Concentration of the Chromatic Number

Radhika Shah

Euler Circle

June 2026

Introduction

Deciding whether the chromatic number $\chi(G)$ is less than k is an NP-complete problem for $k \geq 3$, it takes nondeterministic polynomial to solve but is easy to verify. However, for a random graph $G(n,p)$, $\chi(G)$ is highly predictable.

Main result. The result established by the Shamir-Spencer theorem is that for any fixed $p \in (0,1)$ and any $t > 0$,

$$\mathbb{P}[|\chi(G(n,p)) - \mathbb{E}[\chi(G(n,p))]| \geq t\sqrt{n}] \leq 2e^{-\frac{t^2}{2}}.$$

The chromatic number lies within an interval of width $O(\sqrt{n})$ around its mean.

Introduction

In order to prove this result, we first define several concepts, including Hoeffding's lemma, the Azuma-Hoeffding inequality, McDiarmid's inequality, the vertex exposure martingale, and then the concentration of $\chi(G)$.

Finally, we discuss Alon's improvement to $O(\frac{\sqrt{n}}{\log(n)})$ and extensions to other random graph models.

The Erdős-Rényi Model

The Erdős-Rényi Random Graph definition: Start with n vertices labeled $1, 2, ..., n$. For each pair of vertices $\{i, j\}$, flip a coin that is heads with probability $p \in [0, 1]$, independently of all other coins. Include the edge $\{i, j\}$ if and only if the coin is heads. The result is $G(n, p)$, the Erdős-Rényi model of random graphs.

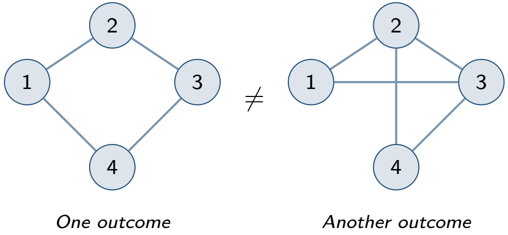


Figure: Two different outcomes of $G(4)$, every run is random.

The Erdős-Rényi Model

Let N be the actual number of edges in $G(n,p)$. $G(n,p)$ has $\binom{n}{2} = \frac{n(n-1)}{2}$ possible edges. Each is present independently with probability p, so the expected number of edges is

$$E\left[\mathrm{N}\right] = p \cdot \binom{n}{2} = p \cdot \frac{n(n-1)}{2}.$$

For instance, for $n = 100$ and $p = 0.5$, $E\left[\mathrm{N}\right] = 0.5 \cdot \frac{100 \cdot 99}{2} = 2475$.

The actual count varies but concentrates tightly by the law of large numbers.

To find $E\left[\chi(G(n,p))\right]$ for fixed $p \in [0,1]$, with probability tending to 1 as $n \to \infty$: $\chi(G(n,p))$ is approximately $\frac{n}{2 \cdot \log_b(n)}$, where $b = \frac{1}{1-p}$.

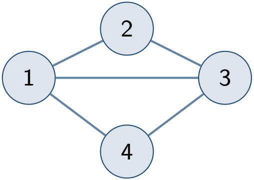
The Erdős-Rényi Model

n	p	$b = \frac{1}{1-p}$	$\log_b(n)$	$E(\chi(G(n,p))) \approx \frac{n}{2 \cdot \log_b(n)}$
100	0.5	2	6.6	8
1000	0.5	2	10	50
10000	0.5	2	13.3	376
1000	0.9	10	3	167
1000	0.1	1.11	68	7

Table: Chromatic number for various n and p, higher p gives denser graphs and higher χ

Graph Coloring

A graph $G = (V, E)$ consists of a set V of vertices and a set E of edges that connect vertices. An edge $\{u, v\}$ means u and v are directly connected, or adjacent.



$V = \{1, 2, 3, 4\}$
$E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 1\}, \{1, 3\}\}$

Figure: A graph with 4 vertices and 5 edges. The line between two vertices is an edge.

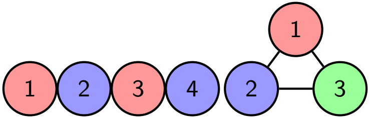
Definition

A proper k-coloring of graph G is an assignment of one of k colors to each vertex in G such that no two adjacent vertices share the same color.

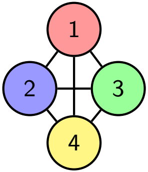
Definition

The chromatic number $\chi(G)$ is the smallest k for which a proper k-coloring exists.

