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

BY BHASWAR B. BHATTACHARYA[*], PERSI DIACONIS[†]
AND SUMIT MUKHERJEE[‡]

University of Pennsylvania[*], Stanford University[†] and Columbia University[‡]

This paper proves limit theorems for the number of monochromatic edges in uniform random colorings of general random graphs. These can be seen as generalizations of the birthday problem (what is the chance that there are two friends with the same birthday?). It is shown that if the number of colors grows to infinity, the asymptotic distribution is either a Poisson mixture or a Normal depending solely on the limiting behavior of the ratio of the number of edges in the graph and the number of colors. This result holds for any graph sequence, deterministic or random. On the other hand, when the number of colors is fixed, a necessary and sufficient condition for asymptotic normality is determined. Finally, using some results from the emerging theory of dense graph limits, the asymptotic (nonnormal) distribution is characterized for any converging sequence of dense graphs. The proofs are based on moment calculations which relate to the results of Erdős and Alon on extremal subgraph counts. As a consequence, a simpler proof of a result of Alon, estimating the number of isomorphic copies of a cycle of given length in graphs with a fixed number of edges, is presented.

CONTENTS

1. Introduction. Suppose the vertices of a finite graph G 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 $\chi_G(c)/c^{|V(G)|}$, where $\chi_G(c)$ denotes the number of proper colorings of G using c-colors and $|V(G)|$ is the number of vertices in G. The function χ_G is the chromatic polynomial of G, and is a central object in graph theory [22, 34, 35]. This paper studies the limiting distribution of the number of monochromatic edges in uniform random colorings of general random graphs.

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 $[n] := \{1, 2, \ldots, n\}$.

Given a graph $G_n \in \mathscr{G}_n$ with adjacency matrix $A(G_n) = ((A_{ij}(G_n)))_{1 \le i,j \le n}$, denote by $V(G_n)$ the set of vertices, and by $E(G_n)$ the set of edges of G_n, respectively. The vertices of G_n are colored with $c = c_n$ colors as follows:

$$(1.1) \qquad \mathbb{P}\big(v \in V(G_n) \text{ is colored with color } a \in \{1, 2, \ldots, c\} \mid G_n\big) = \frac{1}{c},$$

independent from the other vertices. If Y_i is the color of vertex i, then

$$(1.2) \qquad N(G_n) := \sum_{1 \le i < j \le n} A_{ij}(G_n)\mathbf{1}\{Y_i = Y_j\} = \sum_{(i,j) \in E(G_n)} \mathbf{1}\{Y_i = Y_j\},$$

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. When $c = 365$ and $G_n = K_n$ is a complete graph this reduces to the classical birthday problem: $\mathbb{P}(N(K_n) \ge 1)$ is the probability that there are two people with the same birthday in a room with n people. 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 $\mathbb{P}(Y_1 = a) = 1/c$, for all $a \in [c]$. Note that this setup includes the case where $\{G_1, G_2, \ldots\}$ is a deterministic (nonrandom) graph sequence, as well.

An application of the easily available version of Stein's method gives a general limit theorem for $N(G_n)$ that works for all color distributions [5, 11]. For the uniform coloring scheme, using the fact that the random variables are pairwise independent, Barbour, Holst and Janson [5] proved a Poisson approximation for the number of monochromatic edges which works for any sequence of deterministic graphs. The following theorem gives a new proof and slightly extends this result by showing that the same is true for random graphs. Unlike Stein's method, our proof, which is based on the method of moments, does not give convergence rates. However, it illustrates the connections to extremal combinatorics, and builds up to our later results.

THEOREM 1.1. *Let $G_n \in \mathscr{G}_n$ be a random graph sampled according to some probability distribution over $\mathscr{G}_n$ and $c_n \to \infty$. Then under the uniform coloring distribution, the following is true*:

$$N(G_n) \xrightarrow{D} \begin{cases} 0 & \text{if } \dfrac{1}{c_n} \cdot |E(G_n)| \xrightarrow{P} 0, \\[2ex] \infty & \text{if } \dfrac{1}{c_n} \cdot |E(G_n)| \xrightarrow{P} \infty, \\[2ex] W & \text{if } \dfrac{1}{c_n} \cdot |E(G_n)| \xrightarrow{D} Z; \end{cases}$$

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 $|E(G_n)|/c_n$ 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

$$(1.3) \qquad M(G_n) := \sum_{1 \leq i < j \leq n} A_{ij}(G_n) Z_{ij},$$

where $\{Z_{ij}\}_{(i,j) \in E(G_n)}$ are independent $\mathrm{Ber}(1/c)$. The combinatorial quantity that needs to be bounded 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 spectral 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 $\frac{1}{c_n}|E(G_n)| \overset{P}{\to} \infty$, 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 by the mean and scaling by the standard deviation, it is natural to wonder whether the same is true for $N(G_n)$. This is not true in general if $|E(G_n)|/c_n$ goes to infinity, with $c_n = c$ fixed as indicated in Example 6.1 and Example 7.1. On the other hand, if $c_n \to \infty$, an off-the-shelf version of Stein's method can be used to show the normality of $N(G_n)$ under some extra condition on the structure of the graph. However, under the uniform coloring scheme, using the method of moments argument and extremal combinatorics estimates from Alon [2], it can be shown that the normality for $N(G_n)$ is universal whenever $c_n \to \infty$.

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

$$\frac{1}{\sqrt{|E(G_n)|/c_n}} \left(N(G_n) - \frac{|E(G_n)|}{c_n} \right) \overset{D}{\to} N(0, 1).$$

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 shows up. Bounding this quantity requires extensions of Alon's [2, 3] results to multi-graphs and leads to some other results in graph theory which may be of independent interest. Error rates for the above CLT were proved recently by Fang [26].

1.2. *Normality for fixed number of colors.* The limiting distribution of $N(G_n)$ might not be asymptotically normal if $|E(G_n)| \to \infty$, but the number of colors $c_n = c$ is fixed. In fact, in Example 6.1, we construct a graph sequence $\{G_n\}_{n \geq 1}$,

which if uniformly colored with $c = 2$ colors, the limiting distribution of $N(G_n)$ is a mixture of a standard normal and point mass at 0.

However, for many graph sequences the limiting distribution is asymptotically normal. To characterize graph sequences for which asymptotic normality holds, we introduce the following definition.

DEFINITION 1.1.   A deterministic sequence of graphs $\{G_n\}_{n \geq 1}$ is said to satisfy the *asymptotic 4-cycle free (ACF4) condition* if

(1.4) $$N(G_n, C_4) = o(|E(G_n)|^2).$$

A sequence of random graphs $\{G_n\}_{n \geq 1}$ is said to satisfy the *ACF4 condition in probability* if

(1.5) $$N(G_n, C_4) = o_P(|E(G_n)|^2).$$

[The notation $X_n = o_P(a_n)$ means that X_n/a_n converges to zero in probability as $n \to \infty$.]

The following theorem shows that the ACF4 condition is necessary and sufficient for the normality of $N(G_n)$ when the number of colors c is fixed. The proof proceeds along similar lines as in Theorem 1.2. However, in this case, more careful estimates are required to bound the number of multi-subgraphs of G_n.

THEOREM 1.3.   *Let $G_n \in \mathscr{G}_n$ be a random graph sampled according to some probability distribution over $\mathscr{G}_n$. Then for any uniform c-coloring of G_n, with $c \geq 2$, fixed and $|E(G_n)| \xrightarrow{P} \infty$,*

(1.6) $$\frac{1}{\sqrt{|E(G_n)|/c}} \left(N(G_n) - \frac{|E(G_n)|}{c} \right) \xrightarrow{D} N\left(0, 1 - \frac{1}{c} \right),$$

if and only if $\{G_n\}_{n \geq 1}$ satisfies the ACF4 condition in probability.

For the case $c = 2$, the random variable $N(G_n) - |E(G_n)|/c$ can be rewritten as a quadratic form as follows: Let $\underline{X} = (X_1, X_2, \ldots, X_{|V(G_n)|})'$ be a vector of independent Rademacher random variables, then

$$N(G_n) - \frac{|E(G_n)|}{c} \stackrel{D}{=} \frac{1}{2} \sum_{i \leq j} A_{ij}(G_n) X_i X_j.$$

The asymptotic normality of $V_n := \sum_{i \leq j} a_{ij} X_i X_j$ for a general sequence of symmetric matrices $\{A_n = ((a_{ij}))\}$ and i.i.d. real-valued random variables $X_1, X_2, \ldots, X_n$ with zero mean, unit variance and finite fourth moment, is a well-studied problem. The classical sufficient condition for asymptotic normality is (refer to Rotar [43], Hall [32], de Jong [19] for further details)

(1.7) $$\lim_{n \to \infty} \sigma_n^{-4} \mathbb{E}(V_n - \mathbb{E}(V_n))^4 = 3 \quad \text{and} \quad \lim_{n \to \infty} \sigma_n^{-2} \max_{i \in [n]} \sum_{j=1}^{n} a_{ij}^2 = 0,$$

where $\sigma_n^2 = \frac{1}{2} \operatorname{tr}(\boldsymbol{A}_n^2) = \operatorname{Var}(V_n)$. This condition is also necessary when $X_1, X_2,$ $\ldots, X_n$ are i.i.d. $N(0, \tau^2)$, for some $\tau > 0$. Error bounds were obtained by Götze and Tikhomirov [29, 30]. Recently, Nourdin et al. [40] showed that the first condition in (1.7) is sufficient for asymptotic normality whenever $\mathbb{E}(X_1^4) \geq 3$. This is an example of the *Fourth Moment Phenomenon* which asserts that for many sequences of nonlinear functionals of random fields, a CLT is simply implied by the convergence of the corresponding sequence of fourth moments (refer to Nourdin, Peccati and Reinert [42] and the references therein for further details).

For the case where $X_1, X_2, \ldots, X_n$ are Rademacher variables, error bounds were also proved by Chatterjee [12]. Later, Nourdin, Peccati and Reinert [41] showed that in this case, the first condition in (1.7) is necessary and sufficient for asymptotic normality. For the special case where the matrix $\boldsymbol{A}_n$ is the adjacency matrix of a graph, it is easy to see that the fourth moment condition in (1.7) is equivalent to the ACF4 condition, making Theorem 1.3 an instance of the Fourth Moment Phenomenon. For the case of graphs, Theorem 1.3 reconstructs the result about quadratic forms for $c = 2$, and extends it for general $c \geq 3$ when $N(G_n)$ can no longer be written as a single quadratic form.

1.3. *Limiting distribution for converging sequence of dense graphs.* As discussed above, asymptotic normality of the number of monochromatic edges, for a fixed number of colors, does not hold when the ACF4 condition is not satisfied. In particular, no sequence of dense graphs G_n with $\Theta(n^2)$ edges satisfies the ACF4 condition and hence the limiting distribution of $N(G_n)$ is not normal. This raises the question of characterizing the limiting distribution of $N(G_n)$ for dense graphs. Recently, Lovász and coauthors [8, 9, 38] developed a limit theory for dense graphs. Using results from this theory, we obtain the limiting distribution of $N(G_n)$ for any converging sequence of dense graphs.

1.3.1. *Graph limit theory.* Graph limit theory connects various topics such as graph homomorphisms, Szemerédi regularity lemma, quasirandom graphs, graph testing and extremal graph theory, and has even found applications in statistics and related areas [13]. For a detailed exposition of the theory of graph limits refer to Lovász [38]. In the following, we mention the basic definitions about convergence of graph sequences. If F and G are two graphs, then

$$t(F, G) := \frac{|\operatorname{hom}(F, G)|}{|V(G)|^{|V(F)|}},$$

where $|\operatorname{hom}(F, G)|$ denotes the number of homomorphisms of F into G. In fact, $t(F, G)$ denotes the probability that a random mapping $\phi : V(F) \to V(G)$ defines a graph homomorphism. The basic definition is that a sequence G_n of graphs converges if $t(F, G_n)$ converges for every graph F.

There is a natural limit object in the form of a function $W \in \mathscr{W}$, where $\mathscr{W}$ is the space of all measurable functions from $[0, 1]^2$ into $[0, 1]$ that satisfy

$W(x, y) = W(y, x)$ for all x, y. Conversely, every such function arises as the limit of an appropriate graph sequence. This limit object determines all the limits of subgraph densities: if H is a simple graph with $V(H) = \{1, 2, \ldots, |V(H)|\}$, let

$$t(H, W) = \int_{[0,1]^{|V(H)|}} \prod_{(i,j) \in E(H)} W(x_i, x_j) \, dx_1 \, dx_2 \cdots dx_{|V(H)|}.$$

A sequence of graphs $\{G_n\}_{n \geq 1}$ is said to converge to W if for every finite simple graph H,

(1.8)
$$\lim_{n \to \infty} t(H, G_n) = t(H, W).$$

The limit objects, that is, the elements of $\mathcal{W}$, are called *graph limits* or *graphons*. A finite simple graph G on $[n]$ can also be represented as a graph limit in a natural way: Define $f^G(x, y) := \mathbf{1}\{(\lceil nx \rceil, \lceil ny \rceil) \in E(G)\}$, that is, partition $[0, 1]^2$ into n^2 squares of side length $1/n$, and define $f^G(x, y) = 1$ in the (i, j)th square if $(i, j) \in E(G)$ and 0 otherwise. Observe that $t(H, f^G) = t(H, G)$ for every simple graph H and therefore the constant sequence G converges to the graph limit f^G. It turns out that the notion of convergence in terms of subgraph densities outlined above can be suitably metrized using the so-called *cut distance* (refer to [38], Chapter 8).

Every function $W \in \mathcal{W}$ defines an operator $T_W : L_2[0, 1] \to L_2[0, 1]$, by

(1.9)
$$(T_W f)(x) = \int_0^1 W(x, y) f(y) \, dy.$$

T_W is a Hilbert–Schmidt operator, which is compact and has a discrete spectrum, that is, a countable multiset of nonzero real eigenvalues $\{\lambda_i(W)\}_{i \in \mathbb{N}}$. In particular, every nonzero eigenvalue has finite multiplicity and

(1.10)
$$\sum_{i=1}^{\infty} \lambda_i^2(W) = \int_{[0,1]^2} W(x, y)^2 \, dx \, dy := \|W\|_2^2.$$

1.3.2. *Limiting distribution of $N(G_n)$ for dense graphs.* A sequence of random graphs $\{G_n\}_{n \geq 1}$ is said to converge to a random graphon W in distribution, if the sequence

$$\{t(H, G_n) : H \text{ is a finite simple graph}\} \overset{D}{\to} \{t(H, W) : H \text{ is a finite simple graph}\}.$$

The above convergence is defined by the convergence of finite dimensional distributions, that is, for any finite collection of simple graphs $\mathcal{H}$ we have

$$\{t(H, G_n) : H \in \mathcal{H}\} \overset{D}{\to} \{t(H, W) : H \in \mathcal{H}\}.$$

Under the assumption that the sequence of random graphs G_n converges to a random graphon W in distribution, one can derive the limiting distribution of $N(G_n)$

in terms of the eigenvalues of T_W, whenever $\int_{[0,1]^2} W(x, y) \, dx \, dy > 0$ almost surely. Since convergence in cut metric implies

$$(1.11) \qquad \frac{1}{n^2} |E(G_n)| \xrightarrow{P} \frac{1}{2} \int_{[0,1]^2} W(x, y) \, dx \, dy,$$

the positivity condition just ensures that the random graph sequence G_n has $\Theta(n^2)$ edges with high probability, that is, it is dense.

THEOREM 1.4. *Let* $\{G_n\}_{n \geq 1}$ *be a sequence of random graphs converging in distribution to a random graphon* $W \in \mathcal{W}$, *such that almost surely* $\int_{[0,1]^2} W(x, y) \, dx \, dy > 0$. *Then for any uniform c-coloring of* G_n, *with* $c \geq 2$ *fixed and* $|E(G_n)| \xrightarrow{P} \infty$,

$$(1.12) \qquad \frac{1}{\sqrt{2|E(G_n)|}} \left(N(G_n) - \frac{|E(G_n)|}{c} \right) \xrightarrow{D} \frac{1}{2c} \sum_{i=1}^{\infty} \left(\frac{\lambda_i(W)}{(\sum_{j=1}^{\infty} \lambda_j^2(W))^{\frac{1}{2}}} \right) \xi_i,$$

where $\{\xi_i\}_{i \in \mathbb{N}}$ *are i.i.d.* $\chi^2_{(c-1)} - (c-1)$ *random variables independent of* W.

This theorem gives a characterization of the limiting distribution of the number of monochromatic edges for all converging sequences of dense graphs. As before, the main idea of the proof of Theorem 1.4 is moment comparison. However, in this case, the conditional central moments of $N(G_n)$ are compared with the conditional moments of a random variable obtained by replacing the color at every vertex with an independent and appropriately chosen c-dimensional normal, which is then shown to converge in distribution to a weighted sum of independent centered chi-square $\chi^2_{(c-1)}$ random variables. Refer to Beran [6] for more about such distributions.

1.4. *Connections to extremal combinatorics.* The combinatorial quantity that shows up in moment computations for the above theorems 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 [24] 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 [27] extended this result to hypergraphs[1] 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]:

---

[1] A hypergraph is a pair $H = (V(H), E(H))$, where $V(H)$ is a set of vertices, and $E(H)$ is a set of nonempty subsets of $V(H)$ called *hyperedges*. If $E(H)$ consists only of 2-element subsets of $V(H)$, then H is a graph.

THEOREM 1.5 (Theorem B, Alon [2]).   *If H has a spanning subgraph which is a disjoint union of cycles and isolated edges, then*

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

where $|\operatorname{Aut}(H)|$ denotes the number of automorphisms of H.

The above theorem calculates the exact asymptotic behavior of $N(\ell, H)$ for graphs H that 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 [44]. For a graph with no isolated vertex, Alon [3] showed that $\gamma(H) \leq |E(H)|$, and the equality holds if and only if H is a disjoint union of stars. In Observation 4.1, it will be shown that $\gamma(H) \leq |V(H)| - \nu(H)$, where $\nu(H)$ is the number of connected components of H and the condition for equality remains the same. This is 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 for $c \to \infty$ 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. This result is true for simple graphs as well as for 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 follows easily from the definition of $\gamma(H)$, it is a fortunate coincidence, as it is exactly what is needed in the proof of universal normality.

