Lecture 7: October 25

Lecturer: Kuan Yang *Scribe: Weihao Zhu*

7.1 Chromatic Number of $\mathcal{G}(n, 1/2)$

Today, we will discuss more on $\chi(\mathcal{G}(n, 1/2))$. We first focus on the concentration for clique numbers.

Theorem 7.1 (Bollobás-Erdős, 1976 & Matula, 1976) *There exists a $k \approx 2\log_2 n$ such that*

$$\omega(\mathcal{G}(n, 1/2)) \in \{k, k+1\}$$

with high probability.

Proof: Let

$$f(k) = \mathbf{E}[X] = \binom{n}{k} 2^{-\binom{k}{2}}.$$

If $f(k) \to \infty$, then we have $\Delta^* \ll f(k)$ (we omit details temporarily and the full calculation can be found in the following sections), which implies that $\omega(\mathcal{G}(n, 1/2)) \geq k$ w.h.p.

For $k = (1 \pm o(1))2\log_2(n)$, we have

$$\frac{f(k+1)}{f(k)} = \frac{n-k}{k+1} \cdot 2^{-k} = n^{-1+o(1)}.$$

So $f(k)$ decreases rapidly when $k \approx 2\log_2 n$.

Let $k_0 = k_0(n)$ be the value such that $f(k_0) \geq 1 > f(k_0+1)$. For n such that $f(k_0) \to \infty$ and $f(k_0+1) \to 0$, it is known that

$$\omega(\mathcal{G}(n, 1/2)) = k_0$$

with high probability.

If $f(k_0) = O(1)$ (or $f(k_0+1) = O(1)$, then we increacse k_0 by 1), we have $f(k_0-1) \to \infty$ and $f(k_0+1) \to 0$. Thus,

$$\omega((G)(n, 1/2)) \in \{k_0 - 1, k_0\}$$

with high probability. This completes the proof. ■

However, this concentration is not what we want for analyzing chromatic numbers.

For the upper bound, we give a strategy to properly color the graph. Take out an independent set of size approximately $2\log_2 n$, and color them with a new color. Repeat this process until $o(n\log_2 n)$ vertices remaining, and color each of them with a new color. However, after removing independent sets, the distributions of remaining sub-graphs are no longer random graphs. Instead, if we fix a subset S of size m, the distribution of $G[S]$ induced by S is exactly $\mathcal{G}(m, p)$. Then if we could show that for all S of size m, $G[S]$ has an independent set of size $\approx 2\log_2 n$, we can repeat the process: finding an independent set, coloring them with a new color and then erasing them. To show that $\alpha(G[S]) \geq 2\log_2|S|$ for all S, we need the union bound and thus the probability of a "bad" event should be $o(1/\binom{n}{m})$. So this concentration result is not sufficient.

In 1988, to analyze the chromatic number, Bollobás also proved the following "stronger" theorem.

Theorem 7.2 (Bollobás, 1988) *Let k_0 be the largest number such that $f(k_0) \geq 1$, then*

$$\mathbf{Pr}[\omega(\mathcal{G}(n, 1/2)) < k_0 - 3] = e^{-n^{2-o(1)}}.$$

Remark. For a constant p, we have

$$\alpha(\mathcal{G}(n, p)) = \omega(\mathcal{G}(n, 1-p)) \approx 2 \log_{1/p} n$$

with high probability.

Now we can state the result to the chromatic number of $\mathcal{G}(n, 1/2)$.

Theorem 7.3 (Bollobás, 1988)

$$\chi(\mathcal{G}(n, 1/2)) \approx \frac{n}{2 \log_2 n}$$

with high probability.

Proof Sketch. Clearly,

$$\chi(\mathcal{G}(n, 1/2)) \geq \frac{n}{\alpha(\mathcal{G}(n, 1/2))} \geq (1 - o(1)) \frac{n}{2 \log_2 n}$$

with high probability. This provides us a good lower bound of the chromatic number of $\mathcal{G}(n, 1/2)$. We now show that $\chi(\mathcal{G}(n, 1/2)) \leq (1 + o(1))n/(2 \log_2 n)$.

Following the previous idea: finding an independent set, coloring them with a new color and then erasing them, until there are at most m vertices, where we can assign each remaining vertex a distinct color.

