# A local limit theorem for the edge counts of random induced subgraphs of a random graph

Paul Balister[*][†]  Emil Powierski[*]  Alex Scott[*][‡]  Jane Tan[*]

March 28, 2025

## Abstract

Consider a 'dense' Erdős–Rényi random graph model $G = G_{n,M}$ with $n$ vertices and $M$ edges, where we assume the edge density $M/\binom{n}{2}$ is bounded away from 0 and 1. Fix $k = k(n)$ with $k/n$ bounded away from 0 and 1, and let $S$ be a random subset of size $k$ of the vertices of $G$. We show that with probability $1 - \exp(-n^{\Omega(1)})$, $G$ satisfies both a central limit theorem and a local limit theorem for the empirical distribution of the edge count $e(G[S])$ of the subgraph of $G$ induced by $S$, where the distribution is over uniform random choices of the $k$-set $S$.

## 1  Introduction

Consider an Erdős-Rényi random graph model $G = G_{n,M}$ with $n$ vertices and $M$ edges. Set $N := \binom{n}{2}$ and assume $M/N \in [\delta, 1-\delta]$ for some fixed $\delta > 0$, so that $M/N$ is bounded away from both 0 and 1 and both $G$ and its complement are dense. Fix $k = k(n)$ and assume $k/n \in [\delta, 1-\delta]$ is also bounded away from 0 and 1. Letting $S$ be a uniformly chosen random subset of the vertices of $G$ of size $k$, we will study the edge count $e(S) := e(G[S])$ of the subgraph of $G$ induced by $S$. Throughout, we shall always assume that the number of vertices $n$ is sufficiently large, and $M$, $k$ and $S$ are defined as above unless otherwise indicated within the scope of specific statements. We say that an event $E$ happens in $G_{n,M}$ *with very high probability* (w.v.h.p.) if there exists $\varepsilon > 0$ such that, for sufficiently large $n$, we have $\mathbb{P}(E) \geq 1 - \exp(-n^{\varepsilon})$. In this paper, we show that with very high probability (in choice of $G$), the empirical statistics of $e(S)$ satisfy both a central limit theorem and a local limit theorem.

We will be considering probabilities in various probability spaces. Notably, this will involve probabilities with respect to the random graph models $G_{n,p}$ or $G_{n,M}$, as well as probabilities with respect to the random vertex set $S$ that we choose. Where helpful, we

---

[*]Mathematical Institute, University of Oxford, Oxford OX2 6GG, United Kingdom {balister,powierski,scott,jane.tan}@maths.ox.ac.uk

will use subscripts such as $\mathbb{P}_G$ for probabilities over the choice of the random graph, and $\mathbb{P}_S$ for probabilities over the choice of $S$ conditioned on a fixed choice of random graph $G$.

By using a multidimensional version of Stein's method as described in [11], we first deduce in Section 2 a central limit theorem for $e(S)$. Define

$$\lambda := \frac{(n^2 - k^2)k^2}{2n^4} \cdot \frac{M(N - M)}{N^2}. \tag{1}$$

and note that $\lambda = \Theta(1)$. Set $K := \binom{k}{2}$.

**Theorem 1.** *Assume that for some fixed $\delta > 0$, $M/N, k/n \in [\delta, 1 - \delta]$. Then for any $\varepsilon > 0$ and w.v.h.p. in $G = G_{n,M}$ we have for all $z \in \mathbb{R}$,*

$$\left|\mathbb{P}_S(e(S) \leq z) - \mathbb{P}(Z \leq z)\right| \leq n^{-1/4 + \varepsilon}, \tag{2}$$

*where $S \subseteq V(G)$ is a uniformly chosen random subset of size $k$ and $Z \sim N(KM/N, \lambda n^2)$ is a normal random variable with mean $KM/N$ and variance $\lambda n^2$ with $\lambda$ given by (1).*

This central limit theorem serves as a starting point for an iterative smoothing argument that we employ in Section 3 to descend to the following local limit theorem, which is our main result.

**Theorem 2.** *Assume that for some fixed $\delta > 0$, $M/N, k/n \in [\delta, 1 - \delta]$. Then for any $\varepsilon > 0$ and w.v.h.p. in $G = G_{n,M}$ we have that for all integers $z$*

$$n\left|\mathbb{P}_S(e(S) = z) - \varphi(z)\right| \leq n^{-1/14 + \varepsilon},$$

*where $\varphi$ is the density function of a normal random variable $Z \sim N(KM/N, \lambda n^2)$.*

In particular, this shows that the point probabilities of $e(S)$ are of order $1/n$ around its peak, i.e., within $O(n)$ of its mean $KM/N$, and the relative error is at most $n^{-1/14 + \varepsilon}$ in this range.

Somewhat surprisingly, Theorem 1 and Theorem 2 are the first such results on the empirical distribution of edge counts in this very natural setting of subgraphs induced by $k$-subsets, although there is a good deal of related work which we discuss in Section 1.1.

Results that hold w.v.h.p. can in general be easily transferred between the above setting in $G_{n,M}$ and $G_{n,p}$ for $p \in [\delta, 1 - \delta]$; given a statement holding in $G_{n,M}$ and setting $p = M/N$, there is a polynomial probability (on the order of $1/n$) that $e(G_{n,p}) = M$. In that case we have $(G_{n,p} \mid e(G_{n,p}) = M) \sim G_{n,M}$, so the conclusion of the smoothing lemma (Lemma 17 below) transfers to $G_{n,M}$ again with very small failure probability. Consequently, it will be convenient to prove most of our lemmas en route w.v.h.p. for $G_{n,p}$.

However, we remark that the edge count of $G_{n,p}$ can typically vary by order $n$ and so $KM/N$ also varies by order $n$, which is of the same order as the standard deviation of $Z$. Hence, we can't just replace $M$ by $p\binom{n}{2}$ and obtain the same results as above when working in $G_{n,p}$.

We also remark that, due to the exponentially small failure probability, a union bound implies slightly stronger versions of the above results which say that for fixed $\delta > 0$,

w.v.h.p. $G_{n,M}$ satisfies the desired properties for all $k, M$ with $k/n, M/N \in [\delta, 1 - \delta]$. For the most part, it will be more convenient for us to view $k$ and $M$ as fixed, but analogous union bound arguments will later come up for other parameters.

## 1.1 Discussion

Both Theorems 1 and 2 are concerned with induced subgraphs of *fixed* size. We can also consider a *random size* model where we take a subgraph induced by a random subset $S$, where each vertex independently belongs to $S$ with constant probability $r$. In this case, it is straightforward to show that w.v.h.p. $e(S)$ satisfies a central limit theorem. We write $e(S)$ in the form $\sum_{i<j} a_{ij} X_i X_j$, where $A = (a_{ij})$ is the adjacency matrix and the $X_i$ are independent Bernoulli random variables with mean $r$. Then a standard sufficient condition for asymptotic normality is that

$$\frac{\text{Tr}(A^4)}{\sigma^4} \to 0 \qquad \text{and} \qquad \frac{\max_i \sum_j a_{ij}^2}{\sigma^2} \to 0, \tag{3}$$

where $\sigma^2$ is the variance of $e(S)$ (see [5, 6, 9, 12] for more details, including error bounds). This is immediate in the random size model, as $\sigma^2$ is of order $n^3$ w.v.h.p. (driven by copies of $K_{1,2}$), while $\text{Tr}(A^4) = O(n^4)$ and $\max_i \sum_j a_{ij}^2 = \Delta(G) \leq n$. In fact, this holds for any sequence of graphs with density bounded away from 0 (as these contain quadratically many copies of $K_{1,2}$).[1] However, we cannot hope for a local limit theorem when the size of $S$ varies. Indeed, a change of size of $S$ of just 1 will move the mean in Theorem 2 by order $n$, which is the same order as the standard deviation for fixed $|S|$. Thus the distribution will approach a rapidly varying function which is the sum of many small Gaussians (see [10]).

There has been previous work on monochromatic edges in random $c$-colourings of random graphs: for a fixed $c$, consider a function $\phi \colon V(G) \to [c]$ chosen uniformly at random and let $Y = Y_G(\phi)$ be the number of monochromatic edges. For $G \in G_{n,p}$, the variance of $Y$ is typically of order $n^2$ (as copies of $K_{1,2}$ no longer contribute anything), and so (3) is no longer satisfied. In fact, the distribution of $Y$ is typically not Gaussian (a Gaussian distribution requires the graph to have very few copies of $C_4$: see [2] for a detailed analysis). Note that the colour classes in a random colouring do not have fixed size, so we are in a random size model. However, while we do not pursue this further here, our methods should be applicable with fixed sizes of colour classes. For example, with $c = 2$ and colour classes of equal size, we would be considering a random bisection.

A broader class of examples is given by Ramsey graphs. A graph with $n$ vertices is $C$-*Ramsey* if it has no clique or independent set of size $C \log_2 n$. It follows from a result of Erdős and Szemerédi [8] that, for large $n$, $C$-Ramsey graphs have densities bounded away from 0 and 1; thus by (3), in the 'random size' model, $e(S)$ satisfies a central limit theorem. However, the local picture is more subtle. The first question here is whether there are *any* induced subgraphs with the sizes we require. For random graphs, this was shown by Calkin, Frieze and McKay [4], who proved that for $p$ fixed, with high probability a graph in $G_{n,p}$ contains induced subgraphs of all sizes between 0 and $e(G) - cn^{3/2}/\sqrt{\log n}$.

---

[1] In our case, where the size of $S$ is fixed, (3) does not apply, as the $X_i$ are not independent (and, in any case, the first expression in (3) does not tend to 0, as the variance is much smaller, having order $n^2$).

For Ramsey graphs this is much harder: Erdős and McKay (see [7]) conjectured that there are induced subgraphs of all sizes up to $cn^2$. After a long line of subsequent work, a much stronger statement was proved by Kwan, Sah, Sauermann and Sawhney [10]: for every $\varepsilon, C > 0$, every $C$-Ramsey graph with sufficiently many vertices has induced subgraphs of all sizes up to $(1 - \varepsilon)e(G)$. But what about the distribution of random subgraphs? Kwan, Sah, Sauermann and Sawhney showed further that, for any $C > 0$, and $r = r(n)$ bounded away from 0 and 1, in the random size model we have

$$\mathbb{P}[e(S) = x] \leq A/\sigma, \tag{4}$$

while for all $x$ within a bounded number of standard deviations of $\mathbb{E}e(S)$,

$$\mathbb{P}[e(S) = x] \geq a/\sigma,$$

where $a$ and $A$ are positive constants. However, as noted in [10], a local limit need not hold in this setting.

Finally, we note that it would be interesting to prove analogous results for sparse graphs, and for the distribution of other subgraphs (for example, the number of triangles in a random induced subgraph).

# 2 The central limit theorem

The proof of our main central limit theorem is given in two parts. The bulk of the work occurs in the first, where we derive an initial central limit theorem using Stein's method. In the second part, we reformulate this statement to obtain Theorem 1.

It will be helpful to fix some notation that we will use throughout the section. We will be working with a random graph $G = G_{n,M}$ with vertex set $V = V(G)$. For a set $S \subseteq V$ of size $k$, we write $\overline{S} := V \setminus S$ and $\overline{k} := n - k$. We also write $e(S)$ for the number of edges in the induced subgraph $G[S]$. For a vertex $x \in V$, we will write $d_S(x)$ for the number of edges between $x$ and $S$. We also write $d(G)$ for the average degree of $G$.

## 2.1 Stein's method

We will be using Stein's method of exchangeable pairs where we modify $S$ to a new set $S'$ by removing a random element of $S$ and adding a random element of $\overline{S}$. Unfortunately, the expected change $e(S') - e(S)$ depends not just on $e(S)$ but also on the number of edges between $S$ and $\overline{S}$, or equivalently, given that $e(G)$ is fixed, on $e(\overline{S})$. Thus it will be necessary to track both $e(S)$ and $e(\overline{S})$. For this we will need a version of Stein's method that can be applied to the 2-dimensional vector $(e(S), e(\overline{S}))$.

We will use the following $d$-dimensional version of Stein's method which occurs as a special case of a result of Reinert and Röllin [11, Corollary 3.1]. A pair of random variables $(\boldsymbol{W}, \boldsymbol{W}')$ is called *exchangeable* if $(\boldsymbol{W}, \boldsymbol{W}')$ and $(\boldsymbol{W}', \boldsymbol{W})$ have the same distribution.

