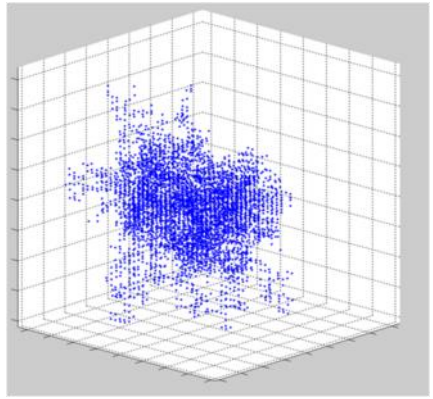
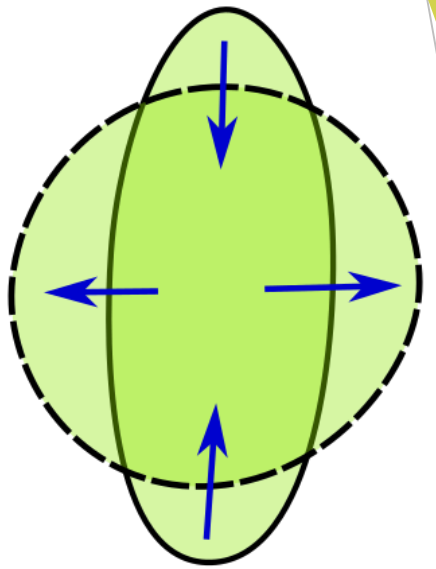
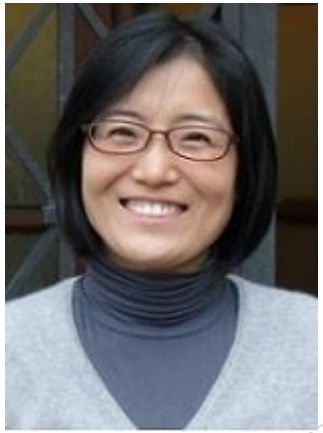


# Percolation through Isoperimetry



Michael Krivelevich
Tel Aviv University

Joint with: Sahar Diskin (Tel Aviv), Joshua Erde (TU Graz), Mihyun Kang (TU Graz)



pictures: Wikipedia, cliparts.co

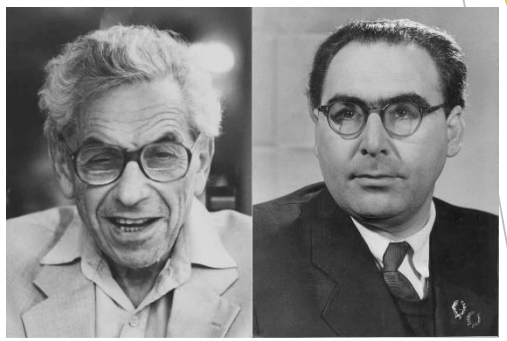
In fact, worked with related model $G(n,m)$

# Let's start with classics…

Erdős, Rényi, 1960 (800+ citations in MathSciNet; 9K+ citations in GS):

Model $G(n,p)$ : $V=[n]$

$$\forall 1 \le i < j \le n, \Pr[(i,j) \in E(G)] = p = p(n), \text{ indep.}$$



Paul Erdős (1913-1996) Alfréd Rényi (1921-1970)

Phase transition at $p=\frac{1}{n}$:

Notation: $|L_1| \ge |L_2| \ge \cdots$ – connected components of $G(n,p)$

$\epsilon > 0$ – small constant

- $p=\frac{1-\epsilon}{n}$ : whp all connected components satisfy: $|L_i| = O_\epsilon(\log n)$

  (subcritical regime)

- $p=\frac{1+\epsilon}{n}$ : whp $\exists$ a unique component $L_1$ of linear size (giant component):

  $|L_1| = \big(1+o(1)\big)y(\epsilon)n$

  all other are small: $|L_i| = O_\epsilon(\log n)\,, i \ge 2$

  (supercritical regime)

$p=\frac{1}{n}$ – critical regime,
won't be addressed in this talk

# Let's start with classics… (cont.)

Definition of $y(\epsilon)$:

$c > 1$ – constant

$y = y(c-1) :=$ unique solution of

$$y = 1 - \exp\{-cy\}$$

in $(0,1)$.