The role of cycle counts is crucial for the asymptotic normality when the number of colors is fixed. As the sequence $N(G_n)$ is uniformly integrable, the fourth moment condition and Theorem 1.3 imply the convergence of all other moments. In the language of graphs, this is equivalent to saying that the ACF4 condition implies that $N(G_n, C_g) = o(|E(G_n)|^{g/2})$, for all $g \geq 3$. This means that if number of 4-cycles in a graph is sub-extremal, then the counts of all other cycles are also sub-extremal. A combinatorial proof of this result and the similarities to results in pseudo-random graphs [15, 17], where the 4-cycle count plays a central role, are discussed in Section 6.1.

1.5. *Other monochromatic subgraphs.*   The above theorems determine the universal asymptotic behavior of the number of monochromatic edges under independent and uniform coloring of the vertices. However, the situation for the number of other monochromatic subgraphs is quite different. Even under uniform coloring, the limit need not be a Poisson mixture. To this end, for $r \geq 1$, denote by $K_{1,r}$ the

complete bipartite graph with one vertex on one side and r vertices on the other side. This will be referred to as the r-*star*. The following proposition shows that the number of monochromatic r-stars in a uniform vertex-coloring of an n-star converges to a polynomial in Poissons, which is not a Poisson mixture.

PROPOSITION 1.6. *Let* $G_n = K_{1,n}$, *be the star graph with* $n + 1$ *vertices.* *Under the uniform coloring distribution, the random variable* $T_{r,n}$ *that counts the* *number of monochromatic* r-*stars in* G_n *satisfies*

$$T_{r,n} \xrightarrow{D} \begin{cases} 0 & \text{if } \dfrac{n}{c_n} \to 0, \\[2mm] \infty & \text{if } \dfrac{n}{c_n} \to \infty, \\[2mm] \dfrac{X(X-1)\cdots(X-r+1)}{r!} & \text{if } \dfrac{n}{c_n} \to \lambda, \end{cases}$$

where $X \sim \text{Poisson}(\lambda)$.

A few 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_n-coloring of G_n converges in distribution to a nontrivial mixture of Poisson variables even when $|N(G_n, C_g)|/c_n^{g-1}$ converges to a fixed number λ. This is in contrast to the situation for edges, where the number of monochromatic edges converges to Poisson(λ) whenever $|E(G_n)|/c_n \to \lambda$. We believe that some sort of 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 a mixture of Poissons, whenever $|N(G_n, C_g)|/c_n^{g-1} \to \lambda > 0$.

1.6. *Literature review on nonuniform colorings.*   A natural generalization of the uniform coloring distribution (1.1) 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 [46], and under the uniform coloring distribution it is precisely the proportion of proper c-colorings of G. Recently, Fadnavis [25] proved that $P_G(\underline{p})$ is Schur-concave 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 that it is properly colored is maximized under the uniform distribution, that is, $p_a = 1/c$ for all $a \in [c]$.

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 grows in an appropriate way compared to the number of certain specific subgraphs in G_n. Arratia, Goldstein and Gordon [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 a complete graph (see also Chatterjee, Diaconis and Meckes [14]). Poisson limit theorems for the number of general monochromatic subgraphs in a random coloring of a graph sequence are also given in Cerquetti and Fortini [11]. 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, most of these results need conditions on the number of certain subgraphs in G_n and the coloring distribution. Moreover, Poisson approximation holds only in the regime where the number of colors c goes to infinity with n. Under the uniform coloring distribution, the random variables in (1.2) have more independence, and using the method of moments and estimates from extremal combinatorics, we show that nice universal limit theorems hold for the number of monochromatic edges.

1.7. *Organization of the paper.* The rest of the paper is organized as follows: Section 2 proves Theorem 1.1 and Section 3 illustrates it with various examples. Section 4 discusses the connections with extremal combinatorics, fractional stable set polytope, and includes a new proof of Theorem 1.5. The proofs of Theorem 1.2 and Theorem 1.3 are in Section 5 and Section 6, respectively. The characterization of the limiting distribution for dense graphs (Theorem 1.4) is detailed in Section 7. Finally, Section 8 proves Proposition 1.6, considers other examples on counting monochromatic cycles, and discusses possible directions for future research. An appendix provides the details on conditional and unconditional convergence of random variables, and proofs of some technical lemmas.

2. Universal Poisson approximation under uniform coloring: Proof of Theorem 1.1.
In this section, we determine 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 $|E(G_n)|/c_n$, 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. Recall the definition of $M(G_n)$ in (1.3). 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 $|E(G_n)|/c_n \xrightarrow{D} Z$. The second part uses this result to complete the proof of Theorem 1.1 using some technical properties of conditional convergence (see Lemma C.1).

2.1. *Computing and comparing moments.* This section is devoted to the computation of conditional moments of $N(G_n)$ and $M(G_n)$, and their comparison. To this end, define for any fixed number k, $A \lesssim_k B$ as $A \leq C(k)B$, 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,

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

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

$$\left| \mathbb{E}\big(N(G_n)^k \mid G_n\big) - \mathbb{E}\big(M(G_n)^k \mid G_n\big) \right|$$

(2.1)
$$\lesssim_k \sum_{\substack{H \in \mathcal{H}_k, \\ H \text{ has a cycle}}} N(G_n, H) \cdot \frac{1}{c_n^{|V(H)|-\nu(H)}},$$

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

PROOF.   Using the multinomial expansion, we have

$$\mathbb{E}\big(N(G_n)^k \mid G_n\big)$$

(2.2)
$$= \sum_{(i_1, j_1) \in E(G_n)} \sum_{(i_2, j_2) \in E(G_n)} \cdots \sum_{(i_k, j_k) \in E(G_n)} \mathbb{E}\left(\prod_{r=1}^{k} \mathbf{1}\{Y_{i_r} = Y_{j_r}\} \,\Big|\, G_n \right)$$

and

$$\mathbb{E}\big(M(G_n)^k \mid G_n\big)$$

(2.3)
$$= \sum_{(i_1, j_1) \in E(G_n)} \sum_{(i_2, j_2) \in E(G_n)} \cdots \sum_{(i_k, j_k) \in E(G_n)} \mathbb{E}\left(\prod_{r=1}^{k} Z_{i_r j_r} \right).$$

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

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

Note that in any graph H, $|E(H)| \geq |V(H)| - \nu(H)$ and equality holds if and only if H is a forest. The result now follows by taking the difference of (2.2) and (2.3) and recalling the definition of $\mathcal{H}_k$. $\square$

Lemma 2.1 shows that bounding the difference of the conditional moments of $N(G_n)$ and $M(G_n)$ entails bounding $N(G_n, H)$, for all graphs H with a cycle. The next lemma estimates the number of copies of a cycle C_g in G_n.

LEMMA 2.2.   *For any positive integer $g \geq 3$ 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 $\sum_{i=1}^{n} \lambda_i^2(G_n) = \operatorname{tr}(A^2) = 2|E(G_n)|$, where $\underline{\lambda}(G_n) = (\lambda_1(G_n), \ldots, \lambda_n(G_n))'$ is the vector of eigenvalues of $A(G_n)$. Also $\operatorname{tr}(A^g)$ counts the number of walks of length g in G_n, and so each cycle in G_n is counted $2g$ times. Thus, for any $g \geq 3$,

$$N(G_n, C_g) \leq \frac{1}{2g} \cdot \operatorname{tr}(A^g) = \frac{1}{2g} \cdot \sum_{i=1}^{n} \lambda_i^g(G_n)$$

(2.4)

$$\leq \frac{1}{2g} \left(\sum_{i=1}^{n} \lambda_i^2(G_n) \right)^{g/2} = \frac{1}{2g} \cdot (2|E(G_n)|)^{g/2}.$$

This completes the proof of the lemma. $\square$

For $a, b \in \mathbb{R}$, $a \lesssim b$, $a \gtrsim b$, and $a \asymp b$ means $a \leq C_1 b$, $a \geq C_2 b$ and $C_2 b \leq a \leq C_1 b$ for some universal constants $C_1, C_2 \in (0, \infty)$, respectively. 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. The following lemma gives a bound on $N(G_n, H)$ in terms of $|E(G_n)|$ for arbitrary subgraphs H of G_n.

LEMMA 2.3.   *For any fixed connected subgraph H, let $N(G_n, H)$ be the set of copies of H in G_n. Then*

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

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

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

PROOF.   The first bound on $N(G_n, H)$ 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 $V(G_n)$, giving the bound

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

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

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

$$N(G_n, H) \lesssim_H N(G_n, C_g) \binom{2|E(G_n)|}{|V(H)| - g}$$

(2.7)
$$\lesssim_H N(G_n, C_g) |E(G_n)|^{|V(H)| - g}$$

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

where the last step uses Lemma 2.2. $\square$

The *girth* of a graph G, to be denoted by $\mathrm{gir}(G)$, is the length of its shortest cycle. A graph with no cycles has infinite girth. If the graph has a cycle, then $\mathrm{gir}(G) \in [3, \infty)$. The above lemmas now imply that the conditional moments of $M(G_n)$ and $N(G_n)$ are asymptotically close, whenever $|E(G_n)|/c \xrightarrow{D} Z$.

LEMMA 2.4. *Let* $M(G_n)$ *and* $N(G_n)$ *be as defined in* (1.2) *and* (1.3), *with* $|E(G_n)|/c \xrightarrow{D} Z$, *then for every fixed* $k \geq 1$,

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

PROOF.   By Lemma 2.1

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

where $v(H)$ 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 $N(G_n, H) = o_P(c_n^{|V(H)| - v(H)})$.

To this end, fix $H \in \mathcal{H}_k$ and let $H_1, H_2, \ldots, H_{v(H)}$ be the connected components of H. Since $g := \mathrm{gir}(H) \geq 3$, assume without loss of generality that H_1 has a cycle of length g. Lemma 2.3 then implies that

$$N(G_n, H) \leq \prod_{i=1}^{v(H)} N(G_n, H_i)$$

(2.8)
$$\lesssim_H |E(G_n)|^{|V(H_1)| - g/2} \prod_{i=2}^{v(H)} E(G_n)^{V(H_i) - 1}$$

$$\lesssim_H |E(G_n)|^{|V(H)| - v(H) + 1 - g/2}$$

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

2.2. *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 $|E(G_n)|/c_n$ are treated separately as follows:

2.2.1. $\frac{1}{c_n} \cdot |E(G_n)| \overset{P}{\to} 0$.   In this case,

$$\mathbb{P}\big(N(G_n) > 0 \mid G_n\big) \leq \mathbb{E}\big(N(G_n) \mid G_n\big) = |E(G_n)|/c_n \overset{P}{\to} 0,$$

and the result follows.

2.2.2. $\frac{1}{c_n} \cdot |E(G_n)| \overset{P}{\to} \infty$.   Using pairwise independence of the collection of random variables $(\mathbf{1}\{Y_i = Y_j\}, 1 \leq i < j \leq n)$ gives

$$\mathbb{E}\big(N(G_n)^2 \mid G_n\big) = \frac{|E(G_n)|^2}{c_n^2} + \frac{|E(G_n)|}{c_n}\left(1 - \frac{1}{c_n}\right),$$

and so

$$\frac{\mathbb{E}(N(G_n)^2 \mid G_n)}{\mathbb{E}(N(G_n) \mid G_n)^2} \overset{P}{\to} 1.$$

This implies that $N(G_n)/\mathbb{E}(N(G_n) \mid G_n)$ converges in probability to 1, and so $N(G_n)$ converges to ∞ in probability, as $\mathbb{E}(N(G_n) \mid G_n) = \frac{1}{c_n} \cdot |E(G_n)| \overset{P}{\to} \infty$.

2.2.3. $\frac{1}{c_n} \cdot |E(G_n)| \overset{D}{\to} Z$, *where Z is some random variable.*   In this regime, the limiting distribution of $N(G_n)$ is a mixture of Poissons. As the Poisson distribution can be uniquely identified by moments, from Lemma 2.4 it follows that conditional on $\{|E(G_n)|/c_n \to \lambda\}$, $N(G_n)$ converges to Poisson(λ) for every $\lambda > 0$. However, this does not immediately imply the unconditional convergence of $N(G_n)$ to a mixture of Poissons. In fact, a technical result, detailed in Lemma C.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 Poisson with mean Z, to be denoted as Poisson(Z), if there exists a nonnegative random variable Z such that

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

The following lemma shows that $M(G_n)$ converges to Poisson(Z) and satisfies the technical condition needed to apply Lemma C.1.

LEMMA 2.5.   *Let $M(G_n)$ be as defined in (1.3) and $\frac{1}{c_n} \cdot |E(G_n)| \overset{D}{\to} Z$. Then* $M(G_n) \overset{D}{\to}$ Poisson(Z), *and further for any $\varepsilon > 0, t \in \mathbb{R}$,*

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

PROOF.   For any $t \in \mathbb{R}$,

$$\mathbb{E}e^{itM(G_n)} = \mathbb{E}\mathbb{E}(e^{itM(G_n)} \mid G_n) = \mathbb{E}\left(1 - \frac{1}{c_n} + \frac{e^{it}}{c_n}\right)^{|E(G_n)|} = \mathbb{E}R_n,$$

where $R_n := (1 - \frac{1}{c_n} + \frac{e^{it}}{c_n})^{|E(G_n)|}$ satisfies $|R_n| \leq 1$. Since

$$\log R_n = |E(G_n)| \log\left(1 - \frac{1}{c_n} + \frac{e^{it}}{c_n}\right)$$

$$= |E(G_n)|\left(\frac{e^{it} - 1}{c_n} + O\left(\frac{1}{c_n^2}\right)\right)$$

$$\xrightarrow{D} (e^{it} - 1)Z,$$

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

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

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

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

as $k \to \infty$. Now applying Fatou's lemma gives

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

from which the lemma follows by taking limit as $k \to \infty$ on both sides.   □

Now, take $X_n = M(G_n)$ and $Y_n = N(G_n)$, and observe that (C.1) and (C.2) hold by Lemma 2.4 and Lemma 2.5, 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 2.1.   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 [31] showed that certain empirical distributions of accidents are well-approximated by a Poisson mixture. Le-Cam and Traxler [37] proved asymptotic properties of random variables distributed as a mixture of Poissons. Poisson mixtures are widely used in modeling count panel data (refer to the recent paper of Burda, Harding and Hausman [10] and the references therein), and have appeared in other applied problems as well [16].

3. 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 3.1 (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, 18, 20, 21] and the references therein.

EXAMPLE 3.2 (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 [28, 36]. 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 the bound $\mathbb{P}(N(G_n) > 0) \geq \frac{\mathbb{E}(N(G_n))^2}{\mathbb{E}(N(G_n)^2)}$, the probability of a match is at least

$1 - \frac{1}{\lambda} = 85\%$. Note that this bound only uses an estimate on the number of edges in the graph. Moreover, assuming the Poisson approximation, the chance of a match is approximately $1 - e^{-\lambda} \approx 99.8\%$, which means that with high probability 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.

This is an example of the *law of truly large numbers* [21], which says that when enormous numbers of events and people and their interactions cumulate over time, almost any outrageous event is bound to occur. The point is that truly rare events are bound to be plentiful in a population of 400 million people. If a coincidence occurs to one person in a million each day, then we expect 400 occurrences a day and close to 140,000 such occurrences a year.

EXAMPLE 3.3 (Galton–Watson trees). An example where the limiting distribution of $N(G_n)$ is indeed a Poisson mixture arises in the uniform coloring of Galton–Watson trees with general offspring distribution. Let G be the Galton Watson tree with offspring distribution Γ, and G_n be the tree obtained by truncating G at height n. Assuming $\mu := \int x \, d\Gamma > 1$, ensures that the Galton Watson tree G is infinite with positive probability.

For $i \geq 0$, let S_i denote the size of the ith generation, and the total progeny up to time n by $Y_n := \sum_{i=0}^{n} S_i$. Assuming that the population starts with one off-spring at time 0, that is, $S_0 \equiv 1$, it is well known that S_n/μ^n is a nonnegative martingale ([23], Lemma 4.3.6). It converges almost surely to a finite valued random variable S_∞, by [23], Theorem 4.2.9, which readily implies Y_n/μ^{n+1} converges almost surely to $S_\infty/(\mu - 1)$. If the graph G is infinite, Theorem 1.1 applies:[2]

$$N(G_n) \xrightarrow{D} \begin{cases} 0 & \text{if } \dfrac{\mu^n}{c_n} \to 0, \\ \infty & \text{if } \dfrac{\mu^n}{c_n} \to \infty, \\ \text{Poisson}\left(\dfrac{b\mu}{\mu - 1} S_\infty\right) & \text{if } \dfrac{\mu^n}{c_n} \to b, \end{cases}$$

where $\text{Poisson}(\frac{b\mu}{\mu-1} S_\infty)$ is a Poisson distribution with the random mean $\frac{b\mu}{\mu-1} S_\infty$ as defined in (2.9). Note that this also includes the case where G is finite and $c_n \to \infty$: In this case, $S_\infty = 0$ and $N(G_n) \to 0$ in probability. It is worth noting

---

[2]In this case, the number of vertices in G_n is also random. However, the proof of Theorem 1.1 goes through as long as $|V(G_n)|$ goes to $+\infty$ in probability, which happens whenever G is infinite.

that $S_\infty \equiv 0$ if and only if $\mathbb{E}(\xi \log \xi) = \infty$ ([23], Theorem 4.3.10). Thus, to get a nontrivial limit the necessary and sufficient condition is $\mathbb{E}(\xi \log \xi) < \infty$.

4. Connections to extremal graph theory. In the method of moment calculations of Lemma 2.1, 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 $H = (V(H), E(H))$, we have

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

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 [24] 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. To state his result requires the following definition.

DEFINITION 4.1. For any graph H, denote the neighborhood of a set $S \subseteq V(H)$ by $N_H(S) = \{v \in V(H) : \exists u \in S \text{ and } (u, v) \in E(H)\}$. Note that elements of S may or may not be in $N_H(S)$: an element $a \in S$ is in $N_H(S)$ if and only if there exists $b \in S$ such that $(a, b) \in E(H)$.

THEOREM 4.1 (Alon [2]). *For a fixed graph H, there exist two positive constants $C_1 = C_1(H)$ and $C_2 = C_2(H)$ such that for all $\ell \geq |E(H)|$,*

$$(4.1) \qquad C_1 \ell^{\gamma(H)} \leq N(\ell, H) \leq C_2 \ell^{\gamma(H)},$$

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

Hereafter, unless mentioned otherwise, we shall only consider graphs H with no isolated vertex. 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 observation gives a slight strengthening of Theorem 1 of [3].

