# Linear sized homogenous sets

Andreas Würfl

Technische Universität München

Frauenwörth, August 2010

*(joint work with Julia Böttcher & Anusch Taraz)*

# In the beginning there was...

## Conjecture

Erdős, Hajnal 1989

For every graph $H$ there is $\varepsilon(H) > 0$ such that all graphs $G \in \textsc{Forb}_n^*(H)$ have $\hom(G) \geqslant n^{\varepsilon(H)}$.

## Definition

For (simple, undirected) graphs $G$, $H$ let

- $\hom(G) = \max\{\alpha(G), \omega(G)\}$,
- $\textsc{Forb}_n^*(H) = \{G : |G| = n, H \not\leq G\}$.

# In the beginning there was...

## Conjecture

ERDŐS, HAJNAL 1989

For every graph $H$ there is $\varepsilon(H) > 0$ such that all graphs $G \in \textsc{Forb}^*_n(H)$ have $\hom(G) \geqslant n^{\varepsilon(H)}$.

## Definition

For (simple, undirected) graphs $G$, $H$ let

- $\hom(G) = \max\{\alpha(G), \omega(G)\}$,
- $\textsc{Forb}^*_n(H) = \{G : |G| = n, H \not\leq G\}$.

**Ramsey:** $\hom(G) = \Theta(\log n)$ for arbitrary $G$

## Conjecture

ERDŐS, HAJNAL 1989

For every graph $H$ there is $\varepsilon(H) > 0$ such that all graphs $G \in \textsc{Forb}^*_n(H)$ have $\hom(G) \geqslant n^{\varepsilon(H)}$.

## Theorem

LOEBL, REED, SCOTT, THOMASON, THOMASSÉ, 2010

For every graph $H$ there is $\varepsilon(H) > 0$ such that almost all graphs $G \in \textsc{Forb}_n^*(H)$ have $\hom(G) \geqslant n^{\varepsilon(H)}$.

## Theorem

Loebl, Reed, Scott, Thomason, Thomassé, 2010

For every graph $H$ there is $\varepsilon(H) > 0$ such that almost all graphs $G \in \textsc{Forb}_n^*(H)$ have $\hom(G) \geqslant n^{\varepsilon(H)}$.

## Question

Loebl, Reed, Scott, Thomason, Thomassé, 2010

Which graphs $H$ have *the property* that there is $\eta(H) > 0$ such that almost all graphs $G \in \textsc{Forb}_n^*(H)$ have $\hom(G) \geqslant \eta(H) \cdot n$, i.e.,

$$\frac{|\textsc{Forb}_{n,\eta(H)}^*(H)|}{|\textsc{Forb}_n^*(H)|} \to 0 \quad (n \to \infty) \ ?$$

$$\textsc{Forb}_{n,\eta}^*(H) = \{G \in \textsc{Forb}_n^*(H) : \hom(G) < \eta n\}$$

# What is known

## An easy case

For $H = P_3$ every graph in $\textsc{Forb}^*_n(H)$ is the disjoint union of cliques.

- $\textsc{Forb}^*_n(H)$ is dominated by graphs $G$ with $\omega(G) = \Theta(\log n)$
- $\Rightarrow P_3$ does not have *the property*

# What is known

## An easy case

For $H = P_3$ every graph in $\textsc{Forb}^*_n(H)$ is the disjoint union of cliques.

- $\textsc{Forb}^*_n(H)$ is dominated by graphs $G$ with $\omega(G) = \Theta(\log n)$
- $\Rightarrow P_3$ does not have *the property*

## Loebl, Reed, Scott, Thomason, Thomassé:

- $P_4$ might not have the property
- all other graphs seem to have *the property*

# The pentagon

$H = C_5$ has *the property* even in a stricter sense:

**Theorem** — Böttcher, Taraz, W, 2010

For every $c \in (0, 1)$ there exists $\eta > 0$ such that

$$\frac{|\textsc{Forb}^*_{n,\eta}(C_5, c)|}{|\textsc{Forb}^*_n(C_5, c)|} \to 0 \quad (n \to \infty).$$