**Theorem 3.** *Suppose that* $(\boldsymbol{W}, \boldsymbol{W}')$ *is an exchangeable pair of* $\mathbb{R}^d$*-valued random column vectors* $\boldsymbol{W} = (W_1, \ldots, W_d)^T$ *and* $\boldsymbol{W}' = (W_1', \ldots, W_d')^T$ *such that*

$$\mathbb{E}[\boldsymbol{W}] = 0 \quad and \quad \mathbb{E}[\boldsymbol{W}\boldsymbol{W}^T] = \Sigma, \tag{5}$$

*where $\Sigma$ is a symmetric positive definite $d \times d$ real matrix. Suppose further that*

$$\mathbb{E}\big[\boldsymbol{W}' - \boldsymbol{W} \mid \boldsymbol{W}\big] = -\Lambda\boldsymbol{W} \tag{6}$$

*for an invertible $d \times d$ real matrix $\Lambda$. For $i = 1, \ldots, d$, let $\lambda^{(i)} = \sum_{m=1}^{d} |(\Sigma^{-1/2}\Lambda^{-1}\Sigma^{1/2})_{m,i}|$,*

$$A = \sum_{i,j} \lambda^{(i)} \Big( \operatorname{Var} \mathbb{E}\Big[ \sum_{k,\ell} \Sigma_{ik}^{-1/2} \Sigma_{j\ell}^{-1/2} (W_k' - W_k)(W_\ell' - W_\ell) \Big| \boldsymbol{W}\Big]\Big)^{1/2},$$

$$B = \sum_{i,j,k} \lambda^{(i)} \, \mathbb{E}\, \Big| \sum_{r,s,t} \Sigma_{ir}^{-1/2} \Sigma_{js}^{-1/2} \Sigma_{kt}^{-1/2} (W_r' - W_r)(W_s' - W_s)(W_t' - W_t)\Big|, \text{ and}$$

$$T = \tfrac{1}{4d}\Big( \tfrac{A}{2} + \sqrt{\sqrt{d}B + \tfrac{A^2}{4}} \Big)^2.$$

*Then, with $Z' \sim N(0,1)$ a standard normal random variable, we have*

$$\big|\mathbb{P}(W_1 \leq \sqrt{\Sigma_{11}}z) - \mathbb{P}(Z' \leq z)\big| \leq \gamma^2 \big( -\tfrac{A}{2}\log T + \tfrac{B}{2\sqrt{T}} + 2\sqrt{d}T\big),$$

*where $\gamma = \gamma(d)$ is a function of $d$ only.*

*Proof.* To deduce this from [11, Corollary 3.1], and following the notation used there, we take the test function

$$h(x) := \mathbb{1}_{(-\infty, \sqrt{\Sigma_{11}}z] \times \mathbb{R}^{d-1}}(x),$$

which is the indicator function of a convex set in $\mathbb{R}^d$. As discussed in [11] (see also [3]), this $h$ satisfies the criteria for non-smooth test functions with constant $a \leq 2\sqrt{d}$. Also, we take $R = 0$ and hence $C' = 0$, $D' = A/2$. Letting $\boldsymbol{Z} = (Z_1, Z_2)$ be a standard bivariate and $Z'$ a standard univariate normal random variable, the supremum in [11, Corollary 3.1] is an upper bound for the expression

$$\mathbb{E}\big[h(\boldsymbol{W}) - h(\Sigma^{1/2}\boldsymbol{Z})\big] = \mathbb{P}\big(W_1 \leq \sqrt{\Sigma_{11}}z\big) - \mathbb{P}\big(Z' \leq z\big),$$

as $(\Sigma^{1/2}\boldsymbol{Z})_1 \sim \sqrt{\Sigma_{11}}Z'$. $\qquad\qquad\square$

In our case, we will apply Theorem 3 with $d = 2$, and $\boldsymbol{W}$ being a normalised version of $(e(S), e(\overline{S}))$. Note that $\mathbb{E}_S(e(S)) = KM/N$ and $\mathbb{E}_S(e(\overline{S})) = \overline{K}M/N$, where $\overline{K} = \binom{\overline{k}}{2}$ and the expectation is over the choice of $S$. We define two random variables $W$ and $\overline{W}$ by

$$W := \frac{e(S) - KM/N}{n}, \qquad \overline{W} := \frac{e(\overline{S}) - \overline{K}M/N}{n},$$

so that $\mathbb{E}[W] = \mathbb{E}[\overline{W}] = 0$. Now pick a uniformly chosen vertex $x$ of $S$ and an independent uniformly chosen vertex $\bar{x}$ of $\overline{S}$. Swap the vertices $x$ and $\bar{x}$ to produce two new subsets $S' := S \setminus \{x\} \cup \{\bar{x}\}$ and $\overline{S}' := \overline{S} \setminus \{\bar{x}\} \cup \{x\}$ (see Fig. 1) and define two more random variables $W'$ and $\overline{W}'$ analogously using the subsets $S'$ and $\overline{S}'$. We note that the random variables $(W, \overline{W})$ and $(W', \overline{W}')$ are exchangeable as $S'$ is uniformly distributed over $k$-sets. Write $\boldsymbol{W}$ for the column vector $(W, \overline{W})^T$ and similarly for $\boldsymbol{W}'$.

In light of Theorem 3, the bulk of the work in this section consists of verifying conditions (5) and (6), which we do in Lemma 5 and Lemma 4 respectively. We will then also
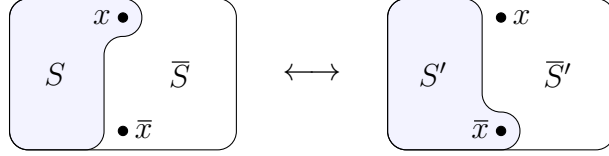
Figure 1: Vertex interchange in $G$.

show that $A$ and $B$ are $O(n^{-1/2+\varepsilon})$, from which a bound of $O(n^{-1/4+\varepsilon})$ will follow. While some of these lemmas are specialised to the random graph setting in which we will apply them, the first few hold in greater generality.

We begin by verifying (6). Given $S$, we perform a random exchange of vertices to obtain $S'$. The expectation $\mathbb{E}[e(S') - e(S) \mid S]$ is then a linear function of $e(S)$ and $e(S, \overline{S})$ which, given that $e(S) + e(\overline{S}) + e(S, \overline{S}) = M$, is also a linear function of $e(S)$ and $e(\overline{S})$.

**Lemma 4.** *For any graph $G$ and $k, \overline{k} > 0$ we have*

$$\mathbb{E}[\boldsymbol{W}' - \boldsymbol{W} \mid \boldsymbol{W}] = -\Lambda \boldsymbol{W}, \quad where \quad \Lambda = \frac{1}{k\overline{k}} \begin{pmatrix} n + \overline{k} - 1 & k - 1 \\ \overline{k} - 1 & n + k - 1 \end{pmatrix}.$$

*Proof.* We fix $S$ and take expectations over the choice of $(x, \overline{x})$. Note that

$$e(S') - e(S) = d_S(\overline{x}) - d_S(x) - \mathbb{1}_{x\overline{x}}, \tag{7}$$

where $d_S(y)$ is the number of edges from $y$ to $S$ and $\mathbb{1}_{xy}$ represents an indicator function of the event that $xy \in E(G)$. Now fixing $G$ and $S$ and taking expectations over $(x, \overline{x})$ we have

$$
\begin{aligned}
\mathbb{E}[e(S') - e(S) \mid S] &= \mathbb{E}_{\overline{x}} d_S(\overline{x}) - \mathbb{E}_x d_S(x) - \mathbb{E}_{x,\overline{x}} \mathbb{1}_{x\overline{x}} \\
&= \frac{e(S,\overline{S})}{\overline{k}} - \frac{2e(S)}{k} - \frac{e(S,\overline{S})}{k\overline{k}} \\
&= \frac{1}{k\overline{k}}\big((k-1)e(S,\overline{S}) - 2\overline{k}e(S)\big) \\
&= \frac{1}{k\overline{k}}\big((k-1)M - (k-1+2\overline{k})e(S) - (k-1)e(\overline{S})\big) \\
&= \frac{1}{k\overline{k}}\big((k-1)M - (n-1+\overline{k})KM/N - (k-1)\overline{K}M/N\big) \\
&\quad - \frac{1}{k\overline{k}}\big((n-1+\overline{k})nW + (k-1)n\overline{W}\big) \\
&= -\frac{1}{k\overline{k}}\big((n-1+\overline{k})nW + (k-1)n\overline{W}\big),
\end{aligned}
$$

where we have used that $k + \overline{k} = n$ and $\mathbb{E}[e(S') - e(S)] = \mathbb{E}[W] = \mathbb{E}[\overline{W}] = 0$, so that the constant term must vanish. Hence (as $\boldsymbol{W}$ is $S$-measurable and the expectation only depends on $\boldsymbol{W}$)

$$\mathbb{E}[W' - W \mid \boldsymbol{W}] = -\frac{n-1+\overline{k}}{k\overline{k}} W - \frac{k-1}{k\overline{k}} \overline{W}.$$

By symmetry, $\mathbb{E}[\overline{W}' - \overline{W} \mid \boldsymbol{W}] = -\frac{n-1+k}{k\overline{k}} \overline{W} - \frac{\overline{k}-1}{k\overline{k}} W$ and the result follows. $\square$

We note for future reference that

$$\Lambda^{-1} = \frac{k\overline{k}}{2n(n-1)} \begin{pmatrix} n + k - 1 & 1 - k \\ 1 - \overline{k} & n + \overline{k} - 1 \end{pmatrix} \tag{8}$$

6

and the entries of $\Lambda^{-1}$ are all $O(n)$.

We now turn to verifying (5). We have already noted that $\mathbb{E}\boldsymbol{W} = 0$, and must check that $\boldsymbol{W}$ has a positive definite covariance matrix $\Sigma$.

Let $P$ be the number of copies of the path on three vertices in $G$, so that

$$P = \sum_{v \in V(G)} \binom{d(v)}{2}.$$

Recall that $d(G) = 2M/n$ is the average degree of $G$ and let $V$ be $n$ times the variance of the degrees of $G$, so that

$$V = \sum_{v \in V(G)} d(v)^2 - nd(G)^2.$$

We note that

$$2Pn = Vn + 4M^2 - 2Mn. \tag{9}$$

**Lemma 5.** *For any graph $G$ and $k, \bar{k} \geq 2$, we have $\mathbb{E}[\boldsymbol{W}\boldsymbol{W}^T] = \Sigma$, where*

$$\Sigma := \frac{2K\overline{K}}{n^2 N^2(n-2)(n-3)} \begin{pmatrix} M(N-M) + \frac{k-2}{k-1}NV & M(N-M) - NV \\ M(N-M) - NV & M(N-M) + \frac{\bar{k}-2}{\bar{k}-1}NV \end{pmatrix}.$$

*In particular, $\Sigma$ is symmetric and positive definite if and only if $V > 0$.*

*Proof.* We first note that

$$\mathbb{E}\big[W^2\big] = \frac{1}{n^2}\mathrm{Var}(e(S)) = \frac{1}{n^2} \sum_{e,f \in E(G)} \left(\mathbb{P}(e, f \subseteq S) - \mathbb{P}(e \subseteq S)\mathbb{P}(f \subseteq S)\right),$$

where the sum is over ordered pairs $(e, f)$ of edges of $G$. We can classify these pairs according to the size of $e \cap f$. Namely, there are $M$ pairs where $e = f$, $2P$ pairs where $|e \cap f| = 1$ and $M^2 - M - 2P$ pairs where $e \cap f = \emptyset$. From this we deduce that

$$\mathbb{E}\big[W^2\big] = \frac{1}{n^2} \left( \frac{(k)_2}{(n)_2} M + \frac{(k)_3}{(n)_3} 2P + \frac{(k)_4}{(n)_4} (M^2 - M - 2P) - \frac{(k)_2^2}{(n)_2^2} M^2 \right),$$

where $(x)_r = x(x-1)\dots(x-r+1)$ denotes the falling factorial. Note that $k, \bar{k} \geq 2$ implies $n \geq 4$, so $(n)_4 > 0$. Using (9) to rewrite $2P$ in terms of $M$, $n$ and $V$ and determining the coefficients of the terms containing $M$, $M^2$ and $V$, respectively, gives

$$\mathbb{E}\big[W^2\big] = \frac{2\bar{k}K}{n^2 N(n)_4} \big(M(N-M)(\bar{k}-1) + (k-2)NV\big),$$

as well as the symmetric expression

$$\mathbb{E}\big[\overline{W}^2\big] = \frac{2k\overline{K}}{n^2 N(n)_4} \big(M(N-M)(k-1) + (\bar{k}-2)NV\big).$$

7

A similar calculation also gives

$$\mathbb{E}\big[W\overline{W}\big] = \frac{1}{n^2} \sum_{e,f \in E(G)} \big(\mathbb{P}(e \subseteq S, f \subseteq \overline{S}) - \mathbb{P}(e \subseteq S)\mathbb{P}(f \subseteq \overline{S})\big)$$

$$= \frac{1}{n^2}\left(\frac{(k)_2(\overline{k})_2}{(n)_4}(M^2 - M - 2P) - \frac{(k)_2(\overline{k})_2}{(n)_2^2}M^2\right)$$

$$= \frac{4K\overline{K}}{n^2 N(n)_4}\big(M(N - M) - NV\big).$$

The expression for $\Sigma$ then follows from these expressions. We note that for any vector $v = (\alpha, \beta)^T$, $v^T \Sigma v$ is the variance of the random variable $\alpha W + \beta \overline{W}$. Thus if $\Sigma$ is not positive definite, then there must be a non-trivial linear relationship $\alpha e(S) + \beta e(\overline{S}) = c$ for all choices of $S$. Assume first that $\alpha \neq -\beta$. Then by interchanging $x \in S$ with $\overline{x} \in \overline{S}$ we deduce from (7) that

$$\alpha\big(d_S(x) - d_S(\overline{x}) - \mathbb{1}_{x\overline{x}}\big) + \beta\big(d_{\overline{S}}(\overline{x}) - d_{\overline{S}}(x) - \mathbb{1}_{x\overline{x}}\big) = 0.$$

