Probability thresholds estimates for coloring properties of random hypergraphs

Alexander Semenov

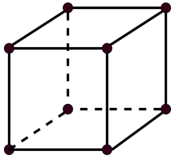
joint work with D. Shabanov

Probabilistic Combinatorics Online

25th September 2020

Independent set

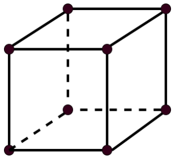
This is a 4-uniform hypergraph H on 8 vertices with 6 edges (surfaces of the cube):



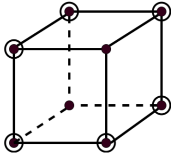
What is the maximum number of vertices we can take such that there's no edge from which we took all 4 of them?

Independent set

This is a 4-uniform hypergraph H on 8 vertices with 6 edges (surfaces of the cube):



What is the maximum number of vertices we can take such that there's no edge from which we took all 4 of them?



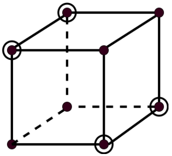
The answer is 6, so is the independence number of this hypergraph.

j-independent set

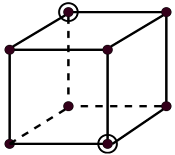
For the same hypergraph what is the maximum number of vertices we can take such that there's no edge from which we took more than 2 vertices? More than 1?

j-independent set

For the same hypergraph what is the maximum number of vertices we can take such that there's no edge from which we took more than 2 vertices? More than 1?



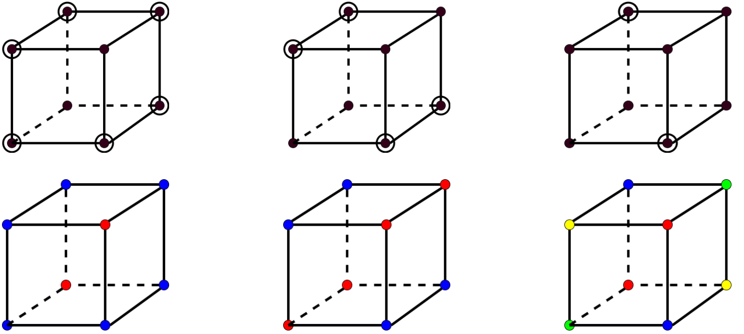
2-independence number of this hypergraph is 4.



1-independence number of this hypergraph is 2.

j-chromatic numbers

How many colors are needed to color vertices of hypergraph such that every color is j-independent set?



$\chi_3(H) = \chi(H) = 2$ $\chi_2(H) = 2$ $\chi_1(H) = 4$

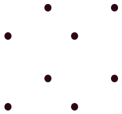
weak *strong*

Definitions

- For an integer j, a j-**independent set** in a hypergraph $H(V, E)$ is a subset $A \subset V$ such that for every edge $e \in E : |e \cap A| \leqslant j$;
- A coloring of hypergraph $H = (V, E)$ is called j-**proper** if every color class is j-independent;
- The j-**chromatic** number $\chi_j(H)$ of hypergraph $H = (V, E)$ is the minimal number of colors in a j-proper coloring of H;
- In the case of k-uniform hypergraph $\chi_{k-1}(H) = \chi(H)$;
- All chromatic numbers $\chi_j(H)$ with $j \geqslant k/2$ are called **weak**.

Binomial model

Let's consider binomial model $H(n, k, p)$ of a random hypergraph.



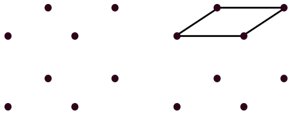
$\chi_3(H) = 1$

$\chi_2(H) = 1$

$\chi_1(H) = 1$

Binomial model

Let's consider binomial model $H(n, k, p)$ of a random hypergraph.



