A loglog step towards the Erdős-Hajnal conjecture

Paul Seymour (Princeton)
Joint work with Matija Bucić, Tung Nguyen and Alex Scott.

Induced subgraph: subgraph obtained by deleting vertices.
H-free: no induced subgraph isomorphic to H.
$\alpha(G)$ = size of the largest stable set in G.
$\omega(G)$ = size of the largest clique in G.

Induced subgraph: subgraph obtained by deleting vertices.
H-free: no induced subgraph isomorphic to *H*.
$\alpha(G)$ = size of the largest stable set in *G*.
$\omega(G)$ = size of the largest clique in *G*.

- There are *n*-vertex graphs with $\max(\alpha(G), \omega(G)) \leq O(\log(n))$.

Induced subgraph: subgraph obtained by deleting vertices.
H-free: no induced subgraph isomorphic to *H*.
$\alpha(G)$ = size of the largest stable set in *G*.
$\omega(G)$ = size of the largest clique in *G*.

- There are *n*-vertex graphs with $\max(\alpha(G), \omega(G)) \leq O(\log(n))$.
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/2}$ for every P_4-free graph *G*.

Induced subgraph: subgraph obtained by deleting vertices.
H-free: no induced subgraph isomorphic to H.
$\alpha(G)$ = size of the largest stable set in G.
$\omega(G)$ = size of the largest clique in G.

- There are n-vertex graphs with $\max(\alpha(G), \omega(G)) \leq O(\log(n))$.
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/2}$ for every P_4-free graph G.
- $\alpha(G) \geq \frac{1}{2}|G|^{1/(t-1)}$ for every K_t-free graph G.

Induced subgraph: subgraph obtained by deleting vertices.
H-free: no induced subgraph isomorphic to H.
$\alpha(G)$ = size of the largest stable set in G.
$\omega(G)$ = size of the largest clique in G.

- There are n-vertex graphs with $\max(\alpha(G), \omega(G)) \leq O(\log(n))$.
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/2}$ for every P_4-free graph G.
- $\alpha(G) \geq \frac{1}{2}|G|^{1/(t-1)}$ for every K_t-free graph G.
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/4}$ for every bull-free graph G. (Chudnovsky, Safra, 2008)

Induced subgraph: subgraph obtained by deleting vertices.
H-free: no induced subgraph isomorphic to H.
$\alpha(G)$ = size of the largest stable set in G.
$\omega(G)$ = size of the largest clique in G.

- There are n-vertex graphs with $\max(\alpha(G), \omega(G)) \leq O(\log(n))$.
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/2}$ for every P_4-free graph G.
- $\alpha(G) \geq \frac{1}{2}|G|^{1/(t-1)}$ for every K_t-free graph G.
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/4}$ for every bull-free graph G. (Chudnovsky, Safra, 2008)
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/3}$ for every C_4-free graph G.

Induced subgraph: subgraph obtained by deleting vertices.
H-free: no induced subgraph isomorphic to H.
$\alpha(G)$ = size of the largest stable set in G.
$\omega(G)$ = size of the largest clique in G.

- There are n-vertex graphs with $\max(\alpha(G), \omega(G)) \leq O(\log(n))$.
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/2}$ for every P_4-free graph G.
- $\alpha(G) \geq \frac{1}{2}|G|^{1/(t-1)}$ for every K_t-free graph G.
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/4}$ for every bull-free graph G. (Chudnovsky, Safra, 2008)
- $\max(\alpha(G), \omega(G)) \geq |G|^{1/3}$ for every C_4-free graph G.

Conjecture (Erdős, Hajnal, 1977)

For every graph H, there exists $c > 0$ such that every H-free graph G has a clique or stable set of size at least $|G|^c$.

H has the EH-property if there exists $c > 0$ such that $\max(\alpha(G), \omega(G)) \geq |G|^c$ for every H-free graph G.

H has the EH-property if there exists $c > 0$ such that $\max(\alpha(G), \omega(G)) \geq |G|^c$ for every H-free graph G.

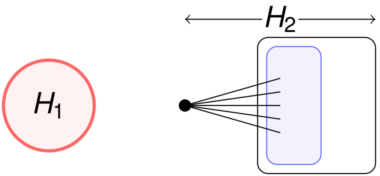
Theorem (Alon, Pach, Solymosi, 2001)

If H_1, H_2 have the EH-property, and H is obtained by substituting H_1 for a vertex of H_2, then H has the EH-property.

H has the EH-property if there exists $c > 0$ such that $\max(\alpha(G), \omega(G)) \geq |G|^c$ for every H-free graph G.

Theorem (Alon, Pach, Solymosi, 2001)

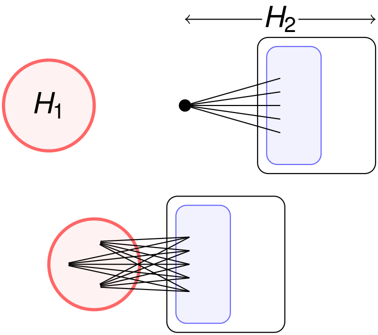
If H_1, H_2 have the EH-property, and H is obtained by substituting H_1 for a vertex of H_2, then H has the EH-property.



H has the EH-property if there exists $c > 0$ such that $\max(\alpha(G), \omega(G)) \geq |G|^c$ for every H-free graph G.

Theorem (Alon, Pach, Solymosi, 2001)

If H_1, H_2 have the EH-property, and H is obtained by substituting H_1 for a vertex of H_2, then H has the EH-property.



Theorem

The following graphs are known to have the EH-property:

- P_4*-free graphs (Erdős, Hajnal, 1989)*

Theorem

The following graphs are known to have the EH-property:

- P_4*-free graphs (Erdős, Hajnal, 1989)*
- *Graphs that can be made from graphs with the EH-property by substitution (Alon, Pach, Solymosi, 2001)*

Theorem

The following graphs are known to have the EH-property:

- *P_4-free graphs (Erdős, Hajnal, 1989)*
- *Graphs that can be made from graphs with the EH-property by substitution (Alon, Pach, Solymosi, 2001)*
- *The bull (Chudnovsky, Safra, 2008)*

Theorem

The following graphs are known to have the EH-property:

- P_4*-free graphs (Erdős, Hajnal, 1989)*
- *Graphs that can be made from graphs with the EH-property by substitution (Alon, Pach, Solymosi, 2001)*
- *The bull (Chudnovsky, Safra, 2008)*
- *The cycle of length five (Chudnovsky, Scott, S., Spirkl, 2021).*

Theorem

The following graphs are known to have the EH-property:

- P_4*-free graphs (Erdős, Hajnal, 1989)*
- *Graphs that can be made from graphs with the EH-property by substitution (Alon, Pach, Solymosi, 2001)*
- *The bull (Chudnovsky, Safra, 2008)*
- *The cycle of length five (Chudnovsky, Scott, S., Spirkl, 2021).*

It is open whether P_5 has the EH-property.