Thus $\alpha d_S(x) - \beta d_{\overline{S}}(x)$ is a constant ($= \alpha d_S(\overline{x}) - \beta d_S(\overline{x})$) for all $x \in S \setminus N(\overline{x})$, and a different constant ($= \alpha d_S(\overline{x}) - \beta d_S(\overline{x}) + (\alpha + \beta)$) for all $x \in S \cap N(\overline{x})$. As this holds for all choices of $\overline{x} \in \overline{S}$ we deduce that $S \cap N(\overline{x})$ is independent of $\overline{x} \in \overline{S}$. But by varying $S$ (and assuming $k, \overline{k} \geq 2$) we deduce that $N(\overline{x}) \setminus \{\overline{x}, \overline{y}\} = N(\overline{y}) \setminus \{\overline{x}, \overline{y}\}$ for any two vertices $\overline{x}, \overline{y} \in V(G)$. This is a contradiction unless $G$ is either the empty graph or the complete graph. On the other hand, $v^T \Sigma v = 0$ for $v = (1, -1)^T$ implies that $d(x) = d_S(x) + d_{\overline{S}}(x)$ is constant for all $x \in S$ which, as this holds for all $S$ and $k \geq 2$, immediately implies that $G$ is a regular graph. Thus in all cases $G$ is regular, which implies $V = 0$. Conversely, if $V = 0$ then $\Sigma$ is clearly singular. $\qquad\square$

In the case that $V = 0$, so $G$ is regular, $W = \overline{W}$ and Lemma 5 reduces to the statement that $\mathbb{E}[W^2] = \frac{2K\overline{K}}{n^2 N^2(n-2)(n-3)}M(N - M)$.

We now move on to bounds for the terms $A$ and $B$ appearing in Theorem 3. We start with the following lemma which implies a suitable bound on $B$.

**Lemma 6.** *Suppose that, for a constant $\varepsilon > 0$, $G$ satisfies*

$$\frac{1}{n} \sum_{v \in V(G)} |d(v) - d(G)|^3 = O(n^{3/2+\varepsilon}). \tag{10}$$

*Then $\mathbb{E}\big[|W' - W|^i |\overline{W}' - \overline{W}|^{3-i}\big] = O(n^{-3/2+\varepsilon})$ for $i = 0, \dots, 3$.*

*Remark 7.* Indeed, for any $\varepsilon > 0$, condition (10) holds w.v.h.p. in $G_{n,p}$ and hence also in $G_{n,M}$ because, for a fixed vertex $v$, w.v.h.p. we have $|d(v) - d(G)| \leq n^{1/2+\varepsilon/3}$, and thus by a union bound this holds w.v.h.p. for all vertices.

*Proof.* It is enough that the third moments $\mathbb{E}[|e(S') - e(S)|^3]$ and $\mathbb{E}[|e(\overline{S}') - e(\overline{S})|^3]$ are both $O(n^{3/2+\varepsilon})$ as $W' - W = (e(S') - e(S))/n$, $\overline{W}' - \overline{W} = (e(\overline{S}') - e(\overline{S}))/n$, and in general $\mathbb{E}|XY^2| \leq (\mathbb{E}|X^3|)^{1/3}(\mathbb{E}|Y^3|)^{2/3}$ and $\mathbb{E}|X^2Y| \leq (\mathbb{E}|X^3|)^{2/3}(\mathbb{E}|Y^3|)^{1/3}$ by Hölder's

inequality. We shall only show that $\mathbb{E}[|e(S') - e(S)|^3] = O(n^{3/2+\varepsilon})$ as the result for $\mathbb{E}[|e(\overline{S}') - e(\overline{S})|^3]$ follows by symmetry.

Using (7), that $\mathbb{E}[(|X| + |Y|)^3] = O(\mathbb{E}|X|^3 + \mathbb{E}|Y|^3)$ for random variables $X$, $Y$, and $d(G) \le n - 1$, we get

$$
\begin{aligned}
\mathbb{E}\big[|e(S') - e(S)|^3\big] &= \mathbb{E}\big[|d_S(\overline{x}) - d_S(x) - \mathbb{1}_{x\overline{x}}|^3\big] \\
&= O\big(\mathbb{E}\big[|d_S(\overline{x}) - \tfrac{k}{n-1}d(G)|^3\big] + \mathbb{E}\big[|d_S(x) - \tfrac{k-1}{n-1}d(G)|^3\big] + 1\big).
\end{aligned}
$$

Thus we need to show that $\mathbb{E}[|d_S(\overline{x}) - \tfrac{k}{n-1}d(G)|^3]$ and $\mathbb{E}[|d_S(x) - \tfrac{k-1}{n-1}d(G)|^3]$ are $O(n^{3/2+\varepsilon})$. Fixing $\overline{x}$ and taking $S$ to be a random $k$-subset of $V(G) \setminus \{\overline{x}\}$ gives

$$
\mathbb{E}\big[|d_S(\overline{x}) - \tfrac{k}{n-1}d(G)|^3 \mid \overline{x}\big] = O\big(\mathbb{E}\big[|X - \mathbb{E}[X]|^3\big] + \big|\mathbb{E}[X] - \tfrac{k}{n-1}d(G)\big|^3\big),
$$

where $X$ is a hypergeometric random variable obtained by selecting $k$ items from a set $V(G) \setminus \{\overline{x}\}$ of size $n - 1$ and counting the number of these which occur in a given set $N(\overline{x})$ of size $d(\overline{x})$. We note that $\mathbb{E}[|X - \mathbb{E}X|^3] = O(n^{3/2})$ and $\mathbb{E}X = \tfrac{k}{n-1}d(\overline{x})$. Then randomising over $\overline{x}$, and using $k \le n$ and (10), we see that

$$
\begin{aligned}
\mathbb{E}\big[|d_S(\overline{x}) - \tfrac{k}{n-1}d(G)|^3\big] &= \tfrac{1}{n}\sum_{\overline{x}} \mathbb{E}\big[|d_S(\overline{x}) - \tfrac{k}{n-1}d(G)|^3 \mid \overline{x}\big] \\
&= O\Big(n^{3/2} + \tfrac{1}{n}\sum_{\overline{x}}|d(\overline{x}) - d(G)|^3\Big) = O\big(n^{3/2+\varepsilon}\big).
\end{aligned}
$$

The proof is completed by an analogous argument for $\mathbb{E}[|d_S(x) - \tfrac{k-1}{n-1}d(G)|^3]$. □

The last missing ingredient for our application of Theorem 3 is a suitable bound on the $A$ appearing there, which will be shown in Lemma 10 below. First, we need a concentration estimate for the variance (and covariance) of various degrees in $G$.

**Lemma 8.** *Let $p = p(n) \in [\delta, 1 - \delta]$ and fix a set $S$ of size $k$. Let $x$ be a random element of $S$ and $\overline{x}$ a random element of $\overline{S}$ as before. Then, for any $\varepsilon > 0$, w.v.h.p. in $G_{n,p}$ we have*

$$
\big|\mathrm{Var}_x d_S(x) - p(1-p)k\big| = O(n^{1/2+\varepsilon}), \tag{11}
$$

$$
\big|\mathrm{Var}_{\overline{x}} d_S(\overline{x}) - p(1-p)k\big| = O(n^{1/2+\varepsilon}), \tag{12}
$$

*and*

$$
\mathrm{Cov}_{\overline{x}}\big(d_S(\overline{x}), d_{\overline{S}}(\overline{x})\big) = \mathbb{E}_{\overline{x}}\big[d_S(\overline{x})d_{\overline{S}}(\overline{x})\big] - \mathbb{E}_{\overline{x}}\big[d_S(\overline{x})\big]\mathbb{E}_{\overline{x}}\big[d_{\overline{S}}(\overline{x})\big] = O(n^{1/2+\varepsilon}). \tag{13}
$$

*Proof.* In the following sums, the indices $v$, $w$, $y$, $z$ run over vertices in $S$, and the indices $\overline{v}$, $\overline{w}$, $\overline{y}$, $\overline{z}$ run over vertices in $\overline{S}$. We use primes on sums to indicate that summation is over tuples of distinct elements only. Let $\mathbb{1}_{vw}$ be the indicator variable that the edge $vw$ is present in $G_{n,p}$ as before; in particular we have $\mathbb{1}_{vv} = 0$ for any vertex $v$. We may then

expand

$$V_1 := k^2 \cdot \mathrm{Var}_x d_S(x) = k \sum_v d_S(v)^2 - \sum_{v,w} d_S(v) d_S(w)$$

$$= k \sum_{v,y,z} \mathbb{1}_{vy} \mathbb{1}_{vz} - \sum_{v,w,y,z} \mathbb{1}_{vy} \mathbb{1}_{wz}$$

$$= (k-2) \sideset{}{'}\sum_{v,y} \mathbb{1}_{vy} + (k-4) \sideset{}{'}\sum_{v,y,z} \mathbb{1}_{vy} \mathbb{1}_{vz} - \sideset{}{'}\sum_{v,w,y,z} \mathbb{1}_{vy} \mathbb{1}_{wz}.$$

Similarly,

$$V_2 := \bar{k}^2 \cdot \mathrm{Var}_{\bar{x}} d_S(\bar{x}) = \bar{k} \sum_{\bar{v}} d_S(\bar{v})^2 - \sum_{\bar{v},\bar{w}} d_S(\bar{v}) d_S(\bar{w})$$

$$= \bar{k} \sum_{\bar{v},y,z} \mathbb{1}_{\bar{v}y} \mathbb{1}_{\bar{v}z} - \sum_{\bar{v},\bar{w},y,z} \mathbb{1}_{\bar{v}y} \mathbb{1}_{\bar{w}z}$$

$$= (\bar{k}-1) \sum_{\bar{v}} \sum_{y} \mathbb{1}_{\bar{v}y} + (\bar{k}-1) \sum_{\bar{v}} \sideset{}{'}\sum_{y,z} \mathbb{1}_{\bar{v}y} \mathbb{1}_{\bar{v}z} - \sideset{}{'}\sum_{\bar{v},\bar{w}} \sum_{y,z} \mathbb{1}_{\bar{v}y} \mathbb{1}_{\bar{w}z},$$

and

$$V_3 := \bar{k}^2 \cdot \mathrm{Cov}_{\bar{x}}(d_S(\bar{x}), d_{\bar{S}}(\bar{x})) = \bar{k} \sum_{\bar{v}} d_S(\bar{v}) d_{\bar{S}}(\bar{v}) - \sum_{\bar{v},\bar{w}} d_S(\bar{v}) d_{\bar{S}}(\bar{w})$$

$$= \bar{k} \sum_{\bar{v},\bar{y},y} \mathbb{1}_{\bar{v}y} \mathbb{1}_{\bar{v}\,\bar{y}} - \sum_{\bar{v},\bar{w},\bar{y},y} \mathbb{1}_{\bar{v}y} \mathbb{1}_{\bar{w}\,\bar{y}}$$

$$= (\bar{k}-2) \sideset{}{'}\sum_{\bar{v},\bar{y}} \sum_{y} \mathbb{1}_{\bar{v}y} \mathbb{1}_{\bar{v}\,\bar{y}} - \sideset{}{'}\sum_{\bar{v},\bar{w},\bar{y}} \sum_{y} \mathbb{1}_{\bar{v}y} \mathbb{1}_{\bar{w}\,\bar{y}}.$$

This allows us to compute

$$\mathbb{E}_G[V_1] = (k-2) \cdot (k)_2 \, p + (k-4) \cdot (k)_3 \, p^2 - 1 \cdot (k)_4 \, p^2 = (k)_3 \, p(1-p),$$
$$\mathbb{E}_G[V_2] = (\bar{k}-1) \cdot \bar{k} k p + (\bar{k}-1) \cdot \bar{k} (k)_2 \, p^2 - 1 \cdot (\bar{k})_2 k^2 p^2 = (\bar{k})_2 k p(1-p),$$
$$\mathbb{E}_G[V_3] = (\bar{k}-2) \cdot (\bar{k})_2 k p^2 - 1 \cdot (\bar{k})_3 k p^2 = 0.$$

