Finding cliques in random graphs by adaptive probing

Based on joint works with

U. Feige, D. Gamarnik, J. Neeman, B. Schiffer, and P. Tetali

Miklos Z. Racz

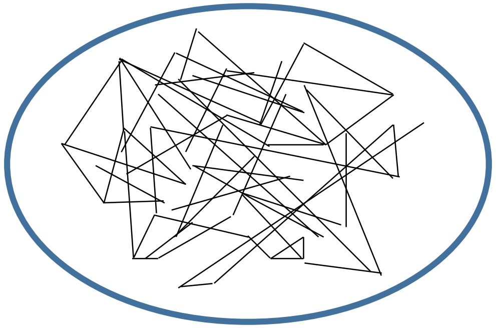
PRINCETON UNIVERSITY

The 19th International Conference on Random Structures and Algorithms

July 16, 2019

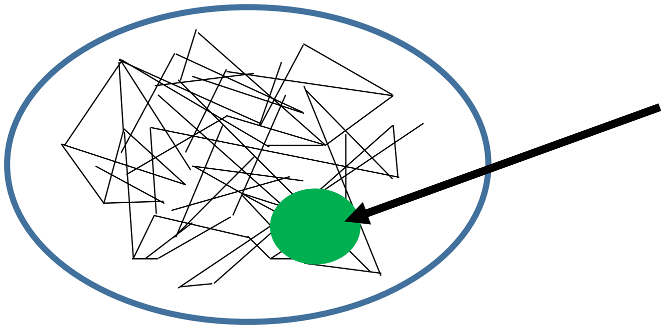
Finding cliques

Erdős-Rényi random graph $G(n, 1/2)$



Finding cliques

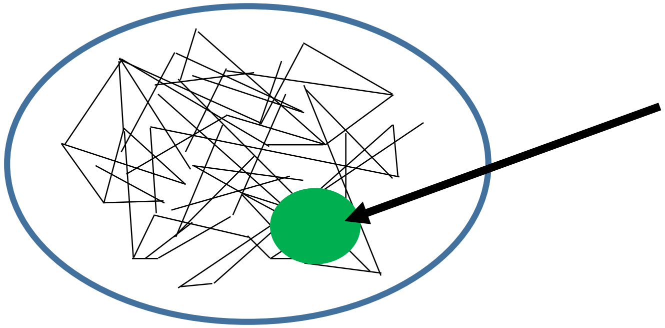
Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2\log n$

Finding cliques

Erdős-Rényi random graph $G(n, 1/2)$

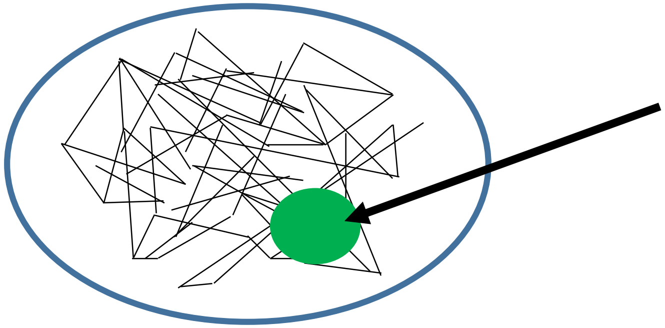


Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique efficiently

Finding cliques

Erdős-Rényi random graph $G(n, 1/2)$



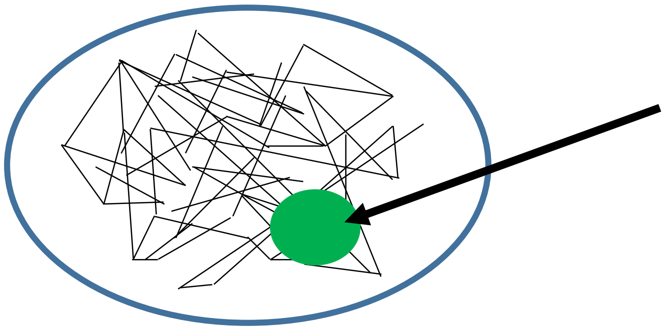
Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique efficiently

Greedy algorithm: finds clique of size $\log n$
Karp (1976): $\geq (1 + \varepsilon) \log n$??

Finding cliques

Erdős-Rényi random graph $G(n, 1/2)$

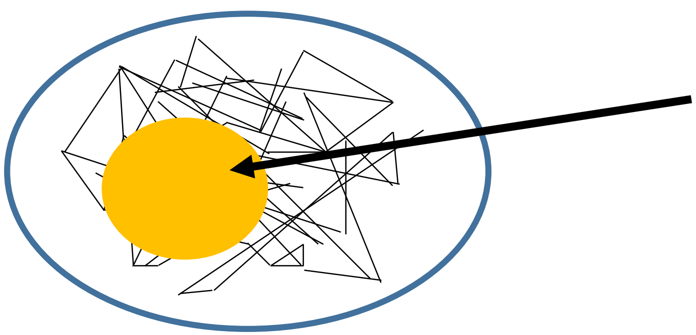


Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique efficiently

Greedy algorithm: finds clique of size $\log n$
Karp (1976): $\geq (1+\varepsilon) \log n$??

Planted clique model $G(n, 1/2, k)$

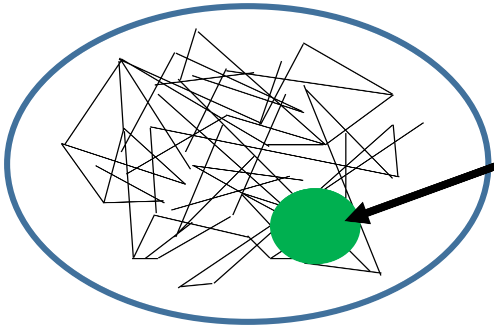


Planted clique
of size k

Challenge:
find planted clique efficiently

Finding cliques

Erdős-Rényi random graph $G(n, 1/2)$

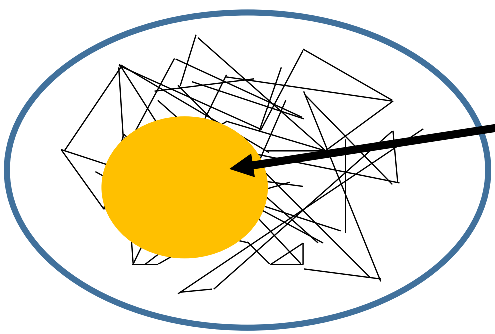


Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique efficiently

Greedy algorithm: finds clique of size $\log n$
Karp (1976): $\geq (1+\varepsilon) \log n$??

