Hypercube percolation

Remco van der Hofstad

TU/e technische universiteit eindhoven

EURANDOM

Networks: Models for Populations and Epidemics, 12-16 September 2011, ICMS, Edinburgh

Joint work with:
- Asaf Nachmias (MIT/UBC)

- Markus Heydenreich (Leiden/CWI)
- Gordon Slade (UBC Vancouver)
- Christian Borgs & Jennifer Chayes (Microsoft Research)
- Joel Spencer (New York University)

Percolation on hypercube

Study random subgraph of hypercube $\mathbb{Q}_n = \{0,1\}^n$, with bonds

$$\{\{x,y\}\colon \exists!\ i \text{ with } x_i \neq y_i\}.$$

Make bonds $\{x,y\}$ independently

occupied	with probability p,
vacant	with probability $1-p$,

where $p \in [0,1]$ is percolation parameter.

Goal: Study percolation phase transition as $n \to \infty$.

Erdős and Spencer (1979): for $p = 1/2 + \varepsilon/2n$, random graph connected with probability $\mathrm{e}^{-\mathrm{e}^{-\varepsilon}(1+o(1))}$.

Inspiration: Erdős-Rényi random graph

Erdős-Rényi random graph is random subgraph of complete graph on V vertices where each of $\binom{V}{2}$ edges is occupied with probab. p.

Phase transition: (Erdős and Rényi (60))
For $p = (1+\varepsilon)/V$, largest connected component $|\mathcal{C}_{\max}|$ is
(a) $\Theta_{\mathbb{P}}(\log V)$ for $\varepsilon < 0$;
(b) $\Theta_{\mathbb{P}}(V)$ for $\varepsilon > 0$;

Scaling window: (Bollobás (84) and Łuczak (90))
For $p = (1/V)(1+\lambda/V^{1/3})$, largest component is $\Theta_{\mathbb{P}}(V^{2/3})$, with expected cluster size $\Theta(V^{1/3})$.

BCHSS05a, 05b: with $\chi(p) = \mathbb{E}_p|\mathcal{C}(v)|$ expected cluster size on hypercube, define critical threshold $p_c(\mathbb{Q}_n)$ by

$$\chi(p) = 2^{n/3} = V^{1/3}.$$

(Sub-)Critical results

Theorem 1 (Subcritical clusters) (BCHSS (05a), (05b)).
For $p = p_c(1+\varepsilon)$, and uniformly in $\varepsilon \leq 0$, as $n \to \infty$,

$$\chi(p) = \frac{O(1)}{|\varepsilon| + 2^{-n/3}},$$

$$\mathbb{P}_p\Big(\chi^2(p) \leq |\mathcal{C}_{\max}| \leq 2\chi^2(p)\log\big(2^n/\chi^3(p)\big)\Big) \geq 1 - \log\big(2^n/\chi^3(p)\big)^{-3/2}.$$

Theorem 2 (Scaling window) (BCHSS (05a), (05b)).
For $p = p_c(1+\varepsilon)$, with $|\varepsilon| \leq \Lambda 2^{-n/3}$, there exists $b_1 = b_1(\Lambda) > 0$ s.t.

$$\mathbb{P}_p\Big(\omega^{-1}2^{2n/3} \leq |\mathcal{C}_{\max}| \leq \omega 2^{2n/3}\Big) \geq 1 - \frac{b_1}{\omega}.$$

vdHHe (11): $|\mathcal{C}_{\max}|2^{-2n/3}$ is tight and non-degenerate:
hallmark of critical behavior.

Supercritical results

Theorem 3 (Supercritical clusters) (vdH+Nachmias (11)).
For $p = p_c(1+\varepsilon)$, with $\varepsilon \gg 2^{-n/3}$,

$$\frac{|\mathcal{C}_{\max}|}{2\varepsilon 2^n} \xrightarrow{\mathbb{P}} 1,$$

while, with $\mathcal{C}_{(2)}$ denoting second largest connected component,

$$\frac{|\mathcal{C}_{(2)}|}{\varepsilon 2^n} \xrightarrow{\mathbb{P}} 0,$$

$$\chi(p) = 4\varepsilon^2 2^n (1 + o(1)).$$

Percolation phase transitions hypercube and complete graph alike:
2ε is asymptotic survival probability of branching process with Poisson($1+\varepsilon$)-offspring distribution.

