

Random subgraphs of the hypercube

Mihyun Kang

Joint work with

Joshua Erde (TU Graz) and Michael Krivelevich (Tel Aviv University)

ADYN Summer School 2022

Talk outline

- I. Random subgraphs
- II. Expansion properties and consequences
- III. Proof ideas of expansion
- IV. Mixing time of lazy random walk on giant component

Part I.

Random subgraphs

Given $p \in (0, 1)$

- $G(n, p)$ = a graph obtained by retaining each edge of complete graph K_n independently with probability p
- = bond percolation on complete graph K_n with edge probability p

Part I.

Random subgraphs

Given $p \in (0, 1)$

$G(n, p)$ = a graph obtained by retaining each edge of complete graph K_n independently with probability p

= bond percolation on complete graph K_n with edge probability p

G_p = a graph obtained by retaining each edge of a given base graph G independently with probability p

= bond percolation on G with edge probability p

The hypercube

Given $d \in \mathbb{N}$, the d -dimensional hypercube Q^d is the graph with

- vertex set

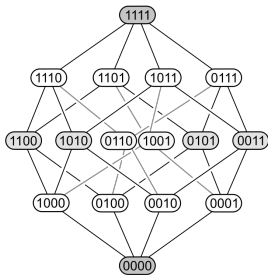
$$V(Q^d) = \{0, 1\}^d = \{x = (x_1, \dots, x_d) : x_i \in \{0, 1\}, 1 \leq i \leq d\}$$

- edge set $E(Q^d) : \forall v = (v_1, \dots, v_d), w = (w_1, \dots, w_d) \in V(Q^d),$
 $\{v, w\} \in E(Q^d)$ iff v and w differ in exactly one coordinate

Obvious facts:

- $|V(Q^d)| = 2^d$
- Q^d is d -regular
- Q^d is bipartite
- diameter of Q^d is d

...



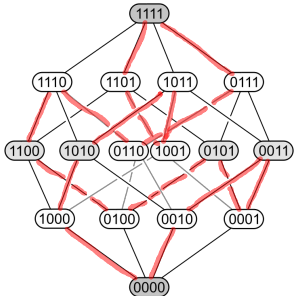
Hasse diagram

A random subgraph of the hypercube

Given $p \in (0, 1)$

Q_p^d = a graph obtained by retaining each edge of Q^d independently with probability p

= bond percolation on Q^d with edge probability p



Typical properties of Q_p^d around $p = \frac{1}{2}$

Connectivity

[SAPOŽENKO 67; BURTIN 67; ERDŐS–SPENCER 79; BOLLOBÁS 83]

$p = \frac{1}{2}$ is a sharp threshold for connectedness: $\forall \varepsilon > 0$

$$\mathbb{P} \left[Q_p^d \text{ is connected} \right] \xrightarrow{d \rightarrow \infty} \begin{cases} 0 & \text{if } p < \frac{1-\varepsilon}{2} \\ 1 & \text{if } p > \frac{1+\varepsilon}{2} \end{cases}$$

Typical properties of Q_p^d around $p = \frac{1}{2}$

Connectivity

[SAPOŽENKO 67; BURTIN 67; ERDŐS–SPENCER 79; BOLLOBÁS 83]

$p = \frac{1}{2}$ is a sharp threshold for connectedness: $\forall \varepsilon > 0$

$$\mathbb{P} \left[Q_p^d \text{ is connected} \right] \xrightarrow{d \rightarrow \infty} \begin{cases} 0 & \text{if } p < \frac{1-\varepsilon}{2} \\ 1 & \text{if } p > \frac{1+\varepsilon}{2} \end{cases}$$

Perfect matching

[BOLLOBÁS 90]

$p = \frac{1}{2}$ is a sharp threshold for the existence of a perfect matching

Is $p = \frac{1}{2}$ a sharp threshold for Hamiltonicity?