Planted clique model $G(n, 1/2, k)$



Planted clique
of size k

Challenge:
find planted clique efficiently

Information-theoretically possible when $k \geq (2+\varepsilon) \log n$
Efficient algorithm known only when $k = \Omega(n^{1/2})$
Conjectured information-computation gap

Adaptive probing

Goal: find max clique
Constraint: computational efficiency

Adaptive probing

~~Goal: find max clique~~
~~Constraint: computational efficiency~~

Goal: find max clique
Constraint: can only look at a small part of the graph

Adaptive probing

~~Goal: find max clique~~
~~Constraint: computational efficiency~~

Goal: find max clique
Constraint: can only look at a small part of the graph

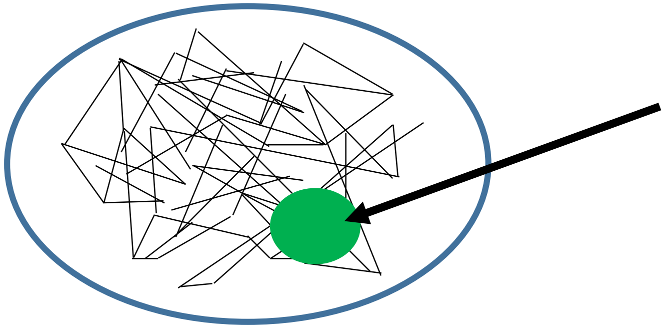
Probe model:
adaptively query pairs of vertices, learn if they are connected by an edge or not

1. $(i_1, j_1) \in E$?
2. $(i_2, j_2) \in E$?
3. $(i_3, j_3) \in E$?
4. Etc.

At most q queries in total.

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$

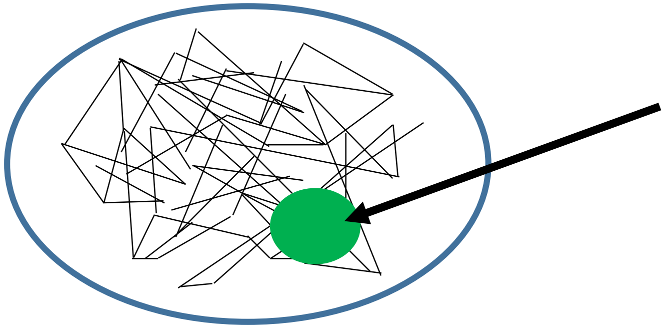


Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries

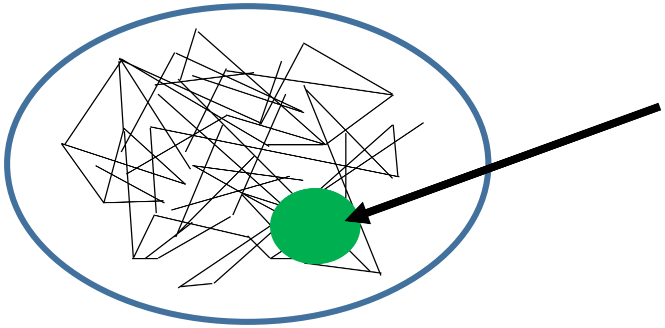
Related work
on finding structure in a random graph
using adaptive edge queries

- Ferber, Krivelevich, Sudakov, Vieira (RSA 2016): Hamilton cycles
- Ferber, Krivelevich, Sudakov, Vieira (RSA 2017): long paths
- Conlon, Fox, Grinshpun, He (2018): target graph H (e.g., small clique)

Main difference: dense vs. sparse random graphs

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$



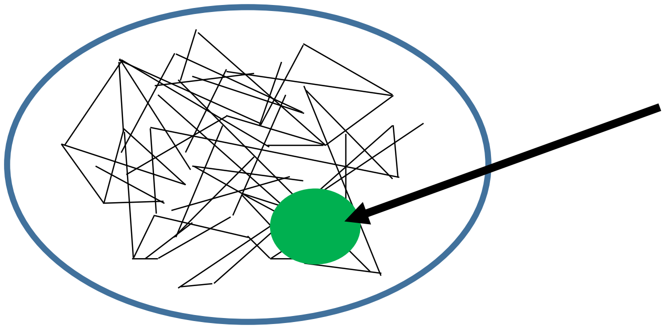
Largest clique
$\omega(G) \approx 2\log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries

Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2 \log n$

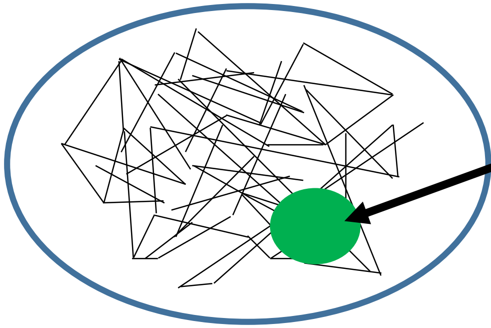
Challenge:
find max clique using
$\leq q$ adaptive edge queries

Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

$\alpha_{\star}(\delta) \coloneqq \sup \alpha$ s.t.
there exists algorithm making $\leq n^{\delta}$ adaptive queries
that finds a clique of size at least $\alpha \log n$ (w/prob. $\geq 1/2$)

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries

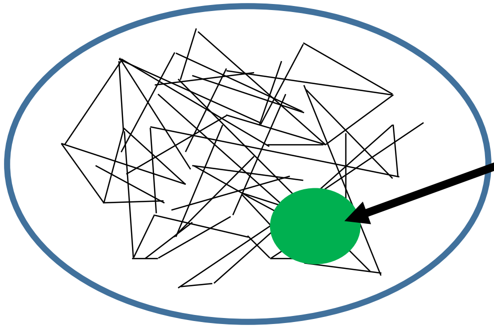
Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

$\alpha_{\star}(\delta) \coloneqq \sup \alpha$ s.t.
there exists algorithm making $\leq n^{\delta}$ adaptive queries
that finds a clique of size at least $\alpha \log n$ (w/prob. $\geq 1/2$)

$\alpha_{\star}(\delta) \leq 2$

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2\log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries

Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

$\alpha_{\star}(\delta) := \sup \alpha$ s.t.
there exists algorithm making $\leq n^{\delta}$ adaptive queries
that finds a clique of size at least $\alpha \log n$ (w/prob. $\geq 1/2$)

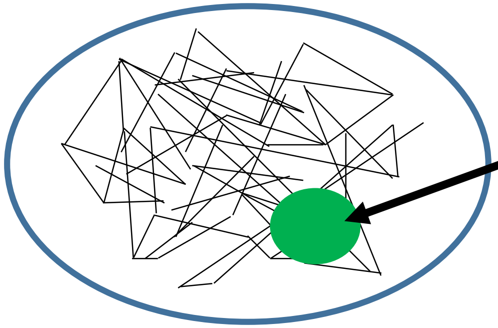
$\alpha_{\star}(\delta) \leq 2$

Algorithms?