Previous work hypercube

Ajtai, Komlós and Szemerédi (82): For $p = (1+\varepsilon)/n$,

$$|\mathcal{C}_{\max}| = \Theta_{\mathbb{P}}(\varepsilon 2^n) \quad \text{for } \varepsilon > 0, \qquad |\mathcal{C}_{\max}| = o(2^n) \quad \text{whp} \quad \text{for } \varepsilon < 0.$$

Bollobás, Kohayakawa and Łuczak (92):

$$|\mathcal{C}_{\max}| = (2\log 2)\frac{n}{\varepsilon^2}\big(1+o(1)\big) \quad \text{whp} \quad \text{for } \varepsilon \leq -(\log n)^2/(\sqrt{n}\log\log n),$$

$$|\mathcal{C}_{\max}| = 2\varepsilon 2^n\big(1+o(1)\big) \quad \text{whp} \quad \text{for } \varepsilon \geq 60(\log n)^3/n.$$

Transition is extremely sharp, critical value close to $1/n$.

Question BKL: Is critical value equal to $1/(n-1)$?

Hierarchy of phase transitions on hypercube

Theorem 4 (Asymptotic expansion p_c). (BCHSS, vdHS05,06, vdHN11) There exist rational numbers a_i with $a_1 = a_2 = 1, a_3 = 7/2,$ s.t., for every $s \geq 1$, if

$$p = \sum_{i=1}^{s} a_i n^{-i} + \delta n^{-s},$$

with $\delta < 0$, then, as $n \to \infty$,

$$|\mathcal{C}_{\max}| \leq (2 \log 2) n^{2s-1} |\delta|^{-2} [1 + o_{\mathbb{P}}(1)] \qquad \text{a.s.,}$$

while for $\delta > 0$, as $n \to \infty$,

$$|\mathcal{C}_{\max}| = 2\delta^{-1} n^{-(s-1)} 2^n (1 + o_{\mathbb{P}}(1)).$$

Extension of AKS to all powers $1/n$, answers question BKL negatively.

Proof AKS: sprinkling

Fix small $\delta > 0$. Write $p_1 = (1+(1-\delta)\varepsilon)/n$ and $p_2 \approx \delta\varepsilon/n$ s.t.

$$(1-p_1)(1-p_2) = 1-p.$$

G_p is union of edges in independent copies G_{p_1} and G_{p_2}.

Step 1: prove weak bound: in G_{p_1} positive fraction vertices has large cluster of size $\geq 2^{c_1 n}$, where $c_1 > 0$ fixed.
Uses: branching process approximation and Azuma-Hoeffding.

Step 2: add sprinkled edges of G_{p_2} and show they connect many large clusters into giant cluster of size $\Theta(2^n)$.

Key tool: isoperimetric inequality for hypercube:
Two disjoint sets of size of order 2^n are connected by at least $2^n/n^{100}$ disjoint paths of length $O(\sqrt{n})$.

Proof AKS: sprinkling (Cont.)

Let V' be vertices whose component has size at least $2^{c_1 n}$ in G_{p_1}, so that $|V'| = \Theta(2^n)$.

If largest connected component in $G_{p_1} \cup G_{p_2}$ has size $\leq \delta 2^n$, then we can partition

$$V' = A \uplus B$$

s.t. $|A|, |B| = \Theta(2^n)$, and *any* path of length at most $\sqrt{n}$ between them contains p_2-closed edge.

Number of such partitions is at most $2^{2^n/2^{c_1 n}}$, probability that path of length k contains p_2-closed edge is $1 - p_2^k$. By isoperimetric inequality

$$2^{c2^n/2^{c_1 n}} \cdot \left(1 - \left(\frac{c\delta\varepsilon}{n}\right)^{\sqrt{n}}\right)^{c2^n/n^{100}} = \mathrm{e}^{-c2^{(1+o(1))n}}.$$

Wasteful: Worst-case, clusters more like random sets.

Proof 1: supercritical cluster tails

Theorem 5 (Supercritical cluster tails). Let $p = p_c(1+\varepsilon)$. Then, for $k_0 = \varepsilon^{-2}(\varepsilon^3 V)^\alpha$ for any $\alpha \in (0, 1/3)$,

$$\mathbb{P}_p(|\mathcal{C}(0)| \geq k_0) = 2\varepsilon(1 + o(1)).$$

Further, with

$$Z_{\geq k} = \Big|\Big\{v \colon |\mathcal{C}(v)| \geq k\Big\}\Big|,$$

denoting number of vertices in large clusters,

$$\frac{Z_{\geq k_0}}{2\varepsilon 2^n} \xrightarrow{\mathbb{P}} 1.$$

Proof: Differential inequalities (as in BA91, BCHSS05a), with careful analysis of constants involved.

Proof 2: using non-backtracking random walk

Let $\mathbb{P}_p(0 \overset{[a,b]}{\longleftrightarrow} x)$ denote probability that 0 is connected to x with shortest path of length $\in [a,b]$.

Take $r \gg 1/\varepsilon$.

We show that $\mathbb{P}_p(0 \overset{[r,2r]}{\longleftrightarrow} x)$ is almost constant in x for large r by comparing percolation paths to non-backtracking random walk, which is random walk conditioned not to reverse immediately.

