Applications of the Sparse Regularity Lemma

Y. Kohayakawa (Emory and São Paulo)

Extremal Combinatorics II

DIMACS 2004

Szemerédi's regularity lemma

1. Works very well for large, dense graphs: n-vertex graphs with $\geq cn^2$ edges, $n \to \infty$

Szemerédi's regularity lemma

1. Works very well for large, dense graphs: n-vertex graphs with $\geq cn^2$ edges, $n \to \infty$
2. Variant for sparse graphs exists (sparse = with $o(n^2)$ edges)

Szemerédi's regularity lemma

1. Works very well for large, dense graphs: n-vertex graphs with $\geq cn^2$ edges, $n \to \infty$
2. Variant for sparse graphs exists (sparse = with $o(n^2)$ edges)
3. Much harder to use

Szemerédi's regularity lemma

1. Works very well for large, dense graphs: n-vertex graphs with $\geq cn^2$ edges, $n \to \infty$
2. Variant for sparse graphs exists (sparse = with $o(n^2)$ edges)
3. Much harder to use
4. **This talk**: some tools to handle difficulties

Outline of the talk

1. Basic definitions and the regularity lemma
2. A simple application of the regularity lemma
3. The difficulty in the sparse setting
4. **Some tools**
5. **Subgraphs of pseudorandom graphs**

ε-regularity

Basic definition. $G = (V, E)$ a graph; U, $W \subset V$ non-empty and disjoint. Say (U, W) is ε*-regular* (in G) if

▷ *for all* $U' \subset U$, $W' \subset W$ *with* $|U'| \geq \varepsilon|U|$ *and* $|W'| \geq \varepsilon|W|$, *we have*

$$\left| \frac{|E(U', W')|}{|U'||W'|} - \frac{|E(U, W)|}{|U||W|} \right| \leq \varepsilon.$$

Szemerédi's regularity lemma

Theorem 1 (The regularity lemma). *For any* $\varepsilon > 0$ *and* $t_0 \geq 1$, *there exist* T_0 *such that any graph* G *admits a partition* $V(G) = V_1 \cup \cdots \cup V_t$ *such that*