1. Greedy: finds clique of size $\log n$

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2\log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries

Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

$\alpha_{\star}(\delta) \coloneqq \sup \alpha$ s.t.
there exists algorithm making $\leq n^{\delta}$ adaptive queries
that finds a clique of size at least $\alpha \log n$ (w/prob. $\geq 1/2$)

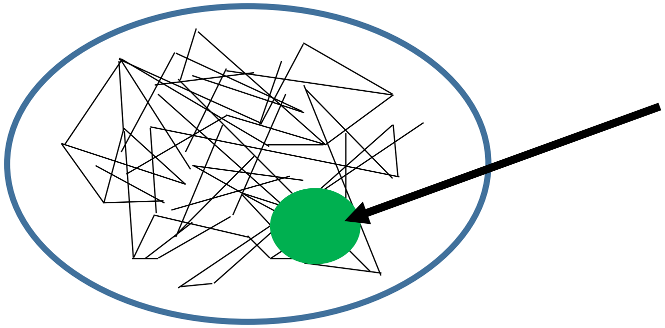
$\alpha_{\star}(\delta) \leq 2$

Algorithms?

1. Greedy: finds clique of size $\log n$
2. Exhaustive search: $2\log\sqrt{q} = \delta \log n$

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries

Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

$\alpha_{\star}(\delta) := \sup \alpha$ s.t.
there exists algorithm making $\leq n^{\delta}$ adaptive queries
that finds a clique of size at least $\alpha \log n$ (w/prob. $\geq 1/2$)

$\alpha_{\star}(\delta) \leq 2$

Algorithms?

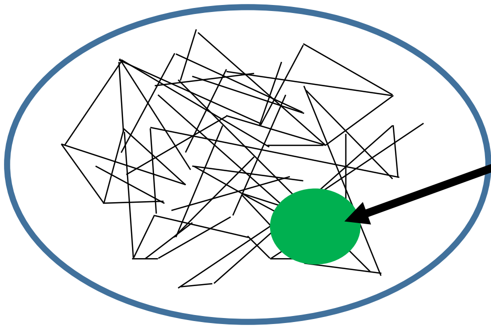
1. Greedy: finds clique of size $\log n$
2. Exhaustive search: $2 \log \sqrt{q} = \delta \log n$
3. Combined: greedy until $\sqrt{q}$ nodes remain,

then switch to exhaustive search:

$$\log(n/\sqrt{q}) + 2 \log \sqrt{q} = \log n + \frac{1}{2} \log q = (1 + \delta/2) \log n$$

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries

Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

$\alpha_{\star}(\delta) \coloneqq \sup \alpha$ s.t.
there exists algorithm making $\leq n^{\delta}$ adaptive queries
that finds a clique of size at least $\alpha \log n$ (w/prob. $\geq 1/2$)

$\alpha_{\star}(\delta) \leq 2$

Algorithms?

1. Greedy: finds clique of size $\log n$
2. Exhaustive search: $2 \log \sqrt{q} = \delta \log n$
3. Combined: greedy until $\sqrt{q}$ nodes remain,

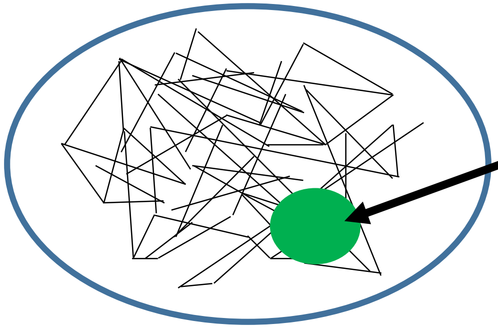
then switch to exhaustive search:

$$\log(n/\sqrt{q}) + 2 \log \sqrt{q} = \log n + \frac{1}{2} \log q = (1 + \delta/2) \log n$$

$$1 + \frac{\delta}{2} \leq \alpha_{\star}(\delta) \leq 2$$

Finding cliques by adaptive probing

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries

Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

$\alpha_{\star}(\delta) := \sup \alpha$ s.t.
there exists algorithm making $\leq n^{\delta}$ adaptive queries
that finds a clique of size at least $\alpha \log n$ (w/prob. $\geq 1/2$)

$\alpha_{\star}(\delta) \leq 2$

Algorithms?

1. Greedy: finds clique of size $\log n$
2. Exhaustive search: $2 \log \sqrt{q} = \delta \log n$
3. Combined: greedy until $\sqrt{q}$ nodes remain,
then switch to exhaustive search:

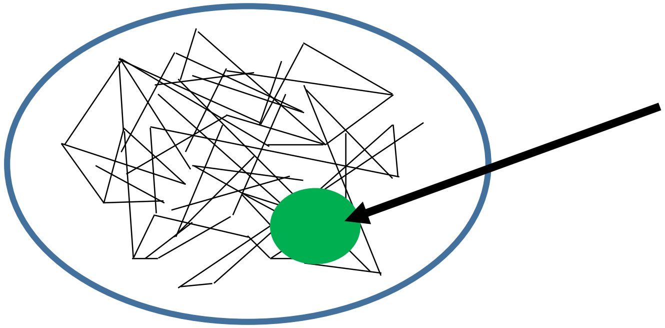
$$\log(n/\sqrt{q}) + 2 \log \sqrt{q} = \log n + \frac{1}{2} \log q = (1 + \delta/2) \log n$$

$$1 + \frac{\delta}{2} \leq \alpha_{\star}(\delta) \leq 2$$

Open problem: find $\alpha_{\star}(\delta)$

Adaptive probing w/ few rounds

Erdős-Rényi random graph $G(n, 1/2)$



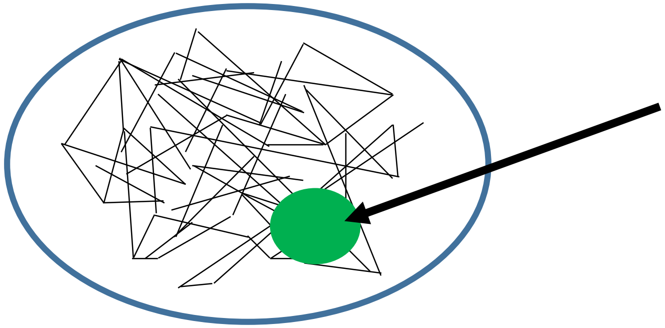
Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries
in at most ℓ rounds

Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

Adaptive probing w/ few rounds

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2 \log n$

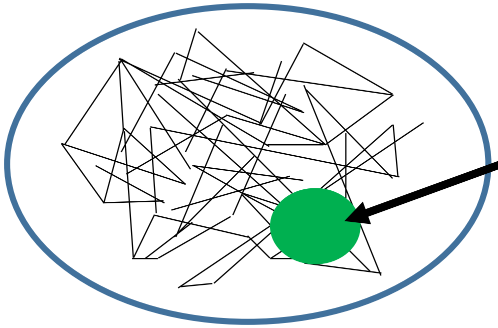
Challenge:
find max clique using
$\leq q$ adaptive edge queries
in at most ℓ rounds

Parametrize $q = n^\delta$, wlog $1 \leq \delta < 2$

$\alpha_\star(\delta, \ell) \coloneqq \sup \alpha$ s.t.
there exists algorithm making $\leq n^\delta$ adaptive queries
in $\leq \ell$ rounds that finds a clique of size at least $\alpha \log n$

Adaptive probing w/ few rounds

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2\log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries
in at most ℓ rounds

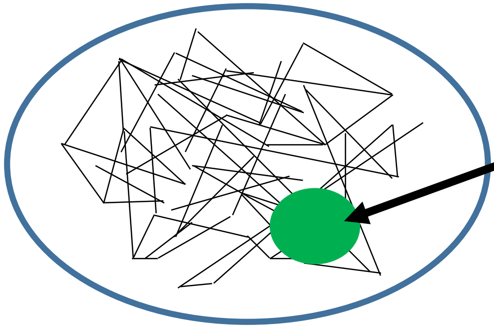
Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

$\alpha_{\star}(\delta, \ell) \coloneqq \sup \alpha$ s.t.
there exists algorithm making $\leq n^{\delta}$ adaptive queries
in $\leq \ell$ rounds that finds a clique of size at least $\alpha \log n$

$\alpha_{\star}(\delta, \ell) \leq \alpha_{\star}(\delta) \leq 2$

Adaptive probing w/ few rounds

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique using
$\leq q$ adaptive edge queries
in at most ℓ rounds

Parametrize $q = n^{\delta}$, wlog $1 \leq \delta < 2$

$\alpha_\star(\delta, \ell) := \sup \alpha$ s.t.
there exists algorithm making $\leq n^{\delta}$ adaptive queries
in $\leq \ell$ rounds that finds a clique of size at least $\alpha \log n$

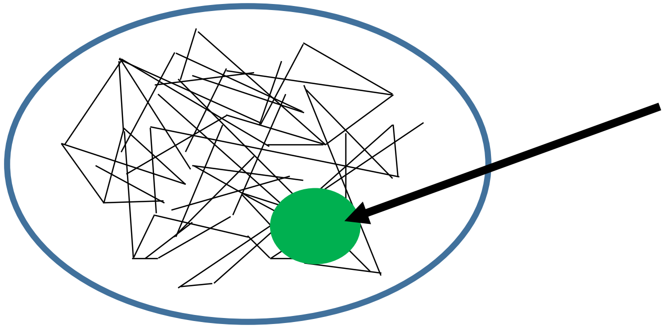
$\alpha_\star(\delta, \ell) \leq \alpha_\star(\delta) \leq 2$

Theorem (Feige, Gamarnik, Neeman, R., Tetali 2018)

For every $\delta < 2$ and constant ℓ we have that $\alpha_\star(\delta, \ell) < 2$.

Specific bounds ($\delta = 1$)

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2\log n$

Challenge:
find max clique using
$O(n)$ adaptive edge queries
in at most ℓ rounds

Theorem (Feige, Gamarnik, Neeman, R., Tetali 2018)

One round: $\alpha_\star(1,1) = 1.$

Two rounds: $4/3 \le \alpha_\star(1,2) \le 2^{2/3} < 1.588.$

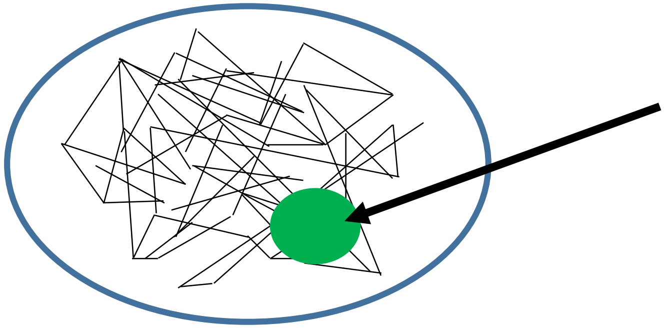
Three rounds: $3/2 \le \alpha_\star(1,3) \le 2^{6/7} < 1.812.$

Four rounds: $\alpha_\star(1,4) \le 2^{14/15} < 1.910.$

ℓ rounds: $\alpha_\star(1,\ell) \le 2^{1-1/(2^\ell-1)}.$

Specific bounds ($\delta = 1$)

Erdős-Rényi random graph $G(n, 1/2)$



Largest clique
$\omega(G) \approx 2 \log n$

Challenge:
find max clique using
$O(n)$ adaptive edge queries
in at most ℓ rounds

Theorem (Feige, Gamarnik, Neeman, R., Tetali 2018)

One round: $\alpha_\star(1,1) = 1.$

Two rounds: $4/3 \leq \alpha_\star(1,2) \leq 2^{2/3} < 1.588.$

Three rounds: $3/2 \leq \alpha_\star(1,3) \leq 2^{6/7} < 1.812.$

Four rounds: $\alpha_\star(1,4) \leq 2^{14/15} < 1.910.$

ℓ rounds: $\alpha_\star(1,\ell) \leq 2^{1-1/(2^\ell - 1)}.$

Q: Is 3 rounds more powerful than 2 rounds?