Asymptotic estimate: $y(\epsilon) = \big(1 + o_\epsilon(1)\big)2\epsilon$

Modern explanation: $p = \frac{c}{n}, c > 1$

$\forall v, \deg(v) \sim Bin(n-1, p) \approx \mathrm{Po}(c)$

$y :=$ survival probability of GW tree with # of children $X \sim \mathrm{Po}(c)$

$\approx \Pr[v \in L_1]$

# Let's start with classics… (cont.)

Definition of $y(\epsilon)$:

$y \coloneqq$ survival probability of GW tree with # of children $X \sim \mathrm{Po}(c)$, $y \in (0,1)$

Recursion:

$$1 - y = \sum_{k=0}^{\infty} \Pr[X = k]\,(1-y)^k$$

$$= \textstyle\sum_{k=0}^{\infty} e^{-c}\,\frac{c^k}{k!}\,(1-y)^k$$

$$= e^{-c} e^{c(1-y)} = e^{-cy}$$

$$\Rightarrow y = 1 - e^{-cy}$$

# Moving along the timeline…

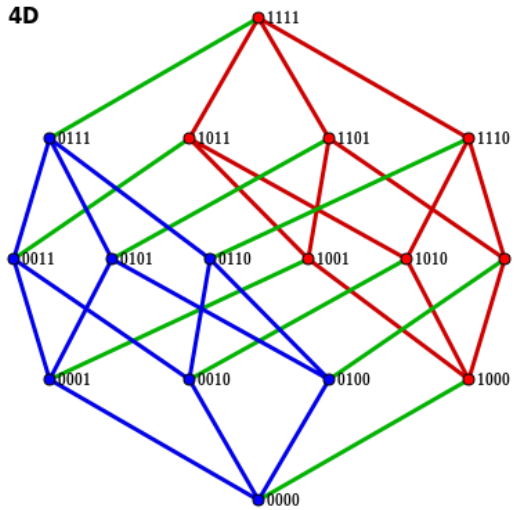
Binary hypercube:

Def.: $Q^d = d$-dimensional hypercube

$V = \{0,1\}^d$

$(\bar{x}, \bar{y}) \in E(Q^d)$ if $\mathrm{dist}(\bar{x}, \bar{y}) = 1$

$-\ n = 2^d$ vertices; $d$ –regular; bipartite




(picture: from http://n-dimensional.de/)

Percolation on $Q^d$:

Model $Q_p^d$: $\forall e \in E(Q^d), \Pr[e \in E(Q_p^d)] = p = p(d)$, indep.

# Moving along the timeline… (cont.)

Percolation on $Q^d$:

Model $Q_p^d$: $\forall e \in E(Q^d), \Pr[e \in E(Q_p^d)] = p = p(d)$, indep.

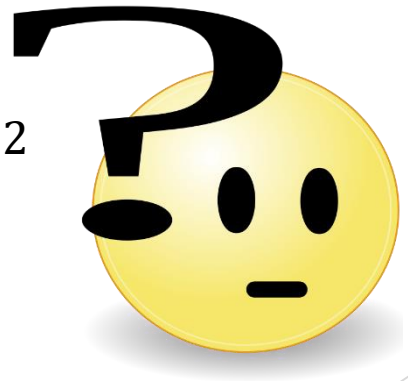
Ajtai, Komlós, Szemerédi'82; Bollobás, Kohayakawa, Łuczak'92

– phase transition at $p = \frac{1}{d}$ :

- $p = \frac{1-\epsilon}{d}$ : whp all connected components satisfy: $|L_i| = O_\epsilon(d) = O_\epsilon(\log n)$ ;
- $p = \frac{1+\epsilon}{d}$ : whp $|L_1| = (1 + o(1))y(\epsilon)n$;

$$|L_i| = O_\epsilon(d) = O_\epsilon(\log n), \quad i \geq 2$$

– results very similar to $G(n,p)$/ER!



# A general question

Model $G_p$:

$G = (V, E)$ – graph on $n$ vertices, $p = p(n)$ – edge probability

$\forall e \in E(G), \Pr[e \in E(G_p)] = p$, indep.