OBSERVATION 4.1. *For any graph H with no isolated vertex,*

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

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 (2.8) and Theorem 4.1 $\gamma(H) \leq |V(H)| - \nu(H) + 1 - g/2 < |V(H)| - \nu(H)$. 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.   □

4.1. *Another proof of Theorem* 1.1.   The main step in the proof of Theorem 1.1 is the moment comparison in Lemma 2.4. In fact, Theorem 4.1 and Observation 4.1 give a direct proof of Lemma 2.4, which does not require the subgraph counting Lemmas 2.2 and 2.3.

With $N(G_n)$ and $M(G_n)$ as defined in (1.2) and (1.3), and $|E(G_n)|/c_n \xrightarrow{D} Z$, for every fixed $k \geq 1$, we have

$$\left| \mathbb{E}\left(N(G_n)^k \mid G_n\right) - \mathbb{E}\left(M(G_n)^k \mid G_n\right) \right|$$

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

$$\lesssim_k \sum_{\substack{H \in \mathcal{H}_k, \\ H \text{ has a cycle}}} \frac{|E(G_n)|^{\gamma(H)}}{c_n^{|V(H)| - \nu(H)}},$$

where the last inequality follows from Theorem 4.1. As the sum is over all graphs H that are not a forest, it follows from Observation 4.1 that $\gamma(H) < |V(H)| - \nu(H)$. Therefore, every term in the sum goes to zero in probability as $n \to \infty$, and, since $H \in \mathcal{H}_k$ is a finite sum, Lemma 2.4 follows. The proof of Theorem 1.1 can now be completed as in Section 2.2.

4.2. *A new proof of Theorem* 1.5 *using Lemma* 2.2.   This section gives a short proof of Theorem 1.5 using Lemma 2.2.

4.2.1. *Proof of Theorem* 1.5.   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 $|E(G)| = \ell$,

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

Case 2. F_i is a cycle of length $g \geq 3$. Then by Lemma 2.2

$$(4.3) \qquad N(G, F_i) \leq \frac{1}{2g} \cdot (2\ell)^{g/2} = \frac{1}{|\operatorname{Aut}(F_i)|} \cdot (2\ell)^{|V(F_i)|/2},$$

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

Now, (4.2) and (4.3) imply that

$$
\begin{aligned}
N(G, F) \leq \prod_{i=1}^{q} N(G, F_i) &\leq \frac{1}{\prod_{i=1}^{q} |\operatorname{Aut}(F_i)|} \cdot (2\ell)^{|V(H)|/2} \\
&= \frac{1}{|\operatorname{Aut}(F)|} \cdot (2\ell)^{|V(H)|/2},
\end{aligned}
$$

(4.4)

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

$$(4.5) \qquad N(\ell, H) \leq \frac{x(H, F)}{N(H, F)} N(\ell, F).$$

Similarly, $N(K_v, H) = \frac{x(H,F)}{N(H,F)} N(K_v, F)$ (see [2], Lemma 6) and it follows from (4.5) that,

$$
\begin{aligned}
N(\ell, H) &\leq \frac{N(K_v, H)}{N(K_v, F)} N(\ell, F) = \frac{|\operatorname{Aut}(F)|}{|\operatorname{Aut}(H)|} N(\ell, F) \\
&\leq \frac{1}{|\operatorname{Aut}(H)|} (2\ell)^{|V(H)|/2},
\end{aligned}
$$

where the last inequality follows from (4.4).

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

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

thus completing the proof. $\square$

4.3. *Connections to fractional stable set and a structural lemma.* A *multi-graph G* is a graph with no self-loops, where there might be more than one edge between two vertices. In this case, the set of edges $E(G)$ is a multi-set where the edges are counted with their multiplicities. Friedgut and Kahn [27] extended Alon's result to hypergraphs, and identified the exponent $\gamma(H)$ as the *fractional stable number* of the hypergraph H, which is the defined as the solution of a linear programming problem. Using this alternative definition, we can define $\gamma(H)$ for any multi-graph as follows:

$$(4.6) \qquad \gamma(H) = \arg \max_{\substack{\phi \in V_H[0,1], \\ \phi(x)+\phi(y)\leq 1 \text{ for } (x,y)\in E(H)}} \sum_{v\in V(H)} \phi(v),$$

where $V_H[0, 1]$ is the collection of all functions $\phi : V(H) \to [0, 1]$. It is clear that $\gamma(H) = \gamma(H_S)$, where H_S is the simple graph obtained from H by replacing the edges between the vertices which occur more than once by a single edge.

The polytope P_H defined by the constraint set of this linear program is called the *fractional stable set polytope* [44] (for more on the fractional stable number refer to [33]). The following lemma, which is an easy consequence of the definitions, relates $\gamma(H)$ to the minimum degree of H, to be denoted by $d_{\min}(H)$. As before, we will only be considering multi-graphs with no isolated vertex.

LEMMA 4.1.   *Let $H = (V(H), E(H))$ be a multi-graph with no isolated vertex and $d_{\min}(H) \geq 2$, and let $\varphi : V(H) \to [0, 1]$ be an optimal solution of the linear program* (4.6). *Then $\gamma(H) \leq \frac{1}{2}|E(H)|$. Moreover, if there exists $v \in V(H)$ such that $\varphi(v) \neq 0$ and $d(v) \geq 3$ then $\gamma(H) < \frac{1}{2}|E(H)|$.*

PROOF.   If $d_{\min}(H) \geq 2$, we have

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

which gives $\gamma(H) \leq \frac{1}{2}|E(H)|$.

Now, suppose there exists $v \in V(H)$ such that $\varphi(v) \neq 0$ and $d(v) \geq 3$. Then

$$|E(H)| \geq \sum_{(x,y)\in E(H)} \{\varphi(x) + \varphi(y)\} = \sum_{x\in V(H)} d(x)\varphi(x)$$

$$\geq 3\varphi(v) + 2 \sum_{x\in V(H)\setminus\{v\}} \varphi(x) = 2\gamma(H) + \varphi(v),$$

and the result follows since $\varphi(v) > 0$.   □

Another important property of the fractional stable set polytope is the following proposition.

PROPOSITION 4.2 ([39]).   *Let $\varphi : V(H) \to [0, 1]$ be any extreme point of the fractional stable set polytope $P(H)$. Then $\varphi(v) \in \{0, \frac{1}{2}, 1\}$, for all $v \in V(H)$.*

Let H be any multi-graph with no isolated vertex, and $\varphi : V(H) \to [0, 1]$ be an extreme point of $P(H)$ that is an optimal solution to the linear program defined in (4.6). If $\varphi(v) = 1/2$ for all $v \in V(H)$, then $\gamma(H) = |V(H)|/2$ and by Alon [2], Lemma 7, H_S, and hence H, has a spanning subgraph which is a disjoint union of cycles or isolated edges. When $\gamma(H) > |V(H)|/2$, the above proposition can be used to prove a structural result for the multi-graph H, depending on the chosen optimal function φ. To this end, partition $V(H) = V_0(H) \cup V_{1/2}(H) \cup V_1(H)$, where $V_a(H) = \{v \in V(H) : \varphi(v) = a\}$, for $a \in \{0, 1/2, 1\}$.

The following lemma gives structural properties of the subgraphs of H induced by this partition of the vertex set. The proof, which is given in Appendix A, closely follows the proof of Lemma 9 of Alon [2], but is re-formulated here in terms of the function φ. The lemma will be used later to prove the normality of $N(G_n)$ in Theorem 1.3.

LEMMA 4.2.   *Let H be a multi-graph with no isolated vertex and $\gamma(H) > |V(H)|/2$. If $\varphi : V(H) \to [0, 1]$ is an optimal solution to the linear program (4.6), then the following holds*:

   (i) *The bipartite graph $H_{01} = (V_0(H) \cup V_1(H), E(H_{01}))$, where $E(H_{01})$ is the set of edges from $V_0(H)$ to $V_1(H)$, has a matching which saturates every vertex in $V_0(H)$.*[3]

   (ii) *The subgraph of H induced by the vertices of $V_{1/2}(H)$ has a spanning subgraph which is a disjoint union of cycles and isolated edges.*

5. Universal normal approximation for uniform coloring.   Theorem 1.1 says that if $\frac{1}{c_n}|E(G_n)| \xrightarrow{P} \infty$, 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. It turns out, as in the Poisson limit theorem, the normality of the standardized random variable $N(G_n)$ is universal whenever both c_n and $|E(G_n)|/c_n$ go to infinity. This will be proved by a similar method of moments argument. For this proof, without loss of generality we can assume that $|E(G_n)| \geq c_n$ almost surely, for every n. This is because $\mathbb{P}(|E(G_n)| < c_n)$ converges to 0 as $n \to \infty$, and so replacing the law of G_n by the conditional law of $(G_n \mid \{|E(G_n)| \geq c_n\})$ does not affect any of the limiting distribution results.

---

[3]A *matching* M in a graph $H = (V(H), E(H))$ is subset of edges of $E(H)$ without common vertices. The matching M is said to saturate $A \subset V(H)$, if, for every vertex $a \in A$, there exists an edge in the matching M incident on a.

5.1. *Proof of Theorem* 1.2.   Let $G_n \in \mathcal{G}_n$ be a random graph sampled according to some probability distribution. This section proves a universal normal limit theorem for

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

(5.1)

$$= \left(\frac{|E(G_n)|}{c_n}\right)^{-\frac{1}{2}} \left(N(G_n) - \frac{|E(G_n)|}{c_n} \right).$$

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$(1/c_n)$, and set

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

(5.2)

$$= \left(\frac{|E(G_n)|}{c_n}\right)^{-\frac{1}{2}} \left(M(G_n) - \frac{|E(G_n)|}{c_n} \right).$$

5.1.1. *Comparing conditional moments.*   A multi-graph H is said to be a multi-subgraph of a simple graph G if the simple graph H_S is a subgraph of G. Recall that for multi-graph $F = (V(F), E(F))$, $\nu(F)$ denotes the number of connected components of F.

OBSERVATION 5.1.   *Let $H = (V(H), E(H))$ be a multi-graph with no isolated vertex. Let F be a multi-graph 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 $|V(F)| = |V(H)| - 2$, the edge removed must be an isolated edge, in which case the number of vertices decreases by two and the number of connected components decreases by one and the result holds.   □

The above observation helps to determine the leading order of the expected central moments for multi-subgraphs of G_n. For any multi-graph $H = (V(H), E(H))$, define

$$Z(H) = \prod_{(i,j)\in E(H)} \left\{ \mathbf{1}\{Y_i = Y_j\} - \frac{1}{c_n} \right\},$$

(5.3)

$$W(H) = \prod_{(i,j)\in E(H)} \left\{ X_{(i,j)} - \frac{1}{c_n} \right\},$$

where duplicate edges (i, j) in the multi-graph H have the same random variable $X_{(i,j)}$. Note if the graph H does not have multiple edges, that is, it is a simple graph, then $\mathbb{E}(W(H)) = 0$.

LEMMA 5.1.   *For any multi-graph $H = (V(H), E(H))$,*

$$\mathbb{E}(Z(H)) \lesssim_H \frac{1}{c_n^{|V(H)|-v(H)}} \quad and \quad \mathbb{E}(W(H)) \lesssim_H \frac{1}{c_n^{|E(H_S)|}}.$$

PROOF.   That both $Z(H)$ and $W(H)$ are independent of the labeling of the vertices is trivial, and so they are well defined. Fixing a labeling of H and expanding the product we have

$$(5.4) \qquad Z(H) = \sum_{b=0}^{|E(H)|} \frac{(-1)^b}{c_n^b} \sum_{\substack{(i_s, j_s)\in E(H), \\ s\in[|E(H)|-b]}} \prod_{s=1}^{|E(H)|-b} \mathbf{1}\{Y_{i_s} = Y_{j_s}\},$$

where the second sum is over all possible choices of $|E(H)| - b$ distinct 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 multi-subgraph of H formed by the edges $(i_1, j_1), (i_2, j_2), \ldots,$ $(i_{|E(H)|-b}, j_{|E(H)|-b})$. Then by Observation 5.1, $|V(F)| - v(F) \geq |V(H)| - v(H) - b$, and

$$(5.5) \qquad \frac{1}{c_n^b}\mathbb{E}\left(\prod_{s=1}^{|E(H)|-b} \mathbf{1}\{Y_{i_s} = Y_{j_s}\}\right) = \frac{1}{c_n^{|V(F)|-v(F)+b}} \leq \frac{1}{c_n^{|V(H)|-v(H)}}.$$

As the number of terms in (5.4) depends only on H, and for every term (5.5) holds, the result follows.

The result for $W(H)$ follows similarly, on noting that $\mathbb{E}W(H) = 0$ if H is a simple graph. The leading order of the expectation comes from the first term

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

and the number of terms depends only on H. $\square$

Expanding the product also shows that the expected central moments of $Z(H)$ and $W(H)$ are equal when the underlying simple graph is a tree.

LEMMA 5.2.   *For any multi-graph $H = (V(H), E(H))$ such that the underlying simple graph H_S is a forest we have $\mathbb{E}(Z(H)) = \mathbb{E}(W(H))$.*

PROOF.   Recall the definitions of $Z(H)$ and $W(H)$ from (5.3). Since both products factorize over connected components of H, without loss of generality

we may assume that H is connected, and the underlying simple graph H_S is a tree. By (5.4)

$$\mathbb{E}\big(Z(H)\big) = \sum_{b=0}^{|E(H)|} \frac{(-1)^b}{c_n^b} \sum_{\substack{(i_s, j_s) \in E(H), \\ s \in [|E(H)|-b]}} \mathbb{E}\left(\prod_{s=1}^{|E(H)|-b} \mathbf{1}\{Y_{i_s} = Y_{j_s}\} \right)$$

$$= \sum_{b=0}^{|E(H)|} \frac{(-1)^b}{c_n^b} \sum_{\substack{(i_s, j_s) \in E(H), \\ s \in [|E(H)|-b]}} \frac{1}{c_n^{|V(F)|-v(F)}}$$

(5.6)

$$= \sum_{b=0}^{|E(H)|} \frac{(-1)^b}{c_n^b} \sum_{\substack{(i_s, j_s) \in E(H), \\ s \in [|E(H)|-b]}} \frac{1}{c_n^{|V(F_S)|-v(F_S)}}$$

$$= \sum_{b=0}^{|E(H)|} \frac{(-1)^b}{c_n^b} \sum_{\substack{(i_s, j_s) \in E(H), \\ s \in [|E(H)|-b]}} \frac{1}{c_n^{|E(F_S)|}},$$

where F is the multi-subgraph of H formed by the edges $(i_1, j_1), (i_2, j_2), \ldots,$ $(i_{|E(H)|-b}, j_{|E(H)|-b})$, and F_S is a forest. The last equality uses $|V(F)| - v(F) = |V(F_S)| - v(F_S) = |E(F_S)|$, because F_S is a forest, since H_S is a tree. The result now follows because every term in (5.6) is equal to every term in the expansion of $\mathbb{E}(W(H))$. $\square$

Using the above lemmas and results about the fractional independence number $\gamma(H)$ from the previous section, the conditional moments of Z_n and W_n can be compared. For a simple graph G and a multi-graph H define

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

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

LEMMA 5.3. *Let Z_n and W_n be as defined in (5.1) and (5.2), respectively. Let $\mathcal{M}_k$ be the set of all unlabeled multi-graphs H with exactly k edges (counting multiplicity) and $d_{\min}(H) \geq 2$ and $\gamma(H) = |E(H)|/2$. Then for every $k \geq 1$,*

$$\left| \mathbb{E}(Z_n^k - W_n^k \mid G_n) - \frac{\sum_{H \in \mathcal{M}_k} M(G_n, H)(\mathbb{E}(Z(H)) - \mathbb{E}(W(H)))}{\left(\frac{|E(G_n)|}{c_n} \right)^{\frac{k}{2}}} \right| \xrightarrow{P} 0,$$

whenever $|E(G_n)|/c_n \xrightarrow{P} \infty$, irrespective of whether $c_n \to \infty$ or $c_n = c$ is fixed.

PROOF.   By the multinomial expansion,

$$\mathbb{E}(Z_n^k|G_n) - \mathbb{E}(W_n^k|G_n)$$

$$(5.7) \qquad = \left(\frac{|E(G_n)|}{c_n}\right)^{-\frac{k}{2}} \sum_{H\in\mathcal{A}_k} M(G_n, H)\big(\mathbb{E}(Z(H)) - \mathbb{E}(W(H))\big),$$

where $\mathcal{A}_k$ is the set of all unlabeled multi-graphs with exactly k edges (counting multiplicity) and no isolated vertex.

For $H \in \mathcal{A}_k$ such that $\gamma(H) < |E(H)|/2$, using Lemma 4.1 and Theorem 4.1, we get

$$\left(\frac{|E(G_n)|}{c_n}\right)^{-\frac{|E(H)|}{2}} M(G_n, H) \cdot \mathbb{E}\big(Z(H)\big)$$

$$(5.8) \qquad \lesssim_H \left(\frac{|E(G_n)|}{c_n}\right)^{-\frac{|E(H)|}{2}} \frac{N(G_n, H_S)}{c^{|V(H)|-\nu(H)}}$$

$$\lesssim_H \frac{|E(G_n)|^{\gamma(H)-\frac{1}{2}|E(H)|}}{c_n^{|V(H)|-\nu(H)-\frac{1}{2}|E(H)|}}$$

$$\lesssim_H \frac{\omega^{\gamma(H)-\frac{1}{2}|E(H)|}}{c_n^{|V(H)|-\nu(H)-\gamma(H)}} \xrightarrow{P} 0,$$

whenever $\omega = |E(G_n)|/c_n \xrightarrow{P} \infty$, since $\gamma(H) \le |V(H)| - \nu(H)$ by Observation 4.1.

Similarly, for $H \in \mathcal{A}_k$ such that $\gamma(H) < |E(H)|/2$,

$$\left(\frac{|E(G_n)|}{c_n}\right)^{-\frac{|E(H)|}{2}} M(G_n, H) \cdot \mathbb{E}\big(W(H)\big)$$

$$(5.9) \qquad \lesssim_H \left(\frac{|E(G_n)|}{c_n}\right)^{-\frac{|E(H)|}{2}} \frac{N(G_n, H_S)}{c_n^{|E(H_S)|}}$$

$$\lesssim_H \left(\frac{|E(G_n)|}{c_n}\right)^{-\frac{|E(H)|}{2}} \frac{N(G_n, H_S)}{c_n^{|V(H)|-\nu(H)}}$$

$$\xrightarrow{P} 0,$$

whenever $\omega = |E(G_n)|/c \xrightarrow{P} \infty$. Here we have used the observation

$$\big|E(H_S)\big| \ge \big|V(H_S)\big| - \nu(H_S) = \big|V(H_S)\big| - \nu(H) = \big|V(H)\big| - \nu(H)$$

in the second line.