Algorithms

Erdős-Rényi random graph $G(n, 1/2)$

Challenge:
find max clique using
$O(n)$ adaptive edge queries
in at most ℓ rounds

<u>1 round:</u> pick $\sqrt{n}$ vertices, probe all pairs

Algorithms

Erdős-Rényi random graph $G(n, 1/2)$

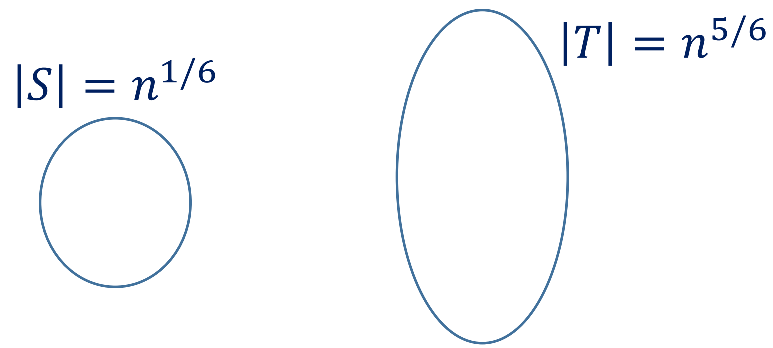
> **Challenge:**
> find max clique using
> $O(n)$ adaptive edge queries
> in at most ℓ rounds

1 round: pick $\sqrt{n}$ vertices, probe all pairs

2 rounds:

Round #1:

- Probe all pairs within S.
- Probe all pairs between S and T.



Algorithms

Erdős-Rényi random graph $G(n, 1/2)$

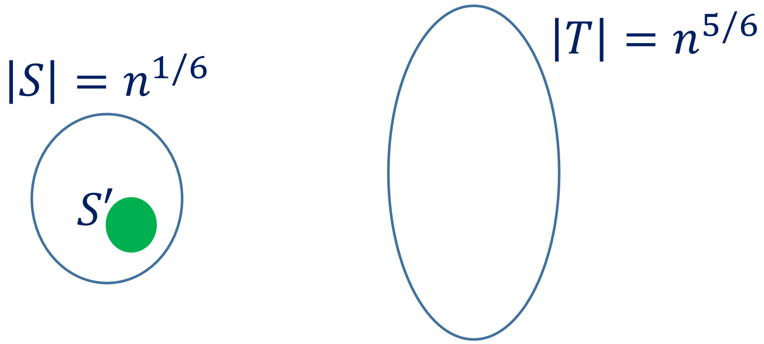
Challenge:
find max clique using
$O(n)$ adaptive edge queries
in at most ℓ rounds

1 round: pick $\sqrt{n}$ vertices, probe all pairs

2 rounds:

Round #1:

- Probe all pairs within S.
- Probe all pairs between S and T.
- $S' :=$ largest clique in S, $|S'| \approx \frac{1}{3}\log n$.

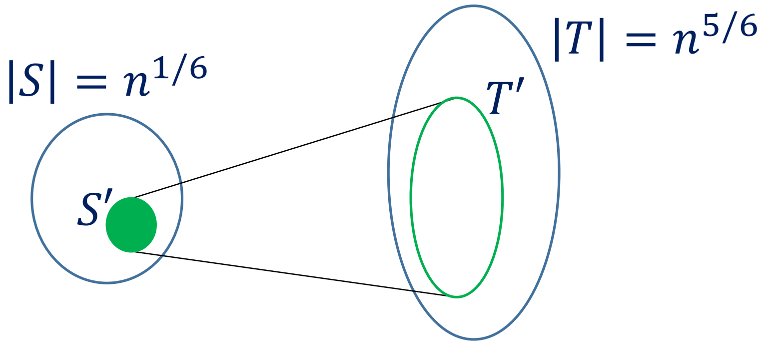


Algorithms

Erdős-Rényi random graph $G(n, 1/2)$

Challenge:
find max clique using
$O(n)$ adaptive edge queries
in at most ℓ rounds

1 round: pick $\sqrt{n}$ vertices, probe all pairs



2 rounds:

Round #1:

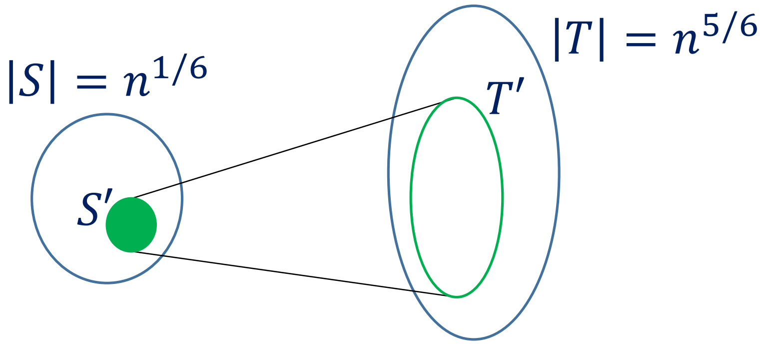
- Probe all pairs within S.
- Probe all pairs between S and T.
- $S' :=$ largest clique in S, $|S'| \approx \frac{1}{3} \log n$.
- $T' :=$ vertices in T which are connected to every vertex in S'.
- $|T'| \approx \frac{n^{5/6}}{2^{(1/3)\log n}} = n^{1/2}$.

Algorithms

Erdős-Rényi random graph $G(n, 1/2)$

Challenge:
find max clique using
$O(n)$ adaptive edge queries
in at most ℓ rounds

1 round: pick $\sqrt{n}$ vertices, probe all pairs



2 rounds:

Round #1:

- Probe all pairs within S.
- Probe all pairs between S and T.
- $S' :=$ largest clique in S, $|S'| \approx \frac{1}{3}\log n$.
- $T' :=$ vertices in T which are connected to every vertex in S'.
- $|T'| \approx \frac{n^{5/6}}{2^{(1/3)\log n}} = n^{1/2}$.