[BOLLOBÁS 80's; FRIEZE 14]

Hamiltonicity

[CONDON-ESPUNY-DÍAZ-GIRÃO-KÜHN-OSTHUS 21]

$p = \frac{1}{2}$ is a sharp threshold the existence of a Hamiltonian cycle.

Emergence of the giant component in Q_p^d

Does the component structure of Q_p^d undergo a phase transition at $p = \frac{1}{d}$?

[ERDŐS-SPENCER 79]

Giant component

[AJTAI-KOMLÓS- SZEMERÉDI 81]

$p = \frac{1}{d}$ is a sharp threshold: $\forall \varepsilon > 0$

- whp all components are of order $O(d)$ if $p < \frac{1-\varepsilon}{d}$
- whp \exists a unique largest component of order $\Theta(2^d)$ if $p > \frac{1+\varepsilon}{d}$

whp = with high probability = with prob tending to one as $d \rightarrow \infty$

Supercritical regime – open questions

$$p = \frac{1+\varepsilon}{d} \text{ for fixed } \varepsilon > 0$$

$L_1 =$ the largest component of Q_p^d

- diameter of L_1 ?
- circumference of L_1 (= length of the longest cycles) ?
- mixing time of lazy simple random walk on L_1 ?

[BOLLOBÁS-KOHAYAKAWA-ŁUCZAK 92]

[PETE 08; VAN DER HOFSTAD-NACHMIAS 17]

Part II.

Expansion properties and consequences

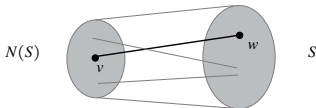
Expanders

[ALON 86; HOORY-LINIAL-WIGDERSON 06; KRIVELEVICH 19; KRIVELEVICH-SUDAKOV 09; SARNAK 04; . . .]

Given a graph G

- $N(S) =$ external neighbourhood of a subset $S \subseteq V(G)$
 $= \{v \in V(G) \setminus S : \exists w \in S \text{ with } \{v, w\} \in E(G)\}$
- G is an α -expander if

$$|N(S)| \geq \alpha|S|, \quad \forall S \subseteq V(G) \text{ with } |S| \leq \frac{|V(G)|}{2}$$



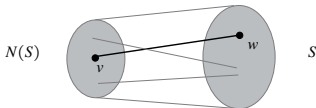
Expanders

[ALON 86; HOORY-LINIAL-WIGDERSON 06; KRIVELEVICH 19; KRIVELEVICH-SUDAKOV 09; SARNAK 04; . . .]

Given a graph G

- $N(S)$ = external neighbourhood of a subset $S \subseteq V(G)$
= $\{v \in V(G) \setminus S : \exists w \in S \text{ with } \{v, w\} \in E(G)\}$
- G is an α -expander if

$$|N(S)| \geq \alpha|S|, \quad \forall S \subseteq V(G) \text{ with } |S| \leq \frac{|V(G)|}{2}$$



Properties of an expander

- small diameter, long cycles, . . .
- edge-expansion for graphs with bounded max degree

Expansion properties and consequences

$L_1 =$ largest component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem

[ERDE-K.-KRIVELEVICH 21+]

whp L_1

- is $c d^{-5}$ -expander

Expansion properties and consequences

$L_1 =$ largest component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem

[ERDE-K.-KRIVELEVICH 21+]

whp L_1

- is $c d^{-5}$ -expander
- contains a $c' d^{-2} (\log d)^{-1}$ -expander on $\geq 0.99 |L_1|$ vertices

Expansion properties and consequences

$L_1 =$ largest component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem

[ERDE-K.-KRIVELEVICH 21+]

whp L_1

- is $c d^{-5}$ -expander
- contains a $c' d^{-2} (\log d)^{-1}$ -expander on $\geq 0.99 |L_1|$ vertices
- has diameter $O(d^3)$
- contains a cycle of length $\Omega(2^d d^{-2} (\log d)^{-1})$

Expansion properties and consequences

$L_1 =$ largest component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem

[ERDE-K.-KRIVELEVICH 21+]

whp L_1

- is $c d^{-5}$ -expander
- contains a $c' d^{-2} (\log d)^{-1}$ -expander on $\geq 0.99 |L_1|$ vertices
- has diameter $O(d^3)$
- contains a cycle of length $\Omega(2^d d^{-2} (\log d)^{-1})$
- has Cheeger constant $\Omega(d^{-5})$

Part III.