Now for the concentration inequalities, enumerate the edges of $K_n$ as $e_1, \ldots, e_{\binom{n}{2}}$ and let $\mathcal{F}_s$ be the $\sigma$-algebra generated by $\mathbb{1}_{e_1}, \ldots, \mathbb{1}_{e_s}$. For $i = 1, 2, 3$, write

$$X_s^{(i)} = \mathbb{E}[V_i \mid \mathcal{F}_s]$$

to denote the standard edge exposure martingale for $V_i$. We show that for $i = 1, 2, 3$, w.v.h.p. $|X_s^{(i)} - X_{s-1}^{(i)}| = O(n^{3/2+\varepsilon})$ for all $s$. Suppose we expose the edge $e_s = ab$ at step $s$. If we write $V_1 = c_{ab} \mathbb{1}_{ab} + c$, where $c_{ab}$ and $c$ do not depend on $\mathbb{1}_{ab}$, then the coefficient $c_{ab}$ is given by

$$c_{ab} = (k-2) \cdot 2 + (k-4) \cdot 2 \sum_{y \neq a,b} (\mathbb{1}_{ay} + \mathbb{1}_{by}) - 4 \sideset{}{'}\sum_{y,z \neq a,b} \mathbb{1}_{yz}$$

$$= 2k d_S(a) + 2k d_S(b) - 4 \sum_y d_S(y) + O(n).$$

In the last step, to obtain a simpler expression, we added some terms of order $O(n)$ that do depend on $\mathbb{1}_{ab}$, but that does not change the fact that $c_{ab}$ does not. We now get

$$
\begin{aligned}
\left|X_s^{(1)} - X_{s-1}^{(1)}\right| &= \left(\mathbb{1}_{ab} - p\right)\mathbb{E}[c_{ab} \mid \mathcal{F}_s] \\
&= 2\left(\mathbb{1}_{ab} - p\right)\left(k\mathbb{E}[d_S(a) \mid \mathcal{F}_s] + k\mathbb{E}[d_S(b) \mid \mathcal{F}_s] - 2\sum_y \mathbb{E}[d_S(y) \mid \mathcal{F}_s]\right) + O(n).
\end{aligned}
$$

The coefficient of $\mathbb{1}_{ab}$ in $V_2$ is 0 if $a$ and $b$ are both in $S$, or both in $\overline{S}$. Otherwise we may assume $a \in S$, $b \in \overline{S}$, and then the coefficient is

$$
c_{ab} = (\overline{k} - 1) + (\overline{k} - 1) \cdot 2\sum_{y \neq a} \mathbb{1}_{by} - 1 \cdot 2 \sum_{y \neq a, \overline{y} \neq b} \mathbb{1}_{\overline{y}y} = 2\overline{k}d_S(b) - 2\sum_{\overline{y}} d_S(\overline{y}) + O(n),
$$

yielding

$$
\left|X_s^{(2)} - X_{s-1}^{(2)}\right| = 2\left(\mathbb{1}_{ab} - p\right)\left(\overline{k}\,\mathbb{E}[d_S(b) \mid \mathcal{F}_s] - \sum_{\overline{y}} \mathbb{E}[d_S(\overline{y}) \mid \mathcal{F}_s]\right) + O(n).
$$

The coefficient of $\mathbb{1}_{ab}$ in $V_3$ is also 0 if $a$ and $b$ are both in $S$. When $a, b \in \overline{S}$, the coefficient is

$$
c_{ab} = (\overline{k} - 2)\sum_y (\mathbb{1}_{ay} + \mathbb{1}_{by}) - 2\sum_{\overline{y} \neq a, b}\sum_y \mathbb{1}_{\overline{y}y} = \overline{k}d_S(a) + \overline{k}d_S(b) - 2\sum_{\overline{y}} d_S(\overline{y}) + O(n),
$$

which gives

$$
\left|X_s^{(3)} - X_{s-1}^{(3)}\right| = \left(\mathbb{1}_{ab} - p\right)\left(\overline{k}\,\mathbb{E}[d_S(a) + d_S(b) \mid \mathcal{F}_s] - 2\sum_{\overline{y}} \mathbb{E}[d_S(\overline{y}) \mid \mathcal{F}_s]\right) + O(n).
$$

When $a \in S$, $b \in \overline{S}$, we have

$$
c_{ab} = (\overline{k} - 2)\sum_{\overline{z}} \mathbb{1}_{b\overline{z}} - {\sum_{\overline{w}, \overline{z} \neq b}}' \mathbb{1}_{\overline{w}\,\overline{z}} = \overline{k}d_{\overline{S}}(b) - \sum_{\overline{w}} d_{\overline{S}}(\overline{w}) + O(n),
$$

and hence

$$
\left|X_s^{(3)} - X_{s-1}^{(3)}\right| = \left(\mathbb{1}_{ab} - p\right)\left(\overline{k}\,\mathbb{E}[d_{\overline{S}}(b) \mid \mathcal{F}_s] - \sum_{\overline{w}} \mathbb{E}[d_{\overline{S}}(\overline{w}) \mid \mathcal{F}_s]\right) + O(n).
$$

Applying the Chernoff bound to the exposed edges, we deduce that for a fixed $s$ and $v \in V(G)$, w.v.h.p., $\mathbb{E}[d_S(v) \mid \mathcal{F}_s] = pk + O(n^{1/2+\varepsilon/2})$. Hence, by a union bound, we have that w.v.h.p. this holds for all $v$ and all $s$. Similarly, we can show that $\mathbb{E}[d_{\overline{S}}(v) \mid \mathcal{F}_s] = p\overline{k} + O(n^{1/2+\varepsilon/2})$ for all $v$ and $s$. This can now be applied to all the expressions derived above to see that, for $i = 1, 2, 3$, we have w.v.h.p. $|X_s^{(i)} - X_{s-1}^{(i)}| = O(n^{3/2+\varepsilon/2})$ for all $s$. The Azuma–Hoeffding inequality (Lemma 22) with $t = n^{5/2+\varepsilon}$, $c_s = O(n^{3/2+\varepsilon/2})$ and hence $\sigma^2 = O(n^{5+\varepsilon})$ now gives the desired bound

$$
\mathbb{P}_G\big(|V_i - \mathbb{E}[V_i]| > t\big) \leq \exp(-n^{\Omega(1)}) + 2\exp\big(-t^2/2\sigma^2\big) = \exp\big(-n^{\Omega(1)}\big)
$$

for $i = 1, 2, 3$. The result follows as $k, \overline{k} = \Theta(n)$. $\qquad\square$

*Remark* 9. The proof of the first part of Lemma 8 (that is, equation (11)), does not require that $\bar{k} \geq \delta n$ and hence the result also holds for $S = V(G)$. This gives precisely that $V = p(1-p)n^2 + O(n^{3/2+\varepsilon})$, which we will use in the following subsection.

**Lemma 10.** *In* $G = G_{n,M}$, *for* $i = 0, 1, 2$ *and all* $\varepsilon > 0$, *w.v.h.p. we have*

$$\mathrm{Var}\big(\mathbb{E}[(W' - W)^i(\overline{W}' - \overline{W})^{2-i} \mid \boldsymbol{W}]\big) = O(n^{-3+\varepsilon}).$$

*Proof.* Let $0 < \varepsilon' < \varepsilon/2$. We work in $G_{n,p}$ with $p = M/\binom{n}{2}$ and start with a fixed set $S$. We then have

$$
\begin{aligned}
\mathbb{E}_{x,\bar{x}}\big[(e(S') - e(S))^2\big] &= \mathbb{E}_{x,\bar{x}}\big[(d_S(\bar{x}) - d_S(x) - \mathbb{1}_{x\bar{x}})^2\big] \\
&= \mathbb{E}_{x,\bar{x}}\big[(d_S(\bar{x}) - d_S(x))^2\big] + \mathbb{E}_{x,\bar{x}}\big[(1 + 2d_S(\bar{x}) - 2d_S(x))\mathbb{1}_{x\bar{x}}\big] \\
&= \mathrm{Var}_{\bar{x}}d_S(\bar{x}) + \mathrm{Var}_x d_S(x) + \big(\mathbb{E}_{\bar{x}}d_S(\bar{x}) - \mathbb{E}_x d_S(x)\big)^2 \\
&\quad + \mathbb{E}_{x,\bar{x}}\big[(1 + 2d_S(\bar{x}) - 2d_S(x))\mathbb{1}_{x\bar{x}}\big],
\end{aligned}
$$

where we have used the independence of $x$ and $\bar{x}$.

By Lemma 8, w.v.h.p. in $G = G_{n,p}$, $\mathrm{Var}_{\bar{x}}d_S(\bar{x})$ and $\mathrm{Var}_x d_S(x)$ are both $p(1-p)k + O(n^{1/2+\varepsilon'})$. We also have w.v.h.p. that $e(S, S') - pk\bar{k}$ and $e(S) - p\binom{k}{2}$ are $O(n^{1+\varepsilon'})$ and thus

$$\mathbb{E}_{\bar{x}}d_S(\bar{x}) - \mathbb{E}_x d_S(x) = \frac{e(S,\overline{S})}{\bar{k}} - \frac{2e(S)}{k} = pk - p(k-1) + O(n^{\varepsilon'}) = O(n^{\varepsilon'}).$$

Finally, w.v.h.p. for all $y \in S$ and $\bar{y} \in \overline{S}$, we have $d_S(\bar{y}), d_S(y) = pk + O(n^{1/2+\varepsilon'})$. Thus

$$\mathbb{E}_{x,\bar{x}}\big[(1 + 2d_S(\bar{x}) - 2d_S(x))\mathbb{1}_{x\bar{x}}\big] = O(n^{1/2+\varepsilon'}).$$

Putting all of this together, we see that for fixed $S$ w.v.h.p. in $G$ we have

$$\left|\mathbb{E}_{x,\bar{x}}\big[(e(S') - e(S))^2\big] - 2p(1-p)k\right| \leq \tfrac{1}{2}n^{1/2+\varepsilon/2}. \tag{14}$$

With $S$ no longer fixed and denoting by $X$ the number of $S$ failing (14), we have $\mathbb{E}[X] \leq \exp(-2n^{\eta})\binom{n}{k}$ for some constant $\eta > 0$ and hence Markov gives

$$\mathbb{P}_G\big(X > \exp(-n^{\eta})\binom{n}{k}\big) \leq \frac{\mathbb{E}[X]}{\exp(-n^{\eta})\binom{n}{k}} \leq \exp(-n^{\eta}),$$

so w.v.h.p. in $G$ there is at most an $\exp(-n^{\eta})$ proportion of choices of $S$ not satisfying (14).

This property holds w.v.h.p. in $G_{n,p}$ and thus in $G_{n,M}$ and we now fix $G$ with this property. We want to condition on $\boldsymbol{W}$ and let $\eta' < \eta$. Say that $\boldsymbol{w}$ is *typical* if $\mathbb{P}_S(\boldsymbol{W} = \boldsymbol{w}) \geq \exp(-n^{\eta'})$. For a typical $\boldsymbol{w}$, since $|(e(S') - e(S))^2|$ is bounded by $n^4$, we have

$$\left|\mathbb{E}\big[(e(S') - e(S))^2 \mid \boldsymbol{W} = \boldsymbol{w}\big] - 2p(1-p)k\right| \leq \exp(-n^{\eta} + n^{\eta'})n^4 + \tfrac{1}{2}n^{1/2+\varepsilon/2} \leq n^{1/2+\varepsilon/2}.$$

On the other hand, there are at most $\binom{n}{2}^2 \leq n^4$ possible values of $(e(S), e(\overline{S}))$ and hence at most $n^4$ possible values of $\overline{W}$. Thus there are at most $n^4 \exp(-n^{\eta'})\binom{n}{k}$ choices of $S$ where $\boldsymbol{W}$ takes a non-typical value, yielding that

$$\mathbb{P}_S\big(|\mathbb{E}\big[(e(S') - e(S))^2 \mid \boldsymbol{W}\big] - 2p(1-p)k| > n^{1/2+\varepsilon/2}\big) \leq n^4 \exp(-n^{\eta'}).$$

Additionally, using that $0 \leq \mathbb{E}[(e(S') - e(S))^2 \mid \boldsymbol{W}] \leq n^4$, we get

$$\mathrm{Var}\big(\mathbb{E}\big[(e(S') - e(S))^2 \mid \boldsymbol{W}\big]\big) \leq n^{12}\exp(-n^{\eta'}) + n^{1+\varepsilon} = O(n^{1+\varepsilon}).$$

Now noting that $W' - W = (e(S') - e(S))/n$, we obtain that w.v.h.p. in $G_{n,M}$ we have

$$\mathrm{Var}\big(\mathbb{E}\big[(W' - W)^2 \mid \boldsymbol{W}\big]\big) = n^{-4}\mathrm{Var}\big(\mathbb{E}\big[(e(S') - e(S))^2 \mid S\big]\big) = O(n^{-3+\varepsilon}).$$