$G$ – $d$-regular graph, $|V(G)| = n, d \to \infty$

$p = \frac{1+\epsilon}{d}, \epsilon > 0$ – small const. (supercritical regime)

whp: $|L_1| = ?$

$|L_2| \leq ?$

In words: how general is the ER picture?

# Remarks

- Assume: $d \to \infty$

  $d = \text{const}$ – a very different regime

  (local structure matters a lot,

  Benjamini-Schramm limit, etc.)

- Subcritical regime

  $p = \frac{1-\epsilon}{d}, \epsilon > 0$ – constant

  whp $|L_i| = O_\epsilon(\log n)$

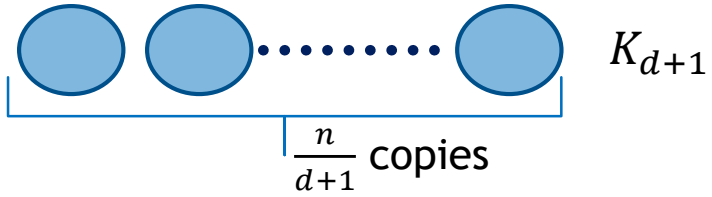
  – fairly simple to prove using, say, graph exploration algorithms

  (BFS/DFS)

# Let's be realistic…

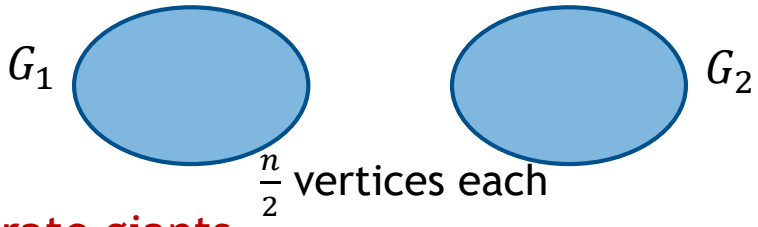
Need to assume something about (edge distribution of) $G$…

Ex.: $G =$ disjoint union of cliques $K_{d+1}$




Clearly: $|L_1| \leq d+1, \ \forall p = p(d)$

Ex.: $G =$ disjoint union of two nice $d$-regular graphs on $n/2$ vertices each




Expect two separate giants…

# Graph isoperimetry

- a key assumption
- fairly natural

Notation: $G = (V, E), \emptyset \neq S \subsetneq V -$ vertex subset

$$\partial_G(S) = \{e : |e \cap S| = |e \cap S^c| = 1\}$$

– edge boundary of $S$ in $G$

Def.: $G = (V, E), |V| = n$

$$i(G) = \min_{S \subset V, 1 \leq |S| \leq \frac{n}{2}} \frac{|\partial_G(S)|}{|S|}$$

– isoperimetric constant of $G$

# Graph isoperimetry – examples

1. $G = K_n$

   $i(G) = \frac{n}{2}$

2. $G = Q^d$

   Harper's edge isoperimetric inequality'64:

   $\forall S \subset V(Q^d), |S| = k$

   $\left|\partial_{Q^d}(S)\right| \geq dk - k\log_2 k$

   worst case: $k = 2^{d-1}$ (half cube): $|\partial(S)| \geq k$

   $\Rightarrow i(Q^d) \geq 1$

– fairly weak expander

# Graph isoperimetry — examples (cont.)

3. $G$ — $d$-regular graph on $n$ vertices

$d = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ — eigenvalues of

(the adjacency matrix of) $G$

Alon, Milman'84:

$$\forall S, |\partial_G(S)| \geq \frac{(d-\lambda_2)|S|(n-|S|)}{n}$$

$$\Rightarrow \lambda_2 \ll d \Rightarrow i(G) = \left(1 - o(1)\right)\frac{d}{2}$$

Comment: for good $d$-regular expanders $G$ can hope:

$$i(G) = \Theta(d)$$

# Main result

Th. 1: $G$ – $d$-regular graph on $n$ vertices, $d \to \infty$

$\epsilon > 0$ – small constant

Assume: $i(G) \geq C$

for some $C \geq C_0(\epsilon)$ [can be constant]

Form $G_p$: $\forall e \in E(G), \Pr[e \in E(G_p)] = p$, indep.; $p = \frac{1+\epsilon}{d}$

Then whp:

$$\left||L_1| - yn\right| \leq \frac{10n}{C^{\frac{1}{20}}}$$

where $y(\epsilon)$ is the unique solution of $y = 1 - e^{-cy}$ in $(0,1)$

Furthermore, $|L_2| \leq (1 + o(1))\frac{10n}{C^{\frac{1}{20}}}$ .