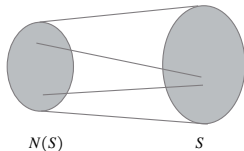
Proof ideas

Theorem

whp L_1 is a $\frac{1}{\text{poly}(d)}$ -expander

i.e., $\forall S \subseteq V(L_1)$ with $|S| \leq \frac{|V(L_1)|}{2}$,

$$|N(S)| \geq \frac{|S|}{\text{poly}(d)}$$



Sprinkling argument

Sprinkling

$$p = \frac{1+\varepsilon}{d} \text{ for } \varepsilon > 0$$

$$q_1 = \frac{1+\delta_1}{d} \text{ and } q_2 = \frac{\delta_2}{d} \text{ s.t. } 1-p = (1-q_1)(1-q_2) \text{ and } 0 < \delta_2 \ll \delta_1$$

$$\mathcal{Q}_p^d \sim \mathcal{Q}_{q_1}^d \cup \mathcal{Q}_{q_2}^d$$

Sprinkling argument

Sprinkling

$$p = \frac{1+\epsilon}{d} \text{ for } \epsilon > 0$$

$$q_1 = \frac{1+\delta_1}{d} \text{ and } q_2 = \frac{\delta_2}{d} \text{ s.t. } 1 - p = (1 - q_1)(1 - q_2) \text{ and } 0 < \delta_2 \ll \delta_1$$

$$Q_p^d \sim Q_{q_1}^d \cup Q_{q_2}^d$$

Largest components before and after sprinkling

L'_1 = largest component in $Q_{q_1}^d$ (before sprinkling)

L_1 = largest component in Q_p^d (after sprinkling)

$\gamma(x)$ = survival probability of $\text{Po}(1+x)$ branching process

Lemma

[AJTAI-KOHLÓŠ- SZEMERÉDI 81]

- whp $L'_1 \sim \gamma(\delta_1) 2^d$
- whp $L_1 \sim \gamma(\epsilon) 2^d$

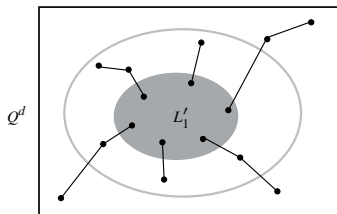
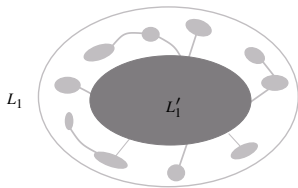
Giant component before and after sprinkling

L'_1 = largest component in $Q_{q_1}^d$ (before sprinkling)

L_1 = largest component in Q_p^d (after sprinkling)

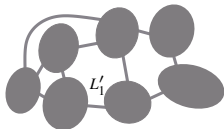
Lemma

- whp \forall connected component in $Q_p^d [L_1 - L'_1]$ is of order $O(d)$
- whp \forall vertex in $V(Q^d)$ is within distance two from $\geq cd^2$ vertices in L'_1



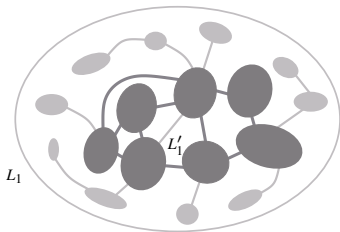
Splitting the largest component into pieces

- L'_1 = largest component (before sprinkling)
= split into a family \mathcal{C} of vertex-disjoint connected subgraphs ('pieces'), each of order $\text{poly}(d)$