By symmetry, we also have $\mathrm{Var}(\mathbb{E}[(\overline{W}' - \overline{W})^2 \mid \boldsymbol{W}]) = O(n^{-3+\varepsilon})$. The proof for $\mathrm{Var}(\mathbb{E}[(W' - W)(\overline{W}' - \overline{W}) \mid \boldsymbol{W}])$ is similar. Indeed, for fixed $S$ we can write

$$\begin{aligned}
\mathbb{E}_{x,\bar{x}}\big[(e(S') - e(S))(e(\overline{S}') - e(\overline{S}))\big] &= \mathbb{E}_{x,\bar{x}}\big[(d_S(\bar{x}) - d_S(x) - \mathbb{1}_{x\bar{x}})(d_{\overline{S}}(x) - d_{\overline{S}}(\bar{x}) - \mathbb{1}_{x\bar{x}})\big] \\
&= -\mathrm{Cov}_{\bar{x}}\big(d_S(\bar{x}), d_{\overline{S}}(\bar{x})\big) - \mathrm{Cov}_x\big(d_S(x), d_{\overline{S}}(x)\big) \\
&\quad + \big(\mathbb{E}_{\bar{x}}[d_S(\bar{x})] - \mathbb{E}_x[d_S(x)]\big)\big(\mathbb{E}_x[d_{\overline{S}}(x)] - \mathbb{E}_{\bar{x}}[d_{\overline{S}}(\bar{x})]\big) \\
&\quad + \mathbb{E}_{x,\bar{x}}\big[\mathbb{1}_{x\bar{x}}(1 - d_S(\bar{x}) + d_S(x) - d_{\overline{S}}(x) + d_{\overline{S}}(\bar{x}))\big].
\end{aligned}$$

Once again applying Lemma 8, we know that w.v.h.p. the covariances $\mathrm{Cov}_{\bar{x}}(d_S(\bar{x}), d_{\overline{S}}(\bar{x}))$ and $\mathrm{Cov}_x(d_S(x), d_{\overline{S}}(x))$ are $O(n^{1/2+\varepsilon'})$ in $G_{n,p}$. Further, w.v.h.p. $e(S, S') - pk\bar{k}$, $e(S) - pk^2/2$ and $e(S, \overline{S}) - p\bar{k}^2/2$ are $O(n^{1+\varepsilon'})$ and so

$$\begin{aligned}
\big(\mathbb{E}_{\bar{x}}[d_S(\bar{x})] - \mathbb{E}_x[d_S(x)]\big)\big(\mathbb{E}_x[d_{\overline{S}}(x)] - \mathbb{E}_{\bar{x}}[d_{\overline{S}}(\bar{x})]\big) \\
= \big(e(S, \overline{S})/\bar{k} - 2e(S)/k\big)\big(2e(\overline{S})/\bar{k} - e(S, \overline{S})/k\big) \\
= O(n^{\varepsilon'})O(n^{\varepsilon'}) = O(n^{2\varepsilon'}).
\end{aligned}$$

Finally, as above w.v.h.p. we have $\mathbb{E}_{x,\bar{x}}[\mathbb{1}_{x\bar{x}}(1 - d_S(\bar{x}) + d_S(x) - d_{\overline{S}}(x) + d_{\overline{S}}(\bar{x}))] = O(n^{1/2+\varepsilon'})$. Putting everything together, we obtain $\mathbb{E}_{x,\bar{x}}[(e(S') - e(S))(e(\overline{S}') - e(\overline{S}))] = O(n^{1/2+\varepsilon'})$. Arguing as above, we can then deduce that w.v.h.p. in $G_{n,M}$ we have

$$\begin{aligned}
\mathrm{Var}\big(\mathbb{E}\big[(W' - W)(\overline{W}' - \overline{W}) \mid \boldsymbol{W}\big]\big) &= n^{-4}\mathrm{Var}\big(\mathbb{E}\big[(e(S') - e(S))(e(\overline{S}') - e(\overline{S})) \mid \boldsymbol{W}\big]\big) \\
&= O(n^{-3+\varepsilon})
\end{aligned}$$

as desired. $\qquad\square$

**Lemma 11.** *For any $\varepsilon > 0$, w.v.h.p. in $G_{n,M}$, for any $z$,*

$$\big|\mathbb{P}_S(W \leq \sqrt{\Sigma_{11}}z) - \mathbb{P}_S(Z' \leq z)\big| = O(n^{-1/4+\varepsilon}),$$

*where $Z' \sim N(0,1)$ is a standard normal random variable,*

$$\Sigma_{11} = \mathrm{Var}(W) = \frac{2K\overline{K}}{n^2 N^2 (n-2)(n-3)}\big(M(N-M) + \tfrac{k-2}{k-1}NV\big)$$

*and $V = \sum_{i=1}^n d(v)^2 - nd(G)^2$.*

*Proof.* We apply Theorem 3 to our $(\boldsymbol{W}, \boldsymbol{W}')$. Lemma 4 shows that (6) holds for the $\Lambda$ given there and Lemma 5 shows that (5) holds with the $\Sigma$ given there. We note that under the conditions of the theorem, the entries of both $\Sigma^{1/2}$ and $\Sigma^{-1/2}$ are $O(1)$.

Indeed, we have $V = p(1-p)n^2 + O(n^{3/2+\varepsilon})$ where $p = M/N$ by Remark 9, and hence $NV = (2 + o(1))M(N-M)$. Thus

$$\Sigma = \Theta(1) \cdot \begin{pmatrix} 1 + 2\kappa + o(1) & -1 + o(1) \\ -1 + o(1) & 1 + 2\kappa^{-1} + o(1) \end{pmatrix}$$

where $\kappa = k/\overline{k} = \Theta(1)$. The entries of $\Lambda^{-1}$ are $O(n)$ by (8), and hence the terms $\lambda^{(i)}$ of Theorem 3 are also $O(n)$. Since the entries of $\Sigma^{-1/2}$ are $O(1)$, Lemma 6 gives $B = O(n^{-1/2+\varepsilon})$ and Lemma 10 gives $A = O(n^{-1/2+\varepsilon})$. This then implies that $T = O(n^{-1/2+\varepsilon})$ and $B/\sqrt{T} = O(B/\sqrt{B}) = O(n^{-1/4+\varepsilon})$. Hence our application of Theorem 3 yields, for any $z$,

$$\left| \mathbb{P}_S(W \leq \sqrt{\Sigma_{11}}z) - \mathbb{P}_S(Z' \leq z) \right| = O(n^{-1/4+\varepsilon}),$$

as desired. $\qquad\square$

## 2.2 Proof of Theorem 1

We will now show that the output from Stein's method, namely Lemma 11, can be reformulated into the statement of Theorem 1. First, note that the $V$ and by extension the $\Sigma_{11}$ appearing in Theorem 1 depend on the graph $G = G_{n,M}$. To get around this dependency, we can use the fact that $V$ is well-concentrated in $G_{n,M}$ which was essentially established in Remark 9. The following calculation is the source of the definition $\lambda = \frac{(n^2-k^2)k^2}{2n^4} \cdot \frac{M(N-M)}{N^2}$ that we made in (1).

**Corollary 12.** *For any $\varepsilon > 0$, w.v.h.p. we have $\Sigma_{11} = \lambda(1 + O(n^{-1/2+\varepsilon}))$.*

*Proof.* By Lemma 8 (or specifically Remark 9) we may assume $V = p(1-p)n^2 + O(n^{3/2+\varepsilon})$ where $p = M/N$ is bounded away from 0 and 1. Now as $k, \overline{k} = \Theta(n)$ and $\overline{k} = n - k$,

$$
\begin{aligned}
\Sigma_{11} &= \tfrac{2K\overline{K}}{n^2 N^2 (n-2)(n-3)} \big( M(N-M) + \tfrac{k-2}{k-1}NV \big) \\
&= \tfrac{k^2\overline{k}^2}{2n^4} p(1-p) \big( 1 + \tfrac{2k}{k} + O(n^{-1/2+\varepsilon}) \big) \\
&= \tfrac{k^2}{2n^4} p(1-p) \overline{k}(\overline{k} + 2k) \big( 1 + O(n^{-1/2+\varepsilon}) \big) \\
&= \tfrac{k^2}{2n^4} \cdot \tfrac{M(N-M)}{N^2} \cdot (n^2 - k^2) \big( 1 + O(n^{-1/2+\varepsilon}) \big) \\
&= \lambda \big( 1 + O(n^{-1/2+\varepsilon}) \big). \qquad\square
\end{aligned}
$$

To obtain our modified statement in Theorem 1, in addition to getting rid of dependency on the graph, we also transfer to an expression given in terms of the edge count $e(S)$ without normalisation.

*Proof of Theorem 1.* It is enough to show that the left hand side of (2) is $O(n^{-1/4+\varepsilon})$. As before, let $Z'$ be a standard normal random variable. Using the shorthand $\rho(x) := (x - KM/N)/\sqrt{\Sigma_{11}}n$ and $c := \sqrt{\Sigma_{11}/\lambda}$, we may write

$$\left| \mathbb{P}_S(e(S) \leq z) - \mathbb{P}_S(Z \leq z) \right| = \left| \mathbb{P}_S(W \leq \sqrt{\Sigma_{11}}\rho(z)) - \mathbb{P}_S(Z' \leq \sqrt{\Sigma_{11}/\lambda}\rho(z)) \right|,$$

and this expression can be bounded by

$$\left| \mathbb{P}_S(W \leq \sqrt{\Sigma_{11}}\rho(z)) - \mathbb{P}_S(Z' \leq \rho(z)) \right| + \left| \mathbb{P}_S(Z' \leq \sqrt{\Sigma_{11}/\lambda}\rho(z)) - \mathbb{P}_S(Z' \leq \rho(z)) \right|.$$

The first of these summands is with very high probability $O(n^{-1/4+\varepsilon})$ by Lemma 11. As for the second summand, we may assume $\rho(z) = O(\log n)$ as otherwise $\mathbb{P}(Z' \leq \rho(z))$ and $\mathbb{P}(Z' \leq c\rho(z))$ are both within $O(n^{-1})$ of 0 or 1, depending on the sign on $\rho(z)$. Since the probability density function of $Z'$ is bounded by 1 and $c = 1 + O(n^{-1/2+\varepsilon})$ by Corollary 12, we obtain

$$\left| \mathbb{P}(Z' \leq \rho(z)) - \mathbb{P}(Z' \leq c\rho(z)) \right| \leq |c\rho(z) - \rho(z)| = O(|c-1|\log n) = O(n^{-1/2+2\varepsilon}),$$

completing the proof. $\qquad\square$

# 3   The local limit theorem

Recall that $S$ is a random subset of the vertex set of $G = G_{n,M}$ of size $k$, where $M/N, k/n \in [\delta, 1-\delta]$ with $N = \binom{n}{2}$. We also define $\lambda = \frac{(n^2-k^2)k^2}{2n^4} \cdot \frac{M(N-M)}{N^2}$ and note that $\lambda = \Theta(1)$. We now derive a local limit theorem for the empirical distribution of the number $e(S)$ of edges induced by $S$, where the distribution is over choices of $k$-set $S$ with $G$ fixed, restated below.

**Theorem 2.** *Assume that for some fixed $\delta > 0$, $M/N, k/n \in [\delta, 1-\delta]$. Then for any $\varepsilon > 0$ and w.v.h.p. in $G = G_{n,M}$ we have that for all integers $z$*

$$n\left| \mathbb{P}_S(e(S) = z) - \varphi(z) \right| \leq n^{-1/14+\varepsilon},$$

*where $\varphi$ is the density function of a normal random variable $Z \sim N(KM/N, \lambda n^2)$.*

The conclusion of this theorem and further ones later take the form that, for all values $z \in \{0, \ldots, \binom{n}{2}\}$ (or all intervals of values of $z$), certain events involving $\mathbb{P}(e(S) = z)$ hold with very high probability. We remark again that when we say an event holds w.v.h.p. for any $z$, this implies that w.v.h.p. the event holds for all $z$ at the same time by a simple union bound argument. We will thus pay no particular attention to the order of "w.v.h.p." and "for all $z$".

The proof of Theorem 2 starts from the central limit theorem (Theorem 1), and from there narrows down to point probabilities via a smoothing argument. For a convenient starting point, we reformulate the central limit theorem to deal with probabilities of $e(S)$ being in an interval $[z_0, z_1]$.

**Corollary 13.** *For any $\varepsilon > 0$, w.v.h.p. in $G = G_{n,M}$ we have for any $z_0$, $z_1$ that*

$$\left| \mathbb{P}_S(e(S) \in [z_0, z_1]) - \mathbb{P}_S(Z \in [z_0, z_1]) \right| \leq n^{-1/4+\varepsilon},$$

*where $Z \sim N(KM/N, \lambda n^2)$.*

*Proof.* This follows from Theorem 1 together with the fact that $\mathbb{P}_S(e(S) \in [z_0, z_1]) = \mathbb{P}(e(S) \leq z_1) - \mathbb{P}(e(S) < z_0) = \mathbb{P}(e(S) \leq z_1) - \mathbb{P}(e(S) \leq z_0')$, where $z_0' < z_0$ is sufficiently close to $z_0$. $\qquad\square$

We also record a (weaker) asymptotic upper bound on these probabilities of intervals, which will come in handy in the smoothing procedure that is central in the argument to come.