Theorem

The following graphs are known to have the EH-property:

- *P_4-free graphs (Erdős, Hajnal, 1989)*
- *Graphs that can be made from graphs with the EH-property by substitution (Alon, Pach, Solymosi, 2001)*
- *The bull (Chudnovsky, Safra, 2008)*
- *The cycle of length five (Chudnovsky, Scott, S., Spirkl, 2021).*

It is open whether P_5 has the EH-property.

Theorem (Blanco, Bucić, 2022)

There exists $c > 0$ such that

$$\max(\alpha(G), \omega(G)) \geq 2^{c(\log|G|)^{2/3}}.$$

for every P_5-free graph G.

Cograph: P_4-free graph. Equivalently, a graph that can be constructed starting from one-vertex graphs by repeatedly taking disjoint unions and complete joins.
Define $\mu(G)$ = size of largest induced cograph in G.

Conjecture (Erdős-Hajnal, equivalent form)

For every graph H, there exists $c > 0$ such that $\mu(G) \geq |G|^c$ for every H-free graph G.

Conjecture (Erdős-Hajnal, equivalent form)

For every graph H, there exists $c > 0$ such that $\mu(G) \geq |G|^c$ for every H-free graph G.

Theorem (Erdős, Hajnal, 1989)

For every H there exists $c > 0$ such that $\mu(G) \geq 2^{c\sqrt{\log |G|}}$ for every H-free graph G.

Conjecture (Erdős-Hajnal, equivalent form)

For every graph H, there exists $c > 0$ such that $\mu(G) \geq |G|^c$ for every H-free graph G.

Theorem (Erdős, Hajnal, 1989)

For every H there exists $c > 0$ such that $\mu(G) \geq 2^{c\sqrt{\log |G|}}$ for every H-free graph G.

Theorem (Bucić, Nguyen, Scott, S., 2022)

For every H there exists $c > 0$ such that

$$\mu(G) \geq 2^{c\sqrt{\log |G| \log \log |G|}}$$

for every H-free graph G.

Theorem

For every H there exists $c > 0$ such that

- $\mu(G) \geq 2^{c \log |G|}$ *?? (The E-H conjecture)*

for every H-free graph G.

Theorem

For every H there exists $c > 0$ such that

- $\mu(G) \geq 2^{c \log |G|}$ *?? (The E-H conjecture)*
- $\mu(G) \geq 2^{c\sqrt{\log |G|}}$ *(Erdős, Hajnal, 1989)*

for every H-free graph G.

Theorem

For every H there exists $c > 0$ such that

- $\mu(G) \geq 2^{c \log |G|}$ *?? (The E-H conjecture)*
- $\mu(G) \geq 2^{c\sqrt{\log |G|}}$ *(Erdős, Hajnal, 1989)*
- $\mu(G) \geq 2^{c\sqrt{\log |G| \log \log |G|}}$ *(Bucić, Nguyen, Scott, S., 2022)*

for every H-free graph G.

Disjoint subsets A, B of $V(G)$ are
complete if every vertex in A is adjacent to every vertex in B;
anticomplete if there are no edges between A, B;
a pure pair if A is either complete or anticomplete to B.

Disjoint subsets A, B of $V(G)$ are
complete if every vertex in A is adjacent to every vertex in B;
anticomplete if there are no edges between A, B;
a pure pair if A is either complete or anticomplete to B.

- If we could prove that every H-free graph has a pure pair (A, B) with $|A|, |B| \geq \Omega(|G|)$, we could prove H has the EH-property. But this is true only for very small graphs H.

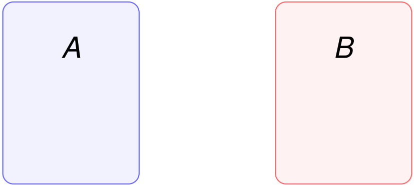
Disjoint subsets A, B of $V(G)$ are
complete if every vertex in A is adjacent to every vertex in B;
anticomplete if there are no edges between A, B;
a pure pair if A is either complete or anticomplete to B.

- If we could prove that every H-free graph has a pure pair (A, B) with $|A|, |B| \geq \Omega(|G|)$, we could prove H has the EH-property. But this is true only for very small graphs H.
- If every H-free graph G has a pure pair (A, B) with $|A|, |B| \geq \Omega(|G|/\mu(G)^k)$, then $\mu(G) \geq 2^{c\sqrt{\log |G|}}$ for H-free graphs.

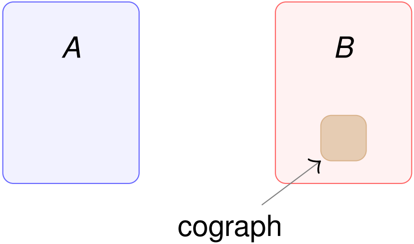
Disjoint subsets A, B of $V(G)$ are
complete if every vertex in A is adjacent to every vertex in B;
anticomplete if there are no edges between A, B;
a pure pair if A is either complete or anticomplete to B.

- If we could prove that every H-free graph has a pure pair (A, B) with $|A|, |B| \geq \Omega(|G|)$, we could prove H has the EH-property. But this is true only for very small graphs H.
- If every H-free graph G has a pure pair (A, B) with $|A|, |B| \geq \Omega(|G|/\mu(G)^k)$, then $\mu(G) \geq 2^{c\sqrt{\log|G|}}$ for H-free graphs.
- If every H-free graph has a pure pair (A, B) with $|A| \geq \Omega(|G|/\mu(G)^k)$ and $|B| \geq \Omega(|G|)$, then $\mu(G) \geq 2^{c\sqrt{\log|G| \log\log|G|}}$ for H-free graphs.

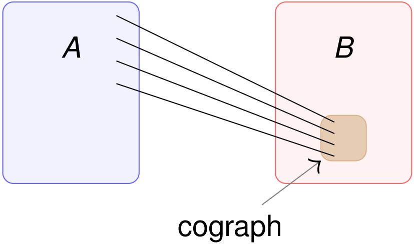
(A, B) is almost-pure if either every vertex in B has at most $|A|/(2\mu(G))$ neighbours in A, or every vertex in B has at most $|A|/(2\mu(G))$ non-neighbours in A.



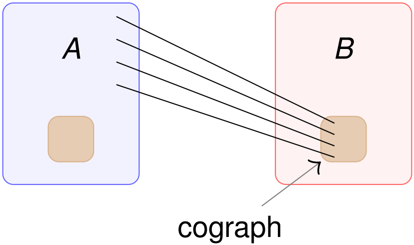
(A, B) is almost-pure if either every vertex in B has at most $|A|/(2\mu(G))$ neighbours in A, or every vertex in B has at most $|A|/(2\mu(G))$ non-neighbours in A.