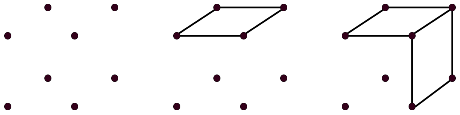
$\chi_3(H) = 1 \quad \chi_3(H) = 2$

$\chi_2(H) = 1 \quad \chi_2(H) = 2$

$\chi_1(H) = 1 \quad \chi_1(H) = 4$

Binomial model

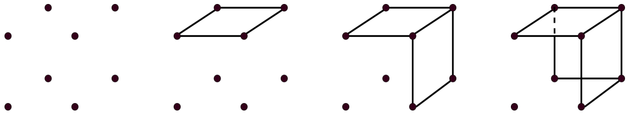
Let's consider binomial model $H(n, k, p)$ of a random hypergraph.



$\chi_3(H) = 1$	$\chi_3(H) = 2$	$\chi_3(H) = 2$
$\chi_2(H) = 1$	$\chi_2(H) = 2$	$\chi_2(H) = 2$
$\chi_1(H) = 1$	$\chi_1(H) = 4$	$\chi_1(H) = 4$

Binomial model

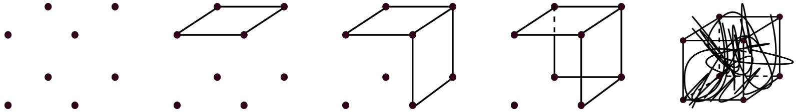
Let's consider binomial model $H(n, k, p)$ of a random hypergraph.



$\chi_3(H) = 1$	$\chi_3(H) = 2$	$\chi_3(H) = 2$	$\chi_3(H) = 2$
$\chi_2(H) = 1$	$\chi_2(H) = 2$	$\chi_2(H) = 2$	$\chi_2(H) = 2$
$\chi_1(H) = 1$	$\chi_1(H) = 4$	$\chi_1(H) = 4$	$\chi_1(H) = 4$

Binomial model

Let's consider binomial model $H(n, k, p)$ of a random hypergraph.



$\chi_3(H) = 1$	$\chi_3(H) = 2$	$\chi_3(H) = 2$	$\chi_3(H) = 2$	$\chi_3(H) = 3$
$\chi_2(H) = 1$	$\chi_2(H) = 2$	$\chi_2(H) = 2$	$\chi_2(H) = 2$	$\chi_2(H) = 4$
$\chi_1(H) = 1$	$\chi_1(H) = 4$	$\chi_1(H) = 4$	$\chi_1(H) = 4$	$\chi_1(H) = 8$

$$= \left\lceil \frac{n}{j} \right\rceil$$

Constant probability

Let's consider j and k are fixed while n tends to infinity. Also let's assume $j < k - 1$ and $p \in (0, 1)$ is some constant.

- Let X be the number of subsets of $j + 1$ vertices that are not contained in any edge of hypergraph $H(n, k, p)$.
- Probability that a set of j + 1 vertices is not contained in any edge of hypergraph $H(n, k, p)$ equals

$$(1 - p)^{\binom{n-j-1}{k-j-1}}.$$

- Therefore,

$$\mathrm{E}X = \binom{n}{j+1}(1 - p)^{\binom{n-j-1}{k-j-1}} \to 0.$$

- Then $\chi_j(H(n, k, p)) = \lceil \frac{n}{j} \rceil$ asymptotically almost surely (a.a.s.).

Dense case

Let's take $1 \leqslant j \leqslant k-1$. Denote

$$d^{(j)} = d^{(j)}(n,p) = j\binom{k-1}{j}\binom{n-1}{k-1}p.$$

- **Shamir, 1989**: $d^{(j)} > n^{j-\varepsilon}, \varepsilon > 0$ and $d^{(j)} = o(n^j \ln n)$;
- **Krivelevich, Sudakov, 1998**: $d^{(j)} = o(n^j)$ and $d^{(j)} \to \infty$;

$$\chi_j(H(n,k,p)) \underset{\mathrm{P}}{\sim} \left(\frac{d^{(j)}}{(j+1)\ln d^{(j)}}\right)^{1/j}, \quad n \to \infty.$$

Sparse case, arbitrary r, $j = k - 1$

Let's now consider the case $p = cn/\binom{n}{k}$.