So we choose m and hope

- with high probability, for any subset S of size m, $\alpha(G[S]) \approx 2 \log_2 n$;
- $n/(2 \log_2 n) + m \leq (1 + o(1))n/(2 \log_2 n)$.

We can show that $m = n/(\log_2 n)^2$ suffices.

Proof: Choose $m = n/(\log_2 n)^2$. Notice that $2 \log_2 m = 2(\log_2 n - 2 \log \log n)$. Fix any subset S of size m, we have

$$\mathbf{Pr}[\alpha(G[S]) < (1 - o(1))2 \log_2 n] \leq e^{-m^{2-o(1)}} \ll e^{-n},$$

which implies that

$$\mathbf{Pr}[\forall S \text{ of size } m, \alpha(G[S]) \geq (1 - o(1))2 \log_2 n] \geq 1 - \binom{n}{m} e^{-n} = 1 - o(1).$$

Set $k = k_0(m) - 3$. While there are at least m vertices remaining, we find an independent set of size k, color them and remove them. Finally, color all remaining vertices with distinct colors. This gives us a proper coloring of the graph. Therefore,

$$\chi \leq \frac{n}{k} + m = (1 + o(1)) \frac{n}{2 \log_2 n},$$

which completes the proof. ■

7.2 Chernoff Bound and Martingale Concentration

Now the remaining task is to show Theorem 7.2, where we need some more tools. We now briefly introduce the Chernoff bound and concentration inequalities for martingales.

7.2.1 Chernoff Bound

Theorem 7.4 (Chernoff bound) *Let $S_n = X_1 + X_2 + \ldots + X_n$ where $X_i \in \{-1, 1\}$ are uniformly i.i.d. Then, for any $\lambda > 0$, we have*

$$\mathbf{Pr}[S_n \geq \lambda\sqrt{n}] \leq e^{-\lambda^2/2}.$$

Proof: Let $t = \lambda/\sqrt{n} \geq 0$. Consider the moment generating function $\mathbf{E}[e^{tS_n}]$. Then, we have

$$\mathbf{Pr}[S_n \geq \lambda\sqrt{n}] \leq \frac{\mathbf{E}[e^{tS_n}]}{e^{t\lambda\sqrt{n}}} \leq e^{-t\lambda\sqrt{n}+t^2 n/2} = e^{-\lambda^2/2},$$

which completes the proof. ■

Remark. Chebyshev inequality only tells us the probability is at most $1/\lambda^2$ since $\mathbf{Var}[S_n] = \sum \mathbf{Var}[X_i] = n$.

The Chernoff bound gives us the following two corollaries.

Corollary 7.5 *Let $X_i \in [-1, 1]$ independently with $\mathbf{E}[X_i] = 0$ (not necessarily i.i.d.). Then, $S_n = X_1 + \ldots + X_n$ has*

$$\mathbf{Pr}[S_n \geq \lambda\sqrt{n}] \leq e^{-\lambda^2/2}.$$

Proof: By convexity, we have

$$e^{tx} \leq \frac{1-x}{2} \cdot e^{-t} + \frac{1+x}{2} \cdot e^{t}.$$

So,

$$\mathbf{E}[e^{tX}] \leq \frac{e^{-t} + e^{t}}{2}.$$

The rest part of the proof is the same. ■

Corollary 7.6 *Let X be the sum of n independent Bernoulli random variables (not necessarily the same). Let $\mu = \mathbf{E}[X]$ and $\lambda \geq 0$. Then,*

$$\mathbf{Pr}[X \geq \mu + \lambda\sqrt{n}] \leq e^{-\lambda^2/2}.$$

Comparison to the normal distribution $N(0, 1)$. As $\mathbf{E}[e^{tX}] = e^{t^2/2}$, we have

$$\mathbf{Pr}[X \geq \lambda] \leq e^{-t\lambda}\mathbf{E}[e^{-tX}] = e^{-t\lambda+t^2/2} = e^{-\lambda^2/2}$$

by setting $t = \lambda$.

Remark. A random variable X with $\mathbf{E}[X] = 0$ and $\mathbf{Pr}[|X| \geq t] \leq 2e^{-ct^2}$ for all $t \geq 0$ and constant $c > 0$ is called a sub-gaussian. Usually, the exact value of c is not significant.

7.2.2 Martingale

We now develop similar sub-gaussian tail bound for other variables.

