UNIVERSAL POISSON AND NORMAL LIMIT THEOREMS IN GRAPH COLORING PROBLEMS WITH CONNECTIONS TO EXTREMAL COMBINATORICS

BHASWAR B. BHATTACHARYA, PERSI DIACONIS, AND SUMIT MUKHERJEE

ABSTRACT. This paper proves limit theorems for the number of monochromatic edges in uniform random colorings of general random graphs. The limit theorems are universal depending solely on the limiting behavior of the ratio of the number of edges in the graph and the number of colors, and works for any graph sequence deterministic or random. The proofs are based on moment calculations which relates to results of Erdős and Alon on extremal subgraph counts. As a byproduct of our calculations a simple new proof of a result of Alon, estimating the number of isomorphic copies of a cycle of given length in graphs with fixed number of edges, is presented.

1. INTRODUCTION

Suppose the vertices of a finite graph $G = (V, E)$, with $|V| = n$, are colored independently and uniformly at random with c colors. The probability that the resulting coloring has no monochromatic edge, that is, it is a proper coloring is $\chi_G(c)/c^n$, where $\chi_G(c)$ denotes the number of proper colorings of G using c-colors. The function χ_G is the chromatic polynomial of G, and is a central object in graph theory [15] with several interesting unsolved problems [23, 24]. A natural generalization of this is to consider a general coloring distribution $\underline{p} = (p_1, p_2, \ldots, p_c)$, that is, the probability a vertex is colored with color $a \in [c]$ is p_a independent from the colors of the other vertices, where $p_a \geq 0$, and $\sum_{a=1}^{c} p_a = 1$. Define $P_G(\underline{p})$ to be the probability that G is properly colored. $P_G(\underline{p})$ is related to Stanley's generalized chromatic polynomial [32], and under the uniform coloring distribution it is precisely the proportion of proper c-colorings of G. Recently, Fadnavis [18] proved that $P_G(\underline{p})$ is Schur-convex for every fixed c, whenever the graph G is claw-free, that is, G has no induced $K_{1,3}$. This implies that for claw-free graphs, the probability it is properly colored is maximized under the uniform distribution, that is, $p_a = 1/c$ for all $a \in [c]$.

The Poisson limit theorems for the number of monochromatic subgraphs in a random coloring of a graph sequence G_n are applicable when the number of colors grow in an appropriate way compared to the number of certain specific subgraphs in G_n. Barbour et al. [5] used Stein's method to show that the number of monochromatic edges for the complete graph converges to a Poisson random variable. Arratia et al. [4] used Stein's method based on dependency graphs to prove Poisson approximation theorems for the number of monochromatic cliques in a uniform coloring of the complete graph (see also Chatterjee et al. [10]). Poisson limit theorems for the number of general monochromatic subgraphs in a random coloring of a graph sequence was studied by Cerquetti and Fortini [9], again using Stein's method. They assumed that the distribution of colors was exchangeable and proved that the number of copies of any particular monochromatic subgraph converges in distribution to a mixture of Poissons. However, all these results require some technical conditions on the subgraph counts and the coloring probabilities for the error terms in

2010 *Mathematics Subject Classification.* 05C15, 60C05, 60F05, 05D99.

Key words and phrases. Combinatorial probability, Extremal combinatorics, Graph coloring, Limit theorems.

Stein's method to vanish. In particular, while counting the number of monochromatic edges, the conditions depend on the number of edges and 2-stars, and the coloring probabilities. In this paper we show that these extra conditions are redundant under the uniform coloring scheme; the limiting behavior of the number of monochromatic edges is solely governed by the limit of the ratio of the number of edges in the graph and the number of colors.

1.1. **Universal Limit Theorems For Monochromatic Edges.** Let $\mathscr{G}_n$ denote the space of all simple undirected graphs on n vertices labeled by $[n] := \{1, 2, \cdots, n\}$. Given a graph $G_n \in \mathscr{G}_n$ with adjacency matrix $A(G_n) = ((A_{ij}(G_n)))_{1\leq i,j\leq n}$, denote by $V(G_n)$ the set of vertices, and by $E(G_n)$ the set of edges of G_n, respectively. Let $\underline{p} = (p_1, p_2, \ldots, p_c)$ be a probability vector, that is, $p_a \geq 0$, and $\sum_{a=1}^{c} p_a = 1$. The vertices of G_n are colored with c colors as follows:

$$\mathbb{P}(v \in V(G_n) \text{ is colored with color } a \in \{1, 2, \ldots, c\}|G_n) = p_a,$$

independent from the other vertices. The coloring distribution is said to be *uniform* whenever $p_a = 1/c$, for all $a \in [c]$. If Y_i is the color of vertex i, then

$$N(G_n) := \sum_{1\leq i<j\leq n} A_{ij}(G_n)\mathbf{1}\{Y_i = Y_j\} = \sum_{(i,j)\in E(G_n)} \mathbf{1}\{Y_i = Y_j\}, \tag{1.1}$$

denotes the number of monochromatic edges in the graph G_n. Note that $\mathbb{P}(N(G_n) = 0)$ is the probability that G_n is properly colored. We study the limiting behavior of $N(G_n)$ as the size of the graph becomes large, allowing the graph itself to be random, under the assumption that the joint distribution of $(A(G_n), \underline{Y_n})$ is mutually independent, where $\underline{Y_n} = (Y_1, Y_2, \ldots, Y_n)$ are i.i.d. random variables with $\mathbb{P}(Y_1 = a) = p_a$, for all $a \in [c]$. Note that this setup includes the case where $\{G_1, G_2, \ldots\}$ is a deterministic (non-random) graph sequence, as well.

An application of the easily available version of Stein's method, similar to that in Cerquetti and Fortini [9], gives a general limit theorem for $N(G_n)$ that works for all color distributions (Theorem 2.1). However, this result, like all other similar results in the literature, requires several conditions involving the number of edges and 2-stars in the graph G_n, even when the coloring scheme is uniform. One of the main contribution of this paper is in showing that these extra conditions are, in fact, redundant under the uniform coloring scheme, and the limiting behavior is solely governed by the limit of $|E(G_n)|/c$.

Theorem 1.1. *Let $G_n \in \mathscr{G}_n$ be a random graph sampled according to some probability distribution over $\mathscr{G}_n$. Then under the uniform coloring distribution, that is, $p_a = 1/c$, for all $a \in [c]$, the following is true:*

$$N(G_n) \xrightarrow{\mathscr{D}} \begin{cases} 0 & \text{if} \quad \frac{1}{c}\cdot|E(G_n)| \xrightarrow{\mathscr{P}} 0, \\ \infty & \text{if} \quad \frac{1}{c}\cdot|E(G_n)| \xrightarrow{\mathscr{P}} \infty, \\ W & \text{if} \quad \frac{1}{c}\cdot|E(G_n)| \xrightarrow{\mathscr{D}} Z; \end{cases}$$

where $\mathbb{P}(W = k) = \frac{1}{k!}\mathbb{E}(e^{-Z}Z^k)$. In other words, W is distributed as a mixture of Poisson random variables mixed over the random variable Z.

Theorem 1.1 is universal because it only depends on the limiting behavior of $|E(G_n)|/c$ and it works for any graph sequence $\{G_n\}_{n\geq 1}$, deterministic or random. The theorem is proved using the method of moments, that is, the conditional moments of $N(G_n)$ are compared with conditional moments of the random variable $M(G_n) := \sum_{1\leq i<j\leq n} A_{ij}(G_n)Z_{ij}$, where $\{Z_{ij}\}_{(i,j)\in E(G_n)}$ are independent $\mathrm{Ber}(1/c)$. The combinatorial quantity to bound during the moment calculations is the number of isomorphic copies of a graph H in another graph G, to be denoted by $N(G, H)$. Using properties of the adjacency matrix of G we estimate $N(G, H)$, when $H = C_g$ is a g-cycle, This

result is then used to show the asymptotic closeness of the conditional moments of $N(G_n)$ and $M(G_n)$.

Theorem 1.1 asserts that if $\frac{1}{c}|E(G_n)| \stackrel{\mathscr{P}}{\to} \infty$, then $N(G_n)$ goes to infinity in probability. Since a Poisson random variable with mean growing to infinity converges to a standard normal distribution after centering and scaling by the mean and the standard deviation, it is natural to wonder whether the same is true for $N(G_n)$. This is not true in general if $|E(G_n)|/c$ goes to infinity, with c fixed. Proposition 6.1 shows that the limiting distribution of the number of monochromatic edges in the complete graph properly scaled is asymptotically a chi-square with $(c-1)$ degrees of freedom, when c is fixed. On the other hand, if $c \to \infty$, Stein's method for normal approximation can be used to prove the asymptotic normality of $N(G_n)$. However, as before, in the off-the-shelf version of Stein's method an extra condition is needed on the structure of the graph, even under the uniform coloring scheme . Nevertheless, as in the Poisson limit theorem, the normality of the standardized random variable $N(G_n)$ is universal and can also be proved by a method of moments argument.

Theorem 1.2. *Let $G_n \in \mathscr{G}_n$ be a random graph sampled according to some probability distribution over $\mathscr{G}_n$, and $N(G_n)$ as defined before. Then for any uniform c-coloring of G_n, with $c \to \infty$ and $|E(G_n)|/c \stackrel{\mathscr{P}}{\to} \infty$,*

$$Z_n := \frac{1}{\sqrt{|E(G_n)|/c}} \left(N(G_n) - \frac{|E(G_n)|}{c} \right) \stackrel{\mathscr{D}}{\to} N(0,1).$$

In the proof of Theorem 1.2 the conditional central moments of $N(G_n)$ are compared with the conditional central moments of $M(G_n)$. In this case a combinatorial quantity involving the number of multi-subgraphs of G_n show up. Bounding this quantity requires extensions of Alon's [2, 3] results to multi-graphs and leads to results in graph theory which may be of independent interest.

1.2. **Connections to Extremal Combinatorics.** The combinatorial quantity that shows up in the moment computations for the Poisson limit theorem is $N(G, H)$, the number of isomorphic copies of a graph H in another graph G. The quantity $N(\ell, H) := \sup_{G:|E(G)|=\ell} N(G, H)$ is a well-known object in extremal graph theory that was first studied by Erdős [17] and later by Alon [2, 3]. Alon [2] showed that for any simple graph H there exists a graph parameter $\gamma(H)$ such that $N(\ell, H) = \Theta(\ell^{\gamma(H)})$. Friedgut and Kahn [19] extended this result to hypergraphs and identified the exponent $\gamma(H)$ as the fractional stable number of the hypergraph H. Alon's result can be used to obtain a slightly more direct proof of Theorem 1.1. However, our estimates of $N(G, C_g)$ using the spectral properties of G lead to a new and elementary proof of the following result of Alon [2]:

Theorem 1.3 (Theorem B, Alon [2])**.** *If H has a spanning subgraph which is a disjoint union of cycles and isolated edges, then*

$$N(\ell, H) = (1 + O(\ell^{-1/2})) \cdot \frac{1}{|Aut(H)|} \cdot (2\ell)^{|V(H)|/2},$$

where $|Aut(H)|$ denotes the number of automorphisms of H.

The above theorem calculates the exact asymptotic behavior of $N(\ell, H)$ for graphs H which have a spanning subgraph consisting of a disjoint union of cycles and isolated edges. There are only a handful of graphs for which such exact asymptotics are known [2, 3]. Alon's proof in [2] uses a series of combinatorial lemmas. We hope the short new proof presented in this paper is of independent interest.

The quantity $\gamma(H)$ is a well studied object in graph theory and discrete optimization and is related to the fractional stable set polytope [30]. While proving Theorems 1.1 and 1.2 we discover several new facts about the exponent $\gamma(H)$, which might be of independent interest in graph theory. Alon [3] showed that $\gamma(H) \leq |E(H)|$, and the equality holds if and only if H is a disjoint union of stars. This is improved to $\gamma(H) \leq |V(H)| - \nu(H)$, where $\nu(H)$ is the number of connected components of H and the condition for equality remains the same. This is proved in Corollary 5.2 and used later to give an alternative proof of Theorem 1.1. In fact, the universality of the Poisson limit necessitates $\gamma(H) < |V(H)| - \nu(H)$ for all graphs with a cycle.

