Combinatorics, Probability and Computing (2011) 20, 413–433. © Cambridge University Press 2011
doi:10.1017/S0963548311000034

Subgraphs of Dense Random Graphs with Specified Degrees

BRENDAN D. McKAY†

School of Computer Science, Australian National University, Canberra ACT 0200, Australia
(e-mail: bdm@cs.anu.edu.au)

Received 27 January 2010; revised 27 November 2010; first published online 27 January 2011

Let $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$ be a vector of non-negative integers with even sum. We prove some basic facts about the structure of a random graph with degree sequence $\boldsymbol{d}$, including the probability of a given subgraph or induced subgraph.

Although there are many results of this kind, they are restricted to the sparse case with only a few exceptions. Our focus is instead on the case where the average degree is approximately a constant fraction of n.

Our approach is the multidimensional saddle-point method. This extends the enumerative work of McKay and Wormald (1990) and is analogous to the theory developed for bipartite graphs by Greenhill and McKay (2009).

1. Introduction

Let $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$ be a vector of non-negative integers with even sum. Let $\boldsymbol{X} = (x_{jk})$ be a symmetric $n \times n$ matrix over $\{0,1\}$ with zero diagonal. Define $G(\boldsymbol{d}, \boldsymbol{X})$ to be the number of $n \times n$ symmetric matrices $\mathcal{A} = (a_{jk})$ over $\{0,1\}$ with zero diagonal, such that

(i) row j sums to d_j, for $1 \leqslant j \leqslant n$;
(ii) $a_{jk} = 0$ whenever $x_{jk} = 1$, for $1 \leqslant j,k \leqslant n$.

Equivalently, $G(\boldsymbol{d}, \boldsymbol{X})$ is the number of labelled simple graphs with n vertices of degree $d_1, d_2, \ldots, d_n$, having no edges in common with the simple graph $\boldsymbol{X}$. The special case where $\boldsymbol{X}$ is the zero matrix $\boldsymbol{0}$ will also be denoted $G(\boldsymbol{d})$. Define $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, where x_j is the sum of the jth row of $\boldsymbol{X}$.

One motive for interest in $G(\boldsymbol{d}, \boldsymbol{X})$ is that the ratio $G(\boldsymbol{d}, \boldsymbol{X})/G(\boldsymbol{d})$ is the probability that a random simple graph with degree sequence $\boldsymbol{d}$ has no edge in common with $\boldsymbol{X}$. Similarly, $G(\boldsymbol{d} - \boldsymbol{x}, \boldsymbol{X})/G(\boldsymbol{d})$ is the probability that $\boldsymbol{X}$ appears as a subgraph. In these cases, and throughout the paper, probability spaces have the uniform distribution.

† Research supported by the Australian Research Council.

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

Define the matrix $\bar{\boldsymbol{X}} = (\bar{x}_{jk})$ over $\{0,1\}$ with $\bar{x}_{jk} = 1$ if and only if $j \neq k$ and $x_{jk} = 0$. For convenience we will adopt the convention that $\sum_{jk\in\boldsymbol{X}}$ means the sum over all $\{j,k\}$ such that $x_{jk} = 1$, and similarly $\sum_{jk\in\bar{\boldsymbol{X}}}$ means the sum over all $\{j,k\}$ such that $\bar{x}_{jk} = 1$. Note that the equal sets $\{j,k\}$ and $\{k,j\}$ do not appear as separate terms in these sums.

Define the following key parameters:

$$\begin{aligned}
E &= \tfrac{1}{2}\sum_{j=1}^{n} d_j && \text{(the number of edges)},\\
d &= \frac{2E}{n} && \text{(the average degree)},\\
\lambda &= \frac{d}{n-1} && \text{(the density ignoring the diagonal)},\\
A &= \tfrac{1}{2}\lambda(1-\lambda), &&\\
X &= \tfrac{1}{2}\sum_{j=1}^{n} x_j && \text{(the number of edges of } \boldsymbol{X}),\\
\delta_j &= d_j - d + \lambda x_j && (1 \leqslant j \leqslant n).
\end{aligned}$$

Direct asymptotic estimation of $G(\boldsymbol{d},\boldsymbol{X})$ for non-zero $\boldsymbol{X}$ has been previously restricted to the sparse range. For representative results with bounded or very slowly growing degrees, see Bollobás and McKay [3] and Wormald [21]. For somewhat higher degrees we have the following. Let $d_{\max} = \max_j d_j$, $x_{\max} = \max_j x_j$ and $\Delta = d_{\max}(d_{\max} + x_{\max})$.

Theorem 1.1 ([15]). *Suppose $d_{\max} \geqslant 1$ and $\Delta = o(E)$. Then, as $n \to \infty$,*

$$G(\boldsymbol{d},\boldsymbol{X}) = \frac{(2E)!}{E!\,2^E \prod_{j=1}^{n} d_j!}\exp\biggl(-\frac{\sum_{j=1}^{n} d_j(d_j-1)}{4E} - \frac{\bigl(\sum_{j=1}^{n} d_j(d_j-1)\bigr)^2}{16E^2} - \frac{\sum_{jk\in\boldsymbol{X}} d_j d_k}{2E} + O(\Delta^2/E)\biggr).$$

The error term in Theorem 1.1 is $o(1)$ only under the stronger condition that $\Delta^2 = o(E)$, which implies that the graphs are quite sparse. The special case $G(\boldsymbol{d})$ was determined by McKay and Wormald [20] under the weaker condition $d_{\max}^3 = o(E)$.

The probability $G(\boldsymbol{d},\boldsymbol{X})/G(\boldsymbol{d})$ of being edge-disjoint from $\boldsymbol{X}$, and the probability $G(\boldsymbol{d}-\boldsymbol{x},\boldsymbol{X})/G(\boldsymbol{d})$ of containing $\boldsymbol{X}$ as a subgraph are easily deduced from Theorem 1.1. They can also be found directly over a sometimes wider range of $\boldsymbol{d}$ values. Let $(a)_b$ denote the falling factorial. The following is a consequence of Theorems 2.9 and 2.10 of McKay [14].

Theorem 1.2 ([14]). *If $\Delta + X = o(E)$ then, as $n \to \infty$,*

$$\frac{G(\boldsymbol{d}-\boldsymbol{x},\boldsymbol{X})}{G(\boldsymbol{d})} = \frac{\prod_{j=1}^{n}(d_j)_{x_j}}{2^X (E)_X}\exp\bigl(O(\Delta X/E)\bigr).$$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

In the case of dense matrices, the first asymptotically precise enumeration result was that of McKay and Wormald [19], who proved Theorem 1.3 (below) in the case of $\boldsymbol{X} = \boldsymbol{0}$ with a slightly weaker error term. This has been extended by Barvinok and Hartigan [1]. They identify a symmetric matrix (λ_{jk}) over $\{0,1\}$ introduced in [19] (and used in greater generality in Section 3) as the matrix that maximizes a certain entropy function. They then express the asymptotic value of $G(\boldsymbol{d})$ as an effectively computable function of (λ_{jk}) provided the values $\{\lambda_{jk}\}_{j\neq k}$ are uniformly bounded away from 0 and 1. This forces the average degree to be $\Theta(n)$ but allows a much greater variation of degrees than we allow. They also show that (λ_{jk}) matches a typical graph with degree sequence $\boldsymbol{d}$ in a sense that we will describe in Section 2. This theme is explored in a somewhat different way by Chatterjee, Diaconis and Sly [5].

Despite the absence of precise enumerative results for densities between $o(n^{-1/2})$ and $c/\log n$, Krivelevich, Sudakov, Vu and Wormald [12] determined several almost-sure properties of random regular graphs over various ranges of density using a combination of switchings and analysis. Other such properties were determined by Boldi and Vigna [2], and Cooper, Frieze, Reed and Riordan [6, 7].

More recently, Krivelevich, Sudakov and Wormald [11] determined the probability of small induced subgraphs in random regular graphs of degree $(n-1)/2$ under some conditions on the order and degree sequence of the subgraph.

The corresponding problems for bipartite graphs and digraphs were studied by Greenhill and McKay [8]; see that paper for a bibliography. The proof method in [8] is quite similar to that here.

We will also have need of the following additional parameters, for $1 \leqslant j \leqslant n$ and $\ell, m \geqslant 1$.

$$B_j = \sum_{k\mid jk\in\boldsymbol{X}} \delta_k, \qquad R = \sum_{j=1}^{n} (d_j - d)^2,$$

$$R_\ell = \sum_{j=1}^{n} \delta_j^\ell, \qquad X_\ell = \sum_{j=1}^{n} x_j^\ell,$$

$$D = \sum_{jk\in\boldsymbol{X}} \delta_j\delta_k, \qquad H = \sum_{jk\in\boldsymbol{X}} x_j x_k,$$

$$L = \sum_{jk\in\boldsymbol{X}} (\delta_j - x_j)(\delta_k - x_k), \qquad C_{\ell,m} = \sum_{j=1}^{n} \delta_j^\ell x_j^m,$$

$$K = \sum_{jk\in\boldsymbol{X}} (d_j - d)(d_k - d).$$

To calibrate and motivate our main enumeration result, we first develop a naïve estimate of $G(\boldsymbol{d}, \boldsymbol{X})$ by extending an idea introduced in [19]. Generate a random graph by independently creating an edge jk with probability λ for each $jk \in \bar{\boldsymbol{X}}$. Each graph with E edges (none in common with $\boldsymbol{X}$) appears with probability $\lambda^E(1-\lambda)^{\binom{n}{2}-X-E}$. Moreover, the event E_j that vertex j has degree d_j has probability $\binom{n-1-x_j}{d_j}\lambda^{d_j}(1-\lambda)^{n-1-x_j-d_j}$ for each j. If we (incorrectly) assume that the events $E_1, \ldots, E_n$ are independent, we obtain a guess

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

for $G(\boldsymbol{d},\boldsymbol{X})$ as follows:

$$\widehat{G}(\boldsymbol{d},\boldsymbol{X}) = (1-\lambda)^{-X}\left(\lambda^{\lambda}(1-\lambda)^{1-\lambda}\right)^{\binom{n}{2}}\prod_{j=1}^{n}\binom{n-1-x_j}{d_j}. \tag{1.1}$$

In [19] it was proved that

$$G(\boldsymbol{d}) = \sqrt{2}\,\widehat{G}(\boldsymbol{d},\boldsymbol{0})\exp\left(\frac{1}{4} - \frac{R^2}{16A^2n^4} + o(1)\right)$$

under certain conditions on $\boldsymbol{d}$. Our main result extends this to non-zero $\boldsymbol{X}$.