$\textsc{Forb}^*_n(H, c) = \{G : H \not\leq G, |G| = n, \|G\| = c\binom{n}{2}\}$

# Sketch of the proof

## Theorem

For every $c \in (0,1)$ there exists $\eta > 0$ such that

$$\frac{|\textsc{Forb}^*_{n,\eta}(C_5, c)|}{|\textsc{Forb}^*_n(C_5, c)|} \to 0 \quad (n \to \infty).$$

1. find a lower bound on the size of $|\textsc{Forb}^*_n(C_5, c)|$
2. apply the regularity lemma to all graphs in $\textsc{Forb}^*_n(C_5, c)$
3. count graphs that share a reduced graph
4. have a closer look at regular pairs
5. count graphs that share a reduced graph again

# Step 1:

Define

$$h(c) := \begin{cases} H(2c)/2 & \text{if } 0 < c < \frac{1}{4}\,, \\ 1/2 & \text{if } \frac{1}{4} \leqslant c \leqslant \frac{3}{4}\,, \\ H(2c-1)/2 & \text{otherwise}\,. \end{cases}$$

Here $H(c) = -c\log c - (1-c)\log(1-c)$ is the binary entropy function.

## Binomial coefficients

For every $c \in (0,1)$

$$\binom{n}{cn} = 2^{H(c)n + o(n)}.$$

# Step 1:

Define

$$h(c) := \begin{cases} H(2c)/2 & \text{if } 0 < c < \frac{1}{4}, \\ 1/2 & \text{if } \frac{1}{4} \leqslant c \leqslant \frac{3}{4}, \\ H(2c-1)/2 & \text{otherwise}. \end{cases}$$

Here $H(c) = -c \log c - (1-c)\log(1-c)$ is the binary entropy function.

## Lemma

For every $c \in (0, 1)$

$$|\text{Forb}_n^*(C_5, c)| \geqslant 2^{h(c)\binom{n}{2} + o(1)\binom{n}{2}}.$$

## Lemma

For every $c \in (0, 1)$

$$|\textsc{Forb}^*_n(C_5, c)| \geqslant 2^{h(c)\binom{n}{2}+o(1)\binom{n}{2}}.$$

## Lemma

For every $c \in (0,1)$

$$|\text{FORB}^*_n(C_5, c)| \geqslant 2^{h(c)\binom{n}{2} + o(1)\binom{n}{2}}.$$

Our Aim:

## Lemma

For every $c \in (0,1)$ there is $\eta > 0$ such that

$$|\text{FORB}^*_{n,\eta}(C_5, c)| \leqslant 2^{(h(c)-r(c))\binom{n}{2} + o(1)\binom{n}{2}},$$

where $r : (0,1) \to \mathbb{R}^+$.

# Step 2:

Regularize all graphs in $\textsc{Forb}^*_n(C_5, c)$ to obtain *coloured reduced graphs*:

## Definition

A graph $G$ with regular partition $V_0 \dot\cup V_1 \dot\cup \ldots \dot\cup V_k$ has the coloured reduced graph $R = (V_R, E_R, \sigma)$ if

$$\sigma : \quad E_R \to \{\blacksquare, {\color{gray}\blacksquare}, \square\}$$

$$\{v_i, v_j\} \mapsto \begin{cases} \blacksquare & \text{if } d(V_i, V_j) > (1-d) \\ {\color{gray}\blacksquare} & \text{if } d(V_i, V_j) \in [d, 1-d] \\ \square & \text{if } d(V_i, V_j) < d. \end{cases}$$

## Step 3:

Let $\mathcal{R}(R, \varepsilon, n, c) = \{G \in \text{Forb}^*_n(C_5, c) : G \text{ has col. red. graph } R\}$.

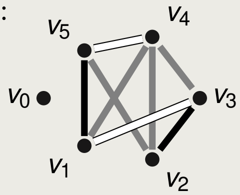
# Step 3:

Let $\mathcal{R}(R, \varepsilon, n, c) = \{G \in \text{FORB}^*_n(C_5, c) : G$ has col. red. graph $R\}$.