Combining (5.8) and (5.9) it follows that the contribution of the sum in the RHS of (5.7) over $H \in \mathcal{A}_k$ such that $\gamma(H) < |E(H)|/2$ is $o_P(1)$.

If $H \in \mathcal{A}_k$ is such that $\gamma(H) > |E(H)|/2$, Lemma 4.1 implies that $d_{\min}(H) = 1$. We will now show that $d_{\min}(H) = 1$ implies $\mathbb{E}Z(H) = \mathbb{E}W(H) = 0$. Label the vertices of H arbitrarily, and without loss of generality assume that the labeled vertex 1 has degree 1, and vertex 2 is the only neighbor of 1. Therefore,

$$\mathbb{E}\big(Z(H) \mid Y_2, \ldots, Y_{|V(H)|}\big)$$

$$= \left(\mathbb{E}\big(\mathbf{1}\{Y_1 = Y_2\} \mid Y_2, \ldots, Y_{|V(H)|}\big) - \frac{1}{c_n} \right)$$

$$\times \prod_{\substack{(i,j) \in E(H), \\ (i,j) \neq (1,2)}} \left\{ \mathbf{1}\{Y_i = Y_j\} - \frac{1}{c_n} \right\}$$

$$= 0,$$

which implies $\mathbb{E}(Z(H)) = 0$. Similarly, $\mathbb{E}(W(H)) = 0$.

Therefore, the sum on the RHS of (5.7) reduces to the set of all $H \in \mathcal{A}_k$ such that $\gamma(H) = |E(H)|/2$ and $d_{\min}(H) \geq 2$. Since this set is precisely $\mathcal{M}_k$, the proof of the lemma is complete. $\square$

A consequence of the proof of the above lemma, is that the moments $\mathbb{E}(Z_n^k \mid G_n)$ are bounded.

COROLLARY 5.1.   *Let Z_n be as defined in* (5.1). *Then for any $k \geq 1$ fixed there exists a constant C_k such that $\mathbb{E}(Z_n^k \mid G_n) \leq C_k$, if $|E(G_n)|/c_n \xrightarrow{P} \infty$, irrespective of whether $c_n \to \infty$ or $c_n = c$ is fixed.*

PROOF.   By the multinomial expansion, we have

$$(5.10) \qquad \mathbb{E}(Z_n^k \mid G_n) = \left(\frac{|E(G_n)|}{c_n} \right)^{-\frac{k}{2}} \sum_{H \in \mathcal{A}_k} M(G_n, H) \mathbb{E}\big(Z(H)\big),$$

where $\mathcal{A}_k$ is the set of all unlabeled multi-graphs with exactly k edges (counting multiplicity) and no isolated vertex.

Now, as in the proof of Lemma 5.3, if $H \in \mathcal{A}_k$ is such that $\gamma(H) > |E(H)|/2$, then $\mathbb{E}(Z(H)) = 0$. Therefore, the sum on the RHS of (5.10) reduces to the set of all $H \in \mathcal{A}_k$ such that $\gamma(H) \leq |E(H)|/2$ and $d_{\min}(H) \geq 2$. For such H, using Lemma 5.1 and Theorem 4.1, we get

$$\left(\frac{|E(G_n)|}{c_n} \right)^{-\frac{|E(H)|}{2}} M(G_n, H) \cdot \mathbb{E}\big(Z(H)\big)$$

$$\lesssim_H \left(\frac{|E(G_n)|}{c_n} \right)^{-\frac{|E(H)|}{2}} \frac{N(G_n, H_S)}{c_n^{|V(H)| - \nu(H)}}$$

$$\lesssim_H \frac{|E(G_n)|^{\gamma(H) - \frac{1}{2}|E(H)|}}{c_n^{|V(H)| - \nu(H) - \frac{1}{2}|E(H)|}}$$

$$\lesssim_H \frac{1}{c_n^{|V(H)| - \nu(H) - \gamma(H)}},$$

if $|E(G_n)|/c \xrightarrow{P} \infty$. Since $\gamma(H) \le |V(H)| - \nu(H)$ by Observation 4.1, the proof is complete. $\square$

Note that the above lemmas are true irrespective of whether $c_n \to \infty$ or $c_n = c$ is fixed. Hence, they will be relevant even when we are dealing with the case c is fixed. In the following lemma, it is shown that the remaining terms in (5.7) are also negligible when $c \to \infty$.

LEMMA 5.4. *Let Z_n and W_n be as defined in (5.1) and (5.2), respectively. If $c_n \to \infty$ and $|E(G_n)|/c_n \xrightarrow{P} \infty$, then for every fixed $k \ge 1$*

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

PROOF. Let $\mathcal{M}_k$ be as defined in part (b) of Lemma 5.3, and $\mathscr{S}_k \subset \mathcal{M}_k$ be the set of all multi-graphs H with $d_{\min}(H) \ge 2$, $|E(H)| = k$ and $\gamma(H) = |E(H)|/2 = |V(H)| - \nu(H)$. By part (b) of Lemma 5.3 and using $c_n \to \infty$ in (5.8) and (5.9) it follows that only the multi-subgraphs of G_n which are in $\mathscr{S}_k$ need to be considered. Now, for $H \in \mathscr{S}_k$,

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

Therefore, H_S is a disjoint union of stars by Observation 4.1. Moreover,

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

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, for any such graph H, $\mathbb{E}(Z(H)) = \mathbb{E}(W(H))$ by Lemma 5.2, and the result follows. $\square$

5.1.2. *Completing the proof of Theorem 1.2.* To complete the proof, it remains to show that W_n satisfies the conditions of Lemma C.1. This is verified in the following lemma.

LEMMA 5.5. *Let W_n be as defined in (5.2). Under the assumption $c \to \infty$ we have $W_n \xrightarrow{D} N(0, 1)$, and for any $\varepsilon > 0$ and $t \in \mathbb{R}$,*

$$(5.11) \qquad \limsup_{k \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\left| \frac{t^k}{k!} \mathbb{E}(W_n^k \mid G_n) \right| > \varepsilon \right) = 0.$$

PROOF.   To prove the first conclusion, let

$$\overline{W}_n := \frac{M(G_n) - \frac{|E(G_n)|}{c_n}}{\sqrt{\frac{|E(G_n)|}{c_n} - \frac{|E(G_n)|}{c_n^2}}}.$$

By the Berry–Esseen theorem

$$\left| \mathbb{P}(\overline{W}_n \leq x \mid G_n) - \Phi(x) \right| \leq K \sqrt{\frac{c_n}{|E(G_n)|}},$$

where $\Phi(x)$ is the standard normal density and $K < \infty$ is some universal constant. Now, the RHS converges to 0 in probability, and so $\overline{W}_n$ converges to $N(0, 1)$ by the Dominated Convergence theorem, giving

$$W_n = \overline{W}_n \left(1 - \frac{1}{c_n} \right)^{1/2} \xrightarrow{D} N(0, 1).$$

For the second conclusion, it suffices to show that

(5.12) $$\sup_{n \geq 1} \mathbb{E}(e^{t W_n}) < \infty$$

for all $t \in \mathbb{R}$. This is because (5.12) implies $\{W_n^{2m}\}_{n \geq 1}$ is uniformly integrable and $\lim_{n \to \infty} \mathbb{E}(W_n^{2m}) = \frac{(2m)!}{2^m m!}$ for all $m \geq 1$. Therefore, by Markov's inequality,

$$\limsup_{k \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\left| \frac{t^k}{k!} \mathbb{E}(W_n^k \mid G_n) \right| > \varepsilon \right) \leq \limsup_{k \to \infty} \limsup_{n \to \infty} \frac{1}{\varepsilon} \mathbb{E}\left(\frac{t^k}{k!} |W_n|^k \right)$$

$$\leq \limsup_{k \to \infty} \limsup_{n \to \infty} \frac{1}{\varepsilon} \frac{t^k}{k!} \mathbb{E}(|W_n|^{2k})^{\frac{1}{2}}$$

$$= 0,$$

which verifies (C.2).

It remains to verify (5.12). Define $\sigma_n^2 := |E(G_n)|/c_n$, and so

$$\mathbb{E}(e^{t W_n} \mid G_n) = e^{-\frac{t|E(G_n)|}{\sigma_n c_n}} \left(1 - \frac{1}{c_n} + \frac{e^t}{c_n} \right)^{|E(G_n)|} := R_n.$$

Therefore, for any $t \in \mathbb{R}$ we have

$$\log R_n = -\frac{t|E(G_n)|}{\sigma_n c_n} + |E(G_n)| \log\left(1 - \frac{1}{c_n} + \frac{e^{\frac{t}{\sigma_n}}}{c_n} \right)$$

$$\leq -\frac{t|E(G_n)|}{\sigma_n c_n} + |E(G_n)| \left(\frac{e^{\frac{t}{\sigma_n}} - 1}{c_n} \right)$$

$$\leq \frac{|E(G_n)|}{c_n} \left(\frac{t^2}{2\sigma_n^2} + e^t \frac{|t|^3}{6\sigma_n^3} \right) \leq \frac{t^2}{2} + \frac{e^t |t|^3}{6},$$

where the last two inequalities follows from Taylor's series and the fact that $|E(G_n)| \geq c_n$. This implies that $\mathbb{E}(R_n) < \infty$, as desired. $\square$

6. Normal limit theorem for fixed number of colors. The condition $c = c_n \to \infty$ in Theorem 1.2 is necessary for the universal normality. The following example demonstrates that $N(G_n)$ is not asymptotically normal, when the number of colors c remains fixed. Recall that, for $r, s \geq 1$, the complete bipartite graph $K_{r,s}$ consists of two disjoint sets of vertices of sizes r and s, with every vertex of the first set is connected to every vertex of the second set.

EXAMPLE 6.1.   Consider coloring the complete bi-partite graph $G_n = K_{2,n}$ with $c = 2$ colors. It is easy to see that $N(G_n)$ is $2nU_n$ or n with probability $\frac{1}{2}$ each, where $nU_n \sim \text{Bin}(n, 1/2)$. This implies that

$$n^{-\frac{1}{2}}(N(G_n) - n) \xrightarrow{D} \frac{1}{2}N(0, 1) + \frac{1}{2}\delta_0,$$

a mixture of a standard normal and point mass at 0.

Note that in the previous example the ACF4 condition is not satisfied: $N(G_n, C_4) = \binom{n}{2}$, and $\lim_{n \to \infty} \frac{N(G_n, C_4)}{|E(G_n)|^2} = 1/8$.

6.1. *The ACF4 condition.* Recall the ACF4 condition $N(G_n, C_4) = o_P(|E(G_n)|^2)$, that is, the number of copies of the 4-cycle C_4 in G_n is subextremal. The following theorem shows this implies that the number of copies of the g-cycle C_g in G_n is also sub-extremal, for all $g \geq 3$.

THEOREM 6.1.   *The ACF4 condition* (1.5) *is equivalent to the condition that* $N(G_n, C_g) = o_P(|E(G_n)|^{g/2})$, *for all* $g \geq 3$.

PROOF.   Let $A(G_n) = ((a_{ij}))$ be the adjacency matrix of the graph G_n. For any two vertices $a, b \in V(G_n)$, let $s_2(a, b)$ be the number of common neighbors of a, b. It is easy to see that

$$(6.1) \qquad \sum_{\substack{a,b \in V(G_n) \\ s_2(a,b) \geq 2}} s_2(a, b)^2 \lesssim \sum_{\substack{a,b \in V(G_n) \\ s_2(a,b) \geq 2}} \binom{s_2(a, b)}{2} \asymp N(G_n, C_4) = o_P(|E(G_n)|^2).$$

Moreover, $\sum_{a,b \in V(G_n)} s_2(a, b) \asymp N(G_n, K_{1,2}) = O(|E(G_n)|^2)$. Finally, observe that for any $m \geq 2$

$$(6.2) \qquad \sum_{\substack{a,b \in V(G_n) \\ s_2(a,b) \geq m}} s_2(a, b) \leq \sum_{\substack{a,b \in V(G_n) \\ s_2(a,b) \geq m}} s_2(a, b)^2/m = o_P(|E(G_n)|^2)/m,$$

where the last step uses (6.1).

Now fix $\varepsilon > 0$ and consider the following two cases depending on whether the length of the cycle is even or odd.

1. Suppose $g = 2h + 1 \geq 3$. It is easy to see that

$$(6.3) \qquad N(G_n, C_g) \lesssim \sum_{i_2, i_3, \ldots, i_g \in V(G_n)} s_2(i_2, i_g) \prod_{j=2}^{g-1} a_{i_j i_{j+1}}.$$

Note that for a path P_{2b-1} with $2b - 1$ edges we have $N(G_n, P_{2b-1}) = O(|E(G_n)|^b)$, as $\gamma(P_{2b-1}) = b$ (see Alon [2], Corollary 1). Then

$$(6.4) \qquad \sum_{\substack{i_2, i_3, \ldots, i_g \in V(G_n), \\ s_2(i_2, i_g) \leq \varepsilon(|E(G_n)|)^{1/2}}} s_2(i_2, i_g) \prod_{j=2}^{g-1} a_{i_j i_{j+1}}$$

$$\lesssim \varepsilon(|E(G_n)|)^{1/2} N(G_n, P_{2h-1}) \lesssim_g \varepsilon |E(G_n)|^{g/2}.$$

Also, using (6.2)

$$\sum_{\substack{i_2, i_3, \ldots, i_g \in V(G_n), \\ s_2(i_2, i_g) > \varepsilon(|E(G_n)|)^{1/2}}} s_2(i_2, i_g) \prod_{j=2}^{g-1} a_{i_j i_{j+1}}$$

$$(6.5) \qquad \leq \sum_{\substack{i_2, i_g \\ s_2(i_2, i_g) \geq \varepsilon |E(G_n)|^{1/2}}} s_2(i_2, i_g) \sum_{i_3, i_4, \ldots, i_{g-1} \in V(G_n)} \prod_{j=3}^{g-2} a_{i_j i_{j+1}}$$

$$= \varepsilon^{-1} o_P(|E(G_n)|^{3/2}) N(G_n, P_{2h-3})$$

$$= \varepsilon^{-1} o_P(|E(G_n)|^{g/2}),$$

where $N(G_n, P_{-1}) := 1$ by definition. Combining (6.4) and (6.5) with (6.3) it follows that there exists a finite constant $C(g)$, depending only on g, such that

$$(6.6) \qquad \frac{N(G_n, C_g)}{|E(G_n)|^{g/2}} \leq C(g)\varepsilon + \varepsilon^{-1} o_P(1).$$

On letting $n \to \infty$ followed by $\varepsilon \to 0$ the RHS above converges to 0 in distribution, thus proving $N(G_n, C_g) = o_P(|E(G_n)|^{g/2})$, for g odd.

2. Suppose $g = 2h \geq 4$. The result will be proved by induction on h. The base case $h = 2$ is true by assumption. Now, suppose $h \geq 2$ and $N(G_n, C_{2h}) = o_P(|E(G_n)|^h)$. For vertices $a, b \in V(G)$, let $s_h(a, b)$ be the number of paths with h edges in G_n with one end point at a and another at b. Therefore, as in (6.1)

$$(6.7) \qquad \sum_{\substack{a, b \in V(G_n) \\ s_h(a, b) \geq 2}} s_h(a, b)^2 \lesssim_h N(G_n, C_{2h}) = o_P(|E(G_n)|^h).$$

Moreover, $\sum_{a, b \in V(G_n)} s_h(a, b) \asymp N(G_n, P_h) = O(|E(G_n)|^h)$. Finally, as in (6.2),

for any $m \geq 2$

$$\tag{6.8} \sum_{\substack{a,b \in V(G_n) \\ s_h(a,b) \geq m}} s_h(a,b) \leq \sum_{a,b \in V(G_n)} s_h(a,b)^2/m = o_P(|E(G_n)|^h)/m,$$

where the last step uses (6.7).

Now, it is easy to see that

$$\tag{6.9} N(G_n, C_{2h+2}) \lesssim_h \sum_{x,y,u,v \in V(G_n)} s_h(x,y)s_h(u,v)a_{xu}a_{yv}.$$

Let $S = \{x, y, u, v \in V(G_n) : s_h(x,y)s_h(u,v) \leq \varepsilon|E(G_n)|^{h-1}\}$. Note that

$$\tag{6.10} \sum_S s_h(x,y)s_h(u,v)a_{xu}a_{yv} \leq \varepsilon|E(G_n)|^{h-1} \sum_{x,y,u,v \in V(G_n)} a_{xu}a_{yv}$$

$$\lesssim \varepsilon|E(G_n)|^{h+1}.$$

Also, using (6.8) and letting S^c denote the complement of S,

$$\sum_{S^c} s_h(x,y)s_h(u,v)a_{xu}a_{yv}$$

$$\leq \sum_{S^c}\sum_r r\mathbf{1}\{r = s_h(x,y)\}s_h(u,v)\mathbf{1}_{\{s_h(u,v) > \frac{\varepsilon|E(G_n)|^{h-1}}{r}\}}$$

$$\leq \sum_{x,y}\sum_r r\mathbf{1}\{r = s_h(x,y)\}\sum_{u,v} s_h(u,v)\mathbf{1}_{\{s_h(u,v) > \frac{\varepsilon|E(G_n)|^{h-1}}{r}\}}$$

$$\tag{6.11} \leq \sum_{x,y}\sum_r r\mathbf{1}\{r = s_h(x,y)\}\left(\frac{\sum_{u,v} s_h^2(u,v)}{\varepsilon|E(G_n)|^{h-1}/r}\right)$$

$$= \left(\frac{\sum_{u,v} s_h^2(u,v)}{\varepsilon E(G_n)^{h-1}}\right)\sum_{x,y}\sum_r r^2\mathbf{1}\{r = s_h(x,y)\}$$

$$= \left(\frac{o_P(|E(G_n)|)}{\varepsilon}\right)\sum_{x,y} s_h(x,y)^2 = \frac{o_P(|E(G_n)|^{h+1})}{\varepsilon}.$$

Combining (6.10) and (6.11) with (6.9) it follows that there exists a finite constant $C(h)$, depending only on h, such that

$$\tag{6.12} \frac{N(G_n, C_{2h+2})}{|E(G_n)|^{h+1}} \leq C(h)\varepsilon + \varepsilon^{-1}o_P(1),$$

from which letting $n \to \infty$ followed by $\varepsilon \to 0$ gives

$$\frac{N(G_n, C_{2h+2})}{|E(G_n)|^{h+1}} \xrightarrow{P} 0$$

as before. This completes the induction step, and hence completes the proof. $\quad\square$

REMARK 6.1.   The above theorem shows that if the number of 4-cycles in a graph is sub-extremal, then the number of copies of any cycle graph is also sub-extremal. This is an illustration of the fourth moment phenomenon: the convergence of all moments of $N(G_n)$ is implied solely by the convergence of the fourth moment of $N(G_n)$. In extremal combinatorics of pseudo-random graphs, 4-cycles play a similar role: the classic result of Chung, Graham and Wilson [15] asserts that if the edge and 4-cycle densities of a graph are as in a binomial random graph, then the graph is essentially pseudo-random and the density of any other subgraph is like that in a binomial random graph. Recently, Conlon, Fox and Zhao [17] proved similar results in the sparse regime.