(A, B) is almost-pure if either every vertex in B has at most $|A|/(2\mu(G))$ neighbours in A, or every vertex in B has at most $|A|/(2\mu(G))$ non-neighbours in A.



(A, B) is almost-pure if either every vertex in B has at most $|A|/(2\mu(G))$ neighbours in A, or every vertex in B has at most $|A|/(2\mu(G))$ non-neighbours in A.



(A, B) is almost-pure if either every vertex in B has at most $|A|/(2\mu(G))$ neighbours in A, or every vertex in B has at most $|A|/(2\mu(G))$ non-neighbours in A.

(A, B) is almost-pure if either every vertex in B has at most $|A|/(2\mu(G))$ neighbours in A, or every vertex in B has at most $|A|/(2\mu(G))$ non-neighbours in A.

- If every H-free graph has an almost-pure pair (A, B) with $|A|, |B| \geq \Omega(|G|)$, then H has the EH-property. But still this is true only for very small graphs H.

(A, B) is almost-pure if either every vertex in B has at most $|A|/(2\mu(G))$ neighbours in A, or every vertex in B has at most $|A|/(2\mu(G))$ non-neighbours in A.

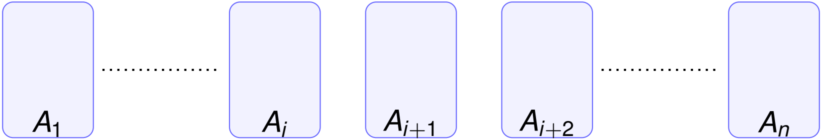
- If every H-free graph has an almost-pure pair (A, B) with $|A|, |B| \geq \Omega(|G|)$, then H has the EH-property. But still this is true only for very small graphs H.
- If every H-free graph has an almost-pure pair (A, B) with $|A|, |B| \geq \Omega(|G|/\mu(G)^k)$, then $\mu(G) \geq 2^{c\sqrt{\log |G|}}$ for H-free graphs. This is true, for all H, and this is how Erdős and Hajnal proved their theorem.

(A, B) is almost-pure if either every vertex in B has at most $|A|/(2\mu(G))$ neighbours in A, or every vertex in B has at most $|A|/(2\mu(G))$ non-neighbours in A.

- If every H-free graph has an almost-pure pair (A, B) with $|A|, |B| \geq \Omega(|G|)$, then H has the EH-property. But still this is true only for very small graphs H.
- If every H-free graph has an almost-pure pair (A, B) with $|A|, |B| \geq \Omega(|G|/\mu(G)^k)$, then $\mu(G) \geq 2^{c\sqrt{\log |G|}}$ for H-free graphs. This is true, for all H, and this is how Erdős and Hajnal proved their theorem.
- If every H-free graph has an almost-pure pair (A, B) with $|A| \geq \Omega(|G|/\mu(G)^k)$ and $|B| \geq \Omega(|G|)$, then $\mu(G) \geq 2^{c\sqrt{\log |G| \log \log |G|}}$ for H-free graphs. This is still open, but related to what we do.

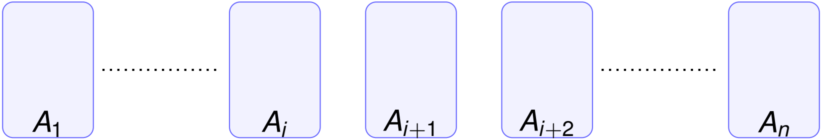
Theorem

For all H, there exist $k > 0$ such that for every H-free graph G and every x with $0 < x \leq \frac{1}{8|H|}$, there is a sequence $A_1, \ldots, A_n$ of disjoint subsets of $V(G)$ with $n \geq \log(1/x)$, and each of cardinality at least $\lfloor x^k |G| \rfloor$, such that for $1 \leq i \leq n$, either
every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ neighbours in A_i, or
every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ non-neighbours in A_i.



Theorem

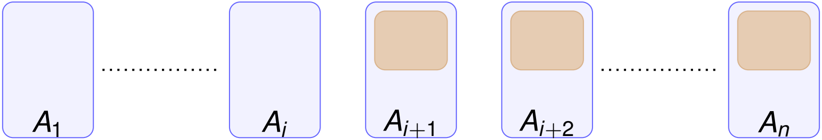
For all H, there exist $k > 0$ such that for every H-free graph G and every x with $0 < x \leq \frac{1}{8|H|}$, there is a sequence $A_1, \ldots, A_n$ of disjoint subsets of $V(G)$ with $n \geq \log(1/x)$, and each of cardinality at least $\lfloor x^k |G| \rfloor$, such that for $1 \leq i \leq n$, either
every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ neighbours in A_i, or
every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ non-neighbours in A_i.



To deduce $\mu(G) \geq 2^{c\sqrt{\log|G| \log\log|G|}}$: take $x = 1/(2\mu(G))$.

Theorem

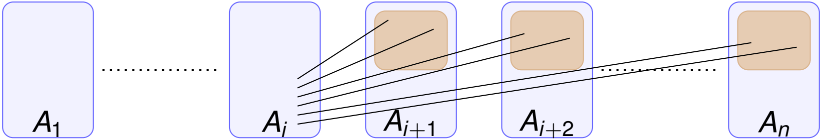
For all H, there exist $k > 0$ such that for every H-free graph G and every x with $0 < x \leq \frac{1}{8|H|}$, there is a sequence $A_1, \ldots, A_n$ of disjoint subsets of $V(G)$ with $n \geq \log(1/x)$, and each of cardinality at least $\lfloor x^k |G| \rfloor$, such that for $1 \leq i \leq n$, either
every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ neighbours in A_i, or
every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ non-neighbours in A_i.



To deduce $\mu(G) \geq 2^{c\sqrt{\log |G| \log\log |G|}}$: take $x = 1/(2\mu(G))$. Assume the union of the brown boxes is a cograph.

Theorem

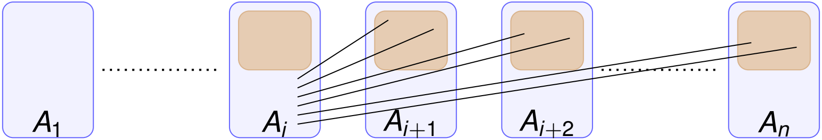
For all H, there exist $k > 0$ such that for every H-free graph G and every x with $0 < x \leq \frac{1}{8|H|}$, there is a sequence $A_1, \ldots, A_n$ of disjoint subsets of $V(G)$ with $n \geq \log(1/x)$, and each of cardinality at least $\lfloor x^k |G| \rfloor$, such that for $1 \leq i \leq n$, either every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ neighbours in A_i, or every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ non-neighbours in A_i.