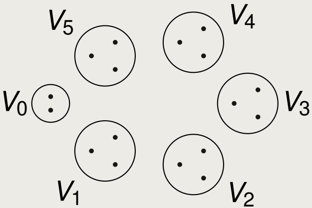
## Remark

An upper bound on $|\mathcal{R}(R, \varepsilon, n, c)|$ is given by the product of the number

$R$ :




$G$ :



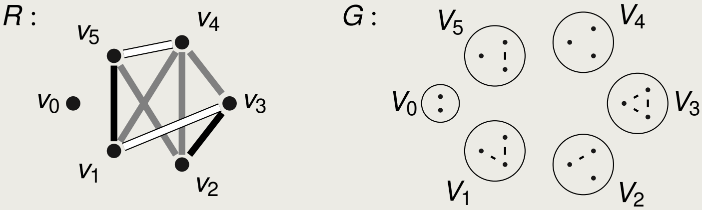
# Step 3:

Let $\mathcal{R}(R, \varepsilon, n, c) = \{G \in \textsc{Forb}^*_n(C_5, c) : G$ has col. red. graph $R\}$.

## Remark

An upper bound on $|\mathcal{R}(R, \varepsilon, n, c)|$ is given by the product of the number

- of choices for edges inside clusters,



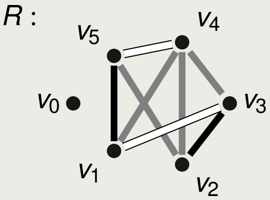
# Step 3:

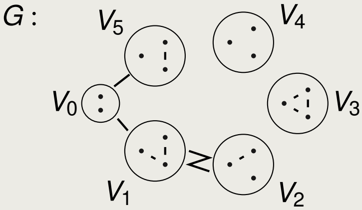
Let $\mathcal{R}(R, \varepsilon, n, c) = \{G \in \text{FORB}_n^*(C_5, c) : G \text{ has col. red. graph } R\}$.

## Remark

An upper bound on $|\mathcal{R}(R, \varepsilon, n, c)|$ is given by the product of the number

- of choices for edges inside clusters,
- of choices for edges in irregular pairs and to $V_0$,



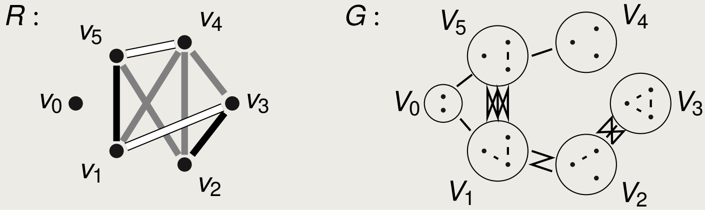
# Step 3:

Let $\mathcal{R}(R, \varepsilon, n, c) = \{G \in \text{FORB}^*_n(C_5, c) : G \text{ has col. red. graph } R\}$.

## Remark

An upper bound on $|\mathcal{R}(R, \varepsilon, n, c)|$ is given by the product of the number

- of choices for edges inside clusters,
- of choices for edges in irregular pairs and to $V_0$,
- of choices for edges in white or black eges of $R$,



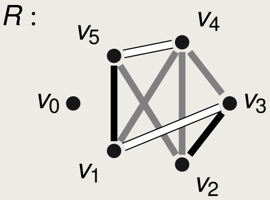
# Step 3:

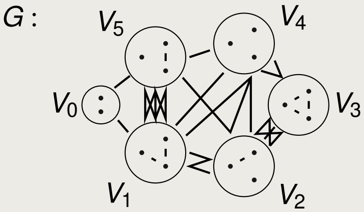
Let $\mathcal{R}(R, \varepsilon, n, c) = \{G \in \text{FORB}^*_n(C_5, c) : G \text{ has col. red. graph } R\}$.

## Remark

An upper bound on $|\mathcal{R}(R, \varepsilon, n, c)|$ is given by the product of the number

- of choices for edges inside clusters,
- of choices for edges in irregular pairs and to $V_0$,
- of choices for edges in white or black eges of $R$,
- of choices for edges in grey edges of $R$.



# Step 3:

## Lemma

$$|\mathcal{R}(R, \varepsilon, n, c)| \leqslant 2^{(f(\varepsilon, k_0)+h(c))\binom{n}{2}+o(1)\binom{n}{2}}$$

And $f(\varepsilon, k_0) \to 0$ as $\varepsilon, \frac{1}{k_0} \to 0$.