Theorem 1.3. *Let $a,b>0$ be constants such that $a+b<\frac{1}{2}$. Then there is a constant $\varepsilon=\varepsilon(a,b)>0$ such that the following holds. Suppose that $d_j-d, x_j = O(n^{1/2+\varepsilon})$ uniformly for $1\leqslant j\leqslant n$, that $X=O(n^{1+2\varepsilon})$, and that*

$$\min\{d, n-d-1\} \geqslant \frac{n}{3a\log n},$$

for sufficiently large n. Then, as $n\to\infty$,

$$G(\boldsymbol{d},\boldsymbol{X}) = \sqrt{2}\,\widehat{G}(\boldsymbol{d},\boldsymbol{X})\exp\left(\frac{1}{4} - \frac{R^2}{16A^2n^4} + \frac{\lambda X^2}{(1-\lambda)n^2} - \frac{D}{2An^2} + O(n^{-b})\right).$$

The proof of Theorem 1.3 is the main task of the paper. Here we will summarize the main phases. The basic idea is to identify $G(\boldsymbol{d},\boldsymbol{X})$ as a coefficient in a multivariable generating function and to extract that coefficient using the saddle-point method. In Section 3, we write $G(\boldsymbol{d},\boldsymbol{X}) = P(\boldsymbol{d},\boldsymbol{X})I(\boldsymbol{d},\boldsymbol{X})$, where $P(\boldsymbol{d},\boldsymbol{X})$ is a rational expression and $I(\boldsymbol{d},\boldsymbol{X})$ is an integral in n complex dimensions. Both depend on the location of the saddle point, which is the solution of some nonlinear equations. Those equations are solved in Section 3.1, and this leads to the value of $P(\boldsymbol{d},\boldsymbol{X})$ in (3.23). In Sections 3.3–3.5, the integral $I(\boldsymbol{d},\boldsymbol{X})$ is estimated in a superset of a small region $\mathcal{R}$ enclosing the origin (defined in (3.25)) and an equivalent small region $\mathcal{R}'$ enclosing $(\pi,\dots,\pi)$. The result is given by Lemma 3.1. Finally, in Section 3.6, we note that the integral restricted to the exterior of $\mathcal{R}\cup\mathcal{R}'$ is negligible.

The proof in various places requires ε to be sufficiently small, but there are only a finite number of such places so we can choose $\varepsilon(a,b)$ to satisfy all of them at once. Note that the theorem remains true if $\varepsilon(a,b)$ is decreased, since the conditions become stronger.

Throughout the paper, the asymptotic notation $O(f(n))$ refers to the passage of n to ∞. We also use a modified notation $\widetilde{O}(f(n))$, which is to be taken as a shorthand for any expression of the form $O(f(n)n^{c\varepsilon})$ with c a numerical constant (perhaps a different constant at each occurrence). Under the assumptions of Theorem 1.3, we have $\lambda^{-1},(1-\lambda)^{-1} = O(\log n)$. This implies, if c_1,c_2,c_3,c_4 are constants, that $\lambda^{c_1}(1-\lambda)^{c_2}n^{c_3+c_4\varepsilon} = \widetilde{O}(n^{c_3})$.

2. Subgraph probabilities

In this section we will derive some important corollaries of Theorem 1.3. The proof of that theorem will be presented in the following section.

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

Define functions $\mathrm{miss}(\boldsymbol{d}, \boldsymbol{X})$ and $\mathrm{hit}(\boldsymbol{d}, \boldsymbol{X})$ such that the probability that a random simple graph with degrees $\boldsymbol{d}$ has no edges in common with $\boldsymbol{X}$ is

$$(1-\lambda)^X \,\mathrm{miss}(\boldsymbol{d}, \boldsymbol{X}),$$

and the probability that it includes $\boldsymbol{X}$ as a subgraph is

$$\lambda^X \,\mathrm{hit}(\boldsymbol{d}, \boldsymbol{X}).$$

In this section, we apply Theorem 1.3 to estimate these probabilities. To avoid unnecessary messiness regarding the value of ε, we will suppose $\varepsilon(a,b)$ in Theorem 1.3 is chosen to be small enough to satisfy the finite number of additional places in this section where a statement is only true if ε is small enough.

Theorem 2.1. *Under the conditions of Theorem 1.3, we have*

$$\mathrm{miss}(\boldsymbol{d}, \boldsymbol{X}) = \exp\Bigg(\frac{\lambda X}{(1-\lambda)n} + \frac{\lambda X_2}{2(1-\lambda)n} + \frac{\lambda(1-2\lambda)X_3}{6(1-\lambda)^2n^2} + \frac{\lambda X^2}{(1-\lambda)n^2} - \frac{D}{\lambda(1-\lambda)n^2} - \frac{C_{1,1}}{(1-\lambda)n} - \frac{(1-2\lambda)C_{1,2}}{2(1-\lambda)^2n^2} - \frac{C_{2,1}}{2(1-\lambda)^2n^2} + O(n^{-b}) \Bigg)$$

and

$$\mathrm{hit}(\boldsymbol{d}, \boldsymbol{X}) = \exp\Bigg(\frac{(1-\lambda)X}{\lambda n} - \frac{(1+\lambda)X_2}{2\lambda n} - \frac{(1+\lambda)(1+2\lambda)X_3}{6\lambda^2n^2} + \frac{(1-\lambda)X^2}{\lambda n^2} - \frac{L}{\lambda(1-\lambda)n^2} + \frac{C_{1,1}}{\lambda n} + \frac{(1+2\lambda)C_{1,2}}{2\lambda^2n^2} - \frac{C_{2,1}}{2\lambda^2n^2} + O(n^{-b}) \Bigg).$$

Proof. Since $(1-\lambda)^X \,\mathrm{miss}(\boldsymbol{d}, \boldsymbol{X}) = G(\boldsymbol{d}, \boldsymbol{X})/G(\boldsymbol{d})$, the first part can be obtained from Theorem 1.3. The second part can be found in similar fashion, or by noting that the probability of a random graph avoiding $\boldsymbol{X}$ is the probability of the complement of the graph having $\boldsymbol{X}$ as a subgraph. □

If $\boldsymbol{X}$ is not too dense, the probabilities in Theorem 2.1 asymptotically match those for an ordinary random graph with edge probability λ. Sufficient conditions are that $\mathrm{miss}(\boldsymbol{d}, \boldsymbol{X}) = 1 + o(1)$ if

$$\lambda X_2 + X \max_j |d_j - d| = o\big((1-\lambda)n\big),$$

and $\mathrm{hit}(\boldsymbol{d}, \boldsymbol{X}) = 1 + o(1)$ if

$$(1-\lambda)X_2 + X \max_j |d_j - d| = o\big(\lambda n\big).$$

Both these sufficient conditions hold, for example, if $X = O(n^{1/2-2\varepsilon})$, or if $d_j - d$ and x_j are uniformly $O(n^{\varepsilon})$ for $1 \leqslant j \leqslant n$ and $X = O(n^{1-2\varepsilon})$.

Since Theorem 2.1 is rather complex, we give some special cases to facilitate its application. We also give the value of $\mathrm{num}(\boldsymbol{d}, \boldsymbol{X})$, which is the exponential factor in Theorem 1.3.

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

Corollary 2.2. *Suppose the conditions of Theorem 1.3 hold, and in addition assume that* $d_1 = \cdots = d_n = d$. *Then*

$$\mathrm{num}(\boldsymbol{d}, \boldsymbol{X}) = \exp\left(\frac{1}{4} + \frac{\lambda(X^2-H)}{(1-\lambda)n^2} + O(n^{-b})\right),$$

$$\mathrm{miss}(\boldsymbol{d}, \boldsymbol{X}) = \exp\left(\frac{\lambda X}{(1-\lambda)n} - \frac{\lambda X_2}{2(1-\lambda)n} - \frac{\lambda(2-\lambda)X_3}{6(1-\lambda)^2n^2} + \frac{\lambda X^2}{(1-\lambda)n^2} - \frac{\lambda H}{(1-\lambda)n^2} + O(n^{-b})\right),$$

$$\mathrm{hit}(\boldsymbol{d}, \boldsymbol{X}) = \exp\left(\frac{(1-\lambda)X}{\lambda n} - \frac{(1-\lambda)X_2}{2\lambda n} - \frac{(1-\lambda^2)X_3}{6\lambda^2n^2} + \frac{(1-\lambda)X^2}{\lambda n^2} - \frac{(1-\lambda)H}{\lambda n^2} + O(n^{-b})\right).$$

Corollary 2.3. *Suppose the conditions of Theorem 1.3 hold, and in addition assume that* $x_1 = \cdots = x_n = x$ (*which implies that* $x = O(n^{2\varepsilon})$). *Then*

$$\mathrm{num}(\boldsymbol{d}, \boldsymbol{X}) = \exp\left(\frac{1}{4} + \frac{\lambda x^2}{4(1-\lambda)} - \frac{K}{2An^2} - \frac{R^2}{16A^2n^4} + O(n^{-b})\right),$$

$$\mathrm{miss}(\boldsymbol{d}, \boldsymbol{X}) = \exp\left(-\frac{\lambda x(x-2)}{4(1-\lambda)} - \frac{xR}{2(1-\lambda)^2n^2} - \frac{K}{2An^2} + O(n^{-b})\right),$$

$$\mathrm{hit}(\boldsymbol{d}, \boldsymbol{X}) = \exp\left(-\frac{(1-\lambda)x(x-2)}{4\lambda} - \frac{xR}{2\lambda^2n^2} - \frac{K}{2An^2} + O(n^{-b})\right).$$

The two parts of Theorem 2.1 have a common generalization. Let $\boldsymbol{Y}$ be a supergraph of $\boldsymbol{X}$. Then the probability that a random graph with degrees $\boldsymbol{d}$ has intersection with $\boldsymbol{Y}$ equal to $\boldsymbol{X}$ is

$$\frac{G(\boldsymbol{d}-\boldsymbol{x}, \boldsymbol{Y})}{G(\boldsymbol{d})}.$$

If the degrees of $\boldsymbol{Y}$ are $y_1, \ldots, y_n$, with $y_j = O(n^{1/2+\varepsilon})$ uniformly over j, and $\sum_{j=1}^n y_j = O(n^{1+2\varepsilon})$, then this probability can be computed using two applications of Theorem 1.3. The resulting general formula is rather complex, so we will be content with presenting the special case where $\boldsymbol{Y}$ consists of a single clique and otherwise isolated vertices. This is the important case of an induced subgraph.