Round #2:

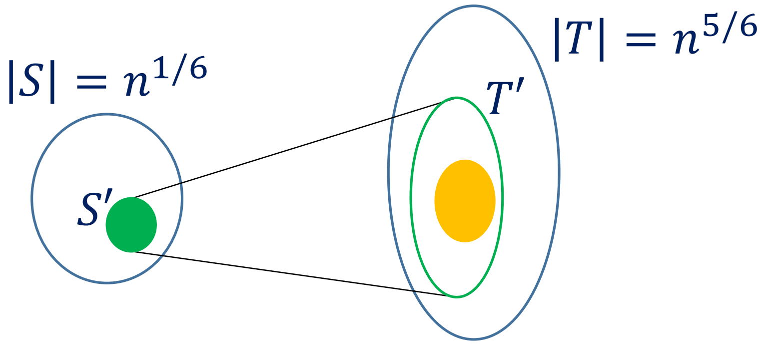
- Probe all pairs within T'.

Algorithms

Erdős-Rényi random graph $G(n, 1/2)$

Challenge:
find max clique using
$O(n)$ adaptive edge queries
in at most ℓ rounds

<u>1 round:</u> pick $\sqrt{n}$ vertices, probe all pairs



<u>2 rounds:</u>

Round #1:

- Probe all pairs within S.
- Probe all pairs between S and T.
- $S' :=$ largest clique in S, $|S'| \approx \frac{1}{3}\log n$.
- $T' :=$ vertices in T which are connected to every vertex in S'.
- $|T'| \approx \frac{n^{5/6}}{2^{(1/3)\log n}} = n^{1/2}$.

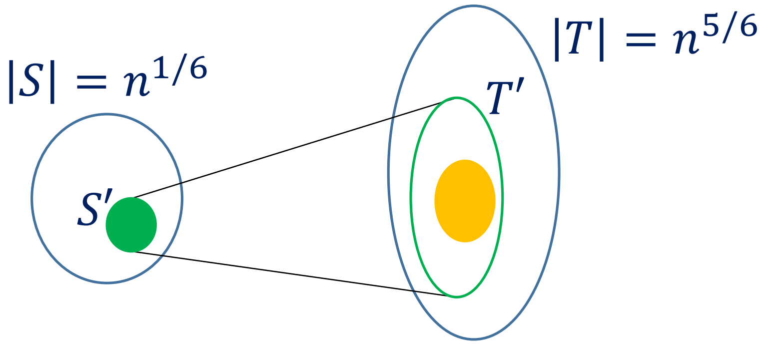
Round #2:

- Probe all pairs within T'.
- Find clique of size $\approx 2\log\sqrt{n} = \log n$.

Algorithms

Erdős-Rényi random graph $G(n, 1/2)$

Challenge:
find max clique using
$O(n)$ adaptive edge queries
in at most ℓ rounds



1 round: pick $\sqrt{n}$ vertices, probe all pairs

2 rounds:

Round #1:

- Probe all pairs within S.
- Probe all pairs between S and T.
- $S' :=$ largest clique in S, $|S'| \approx \frac{1}{3}\log n$.
- $T' :=$ vertices in T which are connected to every vertex in S'.
- $|T'| \approx \frac{n^{5/6}}{2^{(1/3)\log n}} = n^{1/2}$.

Round #2:

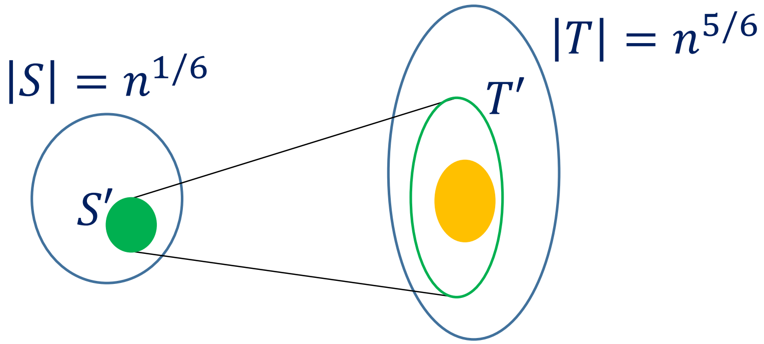
- Probe all pairs within T'.
- Find clique of size $\approx 2\log\sqrt{n} = \log n$.
- Altogether: clique of size $\frac{4}{3}\log n$.

Algorithms

Erdős-Rényi random graph $G(n, 1/2)$

Challenge:
find max clique using
$O(n)$ adaptive edge queries
in at most ℓ rounds

<u>1 round:</u> pick $\sqrt{n}$ vertices, probe all pairs



<u>2 rounds:</u>

Round #1:

- Probe all pairs within S.
- Probe all pairs between S and T.
- $S' :=$ largest clique in S, $|S'| \approx \frac{1}{3}\log n$.
- $T' :=$ vertices in T which are connected to every vertex in S'.
- $|T'| \approx \frac{n^{5/6}}{2^{(1/3)\log n}} = n^{1/2}$.

Round #2:

- Probe all pairs within T'.
- Find clique of size $\approx 2\log\sqrt{n} = \log n$.
- Altogether: clique of size $\frac{4}{3}\log n$.

<u>3 rounds:</u> similar.
Exercise!

Ideas about $\alpha_\star(\delta, \ell) < 2$

Theorem (Feige, Gamarnik, Neeman, R., Tetali 2018)

For every $\delta < 2$ and constant ℓ we have that $\alpha_\star(\delta, \ell) < 2$.

In short: first moment method + some extremal graph theory

Ideas about $\alpha_\star(\delta,\ell) < 2$

Theorem (Feige, Gamarnik, Neeman, R., Tetali 2018)

For every $\delta < 2$ and constant ℓ we have that $\alpha_\star(\delta,\ell) < 2$.

In short: first moment method + some extremal graph theory

In more detail:

- Algorithm takes ℓ rounds, $O(n)$ queries in each round
- $k \coloneqq \alpha \log n$; K a set of vertices of size k
- Fix $\beta_1, \ldots, \beta_\ell \geq 0$ s.t. $\sum_{i=1}^{\ell} \beta_i = 1$; will optimize over later

Def: Round i is **significant** if the # of probes to K

- in rounds 1 to $i-1$ is $\leq \sum_{j=1}^{i-1} \beta_j \binom{k}{2}$, and
- in rounds 1 to i is $\geq \sum_{j=1}^{i} \beta_j \binom{k}{2}$.