To deduce $\mu(G) \geq 2^{c\sqrt{\log|G| \log\log|G|}}$: take $x = 1/(2\mu(G))$. Assume the union of the brown boxes is a cograph.

Theorem

For all H, there exist $k > 0$ such that for every H-free graph G and every x with $0 < x \leq \frac{1}{8|H|}$, there is a sequence $A_1, \ldots, A_n$ of disjoint subsets of $V(G)$ with $n \geq \log(1/x)$, and each of cardinality at least $\lfloor x^k |G| \rfloor$, such that for $1 \leq i \leq n$, either
every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ neighbours in A_i, or
every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ non-neighbours in A_i.



To deduce $\mu(G) \geq 2^{c\sqrt{\log |G| \log \log |G|}}$: take $x = 1/(2\mu(G))$. Assume the union of the brown boxes is a cograph.

Main theorem

$\mathrm{ind}_H(G)$: No of isomorphisms from H to induced subgraphs of G.

Main theorem

$\text{ind}_H(G)$: No of isomorphisms from H to induced subgraphs of G.

Theorem

For all H, there exist $k_1, k_2 > 0$ such that for every graph G and every x with $0 < x \leq \frac{1}{8|H|}$, if $\text{ind}_H(G) < x^{k_1}|G|^{|H|}$, there is a sequence $A_1, \ldots, A_n$ of disjoint subsets of $V(G)$ with $n \geq \log(1/x)$, and each of cardinality at least $\lfloor x^{k_2}|G| \rfloor$, such that for $1 \leq i \leq n$, either every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ neighbours in A_i, or every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ non-neighbours in A_i.

Key lemma

Let $g \in V(H)$. Let $A, B \subseteq V(G)$, and suppose every vertex in A is nonadjacent to at least a moderate amount of B.

Key lemma

Let $g \in V(H)$. Let $A, B \subseteq V(G)$, and suppose every vertex in A is nonadjacent to at least a moderate amount of B. Then either:

- $G[A \cup B]$ contains many copies of H;

Key lemma

Let $g \in V(H)$. Let $A, B \subseteq V(G)$, and suppose every vertex in A is nonadjacent to at least a moderate amount of B. Then either:

- $G[A \cup B]$ contains many copies of H; or
- there exists some $B' \subseteq B$, not too small, such that $G[B']$ contains surprisingly few copies of $H \setminus g$;

Key lemma

Let $g \in V(H)$. Let $A, B \subseteq V(G)$, and suppose every vertex in A is nonadjacent to at least a moderate amount of B. Then either:

- $G[A \cup B]$ contains many copies of H; or
- there exists some $B' \subseteq B$, not too small, such that $G[B']$ contains surprisingly few copies of $H \setminus g$; or
- there exist $A' \subseteq A$ and $B' \subseteq B$, not too small, such that there are very few edges between A' and B'.

Key lemma

Theorem

Let H be a graph and let $g \in V(H)$. Let $b, c > 0$, and let $a := b + (1 + c)|H|$. Let G be a graph, let A, B be disjoint subsets of $V(G)$, and let $0 < x \leq 1/2$. Suppose that every vertex in A has at least $x|B|$ non-neighbours in B. Then either:

- $\mathrm{ind}_H(G) \geq x^a|A| \cdot |B|^{|H|-1}$; *or*
- *there exists $B' \subseteq B$ with $|B'| \geq x|B|$ such that $\mathrm{ind}_{H \setminus g}(G[B']) < x^b|B'|^{|H|-1}$; or*
- *there exists $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq x^a|A|$ and $|B'| \geq x^a|B|$ such that the number of edges between A', B' is at most $2x^c|A'| \cdot |B'|$.*

Key lemma

Theorem

Let H be a graph and let $g \in V(H)$. Let $b, c > 0$, and let $a := b + (1 + c)|H|$. Let G be a graph, let A, B be disjoint subsets of $V(G)$, and let $0 < x \leq 1/2$. Suppose that every vertex in A has at least $x|B|$ non-neighbours in B. Then either:

- *there are at least $x^a|A| \cdot |B|^{|H|-1}$ isomorphisms ϕ from H to induced subgraphs of G where $\phi(g) \in A$ and $\phi(h) \in B$ for all other $h \in V(H)$; or*
- *there exists $B' \subseteq B$ with $|B'| \geq x|B|$ such that $\text{ind}_{H \setminus g}(G[B']) < x^b|B'|^{|H|-1}$; or*
- *there exists $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq x^a|A|$ and $|B'| \geq x^a|B|$ such that the number of edges between A', B' is at most $2x^c|A'| \cdot |B'|$.*

Key lemma

Theorem

Let H be a graph and let $g \in V(H)$. Let $b, c > 0$, and let $a := b + (1 + c)|H|$. Let G be a graph, let A, B be disjoint subsets of $V(G)$, and let $0 < x \leq 1/2$. Suppose that every vertex in A has at least $x|B|$ non-neighbours in B. Then either:

- *there are at least $x^a|A| \cdot |B|^{|H|-1}$ isomorphisms ϕ from H to induced subgraphs of G where $\phi(g) \in A$ and $\phi(h) \in B$ for all other $h \in V(H)$; or*
- *there exists $B' \subseteq B$ with $|B'| \geq x|B|$ such that $\text{ind}_{H \setminus g}(G[B']) < x^b|B'|^{|H|-1}$; or*
- *there exists $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq x^a|A|$ and $|B'| \geq x^a|B|$ such that the number of edges between A', B' is at most $2x^c|A'| \cdot |B'|$.*

Key lemma

Theorem

Let H be a graph and let $g \in V(H)$. Let $b, c > 0$, and let $a := b + (1 + c)|H|$. Let G be a graph, let A, B be disjoint subsets of $V(G)$, and let $0 < x \leq 1/2$. Suppose that every vertex in A has at least $x|B|$ non-neighbours in B. Then either:

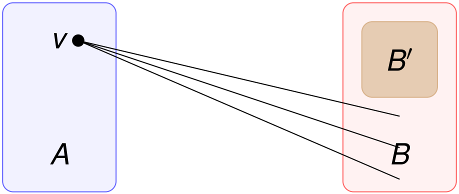
- *there are at least $x^{|H|-1+b+cd}|A| \cdot |B|^{|H|-1}$ isomorphisms ϕ from H to induced subgraphs of G where $\phi(g) \in A$ and $\phi(h) \in B$ for all other $h \in V(H)$, where g has degree d in H; or*
- *there exists $B' \subseteq B$ with $|B'| \geq x|B|$ such that $\text{ind}_{H \setminus g}(G[B']) < x^b|B'|^{|H|-1}$; or*
- *there exists $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq x^a|A|$ and $|B'| \geq x^a|B|$ such that the number of edges between A', B' is at most $2x^c|A'| \cdot |B'|$.*

Key lemma

Proof: Induction on d.
Base case $d = 0$. Let $v \in A$, and let B' be its non-neighbours in B. So $|B'| \geq x|B|$.