Suppose that for some m, we have $x_{m+1} = \cdots = x_n = 0$. Let $\boldsymbol{X}^{[m]}$ be the subgraph of $\boldsymbol{X}$ induced by vertices $1, \ldots, m$ (so $\boldsymbol{X}^{[m]}$ has the same edges as $\boldsymbol{X}$). For $k, \ell \geqslant 0$, define the quantity

$$\omega_{k,\ell} = \sum_{j=1}^{m} (d_j - d)^k (x_j - \lambda(m-1))^\ell.$$

Theorem 2.4. *Assume the conditions of Theorem 1.3 and in addition that* $m = O(n^{1/2+\varepsilon})$ *and* $x_{m+1} = \cdots = x_n = 0$. *Then the probability that a random graph with degree sequence* $\boldsymbol{d}$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

has $\boldsymbol{X}^{[m]}$ *as an induced subgraph is*

$$\lambda^X(1-\lambda)^{\binom{m}{2}-X} \times \exp\left(\frac{2\omega_{1,1}-\omega_{0,2}}{4An}+\frac{m^2}{2n}+\frac{(1-2\lambda)\omega_{0,1}}{4An}+\frac{4\omega_{1,0}\omega_{0,1}-\omega_{0,1}^2-2\omega_{1,0}^2}{8An^2} + \frac{(2\omega_{1,1}-\omega_{2,0}-\omega_{0,2})m}{4An^2}-\frac{(1-2\lambda)(\omega_{0,3}+3\omega_{2,1}-3\omega_{1,2})}{24A^2n^2}+O(n^{-b})\right).$$

Note that, within the stated error term, the probability is independent of $d_{m+1},\dots,d_n$ except inasmuch as they contribute to d and λ.

The factor $\lambda^X(1-\lambda)^{\binom{m}{2}-X}$ in Theorem 2.4 is the probability for an ordinary Erdős–Rényi random graph with edge probability λ. A sufficient condition for the argument of the exponential to be $o(1)$ is $m^2(m+\max_{j=1}^m|d_j-d|)=o(An)$. Relaxing this condition by a factor of m allows us to see the leading terms of the deviation of behaviour from an ordinary random graph.

Corollary 2.5. *Assume the conditions of Theorem 1.3 and also that* $m(m+\max_{j=1}^m|d_j-d|)=o(An)$ *and* $x_{m+1}=\cdots=x_n=0$. *Then the probability that a random graph with degree sequence* $\boldsymbol{d}$ *has* $\boldsymbol{X}^{[m]}$ *as an induced subgraph is*

$$\lambda^X(1-\lambda)^{\binom{m}{2}-X}\exp\left(\frac{\omega_{1,1}}{2An}-\frac{\omega_{0,2}}{4An}+o(1)\right).$$

The term $\omega_{0,2}/(4An)$ was obtained in [11] in the case that $\boldsymbol{d}=((n-1)/2,\dots,(n-1)/2)$, $m=o(\sqrt{n})$, provided that $x_1,\dots,x_m$ do not differ too much from λm. (In the regular case $\omega_{1,1}=0$.)

It has been shown by Barvinok and Hartigan [1] (see also [5]) that an independent-edge model more accurately matching the $\boldsymbol{d}$ model has each edge jk chosen with probability λ_{jk}, where these constants were introduced in [19] and will appear generalized in the following section. We will call this the $\{\lambda_{jk}\}$-*model*. Under our strict constraints on $\boldsymbol{d}$, we have

$$\lambda_{jk}=\lambda+\frac{d_j-d}{n}+\frac{d_k-d}{n}+\frac{(1-2\lambda)(d_j-d)(d_k-d)}{2An^2}+\widetilde{O}(n^{-3/2}). \tag{2.1}$$

We can restate Theorem 2.4 with that model in mind.

Corollary 2.6. *Assume the conditions of Theorem 1.3 and in addition that* $m=O(n^{1/2+\varepsilon})$ *and* $x_{m+1}=\cdots=x_n=0$. *Then the probability that a random graph with degree sequence* $\boldsymbol{d}$ *has* $\boldsymbol{X}^{[m]}$ *as an induced subgraph is*

$$\prod_{jk\in\boldsymbol{X}}\lambda_{jk}\prod_{jk\notin\boldsymbol{X}}(1-\lambda_{jk}) \times \exp\left(-\frac{\omega_{0,2}}{4An}+\frac{m^2}{2n}+\frac{(1-2\lambda)\omega_{0,1}}{4An}+\frac{4\omega_{1,0}\omega_{0,1}-\omega_{0,1}^2}{8An^2} + \frac{(2\omega_{1,1}-\omega_{0,2})m}{4An^2}-\frac{(1-2\lambda)(\omega_{0,3}-3\omega_{1,2})}{24A^2n^2}+O(n^{-b})\right),$$

where the two products are restricted to $1\leqslant j<k\leqslant m$.

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

Note that the $\omega_{1,1}$ term of Corollary 2.5 has disappeared, but the $\omega_{0,2}$ term remains. The exponential factor quantifies how much the $\{\lambda_{jk}\}$-model is in error. However, when $\boldsymbol{X}$ is generated according to the $\{\lambda_{jk}\}$-model, the expectation of the argument of the exponential is $O(n^{-b})$ and the variance is $\widetilde{O}(n^{-1/2})$, so there is some sense in which we can say that the $\{\lambda_{jk}\}$-model gives a very accurate estimate for *typical* subgraphs. The details of this remain to be worked out.

Let $\boldsymbol{Y}$ be any graph on n vertices and let Y be its number of edges. Define the random variable $E_{\boldsymbol{d},\boldsymbol{Y}}$ to be the number of edges that a random graph with degree sequence $\boldsymbol{d}$ has in common with $\boldsymbol{Y}$. Barvinok and Hartigan [1] proved that $E_{\boldsymbol{d},\boldsymbol{Y}}$ is concentrated close to $\sum_{jk\in\boldsymbol{Y}}\lambda_{jk}$ when $Y=\Omega(n^2)$. In fact their estimate is explicit enough to infer this concentration with weaker bounds for $Y=\omega(n^{3/2}\log n)$. Under our strict conditions on $\boldsymbol{d}$, we can obtain such a result for all Y. We have not determined the best result that follows from Theorem 1.3, but will for this paper be content with the following weak corollary of Theorem 2.4.

Corollary 2.7. *Assume the conditions of Theorem 1.3. If* $Y=O(n^{1/5})$, *then*

$$\operatorname{Prob}(E_{\boldsymbol{d},\boldsymbol{Y}}=k)=\binom{Y}{k}\lambda^k(1-\lambda)^{\binom{Y}{2}-k}\bigl(1+\widetilde{O}(n^{-1/5})\bigr) \tag{2.2}$$

uniformly over $k=0,\dots,Y$. *If* $Y=\Omega(n^{1/5})$, *then*

$$E_{\boldsymbol{d},\boldsymbol{Y}}=\lambda Y(1+O(n^{-1/10+\delta}))$$

with probability $1-O(\exp(n^{-\delta}))$ *for any* $\delta\in\bigl(0,\frac{1}{10}\bigr)$.

Proof. If $Y=O(n^{1/5})$, there are $O(n^{1/5})$ vertices incident with edges of $\boldsymbol{Y}$, so Theorem 2.4 shows that Z has a binomial distribution with the precision given by (2.2). If $Y=\Omega(n^{1/5})$, divide the edges of $\boldsymbol{Y}$ into subsets of size $\Theta(n^{1/5})$ and apply (2.2) and Chernoff's inequality to bound the number of edges in each subset that are in common with the random graph. (Clearly the constants in this corollary can be tuned in various ways.) □

The theorems above should be enough to allow transfer of quite a lot of the theory of ordinary random graphs to dense random graphs with given degrees. However, our purpose in this paper is to develop the tools rather than to explore the applications in detail. We will be content with some simple illustrations.

Theorem 2.8. *Let* $\boldsymbol{d}=(d,d,\dots,d)$ *satisfy the conditions of Theorem 1.3. Then, for a random d-regular graph, we have the following.* (*Note that in each case the quantity in front of the exponential is the expectation for ordinary random graphs with edge probability* λ.)

(a) *If n is even, the expected number of perfect matchings is*

$$\frac{\lambda^{n/2}n!}{2^{n/2}(n/2)!}\exp\left(\frac{1-\lambda}{4\lambda}+O(n^{-b})\right).$$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

(b) *If $q = q(n)$ is a integer function such that $3 \leqslant q \leqslant n$, then the expected number of q-cycles is*

$$\frac{\lambda^q n!}{2q(n-q)!} \exp\left(-\frac{(1-\lambda)q(n-q)}{\lambda n^2} + O(n^{-b})\right).$$

(c) *The expected number of spanning trees is*

$$n^{n-2}\lambda^{n-1} \exp\left(\frac{7(1-\lambda)}{2\lambda} + O(n^{-b})\right).$$

Proof. Parts (a) and (b) follow immediately from Corollaries 2.3 and 2.2, respectively.

Part (c) is not so simple since trees have various degree sequences and those with maximum degree greater than $n^{1/2+\varepsilon}$ do not satisfy the requirements of Theorem 1.3. Let $\mathcal{T}$ be the set of all labelled trees with n vertices. If $\boldsymbol{x}$ denotes the degree sequence of a member of $\mathcal{T}$, let $\mathcal{T}_1$ be the subset of $\mathcal{T}$ with $x_{\max} \leqslant n^{\varepsilon}$ and let $\mathcal{T}_2 = \mathcal{T} \setminus \mathcal{T}_1$.

For sufficiently small $\varepsilon > 0$, the following are true.