Cycle counts in graphs are closely related to the sum of powers of eigenvalues of the adjacency matrix. If $\underline{\lambda}(G_n) = (\lambda_1(G_n), \lambda_2(G_n), \ldots, \lambda_n(G_n))'$ is the vector eigenvalues of the adjacency matrix $A(G_n)$, then $\sum_{i=1}^n \lambda_i^g(G_n)$ counts the number of closed walks of length g in the graph G_n. Analogous to the ACF4 condition, a sequence of random graphs $\{G_n\}_{n\geq 1}$ is said to satisfy the *uniform spectral negligibility (USN) condition* in probability if

$$(6.13) \qquad \frac{\max_{i\in[n]}|\lambda_i(G_n)|}{(\sum_{j=1}^n \lambda_j^2(G_n))^{\frac{1}{2}}} = o_P(1).$$

If $\boldsymbol{\lambda}(G_n) = \frac{\underline{\lambda}(G_n)}{\|\underline{\lambda}(G_n)\|_2}$, is the vector of normalized eigenvalues, then USN condition can be rewritten as $\|\boldsymbol{\lambda}(G_n)\|_\infty \xrightarrow{P} 0$.

OBSERVATION 6.1.   *If a sequence of graphs satisfies the USN condition in probability, then it also satisfies the ACF4 condition in probability.*

PROOF.   If a sequence of graphs satisfies the USN condition, then for every $g \geq 3$

$$N(G_n, C_g) \lesssim_g \sum_{i=1}^n \lambda_i^g(G_n) \leq \left(\sum_{i=1}^n \lambda_i^2(G_n)\right)\|\underline{\lambda}(G_n)\|_\infty^{g-2}$$
$$= o_P\big(\|\underline{\lambda}(G_n)\|_2^g\big)$$
$$= o_P\big(|E(G_n)|^{g/2}\big),$$

as $\|\underline{\lambda}(G_n)\|_2^2 = 2|E(G_n)|$.   □

More details about the differences between the ACF4 and USN conditions are presented in Section 6.3.

6.2. *Proof of Theorem* 1.3. We begin with the following lemma, which shows that the ACF4 condition ensures that the count of any graph H with $\gamma(H) = |E(H)|/2$ and a cycle, is sub-extremal. The proof uses the structural Lemma 4.2 to extract the cycle from the graph, which is then controlled using the ACF4 condition. The details of the proof are given in Appendix B.

LEMMA 6.1. *Let H be a multi-graph with no isolated vertex, $d_{\min}(H) \geq 2$, $\gamma(H) = |E(H)|/2$, and at least 3 vertices in one of its connected components. If the ACF4 condition in probability holds for G_n, then*

$$M(G_n, H) = o_P(|E(G_n)|^{|E(H)|/2}),$$

whenever the girth $\mathrm{gir}(H) \geq 3$.

The above result combined with part (b) of Lemma 5.3 shows that the conditional moments of Z_n and W_n, as defined in (5.1) and (5.2), are close whenever the ACF4 condition holds in probability.

LEMMA 6.2. *Let c be fixed and $\{G_n\}_{n \geq 1}$ be sequence of random graphs for which the ACF4 condition in probability holds. With Z_n and W_n as defined in (5.1) and (5.2), for every fixed $k \geq 1$*

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

PROOF. By part (b) of Lemma 5.3,

$$\mathbb{E}(Z_n^k \mid G_n) - \mathbb{E}(W_n^k \mid G_n)$$

$$= \left(\frac{|E(G_n)|}{c}\right)^{-\frac{k}{2}} \sum_{H \in \mathcal{M}_k} M(G_n, H)(\mathbb{E}(Z(H)) - \mathbb{E}(W(H))) + o_P(1),$$

where $\mathcal{M}_k$ is the set of all multi-graphs with exactly k edges (counting multiplicity), no isolated vertex, $d_{\min}(H) \geq 2$, and $\gamma(H) = |E(H)|/2$.

If $H \in \mathcal{M}_k$ is such that the girth $\mathrm{gir}(H) \geq 3$, then by Lemma 6.1 $M(G_n, H) = o_P(|E(G_n)|^{k/2})$. Therefore, the only multi-subgraphs $H \in \mathcal{M}_k$ which remain must be such that H_S is a forest. But, by Lemma 5.2, $\mathbb{E}(Z(H)) = \mathbb{E}(W(H))$ for all such multi-subgraphs, and the result follows. □

6.2.1. *Completing the proof of Theorem* 1.3. As in the proof of Lemma 5.5, it follows that $W_n \xrightarrow{D} N(0, 1 - \frac{1}{c})$, and

$$\limsup_{k \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\left|\frac{t^k}{k!}\mathbb{E}(W_n^k \mid G_n)\right| > \varepsilon\right) = 0,$$

for some $t > 0$. Therefore, $Z_n \xrightarrow{D} N(0, 1 - \frac{1}{c})$ by Lemma C.1.

To prove the necessity, assume that $\frac{N(G_n, C_4)}{|E(G_n)|^2} \overset{P}{\not\to} 0$. Therefore,

$$\limsup_{n \to \infty} \mathbb{E}\left(\frac{N(G_n, C_4)}{|E(G_n)|^2}\right) > 0.$$

By part (b) of Lemma 5.3,

$$\mathbb{E}(Z_n^4 \mid G_n) - \mathbb{E}(W_n^4 \mid G_n)$$

(6.14)
$$= \left(\frac{|E(G_n)|}{c}\right)^{-2}$$

$$\times \sum_{H \in \mathcal{M}_4} M(G_n, H)(\mathbb{E}(Z(H)) - \mathbb{E}(W(H))) + o_P(1),$$

where $\mathcal{M}_4$ be the collection of all multi-graphs with 4 edges such $d_{\min} \geq 2$ and $\gamma(H) = 2$. Also if H_S is a forest then we have $\mathbb{E}Z(H) = \mathbb{E}W(H)$. Now, it is easy to see that the only multi-graph $H \in \mathcal{M}_4$ such that H_S is not a forest is C_4.

By a direct calculation, we have

$$\mathbb{E}(W_n^4 \mid G_n) = 3\left(1 - \frac{1}{c}\right)^2 + o_P(1),$$

$$\mathbb{E}(Z(C_4)) - \mathbb{E}(W(C_4)) = \frac{1}{c^3}\left(1 - \frac{1}{c}\right),$$

which on substituting in (6.14) gives

(6.15) $\quad \mathbb{E}(Z_n^4 \mid G_n) = 3\left(1 - \frac{1}{c}\right)^2 + \frac{1}{c}\left(1 - \frac{1}{c}\right)\left(\frac{N(G_n, C_4)}{|E(G_n)|^2}\right) + o_P(1).$

Note that the $o_P(1)$-term above is bounded, as $\mathbb{E}(Z_n^4 \mid G_n)$ is bounded by Corollary 5.1. Therefore, by taking expectation and limit in (6.15),

(6.16) $$\limsup_{n \to \infty} \mathbb{E}(Z_n^4) > 3\left(1 - \frac{1}{c}\right)^2.$$

Again invoking Corollary 5.1 it follows that $\sup_{n \geq 1} \mathbb{E}(Z_n^8 \mid G_n) < \infty$, and so $\sup_{n \geq 1} \mathbb{E}Z_n^8 < \infty$. Now, suppose $Z_n \overset{D}{\to} Z \sim N(0, 1 - 1/c)$. Therefore, since $\sup_{n \geq 1} \mathbb{E}Z_n^8 < \infty$, Z_n^4 is uniformly integrable, and the convergence in law implies that $\mathbb{E}Z_n^4$ converges to $\mathbb{E}Z^4 = 3(1 - 1/c)^2$, which contradicts (6.16).

6.3. *Connections between the ACF4 and USN conditions.* Recall that for the case $c = 2$, $N(G_n)$ can be rewritten as a quadratic form in terms of the adjacency matrix $A(G_n)$. Note that when $A(G_n) = ((a_{ij}))$, then $\sum_{j=1}^n a_{ij}^2 = \sum_{j=1}^n a_{ij}$ is the degree of the vertex i in G_n. Therefore,

$$\max_{i \in [n]} \sum_{j=1}^n a_{ij}^2 = \Delta(G_n),$$

the maximum degree of a vertex in G_n. Moreover, $\text{tr}(A(G_n)) = \sum_{i,j} a_{ij}^2 = 2|E(G_n)|$, and $\sigma_n^2 = \frac{1}{2}\text{tr}(A(G_n)) = |E(G_n)|$. Therefore, in this case, the classical sufficient condition (1.7) for asymptotic normality of he statistic $V_n = \sum_{i \leq j} a_{ij} X_i X_j$, for $X_1, X_2, \ldots, X_n$ i.i.d. with zero mean and finite fourth moment, can be rewritten as

$$\lim_{n \to \infty} \sigma_n^{-4} \mathbb{E}(V_n - \mathbb{E}(V_n))^4 = 3,$$

(6.17)
$$\lim_{n \to \infty} \sigma_n^{-2} \max_{i \in [n]} \sum_{j=1}^{n} a_{ij}^2 = \lim_{n \to \infty} \frac{\Delta(G_n)}{|E(G_n)|} = 0.$$

From the proof of the necessity part of Theorem 1.3 in the previous section, it can be seen that the ACF4 condition is equivalent to the first condition in (1.7).

For the case when $X_1, X_2, \ldots, X_n$ are i.i.d. Rademacher variables, Nourdin, Peccati and Reinert [41] showed that the second condition in (6.17) is not needed for asymptotic normality. Error bounds were also proved by Chatterjee [12] using Stein's method, under the slightly stronger conditions

$$(6.18) \qquad \lim_{n \to \infty} \|\boldsymbol{\lambda}(A(G_n))\|_4^4 = 0 \quad \text{and} \quad \lim_{n \to \infty} \sigma_n^{-2} \max_{i \in [n]} \sum_{j=1}^{n} a_{ij}^2 = 0.$$

It is easy to see the first condition in (6.18) is equivalent to the USN condition $\lim_{n \to \infty} \|\boldsymbol{\lambda}(A(G_n))\|_\infty = 0$.

Even though the number of cycles in a graph is closely related to the power sum of eigenvalues, there are subtle differences between ACF4 and USN conditions. To this end, consider the following example.

EXAMPLE 6.2. Consider the star graph $K_{1,n}$ with vertices indexed by $\{0, 1, \ldots, n\}$, with the central vertex labeled 0. It is easy to see that $K_{1,n}$ does not satisfy the USN condition:

$$\underline{\lambda}(A(K_{1,n})) = (n^{\frac{1}{2}}, 0, \ldots, 0, -n^{\frac{1}{2}})' \quad \text{and} \quad \lim_{n \to \infty} \|\boldsymbol{\lambda}(A(K_{1,n}))\|_\infty = \frac{1}{\sqrt{2}} \neq 0.$$

Moreover, $\Delta(K_{1,n})/|E(K_{1,n})| = 1$, that is, the second condition in (1.7) and (6.18) is also not satisfied by $K_{1,n}$. Therefore, asymptotic normality of $N(K_{1,n})$ does not follow from results of de Jong [19] or Chatterjee [12].

However, $K_{1,n}$ has no cycles and so the ACF4 condition is satisfied. Therefore, by Theorem 1.3 (also by Nourdin, Peccati and Reinert [41]) it follows that $N(K_{1,n})$ is asymptotically normal. In fact, for $\underline{X} = (X_0, X_1, X_2, \ldots, X_n)'$ i.i.d. Rademacher

$$n^{-\frac{1}{2}} \underline{X}' A(K_{1,n}) \underline{X} = X_0 \left(\frac{\sum_{i=1}^{n} X_i}{\sqrt{n}} \right) \xrightarrow{D} X_0 \cdot Z \sim N(0, 1),$$

where $Z \sim N(0, 1)$.

Consider on the other hand, $\underline{Z} = (Z_0, Z_1, Z_2, \ldots, Z_n)'$ i.i.d. $N(0, 1)$. Then

$$n^{-\frac{1}{2}} \underline{Z} A(K_{1,n}) \underline{Z} = Z_0 \left(\frac{\sum_{i=1}^n Z_i}{\sqrt{n}} \right) \overset{D}{=} Z_0 S_0,$$

where $Z_0, S_0 \sim N(0, 1)$, which is not normally distributed. This is expected because (1.7) is necessary and sufficient when each vertex in the graph is assigned an independent normal random variable.

6.4. *Examples*. Some examples where the ACF4 condition is satisfied and asymptotic normality holds are illustrated below.

EXAMPLE 6.3 (Random regular graphs). $\mathscr{G}_n$ consists of the set all d-regular graphs on n vertices and sampling is done uniformly on this space. In this case, $|E(G_n)| = nd/2$, for all $G_n \in \mathscr{G}_n$. Theorem 1.2 gives

$$(6.19) \qquad \frac{N(G_n) - \frac{1}{c}(\frac{nd}{2})}{\sqrt{\frac{nd}{2c}(1 - \frac{1}{c})}} \overset{D}{\to} N(0, 1) \qquad \text{when } \frac{nd}{c} \to \infty \text{ and } c \to \infty.$$

Moreover, for all $G_n \in \mathscr{G}_n$, $\lambda_{\max}(G_n) = d$ and $\sum_{i=1}^n \lambda_i(G_n)^2 = 2|E(G_n)| = nd$. This implies that $\|\boldsymbol{\lambda}_n\|_\infty = O(\sqrt{d/n}) = o(1)$, whenever $d = o(n)$. Therefore, by Observation 6.1 and Theorem 1.3, (6.19) holds even when c is fixed and $d = o(n)$.

EXAMPLE 6.4 (Hypercube). An important d-regular graph, which is neither sparse nor dense, is the hypercube $Q_n = (V(Q_n), E(Q_n))$, where $V(Q_n) = \{0, 1\}^{\log_2 n}$, where $n = 2^s$ for some $s \geq 2$, and there exists an edge between two vertices if the corresponding binary vectors have Hammimg distance 1. This is a d-regular graph with $d = \log_2 n$ and $|E(Q_n)| = \frac{1}{2} n \log_2 n$. Therefore, $\|\boldsymbol{\lambda}_n\|_\infty = O(\sqrt{d/n}) = o(1)$ and by the previous example,

$$(6.20) \qquad \frac{N(G_n) - \frac{1}{c}(\frac{n \log_2 n}{2})}{\sqrt{\frac{n \log_2 n}{2c}(1 - \frac{1}{c})}} \overset{D}{\to} N(0, 1) \qquad \text{when } \frac{n \log_2 n}{c} \to \infty,$$

irrespective of whether $c \to \infty$ or c is fixed.

EXAMPLE 6.5 (Sparse inhomogeneous random graphs). Consider the following general model for sparse random graphs: every edge (i, j), with $i < j$, 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, Janson and Riordan [7]). By the law of large numbers,

$$\frac{1}{n} |E(G_n)| \overset{P}{\to} \frac{1}{2} \int_0^1 \int_0^1 f(x, y) \, dx \, dy,$$

and it is easy to see that

$$\lim_{n\to\infty} \mathbb{E}N(G_n, C_4) = \frac{1}{2} \int_{[0,1]^4} f(w, x) f(x, y) f(y, z) f(z, x) \, dw \, dx \, dy \, dz.$$

Therefore by Markov's inequality

$$\frac{N(G_n, C_4)}{|E(G_n)|^2} \xrightarrow{P} 0$$

and so by Theorem 1.2 and Theorem 1.3 we have

$$(6.21) \quad \frac{N(G_n) - \frac{|E(G_n)|}{c}}{\sqrt{\frac{n}{c}(1 - \frac{1}{c})}} \xrightarrow{D} N\left(0, \frac{1}{2} \int_0^1 \int_0^1 f(x, y) \, dx \, dy\right) \quad \text{when } \frac{n}{c} \to \infty,$$

irrespective of whether $c \to \infty$ or c is fixed. Note that this model includes as a special case the Erdős–Renyi random graphs $G(n, \lambda/n)$, by taking the function $f(x, y) = \lambda$.

7. Universal nonnormal limit for dense graphs.
The precise conditions required for the normality of the number of monochromatic edges $N(G_n)$ have been determined in the previous sections. It is also shown in Example 6.1 that when these conditions are not met, $N(G_n)$ might have nonstandard limiting distributions. However, in this section it will be shown that there is a universal characterization of the limiting distribution of $N(G_n)$ for a converging sequence of dense graphs G_n.

To this end, consider the following example where the limiting distribution of the number of monochromatic edges is determined for a complete graph.

EXAMPLE 7.1 (Complete graph).   Consider coloring the complete graph K_n with $c = 2$ colors under the uniform distribution, where c is fixed. Let $N(K_n)$ be the number of monochromatic edges of the complete graph K_n. Let U_n be the proportion of vertices of K_n with color 1. Then

$$nU_n \sim \text{Binomial}(n, 1/2) \quad \text{and} \quad n^{1/2}(U_n - 1/2) \xrightarrow{D} N(0, 1/4).$$

In this case, we have $N(K_n) = \binom{nU_n}{2} + \binom{n-nU_n}{2} = \frac{n^2}{2}(U_n^2 + (1 - U_n)^2) - \frac{n}{2}$, and so

$$\frac{1}{n}\left(N(K_n) - \frac{1}{2}\binom{n}{2}\right) = \frac{n}{2}\left(U_n^2 + (1 - U_n)^2 - \frac{1}{2}\right) - \frac{1}{4} \xrightarrow{D} \frac{1}{4}(\chi_{(1)}^2 - 1),$$

where the last convergence follows by an application of Delta method to the function $f(x) = x^2 + (1 - x)^2$, and noting that $f'(1/2) = 0$, $f''(1/2) = 4$.