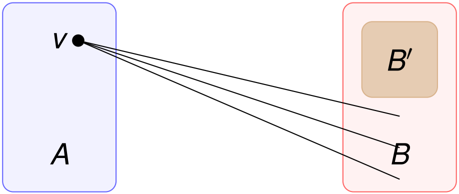
Key lemma

Proof: Induction on d.
Base case $d = 0$. Let $v \in A$, and let B' be its non-neighbours in B. So $|B'| \geq x|B|$.



Key lemma

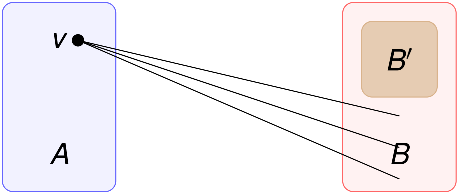
Proof: Induction on d.
Base case $d = 0$. Let $v \in A$, and let B' be its non-neighbours in B. So $|B'| \geq x|B|$.



There are $x^b|B'|^{|H|-1} \geq x^{b+|H|-1}|B|^{|H|-1}$ copies of $H \setminus g$ in $G[B']$ (or else the second outcome holds)

Key lemma

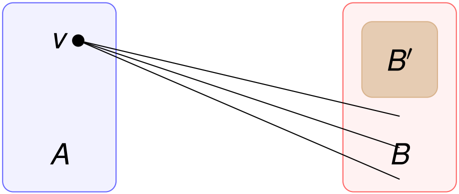
Proof: Induction on d.
Base case $d = 0$. Let $v \in A$, and let B' be its non-neighbours in B. So $|B'| \geq x|B|$.



There are $x^b|B'|^{|H|-1} \geq x^{b+|H|-1}|B|^{|H|-1}$ copies of $H \setminus g$ in $G[B']$ (or else the second outcome holds)
So there are $x^{b+|H|-1}|B|^{|H|-1}$ copies of H where g is mapped to v and all the rest is mapped into B.

Key lemma

Proof: Induction on d.
Base case $d = 0$. Let $v \in A$, and let B' be its non-neighbours in B. So $|B'| \geq x|B|$.



There are $x^b|B'|^{|H|-1} \geq x^{b+|H|-1}|B|^{|H|-1}$ copies of $H \setminus g$ in $G[B']$ (or else the second outcome holds)
So there are $x^{b+|H|-1}|B|^{|H|-1}$ copies of H where g is mapped to v and all the rest is mapped into B.
So there are $x^{b+|H|-1}|A| \cdot |B|^{|H|-1}$ copies of H where g is mapped into A and all the rest is mapped into B.

Key lemma

Inductive case $d > 0$. Let $e = gh$ be an edge incident with g.

Key lemma

Inductive case $d > 0$. Let $e = gh$ be an edge incident with g.

- From the induction, we may assume that there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|^{|H|-1}$ copies of $H \setminus e$ in G where g is mapped into A and all the rest into B. (Call these "good" copies.)

Key lemma

Inductive case $d > 0$. Let $e = gh$ be an edge incident with g.

- From the induction, we may assume that there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|^{|H|-1}$ copies of $H \setminus e$ in G where g is mapped into A and all the rest into B. (Call these "good" copies.)
- There are at most $|B|^{|H|-2}$ copies of $H \setminus \{g, h\}$ in $G[B]$.

Key lemma

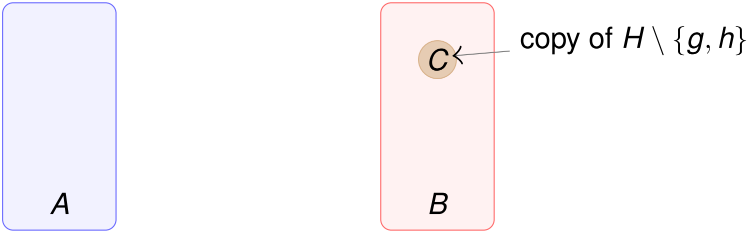
Inductive case $d > 0$. Let $e = gh$ be an edge incident with g.

- From the induction, we may assume that there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|^{|H|-1}$ copies of $H \setminus e$ in G where g is mapped into A and all the rest into B. (Call these "good" copies.)
- There are at most $|B|^{|H|-2}$ copies of $H \setminus \{g, h\}$ in $G[B]$.
- So on average, each such copy extends to $x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ good copies of $H \setminus e$.

Key lemma

Inductive case $d > 0$. Let $e = gh$ be an edge incident with g.

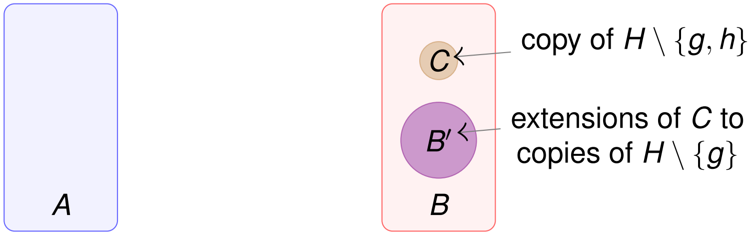
- From the induction, we may assume that there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|^{|H|-1}$ copies of $H \setminus e$ in G where g is mapped into A and all the rest into B. (Call these "good" copies.)
- There are at most $|B|^{|H|-2}$ copies of $H \setminus \{g, h\}$ in $G[B]$.
- So on average, each such copy extends to $x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ good copies of $H \setminus e$.



Key lemma

Inductive case $d > 0$. Let $e = gh$ be an edge incident with g.

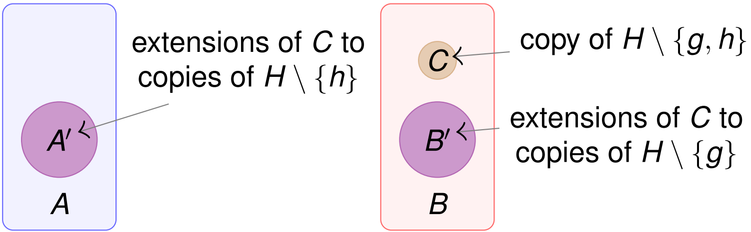
- From the induction, we may assume that there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|^{|H|-1}$ copies of $H \setminus e$ in G where g is mapped into A and all the rest into B. (Call these "good" copies.)
- There are at most $|B|^{|H|-2}$ copies of $H \setminus \{g, h\}$ in $G[B]$.
- So on average, each such copy extends to $x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ good copies of $H \setminus e$.