In a similar manner, the universal normal limit leads to the following interesting observation about $\gamma(H)$. Suppose H has no isolated vertices: if $\gamma(H) > \frac{1}{2}|E(H)|$, then H has a vertex of degree 1 (Lemma 6.2). This result is true for simple graphs as well as multi graphs (with a similar definition of γ for multi-graphs). This result is sharp, in the sense that there are simple graphs with no leaves such that $\gamma(H) = |E(H)|/2$. Even though this result can be proved easily from the definition of $\gamma(H)$, it is a fortunate phenomenon, as it is exactly what is needed for the proof of universal normality.

1.3. **Other Monochromatic Subgraphs.**

Theorems 1.1 and 1.2 determine the universal asymptotic behavior of the number of monochromatic edges under independent and uniform coloring of the vertices. However, the situation while counting copies of other monochromatic subgraphs is quite different. Even under uniform coloring, the limit need not be Poisson mixture. This is illustrated in the following proposition where we show that the number of monochromatic r-stars in a uniform coloring of an n-star converges to a polynomial in Poissons, which is not a Poisson mixture.

Proposition 1.4. *Let $G_n = K_{1,n-1}$, with vertices labeled by $[n]$. Under the uniform coloring scheme, the random variable $T_{r,n}$ which counts the number of monochromatic r-stars in G_n satisfies:*

$$T_{r,n} \stackrel{\mathscr{D}}{\to} \begin{cases} 0 & \text{if } \frac{n}{c} \to 0, \\ \infty & \text{if } \frac{n}{c} \to \infty, \\ \frac{X(X-1)\cdots(X-r+1)}{r!} & \text{if } \frac{n}{c} \to \lambda, \end{cases}$$

where $X \sim Poisson(\lambda)$.

Examples with other monochromatic subgraphs are also considered and several interesting observations are reported. We construct a graph G_n where the number of monochromatic g-cycles ($g \geq 3$) in a uniform c-coloring of G_n converges in distribution to a non-trivial mixture of Poisson even when $|N(G_n, C_g)|/c^{g-1}$ converges to a fixed number λ. This is in contrast to the situation for edges, where the number of monochromatic edges converges to $Poisson(\lambda)$ whenever $|E(G_n)|/c \to \lambda$. We believe that a Poisson-mixture universality holds for cycles as well, that is, the number of monochromatic g-cycles in a uniform random coloring of any graph sequence G_n converges in distribution to a random variable which is mixture of Poisson, whenever $|N(G_n, C_g)|/c^{g-1} \to \lambda > 0$.

1.4. **Organization of the Paper.**

The paper is organized as follows: Section 2 discusses Stein's method approach for studying the limiting behavior of $N(G_n)$. Section 3, proves Theorem 1.1 and Section 4 illustrates it for various ensembles of fixed and random graphs. Section 5 discusses the connections with the subgraph counting problem from extremal combinatorics and gives a new proof of Theorem 1.3. The proof of Theorem 1.2, where the universal normality of $N(G_n)$ is established, is in Section 6. Finally, Section 7 proves Proposition 1.4, considers other examples about counting monochromatic cycles, and discusses possible directions for future research. An appendix gives details on conditional and unconditional convergence of random variables.

2. General Poisson Approximation Using Stein's Method

Cerquetti and Fortini [9] proved a Poisson limit theorem for the number of monochromatic subgraphs in a random coloring of a graph sequence using Stein's method, under the assumption that the distribution of the colors is exchangeable. A similar application of Stein's method gives a general Poisson limit theorem for $N(G_n)$ as well.

Theorem 2.1. *Let $G_n \in \mathscr{G}_n$ be a random graph sampled according to some distribution, and Z be a non-negative random variable. If $T(G_n)$ denotes the set of 2-stars in G_n, then the following is true*

$$\mathbb{P}(N(G_n)=0) \rightarrow \begin{cases} 1 & \text{if} \quad ||\underline{p}||_2|E(G_n)| \xrightarrow{\mathscr{P}} 0, \\ 0 & \text{if} \quad ||\underline{p}||_2|E(G_n)| \xrightarrow{\mathscr{P}} \infty, (||\underline{p}||_3+||\underline{p}||_2^2)|T(G_n)| = o_P(||\underline{p}||_2^2|E(G_n)|), \\ \mathbb{E}(e^{-Z}) & \text{if} \quad ||\underline{p}||_2 \rightarrow 0, ||\underline{p}||_2|E(G_n)| \xrightarrow{\mathscr{D}} Z, (||\underline{p}||_3+||\underline{p}||_2^2)|T(G_n)| \xrightarrow{\mathscr{P}} 0, \end{cases}$$

where $||\underline{p}||_r = \sum_{k=1}^c p_k^r$, with $r \in \mathbb{N}$. More generally, in the third case $N(G_n) \xrightarrow{\mathscr{D}} W$, where W is a non negative integer valued random variable with $\mathbb{P}(W=k)=\frac{1}{k!}\mathbb{E}(e^{-Z}Z^k)$, that is, W is distributed as a Poisson random variable with parameter Z.

Proof. Note that $\mathbb{P}(N(G_n)>0|G_n) \leq \mathbb{E}(N(G_n)|G_n) = ||\underline{p}||_2|E(G_n)|$, from which the first case follows. For the second part, note that

$$\begin{aligned} \mathbb{E}(N(G_n)^2|G_n) &= \sum_{e_1=(i,j),e_2=(k,\ell)} \mathbf{1}\{Y_i=Y_j\}\mathbf{1}\{Y_k=Y_\ell\} \\ &= ||\underline{p}||_2|E(G_n)| + ||\underline{p}||_3|T(G_n)| + ||\underline{p}||_2^2(|E(G_n)|^2 - |T(G_n)| - |E(G_n)|) \\ &= (||\underline{p}||_2 - ||\underline{p}||_2^2)|E(G_n)| + (||\underline{p}||_3 - ||\underline{p}||_2^2)|T(G_n)| + ||\underline{p}||_2^2|E(G_n)|^2 \\ &\leq ||\underline{p}||_2|E(G_n)| + (||\underline{p}||_3 + ||\underline{p}||_2^2)|T(G_n)| + ||\underline{p}||_2^2|E(G_n)|^2 \end{aligned}$$

The result then follows from the inequality

$$\mathbb{P}(N(G_n)>0|G_n) \geq \frac{\mathbb{E}(N(G_n)|G_n)^2}{\mathbb{E}(N(G_n)^2|G_n)}, \tag{2.1}$$

and the given conditions.

For the third case, using [10, Theorem 15] with $Z_n := |E(G_n)|||p||_2$ we have for all $k \in \mathbb{N}$,

$$\left|\mathbb{P}(N(G_n)=k|G_n) - \frac{1}{k!}e^{-Z_n}Z_n^k\right| \leq 2(||\underline{p}||_3 + ||\underline{p}||_2^2)|T(G_n)| + 2||\underline{p}||_2^2|E(G_n)|.$$

The right hand side converges to 0 in probability from given conditions, and by the bounded convergence theorem the conclusion follows. □

For the uniform coloring scheme, $\mathbb{E}(N(G_n)|G_n) = |E(G_n)|/c$ and $\mathbb{P}(N(G_n)>0)$ converges to 0 or 1 depending on whether $|E(G_n)|/c$ converges to 0 or infinity, respectively. The Stein method for Poisson approximation in this case has the following neater form, but requires an extra condition on the number of 2-stars in G_n.

Corollary 2.2. *(Uniform Coloring) Let $G \in \mathscr{G}_n$ be a random graph sampled according to some distribution, and Z be a non-negative random variable. Then under the uniform coloring distribution,* *that is, $p_a = 1/c$, for all $a \in [c]$, the following holds*

$$\mathbb{P}(N(G_n) = 0) \to \begin{cases} 1 & \text{if} \quad \frac{1}{c} \cdot |E(G_n)| \xrightarrow{\mathscr{P}} 0, \\ 0 & \text{if} \quad \frac{1}{c} \cdot |E(G_n)| \xrightarrow{\mathscr{P}} \infty \\ \mathbb{E}(e^{-Z}) & \text{if} \quad \frac{1}{c} \cdot |E(G_n)| \xrightarrow{\mathscr{D}} Z, \frac{1}{c^2} \cdot |T(G_n)| \xrightarrow{\mathscr{P}} 0. \end{cases}$$

□

Remark 2.1. The condition $\frac{1}{c^2} \cdot |T(G_n)|$ converges in probability to 0, is a technical condition that is required for proving the above corollary using Stein's method. For an explicit example where the conclusion holds even though assumptions do not, let $G_n = K_{1,n-1}$ be the star graph with vertices labelled by $[n]$ and the central vertex labelled by 1. Let $c = n$, and $p_a = 1/n$, for $a \in [c]$. Then $\frac{1}{c} \cdot |E(G_n)| = 1 - 1/n \to 1$, and $\frac{1}{c^2} \cdot |T(G_n)| = (n-1)(n-2)/n^2 \to 1$. Thus, in this case the last condition of Corollary 2.2 does not hold. If Y_i denotes the color of vertex $i \in [n]$, then $N(G_n) := \sum_{i=2}^n 1\{Y_1 = Y_i\}$. Observe that all the summands are independent given Y_1, and so for any $s \in (0,1)$

$$\mathbb{E}s^{N(G_n)} = \mathbb{E}\prod_{j=2}^n \mathbb{E}(s^{1\{Y_1=Y_j\}}|Y_1) = \Big(1 - \frac{1}{n} + \frac{s}{n}\Big)^{n-1} \to e^{s-1} = \sum_{k=0}^\infty \frac{1}{k!e}s^k.$$

Therefore, $N(G_n)$ converges to a Poisson distribution with parameter 1, which is not predicted by Corollary 2.2. On the other hand, Theorem 1.1, which is proved in the next section, covers this case and all other cases where a Poisson limit theorem for monochromatic edges holds.

3. Universal Poisson Approximation For Uniform Colorings: Proof of Theorem 1.1

This section determines the limiting behavior of $\mathbb{P}(N(G_n) = 0)$ under minimal conditions. Using the method of moments, we show that $N(G_n)$ has a universal threshold which depends only on the limiting behavior of $|E(G_n)|/c$, and a Poisson limit theorem holds at the threshold.

Let $G_n \in \mathscr{G}_n$ be a random graph sampled according to some probability distribution. For every $n \in \mathbb{N}$ fixed, for $(i,j) \in E(G_n)$ define the collection of random variables

$$\{Z_{ij}, (i,j) \in E(G_n)\}, \text{ where } Z_{ij} \text{ are i.i.d. Ber}(1/c), \text{ and set } M(G_n) := \sum_{1 \le i < j \le n} A_{ij}(G_n) Z_{ij}. \tag{3.1}$$

The proof of Theorem 1.1 is given in two parts: The first part compares the conditional moments of $N(G_n)$ and $M(G_n)$, given the graphs G_n, showing that they are asymptotically close, when $|E(G_n)|/c \xrightarrow{\mathscr{D}} Z$. The second part uses this result to complete the proof of Theorem 1.1 using some technical properties of conditional convergence (see Lemma 8.1).

3.1. Computing and Comparing Moments.

This section is devoted to the computation of conditional moments of $N(G_n)$ and $M(G_n)$, and the comparison of these. To this end, define for any fixed number k, $A \lesssim_k B$ as $A \le C(k)B$, where $C(k)$ is a constant that depends only on k. Let $G_n \in \mathscr{G}_n$ be a random graph sampled according to some probability distribution. For any fixed subgraph H of G_n, let $N(G_n, H)$ be the number of isomorphic copies of H in G_n, that is,

$$N(G_n, H) := \sum_{S \subset E(G_n): |S| = |E(H)|} \mathbf{1}\{G_n[S] = H\},$$

where the sum is over subsets S of $E(G_n)$ with $|S| = |E(H)|$, and $G_n[S]$ is the subgraph of G_n induced by the edges of S.

Lemma 3.1. *Let $G_n \in \mathscr{G}_n$ be a random graph sampled according to some probability distribution. For any $k \geq 1$, let $\mathcal{H}_k$ to be the collection of all graphs with at most k edges and no isolated vertices. Then*

$$|\mathbb{E}(N(G_n)^k|G_n) - \mathbb{E}(M(G_n)^k|G_n)| \quad \lesssim_k \quad \sum_{\substack{H \in \mathcal{H}_k, \\ H \text{ has a cycle}}} N(G_n, H) \cdot \frac{1}{c^{|V(H)|-\nu(H)}}, \tag{3.2}$$

where $\nu(H)$ is the number of connected components of H.