(i) The probability in $\mathcal{T}$ that the maximum degree exceeds k is at most $2n/k!$ for any integer $k \geqslant 0$.
(ii) $X = n - 1$.
(iii) In $\mathcal{T}_1$ we have that $X_2 = 5n + O(n^{1/2+3\varepsilon})$ with probability $1 - O(e^{-n^{\varepsilon}})$.
(iv) In $\mathcal{T}_1$ we have $X_3 = \widetilde{O}(n)$ and $H = \widetilde{O}(n)$.
(v) The sum of $\lambda^{-\sum_{j=1}^{n} \max\{0, x_j - n^{\varepsilon}\}}$ over $\mathcal{T}_2$ is $\widetilde{O}(1)\, n^{n-1}/(n^{\varepsilon})!$.

Facts (i) and (iv) follow from the well-known generating function for labelled trees by degree sequence, which is

$$z_1 z_2 \cdots z_n (z_1 + z_2 + \cdots + z_n)^{n-2}.$$

To obtain (iii), note that the same probabilities occur if we take $x_j = 1 + Y_j$ for $1 \leqslant j \leqslant n$, where $Y_1, \ldots, Y_n$ are independent Poisson variates with mean 1 truncated at n^{ε}, subject to having sum $n - 2$. (This is a standard property of multinomial distributions; any mean will do.) Now we can write

$$\mathrm{Prob}\big(|\textstyle\sum Y_j^2 - 2n| \geqslant n^{1/2+3\varepsilon} \mid \sum Y_j = n-2\big) \leqslant \frac{\mathrm{Prob}\big(|\sum Y_j^2 - 2n| \geqslant n^{1/2+3\varepsilon}\big)}{\mathrm{Prob}\big(\sum Y_j = n-2\big)}.$$

The expectation of Y_j^2 is $2 + O(n^{-1})$. Now bound the numerator by applying a concentration inequality like Hoeffding's [10] and the denominator by noting that $\sum Y_j$ would be a Poisson distribution with mean n except for the truncation. Item (iii) follows.

To obtain (v), note that $\sum_{j=1}^{n} \max\{0, x_j - n^{\varepsilon}\} \geqslant \Delta - n^{\varepsilon}$ for trees with maximum degree Δ, and the number of such trees is bounded by (i) with $k = \Delta - 1$. Summing over $\Delta > n^{\varepsilon}$ gives the desired bound.

Now we bound the expected number of trees in $\mathcal{T}_2$ that appear in a random d-regular graph. For a tree $T \in \mathcal{T}_2$, let $F(T)$ be any forest obtained by deleting all but $\lceil n^{\varepsilon} \rceil$ edges from each vertex that has degree greater than n^{ε}. Then the probability that T appears is bounded by the probability that $F(T)$ appears. Moreover, $F(T)$ satisfies the requirements of Corollary 2.2 and has at least $n - 1 - \sum_{j=1}^{n} \max\{0, x_j - n^{\varepsilon}\}$ edges. Corollary 2.2 gives

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

$\mathrm{hit}(\boldsymbol{d}, F(T)) = e^{O(n^\varepsilon)}$. Applying fact (v), we find that the expected number of these trees easily falls within the error term of part (c) of the theorem.

Finally, the trees in $\mathcal{T}_1$ all satisfy the conditions of Corollary 2.2 and have $\mathrm{hit}(\boldsymbol{d}, \boldsymbol{X}) = O(n^3)$. Those with $X_2 = 5n + O(n^{1/2+3\varepsilon})$ have $\mathrm{hit}(\boldsymbol{d}, \boldsymbol{X}) = \exp\bigl(7(1-\lambda)/(2\lambda) + O(n^{-b})\bigr)$. Part (c) of the theorem now follows from items (i) and (iii). □

The average number of spanning trees in random regular graphs of bounded degree was studied in [13].

3. Proof of the main theorem

In this section we will prove Theorem 1.3. We begin by expressing $G(\boldsymbol{d}, \boldsymbol{X})$ as a product $P(\boldsymbol{d}, \boldsymbol{X})I(\boldsymbol{d}, \boldsymbol{X})$, where $P(\boldsymbol{d}, \boldsymbol{X})$ is a rational expression and $I(\boldsymbol{d}, \boldsymbol{X})$ is an n-dimensional integral. These are defined in (3.1). The value of $P(\boldsymbol{d}, \boldsymbol{X})$ is estimated in (3.23), while the value of $I(\boldsymbol{d}, \boldsymbol{X})$ is given by Lemmas 3.1 and 3.3. Putting these together gives Theorem 1.3.

We will use a shorthand notation for summation over doubly subscripted variables. From the matrix $\boldsymbol{X} = (x_{jk})$, define sets

$$\boldsymbol{X}(j) = \{\, k : 1 \leqslant k \leqslant n,\, x_{jk} = 1 \,\}, \quad \bar{\boldsymbol{X}}(j) = \{\, k : 1 \leqslant k \leqslant n,\, x_{jk} = 0,\, k \neq j \,\}$$

for $1 \leqslant j \leqslant n$. Note that $j \notin \boldsymbol{X}(j), \bar{\boldsymbol{X}}(j)$, also that $|\boldsymbol{X}(j)| = x_j$ and $|\bar{\boldsymbol{X}}(j)| = n - 1 - x_j$. If z_{jk} is a symmetric variable for $1 \leqslant j, k \leqslant n$, we define

$$z_{j*} = \sum_{k=1}^{n} z_{jk}, \qquad z_{**} = \sum_{j=1}^{n}\sum_{k=1}^{n} z_{jk},$$

$$z_{j*|\boldsymbol{X}} = \sum_{k \in \boldsymbol{X}(j)} z_{jk}, \qquad z_{**|\boldsymbol{X}} = \sum_{jk \in \boldsymbol{X}} z_{jk},$$

$$z_{j*|\bar{\boldsymbol{X}}} = \sum_{k \in \bar{\boldsymbol{X}}(j)} z_{jk}, \qquad z_{**|\bar{\boldsymbol{X}}} = \sum_{jk \in \bar{\boldsymbol{X}}} z_{jk}.$$

There is some slight lack of symmetry in the definitions. To clarify, we note that

$$\sum_j z_{j*|\boldsymbol{X}} = 2z_{**|\boldsymbol{X}} \quad \text{and} \quad \sum_j z_{j*|\bar{\boldsymbol{X}}} = 2z_{**|\bar{\boldsymbol{X}}}, \quad \text{but} \quad \sum_j z_{j*} = z_{**}.$$

Firstly, notice that $G(\boldsymbol{d}, \boldsymbol{X})$ is the coefficient of $z_1^{d_1} z_2^{d_2} \cdots z_n^{d_n}$ in the function

$$\prod_{jk \in \bar{\boldsymbol{X}}} (1 + z_j z_k).$$

By Cauchy's theorem this equals

$$G(\boldsymbol{d}, \boldsymbol{X}) = \frac{1}{(2\pi i)^n} \oint \cdots \oint \frac{\prod_{jk \in \bar{\boldsymbol{X}}} (1 + z_j z_k)}{z_1^{d_1+1} \cdots z_n^{d_n+1}}\, dz_1 \cdots dz_n,$$

where each integral is along a simple closed contour enclosing the origin anticlockwise. It will suffice to take each contour to be a circle; specifically, we will write

$$z_j = r_j e^{i\theta_j}$$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

for $1 \leqslant j \leqslant n$. Also, define

$$\lambda_{jk} = \frac{r_j r_k}{1 + r_j r_k}$$

for $1 \leqslant j,k \leqslant n$. Then

$$G(\boldsymbol{d}, \boldsymbol{X}) = \frac{\prod_{jk \in \bar{\boldsymbol{X}}} (1 + r_j r_k)}{(2\pi)^n \prod_{j=1}^n r_j^{d_j}} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \frac{\prod_{jk \in \bar{\boldsymbol{X}}} \left(1 + \lambda_{jk}(e^{i(\theta_j + \theta_k)} - 1)\right)}{\exp(i \sum_{j=1}^n d_j \theta_j)} \, d\boldsymbol{\theta}, \tag{3.1}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$. Write $G(\boldsymbol{d}, \boldsymbol{X}) = P(\boldsymbol{d}, \boldsymbol{X}) I(\boldsymbol{d}, \boldsymbol{X})$, where $P(\boldsymbol{d}, \boldsymbol{X})$ denotes the factor in front of the integral in (3.1) and $I(\boldsymbol{d}, \boldsymbol{X})$ denotes the integral. We will choose the radii r_j so that there is no linear term in the logarithm of the integrand of $I(\boldsymbol{d}, \boldsymbol{X})$ when expanded for small $\boldsymbol{\theta}$. The linear term is

$$\sum_{jk \in \bar{\boldsymbol{X}}} \lambda_{jk}(\theta_j + \theta_k) - \sum_{j=1}^n d_j \theta_j.$$

For this to vanish for all $\boldsymbol{\theta}$, we require

$$\lambda_{j*|\bar{\boldsymbol{X}}} = d_j \quad (1 \leqslant j \leqslant n). \tag{3.2}$$

Although it is not hard to show that (3.2) has an exact solution, we can get by with a near-solution since (3.1) is valid for all positive radii. In Section 3.1 we find such a near-solution and determine to sufficient accuracy the various functions of the radii, such as $P(\boldsymbol{d}, \boldsymbol{X})$, that we require. In Section 3.3 we evaluate the integral $I(\boldsymbol{d}, \boldsymbol{X})$ within a certain region $\mathcal{R}$ defined in (3.25). Section 3.6 notes that the contribution to the integral from the region outside of $\mathcal{R}$ and its translate $\mathcal{R} + (\pi, \ldots, \pi)$ is minor in comparison.

3.1. Locating the saddle-point

In this section we derive a near-solution of (3.2) and record some of the consequences. As with the whole paper, we work under the assumptions of Theorem 1.3.