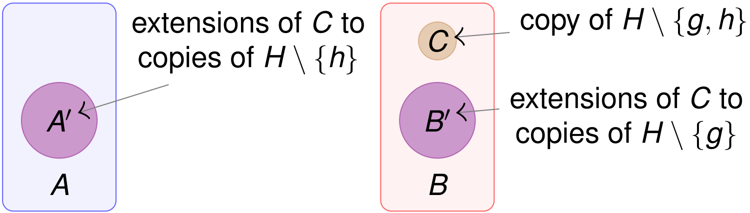


Key lemma

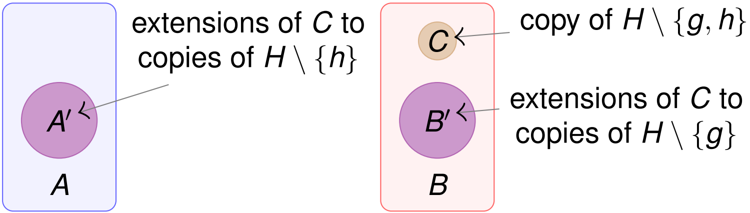
Inductive case $d > 0$. Let $e = gh$ be an edge incident with g.

- From the induction, we may assume that there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|^{|H|-1}$ copies of $H \setminus e$ in G where g is mapped into A and all the rest into B. (Call these "good" copies.)
- There are at most $|B|^{|H|-2}$ copies of $H \setminus \{g, h\}$ in $G[B]$.
- So on average, each such copy extends to $x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ good copies of $H \setminus e$.

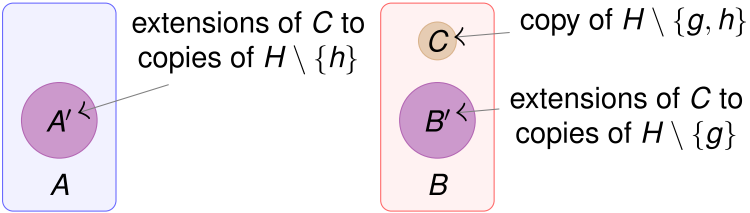




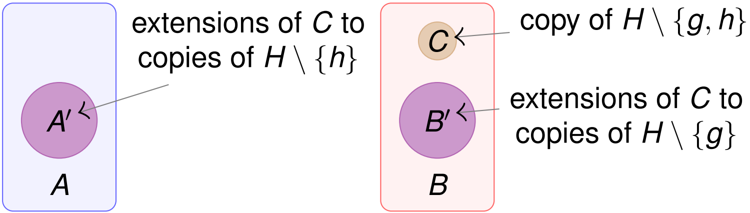
extensions of C to copies of $H \setminus \{h\}$
copy of $H \setminus \{g, h\}$
C
extensions of C to copies of $H \setminus \{g\}$
A'
B'
A
B



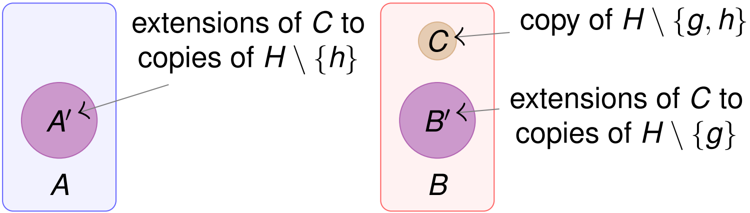
- On average (over the choices of C) there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ nonedges between A' and B'.



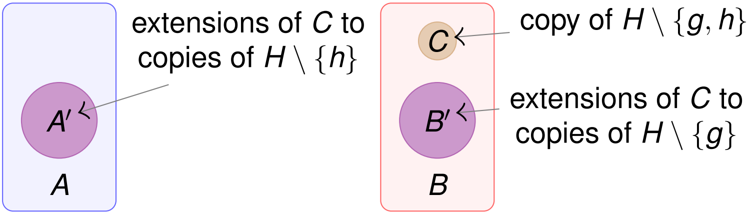
- On average (over the choices of C) there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ nonedges between A' and B'.
- Ignore all choices of C with fewer than $\frac{1}{2}x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ nonedges between A' and B' (this loses at most half of the good copies of $H \setminus e$).



- On average (over the choices of C) there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ nonedges between A' and B'.
- Ignore all choices of C with fewer than $\frac{1}{2}x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ nonedges between A' and B' (this loses at most half of the good copies of $H \setminus e$).
- So for all C, $|A'| \geq \frac{1}{2}x^{|H|-1+b+c(d-1)}|A| \geq x^a|A|$ and $|B'| \geq x^a|B|$.



- On average (over the choices of C) there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ nonedges between A' and B'.
- Ignore all choices of C with fewer than $\frac{1}{2}x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ nonedges between A' and B' (this loses at most half of the good copies of $H \setminus e$).
- So for all C, $|A'| \geq \frac{1}{2}x^{|H|-1+b+c(d-1)}|A| \geq x^a|A|$ and $|B'| \geq x^a|B|$.
- If for some choice of C, there are only $2x^c|A'| \cdot |B'|$ edges between A', B', the third outcome holds.



- On average (over the choices of C) there are at least $x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ nonedges between A' and B'.
- Ignore all choices of C with fewer than $\frac{1}{2}x^{|H|-1+b+c(d-1)}|A| \cdot |B|$ nonedges between A' and B' (this loses at most half of the good copies of $H \setminus e$).
- So for all C, $|A'| \geq \frac{1}{2}x^{|H|-1+b+c(d-1)}|A| \geq x^a|A|$ and $|B'| \geq x^a|B|$.
- If for some choice of C, there are only $2x^c|A'| \cdot |B'|$ edges between A', B', the third outcome holds.
- Otherwise, there are always at least $2x^c|A'| \cdot |B'|$ edges between A', B'; so the number of good copies of H is big and the first outcome holds.