Proof. Using the multinomial expansion and the definition of $\mathcal{H}_k$,

$$\mathbb{E}(N(G_n)^k|G_n) \quad = \quad \sum_{(i_1,j_1)\in E(G_n)} \sum_{(i_2,j_2)\in E(G_n)} \cdots \sum_{(i_k,j_k)\in E(G_n)} \mathbb{E}\left(\prod_{r=1}^k \mathbf{1}\{Y_{i_r} = Y_{j_r}\}\Big| G_n\right) \tag{3.3}$$

Similarly,

$$\mathbb{E}(M(G_n)^k|G_n) = \sum_{(i_1,j_1)\in E(G_n)} \sum_{(i_2,j_2)\in E(G_n)} \cdots \sum_{(i_k,j_k)\in E(G_n)} \mathbb{E}\left(\prod_{r=1}^k Z_{i_r j_r}\right), \tag{3.4}$$

If H is the simple subgraph of G_n induced by the edges $(i_1, j_1), (i_2, j_2), \ldots (i_k, j_k)$. Then

$$\mathbb{E}\left(\prod_{r=1}^k \mathbf{1}\{Y_{i_r} = Y_{j_r}\}\Big| G_n\right) = \frac{1}{c^{|V(H)|-\nu(H)}} \text{ and } \mathbb{E}\left(\prod_{r=1}^k Z_{i_r j_r}\right) = \frac{1}{c^{|E(H)|}}.$$

The result now follows by taking the difference of (3.3) and (3.4), and noting that in any graph H, $|E(H)| \geq |V(H)| - \nu(H)$ and equality holds if and only if H is a forest. □

Remark 3.1. Observe that if G_n is a forest (disjoint union of trees), then by the above lemma, $\mathbb{E}(M(G_n)^k|G_n) = \mathbb{E}(N(G_n)^k|G_n)$, and consequently the laws of $M(G_n)$ and $N(G_n)$ are exactly the same when G_n is a forest. In particular, this means that for a forest

$$\mathbb{P}(N(G_n) = 0|G_n) = \left(1 - \frac{1}{c}\right)^{|E(G_n)|},$$

for every c, n. Note that under the uniform coloring scheme $\mathbb{P}(N(G_n) = 0|G_n) = c^{-n}\chi(G_n, c)$, where $\chi(G_n, c)$ is the number of proper c-coloring of the graph G_n. Therefore, determining exact or asymptotic expressions of $\mathbb{P}(N(G_n) = 0|G_n)$, when c is a fixed constant, for a graph G amounts to counting the number of proper c-colorings of G. Though this is easy for trees, in general it is a notoriously hard problem, and is known to be #P-complete (refer to the survey by Frieze and Vigoda [20] and the references therein for the various hardness results and approximate counting techniques related to proper graph colorings).

Lemma 3.1 shows that bounding the difference of the conditional moments of $N(G_n)$ and $M(G_n)$ entails bounding $N(G_n, H)$, the number of copies of a subgraph H in G_n. The next lemma estimates the number of copies of a cycle C_g in G_n.

Lemma 3.2. *For $g \geq 3$ and $G_n \in \mathscr{G}_n$ let $N(G_n, C_g)$ be the number of g-cycles in G_n. Then*

$$N(G_n, C_g) \leq \frac{1}{2g} \cdot (2|E(G_n)|)^{g/2}.$$

Proof. Let $A := A(G_n)$ be the adjacency matrix of G_n. Note that $\sum_{i=1}^n \lambda_i^2 = \mathrm{tr}(A^2) = 2|E(G_n)|$, where $|\lambda_1| \geq \cdots \geq |\lambda_n|$ are the ordered absolute eigenvalues of A. Note that $\mathrm{tr}(A^g)$ counts the number of walks of length g in G_n, and so each cycle in G_n is counted $2g$ times. Thus, as an upper bound,

$$N(G_n, C_g) \leq \frac{1}{2g} \cdot \mathrm{tr}(A^g) = \frac{1}{2g} \cdot \sum_{i=1}^n \lambda_i^g \leq \frac{1}{2g} \cdot \sup_i |\lambda_i|^{g-2} \sum_{i=1}^n \lambda_i^2 \leq \frac{1}{2g} \cdot (2|E(G_n)|)^{g/2},$$

where the last step uses $\sup_i |\lambda_i|^{g-2} = (\sup_i |\lambda_i|^2)^{g/2-1} \leq (2|E(G_n)|)^{g/2-1}$. □

Remark 3.2. In extremal graph theory, the study of $N(G, H)$, for arbitrary graphs G and H, was initiated by Erdős [17], and later carried forward by Alon [2, 3]. In fact, Lemma 3.2 is a special case of Theorem B of Alon [2]. In Section 5 the proof of Lemma 3.2 is used to give a new and short proof of Theorem B.

For a given simple graph H, the notation $A \lesssim_H B$ will mean $A \leq C(H) \cdot B$, where $C(H)$ is a constant that depends only on H. Lemma 3.2 gives a bound on $N(G_n, H)$ in terms of $|E(G_n)|$ for arbitrary subgraphs H of G_n.

Lemma 3.3. *For any fixed connected subgraph H, let $N(H, G_n)$ be the set of copies of H in G_n. Then*

$$N(G_n, H) \lesssim_H |E(G_n)|^{|V(H)|-1}. \tag{3.5}$$

Furthermore, if H has a cycle of length $g \geq 3$, then

$$N(G_n, H) \lesssim_H |E(G_n)|^{|V(H)|-g/2}. \tag{3.6}$$

Proof. The first bound on $N(H, G_n)$ can be obtained by a crude counting argument as follows: First choose an edge of G_n in $E(G_n)$ which fixes 2 vertices of H. Then the remaining $|V(H)| - 2$ vertices are chosen arbitrarily from the set of allowed $V(G_n)$ vertices, giving the bound

$$N(G_n, H) \leq 2^{\binom{|V(H)|}{2}} |E(G_n)| \binom{|V(G_n)|}{|V(H)|-2} \leq 2^{\binom{|V(H)|}{2}} |E(G_n)| \binom{2|E(G_n)|}{|V(H)|-2} \lesssim_H |E(G_n)|^{|V(H)|-1},$$

where we have used the fact that the number of graphs on $|V(H)|$ vertices is at most $2^{\binom{|V(H)|}{2}}$.

Next, suppose that H has a cycle of length $g \geq 3$. Choosing a cycle of length g arbitrarily from G_n, there are $|V(G_n)|$ vertices from which the remaining $|V(H)| - g$ vertices are chosen arbitrarily. Since the edges among these vertices are also chosen arbitrarily, the following crude upper bound holds

$$\begin{aligned} N(G_n, H) \leq 2^{\binom{|V(H)|}{2} - g} N(G_n, C_g) \binom{2|E(G_n)|}{|V(H)|-g} & \lesssim_H N(G_n, C_g) |E(G_n)|^{|V(H)|-g} \\ & \lesssim_H |E(G_n)|^{|V(H)|-g/2}. \end{aligned} \tag{3.7}$$

where the last step uses Lemma 3.2. □

The above lemmas, imply the most central result of this section: the conditional moments of $M(G_n)$ and $N(G_n)$ are asymptotically close, whenever $|E(G_n)|/c \xrightarrow{\mathscr{D}} Z$.

Lemma 3.4. *Let $M(G_n)$ and $N(G_n)$ be as defined before, with $|E(G_n)|/c \xrightarrow{\mathscr{D}} Z$, then for every fixed $k \geq 1$,*

$$|\mathbb{E}(N(G_n)^k | G_n) - \mathbb{E}(M(G_n)^k | G_n)| \xrightarrow{\mathscr{P}} 0.$$

Proof. By Lemma 3.1

$$|\mathbb{E}(N(G_n)^k|G_n) - \mathbb{E}(M(G_n)^k|G_n)| \leq \sum_{\substack{H\in\mathcal{H}_k,\\ H \text{ has a cycle}}} N(G_n, H) \cdot \frac{1}{c^{|V(H)|-\nu(H)}},$$

where $\nu(H)$ is the number of connected components of H. As the sum over $H \in \mathcal{H}_k$ is a finite sum, it suffices to show that for a given $H \in \mathcal{H}_k$ with a cycle $N(G_n, H) = o_P(c^{|V(H)|-\nu(H)})$.

To this end, fix $H \in \mathcal{H}_k$ and let $H_1, H_2, \ldots, H_{\nu(H)}$ be the connected components of H. Without loss of generality, suppose the girth of G, $g(G) = g(H_1) = g \geq 3$. Lemma 3.3 then implies that

$$\begin{aligned} N(G_n, H) \leq \prod_{i=1}^{\nu(H)} |N(H_i, G_n) &\lesssim_H \quad E(G_n)^{V(H_1)-g/2} \prod_{i=2}^{\nu(H)} E(G_n)^{V(H_i)-1} \\ &\lesssim_H \qquad E(G_n)^{V(H)-\nu(H)+1-g/2} \end{aligned} \tag{3.8}$$

which is $o_p(c^{V(H)-\nu(H)})$ since $g/2 - 1 > 0$. $\square$

Remark 3.3. The above proof shows that the error rate between the difference of conditional moments is better when g is larger, that is, the Poisson approximation is more accurate on graphs with large girth. In his 1981 paper, Alon [2] proved that for every fixed H, $N(G_n, H) \leq c_2(H)|E(G_n)|^{\gamma(H)}$, where $c_2(H)$ and $\gamma(H)$ are constants depending only on the graph H. Section 5 below shows that by plugging in this estimate and using the property of $\gamma(H)$, it is possible to obtain a direct proof 3.4 that obviates the calculations in Lemma 3.2 and Lemma 3.3. This gives slightly better error rates between the difference of the conditional moments.

3.2. **Completing the Proof of Theorem 1.1.** The results from the previous section are used here to complete the proof of Theorem 1.1. The three different regimes of $|E(G_n)|/c$ are treated separately as follows:

3.2.1. $\frac{1}{c} \cdot |E(G_n)| \stackrel{\mathscr{P}}{\to} 0$. The result follows directly from Corollary 2.2.

3.2.2. $\frac{1}{c} \cdot |E(G_n)| \stackrel{\mathscr{P}}{\to} \infty$. It follows from the proof of Theorem 2.1 that $\mathbb{E}(N_n^2|G_n)/\mathbb{E}(N_n|G_n)^2 \to 1$. This implies that $N_n/\mathbb{E}(N_n|G_n)$ converges in probability to 1, and so N_n converges to ∞ in probability, as $\mathbb{E}(N_n|G_n) = \frac{1}{c} \cdot |E(G_n)| \stackrel{\mathscr{P}}{\to} \infty$.

3.2.3. $\frac{1}{c} \cdot |E(G_n)| \stackrel{\mathscr{D}}{\to} Z$*, where* Z *is some random variable.* In this regime the limiting distribution of $N(G_n)$ is a mixture of Poisson. As the Poisson distribution can be uniquely identified by moments, from Lemma 3.4 it follows that conditional on $\{|E(G_n)|/c \to \lambda\}$, $N(G_n)$ converges to $Poisson(\lambda)$ for every $\lambda > 0$. However, this does not immediately imply the unconditional convergence of $N(G_n)$ to a mixture of Poisson. In fact, a technical result, detailed in Lemma 8.1, and convergence of $M(G_n)$ to a Poisson mixture is necessary to complete the proof.

To begin with, recall that a random variable X is a mixture of Poisson with mean Z, to be denoted as $Poisson(Z)$, if there exists a non-negative random variable Z such that

$$\mathbb{P}(X = k) = \mathbb{E}\left(\frac{1}{k!}e^{-Z}Z^k\right).$$

The following lemma shows that $M(G_n)$ converges to $Poisson(Z)$ and satisfies the technical condition needed to apply Lemma 8.1.

