*The Number of Graphs and a Random Graph with a Given Degree Sequence**

Alexander Barvinok,[1] J. A. Hartigan[2]
[1] *Department of Mathematics, University of Michigan, Ann Arbor, Michigan; e-mail: barvinok@umich.edu*
[2] *Department of Statistics, Yale University, New Haven, Connecticut; e-mail: john.hartigan@yale.edu*

Received 1 June 2010; accepted 16 November 2011
Published online 29 February 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/rsa.20409

ABSTRACT: We consider the set of all graphs on n labeled vertices with prescribed degrees $D = (d_1, \ldots, d_n)$. For a wide class of *tame* degree sequences D we obtain a computationally efficient asymptotic formula approximating the number of graphs within a relative error which approaches 0 as n grows. As a corollary, we prove that the structure of a random graph with a given tame degree sequence D is well described by a certain *maximum entropy matrix* computed from D. We also establish an asymptotic formula for the number of bipartite graphs with prescribed degrees of vertices, or, equivalently, for the number of 0-1 matrices with prescribed row and column sums.
 Random Struct. Alg., 42, 301–348, 2013

Keywords: graphs; degree sequences; asymptotic formulas

1. INTRODUCTION AND MAIN RESULTS

1.1. Graphs and Their Degree Sequences

Let $D = (d_1, \ldots, d_n)$ be a vector of positive integers and let $G(D)$ be the set of all graphs (undirected, with no loops or multiple edges) on the set $\{1, \ldots, n\}$ of vertices such that the degree of the k-th vertex is d_k for $k = 1, \ldots, n$. Equivalently, $G(D)$ is the set of all $n \times n$ symmetric matrices with 0-1 entries, zero trace and row (column) sums $d_1, \ldots, d_n$. We assume that

$$d_1 + \cdots + d_n \equiv 0 \mod 2, \tag{1.1.1}$$

since otherwise the set $G(D)$ is empty.

Correspondence to: A. Barvinok
*Supported by NSF (DMS 0856640); United States - Israel BSF (2006377).

The theorem of Erdős and Gallai, see, for example, Theorem 6.3.6 of [6], states the necessary and sufficient conditions for the existence of a graph with the given degree sequence. Without loss of generality, we assume that

$$d_1 \geq d_2 \geq \cdots \geq d_n.$$

Then, the necessary and sufficient condition for $G(D)$ to be non-empty is that (1.1.1) holds and

$$\sum_{i=1}^{k} d_i \leq k(k-1) + \sum_{i=k+1}^{n} \min\{k, d_i\} \quad \text{for } k = 1, \ldots, n. \tag{1.1.2}$$

Our main goal is to estimate the cardinality $|G(D)|$ of $G(D)$. Using the obtained estimate, we deduce a concentration result for a random graph $G \in G(D)$ sampled from the uniform probability measure on $G(D)$.

1.2. The Maximum Entropy Matrix and Tame Degree Sequences

The following matrix plays the crucial role in our construction.

Let us consider the space $\mathbb{R}^{\binom{n}{2}}$ of vectors $x = (\xi_{\{j,k\}})$, where $\{j, k\}$ is an unordered pair of indices $1 \leq j \neq k \leq n$. We consider the polytope $\mathcal{P} \subset \mathbb{R}^{\binom{n}{2}}$, $\mathcal{P} = \mathcal{P}(D)$, defined by the equations

$$\sum_{j: j \neq k} \xi_{\{j,k\}} = d_k \quad \text{for } k = 1, \ldots, n$$

and inequalities

$$0 \leq \xi_{\{j,k\}} \leq 1.$$

The integer points in $\mathcal{P}(D)$ correspond to the labeled graphs with degree sequence D, which we write as

$$G(D) = \mathcal{P}(D) \cap \mathbb{Z}^{\binom{n}{2}}.$$

We assume that $\mathcal{P}(D)$ is non-empty. We consider the following *entropy function* on $\mathcal{P}(D)$:

$$H(x) = \sum_{\{j,k\}} \left(\xi_{\{j,k\}} \ln \frac{1}{\xi_{\{j,k\}}} + (1 - \xi_{\{j,k\}}) \ln \frac{1}{1 - \xi_{\{j,k\}}} \right) \quad \text{for } x = (\xi_{\{j,k\}}).$$

Since H is a strictly concave function, it attains its maximum on $\mathcal{P}$ at a unique point, $z = (\zeta_{\{j,k\}})$, $z = z(D)$, which we call the *maximum entropy matrix* associated with the degree sequence D. Matrix z can be easily calculated by interior point methods, see [19].

For $0 < \delta \leq 1/2$ we say that the degree sequence D is δ-*tame* if the polytope $\mathcal{P}(D)$ is non-empty and if

$$\delta \leq \zeta_{\{j,k\}} \leq 1 - \delta \quad \text{for all } 1 \leq j \neq k \leq n,$$

where $z = (\zeta_{\{j,k\}})$ is the maximum entropy matrix associated with degree sequence D. In Theorem 2.1 we state some sufficient conditions for a degree sequence D to be tame.