Change variables to $\{a_j\}_{j=1}^n$ as follows:

$$r_j = r \frac{1 + a_j}{1 - r^2 a_j}, \tag{3.3}$$

where

$$r = \sqrt{\frac{\lambda}{1 - \lambda}}.$$

From (3.3) we find that

$$\lambda_{jk}/\lambda = 1 + a_j + a_k + Z_{jk}, \tag{3.4}$$

where

$$Z_{jk} = \frac{a_j a_k (1 - r^2 - r^2 a_j - r^2 a_k)}{1 + r^2 a_j a_k}, \tag{3.5}$$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

and that equation (3.2) can be rewritten as

$$\frac{\delta_j}{\lambda} = (n-1)a_j - a_j x_j + \sum_{k\in\bar{\boldsymbol{X}}(j)} a_k + Z_{j^*|\bar{\boldsymbol{X}}}. \tag{3.6}$$

Summing (3.6) over all j, we find that

$$X = \sum_{j=1}^{n}\bigl((n-1)a_j - a_j x_j\bigr) + Z_{**|\bar{\boldsymbol{X}}}. \tag{3.7}$$

Replace the term $\sum_{k\in\bar{\boldsymbol{X}}(j)} a_k$ in (3.6) by $\sum_{k=1}^{n} a_k - \sum_{k\in\boldsymbol{X}(j)} a_k - a_j$, and substitute the value

$$\sum_{k=1}^{n} a_k = \frac{1}{n}\sum_{k=1}^{n}\bigl(a_k + a_k x_k\bigr) + \frac{X}{n} - \frac{1}{n}Z_{**|\bar{\boldsymbol{X}}}$$

implied by (3.7). After some rearrangement, we find that $a_j = \mathbb{A}_j(a_1,\ldots,a_n)$ for each j, where

$$\begin{aligned}\mathbb{A}_j(a_1,\ldots,a_n) = {}& \frac{\delta_j}{\lambda n} + \frac{2a_j + a_j x_j}{n} - \frac{X}{n^2} - \frac{1}{n^2}\sum_{k=1}^{n}\bigl(a_k + a_k x_k\bigr)\\ &+ \frac{1}{n}\sum_{k\in\boldsymbol{X}(j)} a_k - \frac{1}{n}Z_{j^*|\bar{\boldsymbol{X}}} + \frac{1}{n^2}Z_{**|\bar{\boldsymbol{X}}}.\end{aligned} \tag{3.8}$$

In the vicinity of $\boldsymbol{a} = (0,0,\ldots,0)$, the iteration $\boldsymbol{a} := \bigl(\mathbb{A}_1(\boldsymbol{a}),\ldots,\mathbb{A}_n(\boldsymbol{a})\bigr)$ is a contraction mapping that converges to a solution of (3.2), as can be proved using the method demonstrated in [4]. However, as noted above, we do not need to solve (3.2) exactly but will work with an approximate solution. Hopefully without confusing the reader, from now on we will use $\boldsymbol{a}$ to denote the result of four iterations starting at $\boldsymbol{a} = (0,0,\ldots,0)$. We will also write Z_{jk} and λ_{jk} to mean the values implied by (3.4) and (3.5) for our chosen $\boldsymbol{a}$. Applying (3.8) four times, we find

$$a_j = \frac{\delta_j}{\lambda n} + \frac{\delta_j x_j}{\lambda n^2} - \frac{X}{n^2} + \frac{B_j}{\lambda n^2} + \cdots + \widetilde{O}(n^{-5/2}), \tag{3.9}$$

where the ellipsis conceals about 60 terms of order $\widetilde{O}(n^{-3/2})$. Most of the terms involve counts of subgraphs of $\boldsymbol{X}$ up to order 5, with the vertices weighted by powers of the numbers $\{\delta_j\}$. This implies an expansion

$$Z_{jk} = \frac{\delta_j\delta_k(1-2\lambda)}{2\lambda A n^2} + \cdots + \widetilde{O}(n^{-5/2}).$$

The value of λ_{jk} is given by substituting estimate (3.9) into (3.4). In particular, uniformly over j,

$$\lambda_{j^*|\bar{\boldsymbol{X}}} = d_j + \widetilde{O}(n^{-3/2}). \tag{3.10}$$

Define $\alpha_{jk}, \beta_{jk}, \gamma_{jk}$ by $\alpha_{jk} = \beta_{jk} = \gamma_{jk} = 0$ if $j = k$, and

$$\begin{aligned}\tfrac{1}{2}\lambda_{jk}(1-\lambda_{jk}) &= A + \alpha_{jk},\\ \tfrac{1}{6}\lambda_{jk}(1-\lambda_{jk})(1-2\lambda_{jk}) &= A_3 + \beta_{jk},\\ \tfrac{1}{24}\lambda_{jk}(1-\lambda_{jk})(1-6\lambda_{jk}+6\lambda_{jk}^2) &= A_4 + \gamma_{jk},\end{aligned} \tag{3.11}$$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

if $j \neq k$, where

$$A = \tfrac{1}{2}\lambda(1-\lambda), \quad A_3 = \tfrac{1}{6}\lambda(1-\lambda)(1-2\lambda) \quad \text{and} \quad A_4 = \tfrac{1}{24}\lambda(1-\lambda)(1-6\lambda+6\lambda^2).$$

In evaluating the integral $I(\boldsymbol{d}, \boldsymbol{X})$, the following approximations of α_{jk}, β_{jk}, and γ_{jk} for $j \neq k$ will be required:

$$\begin{aligned}\alpha_{jk} = {} & \frac{(1-2\lambda)(\delta_j+\delta_k)}{2n} - \frac{\delta_j^2+\delta_k^2}{2n^2} + \frac{(1-12A)\delta_j\delta_k}{4An^2} + \frac{(1-2\lambda)(B_j+B_k)}{2n^2} \\ & - \frac{\lambda(1-2\lambda)X}{n^2} + \frac{(1-2\lambda)(\delta_j x_j + \delta_k x_k)}{2n^2} + \widetilde{O}(n^{-3/2}),\end{aligned} \tag{3.12}$$

$$\beta_{jk} = \frac{(1-12A)(\delta_j+\delta_k)}{6n} + \widetilde{O}(n^{-1}), \tag{3.13}$$

$$\gamma_{jk} = \widetilde{O}(n^{-1/2}). \tag{3.14}$$

We will also need the following summations:

$$\alpha_{j*} = \tfrac{1}{2}(1-2\lambda)\delta_j - \frac{\delta_j^2}{2n} - \frac{R_2}{2n^2} + \frac{(1-2\lambda)(\delta_j x_j + B_j)}{2n} + \widetilde{O}(n^{-1/2}), \tag{3.15}$$

$$\alpha_{**} = -\frac{R_2}{n} + \lambda(1-2\lambda)X + \widetilde{O}(n^{1/2}), \tag{3.16}$$

$$\beta_{j*} = \tfrac{1}{6}(1-12A)\delta_j + \widetilde{O}(1), \tag{3.17}$$

$$\beta_{**} = \widetilde{O}(n). \tag{3.18}$$

3.2. Estimating the factor $P(\boldsymbol{d}, \boldsymbol{X})$

Let

$$\Lambda = \prod_{jk \in \bar{\boldsymbol{X}}} \lambda_{jk}^{\lambda_{jk}} (1-\lambda_{jk})^{1-\lambda_{jk}}.$$

Then

$$\begin{aligned}\Lambda^{-1} &= \prod_{jk \in \bar{\boldsymbol{X}}} \left(\left(\frac{1+r_j r_k}{r_j r_k} \right)^{\lambda_{jk}} (1+r_j r_k)^{1-\lambda_{jk}} \right) \\ &= \prod_{jk \in \bar{\boldsymbol{X}}} (1+r_j r_k) \prod_{j=1}^{n} r_j^{-\lambda_{j*|\bar{X}}} \\ &= \prod_{jk \in \bar{\boldsymbol{X}}} (1+r_j r_k) \prod_{j=1}^{n} r_j^{-d_j + \widetilde{O}(n^{-3/2})}\end{aligned}$$

using (3.10). Therefore the factor $P(\boldsymbol{d}, \boldsymbol{X})$ in front of the integral in (3.1) is given by

$$P(\boldsymbol{d}, \boldsymbol{X}) = (2\pi)^{-n} \Lambda^{-1} \exp\bigl(\widetilde{O}(n^{-1/2})\bigr). \tag{3.19}$$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

We proceed to estimate Λ. Writing $\lambda_{jk} = \lambda(1+z_{jk})$, we have

$$\log\left(\frac{\lambda_{jk}^{\lambda_{jk}}(1-\lambda_{jk})^{1-\lambda_{jk}}}{\lambda^{\lambda}(1-\lambda)^{1-\lambda}}\right) = \lambda z_{jk}\log\left(\frac{\lambda}{1-\lambda}\right) + \frac{\lambda}{2(1-\lambda)}z_{jk}^2 - \frac{\lambda(1-2\lambda)}{6(1-\lambda)^2}z_{jk}^3 + \frac{\lambda(1-3\lambda+3\lambda^2)}{12(1-\lambda)^3}z_{jk}^4 + \widetilde{O}(n^{-5/2}). \tag{3.20}$$

We know from (3.10) that $\lambda_{**|\bar{\boldsymbol{X}}} = E + \widetilde{O}(n^{-1/2})$, which implies that $z_{**|\bar{\boldsymbol{X}}} = X + \widetilde{O}(n^{-1/2})$, hence the first term on the right side of (3.20) contributes $\lambda^{\lambda X}(1-\lambda)^{-\lambda X}\exp\bigl(\widetilde{O}(n^{-1/2})\bigr)$ to Λ. Now using (3.4), and recalling that $|\bar{\boldsymbol{X}}| = \binom{n}{2} - X$, we can write $z_{jk} = a_j + a_k + Z_{jk}$ and apply the estimates in the previous subsection to obtain

$$\begin{aligned}\Lambda = {} & \bigl(\lambda^{\lambda}(1-\lambda)^{1-\lambda}\bigr)^{\binom{n}{2}}(1-\lambda)^{-X} \\ & \times \exp\left(\frac{(n+2)R_2}{4An^2} - \frac{(1-2\lambda)R_3}{24A^2n^2} + \frac{(1-6A)R_4}{96A^3n^3}\right. \\ & \left. + \frac{C_{2,1}+2D}{4An^2} - \frac{\lambda^2X^2}{2An^2} + \frac{R_2^2}{16A^2n^4} + \widetilde{O}(n^{-1/2})\right).\end{aligned} \tag{3.21}$$