This example motivates the characterization of the limiting distribution for any converging sequence of dense graphs. The limit theory of dense graphs was developed recently by Lovász and coauthors [8, 9, 38]. Using results from this limit theory, the limiting distribution of $N(G_n)$ can be characterized for any dense graph sequence G_n converging to a limit $W \in \mathcal{W}$.

7.1. *Proof of Theorem* 1.4. Write

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

As before, Theorem 1.4 will be proved by comparing the conditional moments of

$$(7.1) \qquad \Gamma_n = \frac{N(G_n) - \frac{|E(G_n)|}{c}}{\sqrt{2|E(G_n)|}}$$

with another random variable for which the asymptotic distribution can be obtained much easily. To this end, define the random variable

$$Q(G_n) := \sum_{(i,j) \in E(G_n)} \sum_{a=1}^{c} S_{ia} S_{ja},$$

with $S_{va} = X_{va} - \overline{X}_{v\cdot}$, where $\{X_{va} : v \in V(G_n), a \in [c]\}$ is a collection of independent $N(0, 1/c)$ random variables and $\overline{X}_{v\cdot} = \frac{1}{c} \sum_{a=1}^{c} X_{va}$. Note that for each $v \in V(G_n)$ the random vector $\underline{S}_v := (S_{v1}, S_{v2}, \ldots, S_{vc})'$ is a multivariate normal $N(\mathbf{0}, \frac{1}{c}\mathbf{I})$ conditioned on $\sum_{a=1}^{c} S_{va} = 0$. Also, $\{\underline{S}_v, v \in V(G_n)\}$ are independent and identically distributed random vectors.

Finally, define

$$(7.2) \qquad \Delta_n := \frac{Q(G_n)}{\sqrt{2|E(G_n)|}}.$$

7.2. *Comparing conditional moments.* The moments of Γ_n and Δ_n involve sums over multi-subgraphs of G_n. We begin with a simple observation about general multi-graphs.

OBSERVATION 7.1. *Let $H = (V(H), E(H))$ be any finite multi-graph with $d_{\min}(H) \geq 2$ and $|V(H)| = |E(H)|$. Then H is a disjoint union of cycles and isolated doubled edges.*

PROOF. Let $H_1, H_2, \ldots, H_\nu$ be the connected components of H. Note that if there exists $i \in [\nu]$ such that $|E(H_i)| < |V(H_i)|$, then H_i must be a tree, which has a vertex of degree 1. Therefore, $|E(H_i)| = |V(H_i)|$ for all $i \in [\nu]$.

Now, let F be any connected component of H, and F_S be the underlying simple graph. Since F is connected either $|E(F_S)| = |V(F)| = |E(F)|$ or $|E(F_S)| = |V(F)| - 1 = |E(F)| - 1$.

If $|E(F_S)| = |V(F)| = |E(F)|$, then F itself is a simple graph with $d_{\min}(F) \geq 2$, which implies that F is a cycle of length $|V(F)|$.

On the other hand, if $|E(F_S)| = |V(F)| - 1 = |E(F)| - 1$, then F_S is a tree. But any tree has at least two degree one vertices, and one extra edge cannot add to

both their degrees unless the tree is just an isolated edge. This implies that F_S is an isolated edge, and F is an isolated doubled edge. $\quad\square$

As in part (b) of Lemma 5.3, the following lemma identifies the set of multi-graphs for which the moments are equal.

LEMMA 7.1. *For any multi-subgraph $H = (V(H), E(H))$ of G_n define*

$$Z(H) = \prod_{(i,j)\in E(H)} \sum_{a=1}^{c} \left(\mathbf{1}\{Y_i = a\} - \frac{1}{c} \right)\left(\mathbf{1}\{Y_j = a\} - \frac{1}{c} \right),$$

$$T(H) = \prod_{(i,j)\in E(H)} \sum_{a=1}^{c} S_{ia} S_{ja}.$$

If $|V(H)| = |E(H)|$, then $\mathbb{E}(Z(H)) = \mathbb{E}(T(H))$.

PROOF. If $d_{\min}(H) = 1$, by arguments similar to Lemma 4.1, $\mathbb{E}(Z(H)) = \mathbb{E}(T(H)) = 0$. Therefore, it suffices to assume $d_{\min}(H) \geq 2$ and $|V(H)| = |E(H)|$. By Observation 7.1, H is a disjoint union of cycles and isolated doubled edges. Since both $Z(H)$ and $T(H)$ factorize over connected components, w.l.o.g. H can be assumed to be either an isolated doubled edge or a cycle. More generally, it suffices to show that $\mathbb{E}(T(H)) = \mathbb{E}(Z(H))$ for any multi-graph H with each vertex having degree 2. Now, it is easy to see that since the random variables corresponding to each vertex are independent in both the cases, it suffices to prove

$$\mathbb{E}\left(\mathbf{1}\{Y_i = a\} - \frac{1}{c} \right)\left(\mathbf{1}\{Y_i = b\} - \frac{1}{c} \right) = \mathbb{E}S_{ia} S_{ib},$$

for any $a, b \in [c]$. This follows on noting that if $a = b$ both sides above equal $\frac{1}{c}(1 - \frac{1}{c})$, whereas for $a \neq b$ both sides above equal $-\frac{1}{c^2}$. $\quad\square$

Using this lemma it can now be shown that the conditional moments of Γ_n and Δ_n are asymptotically close, whenever the random graph sequence G_n converges in distribution to $W \in \mathcal{W}$ such that $\int_{[0,1]^2} W(x, y)\, dx\, dy > 0$ almost surely, that is, G_n is dense.

LEMMA 7.2. *Suppose the sequence of random graphs $\{G_n\}_{n=1}^{\infty}$ converges in distribution to a limit $W \in \mathcal{W}$ with $\int_{[0,1]^2} W(x, y)\, dx\, dy > 0$ almost surely. Then for Γ_n and Δ_n as defined in (7.1) and (7.2), and for every fixed $k \geq 1$,*

$$\left| \mathbb{E}(\Delta_n^k \mid G_n) - \mathbb{E}(\Gamma_n^k \mid G_n) \right| \xrightarrow{P} 0.$$

PROOF.   By Equation (5.7),

$$\left|\mathbb{E}(\Delta_n^k \mid G_n) - \mathbb{E}(\Gamma_n^k \mid G_n)\right|$$

(7.3)

$$\leq \sum_{H \in \mathcal{A}_k} \frac{M(G_n, H)}{(2|E(G_n)|)^{\frac{k}{2}}} \left|\mathbb{E}(T(H)) - \mathbb{E}(Z(H))\right|,$$

where $\mathcal{A}_k$ is the collection of all unlabeled multi-subgraphs of G_n with exactly k edges and no isolated vertex. If $H \in \mathcal{A}_k$ is such that $|V(H)| > |E(H)|$, then H must have a vertex of degree 1, and $\mathbb{E}(T(H)) = \mathbb{E}(Z(H)) = 0$. Moreover, if $H \in \mathcal{A}_k$ is such that $|V(H)| = |E(H)|$, then by Lemma 7.1 $\mathbb{E}(T(H)) = \mathbb{E}(Z(H))$. Therefore, (7.3) simplifies to

$$\left|\mathbb{E}(\Delta_n^k \mid G_n) - \mathbb{E}(\Gamma_n^k \mid G_n)\right| \lesssim_k \sum_{\substack{H \in \mathcal{A}_k \\ |V(H)| < |E(H)|}} \frac{N(G_n, H)}{(2|E(G_n)|)^{\frac{k}{2}}}$$

$$\lesssim_k \sum_{\substack{H \in \mathcal{A}_k \\ |V(H)| < |E(H)|}} \frac{n^{|V(H)|}}{(2|E(G_n)|)^{\frac{|E(H)|}{2}}},$$

where the last term follows from noting that $N(G_n, H) \lesssim_H n^{|V(H)|}$ for any H and G_n. Since $\int_{[0,1]^2} W(x, y)\,dx\,dy > 0$ almost surely (1.11) gives

$$\frac{n^{|V(H)|}}{(2|E(G_n)|)^{\frac{|E(H)|}{2}}} = O_P\left(n^{|V(H)|-|E(H)|}\right),$$

which goes to zero in probability for all $H \in \mathcal{A}_k$ such that $|V(H)| < |E(H)|$. This completes the proof of the lemma. $\square$

7.2.1. *Completing the proof of Theorem* 1.4.   As the conditional moments of Γ_n and Δ_n are asymptotically close, it remains to analyze the limiting distribution of Δ_n. In this section, it will be shown that Δ_n converges to $\frac{1}{2c}\chi_{c-1}^2(W)$, where

(7.4) $$\chi_{c-1}^2(W) := \sum_{i=1}^{\infty} \left(\frac{\lambda_i^2(W)}{(\sum_{j=1}^{\infty} \lambda_j^2(W))^{\frac{1}{2}}}\right)\xi_i := \sum_{j=1}^{\infty} \tilde{\lambda}_j^2(W)\xi_i,$$

where $\{\xi_i\}_{i \in \mathbb{N}}$ are independent $\chi_{(c-1)}^2 - (c-1)$ random variables, and

$$\tilde{\lambda}_j(W) := \frac{\lambda_j(W)}{(\sum_{i=1}^{\infty} \lambda_i(W))^{\frac{1}{2}}}.$$

The first step is to show that the random variable $\chi_{c-1}^2(W)$, which is a infinite sum of centered chi-square random variables, is well defined.

PROPOSITION 7.1.   *Let $(a_1, a_2, \ldots)$ be an infinite sequence of random variables such that $\sum_{j=1}^{\infty} a_j^2 \overset{a.s.}{=} 1$. Given $(\xi_1, \xi_2, \ldots)$ independent $\chi_c^2 - c$ random variables independent of the sequence $(a_1, a_2, \ldots)$, the sum $S := \sum_{j=1}^{\infty} a_j \xi_j$ converges almost surely and in L^1. Further, for $|t| < 1/8$ the moment generating function of S is finite, and is given by*

$$\mathbb{E} e^{tS} = \mathbb{E}\left(\prod_{j=1}^{\infty} \frac{e^{-cta_j}}{(1 - 2ta_j)^{\frac{c}{2}}} \right).$$

PROOF.   By defining $\mathscr{G} := \sigma(\{a_j\}_{j \in \mathbb{N}})$ and $S_n := \sum_{j=1}^{n} a_j \xi_j$ and $\mathscr{F}_n := \sigma(\sigma(\{\xi_j\}_{j=1}^n), \mathscr{G})$, it follows that $(S_n, \mathscr{F}_n)$ is a martingale, with

$$\limsup_n \mathbb{E} S_n^2 = 2c \left(\mathbb{E} \sum_{j=1}^{\infty} a_j^2 \right) = 2c < \infty,$$

and S_n converges almost surely and in L^1 [23].

To compute the moment generating function, note that $h : [-1/2, 1/2] \mapsto \mathbb{R}$ given by $h(z) = -\log(1 - z) - z - z^2$ has a unique global maxima at $z = 0$, and so $-\log(1 - z) - z \leq z^2$ for $|z| \leq 1/2$. Therefore, for any $|t| < 1/8$,

$$\log \mathbb{E}\left(e^{2tS_n} \mid \mathscr{G}\right) = c \sum_{j=1}^{n} \left(\frac{-\log(1 - 4ta_j) - 4ta_j}{2} \right) \leq 8ct^2,$$

and so $\mathbb{E}(e^{2tS_n} \mid \mathscr{G}) \leq e^{8ct^2} < \infty$. It follows that e^{tS_n} and $\mathbb{E}(e^{tS_n} \mid \mathscr{G})$ are both uniformly integrable, and

$$\mathbb{E} e^{tS} = \lim_{n \to \infty} \mathbb{E}\left(\mathbb{E}\left(e^{tS_n} \mid \mathscr{G}\right)\right)$$

$$= \lim_{n \to \infty} \mathbb{E}\left(\prod_{j=1}^{n} \frac{e^{-cta_j}}{(1 - 2ta_j)^{\frac{c}{2}}} \right) = \mathbb{E}\left(\prod_{j=1}^{\infty} \frac{e^{-cta_j}}{(1 - 2ta_j)^{\frac{c}{2}}} \right) < \infty. \qquad \square$$

To prove Theorem 1.4 we now invoke Lemma C.1 with Γ_n and Δ_n. Lemma 7.2 shows that (C.1) holds. The following lemma takes the first step towards (C.2) by showing that the limiting distribution of Δ_n is a weighted sum of chi-square random variables.

LEMMA 7.3.   *If a sequence of random graphs $\{G_n\}_{n=1}^{\infty}$ converges in the cut-metric in distribution to a limit $W \in \mathscr{W}$, then for $|t| < c/4$*

$$\lim_{n \to \infty} \mathbb{E} e^{t\Delta_n} = \mathbb{E}\left(e^{t \frac{\chi_{c-1}^2(W)}{2c}}\right).$$

PROOF.   Let $\{\lambda_j(W)\}_{j \in \mathbb{N}}$ be the eigenvalues of the operator W as in (1.9), and $\{\tilde{\lambda}_j(W)\}_{j \in \mathbb{N}}$ as defined in (7.4). By Proposition 7.1, $\mathbb{E}e^{t\frac{\chi^2_{c-1}(W)}{2c}} < \infty$ for $|t| < c/4$, which implies that

$$\mathscr{E}(W) := \mathbb{E}\left(e^{t\frac{\chi^2_{c-1}(W)}{2c}} \mid W\right) = \left(\prod_{j=1}^{\infty} \frac{e^{-\frac{\tilde{\lambda}_j(W)t(c-1)}{2c}}}{(1 - \frac{\tilde{\lambda}_j(W)t}{c})^{\frac{c-1}{2}}}\right) \overset{a.s.}{<} \infty.$$

Now, by Fubini's theorem,

$$\log \mathscr{E}(W) = \sum_{j=1}^{\infty} \left\{ \frac{1-c}{2} \log\left(1 - \frac{\tilde{\lambda}_j(W)t}{c}\right) - \frac{\tilde{\lambda}_j(W)t(c-1)}{2c} \right\}$$

$$(7.5) \qquad = \frac{c-1}{2} \sum_{x=2}^{\infty} \frac{t^x}{xc^x} \sum_{j=1}^{\infty} \frac{\lambda_j(W)^x}{(\sum_{j=1}^{\infty} \lambda_j(W))^{\frac{x}{2}}}$$

$$= \frac{c-1}{2} \sum_{x=2}^{\infty} \frac{t^x}{xc^x} \frac{t(C_x, W)}{t(C_2, W)^{x/2}} \overset{a.s.}{<} \infty,$$

where the last step uses $t(C_g, W) = \sum_{j=1}^{\infty} \lambda_j^g(W)$ for $g \geq 2$ (see [38], Section 7.5). This implies that

$$(7.6) \qquad \lim_{N \to \infty} \sum_{x=2}^{N} \frac{c-1}{2} \frac{t^x}{xc^x} \frac{t(C_x, W)}{t(C_2, W)^{x/2}} \overset{a.s.}{=} \sum_{x=2}^{\infty} \frac{c-1}{2} \frac{t^x}{xc^x} \frac{t(C_x, W)}{t(C_2, W)^{x/2}}$$

$$= \log \mathbb{E}\left(e^{t\frac{\chi^2_{c-1}(W)}{2c}} \mid W\right).$$

Using the spectral decomposition, write the adjacency matrix $A(G_n)$ as $\sum_{j=1}^{n} \lambda_j(G_n) \underline{p}_j \underline{p}_j'$. For $a \in [c]$ and $j \in [n]$, set $\boldsymbol{S}_a := (S_{1a}, S_{2a}, \dots, S_{|V(G_n)|a})'$ and $y_{aj} = \underline{p}_j' \boldsymbol{S}_a$. Then

$$\sum_{a=1}^{c} \boldsymbol{S}_a' A(G_n) \boldsymbol{S}_a = \sum_{j=1}^{n} \lambda_j(G_n) \sum_{a=1}^{c} y_{aj}^2.$$

Since $\text{Cov}(y_{aj}, y_{bj}) = -1/c^2$ for $1 \leq a < b \leq c$, it follows that $A_j := \sum_{a=1}^{c} y_{aj}^2 \sim \frac{1}{c}\chi^2_{c-1}$. Also, since $\text{Cov}(y_{ai}, y_{aj}) = 0$ for $i \neq j$, it follows that $\{A_j\}_{j=1}^{n}$ are i.i.d., and

$$\mathbb{E}\left(e^{t\Delta_n} \mid G_n\right) = \mathbb{E}\left(\exp\left\{\frac{t\sum_{j=1}^{n} \lambda_j(G_n)A_j}{2\sqrt{\sum_{j=1}^{n} \lambda_j^2(G_n)}}\right\} \mid G_n\right)$$

$$(7.7) \qquad = \prod_{j=1}^{n} \left(1 - \tilde{\lambda}_j(G_n)\frac{t}{c}\right)^{\frac{1-c}{2}},$$

where $\tilde{\lambda}_j(G_n) = \frac{\lambda_j(G_n)}{\sqrt{\sum_{j=1}^n \lambda_j^2(G_n)}}$. Therefore, using $\sum_{j=1}^n \tilde{\lambda}_j(G_n) = 0$,

$$
\log \prod_{j=1}^n \left(1 - \tilde{\lambda}_j(G_n)\frac{t}{c} \right)^{\frac{1-c}{2}}
$$

$$
= \left(\frac{c-1}{2} \right) \sum_{x=2}^{\infty} \frac{t^x}{xc^x} \cdot \frac{t(C_g, G_n)}{t(C_2, G_n)^{\frac{x}{2}}}
$$

(7.8)

$$
= \left(\frac{c-1}{2} \right) \sum_{x=2}^{N} \frac{t^x}{xc^x} \cdot \frac{t(C_g, G_n)}{t(C_2, G_n)^{\frac{x}{2}}}
$$

$$
+ \left(\frac{c-1}{2} \right) \sum_{x=N+1}^{\infty} \frac{t^x}{xc^x} \cdot \frac{t(C_g, G_n)}{t(C_2, G_n)^{\frac{x}{2}}}.
$$

Now, since $G_n \overset{D}{\Rightarrow} W$, fixing a positive integer $N \geq 3$, $\{t(C_g, G_n), g \in [3, N]\} \overset{D}{\to}$ $\{t(C_g, W), g \in [3, N]\}$. This and (7.6) implies that

(7.9) $\lim_{N \to \infty} \lim_{n \to \infty} \left(\frac{c-1}{2} \right) \sum_{x=2}^{N} \frac{t^x}{xc^x} \cdot \frac{t(C_g, G_n)}{t(C_2, G_n)^{\frac{x}{2}}} \overset{D}{\to} \log \mathbb{E}\big(e^{t\frac{\chi_{c-1}^2(W)}{2c}} \mid W\big).$

Moreover, since $t(C_g, G_n) \leq t(C_2, G_n)^{\frac{x}{2}}$ and $|t| < c/4$, the second term in the RHS of (7.8) is bounded by $\sum_{x=N+1}^{\infty} \frac{1}{x4^x}$, which converges to 0 as $N \to \infty$. Combining (7.7), (7.8), and (7.9), it follows that

(7.10) $\mathbb{E}\big(e^{t\Delta_n} \mid G_n\big) = \prod_{j=1}^n \left(1 - \tilde{\lambda}_j(G_n)\frac{t}{c} \right)^{\frac{1-c}{2}} \overset{D}{\to} \mathbb{E}\big(e^{t\frac{\chi_{c-1}^2(W)}{2c}} \mid W\big).$

Finally, using the fact that $-\log(1-x) \leq x + x^2$ for $|x| \leq 1/2$ we have

$$
\log \mathbb{E}\big(e^{2t\Delta_n} \mid G_n\big) = \frac{1-c}{2} \sum_{j=1}^n \log\left(1 - \tilde{\lambda}_j(G_n)\frac{2t}{c} \right)
$$

$$
\leq \frac{2t^2(c-1)}{c^2} < \infty.
$$

This implies $e^{t\Delta_n}$ and $\mathbb{E}(e^{t\Delta_n} \mid G_n)$ are both uniformly integrable. Therefore, from (7.10)

(7.11) $\lim_{n \to \infty} \mathbb{E}\big(e^{t\Delta_n}\big) = \mathbb{E}\big(e^{t\frac{\chi_{c-1}^2(W)}{2c}}\big).$

By Proposition 7.1, the RHS of (7.11) is the moment generating function of $\frac{\chi_{c-1}^2(W)}{2c}$, and the proof is complete. $\square$

Finally, to complete the proof of Theorem 1.4 using Lemma C.1, observe: since $\mathbb{E}e^{t\Delta_n}$ converges to $\mathbb{E}e^{t\frac{\chi^2_{c-1}(W)}{2c}}$, for some $t > 0$, the arguments identical to those used in (5.12) imply condition (C.2) for Δ_n.