Definition 7.1 (martingale) *A martingale is a random variable sequence* $\{Z_0, Z_1, \ldots\}$ *such that for any* n, $\mathbf{E}[Z_n] < \infty$ *and*

$$\mathbf{E}[Z_{n+1}|Z_0, \ldots, Z_n] = Z_n.$$

Remark. Usually, Z_n depends on $X_0, \ldots, X_n$ and satisfies $\mathbf{E}[Z_{n+1}|X_0, \ldots, X_n] = Z_n$.

Definition 7.2 (Doob martingale) *Given an underlying r.v.s.* $X_1, \ldots, X_n$ *and* $f(X_1, \ldots, X_n)$, *then*

$$Z_i = \mathbf{E}[f(X_1, \ldots, X_n)|X_1, \ldots, X_i]$$

is a martingale with respect to $X_1, \ldots, X_n$.

In random graphs, we have two classical martingales:

- Edge-exposure martingale: $\mathbf{E}[f(\mathcal{G}(n,p))|X_0, X_1, \ldots, X_{\binom{n}{2}}]$, where each variable symbolizes an edge;
- Vertex-exposure martingale: $\mathbf{E}[f(\mathcal{G}(n,p))|X_0, X_1, \ldots, X_n]$, where each variable symbolizes a vertex.

Remark. There is a trade-off between the length and the difference bound.

Theorem 7.7 (Azuma's inequality) *Let* $Z_0, Z_1, \ldots, Z_n$ *be a martingale such that* $|Z_i - Z_{i-1}| \leq c_i$ *for any* $i \in [n]$. *Then,*

$$\mathbf{Pr}[Z_n - Z_0 \geq \lambda] \leq e^{-\lambda^2/2(c_1^2+\ldots+c_n^2)}.$$

More generally, if Z_i *conditioned on* $Z_0, \ldots, Z_{i-1}$ *lies inside an interval of length* c_i *(the interval may depends on* $Z_0, \ldots, Z_{i-1}$*, but its length is unpper bounded), then*

$$\mathbf{Pr}[Z_n - Z_0 \geq \lambda] \leq e^{-2\lambda^2/(c_1^2+\ldots+c_n^2)}.$$

Remark. Applying Auzma's inequality to Z_n and $-Z_n$, it gives

$$\mathbf{Pr}[|Z_n - Z_0| \geq \lambda] \leq 2e^{-2\lambda^2/(c_1^2+\ldots+c_n^2)}.$$

Theorem 7.8 (Bounded differences inequality) *Let* $X_1 \in \Omega_1, \ldots, X_n \in \Omega_n$ *be* n *independent r.v.s.. Suppose* $f : \Omega \times \ldots \times \Omega_n \to \mathbb{R}$ *is a function such that*

$$|f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)| \leq c_i.$$

Then the random variable $Z = f(X_1, \ldots, X_n)$ *satisfies that for any* $\lambda \geq 0$,

$$\mathbf{Pr}[Z - \mathbf{E}[Z] \geq \lambda] \leq e^{-2\lambda^2/(c_1^2+\ldots+c_n^2)}.$$

So is $\mathbf{Pr}[Z - \mathbf{E}[Z] \leq -\lambda]$.

In particular, if f *satisfies* $|f(x) - f(y)| \leq c \cdot \|x - y\|_0$, *where the* 0*-norm of a vector* v, *denoted by* $\|v\|_0$, *is the number of nonzero elements in* v *(here we say* f *is c-Lipschitz), then*

$$\mathbf{Pr}[Z - \mathbf{E}[Z] \geq \lambda] \leq e^{-2\lambda^2/(nc^2)},$$

and so is $\mathbf{Pr}[Z - \mathbf{E}[Z] \leq -\lambda]$.

7.2.3 Proof of Theorem 7.2

Now, we use the bounded differences inequality to prove Theorem 7.2.

Proof: Let $k = k_0 - 3$. Define $Y = Y(G)$ as the maximum number of edge-disjoint k-cliques in G. Using the edge-exposure martingale, we have $Y = f(X_{e_1}, \ldots, X_{e_{\binom{n}{2}}})$. Notice that Y changes at most 1 if G changes only one edge. (Warning: This is not true if G changes one vertex!) By the bounded differences inequality, for $G \sim \mathcal{G}(n, 1/2)$, letting $\mu = \mathbf{E}[Y]$, we have

$$\mathbf{Pr}[\omega(G) < k] = \mathbf{Pr}[Y(G) = 0] \leq \mathbf{Pr}[Y - \mu \leq -\mu] \leq e^{-2\mu^2/\binom{n}{2}}.$$

Our goal is to prove

$$\mathbf{Pr}[\omega(G) < k] < e^{-n^{2-o(1)}}.$$

It suffices to show $\mu \geq n^{2-o(1)}$.