Lemma 3.5. *Let $M(G_n)$ be as defined in (3.1) and $\frac{1}{c}\cdot|E(G_n)| \xrightarrow{\mathscr{D}} Z$. Then $M(G_n) \xrightarrow{\mathscr{D}} Poisson(Z)$, and further for any $\epsilon > 0, t \in \mathbb{R}$,*

$$\limsup_{k\to\infty}\limsup_{n\to\infty} \mathbb{P}\left(\left|\frac{t^k}{k!}\mathbb{E}(M(G_n)^k|G_n)\right| > \epsilon\right) = 0.$$

Proof. For any $t \in \mathbb{R}$, $\mathbb{E}e^{itM(G_n)} = \mathbb{E}\mathbb{E}(e^{itM(G_n)}|G_n) = \mathbb{E}\left(1 - \frac{1}{c} + \frac{e^{it}}{c}\right)^{|E(G_n)|} = \mathbb{E}Z_n$, where $Z_n := \left(1 - \frac{1}{c} + \frac{e^{it}}{c}\right)^{|E(G_n)|}$ satisfies $|Z_n| \le 1$. Since

$$\log Z_n = |E(G_n)|\log\left(1 - \frac{1}{c} + \frac{e^{it}}{c}\right) = |E(G_n)|\left(\frac{e^{it}-1}{c} + O\left(\frac{1}{c^2}\right)\right) \xrightarrow{\mathscr{D}} (e^{it}-1)Z,$$

by the dominated convergence theorem $\mathbb{E}e^{itM(G_n)} = \mathbb{E}Z_n \to \mathbb{E}e^{(e^{it}-1)Z}$, which can be easily checked to be the generating function of a random variable with distribution $Poisson(Z)$. Thus, it follows that $M(G_n) \xrightarrow{\mathscr{D}} Poisson(Z)$.

Proceeding to check the second conclusion, recall the standard identity $z^k = \sum_{j=0}^k S(k,j)(z)_j$, where $S(\cdot,\cdot)$ are Stirling numbers of the second kind and $(z)_j = z(z-1)\cdots(z-j+1)$. In the above identity, setting $z = M(G_n)$, taking expectation on both sides conditional on G_n, and using the formula for the Binomial factorial moments,

$$\mathbb{E}(M(G_n)^k|G_n) = \sum_{j=0}^k S(k,j)(|E(G_n)|)_j c^{-j}.$$

The right hand side converges weakly to $\sum_{j=0}^k S(k,j)Z^j$. This is the k-th mean of a Poisson random variable with parameter Z. Using the formula for the Poisson moment generating function, for any $Z \ge 0$ and any $t \in \mathbb{R}$ we have

$$\sum_{k=0}^\infty \frac{t^k}{k!}\sum_{j=0}^k S(k,j)Z^j = e^{Z(e^t-1)} < \infty \implies \frac{t^k}{k!}\sum_{j=0}^k S(k,j)Z^j \xrightarrow{a.s.} 0,$$

as $k \to \infty$. Thus, applying Fatou's Lemma twice gives

$$\limsup_{k\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\left|\frac{t^k}{k!}\mathbb{E}(M(G_n)^k|G_n)\right| > \epsilon\right) \le \limsup_{k\to\infty}\mathbb{P}\left(\left|\frac{t^k}{k!}\sum_{r=0}^k S(k,r)Z^r\right| > \epsilon\right) = 0,$$

and so the proof of the lemma is complete. □

Now, take $U_{n,k} = \mathbb{E}(N(G_n)^k|G_n)$ and $V_{n,k} = \mathbb{E}(M(G_n)^k|G_n)$, and observe that (8.1) and (8.2) hold by Lemma 3.4 and Lemma 3.5, respectively. As $M(G_n)$ converges to $Poisson(Z)$, this implies that $N(G_n)$ converges to $Poisson(Z)$, and the proof of Theorem 1.1 is completed.

Remark 3.4. Theorem 1.1 shows that the limiting distribution of the number of monochromatic edges converges to a Poisson mixture. In fact, Poisson mixtures arise quite naturally in several contexts. It is known that the Negative Binomial distribution is distributed as $Poisson(Z)$, where Z is a Gamma random variable with integer values for the shape parameter. Greenwood and Yule [22] showed that certain empirical distributions of accidents are well-approximated by a Poisson mixture. Le-Cam and Traxler [27] proved asymptotic properties of random variables distributed as mixture of Poisson. Poisson mixtures are widely used in modeling count panel data (refer to the recent paper of Burda et al. [7] and the references therein), and have appeared in other applied problems as well [11].

4. Examples: Applications of Theorem 1.1

In this section we apply Theorem 1.1 to different deterministic and random graph models, and determine the specific nature of the limiting Poisson distribution.

Example 1. (Birthday Problem) When the underlying graph G is the complete graph K_n on n vertices, the above coloring problem reduces to the well-known birthday problem. By replacing the c colors by birthdays, each occurring with probability $1/c$, the birthday problem can be seen as coloring the vertices of a complete graph independently with c colors. The event that two people share the same birthday is the event of having a monochromatic edge in the colored graph. In birthday terms, $\mathbb{P}(N(K_n) = 0)$ is precisely the probability that no two people have the same birthday. Theorem 1.1 says that under the uniform coloring for the complete graph $\mathbb{P}(N(K_n) = 0) \approx e^{-n^2/2c}$. Therefore, the maximum n for which $\mathbb{P}(N(K_n) = 0) \leq 1/2$ is approximately 23, whenever $c = 365$. This reconstructs the classical birthday problem which can also be easily proved by elementary calculations. For a detailed discussion on the birthday problem and its various generalizations and applications refer to [1, 5, 12, 13, 14] and the references therein.

Example 2. (Birthday Coincidences in the US Population) Consider the following question: What is the chance that there are two people in the United States who (a) know each other, (b) have the same birthday, (c) their fathers have the same birthday, (d) their grandfathers have the same birthday, and (e) their great grandfathers have the same birthdays. We will argue that this seemingly impossible coincidence actually happens with almost absolute certainty.

The population of the US is about n=400 million and it is claimed that a typical person knows about 600 people [21, 25]. If the network G_n of 'who knows who' is modeled as an Erdős-Renyi graph, this gives $p = 150 \times 10^{-8}$ and $\mathbb{E}(|E(G_n)|) = 300 \times 4 \times 10^8 = 1.2 \times 10^{11}$. The 4-fold birthday coincidence amounts to $c = (365)^4$ 'colors' and $\lambda = \mathbb{E}(N(G_n)) = \mathbb{E}(|E(G_n)|)/c \approx 6.76$, and using (2.1) the probability of a match is at least $1 - \frac{1}{\lambda} = 85\%$. Moreover, assuming the Poisson approximation, the chance of a match is approximately $1 - e^{-\lambda} \approx 99.8\%$, which means that almost surely there are two friends in the US who have a 4-fold birthday match among their ancestors.

Going back one more generation, we now calculate the probability that there are two friends who have a 5-fold birthday coincidence between their respective ancestors. This amounts to $c = (365)^5$ and Poisson approximation shows that the chance of a match is approximately $1 - e^{-\lambda} \approx 1.8\%$. This implies that even a miraculous 5-fold coincidence of birthdays is actually likely to happen among the people of the US.

Example 3. (Random Regular Graphs) When $\mathscr{G}_n$ consists of the set all d-regular graphs on n vertices and sampling is uniformly on this space, then under the uniform coloring distribution with $c \to \infty$, Theorem 1.1 gives

$$\mathbb{P}(N(G_n) = 0) \to \begin{cases} 1 & \text{if } \frac{nd}{c} \to 0, \\ 0 & \text{if } \frac{nd}{c} \to \infty, \\ e^{-b/2} & \text{if } \frac{nd}{c} \to b. \end{cases}$$

Example 4. (Sparse Inhomogeneous Random Graphs) A general model for sparse random graphs is the following: every edge (i, j) is present independently with probability $\frac{1}{n} \cdot f(\frac{i}{n}, \frac{j}{n})$, for some symmetric continuous function $f : [0, 1]^2 \to [0, 1]$ (see Bollobas et al. [6]). Under the uniform coloring distribution,

$$\frac{1}{n}|E(G_n)| \to \frac{1}{2}\int_0^1 \int_0^1 f(x, y)dxdy.$$

Consequently, theorem 1.1 gives

$$\mathbb{P}(N(G_n)=0) \to \begin{cases} 1 & \text{if } \quad \frac{n}{c} \to 0, \\ 0 & \text{if } \quad \frac{n}{c} \to \infty, \\ e^{-(b/2)\int_0^1\int_0^1 f(x,y)dxdy} & \text{if } \quad \frac{n}{c} \to b. \end{cases}$$

Note that this model includes as a special case the Erdős-Renyi random graphs $G(n, \lambda/n)$, by taking the function $f(x, y) = \lambda$.

Example 5. (Dense graph limits) Limits of dense graphs was developed recently by Lovász and co-authors [28], where a random graph sequence G_n converges in cut-metric to a random symmetric measurable function $W : [0, 1]^2 \mapsto [0, 1]$. Then

$$\frac{1}{n^2}|E(G_n)| \to \frac{1}{2}\int_0^1\int_0^1 W(x,y)dxdy.$$

Thus whenever $c \to \infty$, by Theorem 1.1 we have:

$$\mathbb{P}(N(G_n)=0) \to \begin{cases} 1 & \text{if } \quad \frac{n^2}{c} \to 0, \\ 0 & \text{if } \quad \frac{n^2}{c} \to \infty, \\ \mathbb{E}(e^{-(b/2)\int_0^1\int_0^1 W(x,y)dxdy}) & \text{if } \quad \frac{n^2}{c} \to b. \end{cases}$$

Thus the result holds irrespective of the specific model on random graphs, as long as it converges in the sense of the cut metric. In particular, the result implies directly in the following examples:

- *Inhomogenous random graphs:* Let $f : [0, 1]^2 \to [0, 1]$ be a symmetric continuous function. Consider the random graph model where and edge (i, j) is present with probability $f(\frac{i}{n}, \frac{j}{n})$ and the uniform coloring distribution. Therefore, whenever $c \to \infty$, by Theorem 1.1 we have:

$$\mathbb{P}(N(G_n)=0) \to \begin{cases} 1 & \text{if } \quad \frac{n^2}{c} \to 0, \\ 0 & \text{if } \quad \frac{n^2}{c} \to \infty, \\ e^{-(b/2)\int_0^1\int_0^1 f(x,y)dxdy} & \text{if } \quad \frac{n^2}{c} \to b. \end{cases}$$

  Note that this model includes as a special case the Erdős-Renyi random graphs $G(n, p)$, by taking the function $f(x, y) = p$.
- *Graph Limits:* Let $\{U_i\}_{i=1}^n \stackrel{i.i.d.}{\sim} U(0, 1)$, and let f be a symmetric continuous function, and consider the random graph model where given U the edges are mutually independent, with an edge (i, j) being present with probability $f(U_i, U_j)$. These random graphs also converge to f in probability with respect to the cut metric, and consequently the same conclusions as in the inhomogeneous random graph model hold in this case. Refer to Lovász's recent book [28] for a complete description of the theory of graph limits.

Example 6. (Galton-Watson Trees) Let G_n be a Galton-Watson tree truncated at height n, and let ξ denote a generic random variable from the off-spring distribution. Assume further that $\mu := \mathbb{E}\xi > 1$. This ensures that the total progeny up to time n (which is also the number of edges in G_n) grows with n. Letting $\{Z_i\}_{i=0}^\infty$ denote the size of the i-th generation, the total progeny up to time n is $Y_n := \sum_{i=0}^n Z_i$. Assuming that the population starts with one off-spring at time 0, that is, $Z_0 \equiv 1$, Z_n/μ^n is a non-negative martingale ([16, Lemma 4.3.6]). It converges almost surely to a finite valued random variable Z_∞, by [16, Theorem 4.2.9], which readily implies Y_n/μ^{n+1} converges almost surely to $Z_\infty/(\mu-1)$. Thus, Theorem 1.1 gives

$$\mathbb{P}(N(G_n)=0) \to \begin{cases} 1 & \text{if } \frac{\mu^n}{c} \to 0, \\ 0 & \text{if } \frac{\mu^n}{c} \to \infty, \\ \mathbb{E}e^{-\frac{b\mu}{\mu-1}\cdot Z_\infty} & \text{if } \frac{\mu^n}{c} \to b. \end{cases}$$

Note in passing that $Z_\infty \equiv 0$ if and only if $\mathbb{E}(\xi \log \xi) = \infty$ ([16, Theorem 4.3.10]). Thus, to get a nontrivial limit the necessary and sufficient condition is $\mathbb{E}(\xi \log \xi) < \infty$.

5. Connections to Extremal Graph Theory

In the method of moment calculations of Lemma 3.1, we encounter the quantity $N(G, H)$, the number of isomorphic copies of H in G. More formally, given two graphs $G = (V(G), E(G))$ and $H = (V(H), E(H))$,

$$N(G,H) := \sum_{S \subset E(G): |S| = |E(H)|} \mathbf{1}\{G[S] = H\},$$

where the sum is over subsets S of $E(G)$ with $|S| = |E(H)|$, and $G[S]$ is the subgraph of G induced by the edges of S.