Claim: there is a significant round.
(Proof: induction.)

- Such a K called an i-eligible set.
- Can determine after round $i-1$.
- After round $i-1$,

$$P(K \text{ is a clique}) \leq 2^{-\sum_{j=i}^{\ell} \beta_j \binom{k}{2}}$$

- Union bound over all such K.
- To bound their number: extremal graph theory, next slide

An extremal problem

Def: $N_{n,m,k,\beta} :=$ max # sets of size k that can be β-covered in an n vertex graph with m edges

edges in induced subgraph is $\geq \beta$ fraction of total possible

Theorem (Feige, Gamarnik, Neeman, R., Tetali 2018)

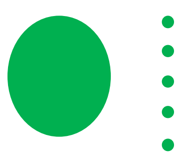
When $\beta \in \left[0, \frac{16}{25}\right]$:

$$N_{n,m,k,\beta} \leq m^{\left(1-\sqrt{1-\beta}\right)k+1} \cdot n^{\left(2\sqrt{1-\beta}-1\right)k+2}.$$

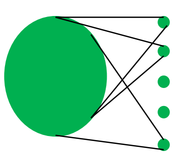
When $\beta \in \left[\frac{16}{25}, 1\right]$:

$$N_{n,m,k,\beta} \leq m^{\left(\sqrt{\beta}/2\right)k+1} \cdot n^{\left(1-\sqrt{\beta}\right)k+2}.$$

Extremal graphs:



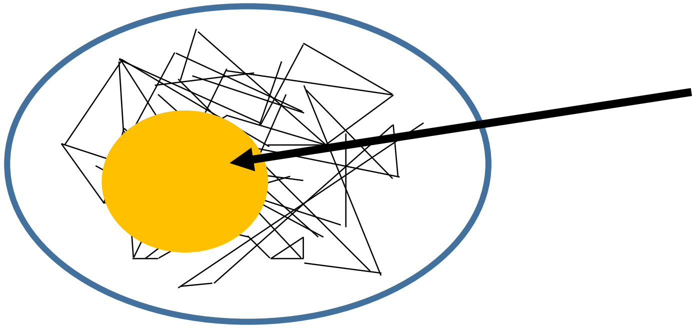
Clique + isolated vertices



``Complete split graph''

Finding a planted clique via probing

Planted clique model $G(n, 1/2, k)$

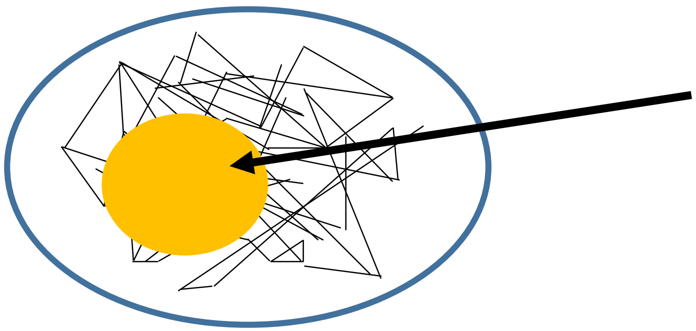


Planted clique
of size k

Challenge:
find planted clique using
$\leq q$ adaptive edge queries

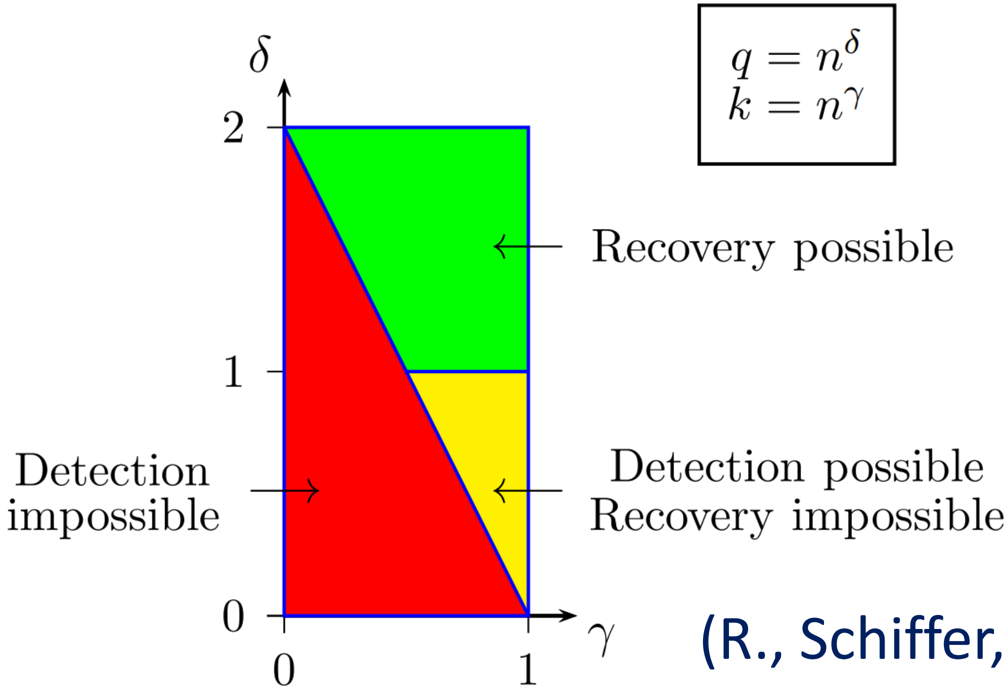
Finding a planted clique via probing

Planted clique model $G(n, 1/2, k)$



Planted clique of size k

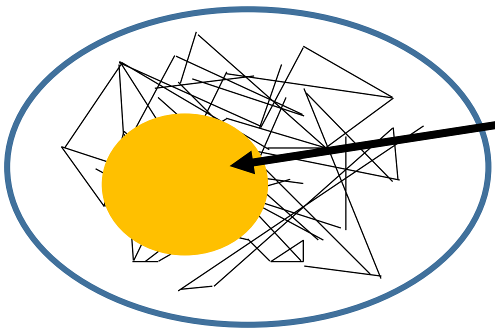
Challenge: find planted clique using $\leq q$ adaptive edge queries