# Main result — interpretation

— the ER picture for $L_1$ holds under very mild assumption on graph isoperimetry:

$$i(G) \geq C(\epsilon)$$

Th. 1.: $G$ — $d$-regular graph on $n$ vertices, $d \to \infty$
$\epsilon > 0$ — small constant
Assume: $i(G) \geq C$
$G_p$: $\forall e \in E(G), \Pr[e \in E(G_p)] = p$, indep.; $p = \frac{1+\epsilon}{d}$
Then whp:

$$\big||L_1| - yn\big| \leq \frac{10n}{C^{\frac{1}{20}}}$$

— other components: small linear/sublinear

Q.: can it be improved? All the way to $|L_2| = O_\epsilon(\log n)$?

# Tightness 1

Q.: How tight is our assumption on $i(G)$?

Can it be further relaxed?

Th. 2: $C > 1$ – constant, $d \to \infty$

$n \gg d^2$, $p = \frac{C}{d}$

$\exists$ a $d$-regular $n$-vertex graph $G$

with $i(G) \geq \frac{1}{40C}$ s.t.

whp all components in $G_p$ are $\leq 3d\log n$ .

Conclusion: in a sense, cannot require less than $i(G) = \Omega(1)$.

# Tightness 2

Q.: What about the second largest component $L_2$ of $G_p$?

Can we get to $|L_2| = O(\log n)$ (like in ER)?

Th. 3: $d \to \infty,\ p = \frac{1+\epsilon}{d},\ d \ll e^{-30C} \cdot n$

$\exists$ a $d$-regular $n$-vertex graph $G$, $i(G) \geq C$ s.t.

whp $|L_2| \geq \epsilon d$ .

Conclusion: $i(G) \geq C$ is not (nearly) enough to deliver:

$$|L_2| = O(\log n)$$

# Local expansion and large components

Q.: What if we assume only expansion of sets $S$ up to $|S| \leq k$?

Th. 4: $k \to \infty, d \leq k, |V(G)| > k$

Assume: $\forall S \subset V(G), |S| \leq k,$

$|\partial_G(S)| \geq d|S|$ [strengthens assumption $\delta(G) \geq d$]

$p = \frac{1+\epsilon}{d}$

Then whp (in $k$) $G_p$ contains a component of size $\geq \frac{k}{2}$ .

Comment: $G = K_n$ (= back to $G(n,p)$ model)

choose $k = \delta n \Rightarrow$ edge expansion $\geq (1-\delta)n|S|$

Can conclude: $pn \geq \frac{1+\epsilon}{1-\delta}$

$\Rightarrow$ whp get a linearly sized (in $n$) component.

# Proof – Theorem 1

Some ingredients:

Lemma 1 (folklore; Beveridge, Frieze, McDiarmid'98): $G = (V, E), \Delta(G) \leq d$

$v \in V(G)$, $k \geq 1$ – integer

$t_k(v) :=$ # of trees of order $k$ in $G$ containing v

$$t_k(v) \leq \frac{k^{k-2} d^{k-1}}{(k-1)!} \leq (ed)^{k-1}$$

– useful when estimating total volume of small components in $G_p$

Lemma 2: $G = (V, E), \Delta(G) \leq d$, $0 < \delta \leq \frac{1}{2}$ – constant

$F \subseteq E(G), |F| =: s = \Omega(d)$

$q := \frac{\delta}{d}$; $F_q :=$random subset of $F$: $\Pr[e \in F_q] = q$, indep.