As in [19], our answer will be simpler when written in terms of binomial coefficients. Using Stirling's formula or otherwise, we find that

$$\begin{aligned}\prod_{j=1}^{n}\binom{n-x_j-1}{d_j} = {} & (2\pi n)^{-n/2}\lambda^{-n/2-\lambda n(n-1)}(1-\lambda)^{-n/2-(1-\lambda)n(n-1)+2X} \\ & \times \exp\left(-\frac{1-14A}{24A} - \frac{R_2}{4An} + \frac{(1-2\lambda)^2R_2}{16A^2n^2} + \frac{(1-2\lambda)R_3}{24A^2n^2}\right. \\ & \left. - \frac{(1-6A)R_4}{96A^3n^3} - \frac{C_{2,1}}{4An^2} + \frac{\lambda X}{(1-\lambda)n} + \widetilde{O}(n^{-1/2})\right).\end{aligned} \tag{3.22}$$

Combining (3.19), (3.21) and (3.22), we find that

$$\begin{aligned}P(\boldsymbol{d},\boldsymbol{X}) = {} & \prod_{j=1}^{n}\binom{n-x_j-1}{d_j}\left(\frac{\pi}{An}\right)^{-n/2}\bigl(\lambda^{\lambda}(1-\lambda)^{1-\lambda}\bigr)^{\binom{n}{2}} \\ & \times \exp\left(\frac{1-14A}{24A} - \frac{R_2}{16A^2n^2} - \frac{R_2^2}{16A^2n^4}\right. \\ & \left. - \frac{\lambda X}{(1-\lambda)n} + \frac{\lambda X^2}{(1-\lambda)n^2} - \frac{D}{2An^2} + O(n^{-b})\right).\end{aligned} \tag{3.23}$$

3.3. Estimating the main part of the integral

Our next task is to evaluate the main part of the integral $I(\boldsymbol{d},\boldsymbol{X})$ given by

$$I(\boldsymbol{d},\boldsymbol{X}) = \int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}\frac{\prod_{jk\in\bar{\boldsymbol{X}}}\bigl(1+\lambda_{jk}(e^{i(\theta_j+\theta_k)}-1)\bigr)}{\exp(i\sum_{j=1}^{n}d_j\theta_j)}\,d\boldsymbol{\theta}. \tag{3.24}$$

It will be established in this section and the next that the value of the integral is concentrated near the places where the integrand has the largest absolute value. This

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

happens at the two points $\boldsymbol{\theta} = (0, 0, \ldots, 0)$ and $\boldsymbol{\theta} = (\pi, \pi, \ldots, \pi)$. These two points are equivalent, since the integrand is unchanged under the mapping $\boldsymbol{\theta} \mapsto \boldsymbol{\theta} + (\pi, \pi, \ldots, \pi)$. (This requires the fact that $\sum d_j$ is even; otherwise the mapping changes the sign of the integrand and the integral is zero as it should be.) Consequently, in this section we will focus on a neighbourhood of $(0, 0, \ldots, 0)$, specifically the hypercube $\mathcal{R}$ defined by

$$\mathcal{R} = \{\boldsymbol{\theta} : |\theta_j| \leqslant n^{-1/2+\varepsilon}, 1 \leqslant j \leqslant n\}. \tag{3.25}$$

Let $F(\boldsymbol{\theta})$ be the integrand of (3.24). We are going to establish the following.

Lemma 3.1. *Under the conditions of Theorem 1.3, there is a region $\mathcal{S}$ with $\mathcal{R} \subseteq \mathcal{S} \subseteq 4\mathcal{R}$ such that*

$$\int_{\mathcal{S}} F(\boldsymbol{\theta})\, d\boldsymbol{\theta} = 2^{-1/2} \left(\frac{\pi}{An}\right)^{n/2} \exp\left(-\frac{1-20A}{24A} + \frac{\lambda X}{(1-\lambda)n} + \frac{R_2}{16A^2n^2} + O(n^{-b})\right). \tag{3.26}$$

In a region $O(1)\mathcal{R}$, we can expand

$$\begin{aligned} F(\boldsymbol{\theta}) &= \exp\Bigg(-\sum_{jk\in\bar{\boldsymbol{X}}} (A+\alpha_{jk})(\theta_j+\theta_k)^2 - i\sum_{jk\in\bar{\boldsymbol{X}}} (A_3+\beta_{jk})(\theta_j+\theta_k)^3 \\ &\quad + \sum_{jk\in\bar{\boldsymbol{X}}} (A_4+\gamma_{jk})(\theta_j+\theta_k)^4 + \widetilde{O}\Bigg(n^{-1/2} + A\sum_{jk\in\bar{\boldsymbol{X}}} |\theta_j+\theta_k|^5\Bigg)\Bigg) \\ &= \exp\Bigg(-\sum_{1\leqslant j<k\leqslant n} (A+\alpha_{jk})(\theta_j+\theta_k)^2 - i\sum_{1\leqslant j<k\leqslant n} (A_3+\beta_{jk})(\theta_j+\theta_k)^3 \\ &\quad + \sum_{1\leqslant j<k\leqslant n} A_4(\theta_j+\theta_k)^4 + \sum_{jk\in\boldsymbol{X}} A\,(\theta_j+\theta_k)^2 + \widetilde{O}(n^{-1/2})\Bigg), \end{aligned}$$

where A, A_3, A_4, α_{jk}, β_{jk}, and γ_{jk} were defined in (3.11). Approximations for α_{jk}, β_{jk}, γ_{jk} were given in (3.12)–(3.14). Note that $\alpha_{jk}, \beta_{jk}, \gamma_{jk} = \widetilde{O}(n^{-1/2})$ uniformly over j, k.

We will transform the integral to diagonalize the quadratic terms, proceeding in two steps. The first step will diagonalize the quadratic form $\sum_{1\leqslant j<k\leqslant n}(\theta_j+\theta_k)^2$, and the second will complete the diagonalization.

3.4. First change of variables

We first adopt from [19] a linear transformation that diagonalizes the quadratic form

$$\sum_{1\leqslant j<k\leqslant n} (\theta_j+\theta_k)^2.$$

Define c and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ by

$$c = 1 - \sqrt{\frac{n-2}{2(n-1)}} = 1 - 2^{-1/2} + O(n^{-1}), \tag{3.27}$$

$$\theta_j = y_j - \frac{c}{n}\sum_{k=1}^{n} y_k \quad (1 \leqslant j \leqslant n). \tag{3.28}$$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

The transformation $\boldsymbol{\theta} = T_1(\boldsymbol{y})$ defined by (3.28) has determinant $1-c$. Also

$$(1+c)\mathcal{R} \subseteq T_1^{-1}\mathcal{R} \subseteq (1-c)^{-1}\mathcal{R}. \tag{3.29}$$

For $\ell \geqslant 1$, define $\mu_\ell = \sum_{j=1}^n y_j^\ell$. We find the following translations:

$$\sum_j \theta_j = (1-c)\mu_1,$$

$$\sum_{1\leqslant j<k\leqslant n} (\theta_j+\theta_k)^2 = (n-2)\mu_2,$$

$$\begin{aligned}\sum_{1\leqslant j<k\leqslant n} (\theta_j+\theta_k)^3 &= (n-4)\mu_3 + \big(3(1-2c)+12c/n\big)\mu_1\mu_2 \\ &\quad + \big((-6c+12c^2-4c^3)/n - 4c^2(3-c)/n^2\big)\mu_1^3,\end{aligned}$$

$$\begin{aligned}\sum_{1\leqslant j<k\leqslant n} (\theta_j+\theta_k)^4 &= (n-8)\mu_4 + 3\mu_2^2 + \big(4(1-2c)+32c/n\big)\mu_1\mu_3 \\ &\quad - \big(24c(1-c)/n + 48c^2/n^2\big)\mu_1^2\mu_2 \\ &\quad + \big(8c^2(1-c)(3-c)/n^2 + 8c^3(4-c)/n^3\big)\mu_1^4,\end{aligned}$$

$$\begin{aligned}\sum_{1\leqslant j<k\leqslant n} \alpha_{jk}(\theta_j+\theta_k)^2 &= \sum_j \big((1-4c/n)\alpha_{j*} + 2c^2\alpha_{**}/n^2\big)y_j^2 \\ &\quad + \sum_{j,k}{}' \big(\alpha_{jk} - 4c\alpha_{j*}/n + 2c^2\alpha_{**}/n^2\big)y_j y_k,\end{aligned}$$

$$\begin{aligned}\sum_{1\leqslant j<k\leqslant n} \beta_{jk}(\theta_j+\theta_k)^3 &= \sum_j \big(1-6c/n+12c^2/n^2)\beta_{j*} - 4c^3\beta_{**}/n^3\big)y_j^3 \\ &\quad + \sum_{j,k}{}' \big((3-12c/n)\beta_{jk} - 6c(1-4c/n)\beta_{k*}/n \\ &\quad + 12c^2\beta_{j*}/n^2 - 12c^3\beta_{**}/n^3\big)y_j y_k^2 \\ &\quad + \sum_{j,k,\ell}{}' \big(-6c\beta_{jk}/n + 12c^2\beta_{j*}/n^2 - 4c^3\beta_{**}/n^3\big)y_j y_k y_\ell,\end{aligned}$$

$$\sum_{jk\in \boldsymbol{X}} (\theta_j+\theta_k)^2 = \sum_{jk\in\boldsymbol{X}} (y_j+y_k)^2 - \frac{4c}{n}\mu_1 \sum_j x_j y_j + \frac{4c^2}{n^2}\mu_1^2 X.$$

In the above, and following, a summation is over $1,2,\ldots,n$ for each index unless otherwise specified. Moreover, a prime on the summation symbol (as $\sum'$) means that only terms where the summation indices have distinct values are included. For example,

$$\sum_{j,k}{}' \quad \text{means} \quad \sum_{\substack{1\leqslant j\leqslant n, 1\leqslant k\leqslant n \\ j\neq k}}.$$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