Graph Coloring



Path P_4: $\chi = 2$ Triangle K_3: $\chi = 3$



K_4: $\chi = 4$

Figure: Examples of proper colorings. Adjacent vertices must receive different colors.

Theorem

$\chi(G) \geq \omega(G)$ where $\omega(G)$ is the clique number, or the size of the largest complete subgraph.

Informal proof: A clique is a set of vertices that are all connected to each other. If G contains a clique of size k, then those k vertices are all directly connected, and therefore all need different colors. Hence $\chi(G) \geq k$.

Graph Coloring

Theorem

$\chi(G) \leq \Delta(G) + 1$, where $\Delta(G)$ is the maximum degree, or the largest number of neighbors any vertex has.

Informal proof: Number the vertices (in any order) and color them one by one, always picking the smallest number not already used by an adjacent vertex. When vertex v is colored, at most $\deg(v) \leq \Delta(G)$ adjacent vertices are already colored, so a color from $\{1, ..., \Delta(G) + 1\}$ is always available.

Martingales

We want to show that $\chi(G(n,p))$ is close to the expected value $E[\chi(G(n,p))]$.

The strategy is to build a random graph one vertex at a time and track the best prediction for χ at each step. Then, we want to show that the predictions are concentrated, so the final value $\chi(G)$ cannot be far from the initial guess $E(\chi(G))$.

The martingale approach is to prove χ concentrates without computing the intermediate estimates.

Definition

A martingale is a sequence of random variables $(X_0, X_1, X_2, ..., X_n)$ where the best prediction of the next value is the current value.

Martingales

A biased coin would not be a martingale. For instance, if the coin is weighted 0.6 chance of heads and 0.4 chance of tails, this is not a martingale: $\mathbb{E}[X_{i+1} \mid X_i = 10] = 0.6(11) + 0.4(9) = 10.2 \neq 10$.

Define a filtration ($F_0 \subseteq F_1 \subseteq F_2 ... \subseteq F_n$) as the amount of information known by the current coin flip. So F_0, there are no vertices placed and no edges revealed. At F_1, vertex v_1 exists. At F_2, vertex v_2 exists, and the coin has flipped to determine whether edge $\{v_1, v_2\}$ exists. By F_i, the entire subgraph with $\{v_1, v_2, ..., v_i\}$ is known (all edges among the i vertices). F_n is the full graph, with all n vertices.

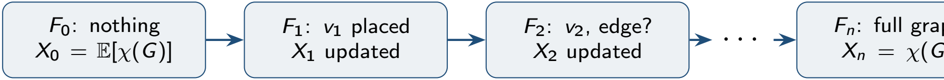


Figure: The Doob martingale: a sequence of improving estimates from $\mathbb{E}[\chi(G)]$ to the exact $\chi(G)$.

Martingales

Definition: A sequence $(X_0, X_1, X_2, ..., X_n)$ is a martingale with respect to a filtration $(F_0 \subseteq F_1 \subseteq F_2 ... \subseteq F_n)$ if:

1. X_i is determined by the information in F_i.
2. $E[X_{i+1} \mid F_i] = X_i$

For step 2, on average, X_{i+1} is the same as X_i. Revealing one more vertex may move the actual prediction up or down, but the expected amount of movement of the prediction is 0 since it's a martingale.

The Doob Martingale. Definition: Let Z be any random variable. In this case, $Z = \chi(G)$. Let (F_i) be a filtration encoding information revealed step by step. Then, the sequence $(X_0, X_1, ..., X_n)$ is always a martingale. It starts at $X_0 = E[Z]$ and ends at $X_n = Z$.

The moment generating function trick

The variable X used in this section denotes a single step of movement, $D_i = X_i - X_{i-1}$. The moment generating function trick is applied to each individual step, not to the whole sequence.

We want to show that the probability that X is greater than t is small to prove that the chromatic number concentrates. Without knowing the exact distribution of X, we can use a workaround by transforming X into e^{sX} for $s > 0, s \in \mathbb{R}$. Then, since e^{sX} is an increasing function, $X \geq t$ means $e^{sX} \geq e^{st}$.

Next, we bring in Markov's inequality. Markov's inequality states that for any non-negative quantity $W, P[W \geq a] \leq \frac{E[W]}{a}$.

Then, applying this to $e^{sX} \geq e^{st}$, we get that $P[e^{sX} \geq e^{st}] \leq \frac{E[e^{sX}]}{e^{st}}$. We choose s to make the bound as tight as possible.