# Step 3:

## Lemma

$$|\mathcal{R}(R, \varepsilon, n, c)| \leqslant 2^{(f(\varepsilon,k_0)+h(c))\binom{n}{2}+o(1)\binom{n}{2}}$$

And $f(\varepsilon, k_0) \to 0$ as $\varepsilon, \frac{1}{k_0} \to 0$.

## Structure of $R$

As any grey triangle would force an induced $C_5$:

- At most half of the edges in $E_R$ are grey.
- The choices in grey edges are at most $2^{h(c)\binom{n}{2}+o(1)\binom{n}{2}}$.

# Step 4:

## Lemma

Let $G$ be a graph on $n$ vertices. Then one of the following is true:

- there is a homogenous set of size $n/6$ in $G$, or
- $G$ has $n/6$ disjoint copies of $P_3$, or
- $G$ has $n/6$ disjoint copies of $\overline{P_3}$.

# Step 4:

## Lemma

Let $G$ be a graph on $n$ vertices. Then one of the following is true:

- there is a homogenous set of size $n/6$ in $G$, or
- $G$ has $n/6$ disjoint copies of $P_3$, or
- $G$ has $n/6$ disjoint copies of $\overline{P_3}$.

- if $G$ has few disjoint copies of $P_3$ there is a large subgraph $G' \subseteq G$ with $G' \in \textsc{Forb}(P_3)$
- $G'$ is the disjoint union of cliques an thus has
  - a large homogenous set, or
  - many disjoint copies of $\overline{P_3}$.

# Step 5:

## Lemma

The number of choices for edges in grey pairs of the coloured reduced graph that do not give an induced $C_5$ is bounded by

$$2^{(h(c)-r(c))\binom{n}{2}+o(1)\binom{n}{2}}.$$

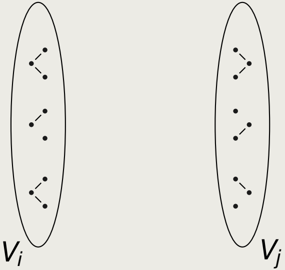
# Step 5:

## Lemma

The number of choices for edges in grey pairs of the coloured reduced graph that do not give an induced $C_5$ is bounded by

$$2^{(h(c)-r(c))\binom{n}{2}+o(1)\binom{n}{2}}.$$

- will select edges for grey pairs randomly
- $\Theta(n^2)$ pairs of $P_3$, $\overline{P_3}$ in each grey edge



# Step 5:

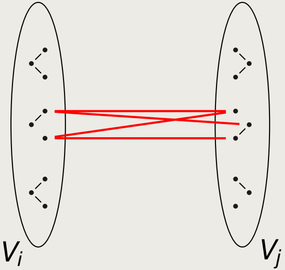
## Lemma

The number of choices for edges in grey pairs of the coloured reduced graph that do not give an induced $C_5$ is bounded by

$$2^{(h(c)-r(c))\binom{n}{2}+o(1)\binom{n}{2}}.$$

- will select edges for grey pairs randomly
- $\Theta(n^2)$ pairs of $P_3$, $\overline{P_3}$ in each grey edge
- each with a positive probability to form an induced $C_5$
- global probability to have no induced $C_5$:

$$2^{-r(c)\binom{n}{2}}$$



# Step 5:

## Finally…

- recall $|\mathcal{R}(R, \varepsilon, n, c)| \leqslant 2^{(f(\varepsilon, k_0)+h(c))\binom{n}{2}+o(1)\binom{n}{2}}$

# Step 5:

## Finally...

- recall $|\mathcal{R}(R, \varepsilon, n, c)| \leqslant 2^{(f(\varepsilon, k_0)+h(c))\binom{n}{2}+o(1)\binom{n}{2}}$
- $|\mathcal{R}(R, \varepsilon, n, c) \cap \textsc{Forb}^*_{n,\eta}(C_5, c)| \leqslant 2^{(f(\varepsilon, k_0)+h(c)-r(c))\binom{n}{2}+o(1)\binom{n}{2}}$