Approximate blowups

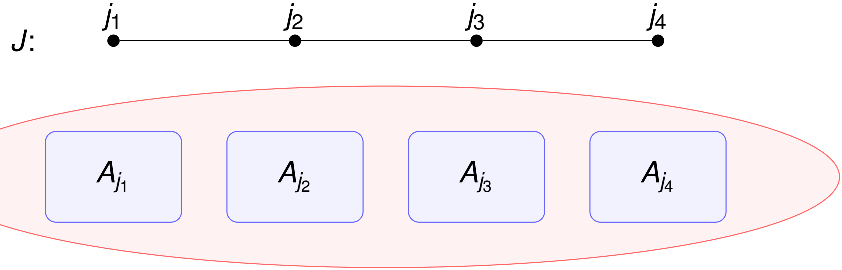
J is a graph, $t > 0$ an integer, and $q \leq 1$ a real number. A (t, q)-blowup of J in G means a family A_j ($j \in V(J)$) of pairwise disjoint subsets of $V(G)$, all of size t, such that for all distinct $i, j \in V(J)$,

- if $ij \notin E(J)$ then every vertex in A_i has at most $q|A_j|$ neighbours in A_j and vice versa;
- if $ij \in E(J)$ then every vertex in A_i has at most $q|A_j|$ non-neighbours in A_j and vice versa.

Approximate blowups

J is a graph, $t > 0$ an integer, and $q \leq 1$ a real number. A (t, q)-blowup of J in G means a family A_j ($j \in V(J)$) of pairwise disjoint subsets of $V(G)$, all of size t, such that for all distinct $i, j \in V(J)$,

- if $ij \notin E(J)$ then every vertex in A_i has at most $q|A_j|$ neighbours in A_j and vice versa;
- if $ij \in E(J)$ then every vertex in A_i has at most $q|A_j|$ non-neighbours in A_j and vice versa.



Proof of the main theorem

Theorem

For all H, there exist $k_1, k_2 > 0$ such that for every graph G and every x with $0 < x \leq \frac{1}{8|H|}$, if $\mathrm{ind}_H(G) < x^{k_1}|G|^{|H|}$, there is a sequence $A_1, \ldots, A_n$ of disjoint subsets of $V(G)$ with $n \geq \log(1/x)$, and each of cardinality at least $\lfloor x^{k_2}|G| \rfloor$, such that for $1 \leq i \leq n$, either every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ neighbours in A_i, or every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ non-neighbours in A_i.

Proof of the main theorem

Theorem

For all H, there exist $k_1, k_2 > 0$ such that for every graph G and every x with $0 < x \leq \frac{1}{8|H|}$, if $\mathrm{ind}_H(G) < x^{k_1}|G|^{|H|}$, there is a sequence $A_1, \ldots, A_n$ of disjoint subsets of $V(G)$ with $n \geq \log(1/x)$, and each of cardinality at least $\lfloor x^{k_2}|G| \rfloor$, such that for $1 \leq i \leq n$, either every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ neighbours in A_i, or every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ non-neighbours in A_i.

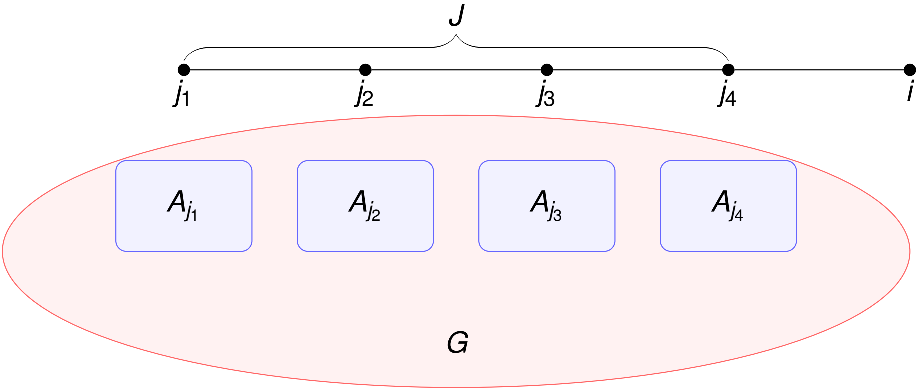
- Choose an induced subgraph J of H maximal such that there is an approximate blowup of J in G. (ie a (t, q)-blowup where $t = \lfloor x^{r_1}|G| \rfloor$ and $q = x^{r_2}$ for appropriate r_1, r_2 depending on J.)

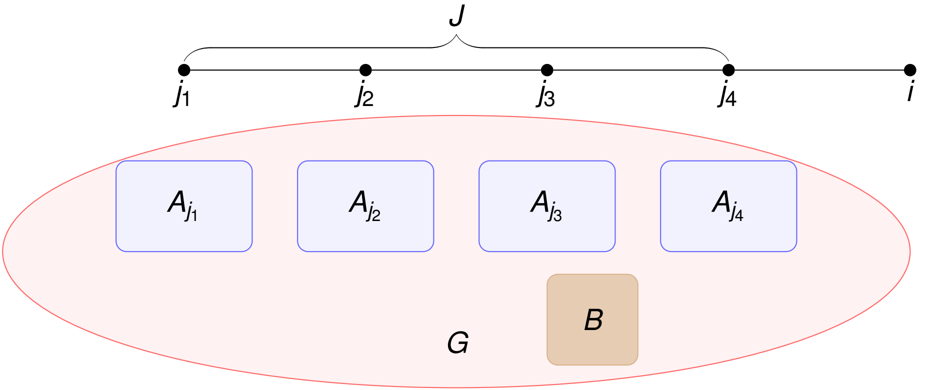
Proof of the main theorem

Theorem

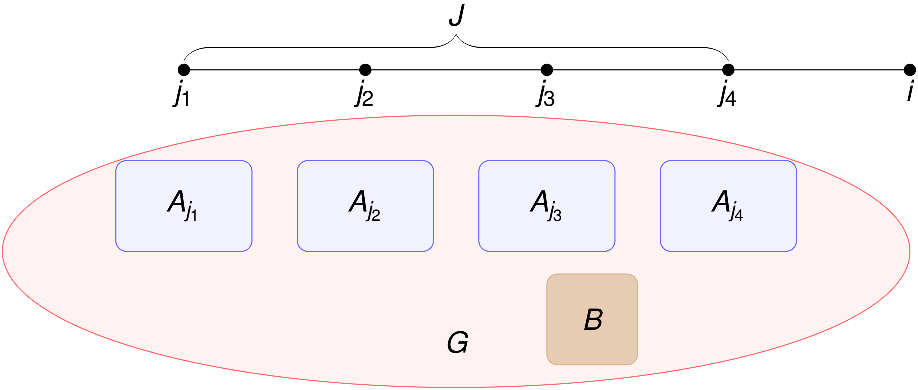
For all H, there exist $k_1, k_2 > 0$ such that for every graph G and every x with $0 < x \le \frac{1}{8|H|}$, if $\mathrm{ind}_H(G) < x^{k_1}|G|^{|H|}$, there is a sequence $A_1, \ldots, A_n$ of disjoint subsets of $V(G)$ with $n \ge \log(1/x)$, and each of cardinality at least $\lfloor x^{k_2}|G| \rfloor$, such that for $1 \le i \le n$, either every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ neighbours in A_i, or every vertex of $A_{i+1} \cup \cdots \cup A_n$ has at most $x|A_i|$ non-neighbours in A_i.