Hoeffding's Lemma

Hoeffding's lemma gives a bound on the moment generating function of a bounded, mean-zero random variable. In our application, X is a single step $D_i = X_i X_{i-1}$. The martingale property gives $E[D_i|F_{i-1}] = 0$, and X is also bounded by $[-1, 1]$ since adding one vertex changes χ by at most one. Hoeffding's lemma gives the upper bound on $E[e^{sX}]$, $E[e^{sX}] \leq e^{(\frac{s^2}{2})}$.

From the previous section we have $P[X \geq t] \leq \frac{E[e^{sX}]}{e^{st}}$. By Hoeffding's lemma, we have the explicit value to plug in for $E[e^{sX}]$, so it is possible to get an actual number bounding $P[X \geq t]$.

Once we can bound the moment generating function of each individual step, we can bound the total drift of the martingale over all n steps.

Azuma-Hoeffding Inequality

The Azuma–Hoeffding inequality extends Hoeffding's lemma from a single bounded variable to a martingale, bounding how far the final value can drift from its starting value.

Theorem

Let $(X_0, X_1, \ldots, X_n)$ be a martingale with $|X_i - X_{i-1}| \leq 1$. Then for all $t > 0$:

$$\mathbb{P}[|X_n - X_0| \geq t] \leq 2e^{\frac{-t^2}{2n}}.$$

Here $X_n = \chi(G)$ and $X_0 = \mathbb{E}[\chi(G)]$, so $|X_n - X_0|$ is exactly how far $\chi(G)$ is from its expected value.

Proof of the Azuma-Hoeffding Inequality

Write $X_n - X_0 = D_1 + D_2 + \cdots + D_n$ where $D_i = X_i - X_{i-1}$ are the martingale steps.
Step 1: Apply the moment-generating function trick:
$\mathbb{P}[X_n - X_0 \geq t] \leq e^{-st} \cdot \mathbb{E}[e^{s(X_n - X_0)}]$.
Step 2: Use $e^{a+b} = e^a \cdot e^b$ to factor:

$$\mathbb{E}[e^{s(D_1 + \cdots + D_n)}] = \mathbb{E}[e^{sD_1} \cdots e^{sD_n}].$$

Step 3: Apply Hoeffding's lemma to each D_i separately (each has $\mathbb{E}[D_i \mid F_{i-1}] = 0$ and $D_i \in [-1, 1]$):

$$\mathbb{E}[e^{sD_i}] \leq e^{s^2/2}.$$

Multiplying over all n steps: $\mathbb{E}[e^{s(X_n - X_0)}] \leq e^{ns^2/2}$.
Step 4: Plug back in: $\mathbb{P}[X_n - X_0 \geq t] \leq e^{-st} \cdot e^{ns^2/2}$. Setting $s = t/n$ minimizes the right-hand side, giving $e^{-t^2/(2n)}$. The two-sided bound $\mathbb{P}[|X_n - X_0| \geq t]$ gives the factor of 2. $\square$

McDiarmid's Inequality

McDiarmid's inequality restates the same idea directly in terms of a function of independent inputs, which is often more convenient to apply than working with the martingale sequence itself.

Theorem

Let $\chi(G(X_1, \ldots, X_n))$ satisfy: changing any single input X_i to any other value in $(X_1, \ldots, X_n)$ changes χ by at most 1 since adding one vertex can change χ by at most 1. Then for any $t > 0$:

$$\mathbb{P}[\chi(G(X_1, \ldots, X_n)) - \mathbb{E}[\chi] \geq t] \leq e^{\frac{-2t^2}{n}}.$$

How McDiarmid's follows from Azuma: Define the Doob martingale $M_k = \mathbb{E}[\chi(G) \mid X_1, \ldots, X_k]$. The bounded differences property ensures $|M_k - M_{k-1}| \leq 1$. Apply Azuma. $\square$

Edge-Exposure and Vertex-Exposure Martingales

There are two ways to reveal $G(n, p)$ step by step: edge by edge, or vertex by vertex.

Fix an arbitrary ordering of all $\binom{n}{2} = \frac{n(n-1)}{2}$ potential edges. Reveal them one at a time, and let $F_i =$ the information from the first i edges. Let $X_i = \mathbb{E}[\chi(G) \mid F_i]$.

Revealing one more edge changes χ by at most 1, since adding an edge can require at most one new color, and removing an edge can free at most one color. So $|X_i - X_{i-1}| \leq 1$ for all i.

Applying Azuma with $\frac{n(n-1)}{2}$ steps:

$$\mathbb{P}[|\chi(G) - \mathbb{E}[\chi(G)]| \geq t] \leq 2e^{\frac{-t^2}{n(n-1)}}.$$

Since $\chi(G) \leq n$ always, this bound only guarantees concentration within $O(n)$ of the mean, which we already knew. So we must use vertex exposure instead.