**Corollary 14.** *For $p = p(n) \in [\delta, 1 - \delta]$ and $\varepsilon > 0$, there is a constant $C > 0$ such that w.v.h.p. in $G_{n,p}$, for any $z$ and any $r$ with $r \geq n^{3/4+\varepsilon}$,*

$$\mathbb{P}_S\big(e(S) \in [z, z + r]\big) \leq Cr/n.$$

*Proof.* With $Z$ as before, Corollary 13 gives that for a fixed $M = M(n)$ w.v.h.p. in $G_{n,M}$, we have $\mathbb{P}(e(S) \in [z, z + r]) \leq \mathbb{P}(Z \in [z, z + r]) + n^{-1/4+\varepsilon}$. Thus, as the pdf of $Z$ is bounded by $1/(n\sqrt{2\pi\lambda})$ (see Lemma 25),

$$\mathbb{P}\big(e(S) \in [z, z + r]\big) \leq \frac{r}{n\sqrt{2\pi\lambda}} + n^{-1/4+\varepsilon} \leq Cr/n$$

for a suitable constant $C$ and provided $r \geq n^{3/4+\varepsilon}$. Since this holds in $G_{n,M}$ for any $M$ with the same constant $C$, revealing the number of edges in $G_{n,p}$, yields the same assertion in $G_{n,p}$, noting that the probability that $M/N \notin [\delta/2, 1 - \delta/2]$ is $o(1/n)$. $\qquad \square$

In addition to the preceding corollaries, the other main ingredients needed for our smoothing method are certain concentration results which we prove in Section 3.1. The smoothing procedure itself is then described and established in Section 3.2, and we apply it to prove Theorem 2 in Section 3.3. While the target local limit theorem is for $G_{n,M}$, it is convenient to work for the most part in $G_{n,p}$ for some $p = p(n) \in [\delta, 1 - \delta]$. We only return to $G_{n,M}$ for the final proof, at which point we transfer results between models as needed.

## 3.1 Concentration of $e(S) - e(S')$ in $G_{n,p}$

With the justification at the end of the preceding section, we now work in $G_{n,p}$ for some $p = p(n) \in [\delta, 1 - \delta]$ and assume $V(G_{n,p}) = [n]$.

To analyse our random model, we view our random set $S$ as being constructed by taking a union $S = S' \cup T$ of a random $(k-t)$-set $S' \subseteq [n]$ and a random $t$-set $T \subseteq [n] \setminus S'$, where $t = t(n)$ will be of order $1 \ll t \ll n$. We will condition on $S'$ and compare probabilities involving $e(S) - e(S')$ to a better-behaved binomial random variable $Y$. Henceforth let $Y \sim \text{Bin}((k - t)t + \binom{t}{2}, p)$ and note that for fixed $S = S' \cup T$, in $G_{n,p}$, $e(S) - e(S')$ is distributed like $Y$.

The next two lemmas show that in $G_{n,p}$, for any interval $A$ of appropriate length and for almost all $S'$, $\mathbb{P}_T(e(S) - e(S') \in A)$ is concentrated around its mean $\mathbb{P}(Y \in A)$. We first do this in Lemma 15 with $S'$ fixed and $T$ chosen from a family of disjoint sets in $[n] \setminus S'$. In this setting, the relevant edge sets are independent allowing us to use simple concentration techniques. The more general case is then contained in Lemma 16.

**Lemma 15** ($T$ from disjoint family). *Assume $p, k/n \in [\delta, 1-\delta]$. Let $\beta, \varepsilon > 0$ be constants and let $a, t \in \mathbb{N}$ be such that $t \leq k/2$ and*

$$a \leq t^{3/2} n^{2\beta - 1/2}. \tag{15}$$

*Let $S' \subseteq [n]$ be fixed with $|S'| = k - t$. Define $m = \lfloor \frac{n-k+t}{t} \rfloor$ and let $\mathcal{T} = \{T_i\}_{i \in [m]}$ be a family of pairwise disjoint $t$-sets in $[n] \setminus S'$. Let $T$ be a uniformly chosen random element*

*of $\mathcal{T}$ and $A = [z, z + a)$ an interval. If $n$ is sufficiently large, then with probability at least $1 - e^{-n^\varepsilon}$ in $G_{n,p}$ we have*

$$\left|\mathbb{P}_T\big(e(S' \cup T) - e(S') \in A\big) - \mathbb{P}(Y \in A)\right| \leq \Delta_A,$$

*where*

$$\Delta_A = \begin{cases} e^{-n^\varepsilon}/m, & \text{if } \mathbb{P}(Y \in A) \leq e^{-n^\varepsilon}/m; \\ n^{\beta+\varepsilon}/m, & \text{otherwise.} \end{cases} \tag{16}$$

*Proof.* In this proof, it will be easier to work with counts of sets rather than probabilities. To that end, let $N_A$ be the number of sets $T \in \mathcal{T}$ such that $e(S' \cup T) - e(S') \in A$. Then we have $\mu_A = \mathbb{E}_G[N_A] = m\mathbb{P}(Y \in A)$, and hence

$$\mathbb{P}_G\big(|\mathbb{P}_T\big(e(S' \cup T) - e(S') \in A\big) - \mathbb{P}(Y \in A)| > \Delta_A\big) = \mathbb{P}_G\big(|N_A - \mu_A| > m\Delta_A\big).$$

We proceed by working with the expression on the right hand side. If $\mathbb{P}(Y \in A) \leq e^{-n^\varepsilon}/m$, then $\mu_A \leq e^{-n^\varepsilon}$, so Markov's inequality gives

$$\mathbb{P}_G(|N_A - \mu_A| > e^{-n^\varepsilon}) \leq \mathbb{P}_G(N_A \geq 1) \leq \mathbb{E}_G(N_A) \leq e^{-n^\varepsilon}.$$

We now turn to the general case. Clearly we may assume $\delta \leq \frac{1}{2}$, so that by our assumptions on $p$ and $k$ and using that $t \leq k/2$ and $k \geq \delta n$ we have

$$\text{Var}(Y) = p(1-p)\big((k-t)t + \binom{t}{2}\big) \geq \delta(1-\delta) \cdot \tfrac{kt}{2} \geq \tfrac{1}{4}\delta^2 nt. \tag{17}$$

Then, by Lemma 23,

$$p' := \mathbb{P}(Y \in A) \leq \sqrt{\frac{\pi}{8\text{Var}(Y)}} \cdot a \leq \frac{2a}{\delta\sqrt{nt}}$$

and we have $N_A \sim \text{Bin}(m, p')$. Hence, by our assumption (15),

$$\mu_A = mp' \leq \frac{n}{t} \cdot \frac{2a}{\delta\sqrt{nt}} \leq 2\delta^{-1}n^{2\beta}.$$

It then follows from Bernstein's inequality (Theorem 21) that, for large enough $n$,

$$\mathbb{P}_G\big(|N_A - \mu_A| > n^{\beta+\varepsilon}\big) \leq 2\exp\left(\frac{-n^{2\beta+2\varepsilon}}{2(\mu_A + n^{\beta+\varepsilon}/3)}\right) \leq e^{-n^\varepsilon}. \qquad \square$$

We now show that the assertion of Lemma 15 still holds for almost all $S'$ when we take a fully random $T \subseteq [n] \setminus S'$. This is a straightforward application of Markov's inequality.

**Lemma 16** (Concentration). *Assume $p$, $k$, $\beta$, $\varepsilon$, $a$ and $t$ are as in Lemma 15. For a fixed $(k-t)$-set $S'$ let $T$ be a random $t$-set in $[n] \setminus S'$. Then w.v.h.p. in $G_{n,p}$, for any interval $A$ as above, at most an $e^{-n^\varepsilon/4}$ proportion of the $(k-t)$-sets $S'$ fail to satisfy*

$$\left|\mathbb{P}_T(e(S', T) + e(T) \in A) - \mathbb{P}(Y \in A)\right| \leq \Delta_A + e^{-n^\varepsilon/2}. \tag{18}$$

17

*Proof.* We begin with a fixed set $S'$. Picking a random $T$ is equivalent to choosing uniformly a random (almost complete) partition of $[n] \setminus S'$ into $t$-sets $\mathcal{T} = \{T_i\}_{i \in [m]}$ and a uniformly random $T \in \mathcal{T}$. Say a family $\mathcal{T}$ is *good* if, when choosing $T$ u.a.r. from $\mathcal{T}$, we have $|\mathbb{P}_T(e(S', T) + e(T) \in A) - \mathbb{P}(Y \in A)| \leq \Delta_A$, otherwise we say $\mathcal{T}$ is *bad*.

When choosing a random $\mathcal{T}$, we have $\mathbb{E}_G[\mathbb{P}_{\mathcal{T}}(\mathcal{T} \text{ is bad})] < e^{-n^\varepsilon}$ by Lemma 15. Applying Markov's inequality then gives

$$\mathbb{P}_G\big(\mathbb{P}_{\mathcal{T}}(\mathcal{T} \text{ is bad}) \geq e^{-n^\varepsilon/2}\big) \leq \mathbb{E}_G[\mathbb{P}_{\mathcal{T}}(\mathcal{T} \text{ is bad})]/e^{-n^\varepsilon/2} \leq e^{-n^\varepsilon/2}.$$

Thus, when choosing $\mathcal{T}$ and $T \in \mathcal{T}$ as described above, with probability at least $1 - e^{-n^\varepsilon/2}$ in $G$ we have

$$\big|\mathbb{P}_T(e(S', T) + e(T) \in A) - \mathbb{P}(Y \in A)\big| \leq \mathbb{P}_{\mathcal{T}}(\mathcal{T} \text{ is good})\Delta_A + \mathbb{P}_{\mathcal{T}}(\mathcal{T} \text{ is bad})$$
$$\leq \Delta_A + e^{-n^\varepsilon/2}.$$

Writing $N_e$ for the number of $S'$ that violate (18), this gives $\mathbb{E}_G[N_e] \leq e^{-n^\varepsilon/2}\binom{n}{k-t}$. By a second application of Markov's inequality, we now have

$$\mathbb{P}_G\big(N_e > e^{-n^\varepsilon/4}\binom{n}{k-t}\big) \leq e^{-n^\varepsilon/4}$$

as desired. $\qquad\square$

## 3.2 Smoothing

For a (fixed) random graph $G$, the only properties that we will now need to analyse the distribution of $e(S)$ are the bounds from the CLT reformulation in Corollary 14 and the concentration property given in Lemma 16. We see directly from these two statements that w.v.h.p. $G_{n,p}$ (with $p \in [\delta, 1 - \delta]$) satisfies both of these. The goal of this subsection is to show that for $G$ with these properties the probability that $e(S) = z$, or the *weight* of $z$, is evenly distributed for points $z$ within intervals of a suitable length in the following sense.

**Lemma 17** (Smoothing Lemma). *Fix $\beta \in (0, \frac{1}{10})$ and $\varepsilon \in (0, \frac{1-10\beta}{4})$. Then there exists an integer $a \in [n^{1-5\beta/2-\varepsilon}, n^{1-5\beta/2}]$ such that w.v.h.p. in $G_{n,p}$, for all $z$, we have*

$$\big|\mathbb{P}_S\big(e(S) = z\big) - a^{-1}\mathbb{P}_S\big(e(S) \in [z, z + a]\big)\big| \leq n^{-1-\beta+2\varepsilon}.$$

To see how this fits into our main outline, note that the central limit theorem gives us the measure of any interval with error $O(n^{-1/4+\varepsilon})$. If we then work on an interval of length about $n^{1-5\beta/2}$, the smoothing argument spreads out this error to $O(n^{-5/4+5\beta/2+\varepsilon})$ at each point and introduces an additional $O(n^{-1-\beta+2\varepsilon})$ error. There is also an additional small error from the variation of the Gaussian, and altogether this leads to the claimed local limit theorem when $\beta$ is chosen appropriately.

Lemma 17 will be established using an iterative procedure to gradually prove smoothing statements for intervals of decreasing length (descending to intervals of length 1). In each step, we consider any interval $A$ of length $r$ from the previous iteration, and show that for any two subintervals $A_1$ and $A_2$ of $A$ with the same length $a \ll r$, their weights cannot

differ too much from each other. This is the content of Lemma 19. Combining all these smoothing statements, we obtain that the probabilities at two single points in any interval of length given by the first iteration step do not differ too much, giving Lemma 17. The condition that Lemma 19 imposes on intervals of length $r$ to get an iteration started is rather weak – we ask that weights of those intervals are at most $O(r/n)$. As we will see in Corollary 20, this condition follows directly from the assertion of the previous application of Lemma 19; at the first iteration, it follows from the CLT in the form of Corollary 14.