For a positive integer $\ell \geq |E(H)|$, define $N(\ell, H) := \sup_{G:|E(G)|=\ell} N(G, H)$. For the complete graph K_h, Erdős [17] determined $N(\ell, K_h)$, which is also a special case of the Kruskal-Katona theorem, and posed the problem of estimating $N(\ell, H)$ for other graphs H. This was addressed by Alon [2] in 1981 in his first published paper. Alon studied the asymptotic behavior of $N(\ell, H)$ for fixed H, as ℓ tends to infinity. He identified the correct order of $N(\ell, H)$, for every fixed H, by proving that:

Theorem 5.1 (Alon [2])**.** *For a fixed graph H, there exists two positive constants $c_1 = c_1(H)$ and $c_2 = c_2(H)$ such that for all $\ell \geq |E(H)|$,*

$$c_1 \ell^{\gamma(H)} \leq N(\ell, H) \leq c_2 \ell^{\gamma(H)}, \tag{5.1}$$

where $\gamma(H) = \frac{1}{2}(|V(H)| + \delta(H))$, and $\delta(H) = \max\{|S| - |N_H(S)| : S \subset V(H)\}$.

Friedgut and Kahn [19] extended this result to hypergraphs, and identified the correct exponent $\gamma(H)$ as the fractional stable number of the hypergraph H. Using the above theorem or the definition of $\gamma(H)$ it is easy to show that $\gamma(H) \leq |E(H)|$, and the equality holds if and only if H is a disjoint union of stars (Theorem 1, Alon [3]). The following corollary gives a sharpening of Theorem 1 of [3]:

Corollary 5.2. *For every graph H,*

$$\gamma(H) \leq |V(H)| - \nu(H),$$

where $\nu(H)$ is the number of connected components of H. Moreover, the equality holds if and only if H is a disjoint union of stars.

Proof. Suppose $H_1, H_2, \ldots, H_{\nu(H)}$ are the connected components of H. Fix $i \in \{1, 2, \ldots, \nu(H)\}$. Since H_i is connected, for every $S \subset V(H_i)$, $|S| - |N_{H_i}(S)| \leq |V(H_i)| - 2$. This implies that $\delta(H) = \sum_{i=1}^{\nu(H)} \delta(H_i) \leq |V(H)| - 2\nu(H)$, and $\gamma(H) \leq |V(H)| - \nu(H)$.

Now, if H is a disjoint union of stars with $\nu(H)$ connected components, then by Theorem 1 of Alon [3], $\gamma(H) = |E(H)| = |V(H)| - \nu(H)$.

Conversely, suppose that $\gamma(H) = |V(H)| - \nu(H)$. If H has a cycle of length $g \geq 3$, then from (3.8) and Theorem 5.1 $\gamma(H) \leq |V(H)| - \nu(H) + 1 - g/2 < |V(H)| - \nu(H)$. Therefore, H has no cycle, that is, it is a disjoint union of trees. This implies that $\gamma(H) = |V(H)| - \nu(H) = |E(H)|$, and from Theorem 1 of Alon [3], H is a disjoint union of stars. $\square$

5.1. **Another Proof of Theorem 1.1 Without Lemma 3.1.** Theorem 5.1 and Corollary 5.2 give a direct proof of Lemma 3.4, which does not require the subgraph counting Lemma 3.1.

With $M(G_n)$ and $N(G_n)$ be as defined before, and $|E(G_n)|/c \xrightarrow{\mathscr{D}} Z$, for every fixed $k \geq 1$,

$$|\mathbb{E}(N(G_n)^k|G_n) - \mathbb{E}(M(G_n)^k|G_n)| \lesssim_k \sum_{\substack{H \in \mathcal{H}_k, \\ H \text{ has a cycle}}} N(G_n, H) \cdot \frac{1}{c^{|V(H)|-\nu(H)}} \lesssim_k \sum_{\substack{H \in \mathcal{H}_k, \\ H \text{ has a cycle}}} \cdot \frac{E(G_n)^{\gamma(H)}}{c^{|V(H)|-\nu(H)}},$$

where the last inequality follows from Theorem 5.1. As the sum is over all graphs H which are not a forest, it follows from Corollary 5.2 that $\gamma(H) < |V(H)| - \nu(H)$. Therefore, every term in the sum goes to zero as $n \to \infty$, and, since $H \in \mathcal{H}_k$ is a finite sum, Lemma 3.4 follows.

5.2. **A New Proof of Theorem 1.3 Using Lemma 3.2.** This section gives a short proof of Theorem 1.3 using Lemma 3.2. The proof uses spectral techniques and is quite different from the proof in Alon [2].

5.2.1. *Proof of Theorem 1.3.* Let F be the spanning subgraph H, and let $F_1, F_2, \ldots, F_q$, be the connected components of F, where each F_i is a cycle or an isolated edge, for $i \in \{1, 2, \ldots, q\}$. Consider the following two cases:

Case **1:** F_i is an isolated edge. Then for any graph G with $|E(G)| = \ell$,

$$N(G, F_i) = \ell = \frac{1}{|Aut(F_i)|} \cdot (2\ell)^{|V(F_i)|/2}. \tag{5.2}$$

Case **2:** F_i is a cycle of length $g \geq 3$. Then by Lemma 3.2

$$N(G, F_i) \leq \frac{1}{2g} \cdot (2\ell)^{g/2} = \frac{1}{|Aut(F_i)|} \cdot (2\ell)^{|V(F_i)|/2}, \tag{5.3}$$

for any graph G with $|E(G)| = \ell$.

Now, (5.2) and (5.3) implies that

$$N(G, F) \leq \prod_{i=1}^{q} N(G, F_i) \leq \frac{1}{\prod_{i=1}^{q} |Aut(F_i)|} \cdot (2\ell)^{|V(H)|/2} = \frac{1}{|Aut(F)|} \cdot (2\ell)^{|V(H)|/2}, \tag{5.4}$$

for all graphs G with $|E(G)| = \ell$.

Let $v = |V(H)| = |V(F)|$ and define $x(H, F)$ to be the number of subgraphs of K_v, isomorphic to H, that contain a fixed copy of F in K_v. Given a graph G with $|E(G)| = \ell$, every F in G can be completed (by adding edges) to an H in G, in at most $x(H, F)$ ways, and in this fashion each H in G is obtained exactly $N(H, F)$ times (see [2, Lemma 3]). This implies that

$$N(\ell, H) \leq \frac{x(H, F)}{N(H, F)} N(\ell, F) \tag{5.5}$$

Similarly, $N(K_v, H) = \frac{x(H,F)}{N(H,F)} N(K_v, F)$ (see [2, Lemma 6]) and it follows from (5.5) that,

$$N(\ell, H) \leq \frac{N(K_v, H)}{N(K_v, F)} N(\ell, F) = \frac{|Aut(F)|}{|Aut(H)|} N(\ell, F) \leq \frac{1}{|Aut(H)|} (2\ell)^{|V(H)|/2}, \tag{5.6}$$

where the last inequality follows from (5.4).

For the lower bound, let $s = \lfloor \sqrt{2\ell} \rfloor$ and note that,

$$\begin{aligned} N(\ell, H) \geq N(K_s, H) &= \binom{s}{|V(H)|} N(K_{|V(H)|}, H) \\ &= \frac{s^{|V(H)|} + O(s^{|V(H)|-1})}{|V(H)|!} N(K_{|V(H)|}, H) \\ &= \frac{(2\ell)^{|V(H)|/2}}{|Aut(H)|} + O(\ell^{|V(H)|/2-1/2}), \end{aligned}$$

thus completing the proof. □

Remark 5.1. Theorem 1.3 calculates $N(\ell, H)$ asymptotically exactly, whenever H has a spanning subgraph which is a disjoint union of cycles or isolated edges. The proof also shows that if H is such a graph then the bound is asymptotically attained by a complete graph, that is, the complete graph maximizes the number of H-subgraphs over the set of all graphs with fixed number of edges. However, this is not true for general subgraphs. For example, if the number of edges is ℓ, a complete graph with $\sqrt{2\ell}$ vertices has $O(\ell^{3/2})$ 2-stars, whereas an $(\ell-1)$-star has $O(\ell^2)$ 2-stars. Thus, a complete graph does not maximize the number of 2-stars for a fixed number of edges. In fact, Alon [3] showed that $\lim_{\ell\to\infty} N(\ell, H)/\ell^{|V(H)|}$ is finite, and it is non-zero if and only if H is a disjoint union of stars. Moreover, he determined $N(\ell, H)$ precisely when H is a disjoint union of 2-stars, and also for some other special types of stars.

6. Universal Normal Approximation For Uniform Colorings

Theorem 1.1 says that if $\frac{1}{c}|E(G_n)| \stackrel{\mathscr{P}}{\to} \infty$, then $N(G_n)$ converges to infinity as well. Since a Poisson random variable with mean growing to ∞ converges to a standard normal distribution after standardizing (centering by mean and scaling by standard deviation), one possible question of interest is whether $N(G_n)$ properly standardized converges to a standard normal distribution. Such a limit theorem can be proved using a direct application of Stein method based on exchangeable pairs [33, Theorem 1]. However, as before, it turns out that even under the uniform coloring scheme an extra condition is needed on the structure of the graph for applying it. Nevertheless, as in the Poisson limit theorem of the previous section, the normality of the standardized random variable $N(G_n)$ is universal and can be proved by a method of moments argument.

This section proves that $N(G_n)$ properly standardized converges to a standard normal whenever both c and $|E(G_n)|/c$ goes to infinity. The calculation of moments in this regime require extensions of Alon's results to multi-graphs and more insights about the exponent $\gamma(H)$.

6.1. Proof of Theorem 1.2.

Let $G_n \in \mathscr{G}_n$ be a random graph sampled according to some probability distribution. This section proves a universal normal limit theorem for

$$Z_n := \left(\frac{|E(G_n)|}{c}\right)^{-\frac{1}{2}} \sum_{(i,j)\in E(G_n)} \left\{\mathbf{1}\{Y_i = Y_j\} - \frac{1}{c}\right\} = \left(\frac{|E(G_n)|}{c}\right)^{-\frac{1}{2}} \left(N(G_n) - \frac{E(G_n)}{c}\right).$$

Associated with every edge of G_n define the collection of random variables $\{X_{ij}, (i,j) \in E(G_n)\}$, where X_{ij} are i.i.d. Ber$(1/c)$, and set

$$W_n := \left(\frac{|E(G_n)|}{c}\right)^{-\frac{1}{2}} \sum_{(i,j)\in E(G_n)} \left\{X_{ij} - \frac{1}{c}\right\} = \left(\frac{|E(G_n)|}{c}\right)^{-\frac{1}{2}} \left(M(G_n) - \frac{E(G_n)}{c}\right).$$

6.1.1. *Comparing Conditional Moments.* Begin with two lemmas which will be used to compare the conditional moment of Z_n and W_n. However, unlike in previous sections, non-simple graphs are needed. To this end, define a multi-graph $G = (V, E)$ to be graph where multiple edges are allowed but there are no self loops. For a multi-graph G denote by G_S the simple graph obtained from G by removing all multiple edges. A multi-graph H is said to be a multi-subgraph of G if the simple graph H_S is a subgraph of G.

Observation 6.1. *Let $H = (V(H), E(H))$ be a multigraph with no isolated vertices. Let F be a multigraph obtained by removing one edge from H and removing all isolated vertices formed. Then $|V(F)| - \nu(F) \geq |V(H)| - \nu(H) - 1$.*

Proof. Observe that $\nu(F) \leq \nu(H) + 1$ and $|V(H)| - 2 \leq |V(F)| \leq |V(H)|$. If $|V(F)| = |V(H)|$ the result is immediate. Now, if $|V(F)| = |V(H)| - 1$, then the vertex removed must have degree 1 and so $\nu(F) = \nu(H)$, and the inequality still holds. Finally, if $\nu(F) = \nu(H) - 2$, the edge removed must be an isolated edge, in which case the number of vertices decrease by 2 and the number of connected components decrease by 1 and the result holds. □

The above observation helps determine the leading order of the expected central moments for multi-subgraph of G_n.