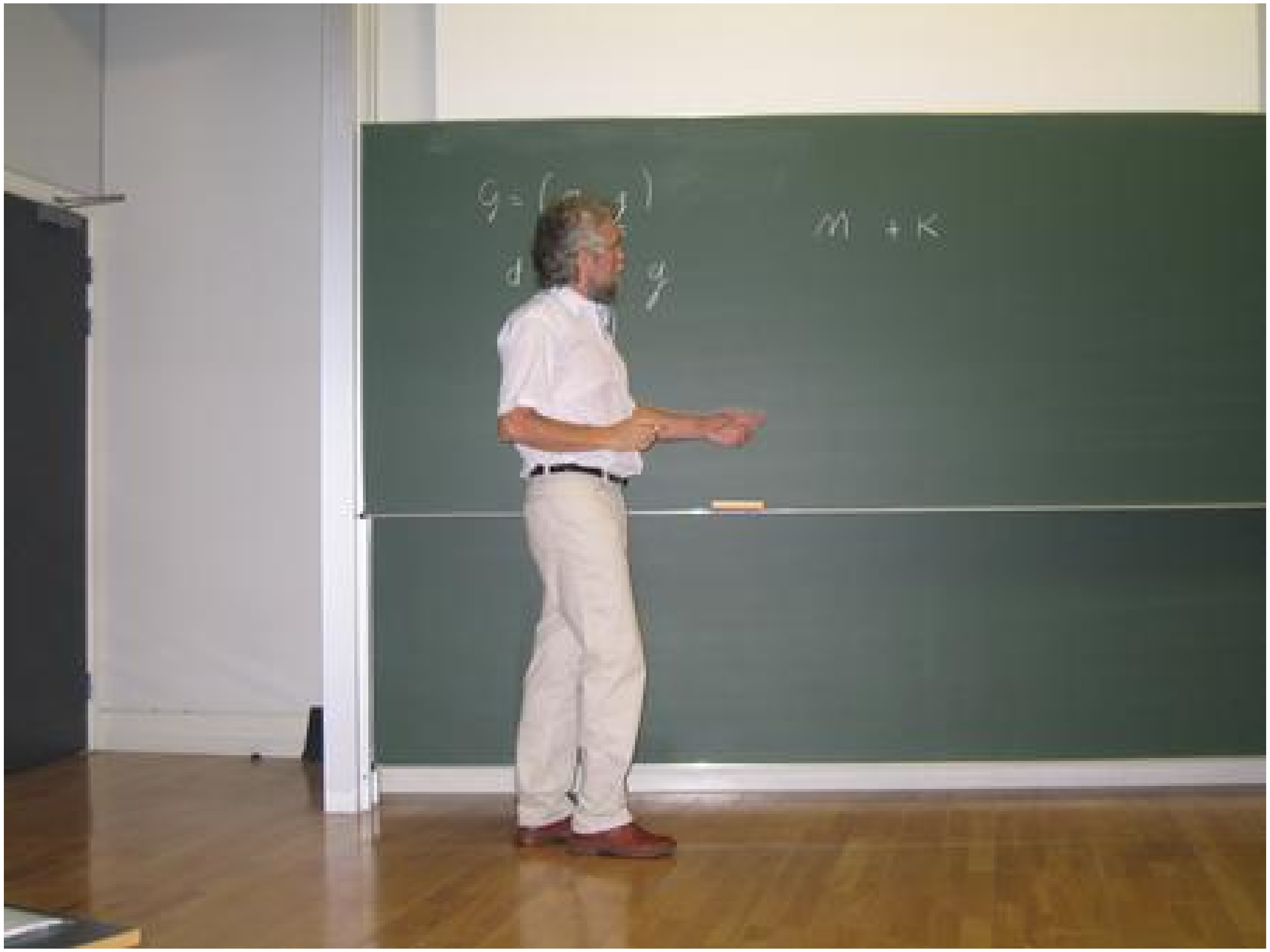
(i) $|V_1| \leq \cdots \leq |V_t| \leq |V_1| + 1$

(ii) $t_0 \leq t \leq T_0$

(iii) at least $(1 - \varepsilon)\binom{t}{2}$ *pairs* (V_i, V_j) $(i < j)$ *are* ε*-regular.*

▷ Myriads of applications

Endre Szemerédi



Endre Szemerédi

Outline of the talk

1. Basic definitions and the regularity lemma
2. A simple application of the regularity lemma
3. The difficulty in the sparse setting
4. Some tools
5. Subgraphs of pseudorandom graphs

Outline of the talk

1. Basic definitions and the regularity lemma
2. A simple application of the regularity lemma
3. The difficulty in the sparse setting and a workaround
4. Some tools
5. Subgraphs of pseudorandom graphs

ε-regularity revisited

The pair (U, W) is ε-regular if for all $U' \subset U$, $W' \subset W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$, we have

$$|E(U', W')| = |U'||W'| \left(\frac{|E(U, W)|}{|U||W|} \pm \varepsilon \right)$$

Clearly, no information if

$$\frac{|E(U, W)|}{|U||W|} \to 0$$

and ε is fixed. (We think of $G = (V, E)$ with $n = |V| \to \infty$.)

ε-regularity; scaled version

▷ Roughly: scale by the global density of the graph

Actual condition is

- *for all* $U' \subset U$, $W' \subset W$ *with* $|U'| \geq \varepsilon|U|$ *and* $|W'| \geq \varepsilon|W|$, we have

$$\left| \frac{|E(U',W')|}{p|U'||W'|} - \frac{|E(U,W)|}{p|U||W|} \right| \leq \varepsilon,$$

where $p = |E(G)|\binom{n}{2}^{-1}$.

OK even if $p \to 0$. [Terminology: (ε, p)-regular pair]

Szemerédi's regularity lemma, sparse version

Any graph with no 'dense patches' admits a Szemerédi partition with the new notion of ε-regularity.

Szemerédi's regularity lemma, sparse version

Any graph with no 'dense patches' admits a Szemerédi partition with the new notion of ε-regularity.