Then,

$$\mathbb{P}_p(0 \overset{[r,2r]}{\longleftrightarrow} x) \approx \frac{1}{V} \sum_x \mathbb{P}_p(0 \overset{[r,2r]}{\longleftrightarrow} x) \le \mathbb{E}_p|B_0(r)|/V.$$

This makes performing complicated sums relatively easy.

Proof 3: large clusters share large boundary

For $x, y \in \mathbb{Q}_n$ and ℓ, let

$$S_\ell(x,y) = \left|\left\{(u,u') \in E \colon x \overset{\ell}{\longleftrightarrow} u \,, y \overset{\ell}{\longleftrightarrow} u'\right\}\right| .$$

Pair x, y is good when $|\mathcal{C}(x)|, |\mathcal{C}(y)| \geq K\varepsilon^{-2}$, and

$$S_{2r+r_0}(x,y) \geq n2^{-n}\varepsilon^{-2}(\varepsilon 2^n)/(\log(\varepsilon^3 2^n))^2,$$

Write $P_{r,r_0,K}$ for number of (r, r_0, K)-good pairs.

Theorem 6 (Most large clusters share many boundary edges).
Assume $\varepsilon^3 2^n \to \infty, \varepsilon \leq n^{-2}$. Take $M = \log\log\log(\varepsilon^3 2^n)$, $r = M/\varepsilon$, and $r_0 = \varepsilon^{-1}\big[\log(\varepsilon^3 2^n) - \log\log(\varepsilon^3 2^n)\big]/2$. Then, there exists $K \to \infty$ s.t.

$$\frac{P_{r,r_0,K}}{(2\varepsilon V)^2} \xrightarrow{\mathbb{P}} 1 \,.$$

Proof 4: improved sprinkling

Take $p_+ = \theta\varepsilon/n$ for our sprinkling probability, where $\theta > 0$ is small. Let p_1 satisfy

$$p = p_1 + (1 - p_1)p_+ \,,$$

Given G_{p_1}, construct auxiliary simple graph $H = (\mathcal{V}, \mathcal{E})$ with

$$\mathcal{V} = \left\{x \in G_{p_1} : |\mathcal{C}(x)| \geq K\varepsilon^{-2}\right\}, \;\; \mathcal{E} = \left\{(x, y) \in \mathcal{V}^2 \; (\gamma, r, r_0, K)\text{-}\mathbf{good}\right\},$$

$$\mathbb{P}_{p_1}\left(P_{r,r_0,K} \geq (1-\alpha)4\varepsilon^2 2^{2n}, |\mathcal{V}| \in [(2-\theta)\varepsilon, (2+\theta)\varepsilon]\right) \geq 1 - \delta.$$

Claim: Probability there exists partition $\mathcal{V} = M_1 \uplus M_2$ with $|M_1| \geq \theta v$ and $|M_2| \geq \theta v$ s.t. sprinkled edges do not connect M_1 to M_2 is small.

Step 1: Number of partitions is at most $2^{3K^{-1}\varepsilon^3 V}$.

Step 2: Given such partition, number of edges (u, u') between $\mathcal{C}(M_1)$ and $\mathcal{C}(M_2)$ is at least $c\varepsilon^2 n 2^n$.

Open problems and extension

(a) Study second largest component in supercritical phase on hypercube, and prove discrete duality principle.

(b) CLT for giant component.

(c) Investigate whether scaling limit of largest clusters is same as for ERRG in Aldous (97).

(d) Extension: Hamming graph= product of d complete graphs.

Results apply for $p = p_c(1 + \varepsilon)$, where $\varepsilon^3 V \to \infty$ with $\varepsilon \ll 1/\log V$, where volume is $V = n^d$.
$d = 2:$ Together with BCHSS05a, BCHSS05b, vdH+Luczak (10), this completes barely supercritical regime.

Literature percolation hypercube

[1-3] Borgs, Chayes, van der Hofstad, Slade, and Spencer.
Random subgraphs of finite graphs. I. The scaling window under the triangle condition. *RSA*, **27**(2):137–184, (2005). II. The lace expansion and the triangle condition. *AoP*, **33**(5):1886–1944, (2005). III. The phase transition for the n-cube. *Combinatorica*, **26**(4):395–410, (2006).

[4-5] van der Hofstad and Slade. Asymptotic expansion in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$. *RSA* **27**: 331-357, (2005).
Expansion in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$: the first three terms. *CPC* **15**(5): 695–713, (2006).

[6-7] Heydenreich and van der Hofstad.
Random graph asymptotics on high-dimensional tori. *CMP*, **270**(2):335–358, (2007). II. Volume, diameter and mixing time. *PTRF* **149**:397–415, (2011).

[8] van der Hofstad and Nachmias. Hypercube percolation. In preparation.