Lemma 6.1. *For any multi-subgraph $H = (V(H), E(H))$ of G_n define*

$$Z(H) = \prod_{(i,j) \in E(H)} \left\{ \mathbf{1}\{Y_i = Y_j\} - \frac{1}{c} \right\}, \quad \text{and} \quad X(H) = \prod_{(i,j) \in E(H)} \left\{ X_{ij} - \frac{1}{c} \right\}.$$

Then $\mathbb{E}(Z(H)) \lesssim_H \frac{1}{c^{|V(H)| - \nu(H)}}$ and $\mathbb{E}(X(H)) \lesssim_H \frac{1}{c^{|E(H_S)|}}$.

Proof. By expanding out the product,

$$Z(H) = \sum_{b=0}^{|E(H)|} \frac{(-1)^b}{c^b} \sum_{\substack{(i_s, j_s) \in E(H), \\ s \in [|E(H)| - b]}} \prod_{s=1}^{|E(H)| - b} \mathbf{1}\{Y_{i_s} = Y_{j_s}\}, \tag{6.1}$$

where the second sum is over all possible choices of $|E(H)| - b$ distinct multi-edges $(i_1, j_1), (i_2, j_2) \ldots (i_{|E(H)|-b}, j_{|E(H)|-b})$ from the multiset $E(H)$.

Let F be the subgraph of H formed by $(i_1, j_1), (i_2, j_2) \ldots (i_{|E(H)|-b}, j_{|E(H)|-b})$. Then by Observation 6.1, $|V(F)| - \nu(F) \geq |V(H)| - \nu(H) - b$, and

$$\frac{1}{c^b} \mathbb{E} \left(\prod_{s=1}^{|E(H)|-b} \mathbf{1}\{Y_{i_s} = Y_{j_s}\} \right) = \frac{1}{c^{|V(F)| - \nu(F) + b}} \leq \frac{1}{c^{|V(H)| - \nu(H)}}. \tag{6.2}$$

As the number of terms in (6.1) depends only on H, and for every term (6.2) holds, the result follows.

The result for $X(H)$ follows similarly. The leading order of the expectation comes from the first term

$$\mathbb{E} \left(\prod_{(i,j) \in E(H)} X_{ij} \right) = \frac{1}{c^{|E(H_S)|}},$$

and the number of terms depends only on H. □

The quantity $\gamma(H)$ was defined for a simple graph by Alon [2]. Friedgut and Kahn [19] showed that $\gamma(H)$ is the fractional stable number of H, which is the solution of a linear programming problem. Using this alternative definition, we can define $\gamma(H)$ for any multigraph as follows:

$$\gamma(H) = \arg \max_{\phi \in V_H[0,1]} \sum_{v \in V(H)} \phi(v) \text{ subject to } \phi(x)+\phi(y) \leq 1 \text{ for } (x,y) \in E(H),$$

where $V_H[0,1]$ is the collection of all functions $\phi : V(H) \to [0,1]$. It is clear that $\gamma(H) = \gamma(H_S)$. The polytope defined by the constraint set of this linear program is called the *fractional stable set polytope* which is a well studied object in combinatorial optimization [30]. With this definition, we now have the following lemma:

Lemma 6.2. *If for any multi-graph $H = (V(H), E(H))$ with no isolated vertices $\gamma(H) > \frac{1}{2}|E(H)|$, then H has a vertex of degree 1. Moreover, if H is a multi-subgraph of G_n which has a vertex of degree 1, then $\mathbb{E}(Z(H)|G_n) = \mathbb{E}(X(H)|G_n) = 0$.*

Proof. Suppose that $\gamma(H) > \frac{1}{2}|E(H)|$, and $d_{\min}(H) \geq 2$. Then for any $\phi : V(H) \to [0,1]$ such that $\phi(x) + \phi(y) \leq 1$ for $(x,y) \in E(H)$,

$$\sum_{x \in V(H)} \phi(x) \leq \frac{1}{d_{\min}(H)} \sum_{(x,y) \in E(H)} \{\phi(x)+\phi(y)\} \leq \frac{1}{2}|E(H)|,$$

which is a contradiction.

Now, without loss of generality assume that vertex 1 has degree 1. Suppose vertex $s \in [n] \backslash \{1\}$ is the only neighbor of 1. Therefore,

$$\mathbb{E}(Z(H)|Y_1, G_n) = \left(\mathbb{E}(\mathbf{1}\{Y_1 = Y_s\}|Y_1, G_n) - \frac{1}{c}\right) \prod_{\substack{(i,j) \in E(H), \\ (i,j) \neq (1,s)}} \left\{\mathbf{1}\{Y_i = Y_j\} - \frac{1}{c}\right\} = 0,$$

which implies $\mathbb{E}(Z(H)|G_n) = 0$. The result for $X(H)$ can be proved similarly. □

With the above lemmas, the conditional moments of Z_n and W_n can be compared. For a simple graph G and a multigraph H define

$$M(G,H) = \sum_{e_1 \in E(G)} \sum_{e_2 \in E(G)} \cdots \sum_{e_{|E(H)|} \in E(G)} \mathbf{1}\{G[e_1, e_2, \ldots e_{|E(H)|}] = H\},$$

where $G[e_1, e_2, \ldots e_{|E(H)|}]$ is the multi-subgraph of G formed by the edges $e_1, e_2, \ldots, e_{|E(H)|}$. It is easy to see that $M(G,H) \lesssim_H N(G, H_S)$.

Lemma 6.3. *Let W_n and Z_n be as defined before, with $c \to \infty$ and $|E(G_n)|/c \xrightarrow{\mathscr{P}} \infty$, then for every fixed $k \geq 1$ we have*

$$|\mathbb{E}(Z_n^k|G_n) - \mathbb{E}(W_n^k|G_n)| \xrightarrow{\mathscr{P}} 0.$$

Proof. Let $\mathcal{M}_k$ be the set of all multi-graphs with exactly k multi edges and $d_{min}(H) \geq 2$. Note that by Lemma 6.2 any $H \in \mathcal{M}_k$ must satisfy $\gamma(H) \leq E(H)/2$. Expanding the product and using Lemma 6.2,

$$\mathbb{E}(Z_n^k|G_n) \quad = \quad \left(\frac{|E(G_n)|}{c}\right)^{-\frac{k}{2}} \sum_{H \in \mathcal{M}_k} M(G_n, H) \cdot \mathbb{E}(Z(H)), \tag{6.3}$$

By similar argument with Z_n replaced by X_n,

$$\mathbb{E}(W_n^k|G_n) = \left(\frac{|E(G_n)|}{c}\right)^{-\frac{k}{2}} \sum_{H \in \mathcal{M}_k} M(G_n, H) \cdot \mathbb{E}(X(H)), \tag{6.4}$$

Now, let $\mathscr{S}_k \subset \mathcal{M}_k$ be the set of all multi-graphs H with $d_{\min}(H) \geq 2, |E(H)| = k$ and $\gamma(H) = |E(H)|/2 = |V(H)| - \nu(H)\}$. Let $\omega = |E(G_n)|/c$. Now by Lemma 6.1 and Theorem 5.1, for any $H \in \mathcal{M}_k \backslash \mathscr{S}_k$,

$$\begin{aligned}
\left(\frac{|E(G_n)|}{c}\right)^{-\frac{|E(H)|}{2}} M(G_n, H) \cdot \mathbb{E}(Z(H)) &\lesssim_H \left(\frac{|E(G_n)|}{c}\right)^{-\frac{|E(H)|}{2}} \frac{N(G_n, H_S)}{c^{|V(H)|-\nu(H)}} \\
&\lesssim_H \frac{|E(G_n)|^{\gamma(H)-\frac{1}{2}|E(H)|}}{c^{|V(H)|-\nu(H)-\frac{1}{2}|E(H)|}} \\
&\lesssim_H \frac{\omega^{\gamma(H)-\frac{1}{2}|E(H)|}}{c^{|V(H)|-\nu(H)-\gamma(H)}} \rightarrow 0,
\end{aligned} \tag{6.5}$$

Similarly, for $H \in \mathcal{M}_k \backslash \mathscr{S}_k$,

$$\begin{aligned}
\left(\frac{|E(G_n)|}{c}\right)^{-\frac{|E(H)|}{2}} M(G_n, H) \cdot \mathbb{E}(X(H)) &\lesssim_H \left(\frac{|E(G_n)|}{c}\right)^{-\frac{|E(H)|}{2}} \frac{N(G_n, H_S)}{c^{|E(H_S)|}} \\
&\lesssim_H \left(\frac{|E(G_n)|}{c}\right)^{-\frac{|E(H)|}{2}} \frac{N(G_n, H_S)}{c^{|V(H)|-\nu(H)}} \rightarrow 0.
\end{aligned} \tag{6.6}$$

The limits in (6.5) and (6.6), together with Equations (6.3) and (6.4) give

$$\lim_{n\to\infty} |\mathbb{E}(Z_n^k|G_n) - \mathbb{E}(W_n^k|G_n)| \leq \lim_{n\to\infty} \left(\frac{|E(G_n)|}{c}\right)^{-\frac{k}{2}} \sum_{H \in \mathscr{S}_k} M(G_n, H) \cdot |\mathbb{E}(Z(H)) - \mathbb{E}(X(H))|. \tag{6.7}$$

Therefore, only multi-subgraphs of G_n which are in $\mathscr{S}_k$ need to considered. As

$$\gamma(H_S) = \gamma(H) = |V(H)| - \nu(H) = |V(H_S)| - \nu(H_S),$$

H_S is a disjoint union of stars by Corollary 5.2. Therefore

$$|E(H_S)| = |V(H_S)| - \nu(H_S) = |V(H)| - \nu(H) = |E(H)|/2.$$

This, along with the fact that H cannot have any vertex of degree 1 gives that any $H \in \mathscr{S}_k$ is a disjoint union of stars, where every edge is repeated twice. Now, it is easy to see that for any such graph H, $\mathbb{E}(Z(H)) = \mathbb{E}(X(H))$, and the result follows from (6.7). □

6.1.2. *Completing the Proof of Theorem 1.2.* To complete the proof of the normal approximation the following lemma, which shows that W_n satisfies the conditions required in Lemma 8.1, is needed.

Lemma 6.4. *Let W_n be as defined before. Then $W_n \xrightarrow{\mathscr{D}} N(0,1)$, and further for any $\epsilon > 0, t \in \mathbb{R}$,*

$$\limsup_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\left|\frac{t^k}{k!}\mathbb{E}(W_n^k|G_n)\right| > \epsilon\right) = 0. \tag{6.8}$$

□

Proof. To prove the first conclusion, let

$$T_n := \frac{M(G_n) - \frac{|E(G_n)|}{c}}{\sqrt{\frac{|E(G_n)|}{c} - \frac{|E(G_n)|}{c^2}}}.$$

By the Berry-Esseen theorem and the dominated convergence theorem it follows that T_n converges to $N(0,1)$. Moreover, $W_n - T_n \xrightarrow{\mathscr{P}} 0$, which implies W_n converges to $N(0,1)$ by Slutsky's theorem.

To prove the second conclusion, it suffices to show that for any $\varepsilon > 0, k \geq 1$,

$$\limsup_{n\to\infty} \mathbb{P}(|\mathbb{E}(W_n^k|G_n)| > \mu_k + \varepsilon) = 0, \tag{6.9}$$

where $\mu_k = \mathbb{E}(Z^k)$ and Z is a standard normal random variable. This is because $\mathbb{E}e^{|t|Z} < \infty$ for any t, so (6.8) follows by applying Fatou's lemma twice as in the proof of Lemma 3.5.