- Choose an induced subgraph J of H maximal such that there is an approximate blowup of J in G. (ie a (t, q)-blowup where $t = \lfloor x^{r_1}|G| \rfloor$ and $q = x^{r_2}$ for appropriate r_1, r_2 depending on J.)
- $J \ne H$ since $\mathrm{ind}_H(G) < x^{k_1}|G|^{|H|}$. Choose $i \in V(H) \setminus V(J)$.

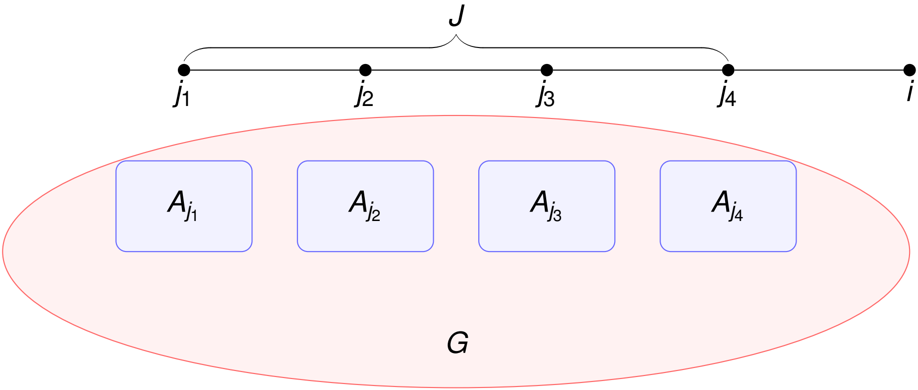

J
j1
j2
j3
j4
i
A j1
A j2
A j3
A j4
G



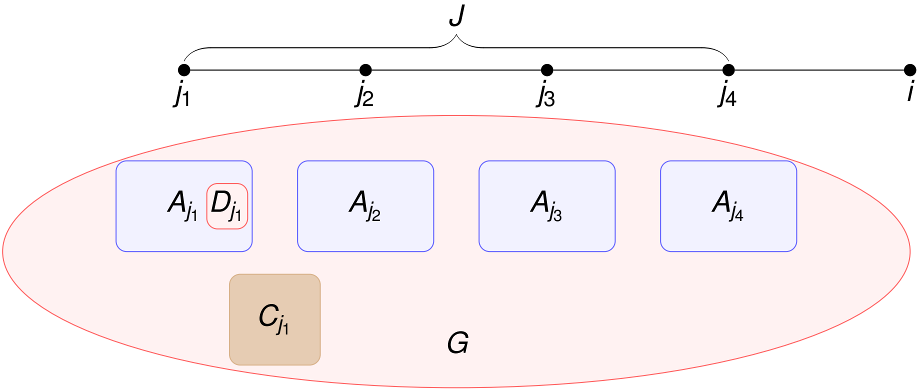
Case 1: there is a subset B disjoint from the A_j's, that is very sparse to some A_j, and has size $c|G|$.



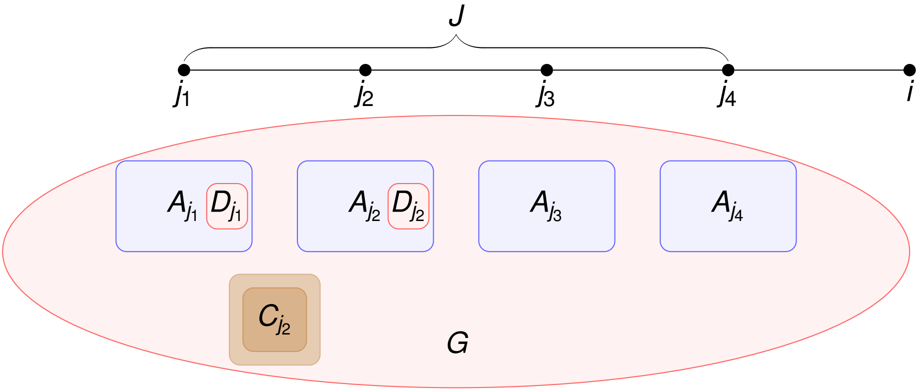
Case 1: there is a subset B disjoint from the A_j's, that is very sparse to some A_j, and has size $c|G|$. Start again, working completely inside B. If this happens many times we generate the sequence of subsets of the theorem.



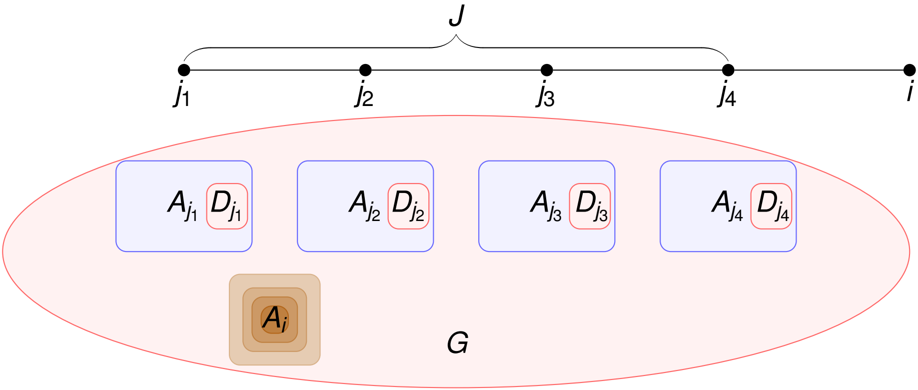
So most vertices in $V(G) \setminus \bigcup_{j \in V(J)} A_j$ are adjacent to at least a small fraction of each A_j, and also nonadjacent to at least a small fraction of each A_j.



So most vertices in $V(G) \setminus \bigcup_{j \in V(J)} A_j$ are adjacent to at least a small fraction of each A_j, and also nonadjacent to at least a small fraction of each A_j. Use the key lemma to get a subset C_{j_1} of $V(G) \setminus \bigcup_{j \in V(J)} A_j$, not too small, that is very dense or very sparse (whichever we want) to a subset $D_{j_1} \subseteq A_{j_1}$ that is not too small.



Repeat to get $C_{j_2} \subseteq C_{j_1}$ not too small, that is dense or sparse to a subset $D_{j_2} \subseteq A_{j_2}$ that is not too small.



Repeat to get $C_{j_2} \subseteq C_{j_1}$ not too small, that is dense or sparse to a subset $D_{j_2} \subseteq A_{j_2}$ that is not too small.
Repeat for all other A_j. This give an approximate blowup of J+i, contrary to the choice of J.