- **Theorem (Dyer, Frieze, Greenhill, 2015)**:
  If $c > r^{k-1}\ln r - \frac{\ln r}{2}$ then a.a.s. $\chi(H(n,k,p)) > r$.
- **Theorem (Dyer, Frieze, Greenhill, 2015)**:
  If $c < r^{k-1}\ln r - \frac{r-1}{r}(1+\ln r) + O(\frac{k^2\ln r}{r^{k-1}})$ then a.a.s. $\chi(H(n,k,p)) \leqslant r$.
- **Theorem (Ayre, Coja-Oghlan, Greenhill, 2015)**:
  For $r > r_0(k)$, if $c < r^{k-1}\ln r - \frac{\ln r}{2} - \ln 2 + o_r(1)$ then a.a.s. $\chi(H(n,k,p)) \leqslant r$.
- **Theorem (Shabanov, 2017)**:
  For $k \geqslant 4$, if $c < r^{k-1}\ln r - \frac{\ln r}{2} - \frac{r-1}{r} + o_{r,k}(1)$ then a.a.s. $\chi(H(n,k,p)) \leqslant r$.

Sparse case, $r = 2$, $j = k - 1$

- **Theorem (Coja-Oghlan, Panagiotou, 2012)**:
  There exists $\varepsilon_k = 2^{-k(1+o_k(1))}$, such that if

$$c < 2^{k-1}\ln 2 - \frac{\ln 2}{2} - \frac{1}{4} - \varepsilon_k,$$

  then a.a.s. $\chi(H(n,k,p)) \leqslant 2$. On the other hand, if

$$c > 2^{k-1}\ln 2 - \frac{\ln 2}{2} - \frac{1}{4} + \varepsilon_k$$

  then a.a.s. $\chi(H(n,k,p)) > 2$.

Sparse case, $r = 2$, $k - j = o(\sqrt{k})$

- **Theorem (Semenov, 2017)**:
  There exist C_l, C_u and k_0, such that if $k > k_0$, $2 \leqslant k - j = o(\sqrt{k})$ and

$$c > \frac{2^{k-1} \ln 2}{\sum_{s=0}^{k-j-1} \binom{k}{s}} - \frac{\ln 2}{2} + C_u \cdot k^{k-j-1} \cdot 2^{-k},$$

  then a.a.s. $\chi_j(H(n, k, p)) > 2$. On the other hand, if

$$c < \frac{2^{k-1} \ln 2}{\sum_{s=0}^{k-j-1} \binom{k}{s}} - \frac{\ln 2}{2} - C_l \cdot (k/8)^{j+1-k},$$

  then a.a.s. $\chi_j(H(n, k, p)) \leqslant 2$.

Sparse case, arbitrary r, $k - j = o(k^{1/4})$

- **Theorem (Semenov, Shabanov, 2020)**: For any $r > 2$, there exist C_l, C_u and $k_0 = k_0(r)$, such that if $k > k_0$, $2 \leqslant k - j = o(k^{1/4})$ and

$$c > \frac{r^{k-1} \ln r}{\sum_{s=0}^{k-j-1} \binom{k}{s}(r-1)^s} - \frac{\ln r}{2} + C_u \cdot \binom{k}{j+1} \cdot r^{-j},$$

then a.a.s. $\chi_j(H(n,k,p)) > r$. On the other hand, if

$$c < \frac{r^{k-1} \ln r}{\sum_{s=0}^{k-j-1} \binom{k}{s}(r-1)^s} - \frac{\ln r}{2} - C_l \cdot k^{j-k+1},$$

then a.a.s. $\chi_j(H(n,k,p)) \leqslant r$.