To prove (6.9), note from the proof of Lemma 6.3 that $\mathbb{E}(W_n^k|G_n) = o_P(1)$ for odd k. Therefore, assume that $k = 2m$ is even. Recall that $\mathscr{S}_{2m}$ is the number of multi-subgraphs of G_n with m double edges, where the underlying simple graph is a disjoint union of stars. Denote by $\mathcal{A}_{2m}$ the collection of all multi-subgraphs of G_n with m double edges. Note that

$$\sum_{H\in\mathscr{S}_{2m}} M(G_n, H) \leq \sum_{H\in\mathcal{A}_{2m}} M(G_n, H) \leq \binom{|E(G_n)|}{m}\frac{(2m)!}{2^m},$$

where in the last step we use the fact that any such graph in $\mathcal{A}_{2m}$ can be produced by choosing m out of the $|E(G_n)|$ edges and then permuting the $2m$ edges (each chosen edge doubled) within themselves. Therefore, from the proof of Lemma 6.3

$$\begin{aligned}
\mathbb{E}(W_n^{2m}|G_n) &= \left(\frac{|E(G_n)|}{c}\right)^{-m} \sum_{H\in\mathscr{S}_{2m}} M(G_n, H)\cdot\mathbb{E}(X(H)) + o_P(1)\\
&= \frac{\sum_{H\in\mathscr{S}_{2m}} M(G_n, H)}{|E(G_n)|^m}\left(1 - \frac{1}{c}\right)^m + o_P(1)\\
&\leq \binom{|E(G_n)|}{m}\frac{(2m)!}{2^m}\frac{1}{|E(G_n)|^m}\left(1-\frac{1}{c}\right)^m + o_P(1)\\
&\leq \frac{(2m)!}{2^m m!} + o_P(1),
\end{aligned}$$

which establishes (6.9), hence completing the proof of Lemma 6.4. □

Remark 6.1. Lemma 6.2 implies that a graph H has a vertex of degree 1, whenever $\gamma(H) > |E(H)|/2$. In fact, this result is tight, that is, there are graphs, like the cycle C_g or the complete bipartite graph $K_{2,s}$, with $\gamma(H) = |E(H)|/2$ and no isolated vertices. Even though this result can be proved easily from definitions, it plays a crucial part in our proof of Theorem 1.2. As the number of copies of H in G_n is small whenever $\gamma(H) < |E(H)|/2$, these graphs asymptotically do not contribute to the expectation. The fact that $\gamma(H) > \frac{1}{2}|E(H)|$, implies that H has a vertex of degree 1 ensures that the expected central moments $Z(H)$ vanish. Therefore, the only graphs that contribute in the moments are those where $\gamma(H) = |E(H)|/2$. That the threshold $|E(H)|/2$, which is obtained from probabilistic calculations, is also the threshold where graphs have degree 1 vertices is a fortunate coincidence which illustrates a nice interplay between probability and graph theory in this problem.

6.2. **Non-Normal Limit for Fixed Colors.** The assumption that the number of colors c goes to infinity is essential for the normality in Theorem 1.2. If $|E(G_n)|/c$ goes to infinity and c is fixed, then the limiting distribution of the number of monochromatic edges might not be normal, as demonstrated in the following proposition:

Proposition 6.1. *For c fixed and the uniform coloring distribution, the number of monochromatic edges $N(G_n)$ of the complete graph K_n satisfies:*

$$\frac{1}{n}\left(N(K_n) - c\binom{n/c}{2}\right) \xrightarrow{\mathscr{D}} \chi^2_{(c-1)}.$$

Proof. For $a \in [q]$, define X_a be the number of vertices of K_n with color a. Then for $\underline{X} = (X_1, X_2, \ldots, X_q)$ and $\underline{p} = (1/c, 1/c, \ldots, 1/c)$,

$$\underline{X} \sim \text{Mutlinomial}\left(n, \underline{p}\right), \text{ and } n^{-\frac{1}{2}}\left(\underline{X} - n\underline{p}\right) \xrightarrow{\mathscr{D}} N(0, \Sigma),$$

as $n \to \infty$, where $\Sigma = \frac{1}{c}\mathbf{I} - \frac{1}{c^2}\mathbf{1}\mathbf{1}'$. This implies that $n^{-1}(\underline{X} - n\underline{p})'(\underline{X} - n\underline{p}) \xrightarrow{\mathscr{D}} \chi^2_{(c-1)}$, as $rank(\Sigma) = c - 1$.

Now, by a second order Taylor series expansion, the number of monochromatic edges of K_n is

$$N(K_n) = \sum_{a=1}^{c}\binom{X_a}{2} = c\binom{n/c}{2} + \sum_{a=1}^{c}(X_a - n/c)^2 + o_P(n),$$

and the result follows. □

7. Extremal Examples: Stars and Cycles

Example 1 of Cerquetti and Fortini [9] shows that the conditions in Theorem 2.1 cannot be entirely relaxed for general non-uniform coloring distributions, that is, that there exists a graph and a distribution $\underline{p} = (p_1, \cdots, p_c)$, with $c = |E(G_n)|$, which do not satisfy the conditions of Theorem 2.1 and $N(G_n)$ does not converge to a mixture of Poisson. This example indicates that universality cannot be extended very much beyond the uniform coloring distribution.

Another relevant question is whether it is possible to expect a similar Poisson universality result for other subgraphs, under uniform coloring scheme? This section begins by proving Proposition 1.4 which shows that we may not get Poisson mixtures in the limit while counting monochromatic r-stars, in a uniform c-coloring of an n-star.

7.1. **Monochromatic Stars.** Consider the $(n-1)$-star, $K_{1,n-1}$ with vertices labelled by $[n]$, with the central vertex labeled 1. Color the vertices of $K_{1,n-1}$, uniformly from $[c]$, independent from other vertices. Consider the limiting distribution of the number of monochromatic r-stars $K_{1,r-1}$ generated by this random coloring, where r is a fixed constant. If Y_i denotes the color of the vertex i, the random variable is

$$T_{r,n} = \sum_{\substack{S \subseteq [n]\backslash\{1\} \\ |S| = r-1}} \prod_{j \in S} \mathbf{1}\{Y_1 = Y_j\}.$$

Proposition 1.4 shows that the limiting behavior of $T_{r,n}$ cannot converge to a mixture of Poissons. This illustrates that the phenomenon of universality of the Poisson approximation that holds for the number of monochromatic edges, does not extend to arbitrary subgraphs. In particular, it is not even true for the 2-star, which is the simplest extension of an edge.

7.1.1. *Proof of Proposition 1.4.* Note that if the number of monochromatic edges in $G_n = K_{1,n-1}$ is $N(G_n)$, then

$$T_{r,n} \stackrel{\mathscr{D}}{=} \binom{N(G_n)}{r}.$$

If $n/c \to 0$, then from Theorem 1.1 $N(G_n) \stackrel{\mathscr{P}}{\to} 0$ and so $T_{r,n} \stackrel{\mathscr{P}}{\to} 0$. Similarly, if $n/c \to \infty$, $T_{r,n} \stackrel{\mathscr{P}}{\to} \infty$.

Finally, if $\frac{n}{c} \to \lambda$, the number of monochromatic edges $N(G_n)$ in $K_{1,n-1}$ converges to $X \sim Poisson(\lambda)$, by Theorem 1.1. This implies that

$$T_{r,n} \stackrel{\mathscr{D}}{=} \binom{N(G_n)}{r} \stackrel{\mathscr{D}}{\to} \binom{X}{r} = \frac{X(X-1)\cdots(X-r+1)}{r!}.$$

This random variable does not assign positive masses at all non-negative integers, and so it cannot be a mixture of Poisson variates.

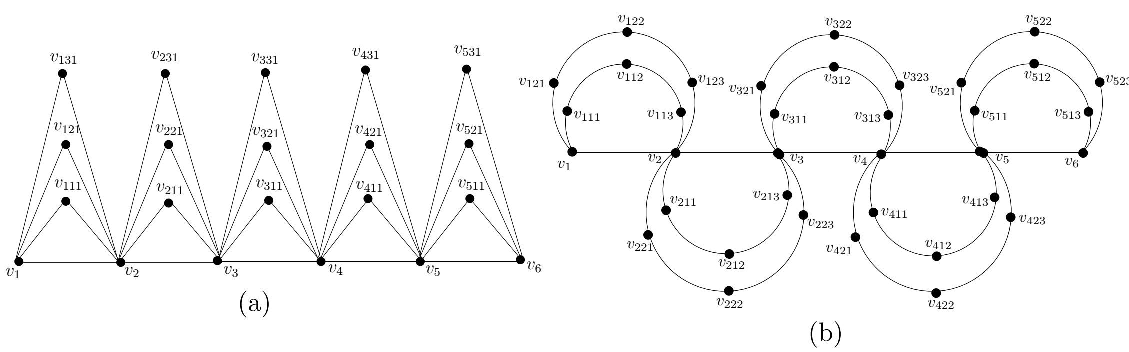

FIGURE 1. (a) The graph $G_{5,3,3}$, and (b) the graph $G_{5,2,5}$

7.2. **Monochromatic Cycles.** Recall that the number of monochromatic edges $N(G_n)$ converges to $Poisson(\lambda)$ whenever $|E(G_n)|/c \to \lambda$. The limiting distribution of the number of edges can only be a non-trivial mixture of Poissons when $E(G_n)|/c \to Z$, and Z has a non-degenerate distribution. We now construct a graph G_n where the number of monochromatic g-cycles in a uniform c-coloring of G_n converges in distribution to non-trivial mixture of Poissons even when $|N(G_n, C_g)|/c^{g-1}$ converges to a fixed number λ. This phenomenon, which cannot happen in the case of edges, makes the problem of finding the limiting distribution of the number of monochromatic cycles, much more challenging.

For a, b positive integers and $g \geq 3$, define a graph $G_{a,b,g}$ as follows: Let

$$V(G_{a,b,g}) = \{v_1, v_2, \ldots v_{a+1}\} \bigcup_{i=1}^{a} \bigcup_{j=1}^{b} \{v_{ijk} : k \in \{1, 2, \ldots g-2\}\}.$$

The edges are such that vertices $v_1, v_2, \ldots v_{a+1}$ form a path of length a, and for every $i \in [a]$ and $j \in [b]$, $v_i, v_{ij1}, v_{ij2}, \ldots v_{ijg-2}, v_{i+1}$ form a cycle of length g (Figure 1 shows the structure of graphs $G_{5,3,3}$ and $G_{5,2,5}$, and the corresponding vertex labelings.). Note that graph $G_{a,b,g}$ has $b(g-2)+a+1$ vertices, $b(g-1)+a$ edges, and ab cycles of length g.

We consider a uniform c-coloring of the vertices of $G_{a,b,g}$ and count the number of monochromatic g-cycles. Let Y_i be the color of the vertex v_i and Y_{ijk} the color of the vertex v_{ijk}, for $i \in [a]$ and

$j \in [b]$. The random variable

$$Z_{a,b,g} := Z(G_{a,b,g}) := \sum_{i=1}^{a} \sum_{j=1}^{b} \prod_{k=1}^{g-2} \mathbf{1}\{Y_i = Y_{i+1} = Y_{ijk}\},$$

which counts the number of monochromatic g-cycles in the graph $G_{a,b,g}$. The following proposition shows that there exists a choice of parameters a, b, c such that $|N(G_n, C_g)|/c^{g-1} \to \lambda$ and $Z_{a,b,g}$ converges in distribution to a non-trivial mixture of Poissons.

Proposition 7.1. *For $a = \lambda n$ and $b = n^{g-2}$ and $c = n$, $Z_{a,b,g} \xrightarrow{\mathscr{D}}$ Poisson(W), where $W \sim$ Poisson(λ).*

Proof. Let $\underline{Y} = (Y_1, Y_2, \cdots Y_{a+1})$ and note that

$$\prod_{k=1}^{g-2} \mathbf{1}\{Y_i = Y_{i+1} = Y_{ijk}\}\Big|\underline{Y} \sim \mathrm{Ber}(1/c^{g-2}) \text{ and } \sum_{j=1}^{b} \prod_{kj=1}^{g-2} \mathbf{1}\{Y_i = Y_{i+1} = Y_{ijk}\}\Big|\underline{Y} \sim \mathrm{Bin}(b, 1/c^{g-2}).$$

This implies that

$$\mathbb{E}\left(e^{itZ_{a,b,g}}\right) = \mathbb{E}\left(\prod_{i=1}^{a} \mathbb{E}\left(e^{it\sum_{j=1}^{b}\prod_{k=1}^{g-2}\mathbf{1}\{Y_i=Y_{i+1}=Y_{ijk}\}}\Big|\underline{Y}\right)\right) = \mathbb{E}\left(1 - \frac{1}{c^{g-2}} + \frac{e^{it}}{c^{g-2}}\right)^{bN_a}, \tag{7.1}$$

where $N_a = \sum_{i=1}^{a} \mathbf{1}\{Y_i = Y_{i+1}\}$, is number of monochromatic edges in the path $v_1, v_2, \ldots, v_{a+1}$.