Definition. Say $G = (V, E)$ is *locally* (η, b)*-bounded* if for all $U \subset V$ with $|U| \geq \eta |V|$, we have

$$\#\{\text{edges within } U\} \leq b|E|\binom{|U|}{2}\binom{|V|}{2}^{-1}.$$

Szemerédi's regularity lemma, sparse version (cont'd)

Theorem 2 (The regularity lemma). *For any* $\varepsilon > 0$, $t_0 \geq 1$, *and* b, *there exist* $\eta > 0$ *and* T_0 *such that any locally* (η, b)*-bounded graph* G *admits a partition* $V(G) = V_1 \cup \cdots \cup V_t$ *such that*

(i) $|V_1| \leq \cdots \leq |V_t| \leq |V_1| + 1$

(ii) $t_0 \leq t \leq T_0$

(iii) at least $(1-\varepsilon)\binom{t}{2}$ *pairs* (V_i, V_j) $(i < j)$ *are* (ε, p)*-regular, where* $p = |E(G)|\binom{n}{2}^{-1}$.

Vojta Rödl

Simple example: $K^3 \hookrightarrow G$? (G dense)

1. *Regularize* G: apply Szemerédi's regularity lemma to G

2. Analyse the 'cleaned-up graph' G^* (Definition 7) and search for $G_3^{(\varepsilon)}(m, (\rho_{ij})) \subset G$ (Notation 8)

3. If found, OK. Can even estimate $\#\{K^3 \hookrightarrow G_3^{(\varepsilon)}(m, (\rho_{ij}))\}$ using the 'Counting Lemma' (Lemma 9)

Simple example: $K^3 \hookrightarrow G$? (G sparse)

1. *Regularize* G: apply Szemerédi's regularity lemma to G
2. Analyse the 'cleaned-up graph' G^* (Definition 7) and search for $G_3^{(\varepsilon)}(m, (\rho_{ij})) \subset G$ (Notation 8)
3. If found, OK?

Miserable: Counting Lemma is false if $\rho \to 0$

Fact 3. $\forall \varepsilon > 0\ \exists \rho > 0$, $m_0\ \forall m \geq m_0\ \exists G_3^{(\varepsilon)}(m, \rho)$ *with*

$$K^3 \not\subset G_3^{(\varepsilon)}(m, \rho).$$

[cf. Lemma 9]

An observation

Counterexamples to the embedding lemma in the sparse setting do exist (Fact 3), but

are extremely rare.

An observation

Counterexamples to the embedding lemma in the sparse setting do exist (Fact 3), but

are extremely rare.

Workaround:

- ▷ An asymptotic enumeration lemma [Lemma 10]
- ▷ Consequence for random graphs: can recover an embedding lemma for K^3 for subgraphs of random graphs [Corollary 11].

An observation

Counterexamples to the embedding lemma in the sparse setting do exist (Fact 3), but

are extremely rare.

Workaround:

- ▷ An asymptotic enumeration lemma [Lemma 10]
- ▷ Consequence for random graphs: can recover an embedding lemma for K^3 for subgraphs of random graphs [Corollary 11]. Conjecture for general graphs H [Conjecture 13].

An application

Asymptotic enumeration lemma above for K^3:

An application

Asymptotic enumeration lemma above for K^3: used in the proof of a random version of Roth's theorem (Szemerédi's theorem for $k = 3$). [Theorem 12]

Outline of the talk

1. Basic definitions and the regularity lemma
2. A simple application of the regularity lemma
3. The difficulty in the sparse setting and a workaround (enumeration)
4. Some tools
5. Subgraphs of pseudorandom graphs

Outline of the talk

1. Basic definitions and the regularity lemma
2. A simple application of the regularity lemma
3. The difficulty in the sparse setting and a workaround (enumeration)
4. Some further tools
5. Subgraphs of pseudorandom graphs

Hereditary nature of regularity

Setup. $B = (U, W; E)$ an ε-regular bipartite graph with $|U| = |W| = m$ and $|E| = \rho m^2$, $\rho > 0$ constant, and an integer d. Sample $N \subset U$ and $N' \subset W$ with $|N| = |N'| = d$ uniformly at random.