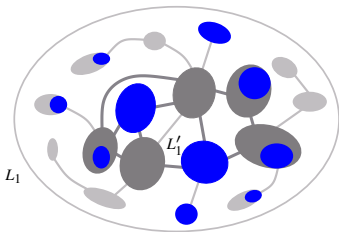
Splitting the largest component into pieces

- L'_1 = largest component (before sprinkling)
= split into a family \mathcal{C} of vertex-disjoint connected subgraphs ('pieces'), each of order $\text{poly}(d)$
- L_1 = largest component (after sprinkling)



Splitting the largest component into pieces

- L'_1 = largest component (before sprinkling)
= split into a family \mathcal{C} of vertex-disjoint connected subgraphs ('pieces'), each of order $\text{poly}(d)$
- L_1 = largest component (after sprinkling)
- S = arbitrary subset of $V(L_1)$ with $|S| \leq \frac{|V(L_1)|}{2}$

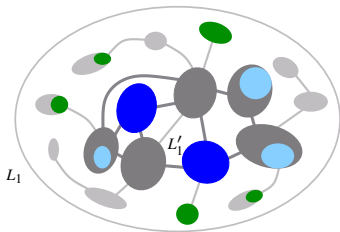


Splitting the largest component into pieces

L'_1 = largest component (before sprinkling)
= split into a family \mathcal{C} of vertex-disjoint connected subgraphs ('pieces'), each of order $\text{poly}(d)$

L_1 = largest component (after sprinkling)

S = arbitrary subset of $V(L_1)$ with $|S| \leq \frac{|V(L_1)|}{2}$



$S_1 = S - L'_1$

$S_2 =$ vertices in pieces $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

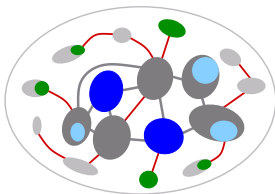
$S_3 =$ vertices in pieces $C \in \mathcal{C}$ with $C \subseteq S$

Contribution of S_1 to $N(S)$

$$S_1 = S - L'_1$$

$S_2 =$ vertices in pieces $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

$S_3 =$ vertices in pieces $C \in \mathcal{C}$ with $C \subseteq S$



With sprinkling, each component in $Q_p^d[L_1 - L'_1]$ which intersects with S_1

- contributes **at least one edge** to $N(S)$
- or is **connected to** $S_2 \cup S_3$

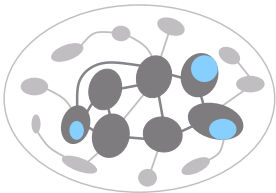
Thus $|N(S)| \geq \frac{c|S_1|}{d}$ or $e(S_1, S_2 \cup S_3) \geq \frac{c|S_1|}{d}$ and thus $|S_2 \cup S_3| \geq \frac{c|S_1|}{d^2}$

Contribution of S_2 to $N(S)$

L'_1 = split into a family \mathcal{C} of pieces, each of order $\text{poly}(d)$

S_2 = vertices in pieces $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

S_3 = vertices in pieces $C \in \mathcal{C}$ with $C \subseteq S$



Each piece $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

● contributes at least one edge to $N(S)$

and each piece is of order $\text{poly}(d)$

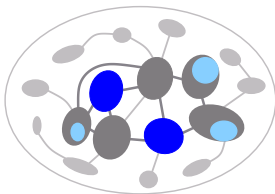
Thus $|N(S)| \geq \frac{|S_2|}{\text{poly}(d)}$

Contribution of S_3 to $N(S)$

L'_1 = split into a family \mathcal{C} of pieces, each of order $\text{poly}(d)$

S_2 = vertices in pieces $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

S_3 = vertices in pieces $C \in \mathcal{C}$ with $C \subseteq S$



(1) Partition the family \mathcal{C} of pieces into two disjoint families $\{\mathcal{A}, \mathcal{B}\}$

$$\mathcal{A} := \{C \in \mathcal{C} : C \subseteq S\} \quad \text{and} \quad \mathcal{B} := \mathcal{C} - \mathcal{A}$$