Consider an auxiliary graph H whose vertices are k-cliques in G, and $(u, v) \in E(H)$ if clique u and clique v overlap in at least 2 vertices in G. Then, based on Caro-Wei inequality, we have

$$Y = \alpha(H) \geq \frac{|V(H)|^2}{|V(H)| + 2|E(H)|}.$$

Now, we use second moment method to compute $|V(H)|$ and $|E(H)|$.

As

$$\mu_v = \mathbf{E}[|V(H)|] = \binom{n}{k} \cdot 2^{-\binom{k}{2}} \geq n^{3-o(1)} \to \infty,$$

by the second moment method, we have $V(H) = (1 \pm o(1))\mu_v$ with high probability.

For $|E(H)|$, we have

$$\mu_e = \mathbf{E}[|E(H)|] = \frac{\Delta}{2} = \frac{\mu_v}{2}\Delta^* = \frac{\mu_v}{2} \sum_{\ell=2}^{k-1} \binom{k}{\ell}\binom{n-k}{k-\ell} 2^{\binom{\ell}{2} - \binom{k}{2}}.$$

Let $g(\ell) = \binom{k}{\ell}\binom{n-k}{k-\ell} \cdot 2^{\binom{\ell}{2} - \binom{k}{2}}$, then

$$\frac{g(\ell)}{g(\ell+1)} = \frac{(\ell+1)(n-2k+\ell+1)}{(k-\ell)^2} \cdot 2^{-\ell}.$$

Note that $k \approx 2\log_2 n$. This implies that if $\ell \geq \frac{3}{4}k$, then $g(\ell) \leq g(\ell+1)$. Therefore,

$$\sum_{\frac{3}{4}k \leq \ell < k} g(\ell) \leq \frac{k}{4} g(k-1) = \frac{k}{4} \cdot k \cdot (n-k) \cdot 2^{-(k-1)} = O(k^2/n).$$

If $\ell \leq \frac{3}{4}k$, then

$$\begin{aligned}
\frac{g(\ell)}{\mu_v} &= \frac{\binom{k}{\ell}\binom{n-k}{k-\ell}}{\binom{n}{k}} \cdot 2^{\binom{\ell}{2}} \\
&\leq \frac{(n-k)^{k-\ell}/(k-\ell)!}{(n-k)^k/k!} \cdot \binom{k}{\ell} \cdot 2^{\ell(\ell-1)/2} \\
&\leq \frac{k!/(k-\ell)!}{(n-k)^\ell} \cdot \binom{k}{\ell} \cdot 2^{\ell(\ell-1)/2} \\
&\leq \frac{k^{2\ell}}{(n-k)^\ell} \cdot 2^{\ell(\ell-1)/2} \\
&= \Big(\frac{k^2 \cdot 2^{(\ell-1)/2}}{n-k}\Big)^\ell \\
&= \begin{cases} O(k^4/n^2) & \ell = 2 \\ o(k^4/n^2) & \ell > 2 \end{cases}
\end{aligned}$$

Therefore, if $\mu_v \geq n^{3-o(1)}$, then

$$\mu_e = \frac{\mu_v}{2} \sum g(\ell) = c \cdot \mu_v^2 \cdot \frac{k^4}{n^2} \gg \mu_v.$$

So, we have

$$\begin{aligned}
\mathbf{E}[Y] &\geq \mathbf{E}\Big[\frac{v^2}{v+2e}\Big] \\
&\geq \mathbf{E}\Big[\frac{v^2}{v+2e} | v \geq (1-o(1))\mu_v\Big] \cdot \mathbf{Pr}[v \geq (1-o(1))\mu_v] \\
&= (1-o(1))\mathbf{E}\Big[\frac{\mu_v^2}{\mu_v+2e}\Big] \\
&\geq (1-o(1))\frac{\mu_v^2}{\mu_v+2\mu_e} && \text{(by Jensen's inequality)} \\
&= O(n^2/k^4).
\end{aligned}$$