(R., Schiffer, 2019)

Finding a planted clique via probing

Planted clique model $G(n, 1/2, k)$

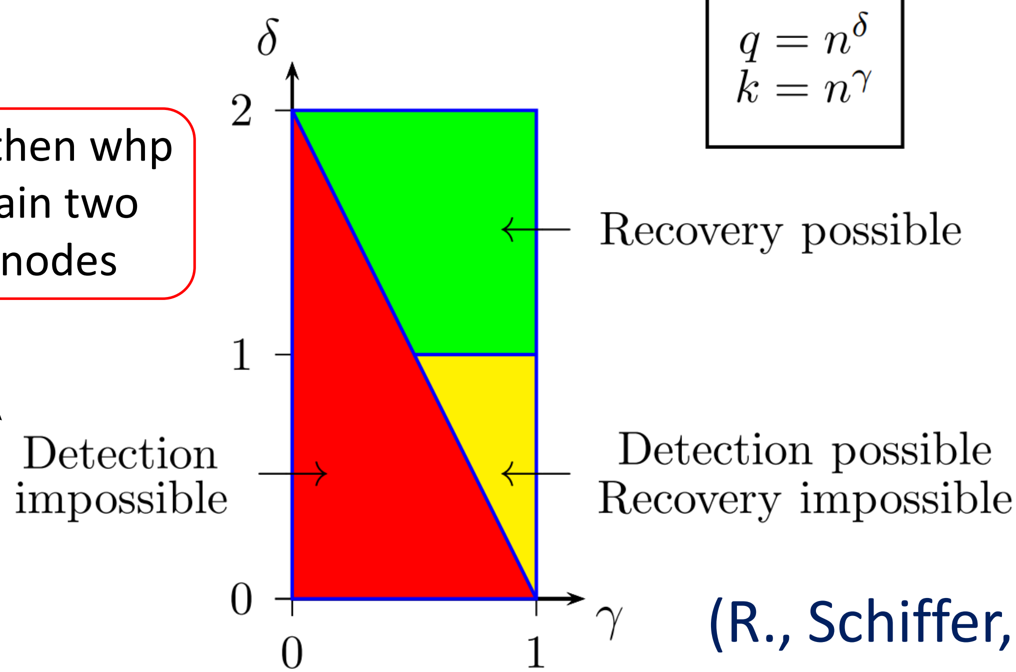


Planted clique of size k

Challenge:
find planted clique using $\leq q$ adaptive edge queries

If $q = o(n^2/k^2)$ then whp no queries contain two planted clique nodes

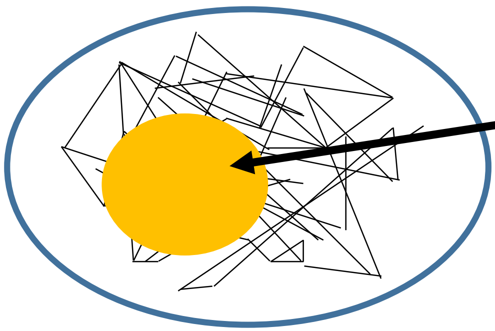
$q = n^\delta$
$k = n^\gamma$



(R., Schiffer, 2019)

Finding a planted clique via probing

Planted clique model $G(n, 1/2, k)$



Planted clique of size k

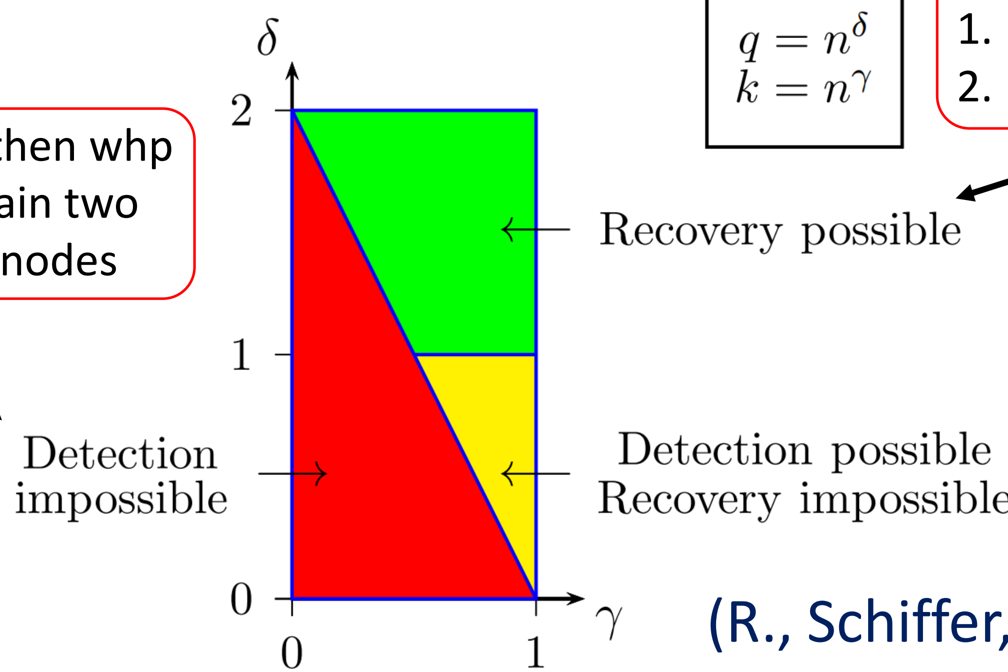
Challenge: find planted clique using $\leq q$ adaptive edge queries

$$q = n^{\delta}$$
$$k = n^{\gamma}$$

Algorithm:

1. Sample, find large clique
2. Extend to planted clique

If $q = o(n^2/k^2)$ then whp no queries contain two planted clique nodes



(R., Schiffer, 2019)

Summary

- Adaptive edge query model: constraint worth exploring
- **Main result:** cannot find the largest clique w/ constant rounds
- **Open problem:** compute $\alpha_\star(\delta)$. Is $\alpha_\star(\delta) < 2$?
- Is three rounds more powerful than two rounds?

Summary

- Adaptive edge query model: constraint worth exploring
- **Main result:** cannot find the largest clique w/ constant rounds
- **Open problem:** compute $\alpha_\star(\delta)$. Is $\alpha_\star(\delta) < 2$?
- Is three rounds more powerful than two rounds?

Thank you!