Substituting $a = \lambda n := a_n$ and $b = n^{g-2} := b_n$ and $c = n := c_n$, we have $N(G, C_g)/c_n^{g-1} = a_n b_n / c_n = \lambda$. With this choice a_n, b_n, c_n, we have by Theorem 1.1, N_{a_n} converges in distribution to $W :=$ Poisson(λ), as $a_n/c_n = \lambda$. Therefore,

$$\left(1 - \frac{1}{c_n^{g-2}} + \frac{e^{it}}{c_n^{g-2}}\right)^{b_n N_{a_n}} = e^{b_n N_{a_n} \log\left(1 - \frac{1}{c_n^{g-2}} + \frac{e^{it}}{c_n^{g-2}}\right)} \xrightarrow{\mathscr{D}} e^{(e^{it}-1)W}.$$

As $|e^{itZ_{a_n,b_n,g}}| \leq 1$, from (7.1) and the dominated convergence theorem we have

$$\mathbb{E}\left(e^{itZ_{a_n,b_n,g}}\right) \xrightarrow{\mathscr{D}} \mathbb{E}\left(e^{(e^{it}-1)W}\right),$$

which the characteristic function of Poisson(W), where $W \sim$ Poisson(λ). □

Remark 7.1. We were unable to construct an example of a graph for which the number of monochromatic triangles converges to some distribution which is not a mixture of Poissons, when $N(G_n, C_3)/c^2 \to \lambda$. The above construction and the inability to construct any counterexamples, even for triangles, lead us to believe that some kind of Poisson universality holds for cycles as well. More formally, we conjecture that the number of monochromatic g-cycles in a uniform random coloring of any graph sequence G_n converges in distribution to a random variable which is mixture of Poissons, whenever $|N(G_n, C_g)|/c^{g-1} \to \lambda$, for some fixed $\lambda > 0$.

Acknowledgement: The authors are indebted to Noga Alon for pointing out an error in the proof of Theorem 1.3 in an earlier draft. The authors also thank Sourav Chatterjee and Bjarne Toft for helpful comments.

References

[1] D. Aldous, *Probability approximations via the Poisson clumping heuristic*, Applied Mathematical Sciences, Vol. 77, Springer-Verlag, New York, 1989.

[2] N. Alon, On the number of subgraphs of prescribed type of graphs with a given number of edges, *Israel Journal of Mathematics*, Vol. 38, 116–130, 1981.

[3] N. Alon, On the number of certain subgraphs contained in graphs with a given number of edges, *Israel Journal of Mathematics*, Vol. 53, 97–120, 1986.

[4] R. Arratia, L. Goldstein, and L. Gordon, Poisson approximation and the Chen-Stein method, *Statistical Science*, Vol. 5 (4), 403–424, 1990.

[5] A. D. Barbour, L. Holst, L. and S. Janson, *Poisson Approximations*, Oxford University Press, Oxford, 1992.

[6] B. Bollobás, S. Janson, and O. Riordan, The phase transition in inhomogeneous random graphs, *Random Structures and Algorithms*, Vol. 31 (1), 3–122, 2007.

[7] M. Burda, M. Harding, and J. Hausman, A Poisson mixture model of discrete choice, *Journal of Econometrics*, Vol. 166, 184–203, 2012.

[8] M. Camarri and J. Pitman, Limit distributions and random trees derived from the birthday problem with unequal probabilities, *Electron. J. Probab.*, Vol. 5, Paper 2, 1–18, 2000.

[9] A. Cerquetti and S. Fortini, A Poisson approximation for colored graphs under exchangeability, *Sankhya: The Indian Journal of Statistics* Vol. 68(2), 183–197, 2006.

[10] S. Chatterjee, P. Diaconis, E. Meckes, Exchangeable pairs and Poisson approximation, *Electron. Encyclopedia Probab.*, 2004.

[11] K. Church and W. A. Gale, Poisson mixtures, *Natural Language Engineering*, Vol. 1(2), 163–190, 1995.

[12] A. DasGupta, The matching, birthday and the strong birthday problem: a contemporary review, *J. Statist. Plann. Inference*, Vol. 130, 377–389, 2005.

[13] P. Diaconis and S. Holmes, A bayesian peek into Feller I, *Sankhya, Series A*, Vol. 64 (3), 820–841, 2002.

[14] P. Diaconis and F. Mosteller, Methods for studying coincidences, *The Journal of the American Statistical Association*, Vol. 84(408), 853–861, 1989.

[15] F. M Dong, K. M. Koh, and K. L. Teo, *Chromatic polynomials and chromaticity of graphs*, World Scientific Publishing Company, 2005.

[16] R. Durrett, *Probability: Theory and Examples*, 4th edition, Cambridge U. Press, 2010.

[17] P. Erdős, On the number of complete subgraphs contained in certain graphs, *Publ. Math. Inst. Hungar. Acad. Sci.*, Vol. 7, 459-464, 1962.

[18] S. Fadnavis, A generalization of the birthday problem and the chromatic polynomial, *arXiv:1105.0698v2 [math.CO]*, 2011.

[19] E. Friedgut and J. Kahn, On the number of copies of one hypergraph in another, *Israel Journal of Mathematics*, Vol. 105 (1), 251–256, 1998.

[20] A. Frieze and E. Vigoda, A survey on the use of Markov chains to randomly sample colorings, *Combinatorics, Complexity and Chance*, Oxford University Press, 2007.

[21] A. Gelman, *600: The Average American Knows How Many People?*, New York Times, Page D7, Feb. 19, 2013.

[22] M. Greenwood and G. V. Yule, An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attack of disease or of repeated accidents, *J. Roy. Statist. Soc.*, Vol. 83, 255–279, 1920.

[23] T. R. Jensen and B. Toft, *Graph Coloring Problems*, Wiley-Interscience Series in Discrete Mathematics and Optimization, 1995.

[24] T. R. Jensen and B. Toft, Unsolved Graph Coloring Problems, *manuscript in preparation*, 2013.

[25] P. Killworth and H. Bernard, The reversal small-world experiment, *Social Networks*, Vol. 1 (2), 1979.

[26] J. B. Lasserre, Quasi-Feller Markov chains, *Journal of Applied Mathematics and Stochastic Analysis*, Vol. 13(1), 15–24, 2000.

[27] L. Le-Cam and R. Traxler, On the asymptotic behavior of mixtures of Poisson distributions, *Zeitschrift fur Wahrscheinlichkeitstheorie*, Vol. 44, 1–45, 1978.

[28] L. Lovász, *Large Networks and Graph Limits*, Colloquium Publications, Vol. 60, 2012.

[29] H. L. Royden, *Real Analysis*, Pearson Education, Third Edition, 1988.

[30] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, Algorithms and Combinatorics, Vol. 24, Berlin, Springer, 2003.

[31] J. Sethuraman, Some limit theorems for joint distributions, *Sankhya: The Indian Journal of Statistics, Series A*, Vol. 23 (4), 1961.

[32] R. P. Stanley, A symmetric function generalization of the chromatic polynomial of a graph, *Advances in Mathematics*, Vol. 111 (1), 166–194, 1995.
[33] C. Stein, Approximate computation of expectations, *Institute of Mathematical Statistics Lecture Notes*, Monograph Series, Vol. 7, 1986.
[34] T. J. Sweeting, On conditional weak convergence, *Journal of Theoretical Probability*, Vol. 2 (4), 461–474, 1989.

8. Appendix: Conditional Convergence to Unconditional Convergence

There are many conditions on modes on convergence which ensure the convergence of a sequence of joint distributions when it is known that the associated sequence of marginal and conditional distributions converge [31, 34]. This section gives a proof of a technical lemma which allows conclusions about the limiting distribution of a random variable from its conditional moments. The lemma is used twice in the paper in the final steps of our proofs of the universal Poisson and the Normal limit theorems.

Lemma 8.1. *Let $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ be a sequence of probability spaces, and let $\mathcal{G}_n \subset \mathcal{F}_n$ be a sequence of sub sigma fields. Also let (X_n, Y_n) be a sequence of random variables on $(\Omega_n, \mathcal{F}_n)$, and assume that for any $k \geq 1$ the conditional expectations $U_{n,k} := \mathbb{E}(X_n^k|\mathcal{G}_n), V_{n,k} := \mathbb{E}(Y_n^k|\mathcal{G}_n)$ exist as finite random variables. Suppose the following two conditions hold:*

$$\limsup_{n\to\infty} \mathbb{P}(|U_{n,k} - V_{n,k}| > \epsilon) = 0, \tag{8.1}$$

$$\limsup_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\left|\frac{t^k}{k!} U_{n,k}\right| > \epsilon\right) = 0. \tag{8.2}$$

Then for any $t \in \mathbb{R}$ we have $\mathbb{E}e^{itX_n} - \mathbb{E}e^{itY_n} \to 0$.

Proof. Without loss of generality assume $t \neq 0$. Note that

$$\mathbb{P}\left(\left|\frac{t^k}{k!} V_{n,k}\right| > \epsilon\right) \leq \mathbb{P}\left(\left|\frac{t^k}{k!} U_{n,k}\right| > \frac{\epsilon}{2}\right) + \mathbb{P}\left(|U_{n,k} - V_{n,k}| > \frac{\epsilon k!}{2|t|^k}\right).$$

Taking limits as $n \to \infty$, and using (8.1) and (8.2), it follows that

$$\limsup_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\left|\frac{t^k}{k!} V_{n,k}\right| > \epsilon\right) = 0. \tag{8.3}$$

By a Taylor's series expansion, for any $k \in \mathbb{N}$, $\left|e^{it} - \sum_{r=0}^{k-1} \frac{(it)^r}{r!}\right| \leq \frac{t^k}{k!}$, and so

$$|\mathbb{E}(e^{itX_n}|\mathcal{G}_n) - \mathbb{E}(e^{itY_n}|\mathcal{G}_n)| \leq \sum_{r=0}^{k-1} \frac{|t|^r}{r!} |U_{n,r} - V_{n,r}| + \frac{|t|^k}{k!} U_{n,k} + \frac{|t|^k}{k!} V_{n,k}. \tag{8.4}$$

From (8.4) and taking limits as $n \to \infty$ followed by $k \to \infty$, using (8.1), (8.2) and (8.3) gives $\limsup_{n\to\infty} \mathbb{P}(|\mathbb{E}(e^{itX_n}|\mathcal{G}_n) - \mathbb{E}(e^{itY_n}|\mathcal{G}_n)| > \epsilon) = 0$. This implies that $|\mathbb{E}(e^{itX_n}|\mathcal{G}_n) - \mathbb{E}(e^{itY_n}|\mathcal{G}_n)|$ converges to 0 in probability. Since $|\mathbb{E}(e^{itX_n}|\mathcal{G}_n) - \mathbb{E}(e^{itY_n}|\mathcal{G}_n)|$ is also bounded by 2 in absolute value, dominated convergence gives

$$\lim_{n\to\infty} \mathbb{E}(e^{itX_n}) - \mathbb{E}(e^{itY_n}) = 0,$$

thus completing the proof of Lemma 8.1. ∎

Remark 8.1. As mentioned earlier, the reason for Lemma 8.1 is that separate convergences of the conditional distribution and the marginal distribution do not in general, imply the unconditional convergence. Consider the following example: Let $c_n \downarrow 0$ be a sequence of constants with $nc_n \to 0$. Define

$$K_n(x,\cdot) = \begin{cases} \boldsymbol{\delta}_{c_n x} & \text{if } x > 0, \\ 1 & \text{if } x = 0; \end{cases} \qquad K(x,\cdot) = \begin{cases} \boldsymbol{\delta}_0 & \text{if } x > 0, \\ 1 & \text{if } x = 0. \end{cases}$$

It is easy to check that for all $f \in C[0,1]$, $K_n(f) \to K(f)$. However, if $\pi_n \sim Uniform(0, 1/n)$, then $\pi_n \stackrel{\mathscr{P}}{\to} \boldsymbol{\delta}_0 := \pi$, but $\pi_n K_n(f) \to f(0)$ and $\pi K(f) = f(1)$. Some conditions like the *weak-Feller* property of the kernel [26] or *set-wise* convergence of the marginals (via the Vitali-Hahn-Saks theorem [29]) are required for joint convergence to be true. Nevertheless, in our case the proof of the unconditional convergence follows without having to invoke any such general theorems as sums of i.i.d. random variables can be dealt with directly, and exponential moments of Poisson and Normal distributions are finite.

Department of Statistics, Stanford University, California, USA, bhaswar@stanford.edu

Department of Mathematics and Statistics, Stanford University, California, USA, diaconis@math.stanford.edu

Department of Statistics, Stanford University, California, USA, sumitm@stanford.edu