Using the size of the hypercube $\mathcal{R}$ together with the bounds (3.15)–(3.18), we find that whenever $\boldsymbol{\theta} \in O(1)\mathcal{R}$, we have $F(\boldsymbol{\theta}) = G(\boldsymbol{y})$, where

$$\begin{aligned} G(\boldsymbol{y}) = &- \sum_j \big((n-2)A + \alpha_{j*} - Ax_j\big) y_j^2 \\ &+ \sum_{j,k}{}' \big(-\alpha_{jk} + 2c\alpha_{j*}/n + 2c\alpha_{k*}/n - 2c^2\alpha_{**}/n^2 \\ &+ Ax_{jk} - 2Acx_j/n - 2Acx_k/n + 4Ac^2X/n^2\big) y_j y_k \\ &- i\sum_j \big(nA_3 + \beta_{j*}\big) y_j^3 \\ &- i\sum_{j,k}{}' \big(A_3(3-6c) + 3\beta_{jk} - 6c\beta_{k*}/n\big) y_j y_k^2 \\ &- i\sum_{j,k,\ell}{}' \big(A_3(-6c + 12c^2 - 4c^3)/n - 6c\beta_{jk}/n + 12c^2\beta_{j*}/n^2\big) y_j y_k y_\ell \\ &+ nA_4 \sum_j y_j^4 + 3A_4 \sum_{j,k}{}' y_j^2 y_k^2 \\ &+ 4A_4(1-2c) \sum_{j,k}{}' y_j y_k^3 - 24A_4 c(1-c)/n \sum_{j,k,\ell}{}' y_j y_k y_\ell^2 \\ &+ 8c^2(1-c)(3-c)A_4/n^2 \sum_{j,k,\ell,m}{}' y_j y_k y_\ell y_m + \widetilde{O}(n^{-1/2}). \end{aligned} \tag{3.30}$$

3.5. Completing the diagonalization

We now make a second change of variables, $\boldsymbol{y} = T_2(\boldsymbol{z})$, that diagonalizes the quadratic part of $G(\boldsymbol{y})$, where $\boldsymbol{z} = (z_1, \ldots, z_n)$. We will use the method from [8] that is a slight extension of [18, Lemma 3.2].

Lemma 3.2. *Let $\boldsymbol{U}$ and $\boldsymbol{Y}$ be square matrices of the same order, such that $\boldsymbol{U}^{-1}$ exists and all the eigenvalues of $\boldsymbol{U}^{-1}\boldsymbol{Y}$ are less than 1 in absolute value. Then*

$$(\boldsymbol{I} + \boldsymbol{Y}\boldsymbol{U}^{-1})^{-1/2}(\boldsymbol{U} + \boldsymbol{Y})(\boldsymbol{I} + \boldsymbol{U}^{-1}\boldsymbol{Y})^{-1/2} = \boldsymbol{U},$$

where the fractional powers are defined by the binomial expansion. □

If we also have that both $\boldsymbol{U}$ and $\boldsymbol{Y}$ are symmetric, then $(\boldsymbol{I} + \boldsymbol{Y}\boldsymbol{U}^{-1})^{-1/2}$ is the transpose of $(\boldsymbol{I} + \boldsymbol{U}^{-1}\boldsymbol{Y})^{-1/2}$, as proved in [4].

Let $\boldsymbol{V} = (v_{jk})$ be the symmetric matrix such that the quadratic terms of (3.30) are $-\boldsymbol{y}\boldsymbol{V}\boldsymbol{y}^T$. We have for all $j \neq k$ that

$$\begin{aligned} v_{jj} &= An + \widetilde{O}(n^{1/2}), \\ v_{jk} &= Ax_{jk} + \widetilde{O}(n^{-1/2}). \end{aligned}$$

Apply Lemma 3.2 with $\boldsymbol{V} = \boldsymbol{U} + \boldsymbol{Y}$, where $\boldsymbol{U}$ is the diagonal matrix with the same diagonal entries as $\boldsymbol{V}$. The matrix $\boldsymbol{U}^{-1}\boldsymbol{Y}$ has jk-entry equal to $n^{-1}x_{jk} + \widetilde{O}(n^{-3/2})$. Therefore, since the ∞-norm (maximum row sum of absolute values) of $\boldsymbol{U}^{-1}\boldsymbol{Y}$ is $\widetilde{O}(n^{-1/2})$, the eigenvalues of $\boldsymbol{U}^{-1}\boldsymbol{Y}$ are all $\widetilde{O}(n^{-1/2})$.

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

Let T_2 be the transformation given by $T_2(\boldsymbol{y}) = \boldsymbol{z}$, where $\boldsymbol{z}^T = (\boldsymbol{I} + \boldsymbol{U}^{-1}\boldsymbol{Y})^{-1/2}\boldsymbol{y}^T$. By [4, Lemma 2], the Jacobian of T_2 is $1 + \widetilde{O}(n^{-1/2})$. Expanding $(\boldsymbol{I} + \boldsymbol{U}^{-1}\boldsymbol{Y})^{-1/2}$, we find that for $\boldsymbol{y} \in O(1)\mathcal{R}$,

$$y_j = z_j + \sum_{k=1}^{n} \big(\widetilde{O}(n^{-3/2})z_k + \widetilde{O}(n^{-1})x_{jk}\big)z_k, \tag{3.31}$$

for each j, where the coefficients are uniform and independent of $\boldsymbol{z}$. An expression of identical form writes $\boldsymbol{z}$ in terms of $\boldsymbol{y}$. For $\boldsymbol{y} \in O(1)\mathcal{R}$, we find that $G(\boldsymbol{y}) = H(\boldsymbol{z})$ where

$$\begin{aligned} H(\boldsymbol{z}) = &-\sum_j \big((n-2)A + \alpha_{j*} - Ax_j\big)z_j^2 - i\sum_j \big(nA_3 + \widetilde{O}(n^{1/2})\big)z_j^3 \\ &- i\sum_{j,k}{}' \big(3A_3(1-2c) + \widetilde{O}(n^{-1/2})\big)z_j z_k^2 - i\sum_{j,k,\ell}{}' \widetilde{O}(n^{-1})z_j z_k z_\ell \\ &+ nA_4\sum_j \big(1 + \widetilde{O}(n^{-1})\big)z_j^4 + 3A_4\sum_{j,k}{}' \big(1 + \widetilde{O}(n^{-1})\big)z_j^2 z_k^2 \\ &+ \sum_{j,k}{}' \widetilde{O}(1)z_j z_k^3 - \sum_{j,k,\ell}{}' \widetilde{O}(n^{-1})z_j z_k z_\ell^2 + \sum_{j,k,\ell,m}{}' \widetilde{O}(n^{-2})z_j z_k z_\ell z_m + \widetilde{O}(n^{-1/2}), \end{aligned} \tag{3.32}$$

with only the final expression of the form $\widetilde{O}(\,)$ being a function of $\boldsymbol{z}$.

Now define $\mathcal{S} = T_1^{-1}(T_2^{-1}(2\mathcal{R}))$. By (3.29) and (3.31), $\mathcal{R} \subseteq \mathcal{S} \subseteq 4\mathcal{R}$. Consequently the conditions for our approximations are satisfied and (3.32) is valid for $\boldsymbol{z} \in 2\mathcal{R}$.

We can now apply Theorem A.1 (see Appendix) to estimate the integral of $H(\boldsymbol{z})$ over $2\mathcal{R}$. We list the coefficients required:

$$\begin{aligned} \hat{A} &= A, & \hat{D}_{jk\ell} &= \widetilde{O}(1), \\ N &= n, & \hat{E}_j &= A_4 + \widetilde{O}(n^{-1}), \\ \hat{J}_j &= 0, & \hat{F}_{jk} &= 3A_4 + \widetilde{O}(n^{-1}), \\ \hat{a}_j &= (2A - \alpha_{j*} + Ax_j)n^{-1/2}, & \hat{G}_{jk} &= \widetilde{O}(n^{-1/2}), \\ \hat{B}_j &= -iA_3 + \widetilde{O}(n^{-1/2}), & \hat{H}_{jk\ell} &= \widetilde{O}(n^{-1/2}), \\ \hat{C}_{jk} &= -3(1-\sqrt{2})iA_3 + \widetilde{O}(n^{-1/2}), & \hat{I}_{jk\ell m} &= \widetilde{O}(n^{-1/2}). \end{aligned}$$

We can take $\Delta = \frac{3}{4}$ and have $\delta(\boldsymbol{z}) = \widetilde{O}(n^{-1/2})$. Applying Theorem A.1, we find that

$$\int_{2\mathcal{R}} H(\boldsymbol{z})\,d\boldsymbol{z} = \left(\frac{\pi}{An}\right)^{n/2} \exp\left(-\frac{1-20A}{24A} + \frac{\lambda X}{(1-\lambda)n} + \frac{R_2}{16A^2n^2} + \widetilde{O}(n^{-1/2})\hat{Z}\right), \tag{3.33}$$

where

$$\hat{Z} = \exp\left(\frac{(1-2\lambda)^2}{6A}\right).$$

From the conditions of Theorem 1.3, we have that $\widetilde{O}(n^{-1/2})\hat{Z} = O(n^{-b})$. Lemma 3.1 now follows on recalling that the Jacobian determinants of T_1 and T_2 are $\sqrt{2} + \widetilde{O}(n^{-1/2})$ and $1 + \widetilde{O}(n^{-1/2})$, respectively.

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

3.6. Bounding the remainder of the integral

In the previous section, we estimated the value of the integral $I(\boldsymbol{d}, \boldsymbol{X})$ restricted to a small region $\mathcal{S} \supseteq \mathcal{R}$. As mentioned earlier, the integral over $\mathcal{S} + (\pi, \dots, \pi)$ is the same. It remains to bound the integral over the remaining parts of $[-\pi, \pi]^n$. Define $\mathcal{R}^c = [-\pi, \pi]^n \setminus \big(\mathcal{R} \cup (\mathcal{R} + (\pi, \dots, \pi))\big)$. By employing the same technique as in [8], but with the dissection of the region utilized in [19], we can establish the following. We will omit the proof since no new techniques are required, but we note for convenience that

$$|F(\boldsymbol{\theta})| = \prod_{jk \in \bar{\boldsymbol{X}}} f_{jk}(A + \alpha_{jk}, \theta_j + \theta_k),$$

where

$$f_{jk}(q, z) = \sqrt{1 - 4q(1 - \cos z)} \leqslant \exp\left(-qz^2 + \frac{1}{12} q z^4\right).$$

Lemma 3.3. *Under the conditions of Theorem 1.3,*

$$\int_{\mathcal{R}^c} |F(\boldsymbol{\theta})| \, d\boldsymbol{\theta} = O(n^{-1}) \int_{\mathcal{S}} F(\boldsymbol{\theta}) \, d\boldsymbol{\theta}.$$

Appendix: The value of an integral

In this appendix we give the value of a certain multi-dimensional integral. A similar integral appeared in [19] and variations of it appeared in [9], [16], [17] and [18].