Theorem 4. *For any $\beta > 0$, $\rho > 0$, and $\varepsilon' > 0$, if $\varepsilon \leq \varepsilon_0(\beta, \rho, \varepsilon')$, $d \geq d_0(\beta, \rho, \varepsilon')$, and $m \geq m_0(\beta, \rho, \varepsilon')$, then*

$$\mathbb{P}\Big((N, N') \textit{ bad}\Big) \leq \beta^d,$$

where (N, N') is bad if $\Big||E(N, N')|d^{-2} - \rho\Big| > \varepsilon'$ or else (N, N') is not ε'-regular.

A result similar to Theorem 4 was proved by Duke and Rödl, ’85.

Hereditary nature of regularity (cont'd)

Roughly speaking, Theorem 4 is true for subgraphs of $G(n, p)$, if

$$dp^2 \gg (\log n)^4.$$

Hereditary nature of regularity (cont'd^2)

Applicable version: suppose U, W, U', $W' \subset V(G(n,p))$, pairwise disjoint, with $|U| = |W| = |U'| = |W'| = m$. Suppose (U,W) (ε,p)-regular for $H \subset G$; interested in the pair $(N_H(u')\cap U, N_H(w')\cap W)$, where $N_H(u')$ is the nbhd of $u' \in U'$ in H, &c. Suppose $p^3 m \gg (\log n)^{100}$.

Theorem 5. $\forall \varepsilon' > 0\ \exists \varepsilon > 0$: *with probability* $\to 1$ *as* $n \to \infty$ *have:* $\forall U$, W, U', $W' \subset V(G(n,p))$, $\exists U'' \subset U'$, $W'' \subset W'$ *with* $|U''|$, $|W''| \geq (1-\varepsilon')m$, *so that* $\forall u'' \in U''$, $w'' \in W''$,

$$(N_H(u'') \cap U, N_H(w'') \cap W) \text{ is } (\varepsilon', p)\text{-regular,}$$
$$\text{with density } (1 \pm \varepsilon')|E_H(U,W)|/|U||W|.$$

[K. and Rödl, 2003]

Local characterization for regularity

Setup. $B = (U, W; E)$, a bipartite graph with $|U| = |W| = m$. Consider the properties

(PC) for some constant p, have $m^{-1} \sum_{u \in U} |\deg(u) - pm| = o(m)$ and

$$\frac{1}{m^2} \sum_{u,u' \in U} |\deg(u,u') - p^2 m| = o(m).$$

(R) (U, W) is $o(1)$-regular (classical sense).

Theorem 6. (PC) *and* (R) *are equivalent.*

Local characterization for regularity (cont'd)

Roughly speaking, Theorem 6 holds for subgraphs of $G(n, p)$, as long as

$$p^2 m \gg (\log n)^{100}.$$

[K. and Rödl, 2003]

Outline of the talk

1. Basic definitions and the regularity lemma
2. A simple application of the regularity lemma
3. The difficulty in the sparse setting and a workaround (enumeration)
4. Some further tools (hereditary nature; local characterization)
5. Subgraphs of pseudorandom graphs

Outline of the talk

1. Basic definitions and the regularity lemma
2. A simple application of the regularity lemma
3. The difficulty in the sparse setting and a workaround (enumeration)
4. Some further tools (hereditary nature; local characterization)
5. Subgraphs of pseudorandom graphs

Subgraphs of pseudorandom graphs

Roughly speaking:

Subgraphs of pseudorandom graphs

Roughly speaking: the local characterization of regularity (Theorem 6) holds for subgraphs of ‘strongly pseudorandom’ graphs, e.g., Ramanujan graphs (enough: $\lambda \ll d^2/n$).

Subgraphs of pseudorandom graphs

Roughly speaking: the local characterization of regularity (Theorem 6) holds for subgraphs of 'strongly pseudorandom' graphs, e.g., Ramanujan graphs (enough: $\lambda \ll d^2/n$).

▷ Need somewhat higher densities than in the r.gs case

Subgraphs of pseudorandom graphs

Roughly speaking: the local characterization of regularity (Theorem 6) holds for subgraphs of 'strongly pseudorandom' graphs, e.g., Ramanujan graphs (enough: $\lambda \ll d^2/n$).