We shall use an iterative sequence for the above procedure. Each step requires a suitable choice of parameters $a$ and $r$ as above, as well as $t$ from Section 3.1. The conditions of Lemma 19 impose the following constraints. For a given constant $\beta \in (0, \frac{1}{10})$, let us say that the triple $(a, r, t)$ of integers is *valid (for $\beta$)* provided that

(i) $a \leq r$,

(ii) $a = ct^{3/2}n^{2\beta-1/2}$ for some $\frac{1}{2} \leq c \leq 1$, and

(iii) $r \leq n^{-\beta}\sqrt{nt}$.

*Remark* 18. In Lemmas 15 and 16 we already required the $c \leq 1$ condition in (ii) explicitly while the requirement that $t \leq k/2$ follows for large $n$ from the $c \geq \frac{1}{2}$ part of (ii) (as (i)–(iii) imply $ct \leq n^{1-3\beta}$ while $k \geq \delta n$). We also note for later that if $r^3 n^{-2+5\beta} \leq a \leq r$ then one can set $t = \lceil (a^2 n^{1-4\beta})^{1/3} \rceil$ to satisfy (i)–(iii) when $n$ is large.

**Lemma 19** (Difference Lemma). *Let $\beta \in (0, \frac{1}{10})$, $\varepsilon > 0$ and let $(a, r, t)$ be valid for $\beta$. Let $S'$ and $S$ be uniformly and independently chosen random subsets of $[n]$ of sizes $k-t$ and $k$ respectively. Suppose that there is a constant $C$ such that w.v.h.p. in $G_{n,p}$ we have*

$$\mathbb{P}_{S'}\big(e(S') \in [z, z+r)\big) \leq \frac{Cr}{n} \tag{19}$$

*for all $z$. Then w.v.h.p. in $G_{n,p}$ we have*

$$\left|\mathbb{P}_S\big(e(S) \in [z_1, z_1+a)\big) - \mathbb{P}_S\big(e(S) \in [z_2, z_2+a)\big)\right| \leq an^{-1-\beta+\varepsilon} \tag{20}$$

*whenever $|z_1 - z_2| \leq r$.*

*Proof.* We can assume that $n$ is large, so that $(a, r, t)$ being valid for $\beta$ implies $t \leq k/2$. To rewrite the quantity we wish to bound, we couple our choices of $S$ and $S'$: generate a random set $S$ in parts by taking a random set $S'$ of size $k' := k-t$ as well as a random set $T \subseteq [n] \backslash S'$ of size $t$, and then defining $S := S' \cup T$ (as in Section 3.1). Let $A_1 := [z_1, z_1+a)$ and $A_2 := [z_2, z_2 + a)$ and note that $|z_1 - z_2| \leq r$ by assumption. We call $S'$ *good* if (18) holds for both $A_1$ and $A_2$, otherwise $S'$ will be called *bad*. If we condition on $S'$ and $S'$ is good, then

$$\left|\mathbb{P}(e(S) \in A_1 \mid S') - \mathbb{P}(e(S) \in A_2 \mid S')\right|$$
$$\leq \left|\mathbb{P}_T\big(e(S) - e(S') \in A_1 - e(S')\big) - \mathbb{P}_Y\big(Y \in A_1 - e(S')\big)\right|$$
$$+ \left|\mathbb{P}_T\big(e(S) - e(S') \in A_2 - e(S')\big) - \mathbb{P}_Y\big(Y \in A_2 - e(S')\big)\right|$$
$$+ \left|\mathbb{P}_Y\big(Y \in A_1 - e(S')\big) - \mathbb{P}_Y\big(Y \in A_2 - e(S')\big)\right|$$
$$\leq \Delta_{A_1-e(S')} + \Delta_{A_2-e(S')} + 2e^{-n^\varepsilon/2} + P_Y(e(S')),$$

where $Y \sim \mathrm{Bin}(t(k-t) + \binom{t}{2}, p)$ and

$$P_Y(e) := \big|\mathbb{P}(Y \in A_1 - e) - \mathbb{P}(Y \in A_2 - e)\big|.$$

The conclusion of Lemma 16 means that w.v.h.p. in $G_{n,p}$, $\mathbb{P}_{S'}(S' \text{ is bad}) \le 2e^{-n^\varepsilon/4}$, so

$$
\begin{aligned}
\big|\mathbb{P}(e(S) \in A_1) &- \mathbb{P}(e(S) \in A_2)\big| \\
&\le \mathbb{P}(S' \text{ is bad}) + \mathbb{E}_{S'}\big[\Delta_{A_1-e(S')} + \Delta_{A_2-e(S')} + 2e^{-n^\varepsilon/2} + P_Y(e(S'))\big] \\
&\le 4e^{-n^\varepsilon/4} + \mathbb{E}_{S'}\big[\Delta_{A_1-e(S')} + \Delta_{A_2-e(S')} + P_Y(e(S'))\big] \\
&\le 4e^{-n^\varepsilon/4} + \sum_{0 \le e \le \binom{k'}{2}} \mathbb{P}(e(S') = e)\big(\Delta_{A_1-e} + \Delta_{A_2-e} + P_Y(e)\big). \quad (21)
\end{aligned}
$$

Let $I$ be the set of $e$ such that there is a $z \in A_1 \cup A_2$ with

$$|z - e - \mathbb{E}[Y]| \le \sqrt{kt}\, n^\varepsilon. \quad (22)$$

When $e \notin I$, we have exponential bounds for all terms of (21): using Theorem 21 we have for $i = 1, 2$ and sufficiently large $n$

$$
\begin{aligned}
\mathbb{P}(Y \in A_i - e) &\le \mathbb{P}\big(|Y - \mathbb{E}[Y]| > \sqrt{kt}\, n^\varepsilon\big) \\
&\le 2\exp\left(-\frac{ktn^{2\varepsilon}}{2\big(((k-t)t + \binom{t}{2})p(1-p) + \sqrt{kt}\, n^\varepsilon/3\big)}\right) \\
&\le 2\exp\left(-\frac{ktn^{2\varepsilon}}{kt + \sqrt{kt}\, n^\varepsilon}\right) \\
&\le \exp(-n^\varepsilon)/m,
\end{aligned}
$$

where $m = \lfloor \frac{n-k+t}{t} \rfloor$ as in Lemma 15. Thus we are in the first case of (16), and so $\Delta_{A_1-e}, \Delta_{A_2-e} = \exp(-n^\varepsilon)/m$, and also $P_Y(e) \le \exp(-n^\varepsilon)/m$. Hence

$$4e^{-n^\varepsilon/4} + \sum_{e \notin I} \mathbb{P}(e(S') = e)\big(\Delta_{A_1-e} + \Delta_{A_2-e} + P_Y(e)\big) \le 4e^{-n^\varepsilon/4} + 3e^{-n^\varepsilon}/m \le 5e^{-n^\varepsilon/4}.$$

We now turn to the contribution from the $e \in I$. For these $e$, in (16) we have $\Delta_{A_i-e} \le n^{\beta+\varepsilon}/m$. In addition, Lemma 23 gives a bound on $P_Y(e)$, so that by (17) and the fact that $m \ge (n-k)/t \ge \delta n/t$ we have

$$
\begin{aligned}
\Delta_{A_1-e} + \Delta_{A_2-e} + P_Y(e) &\le 2n^{\beta+\varepsilon}/m + \frac{\pi a|z_2 - z_1|}{4\mathrm{Var}(Y)} \le 2\delta^{-1}n^{\beta+\varepsilon-1}t + \frac{\pi ar}{\delta^2 nt} \\
&\le \frac{a}{\sqrt{nt}}\big(4\delta^{-1}n^{-\beta+\varepsilon} + \pi\delta^{-2}n^{-\beta}\big) \le \frac{a}{\sqrt{nt}} \cdot 5\delta^{-1}n^{-\beta+\varepsilon},
\end{aligned}
$$

where we have used conditions (ii) and (iii) and assumed $n$ sufficiently large.

To obtain a bound on the total weight of all $e \in I$, we first observe that, for such $e$, the condition in (22) is in fact satisfied by one of $z_1$, $z_1 + a$, $z_2$ or $z_2 + a$. Indeed, when $e + \mathbb{E}[Y] \notin A_1 \cup A_2$, one of $z_1$, $z_1 + a$, $z_2$, $z_2 + a$ is at least as close to $e + \mathbb{E}[Y]$ as any element of $A_1 \cup A_2$. Else, if $e + \mathbb{E}[Y]$ is an element of, say, $A_1$, the validity of $(a, r, t)$

(conditions (i) and (iii)) implies that $a \leq r \leq \sqrt{nt} \leq \sqrt{kt}\, n^\varepsilon$, so it follows that any $z \in A_1$, and in particular $z_1$, satisfies (22). We thus get

$$\mathbb{P}_{S'}(e(S') \in I) \leq \sum_{z \in \{z_1, z_1+a, z_2, z_2+a\}} \mathbb{P}_{S'}\big(|z - e(S') - \mathbb{E}[Y]| \leq \sqrt{kt}\, n^\varepsilon\big)$$

$$\leq 8C(2\sqrt{kt}\, n^\varepsilon + 1)/n = O(\sqrt{nt} \cdot n^{\varepsilon-1}).$$

For the last inequality, observe that (19) implies that for all $r' \geq r$, $\mathbb{P}(e(S') \subseteq [z, z+r')) \leq 2Cr'/n$ and that this can be applied as $r \leq \sqrt{kt}\, n^\varepsilon$.

Putting all of this together and using that $(a, r, t)$ is valid (conditions (ii) and (iii)) in the penultimate step, we obtain

$$|\mathbb{P}(e(S) \in A_1) - \mathbb{P}(e(S) \in A_2)| \leq 5e^{-n^\varepsilon/4} + \sum_{e \in I} \mathbb{P}(e(S') = e)\big(\Delta_{A_1-e} + \Delta_{A_2-e} + P_Y(e)\big)$$

$$= 5e^{-n^\varepsilon/4} + O(\sqrt{nt} \cdot n^{\varepsilon-1}) \cdot \tfrac{a}{\sqrt{nt}} \cdot 5\delta^{-1}n^{-\beta+\varepsilon}$$

$$= O(an^{-\beta-1+2\varepsilon}).$$

As this holds for any $\varepsilon > 0$, the result follows for sufficiently large $n$ on replacing $\varepsilon$ by $\varepsilon/3$, say. $\qquad\square$

We now observe that the conclusion of Lemma 19 yields (19) for the smaller interval of size $a$, which will allow us to kick off the next iteration step.

**Corollary 20.** *Under the notation and assumptions of Lemma 19, there is a constant $C' > 0$ such that w.v.h.p. in $G_{n,p}$ we also have $\mathbb{P}\big(e(S) \in [z, z+a)\big) \leq \frac{C'a}{n}$ for all $z$.*

*Proof.* We look at $[z, z+a)$ for any fixed $z$. Consider the interval $[z, z+r)$, and let $\mathcal{I}$ be a maximal collection of disjoint intervals of length $a$ in $[z, z+r)$, so that $|\mathcal{I}| = \lfloor r/a \rfloor \geq r/2a$. W.v.h.p., for each $A \in \mathcal{I}$, we have by Lemma 19 that $|\mathbb{P}(e(S) \in [z, z+a)) - \mathbb{P}(e(S) \in A)| \leq a/n$ (as without loss of generality $\varepsilon < \beta$). Thus this assertion and $\mathbb{P}(e(S') \in [z, z+r)) \leq \frac{Cr}{n}$ hold at the same time w.v.h.p. and they imply $\mathbb{P}(e(S) \in [z, z+a)) \leq \frac{C'a}{n}$ for $C' = 2C + 1$. Indeed, supposing otherwise gives

$$\mathbb{P}\big(e(S) \in [z, z+r)\big) \geq \sum_{A \in \mathcal{I}} \mathbb{P}(e(S) \in A) > \big(\tfrac{C'a}{n} - \tfrac{a}{n}\big) \cdot \tfrac{r}{2a} = \tfrac{Cr}{n},$$

which contradicts our assumption. $\qquad\square$

We are now ready to prove the main smoothing lemma.

*Proof of Lemma 17.* Set $a_0 = 1$ and inductively define for $j \geq 0$, $t_j = \lceil (a_j^2 n^{1-4\beta})^{1/3} \rceil$ and $a_{j+1} = \lfloor (a_j n^{2-5\beta})^{1/3} \rfloor$. It can be easily verified that for all $j$, $(a_j, a_{j+1}, t_j)$ is valid for $\beta$. Indeed, $a_j$ will be (just less than) $n^{(1-5\beta/2)(1-3^{-j})}$, so $a_j \leq a_{j+1}$. Also $n^{-\beta}\sqrt{nt_j} \geq (a_j n^{2-5\beta})^{1/3} \geq a_{j+1}$, and $a_j \leq t_j^{3/2} n^{2\beta-1/2} = (1 + o(1))a_j$. Also, there exists some $j_0$, depending only on $\varepsilon$, such that for sufficiently large $n$, $a_{j_0} \in [n^{1-5\beta/2-\varepsilon}, n^{1-5\beta/2}]$. This will be our $a$.