Edge-Exposure and Vertex-Exposure Martingales

Approach	Number of steps	Window from Azuma
Edge exposure	$\frac{n(n-1)}{2}$	$O(n)$
Vertex exposure	n	$O(\sqrt{n})$

Table: The improvement comes from reducing the number of steps from $O(\frac{n(n-1)}{2})$ to n.

Vertex-Exposure Martingale. Reveal the graph vertex by vertex. At step i, we add vertex v_i and reveal all $i-1$ edges between v_i and the previous vertices $v_1, \ldots, v_{i-1}$. Let $G_i = G[\{v_1, \ldots, v_i\}]$ (the induced subgraph on the first i vertices). Define:

$$Y_i = \mathbb{E}[\chi(G) \mid G_i].$$

Vertex-Exposure Martingale

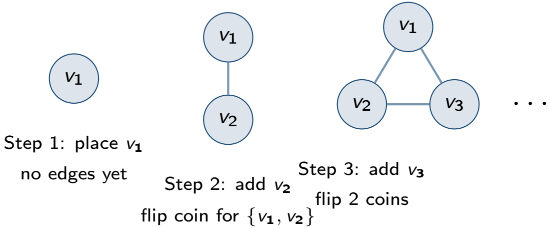


Figure: Vertex exposure: the graph is revealed one vertex at a time, with all its edges to prior vertices.

Theorem

Bounded Differences Lemma for Vertex Exposure: $|Y_i - Y_{i-1}| \leq 1$ for all $i = 1, \ldots, n$.

Vertex-Exposure Martingale

Theorem

For *any* graph H and any vertex v, $|\chi(H) - \chi(H - v)| \leq 1$, where $H - v$ denotes H with v isolated.

Proof

1. $\chi(H) \leq \chi(H - v) + 1$: Take an optimal coloring of $H - v$ using $\chi(H - v)$ colors. When v is restored, it may conflict with all $\chi(H - v)$ colors among its neighbors, requiring at most one additional color.
2. $\chi(H - v) \leq \chi(H)$: Any proper coloring of H restricts to a proper coloring of $H - v$ by ignoring v.

The step from Y_{i-1} to Y_i reveals the edges from v_i to $\{v_1, \ldots, v_{i-1}\}$. For any two outcomes of these edges, the resulting full graphs differ only in edges adjacent to v_i. By the bound above, their chromatic numbers differ by at most 1, so the conditional expectations Y_i differ by at most 1. $\square$

Main Theorem (Shamir-Spencer)

Theorem (Shamir–Spencer)

For fixed $p \in (0, 1)$ and any $t > 0$:

$$\mathbb{P}\left[|\chi(G(n, p)) - \mathbb{E}[\chi(G(n, p))]| \geq t\sqrt{n}\right] \leq 2e^{-t^2/2}.$$

Proof:

1. Let $(Y_i)_{i=0}^n$ be the vertex-exposure Doob martingale, so $Y_0 = \mathbb{E}[\chi(G)]$ and $Y_n = \chi(G)$.
2. By the Bounded Differences Lemma: $|Y_i - Y_{i-1}| \leq 1$ with $c_i = 1$ for all i.
3. Apply Azuma–Hoeffding with $\sum c_i^2 = n$:

$$\mathbb{P}[|Y_n - Y_0| \geq t] \leq 2e^{-t^2/(2n)}.$$

4. Substitute $t \mapsto t\sqrt{n}$: $\mathbb{P}[|\chi(G) - \mathbb{E}[\chi(G)]| \geq t\sqrt{n}] \leq 2e^{-t^2 n/(2n)} = 2e^{-t^2/2}$.
   $\square$

Alon's Improvement

The Shamir–Spencer bound of $O(\sqrt{n})$ is not the tightest possible. In the 1990s, Alon improved it to $O(\sqrt{n}/\log n)$ for fixed $p \in (0, 1)$.

Theorem

Alon's improvement: For fixed $p \in (0, 1)$, $\chi(G(n, p))$ is concentrated in an interval of width $O(\sqrt{n}/\log n)$.

The Shamir–Spencer proof uses only the fact that adding one vertex changes χ by at most 1. This is true, but for a *typical* $G(n, p)$, adding a new vertex almost never requires a new color. The new vertex can almost always be accommodated by an existing color class, because a random graph has large independent sets available. Only a small fraction of vertices, roughly $O(\sqrt{n}/\log n)$ of them, actually force χ to increase.