- ▷ Need somewhat higher densities than in the r.gs case
- ▷ Good news: should have *constructive* versions of previous results involving random graphs

[K., Rödl, Schacht, Sissokho, Skokan, 2004+]

A class of strongly pseudorandom graphs

Say G satisfies $\mathrm{STRONG\text{-}DISC}(\gamma)$ if

▷ For all disjoint U and $W \subset V(G)$, we have

$$\Big| e_G(U, W) - p_G|U||W| \Big| < \gamma p_G^2 n \sqrt{|U||W|},$$

where $p_G = |E(G)|\binom{n}{2}^{-1}$.

Roughly: graphs satisfying $\mathrm{STRONG\text{-}DISC}(o(1))$ are such that any proportional subgraph $H \subset G$ satisfying (R) satisfies (PC).

Concrete application

Theorem 6 generalizes to proportional subgraphs of (n, d, λ)-graphs with $\lambda \ll d^2/n$.

Concrete application

Theorem 6 generalizes to proportional subgraphs of (n, d, λ)-graphs with $\lambda \ll d^2/n$.

Can use this, e.g.,

1. to develop a constructive version of the regularity lemma for subgraphs of (n, d, λ)-graphs,
2. to prove counting lemmas for subgraphs of such graphs,
3. to prove Turán type results for such graphs.

Postponed stuff and others

1. Definition 7: cleaned-up graph
2. Notation 8: $G_3^{(\varepsilon)}(m, (\rho_{ij}))$
3. Lemma 9: Counting Lemma
4. Theorem 12: AP3s
5. Theorem 14: Turán problem
6. Theorem 15 and Corollary 16: fault-tolerance
7. Theorem 18: size-Ramsey numbers

Terminology: *Cleaned-up graph* G^*

Definition 7. *After regularization of G, have $V = V_1 \cup \cdots \cup V_t$. Remove all edges in $G[V_i, V_j]$ for all i and j such that*

1. (V_i, V_j) *is not ε-regular,*
2. $|E(V_i, V_j)| \leq f(\varepsilon)m^2$ (*suitable f with $f(\varepsilon) \to 0$ as $\varepsilon \to 0$*).

Resulting graph: cleaned-up graph G^*.

In G^*, every $G^*[V_i, V_j]$ is regular and 'dense'. Usually, lose very little.

Notation: $G_3^{(\varepsilon)}(m, (\rho_{ij}))$

Notation 8. *Suppose* $G = (V_1, V_2, V_3; E)$ *tripartite is such that*

1. $|V_i| = m$ *for all* i*,*
2. (V_i, V_j) ε*-regular for all* $i < j$*,*
3. $|E(V_i, V_j)| = \rho_{ij} m^2$ *for all* $i < j$*.*

Write $G_3^{(\varepsilon)}(m, (\rho_{ij}))$ *for a graph as above.*

▷ 'ε-regular triple'

A counting lemma (simplest version)

Setup. $G = (V_1, V_2, V_3; E)$ tripartite with

1. $|V_i| = m$ for all i
2. (V_i, V_j) ε-regular for all $i < j$
3. $|E(V_i, V_j)| = \rho m^2$ for all $i < j$

That is, $G = G_3^{(\varepsilon)}(m, \rho)$, i.e., G is an ε-regular triple with density ρ.

Just like random:

Lemma 9 (Counting Lemma). $\forall \rho > 0, \delta > 0\ \exists \varepsilon > 0, m_0$:
if $m \geq m_0$, *then*

$$\left|\#\{K^3 \hookrightarrow G\} - \rho^3 m^3\right| \leq \delta m^3.$$

An asymptotic enumeration lemma

Lemma 10 (K., Łuczak, Rödl, '96). $\forall\beta > 0\ \exists\varepsilon > 0,\ C > 0,\ m_0$: *if* $T = \rho m^2 \geq Cm^{3/2}$, *then*

$$\#\{G_3^{(\varepsilon)}(m,\rho) \not\supset K^3\} \leq \beta^T \binom{m^2}{T}^3.$$

Observe that $\rho \geq C/\sqrt{m} \to 0$.