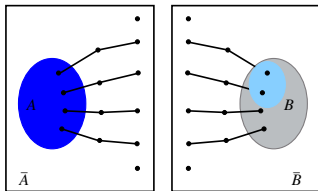
This partitions $V(L'_1)$ into two sets A, B where

$$A := V(\mathcal{A}) = S_3 \quad \text{and} \quad B := V(\mathcal{C} - \mathcal{A})$$

Contribution of S_3 to $N(S)$ – extending and connecting

(2) Extending the partition $V(L'_1) = A \dot{\cup} B$ to a partition $V(Q^d) = \bar{A} \dot{\cup} \bar{B}$ s.t.

- every vertex in \bar{A} is within distance 2 of A
- every vertex in \bar{B} is within distance 2 of B

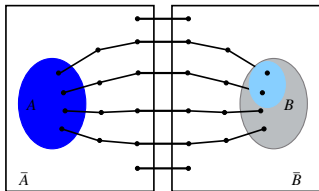


whp every vertex in $V(Q^d)$ is within distance two from vertices in L'_1

Contribution of S_3 to $N(S)$ – extending and connecting

(2) Extending the partition $V(L'_1) = A \dot{\cup} B$ to a partition $V(Q^d) = \bar{A} \dot{\cup} \bar{B}$ s.t.

- every vertex in \bar{A} is within distance 2 of A
- every vertex in \bar{B} is within distance 2 of B



Edge-isoperimetry in Q^d

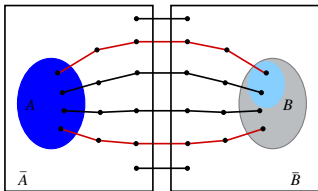
[HARPER 64; LINDSEY 64; BERNSTEIN 67; HART 76]

$$|E(X, X^c)| \geq |X| (d - \log_2 |X|), \quad \forall X \subseteq V(Q^d) \quad \text{with} \quad |X| \leq 2^{d-1}$$

Contribution of S_3 to $N(S)$ – extending and connecting

(2) Extending the partition $V(L'_1) = A \dot{\cup} B$ to a partition $V(Q^d) = \bar{A} \dot{\cup} \bar{B}$ s.t.

- every vertex in \bar{A} is within distance 2 of A
- every vertex in \bar{B} is within distance 2 of B



(3) Sprinkle with $q_2 = \frac{\delta_2}{d}$

Lemma

whp \exists at least $\frac{|A|}{\text{poly}(d)}$ vertex-disjoint A - B -paths of length at most 5 in $Q_{q_2}^d$

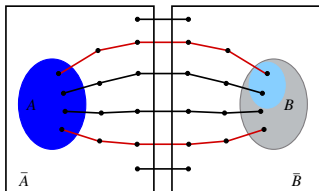
Contribution of S_3 to $N(S)$

L'_1 = split into a family \mathcal{C} of 'pieces', each of order $\text{poly}(d)$

S_2 = vertices in pieces $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

S_3 = vertices in pieces $C \in \mathcal{C}$ with $C \subseteq S$

= A



Each A - B -path in $\mathcal{Q}_{q_2}^d$ contributes at least one edge to $N(S)$, unless it goes to S_2

Thus $|N(S)| \geq \frac{|S_3|}{\text{poly}(d)} - d|S_2|$

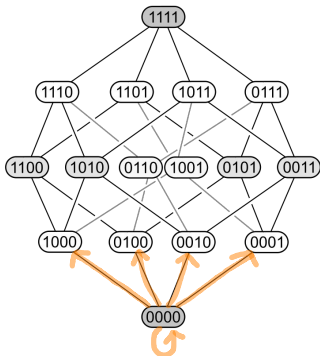
Part IV.