7.3. *More examples.* The limiting chi-square distribution of the complete graph was illustrated before in Example 7.1. A few other simple examples where Theorem 1.4 can be used to determine the limiting distributions are given below.

EXAMPLE 7.2 (Complete bipartite graph). Consider the complete bipartite graph $K_{\frac{n}{2}, \frac{n}{2}}$, which converges to the limit $W(x, y) = \mathbf{1}\{(x - 1/2)(y - 1/2) < 0\}$. It is easy to see that the only nonzero eigenvalues of W are $\pm \frac{1}{2}$. Therefore, by Theorem 1.4 it follows that

$$\frac{1}{n}\left(N(K_{\frac{n}{2}, \frac{n}{2}}) - \frac{n^2}{4c}\right) \xrightarrow{D} \frac{1}{4c}(\xi_1 - \xi_2),$$

where ξ_1 and ξ_2 are independent $\chi^2_{(c-1)}$ random variables.

EXAMPLE 7.3 (Inhomogeneous random graphs). Let $f : [0, 1]^2 \to [0, 1]$ be a symmetric continuous function. Consider the random graph model where an edge (i, j) is present with probability $f(\frac{i}{n}, \frac{j}{n})$. Therefore, whenever $\int_{[0,1]^2} f(x, y)\,dx\,dy > 0$ the limit theorem in (1.12) holds. In particular, the Erdős–Rényi random graph $G(n, p)$ can be obtained by taking the function $f(x, y) = p$. In this case, p is the only nonzero eigenvalue of the operator f and

$$\frac{1}{n}\left(N(G(n, p)) - \frac{|E(G_n)|}{c}\right) \xrightarrow{D} \frac{p^{\frac{1}{2}}}{2c} \cdot \left(\chi^2_{(c-1)} - (c - 1)\right).$$

Note that this reduces to Example 7.1, for $c = 2$ and $p = 1$.

8. Extremal examples: Stars and cycles. Another relevant question is whether it is possible to expect a similar Poisson universality result for other subgraphs, under the uniform coloring scheme? This section begins by proving Proposition 1.6 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.

8.1. *Monochromatic stars.* Consider the n-star, $K_{1,n}$ with vertices labeled by $[n]$, with the central vertex labeled 0. Color the vertices of $K_{1,n}$, uniformly from $[c]$, independently. Consider the limiting distribution of the number of monochromatic r-stars $K_{1,r}$ generated by this random coloring, where r is a fixed constant. If Y_i denotes the color of vertex i, the random variable is

$$T_{r,n} = \sum_{\substack{S \subseteq [n] \\ |S|=r-1}} \prod_{j \in S} \{Y_0 = Y_j\}.$$

Proposition 1.6 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.

8.1.1. *Proof of Proposition* 1.6.   Note that if the number of monochromatic edges in $G_n = K_{1,n}$ is $N(G_n)$, then

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

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

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

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

This random variable does not assign positive mass at all nonnegative integers, and so it cannot be a mixture of Poisson variates.

8.2. *Monochromatic cycles*.   Recall that the number of monochromatic edges $N(G_n)$ converges to Poisson(λ) whenever $|E(G_n)|/c_n \to \lambda$. The limiting distribution of the number of edges can only be a nontrivial mixture of Poissons when $|E(G_n)|/c_n \to Z$, and Z has a nondegenerate 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 a nontrivial mixture of Poissons even when $|N(G_n, C_g)|/c_n^{g-1}$ 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 graph $G_{5,2,5}$ and its 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_n-coloring of the vertices of $G_{a,b,g}$ and count the number of monochromatic g-cycles. Let Y_i be the color of vertex v_i and Y_{ijk} the color

FIG. 1.   *The graph $G_{5,2,5}$.*

of vertex v_{ijk}, for $i \in [a]$ and $j \in [b]$. The random variable

$$(8.1) \qquad Z_{a,b,g} := Z(G_{a,b,g}) := \sum_{i=1}^{a} \sum_{j=1}^{b} \prod_{k=1}^{g-2} \mathbf{1}\{Y_i = Y_{i+1} = Y_{ijk}\},$$

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_n such that $|N(G_n, C_g)|/c_n^{g-1} \to \lambda$ and $Z_{a,b,g}$ converges in distribution to a nontrivial mixture of Poissons.

PROPOSITION 8.1.   *For $a = \lambda n$ and $b = n^{g-2}$ and $c_n = n$, $Z_{a,b,g} \xrightarrow{D}$ Poisson(W), where $W \sim$ Poisson(λ).*

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

$$\prod_{k=1}^{g-2} \mathbf{1}\{Y_i = Y_{i+1} = Y_{ijk}\} \mid \underline{Y} \sim \text{Ber}(1/c_n^{g-2})$$

and

$$\sum_{j=1}^{b} \prod_{k=1}^{g-2} \mathbf{1}\{Y_i = Y_{i+1} = Y_{ijk}\} \mid \underline{Y} \sim \text{Bin}(b, 1/c_n^{g-2}).$$

This implies that

$$\mathbb{E}(e^{it Z_{a,b,g}}) = \mathbb{E}\left(\prod_{i=1}^{a} \mathbb{E}(e^{it \sum_{j=1}^{b} \prod_{k=1}^{g-2} \mathbf{1}\{Y_i = Y_{i+1} = Y_{ijk}\}} \mid \underline{Y})\right)$$

$$(8.2) \qquad = \mathbb{E}\left(1 - \frac{1}{c_n^{g-2}} + \frac{e^{it}}{c_n^{g-2}}\right)^{bN_a},$$

where $N_a = \sum_{i=1}^{a} \mathbf{1}\{Y_i = Y_{i+1}\}$, is the 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 $c_n = n$, 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 $W := \text{Poisson}(\lambda)$, as $a_n/c_n = \lambda$. Therefore,

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

By the dominated convergence theorem, we have

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

which the characteristic function of $\text{Poisson}(W)$, where $W \sim \text{Poisson}(\lambda)$. $\square$

REMARK 8.1.   We are not aware of any example for which the number of monochromatic triangles converges to a distribution that is not a mixture of Poissons or a polynomial in Poissons, when $N(G_n, C_3)/c_n^2 \to \lambda$. In fact, we believe that the number of monochromatic triangles in a uniform random coloring of any graph sequence G_n, converges in distribution to a random variable that is a mixture of Poissons or a polynomial in Poissons, whenever $|N(G_n, C_3)|/c_n^2 \to \lambda$, for some fixed $\lambda > 0$.

APPENDIX A: PROOF OF LEMMA 4.2

Recall, from the statement the lemma, that $V(H)$ is partitioned into three sets $V_0(H) \cup V_{1/2}(H) \cup V_1(H)$, where $V_a(H) = \{v \in V(H) : \varphi(v) = a\}$, for $a \in \{0, 1/2, 1\}$ for a fixed optimal solution φ of (4.6), and $H_{01} = (V_0(H) \cup V_1(H), E(H_{01}))$, where $E(H_{01})$ is set of edges from $V_0(H)$ to $V_1(H)$.

By Hall's marriage theorem, the bipartite graph H_{01} has a matching that saturates every vertex in $V_0(H)$, if and only if for all $S \subset V_0(H)$, $|N_{H_{01}}(S)| \geq |S|$. Suppose this is false and there exists $A \subset V_0(H)$ such that $|N_{H_{01}}(A)| < |A|$. Let $B = V_1(H) \setminus N_{H_{01}}(A)$. By assumption, $\gamma(H) = \frac{1}{2}|V_{1/2}(H)| + |V_1(H)| > \frac{1}{2}(|V_0(H)| + |V_{1/2}(H)| + |V_1(H)|)$, which implies that $|V_1(H)| > |V_0(H)|$, and so

$$|B| = |V_1(H)| - |N_{H_{01}}(A)| > |V_1(H)| - |A| \geq |V_1(H)| - |V_0(H)| > 0.$$

In particular, this implies that the set B is not empty.

Now, define a function $\tilde{\varphi} : V(H) \to [0, 1]$ such that

$$\tilde{\varphi}(v) := \begin{cases} 1 & \text{if } v \in B, \\ 0 & \text{if } v \in N_H(B), \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

Note that $\tilde{\varphi} \in P(H)$, and $\sum_{v \in V} \tilde{\varphi}(v) = \frac{1}{2}(|V(H)| + |B| - |N_H(B)|)$. Now, as $N_H(B) \subseteq V_0(H) \setminus A$,

$$\sum_{v \in V} \tilde{\varphi}(v) = \frac{1}{2}(|V(H)| + |B| - |N_H(B)|)$$

$$= \frac{1}{2}(|V(H)| + |V_1(H)| - |N_{H_{01}}(A)| - |N_H(B)|)$$

(A.1)
$$\geq \frac{1}{2}(|V(H)| + |V_1(H)| - |V_0(H)| - |N_{H_{01}}(A)| + |A|)$$

$$> \frac{1}{2}(|V(H)| + |V_1(H)| - |V_0(H)|)$$

$$= \frac{1}{2}|V_{1/2}(H)| + |V_1(H)| = \gamma(H).$$

This contradicts the maximality of $\gamma(H)$ and proves that H_{01} has a matching that saturates every vertex in $V_0(H)$.

Next, denote the multi-graph induced by the vertices of $V_{1/2}(H)$ by F. To prove that F has a spanning subgraph which is a disjoint union of cycles and isolated edges, it suffices to show that $\delta(F_S) = 0$ (Lemma 7, Alon [2]). To this end we need to show that for all $S \subset V(F)$, $|N_F(S)| \geq |S|$. Assuming this is false, there exists $C \subset V(F)$, such that $|N_F(C)| < |C|$. Define $C_1 = C \setminus N_F(C)$. Note that $C_1 \cap N_F(C_1) = \varnothing$: if there exists $x \in C_1 \cap N_F(C_1)$ then there exists $y \in C_1$ such that $(x, y) \in E(F)$, which means $x \in C_1 \cap N_F(C)$, a contradiction. Moreover, $|N_F(C)| \geq |N_F(C) \cap C| + |N_F(C_1)|$, and

$$|C_1| = |C| - |C \cap N_F(C)|$$

$$> |N_F(C)| - |C \cap N_F(C)|$$

$$\geq |N_F(C_1)|.$$

Let $D = C_1 \cup V_1(H)$. Note that $N_H(D) = N_H(C_1) \cup N_H(V_1(H)) = N_F(C_1) \cup N_H(V_1(H))$ and $D \cap N_H(D) = \varnothing$. Also,

$$|D| - |N_H(D)| \geq |V_1(H)| - |N_H(V_1(H))| + (|C_1| - |N_F(C_1)|)$$

$$> |V_1(H)| - |V_0(H)|.$$

Define $\tilde{\varphi} : V(H) \to [0, 1]$

(A.2)
$$\tilde{\varphi}(v) := \begin{cases} 1 & \text{if } v \in D, \\ 0 & \text{if } v \in N_H(D), \\ \dfrac{1}{2} & \text{otherwise.} \end{cases}$$

Note that $\tilde{\varphi} \in P(H)$, and $\sum_{v \in V} \tilde{\varphi}(v) = \frac{1}{2}(|V(H)| + |D| - |N_H(D)|)$. Now, a contradiction can be obtained as before.

APPENDIX B: PROOF OF LEMMA 6.1

Let H be a multigraph with no isolated vertex, $d_{\min}(H) \geq 2$, $\gamma(H) = |E(H)|/2$, and at least 3 vertices in one of its connected components. Recall, we need to show that $M(G_n, H) = o_P(|E(G_n)|^{|E(H)|/2})$, whenever the ACF4 condition in probability holds for G_n and H contains a cycle C_g, for some $g \geq 3$.

To begin with assume that H is connected. If $|E(H)| = |V(H)| - 1$, the graph H is a tree and $d_{\min}(H) = 1$. Also, if $|E(H)| = |V(H)|$, then since $|V(H)| \geq 3$ and $d_{\min}(H) \geq 2$, the only possibility is that $H = C_g$ for some $g \geq 3$, and $M(G_n, H) = o_P(|E(G_n)|^{|E(H)|/2})$ by the ACF4 condition in probability and Theorem 6.1.

Therefore, it suffices to consider a subgraph H of G_n such that $|E(H)| > |V(H)|$. If $\gamma(H) \leq |V(H)|/2$,

$$
\begin{aligned}
M(G_n, H) &\lesssim_H |E(G_n)|^{\gamma(H)} \\
&\lesssim_H |E(G_n)|^{|V(H)|/2} \\
&= o_P(|E(G_n)|^{|E(H)|/2}).
\end{aligned}
$$

Therefore, assume that $\gamma(H) > |V(H)|/2$. As in the proof of Lemma 4.2, let $\varphi : V(H) \to [0, 1]$ be an extreme point of $P(H)$ that is an optimal solution of the linear program defined in (4.6). Partition $V(H) = V_0(H) \cup V_{1/2}(H) \cup V_1(H)$, where $V_a(H) = \{v \in V(H) : \varphi(v) = a\}$, for $a \in \{0, 1/2, 1\}$. Note that $\gamma(H) > |V(H)|/2$ implies that φ is not identically equal to 1/2. Depending on the size of $V_{1/2}(H)$ the following cases arise:

$|V_{1/2}(H)| \neq 0$. Let H_{01} be the graph with vertex set $V_0(H) \cup V_1(H)$ and edge set $E(H_{01})$, where $E(H_{01})$ is the set of edges from $V_0(H)$ to $V_1(H)$. Let $H_{1/2}$ be the subgraph of H induced by the vertices of $V_{1/2}(H)$. Decompose H into subgraphs H_{01} and $H_{1/2}$. By Lemma 4.2, H_{01}, has a matching which saturates every vertex in $V_0(H)$. Therefore,

$$(B.1) \qquad M(G_n, H_{01}) \lesssim_H |E(G_n)|^{|V_1(H)|} \lesssim_H |E(G_n)|^{|E(H_{01})|/2},$$

since $d_{\min}(H) \geq 2$ implies $|E(H_{01})| \geq 2|V_1(H)|$. Moreover, the subgraph F of H induced by the vertices of $V_{1/2}(H)$ has a spanning subgraph which is a disjoint union of cycles and isolated edges. If $F_1, F_2, \ldots, F_\nu$ are the connected components of F, then by Theorem 4 of Alon [2], $\gamma(F_i) = |V(F_i)|/2$ (note that this is true even if F_i is a multigraph), and

$$(B.2) \qquad\qquad M(G_n, F_i) \lesssim_H |E(G_n)|^{|V(F_i)|/2}.$$

Denote by $E(V(F_i), V_0(H))$ the subset of edges in H with one vertex in $V(F_i)$ and another in $V_0(H)$, for $i \in [\nu]$. Using estimates (B.1) and (B.2),

$$
\begin{aligned}
\frac{M(G_n, H)}{|E(G_n)|^{|E(H)|/2}} &\lesssim_H \frac{M(G_n, H_{01}) \prod_{i=1}^{\nu} M(G_n, F_i)}{|E(G_n)|^{|E(H)|/2}} \\
&\lesssim_H \prod_{i=1}^{\nu} \frac{|E(G_n)|^{|V(F_i)|/2 - |E(F_i)|/2}}{|E(G_n)|^{|E(V(F_i), V_0(H))|/2}} \\
&= \prod_{i=1}^{\nu} |E(G_n)|^{\lambda(F_i)/2},
\end{aligned}
$$

(B.3)

where

(B.4) $$\lambda(F_i) := |V(F_i)| - |E(F_i)| - |E(V(F_i), V_0(H))|.$$

Note that $|E(V(F_i), V_0(H))| > 0$, since H is connected. Therefore, $|E(G_n)|^{\lambda(F_i)/2} = o_P(1)$ whenever $|V(F_i)| \le |E(F_i)|$. Otherwise F_i is a tree and has at least 2 vertices of degree 1. The degree 1 vertices must be connected to some vertex in H_0, which implies that $|E(V(F_i), V_0(H))| \ge 2$ and again $|E(G_n)|^{\lambda(F_i)/2} = o_P(1)$.