Consequence for random graphs

Easy expectation calculations imply

▷ if $p \gg 1/\sqrt{n}$, then almost every $G(n,p)$ is such that

$$\left(K^3\text{-free } G_3^{(\varepsilon)}(m,\rho)\right) \not\subset G(n,p),$$

if (*) $mp \gg \log n$ and $\rho \geq \alpha p$ for some fixed α.

Conclusion. Recovered an 'embedding lemma' in the sparse setting, *for subgraphs of random graphs*.

Corollary 11 (EL for subgraphs of r.gs). *If $p \gg 1/\sqrt{n}$ and (*) holds, then almost every $G(n,p)$ is such that if $G_3^{(\varepsilon)}(m,\rho) \subset G(n,p)$, then*

$$\exists K^3 \hookrightarrow G_3^{(\varepsilon)}(m,\rho) \subset G(n,p).$$

Superexponential bounds

Suppose we wish to prove a statement about all subgraphs of $G(n, p)$.

- Too many such subgraphs: about $2^{p\binom{n}{2}}$
- $G(n, p)$ has no edges with probability $(1-p)^{\binom{n}{2}} \geq \exp\{-2pn^2\}$, if, say, $p \leq 1/2$.
- Bounds of the form

$$o(1)^T \binom{\binom{m}{2}}{T}$$

for the cardinality of a family of 'undesirable subgraphs' $U(m, T)$ do the job. Use of such bounds goes back to Füredi, '94.

An application

The above asymptotic enumeration lemma is used in the proof of the following result.

Theorem 12 (K., Łuczak, Rödl, '96). $\forall \eta > 0\ \exists C:$ *if randomly select* $R \subset \{1, \ldots, n\}$ *with* $|R| = C\sqrt{n}$*, then a.a.s.*

$$R \rightarrow_\eta \textit{AP3}.$$

$R \rightarrow_\eta$ AP3 means *any* $S \subset R$ *with* $|S| \geq \eta|R|$ *contains an AP3* (arithmetic progression of 3 terms)

General graphs H?

Let us state our conjecture for $H = K^k$.

Conjecture 13 (K., Łuczak, Rödl, '97). $\forall k \geq 4,\ \beta > 0\ \exists \varepsilon > 0,\ C > 0,\ m_0$: *if* $T = \rho m^2 \geq C m^{2-2/(k+1)}$*, then*

$$\#\{G_k^{(\varepsilon)}(m, \rho) \not\supset K^k\} \leq \beta^T \binom{m^2}{T}^{\binom{k}{2}}.$$

For general H, the conjecture involves the 2-*density* of H.

Best known so far: $k = 5$, by Gerke, Prömel, Schickinger, Steger, and Taraz, 2004.

If true, Conjecture 13 implies . . .

1. The Rödl–Ruciński theorem on threshold for Ramsey properties of random graphs and the Turán counterpart.

2. Łuczak, 2000: almost all triangle-free graphs are very close to being bipartite ($e(G^n) \gg n^{3/2}$). Conjecture 13 is the 'only' missing ingredient for the general K^{k+1}-free $\Rightarrow$ very close to k-partite.

Turán type results for subgraphs of random graphs

Theorem 14 (K., Rödl, and Schacht, '04). *Let* H *be a graph with maximum degree* $\Delta = \Delta(H)$*, and suppose*

$$np^{\Delta} \gg (\log n)^4.$$

Then

$$\mathrm{ex}(G(n,p), H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) p \binom{n}{2}$$

with probability $\to 1$ *as* $n \to \infty$.

Conjectured threshold for p:

$$np^{d_2(H)} \to \infty$$

should suffice. [If $H = K^k$, have $d_2(H) = (k+1)/2$.]