Mixing time of lazy random walk

Mixing time of lazy random walk on Q^d

In each step,

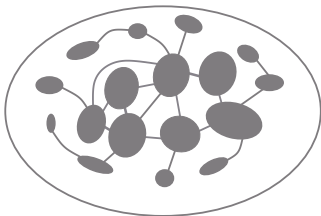
- it remains at the current position with prob $\frac{1}{2}$
- it moves to a uniformly chosen random neighbour with prob $\frac{1}{2}$



Mixing time: $O(d \log d)$

Mixing time of lazy random walk on giant comp of Q_p^d

$L_1 =$ giant component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

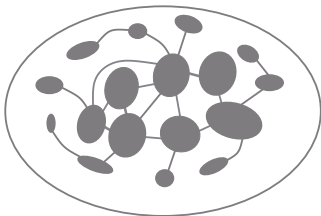


Mixing time of lazy random walk on giant comp of Q_p^d

$L_1 =$ giant component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

What is the mixing time of the lazy random walk on L_1 ?

[PETE 08; VAN DER HOFSTAD-NACHMIAS 17]

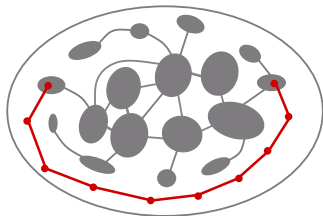


Mixing time of lazy random walk on giant comp of Q_p^d

$L_1 =$ giant component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

What is the mixing time of the lazy random walk on L_1 ?

[PETE 08; VAN DER HOFSTAD-NACHMIAS 17]



whp L_1 contains **bare paths** of length $\Omega(d)$

\implies mixing time: $\Omega(d^2)$

Mixing time of lazy random walk

Given a graph G ,

$t_{\text{mix}}(G)$ = mixing time of a lazy random walk on a graph G

$\Phi(G)$ = Cheeger constant of G (= bottleneck ratio)

$\pi_{\min}(G) = \min\left\{\frac{d_G(x)}{2|E(G)|} : x \in V(G)\right\}$

[LAWLER-SOKAL 88; JERRUM-SINCLAIR 89; LEVIN-PERES-WILMER 07]

$$t_{\text{mix}}(G) \leq \frac{2}{\Phi(G)^2} \log\left(\frac{4}{\pi_{\min}(G)}\right)$$

Mixing time of lazy random walk on giant comp of Q_p^d

Given a graph G ,

$t_{\text{mix}}(G)$ = mixing time of a lazy random walk on a graph G

$\Phi(G)$ = Cheeger constant of G (= bottleneck ratio)

$\pi_{\text{min}}(G) = \min\left\{\frac{d_G(x)}{2|E(G)|} : x \in V(G)\right\}$

[LAWLER-SOKAL 88; JERRUM-SINCLAIR 89; LEVIN-PERES-WILMER 07]

$$t_{\text{mix}}(G) \leq \frac{2}{\Phi(G)^2} \log\left(\frac{4}{\pi_{\text{min}}(G)}\right)$$

L_1 = giant component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

[ERDE-K.-KRIVELEVICH 21+]

whp

$$\Phi(L_1) = \Omega\left(d^{-5}\right) \quad \text{and} \quad \pi_{\text{min}}(L_1) = \Omega\left(2^{-d}\right)$$
$$t_{\text{mix}}(L_1) = O\left(d^{11}\right)$$

Summary

$L_1 =$ largest component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem

[ERDE-K.-KRIVELEVICH 21+]

whp L_1

- is $c d^{-5}$ -expander
- contains a $c' d^{-2} (\log d)^{-1}$ -expander on $\geq 0.99 |L_1|$ vertices
- has diameter $O(d^3)$
- contains a cycle of length $\Omega(2^d d^{-2} (\log d)^{-1})$
- has Cheeger constant $\Omega(d^{-5})$

whp the mixing time of the lazy simple random walk on L_1 is $O(d^{11})$.

Open problems

$L_1 =$ largest component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Correct order of

- diameter of L_1 : $\Theta(d^2)$?
- circumference of L_1 : $\Omega(2^d)$?
- mixing time of lazy random walk on L_1 ?