$\Pr[F_q$ contains a matching of size $\geq \frac{\delta^2 s}{d}] \geq 1 - \exp\{-\frac{\delta^2 s}{d}\}$

– useful when connecting large pieces of $G_{p_1}$ through sprinkling

# Proof – Theorem 1 (cont.)

Sprinkling: $1-p=(1-p_1)(1-p_2)$

$p_2=\frac{\delta}{d},\ \delta \ll \epsilon,\ p_1 \geq \frac{1+\epsilon-\delta}{d}$

$G_1 \sim G_{p_1},\ G_2 \sim G_1 \cup G_{p_2} \sim G_p$

Key lemma: large components are well spread

$V_L(G_1) \coloneqq \{v: |C_{G_1}(v)| \geq \delta d\}$ – large components

$W_L(G_1) \coloneqq \{w: |N_G(w) \cap V_L(G_1)| \geq \delta^5 d\}$ – many neighbors in large components

Expect: $\forall v, \Pr[v \in V_L(G_1)] \approx y(\epsilon-\delta)$

$\Rightarrow \Omega(\epsilon)d$ neighbors of $w$ to fall into $V_L(G_1)$

Lemma 3: whp $|W_L(G_1)| \geq (1-\exp\{-\delta^6 d\})n$

– useful in patching large components of $G_{p_1}$ through sprinkling

# Proof – Theorem 1 (cont.)

Proof steps

Sprinkling: $G_p = G_{p_1} \cup G_{p_2}$, $p_2 = \frac{\delta}{d}$

1. Counting small components (through Lemma 1, á la ER)

$\Rightarrow$ whp: $|V_L(G_1)| \geq \big(1 - o(1)\big)y(\epsilon - \delta)n$

$$\big|V_L\big(G_p\big)\big| \leq \big(1 + o(1)\big)y(\epsilon)n$$

2. Merging most of the volume in $V_L(G_1)$ (but $\leq \delta n$ vertices)

by sprinkling (through Lemma 2, á la AKS)

$\Rightarrow$ whp get a giant $L_1, |L_1| \geq (y(\epsilon) - c\delta)n$

$$|L_1| \leq \big(1 + o(1)\big)y(\epsilon)n$$

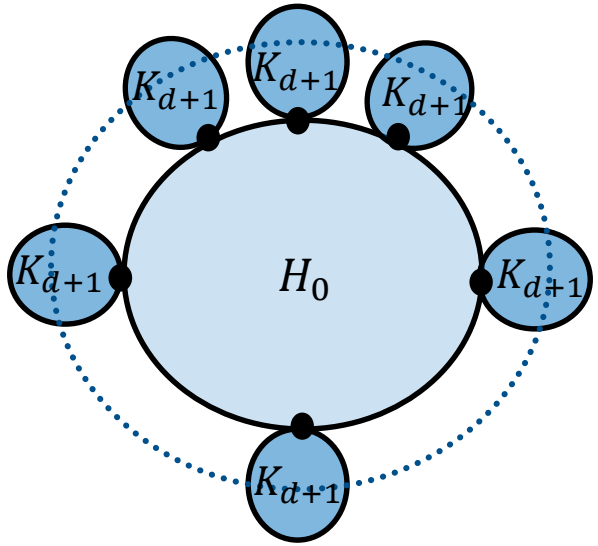
all but $O(\delta)n$ of remaining vertices – reside in small components. ■

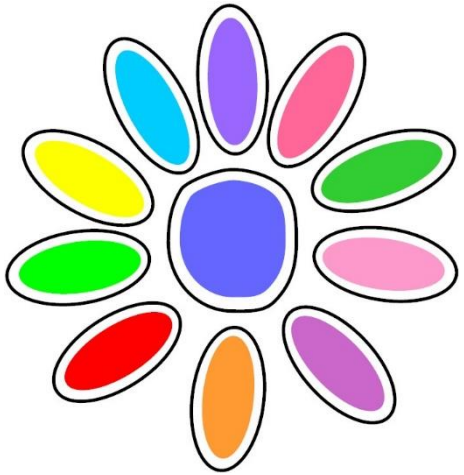
# Proof — Theorem 2

[ Need: construction of $G$ — $d$-regular, $i(G) = \Theta(1)$, $p = \frac{c}{d}$,

whp in $G_p$ all components are small]

Construction



- $|V(H_0)| = \frac{n}{d+2}$, $H_0 = d_0$-regular, excellent expander, $d_0 = \Theta(d)$
- attach a clique of size $d + 1$ to each vertex of $H_0 =: G$

  [+ massage it a bit to make $G$ $d$-regular]