# Step 5:

## Finally...

- recall $|\mathcal{R}(R, \varepsilon, n, c)| \leqslant 2^{(f(\varepsilon, k_0)+h(c))\binom{n}{2}+o(1)\binom{n}{2}}$
- $|\mathcal{R}(R, \varepsilon, n, c) \cap \text{FORB}^*_{n,\eta}(C_5, c)| \leqslant 2^{(f(\varepsilon, k_0)+h(c)-r(c))\binom{n}{2}+o(1)\binom{n}{2}}$
- for $c \in (0, 1)$: $r(c) > 0$
  $\Rightarrow$ select $\varepsilon, k_0$ s.t. $f(\varepsilon, k_0) < \frac{1}{2}r(c)$

# Step 5:

## Finally...

- recall $|\mathcal{R}(R, \varepsilon, n, c)| \leqslant 2^{(f(\varepsilon,k_0)+h(c))\binom{n}{2}+o(1)\binom{n}{2}}$
- $|\mathcal{R}(R, \varepsilon, n, c) \cap \text{FORB}^*_{n,\eta}(C_5, c)| \leqslant 2^{(f(\varepsilon,k_0)+h(c)-r(c))\binom{n}{2}+o(1)\binom{n}{2}}$
- for $c \in (0, 1)$: $r(c) > 0$
  $\Rightarrow$ select $\varepsilon, k_0$ s.t. $f(\varepsilon, k_0) < \frac{1}{2}r(c)$
- number of coloured reduced graphs is constant

# Step 5:

## Finally…

- recall $|\mathcal{R}(R, \varepsilon, n, c)| \leqslant 2^{(f(\varepsilon,k_0)+h(c))\binom{n}{2}+o(1)\binom{n}{2}}$
- $|\mathcal{R}(R, \varepsilon, n, c) \cap \text{FORB}^*_{n,\eta}(C_5, c)| \leqslant 2^{(f(\varepsilon,k_0)+h(c)-r(c))\binom{n}{2}+o(1)\binom{n}{2}}$
- for $c \in (0, 1)$: $r(c) > 0$
  $\Rightarrow$ select $\varepsilon, k_0$ s.t. $f(\varepsilon, k_0) < \frac{1}{2}r(c)$
- number of coloured reduced graphs is constant
- $\frac{|\text{FORB}^*_{n,\eta}(C_5,c)|}{|\text{FORB}^*_{n}(C_5,c)|} \leqslant \frac{2^{(h(c)-0.5r(c))\binom{n}{2}+o(1)\binom{n}{2}}}{2^{h(c)\binom{n}{2}+o(1)\binom{n}{2}}} \to 0$ as $n \to \infty$

# Step 5:

## Finally…

- recall $|\mathcal{R}(R, \varepsilon, n, c)| \leqslant 2^{(f(\varepsilon,k_0)+h(c))\binom{n}{2}+o(1)\binom{n}{2}}$
- $|\mathcal{R}(R, \varepsilon, n, c) \cap \textsc{Forb}^*_{n,\eta}(C_5, c)| \leqslant 2^{(f(\varepsilon,k_0)+h(c)-r(c))\binom{n}{2}+o(1)\binom{n}{2}}$
- for $c \in (0, 1)$: $r(c) > 0$
  $\Rightarrow$ select $\varepsilon, k_0$ s.t. $f(\varepsilon, k_0) < \frac{1}{2}r(c)$
- number of coloured reduced graphs is constant
- $\frac{|\textsc{Forb}^*_{n,\eta}(C_5,c)|}{|\textsc{Forb}^*_{n}(C_5,c)|} \leqslant \frac{2^{(h(c)-0.5r(c))\binom{n}{2}+o(1)\binom{n}{2}}}{2^{h(c)\binom{n}{2}+o(1)\binom{n}{2}}} \to 0$ as $n \to \infty$
- $\eta$ only depends on the number of clusters

# What's next?

- Does this method work for other graphs $H$?
- Are there better bounds for $\eta(H)$?

# Bibliography

Loebl, Reed, Scott, Thomason, Thomassé
Almost all *H*-free graphs have the Erdős-Hajnal property
*submitted*, 2010