This appendix is notationally independent of the rest of the paper. Summations without explicit limits are over $1, 2, \dots, N$ for each of the summation indices. A prime on the summation symbol (as $\sum'$) indicates that only terms with distinct values of the summation indices are included.

Theorem A.1. *Let $\varepsilon', \varepsilon'', \varepsilon''', \bar{\varepsilon}, \check{\varepsilon}, \Delta$ be constants such that $0 < \varepsilon' < \varepsilon'' < \varepsilon'''$, $\check{\varepsilon} > 0$, $\bar{\varepsilon} \geqslant 0$, and $0 < \Delta < 1$. The following is true if ε''' and $\bar{\varepsilon}$ are sufficiently small.*

Let $\hat{A} = \hat{A}(N)$ be a real-valued function such that $\hat{A}(N) = \Omega(N^{-\varepsilon'})$. For $1 \leqslant j, k, \ell, m$, let $\hat{a}_j$, $\hat{B}_j$, $\hat{C}_{jk}$, $\hat{D}_{jk\ell}$, $\hat{E}_j$, $\hat{F}_{jk}$, $\hat{G}_{jk}$, $\hat{H}_{jk\ell}$, $\hat{I}_{jk\ell m}$, and $\hat{J}_j$ be complex-valued functions of N such that $\hat{a}_j, \hat{B}_j, \dots, \hat{J}_j = O(N^{\bar{\varepsilon}})$ uniformly over j, k, ℓ, m. Suppose that

$$\begin{aligned} f(\boldsymbol{z}) = \exp\Bigg(& -\hat{A}N \sum_j z_j^2 + \sum_j \hat{J}_j z_j + N^{1/2} \sum_j \hat{a}_j z_j^2 + N \sum_j \hat{B}_j z_j^3 + \sum_{j,k}{}' \hat{C}_{jk}\, z_j z_k^2 \\ & + N^{-1} \sum_{j,k,\ell}{}' \hat{D}_{jk\ell}\, z_j z_k z_\ell + N \sum_j \hat{E}_j z_j^4 + \sum_{j,k}{}' \hat{F}_{jk}\, z_j^2 z_k^2 + N^{1/2} \sum_{j,k}{}' \hat{G}_{jk} z_j z_k^3 \\ & + N^{-1/2} \sum_{j,k,\ell}{}' \hat{H}_{jk\ell}\, z_j z_k z_\ell^2 + N^{-3/2} \sum_{j,k,\ell,m}{}' \hat{I}_{jk\ell m}\, z_j z_k z_\ell z_m + \delta(\boldsymbol{z}) \Bigg) \end{aligned}$$

is integrable for $\boldsymbol{z} = (z_1, z_2, \dots, z_N) \in U_N$ and $\delta(N) = \max_{\boldsymbol{z} \in U_N} |\delta(\boldsymbol{z})| = o(1)$, where

$$U_N = \{\boldsymbol{z} \subseteq \mathbb{R}^N : |z_j| \leqslant N^{-1/2+\hat{\varepsilon}} \text{ for } 1 \leqslant j \leqslant N\},$$

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

where $\hat{\varepsilon} = \hat{\varepsilon}(N)$ *satisfies* $\varepsilon'' \leqslant 2\hat{\varepsilon} \leqslant \varepsilon'''$. *Then, provided the* $O(\,)$ *term in the following converges to zero,*

$$\int_{U_N} f(z)\,dz = \left(\frac{\pi}{\hat{A}N}\right)^{N/2} \exp\bigl(\Theta_1 + O\bigl(N^{-1/2+\hat{\varepsilon}} + (N^{-\Delta} + \delta(N))\hat{Z}\bigr)\bigr),$$

where

$$\begin{aligned}
\Theta_1 = {} & \frac{1}{2\hat{A}N^{1/2}} \sum_j \hat{a}_j + \frac{1}{4\hat{A}^2N} \sum_j \hat{a}_j^2 + \frac{15}{16\hat{A}^3N} \sum_j \hat{B}_j^2 + \frac{3}{8\hat{A}^3N^2} \sum_{j,k}{}' \hat{B}_j\hat{C}_{jk} \\
& + \frac{1}{16\hat{A}^3N^3} \sum_{j,k,\ell}{}' \hat{C}_{jk}\hat{C}_{j\ell} + \frac{3}{4\hat{A}^2N} \sum_j \hat{E}_j + \frac{1}{4\hat{A}^2N^2} \sum_{j,k}{}' \hat{F}_{jk} \\
& + \frac{4}{\hat{A}N} \sum_j \hat{J}_j^2 + \frac{3}{4\hat{A}^2N} \sum_j \hat{B}_j\hat{J}_j + \frac{1}{4\hat{A}^2N^2} \sum_{j,k}{}' \hat{C}_{jk}\hat{J}_k,
\end{aligned}$$

$$\begin{aligned}
\hat{Z} = \exp\Bigg(& \frac{1}{4\hat{A}^2N} \sum_j \mathrm{Im}(\hat{a}_j)^2 + \frac{15}{16\hat{A}^3N} \sum_j \mathrm{Im}(\hat{B}_j)^2 + \frac{3}{8\hat{A}^3N^2} \sum_{j,k}{}' \mathrm{Im}(\hat{B}_j)\,\mathrm{Im}(\hat{C}_{jk}) \\
& + \frac{1}{16\hat{A}^3N^3} \sum_{j,k,\ell}{}' \mathrm{Im}(\hat{C}_{jk})\,\mathrm{Im}(\hat{C}_{j\ell}) + \frac{1}{4\hat{A}N} \sum_j \mathrm{Im}(\hat{J}_j)^2 \\
& + \frac{3}{4\hat{A}^2N} \sum_j \mathrm{Im}(\hat{B}_j)\,\mathrm{Im}(\hat{J}_j) + \frac{1}{4\hat{A}^2N^2} \sum_{j,k}{}' \mathrm{Im}(\hat{B}_{jk})\,\mathrm{Im}(\hat{J}_k)\Bigg).
\end{aligned}$$

Proof. The method of proof is the same as in [4], with extra terms added. To simplify the process, we did not explicitly compute the lower-order terms which are presented as Θ_2 in [4]. The details will be omitted. □

References

[1] Barvinok, A. and Hartigan, J. A. (2010) The number of graphs and a random graph with a given degree sequence. Preprint, available at `arxiv.org/abs/1003.0356`.

[2] Boldi, P. and Vigna, S. (2003) Lower bounds for sense of direction in regular graphs. *Distrib. Comput.* **16** 279–286.

[3] Bollobás, B. and McKay, B. D. (1986) The number of matchings in random regular graphs and bipartite graphs. *J. Combin. Theory Ser. B* **41** 80–91.

[4] Canfield, E. R., Greenhill, C. and McKay, B. D. (2008) Asymptotic enumeration of dense 0–1 matrices with specified line sums. *J. Combin. Theory Ser. A* **115** 32–66.

[5] Chatterjee, S., Diaconis, P. and Sly, A. (2010) Random graphs with a given degree sequence. Preprint, available at `arxiv.org/abs/1005.1136`.

[6] Cooper, C., Frieze, A. and Reed, B. (2002) Random regular graphs of non-constant degree: Connectivity and Hamiltonicity. *Combin. Probab. Comput.* **11** 249–261.

[7] Cooper, C., Frieze, A., Reed, B. and Riordan, O. (2002) Random regular graphs of non-constant degree: Independence and chromatic number. *Combin. Probab. Comput.* **11** 323–341.

[8] Greenhill, C. and McKay, B. D. (2009) Random dense bipartite graphs and directed graphs with specified degrees. *Random Struct. Alg.* **35** 222–249.

[9] Greenhill, C., McKay, B. D. and Wang, X. (2006) Asymptotic enumeration of sparse irregular bipartite graphs. *J. Combin. Theory Ser. A* **113** 291–324.

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press

[10] Hoeffding, W. (1963) Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.
[11] Krivelevich, M., Sudakov, B. and Wormald, N. C. (2010) Regular induced subgraphs of a random graph. *Random Struct. Alg.*, published online, doi: 10.1002/rsa.20324.
[12] Krivelevich, M., Sudakov, B., Vu, V. and Wormald, N. C. (2001) Random regular graphs of high degree. *Europ. J. Combin.* **18** 346–363.
[13] McKay, B. D. (1981) Spanning trees in random regular graphs. In *Third Caribbean Conference on Combinatorics and Computing* (University of West Indies 1981), pp. 139–143.
[14] McKay, B. D. (1981) Subgraphs of random graphs with specified degrees. *Congr. Numer.* **33** 213–223.
[15] McKay, B. D. (1985) Asymptotics for symmetric 0–1 matrices with prescribed row sums. *Ars Combin.* **19A** 15–26.
[16] McKay, B. D. (1990) The asymptotic numbers of regular tournaments, Eulerian digraphs and Eulerian oriented graphs. *Combinatorica* **10** 367–377.
[17] McKay, B. D. and Robinson, R. W. (1998) Asymptotic enumeration of Eulerian circuits in the complete graph. *Combin. Probab. Comput.* **7** 437–449.
[18] McKay, B. D. and Wang, X. (1996) Asymptotic enumeration of tournaments with a given score sequence. *J. Combin. Theory Ser. A* **73** 77–90.
[19] McKay, B. D. and Wormald, N. C. (1990) Asymptotic enumeration by degree sequence of graphs of high degree. *European J. Combin.* **11** 565–580.
[20] McKay, B. D. and Wormald, N. C. (1991) Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica* **11** 369–382.
[21] Wormald, N. C. (1999) Models of random regular graphs. In *Surveys in Combinatorics 1999* (J. D. Lamb and D. A. Preece, eds), Cambridge University Press, pp. 239–298.

https://doi.org/10.1017/S0963548311000034 Published online by Cambridge University Press