Some applications of the hereditary nature &c

1. Turán type results for subgraphs of random graphs [Theorem 14]
2. Small fault-tolerant networks [Theorem 15 and Corollary 16]
3. Size-Ramsey numbers [Theorem 18]

Some applications of the hereditary nature &c

1. Turán type results for subgraphs of random graphs
2. Small fault-tolerant networks [Theorem 15 and Corollary 16]
3. Size-Ramsey numbers

Small fault-tolerant networks

$\mathcal{B}(m, m; \Delta)$: family of m by m bipartite graphs with maximum degree $\leq \Delta$

Theorem 15 (Alon, Capalbo, K., Rödl, Ruciński, Szemerédi, ’00). *For all $\eta > 0$ and Δ, there is C such that if*

$$p = C \left(\frac{\log n}{n} \right)^{1/2\Delta} \quad \textit{and} \quad m = \lfloor n/C \rfloor,$$

then

$$G(n, n; p) \to_\eta \mathcal{B}(m, m; \Delta)$$

with probability $\to 1$ as $n \to \infty$.

Small fault-tolerant networks (cont'd)

Corollary 16. *There is an* η*-fault-tolerant graph* Γ *for* $\mathcal{B}(m, m; \Delta)$ *with* $\tilde{O}(m^{2-1/2\Delta})$ *edges.*

Remark. If $\tilde{\Gamma} \supset B$ any $B \in \mathcal{B}(m, m; \Delta)$, then

$$|E(\tilde{\Gamma})| \geq cm^{2-2/\Delta}.$$

Size-Ramsey numbers for bounded degree graphs

The *size-Ramsey number* of H is

$$r_e(H) = \min\{|E(\Gamma)| \colon \Gamma \to (H, H)\}.$$

Known that $r_e(H)$ is linear in $|V(H)|$ if H is a path (Beck, '83), tree with bounded degree (Friedman and Pippenger, '87), cycle (Haxell, K., and Łuczak, '95), and (almost linear if) H is a long subdivision (Pak, '01).

Size-Ramsey numbers for bounded degree graphs (cont'd)

Theorem 17 (Rödl and Szemerédi, '00). *$r_e(H) \geq cn(\log n)^{\alpha}$ for a certain cubic, n-vertex graph H (c and $\alpha > 0$ universal constants).*

Theorem 18 (K., Rödl and Szemerédi, '0?). *For any Δ there is $\varepsilon = \varepsilon(\Delta) > 0$ for which we have*

$$r_e(H) \leq n^{2-\varepsilon}$$

for any n-vertex graph H with $\Delta(H) \leq \Delta$.

[$\varepsilon \leq 1/2\Delta$?]

Remark

Meaning of '*assertion* $\mathbb{A}$ *is true for (proportional) subgraphs of pseudorandom graphs*' is roughly clear.

Remark

Meaning of ‘*assertion* $\mathbb{A}$ *is true for (proportional) subgraphs of pseudorandom graphs*’ is roughly clear.

However, shall often say ‘*assertion* $\mathbb{A}$ *is true for (proportional) subgraphs of random graphs*’,

Remark

Meaning of '*assertion* $\mathbb{A}$ *is true for (proportional) subgraphs of pseudorandom graphs*' is roughly clear.

However, shall often say '*assertion* $\mathbb{A}$ *is true for (proportional) subgraphs of random graphs*', which means. . .

Remark

Assertion A is true for (proportional) subgraphs of random graphs:

$\triangleright$ with probability $\to 1$ as $n \to \infty$, assertion A holds for *any* subgraph of $G(n, p)$.

Assertion A will often be an implication $P \Rightarrow Q$

$P \Rightarrow Q$ will often be true for dense graphs, *i.e.*, with $\geq cn^2$ edges,
and false for sparse graphs in general

Recent result: properties that will make our results hold for deterministic classes of graphs.

Remark

Assertion A is true for (proportional) subgraphs of random graphs:

$\triangleright$ with probability $\to 1$ as $n \to \infty$, assertion A holds for *any* subgraph of $G(n, p)$.

Assertion A will often be an implication $P \Rightarrow Q$

$P \Rightarrow Q$ will often be true for dense graphs, *i.e.*, with $\geq cn^2$ edges,
and false for sparse graphs in general

Recent result: properties that will make our results hold for deterministic classes of graphs. Turns out that, e.g., Ramanujan graphs will do (eigen-value conditions).