Now let $T_j = \sum_{i=0}^{j-1} t_i$ and note that as $j_0$ is a constant and $t_j \leq \lceil n^{1-3\beta} \rceil$, we have that for sufficiently large $n$ that $T_{j_0+1} \leq k/2$. We now prove by reverse induction on $j$ that there is a constant $C_j$ such that for any $k' \in [k - T_j, k]$, if we choose a set $S$ of size $k'$ uniformly at random, then w.v.h.p. in $G_{n,p}$ we have for all $z$,

$$\mathbb{P}_S\big(e(S) \in [z, z + a_j)\big) \leq C_j a_j / n \tag{23}$$

and (for $j < j_0$)

$$\left| a_j^{-1} \mathbb{P}_S\big(e(S) \in [z, z + a_j)\big) - a_{j+1}^{-1} \mathbb{P}_S\big(e(S) \in [z, z + a_{j+1})\big) \right| \leq n^{-1-\beta+\varepsilon}. \tag{24}$$

Indeed (23) holds for $j = j_0$ by Theorem 1 provided $a_{j_0} \geq n^{3/4+\varepsilon'}$ for some $\varepsilon' > 0$, and this holds as $1 - 5\beta/2 - \varepsilon > 3/4$. Then for $j = j_0 - 1, \ldots, 0$ in turn, if $S$ is a uniformly chosen set of size $k' \in [k - T_j, k]$ and $S'$ is a uniformly chosen set of size $k' - t_j \in [k - T_{j+1}, k]$, then (23) applied with $S$ replaced with $S'$ and $j$ replaced with $j + 1$ implies (24) by Lemma 19. Corollary 20 then implies (23) for $S$ and $j$.

Hence we have for all $z$, and $S$ uniformly random of size $k$,

$$\left| \mathbb{P}\big(e(S) = z\big) - a_{j_0}^{-1} \mathbb{P}\big(e(S) \in [z, z + a_{j_0})\big) \right| \leq j_0 n^{-1-\beta+\varepsilon} \leq n^{-1-\beta+2\varepsilon}. \qquad \square$$

## 3.3 Proof of Theorem 2

Combining the control we have from Lemma 17, which relates single-point weights to weights of long intervals, with Corollary 13 which gives us the weight of long intervals, we can now obtain the single-point weights up to a small error term as desired.

*Proof of Theorem 2.* Lemma 17 states that w.v.h.p. $G_{n,p}$ satisfies

$$\left| \mathbb{P}\big(e(S) = z\big) - a^{-1} \mathbb{P}\big(e(S) \in [z, z + a)\big) \right| \leq n^{-1-\beta+\varepsilon},$$

for all $z$, where $a$ is some integer satisfying

$$n^{1-5\beta/2-\varepsilon} \leq a \leq n^{1-5\beta/2}.$$

Setting $p = M/N$, there is a polynomial probability (on the order of $1/n$) that $e(G_{n,p}) = M$. In that case we have $(G_{n,p} \mid e(G_{n,p}) = M) \sim G_{n,M}$, so the conclusion of Lemma 17 transfers to $G_{n,M}$ again with very small failure probability. Using this in the first step and Corollary 13 in the second, and with $Z$ and $\phi$ as in the theorem statement, we obtain

$$\begin{aligned}
\mathbb{P}\big(e(S) = z\big) &= a^{-1} \mathbb{P}\big(e(S) \in [z, z + a)\big) + O(n^{-1-\beta+\varepsilon}) \\
&= a^{-1} \mathbb{P}\big(Z \in [z, z + a)\big) + O(a^{-1} n^{-1/4+\varepsilon} + n^{-1-\beta+\varepsilon}) \\
&= a^{-1} \int_z^{z+a} \varphi(t)\, dt + O(a^{-1} n^{-1/4+\varepsilon} + n^{-1-\beta+\varepsilon}) \\
&= \varphi(z) + O(a/n^2 + a^{-1} n^{-1/4+\varepsilon} + n^{-1-\beta+\varepsilon}) \\
&= \varphi(z) + O(n^{-1-5\beta/2} + n^{-5/4+5\beta/2+2\varepsilon} + n^{-1-\beta+\varepsilon}),
\end{aligned}$$

where the penultimate step follows from Lemma 25 and the fact that $\mathrm{Var}(Z) = \Theta(n^2)$. To optimise the exponent, we take $\beta = 1/14$, giving

$$\mathbb{P}(e(S) = z) = \varphi(z) + O(n^{-1-1/14+2\varepsilon}). \qquad \square$$

# References

[1] S. N. Bernstein. On a modification of Chebyshev's inequality and of the error formula of Laplace. *Ann. Sci. Inst. Sav. Ukraine, Sect. Math.*, **4**:38–48, 1924.

[2] B. B. Bhattacharya, P. Diaconis, and S. Mukherjee. Universal limit theorems in graph coloring problems with connections to extremal combinatorics. *The Annals of Applied Probability*, pages 337–394, 2017.

[3] E. Bolthausen and F. Götze. The rate of convergence for multivariate sampling statistics. *The Annals of Statistics*, **21**:1692 – 1710, 1993.

[4] N. Calkin, A. Frieze, and B. D. McKay. On subgraph sizes in random graphs. *Combinatorics, Probability and Computing*, 1(2):123–134, 1992.

[5] S. Chatterjee. A new method of normal approximation. *The Annals of Probability*, pages 1584–1610, 2008.

[6] P. De Jong. A central limit theorem for generalized quadratic forms. *Probability Theory and Related Fields*, 75(2):261–277, 1987.

[7] P. Erdős. Some of my favourite problems in various branches of combinatorics. *Le Matematiche*, 47(2):231–240, 1992.

[8] P. Erdős and A. Szemerédi. On a Ramsey type theorem. *Periodica Mathematica Hungarica*, 2(1-4):295–299, 1972.

[9] P. Hall. Central limit theorem for integrated square error of multivariate nonparametric density estimators. *Journal of multivariate analysis*, 14(1):1–16, 1984.

[10] M. Kwan, A. Sah, L. Sauermann, and M. Sawhney. Anticoncentration in Ramsey graphs and a proof of the Erdős–McKay conjecture. In *Forum of Mathematics, Pi*, volume 11, page e21, 2023.

[11] G. Reinert and A. Rollin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *The Annals of Probability*, **37**:2150–2173, 2009.

[12] V. Rotar'. Some limit theorems for polynomials of second degree. *Theory of Probability & Its Applications*, 18(3):499–507, 1974.

[13] T. Tao and V. Vu. Random matrices: universality of local spectral statistics of non-Hermitian matrices. *The Annals of Probability*, **43**(2):782–874, 2015.

# A    Technical tools

In this section, we record several technical results that will be used in proofs of our main theorems. While none of them are new, we provide proofs for completeness in cases where a convenient reference is lacking.

We start with two concentration inequalities, namely Bernstein's inequality and a version of the Azuma–Hoeffding inequality with a relaxed difference bound condition.

**Theorem 21** (Bernstein [1])**.** *Let $X_1, \ldots, X_n$ be independent random variables taking values in $[0,1]$ and let $X := \sum_{i=1}^{n} X_i$ and $\sigma^2 := \mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$. Then for any $t \geq 0$ we have*

$$\mathbb{P}\big(|X - \mathbb{E}[X]| > t\big) \leq 2 \exp\Big(-\frac{t^2}{2(\sigma^2 + t/3)}\Big).$$

**Lemma 22** (Azuma–Hoeffding, see e.g. [13], Proposition 34)**.** *Let $X = X(Z_1, \ldots, Z_n)$ be a random variable depending on independent variables $Z_1, \ldots, Z_n$ and suppose $c_1, \ldots, c_n > 0$ are such that*

$$\mathbb{P}\big(\exists i \colon \big|\mathbb{E}[X \mid Z_1, \ldots Z_i] - \mathbb{E}[X \mid Z_1, \ldots Z_{i-1}]\big| > c_i\big) \leq \varepsilon.$$

*Then*

$$\mathbb{P}\big(|X - \mathbb{E}[X]| \geq t\big) \leq \varepsilon + 2 \exp\big(-t^2/2\sigma^2\big),$$

*where $\sigma^2 := \sum_{i=1}^{n} c_i^2$.*

The following is more general that we need as we will only be applying it in the case when $X$ is a Binomial random variable. Nevertheless, we state it here in greater generality as the proof is no harder.

**Lemma 23.** *Suppose $X = \sum_{i=1}^{n} X_i$ where $X_1, \ldots, X_n$ are independent Bernoulli random variables. Then, for any $m, m' \in \mathbb{Z}$,*

$$\mathbb{P}(X = m) \leq \sqrt{\frac{\pi}{8\mathrm{Var}(X)}} \quad and \quad \big|\mathbb{P}(X = m') - \mathbb{P}(X = m)\big| \leq \frac{\pi}{4\mathrm{Var}(X)}|m' - m|.$$

*Proof.* For $p \in [0,1]$ and $\theta \in [-\pi, \pi]$ we have the inequality

$$\big|(1-p) + pe^{i\theta}\big|^2 = (1-p)^2 + p^2 + 2p(1-p)\cos\theta = 1 - 4p(1-p)\sin^2\tfrac{\theta}{2} \leq e^{-4p(1-p)\theta^2/\pi^2},$$

where we used $|\sin\phi| \geq \frac{2}{\pi}|\phi|$ for $|\phi| \leq \frac{\pi}{2}$ and $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$. Now, as the $X_i$ are independent,

$$\mathbb{P}(X = m) = \mathbb{E}[\mathbb{1}_{\{X=m\}}] = \mathbb{E}\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\theta(X-m)}\, d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-im\theta} \prod_j \mathbb{E}e^{i\theta X_j}\, d\theta.$$

Hence, writing $p_j := \mathbb{P}(X_j = 1)$ and $\sigma^2 := \mathrm{Var}(X) = \sum p_j(1 - p_j)$,

$$\mathbb{P}(X = m) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-im\theta} \prod_j \big((1-p_j) + p_j e^{i\theta}\big)\, d\theta \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \prod_j e^{-2p_j(1-p_j)\theta^2/\pi^2}\, d\theta$$

$$\leq \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-2\sigma^2\theta^2/\pi^2}\, d\theta = \frac{1}{2\pi}\cdot\sqrt{\frac{\pi}{2\sigma^2/\pi^2}} = \sqrt{\frac{\pi}{8\sigma^2}}.$$

For the second inequality, we have

$$\mathbb{P}(X = m') - \mathbb{P}(X = m) = \frac{1}{2\pi}\int_{-\pi}^{\pi} (e^{-im'\theta} - e^{-im\theta}) \prod_j \mathbb{E}e^{i\theta X_j}\, d\theta.$$

Thus, as $|e^{-im'\theta} - e^{-im\theta}| \leq |(m' - m)\theta|$,

$$|\mathbb{P}(X = m') - \mathbb{P}(X = m)| \leq \frac{|m' - m|}{2\pi} \int_{-\pi}^{\pi} |\theta| \prod_j e^{-2p_j(1-p_j)\theta^2/\pi^2} \, d\theta$$

$$\leq \frac{|m' - m|}{\pi} \int_0^{\infty} \theta e^{-2\sigma^2\theta^2/\pi^2} \, d\theta$$

$$= \frac{|m' - m|}{\pi} \cdot \frac{\pi^2}{4\sigma^2} = \frac{\pi}{4\sigma^2} |m' - m|,$$

as desired. □

*Remark* 24. The constants in Lemma 23 are not best possible, however it can be seen by taking approximations to suitable Poisson distributions that we can't replace them with the corresponding constants for the normal distribution given by the following.

**Lemma 25.** *If $\varphi$ is a density function of a normal random variable with mean $\mu$ and variance $\sigma^2$ then, for all $x, y \in \mathbb{R}$ we have*

$$\varphi(x) \leq \frac{1}{\sigma\sqrt{2\pi}} \qquad and \qquad |\varphi(x) - \varphi(y)| \leq \frac{|x - y|}{\sigma^2\sqrt{2\pi e}}.$$

*Proof.* The first statement is immediate as $\varphi(x) = e^{-(x-\mu)^2/2\sigma^2}/(\sigma\sqrt{2\pi})$. For the latter, again note that the maximum of $x \mapsto |xe^{-x^2/2}|$ is $e^{-1/2}$ and occurs at $x = \pm 1$ and hence

$$|\varphi'(x)| = \left| -\frac{1}{\sigma^2\sqrt{2\pi}} \frac{x - \mu}{\sigma} \exp\left( -\left(\frac{x - \mu}{\sigma}\right)^2/2 \right) \right| \leq \frac{1}{\sigma^2\sqrt{2\pi e}},$$

with the maximum occurring at $(x - \mu)/\sigma = \pm 1$. The statement then follows from the mean value theorem. □