Alternative proof: Without strong concentration, use alteration method. Pick each $v \in H$ with probability q. Then,

$$\mathbf{E}[Y] \geq \mathbf{E}[q|V(H)| - q^2|E(H)|] = q\mu_v - q^2\mu_e.$$

Choose $q = \frac{\mu_v}{2\mu_e}$, and we have

$$\mathbf{E}[Y] \geq \mu_v^2/4\mu_e = O(n^2/k^4) = n^{2-o(1)},$$

which completes the proof. ∎

Finally, if $\mathbf{E}[\chi(\mathcal{G}(n,p))]$ is known, using vertex-exposure martingale, it gives us the following theorem.

Theorem 7.9 (Shamir & Spencer, 1987) *For any* $\lambda \geq 0$,

$$\mathbf{Pr}[\chi - \mathbf{E}[\chi] \geq \lambda \cdot \sqrt{n-1}] \leq e^{-2\lambda^2}.$$

So is $\mathbf{Pr}[\chi - \mathbf{E}[\chi] \leq -\lambda \cdot \sqrt{n-1}]$.

Remark. For $p = 1/2$, χ does not concentrate on any interval of length no larger than $n^{1/4}$. But for sparse random graphs where $p = n^{-\alpha}$ for all $\alpha > 1/2$, χ has a two-point concentration. We leave a simple version as homework.

7.3 An Introduction to the Lovász Local Lemma

Suppose we have a set of events $A_1, \ldots, A_n$, each with probability p_i. If $\sum p_i < 1$, then by the union bound (or Markov's inequality), we know that $\mathbf{Pr}[\cap \overline{A}_i] > 0$ or even almost surely if $\sum p_i = o(1)$. If $\sum p_i = O(1)$ or even $\sum p_i \to \infty$, then we know nothing about $\mathbf{Pr}[\cap \overline{A}_i]$. Let X_i be the indicator of A_i. If $\mathbf{Var}[X] = o(\mathbf{E}[X]^2)$, then $\mathbf{Pr}[\cap \overline{A}_i] = \mathbf{Pr}[X = 0] = o(1)$. However, what do we need if we want to prove that $\mathbf{Pr}[\cap \overline{A}_i] > 0$?

In this section, we will introduce the celebrated Lovász local lemma. We start from the definition of dependency.

Definition 7.3 (Dependency) *Suppose we have n "bad events" $A_1, \ldots, A_n$. For each A_i, there is some subset $N(i) \subseteq [n]$ such that A_i is independent from $\{A_j : j \neq i, j \notin N(i)\}$. We say an event A_0 is independent from $\{A_1, \ldots, A_m\}$ if for any $B_i \in \{A_i, \overline{A}_i\}$, $\mathbf{Pr}[A_0 | B_1, B_2, \ldots, B_m] = \mathbf{Pr}[A_0]$.*

Remark. We usually represent above relations by a dependency (di-)graph whose vertices are events, and $A_i \to A_j$ if and only if $j \in N(i)$.

Important Remark. Pay attention that pairwise independence does not implies mutually independence. For the local lemma we need a stronger notion of independence. Consider $x_1, x_2, x_3 \in \{0, 1\}$ uniformly and A_i is the event that $\sum_{j \neq i} x_j = 0$. Then any two events are pairwise independent but are not independent if we consider the third event. Thus, the empty graph is not a valid dependency graph. But, any graph with at least two edges is a valid dependency graph.

Theorem 7.10 (Lovász Local Lemma, symmetric version) *Let $A_1, \ldots, A_n$ be events with $\mathbf{Pr}[A_i] \leq p$. Suppose that each A_i is independent from all other A_j except at most d of them. If $ep(d + 1) \leq 1$, then* $\mathbf{Pr}[\cap \overline{A}_i] > 0$.

Let's take an example of hypergraph coloring. Let $H = (V, E)$ be a hyper-graph. A coloring c is proper if there doesn't exist a monochromatic edge. We can see that for any two edges $e, f \in E$, $e \sim f$ if $e \cap f = \emptyset$. According to Lovász local lemma, if the hypergraph is k-uniform, maximum vertex degree is at most Δ, and $ek\Delta q^{1-k} \leq 1$, then H is q-colorable.

(To be continued...)