets of size $A/2$!

to recap

Our question: Find a large ind.set in a K_r-free graph, where $r = \log^2 d$.

Thm:  If K_r-free, $\alpha(G) \geq \tilde{\Omega}\left(\left(\frac{n}{d}\right) \cdot \frac{\log d}{r}\right)$

Conjecture: $\alpha(G) \geq \Omega\left(\left(\frac{n}{d}\right) \cdot \frac{\log d}{\log r}\right)$

Theorem 3: $\alpha(G) \geq \tilde{\Omega}\left(\left(\frac{n}{d}\right) \cdot \sqrt{\frac{\log d}{\log r}}\right)$

Theorem 2:
The integrality gap of the Lovasz ϑ-function is at most

$$\tilde{O}\left(\frac{d}{\log^{3/2} d}\right)$$

$\Rightarrow$ estimate MaxIS size to within this factor in poly-time.

thanks!