$|V_{1/2}(H)| = 0$. In this case, by Lemma 4.2, $V(H) = V_0(H) \cup V_1(H)$, that is, every vertex is assigned the value 0 or 1 by the chosen optimal function φ. By Lemma 4.1, for every vertex $v \in V_1(H)$, $d(v) = 2$. Therefore,

(B.5) $$|E(H)| = 2\gamma(H) = 2|V_1(H)|,$$

and the graph H is bi-partite. By Lemma 4.2, H then has a matching which saturates every vertex in $V_0(H)$. By assumption, the girth $g := \mathrm{gir}(H) \ge 3$, and so there exists a subgraph F of H which is isomorphic to C_g. Let $H^- = (V(H^-), E(H^-))$ be the graph obtained by removing all the vertices of the cycle F and the all the edges of H incident on these vertices. Note that if $V_0(H^-) := V_0(H) \setminus V(F)$ is empty, then H^- is a graph with isolated vertices and no edges.[4] Otherwise $V_0(H^-)$ is nonempty, and we will show that, H^- also has a matching which saturates vertices of $V_0(H^-)$. To this end, let $A \subset V_0(H^-)$. By the saturating matching in H, $|N_H(A)| \ge |A|$. Also, observe that $|N_{H^-}(A)| = |N_H(A)|$, since removing F from H leaves the vertices in A and its neighbors unchanged. Therefore, $|N_{H^-}(A)| \ge |A|$ for all $A \subset V_0(H^-)$, that is, even after removing the cycle F

---

[4]For example, when $H = K_{2,t}$ is the complete bi-partite graph with two vertices on one side and t on the other, then H^- consists of $t - 2$ isolated vertices. In this case, $\gamma(H) = |E(H)|/2 = t$ and

$$M(G_n, H) \lesssim_t M(G_n, C_4)|E(G_n)|^{t-2} = o_P(|E(G_n)|^t),$$

by the ACF4 condition, as required.

from H, there is a matching in H^- which saturates every vertex in $V_0(H^-)$. This implies, by the ACF4 condition,

$$\frac{M(G_n, H)}{|E(G_n)|^{|E(H)|/2}} \lesssim_H \frac{M(G_n, F)|E(G_n)|^{|V(H^-) \cap V_1(H)|}}{|E(G_n)|^{|E(H)|/2}}$$

$$\lesssim_H \frac{M(G_n, F)}{|E(G_n)|^{g/2}} = o_P(1),$$

because $|V(H^-) \cap V_1(H)| = |V_1(H)| - g/2 = |E(H)|/2 - g/2$.

Finally, if H is not connected, let $H_1, H_2, \ldots, H_r$ be the connected components of H. There exists $j \in [r]$ such that $d_{\min}(H_j) \geq 2$ and $\mathrm{gir}(H_j) \geq 3$. By applying the above argument for H_j, it follows that $M(G_n, H_j) = o_P(|E(G_n)|^{|E(H_j)|/2})$. Also, for all $i \in [r]$, $M(G_n, H_i) \lesssim_H |E(G_n)|^{\gamma(H_i)} \leq |E(G_n)|^{|E(H_i)|/2}$, by Lemma 4.1. This implies

$$M(G_n, H) \lesssim_H \prod_{i=1}^r M(G_n, H_i)$$

$$= o_P(|E(G_n)|^{|E(H_j)|/2}) O_P(|E(G_n)|^{|E(H_i)|/2})$$

$$= o_P(|E(G_n)|^{|E(H)|/2}),$$

and the result follows.

APPENDIX C: FROM MOMENTS TO DISTRIBUTIONS

In this section, we show how to conclude about convergence in distribution from the convergence of moments, and the convergence of conditional distributions. The following lemma shows the convergence in distribution of a sequence of random variables from the convergence of the conditional moments. The lemma is used in repeatedly in the paper in the final steps of our proofs of all the main theorems.

LEMMA C.1. *Let $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ be a sequence of probability spaces, and $\mathcal{A}_n \subset \mathcal{F}_n$ 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 $U_{n,k} := \mathbb{E}(X_n^k \mid \mathcal{A}_n)$, $V_{n,k} := \mathbb{E}(Y_n^k \mid \mathcal{A}_n)$ exist as finite random variables, such that*

(C.1) $$\limsup_{n \to \infty} \mathbb{P}(|U_{n,k} - V_{n,k}| > \varepsilon) = 0.$$

Moreover, if there exists $\eta > 0$ such that

(C.2) $$\limsup_{k \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\left|\frac{\eta^k}{k!} U_{n,k}\right| > \varepsilon\right) = 0,$$

then for any $t \in \mathbb{R}$, $\mathbb{E}e^{itX_n} - \mathbb{E}e^{itY_n} \to 0$.

PROOF.   First, note that

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

Taking limits as $n \to \infty$, and using (C.1) and (C.2), it follows that

$$(C.3) \qquad\qquad \limsup_{k\to\infty}\limsup_{n\to\infty} \mathbb{P}\left(\left|\frac{\eta^k}{k!}V_{n,k}\right| > \varepsilon\right) = 0.$$

To prove the lemma it suffices to show that for any $t \in \mathbb{R}$ and an nonnegative integer ℓ,

$$(C.4) \qquad\qquad \left|\mathbb{E}(e^{itX_n}X_n^\ell \mid \mathcal{A}_n) - \mathbb{E}(e^{itY_n}Y_n^\ell \mid \mathcal{A}_n)\right| \xrightarrow{P} 0.$$

Indeed the lemma follows immediately from (C.4), as follows: Setting $\ell = 0$, $|\mathbb{E}(e^{itX_n} \mid \mathcal{A}_n) - \mathbb{E}(e^{itY_n} \mid \mathcal{A}_n)|$ converges to 0 in probability. Since $|\mathbb{E}(e^{itX_n} \mid \mathcal{A}_n) - \mathbb{E}(e^{itY_n} \mid \mathcal{A}_n)|$ is also bounded by 2 in absolute value, the dominated convergence theorem gives the desired result.

Therefore, it remains to prove the claim (C.4). For this, first assume $|t| \leq \eta$, and let ℓ be fixed but arbitrary nonnegative integer. By a Taylor's series expansion, for any $k \in \mathbb{N}$,

$$\left|e^{it} - \sum_{r=0}^{k-1} \frac{(it)^r}{r!}\right| \leq \frac{\eta^k}{k!},$$

and so

$$\left|\mathbb{E}(e^{itX_n}X_n^\ell \mid \mathcal{A}_n) - \mathbb{E}(e^{itY_n}Y_n^\ell \mid \mathcal{A}_n)\right|$$

$$(C.5) \qquad \leq \sum_{r=0}^{k-1} \frac{|t|^r}{r!}|U_{n,r+\ell} - V_{n,r+\ell}| + \frac{|\eta|^k}{k!}U_{n,k} + \frac{|\eta|^k}{k!}V_{n,k}.$$

From (C.5) taking limits as $n \to \infty$ followed by $k \to \infty$, and using (C.1), (C.2) and (C.3) gives $|\mathbb{E}(e^{itX_n}X_n^\ell \mid \mathcal{A}_n) - \mathbb{E}(e^{itY_n}Y_n^\ell \mid \mathcal{A}_n)| \xrightarrow{P} 0$, for $|t| \leq \eta$.

The proof of (C.4) is now completed by induction. Suppose the result holds for any $|t| \leq p\eta$, for some $p \in \mathbb{N}$, and let $|t| \in (p\eta, (p+1)\eta]$. Then setting $t_0 := t - \frac{t}{|t|}\eta$ we have that $|t - t_0| = \eta$, and $|t_0| \leq p\eta$. Expanding in a Taylor's series around t_0

$$\left|\mathbb{E}(e^{itX_n}X_n^\ell \mid \mathcal{A}_n) - \mathbb{E}(e^{itY_n}Y_n^\ell \mid \mathcal{A}_n)\right|$$

$$\leq \sum_{r=0}^{k-1} \frac{\eta^r}{r!}\left|\mathbb{E}(e^{it_0X_n}X_n^{r+\ell} \mid \mathcal{A}_n) - \mathbb{E}(e^{it_0Y_n}Y_n^{r+\ell} \mid \mathcal{A}_n)\right|$$

$$+ \frac{\eta^k}{k!}U_{n,k} + \frac{\eta^k}{k!}V_{n,k}.$$

Since $|t_0| \leq p\eta$, letting $n \to \infty$ followed by $k \to \infty$ it follows by the induction hypothesis and (C.1) and (C.2) that $|\mathbb{E}(e^{itX_n}X_n^\ell \mid \mathcal{A}_n) - \mathbb{E}(e^{itY_n}Y_n^\ell \mid \mathcal{A}_n)| \xrightarrow{P} 0$. This completes the proof of (C.4) by induction. $\square$

REMARK C.1. There are many conditions on modes of 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 [45, 47]. As pointed out by one of the referee, another alternative strategy to prove the convergence of the number of monochromatic edges $N(G_n)$, for a sequence of random graphs G_n, is to first prove them for deterministic graphs using method of moments as done in the paper, and then lift the result to random graphs. Below we explain how this program can be carried out for Poisson approximation (Theorem 1.1):

For a *fixed* sequence of graphs $G_n \in \mathcal{G}_n$, let $N(G_n)$ be a random variable which counts the number of monochromatic edges in G_n divided by c_n. Using the Poisson approximation for deterministic graphs, if $\lim_{n\to\infty} \frac{|E(G_n)|}{c_n} = \lambda$ then

$$N_n(G_n) \xrightarrow{D} \text{Poisson}(\lambda) \quad \Longrightarrow \quad \lim_{n\to\infty} \mathbb{E}e^{itN_n(G_n)} = \exp\{\lambda(e^{it}-1)\}.$$

To show this we would need to invoke Lemma C.1 above for deterministic graphs, to argue about the convergence in distribution from the comparison of moments. Now, using the universality of this approximation one can show the stronger statement

$$\sup_{G_n\in\mathcal{G}_n:E(G_n)\leq c_n\lambda}\left|\mathbb{E}e^{itN_n(G_n)} - \exp\left\{\frac{|E(G_n)|}{c_n}(e^{it}-1)\right\}\right| = 0$$

for any $\lambda < \infty$. Thus if G_n is a random graph independent of the coloring such that $\frac{|E(G_n)|}{c_n}$ converges to Z, then for λ large,

$$\mathbb{E}e^{itN(G_n)} \approx \mathbb{E}\big(e^{itN(G_n)}\mathbf{1}\{|E(G_n)| \leq \lambda c_n\}\big)$$

$$\approx \mathbb{E}\left(\exp\left\{\frac{|E(G_n)|}{c_n}(e^{it}-1)\right\}\mathbf{1}\{|E(G_n)| \leq \lambda c_n\}\right)$$

$$\approx \mathbb{E}(\exp\{Z(e^{it}-1)\}\mathbf{1}\{Z \leq \lambda\})$$

$$\approx \mathbb{E}(\exp\{Z(e^{it}-1)\}),$$

thus proving Theorem 1.1. A similar heuristic argument can be made for the other theorems.

REFERENCES

[1] ALDOUS, D. (1989). *Probability Approximations via the Poisson Clumping Heuristic*. *Applied Mathematical Sciences* **77**. Springer, New York. MR0969362

[2] ALON, N. (1981). On the number of subgraphs of prescribed type of graphs with a given number of edges. *Israel J. Math*. **38** 116–130. MR0599482

[3] ALON, N. (1986). On the number of certain subgraphs contained in graphs with a given number of edges. *Israel J. Math*. **53** 97–120. MR0861901

[4] ARRATIA, R., GOLDSTEIN, L. and GORDON, L. (1990). Poisson approximation and the Chen–Stein method. *Statist. Sci*. **5** 403–434. MR1092983

[5] BARBOUR, A. D., HOLST, L. and JANSON, S. (1992). *Poisson Approximation*. *Oxford Studies in Probability* **2**. Oxford Univ. Press, New York. MR1163825

[6] BERAN, R. (1975). Tail probabilities of noncentral quadratic forms. *Ann. Statist*. **3** 969–974. MR0381122

[7] BOLLOBÁS, B., JANSON, S. and RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31** 3–122. MR2337396

[8] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2008). Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math*. **219** 1801–1851. MR2455626

[9] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2012). Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math.* (2) **176** 151–219. MR2925382

[10] BURDA, M., HARDING, M. and HAUSMAN, J. (2012). A Poisson mixture model of discrete choice. *J. Econometrics* **166** 184–203. MR2862959

[11] CERQUETTI, A. and FORTINI, S. (2006). A Poisson approximation for coloured graphs under exchangeability. *Sankhyā* **68** 183–197. MR2303080

[12] CHATTERJEE, S. (2008). A new method of normal approximation. *Ann. Probab*. **36** 1584–1610. MR2435859

[13] CHATTERJEE, S. and DIACONIS, P. (2013). Estimating and understanding exponential random graph models. *Ann. Statist*. **41** 2428–2461. MR3127871

[14] CHATTERJEE, S., DIACONIS, P. and MECKES, E. (2005). Exchangeable pairs and Poisson approximation. *Probab. Surv*. **2** 64–106. MR2121796

[15] CHUNG, F. R. K., GRAHAM, R. L. and WILSON, R. M. (1989). Quasi-random graphs. *Combinatorica* **9** 345–362. MR1054011

[16] CHURCH, K. and GALE, W. A. (1995). Poisson mixtures. *Nat. Lang. Eng*. **1** 163–190.

[17] CONLON, D., FOX, J. and ZHAO, Y. (2014). Extremal results in sparse pseudorandom graphs. *Adv. Math*. **256** 206–290. MR3177293

[18] DASGUPTA, A. (2005). The matching, birthday and the strong birthday problem: A contemporary review. *J. Statist. Plann. Inference* **130** 377–389. MR2128015

[19] DE JONG, P. (1987). A central limit theorem for generalized quadratic forms. *Probab. Theory Related Fields* **75** 261–277. MR0885466

[20] DIACONIS, P. and HOLMES, S. (2002). A Bayesian peek into Feller volume I. *Sankhyā Ser. A* **64** 820–841. Special issue in memory of D. Basu. MR1981513

[21] DIACONIS, P. and MOSTELLER, F. (1989). Methods for studying coincidences. *J. Amer. Statist. Assoc*. **84** 853–861. MR1134485

[22] DONG, F. M., KOH, K. M. and TEO, K. L. (2005). *Chromatic Polynomials and Chromaticity of Graphs*. World Scientific Co. Pte. Ltd., Hackensack, NJ. MR2159409

[23] DURRETT, R. (2010). *Probability*: *Theory and Examples*, 4th ed. Cambridge Univ. Press, Cambridge. MR2722836

[24] ERDŐS, P. (1962). On the number of complete subgraphs contained in certain graphs. *Magyar Tud*. *Akad*. *Mat*. *Kutató Int*. *Közl*. **7** 459–464. MR0151956

[25] FADNAVIS, S. (2011). A generalization of the birthday problem and the chromatic polynomial. Preprint. Available at arXiv:1105.0698v2.

[26] FANG, X. (2015). A universal error bound in the CLT for counting monochromatic edges in uniformly colored graphs. *Electron*. *Commun*. *Probab*. **20** 21. MR3320409

[27] FRIEDGUT, E. and KAHN, J. (1998). On the number of copies of one hypergraph in another. *Israel J*. *Math*. **105** 251–256. MR1639767

[28] GELMAN, A. (2013). 600: The average American knows how many people? *New York Times* February 19, page D7.

[29] GÖTZE, F. and TIKHOMIROV, A. (2002). Asymptotic distribution of quadratic forms and applications. *J*. *Theoret*. *Probab*. **15** 423–475. MR1898815

[30] GÖTZE, F. and TIKHOMIROV, A. N. (1999). Asymptotic distribution of quadratic forms. *Ann*. *Probab*. **27** 1072–1098. MR1699003

[31] GREENWOOD, M. and YULE, G. V. (1920). 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*. *Ser*. *A* **83** 255–279.

[32] HALL, P. (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *J*. *Multivariate Anal*. **14** 1–16. MR0734096

[33] JANSON, S., OLESZKIEWICZ, K. and RUCIŃSKI, A. (2004). Upper tails for subgraph counts in random graphs. *Israel J*. *Math*. **142** 61–92. MR2085711

[34] JENSEN, T. R. and TOFT, B. (1995). *Graph Coloring Problems*. Wiley, New York. MR1304254

[35] JENSEN, T. R. and TOFT, B. (2013). Unsolved graph coloring problems. Manuscript in preparation.

[36] KILLWORTH, P. and BERNARD, H. (1979). The reversal small-world experiment. *Social Networks* **1** 159–192.

[37] LE CAM, L. and TRAXLER, R. (1978). On the asymptotic behavior of mixtures of Poisson distributions. *Z*. *Wahrsch*. *Verw*. *Gebiete* **44** 1–45. MR0503463

[38] LOVÁSZ, L. (2012). *Large Networks and Graph Limits*. *American Mathematical Society Colloquium Publications* **60**. Amer. Math. Soc., Providence, RI. MR3012035

[39] NEMHAUSER, G. L. and TROTTER, L. E. JR. (1974). Properties of vertex packing and independence system polyhedra. *Math*. *Program*. **6** 48–61. MR0379294

[40] NOURDIN, I., PECCATI, G., POLY, G. and SIMONE, R. (2016). Classical and free fourth moment theorems: Universality and thresholds. *J*. *Theoret*. *Probab*. **29** 653–680. MR3500415

[41] NOURDIN, I., PECCATI, G. and REINERT, G. (2010). Stein's method and stochastic analysis of Rademacher functionals. *Electron*. *J*. *Probab*. **15** 1703–1742. MR2735379

[42] NOURDIN, I., PECCATI, G. and REINERT, G. (2010). Invariance principles for homogeneous sums: Universality of Gaussian Wiener chaos. *Ann*. *Probab*. **38** 1947–1985. MR2722791

[43] ROTAR, V. I. (1973). Some limit theorems for polynomials of second degree. *Theory Probab*. *Appl*. **18** 499–507.

[44] SCHRIJVER, A. (2003). *Combinatorial Optimization*: *Polyhedra and Efficiency*. *Algorithms and Combinatorics* **24**. Springer, Berlin.

[45] SETHURAMAN, J. (1961). Some limit theorems for joint distributions. *Sankhyā Ser*. *A* **23** 379–386. MR0139190

[46] STANLEY, R. P. (1995). A symmetric function generalization of the chromatic polynomial of a graph. *Adv*. *Math*. **111** 166–194. MR1317387

[47] SWEETING, T. J. (1989). On conditional weak convergence. *J. Theoret. Probab.* **2** 461–474. MR1011199

B. B. BHATTACHARYA
DEPARTMENT OF STATISTICS
THE WHARTON SCHOOL
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA 19104-6302
USA
E-MAIL: bhaswar@wharton.upenn.edu

P. DIACONIS
DEPARTMENT OF MATHEMATICS
  AND STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: diaconis@math.stanford.edu

S. MUKHERJEE
DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027
USA
E-MAIL: sm3949@columbia.edu