# Proof – Theorem 2 (cont.)

Analysis

- $i(G) = \Theta(1)$ – tedious but straightforward
- $G_p$ has a component of size $k$

  $\Rightarrow G_p[V(H_0)]$ has a component of size $\geq \frac{k}{d+1}$

  (to connect pieces of attached cliques)
- $G_p[V(H_0)]$ – subcritical $(pd_0 < 1)$

  $\Rightarrow$ whp all connected components inside $V(H_0)$ are $O(\log n)$

---

$\Rightarrow$ whp $|L_1(G_p)| = O(d\log n)$ . ■

# Proof – Theorem 3

[ Need: construction of $G - d$-regular, $i(G) = \Theta(1)$, $p = \frac{1+\epsilon}{d}$,

whp $|L_2|$ large]

Construction

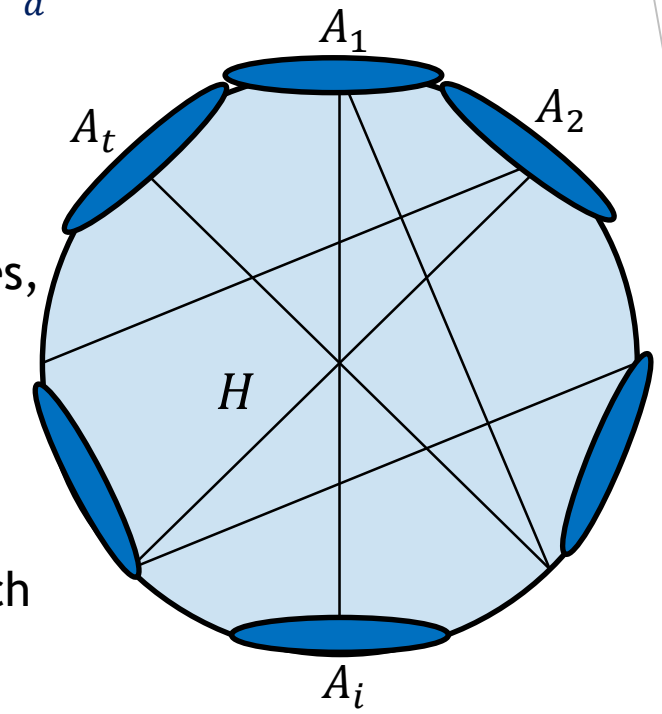
- $H = C_1$-regular expander on $n = t(d + 1 - C_1)$ vertices,

  $C_1$ – large constant, $i(H) \geq \frac{C_1}{3}$

- Equicolor $H$ in $t$ colors $A_1, \ldots, A_t$ of size $d + 1 - C_1$ each

- Put a clique of size $d + 1 - C_1$ in each color class =: $G$

  $i(G) \geq i(H) \geq \frac{C_1}{3}$




$G[A_i] \cong K_{d+1-C_1}$

# Proof – Theorem 3 (cont.)

Analysis

$G_p[A_i]$ – supercritical ($G[A_i] \cong K_{d+1-C_1}$, $dp \approx 1+\epsilon$)

⇒ whp get $\Theta(t)$ pieces of size $\geq \epsilon d$

$\Pr[A_i$ is detached from all other $A_j$'s in $G_p] = (1-p)^{C_1(d+1)} \approx \exp\{-\Theta(C_1)\}$

⇒ whp get $\exp\{-\Theta(C_1)\} \cdot t$ detached components of size $\Theta(d)$ (inside $A_i$'s) in $G_p$

---

⇒ whp many components of size $\Theta(d)$ in $G_p$ . ■

# Open questions

- ? Even tighter requirements on $i(G)$

  to guarantee whp $\left|L_1\left(G_p\right)\right| \approx y(\epsilon)n$?

  Have shown: $i(G) \geq \mathrm{poly}(\frac{1}{\epsilon})$ suffices;

  $i(G) =$ absolute constant might not suffice

- ? Getting ER phenomenon ($|L_1|; |L_i|, i \geq 2$) in full generality?
  - conditions guaranteeing:

    whp $|L_2| = O_\epsilon(\log n)$?

- ? Barely supercritical regime? $\epsilon \to 0$ slowly?


THE END