Proof (upper bound)

- **Different model $H_1(n,k,m)$:** we randomly choose $m = \lceil cn \rceil$ edges with replacement from the set of all possible edges (some edges may repeat); For $c' > c$ and $p' = c'n/\binom{n}{k}$

$$P(\chi_j(H_1(n,k,\lceil cn \rceil)) > r) \leqslant P(\chi_j(H(n,k,p') > r) + o_n(1).$$

- **Counting proper colorings:** let X_n be a random variable corresponding to the number of j-proper colorings of $H_1(n,k,\lceil cn \rceil)$;
- **First moment method:**

$$P(X_n > 0) \leqslant EX_n.$$

Proof (lower bound)

- **Different model $H_2(n,k,m)$:** we choose $m = \lceil cn \rceil$ independent edges and in every edge choose k vertices also randomly and independently (some edges may repeat and be non-proper); For $c' < c$ and $p' = c'n/\binom{n}{k}$

$$\mathrm{P}(\chi_j(H_2(n,k,\lceil cn \rceil)) \leqslant r) \leqslant \mathrm{P}(\chi_j(H(n,k,p') \leqslant r) + o_n(1).$$

- **Sharp threshold:** Based on the result of Hatami and Molloy, there exists a function $\hat{p} = \hat{p}(n)$ such that for every $\varepsilon > 0$,

$$\mathrm{P}(\chi_j(H(n,k,p)) \leqslant r) \to \begin{cases} 1, & p < (1-\varepsilon)\hat{p}; \\ 0, & p > (1+\varepsilon)\hat{p}. \end{cases}$$

  Hence, we need to show that $\mathrm{P}(\chi_j(H_2(n,k,m) \leqslant r)$ is bounded from zero.

Second moment method

- We can consider only the case of n divisible by r.
- **Counting proper balanced colorings:** let X_n be a random variable corresponding to the number of j-proper balanced colorings of $H_2(n, k, \lceil cn \rceil)$, i.e. color classes are of size n/r;

$$\mathrm{P}(\chi_j(H_2(n, k, \lceil cn \rceil)) \leqslant r) \geqslant \mathrm{P}(X_n > 0);$$

- **Second moment method:** The Paley-Zygmund inequality says that

$$\mathrm{P}(X_n > 0) \geqslant \frac{(\mathrm{E}X_n)^2}{\mathrm{E}X_n^2};$$

So, the final step is to show that $\mathrm{E}X_n^2 = O_k((\mathrm{E}X_n)^2)$.

Second moment calculation

- To calculate the second moment we use matrices $A \in \mathcal{A}$ of size r by r with the property:

$$\sum_{i=1}^{r} a_{iu} = n/r, \sum_{u=1}^{r} a_{iu} = n/r.$$

- Let's also denote $J_r \in \mathcal{A}$ as a matrix with entries that are all the same and equal n/r^2.

Helper functions

Then, we introduce functions:

$$\mathcal{H}(A) = -\sum_{i,u=1}^{r} \frac{a_{iu}}{n} \ln \frac{ra_{iu}}{n};$$

$$\mathcal{E}(A) = \ln\Bigg(1 - 2r^{1-k} \sum_{s=j+1}^{k} \binom{k}{s} (r-1)^{k-s} + \sum_{i,u=1}^{r} \sum_{s=j+1}^{k} \binom{k}{s} \sum_{h=0}^{k-s} \sum_{t=0}^{s-j-1+h} \binom{k-s}{h} \binom{s}{t} \cdot \left(\frac{\frac{n}{r} - a_{iu}}{n}\right)^{h+t} \left(\frac{a_{iu}}{n}\right)^{s-t} \left(\frac{\frac{n(r-2)}{r} + a_{iu}}{n}\right)^{k-h-s}\Bigg).$$

Crucial technical lemma

For c that satisfies conditions of the theorem, the expression $\mathcal{G}_c(A) = \mathcal{H}(A) + c \cdot \mathcal{E}(A)$ takes it's minimum value when $A = J_r$. Which comes from the following

Lemma: There exist $b = b(k, r) > 0$, such that for every $A = (a_{iu}, i, u = 1, \ldots, r)$ from $\mathcal{A}$

$$\mathcal{G}_c(A) - \mathcal{G}_c(J_r) \geqslant b \sum_{i,u=1}^{r} \left(\frac{a_{iu}}{n} - \frac{1}{r^2} \right)^2 .$$