1.3. Quadratic Form *q* and Related Quantities

Let $z = (\zeta_{\{j,k\}})$ be the maximum entropy matrix associated with a tame degree sequence D. We consider the following quadratic form $q : \mathbb{R}^n \longrightarrow \mathbb{R}$,

$$q(t) = \frac{1}{2} \sum_{\{j,k\}} \left(\zeta_{\{j,k\}} - \zeta^2_{\{j,k\}}\right)(\tau_j + \tau_k)^2 \quad \text{for } t = (\tau_1, \ldots, \tau_n). \tag{1.3.1}$$

It is easy to see that q is positive definite for $n > 2$. Let us consider the Gaussian probability measure on $\mathbb{R}^n$ with density proportional to e^{-q}. We define the following random variables $f, h : \mathbb{R}^n \longrightarrow \mathbb{R}$,

$$\begin{aligned} f(t) &= \frac{1}{6} \sum_{\{j,k\}} \zeta_{\{j,k\}}(1 - \zeta_{\{j,k\}})(2\zeta_{\{j,k\}} - 1)(\tau_j + \tau_k)^3 \quad \text{and} \\ h(t) &= \frac{1}{24} \sum_{\{j,k\}} \zeta_{\{j,k\}}(1 - \zeta_{\{j,k\}})\left(6\zeta^2_{\{j,k\}} - 6\zeta_{\{j,k\}} + 1\right)(\tau_j + \tau_k)^4 \quad \text{for } t = (\tau_1, \ldots, \tau_n). \end{aligned} \tag{1.3.2}$$

Let

$$\mu = \mathbf{E} f^2 \quad \text{and} \quad \nu = \mathbf{E} h.$$

Our main result is as follows.

Theorem 1.4. *Let us fix $0 < \delta < 1/2$. Let $D = (d_1, \ldots, d_n)$ be a δ-tame degree sequence such that $d_1 + \cdots + d_n \equiv 0 \mod 2$, let $z = (\zeta_{\{j,k\}})$ be the maximum entropy matrix as defined in Section 1.2 and let the quadratic form q and values of μ and ν be as defined in Section 1.3. Let us define an $n \times n$ symmetric matrix $Q = (\omega_{jk})$ by*

$$\begin{aligned} \omega_{jk} &= \zeta_{\{j,k\}}(1 - \zeta_{\{j,k\}}) \quad \textit{for } j \neq k \quad \textit{and} \\ \omega_{jj} &= d_j - \sum_{k:\ k \neq j} \zeta^2_{\{j,k\}} \quad \textit{for } j = 1, \ldots, n. \end{aligned}$$

Then Q is positive definite and the value of

$$\frac{2e^{H(z)}}{(2\pi)^{n/2}\sqrt{\det Q}} \exp\left\{-\frac{\mu}{2} + \nu\right\} \tag{1.4.1}$$

approximates the number of graphs $|G(D)|$ with degree sequence D within a relative error which approaches 0 as $n \longrightarrow +\infty$.

More precisely, for any $0 < \epsilon \leq 1/2$ the value of (1.4.1) approximates $|G(D)|$ within relative error ϵ provided

$$n \geq \left(\frac{1}{\epsilon}\right)^{\gamma(\delta)},$$

where $\gamma = \gamma(\delta)$ is a positive constant.

The main term

$$\frac{2e^{H(z)}}{(2\pi)^{n/2}\sqrt{\det Q}} \tag{1.4.2}$$

of formula (1.4.1) is the "Gaussian approximation" formula of [3], whose appearance, as is discussed in [3], is explained by the Local Central Limit Theorem, see also the discussion below. The factor

$$\exp\left\{-\frac{\mu}{2}+\nu\right\}$$

is the "Edgeworth correction" factor, see [5]. In the course of the proof of Theorem 1.4, we establish a two-sided bound

$$\gamma_1(\delta) \leq \exp\left\{-\frac{\mu}{2}+\nu\right\} \leq \gamma_2(\delta)$$

for some constants $\gamma_1(\delta), \gamma_2(\delta) > 0$, as long as the degree sequence D remains δ-tame.

We note that computing the expectation of a polynomial with respect to the Gaussian probability measure is a linear algebra problem, cf. also Section 5.2. Hence apart from computing the maximum entropy matrix z, which can be done by interior point methods, computing the value of (1.4.1) is a linear algebra problem which can be solved in $O(n^4)$ time in the unit cost model.

1.5. Random Graphs with Prescribed Degree Sequences

Let us consider the set $G(D)$ of all labeled graphs with degree sequence D as a finite probability space with the uniform measure. It is convenient to think of $G \in G(D)$ as of a subgraph of the complete graph K_n with the set

$$V = \{1, \ldots, n\}$$

of vertices and the set

$$E = \{\{j,k\}: \ 1 \leq j \neq k \leq n\}$$

of edges.

Let us sample a graph $G \in G(D)$ at random. What G is likely to look like?

As a corollary of Theorem 1.4, we prove that with overwhelming probability, for a random graph $G \in G(D)$ the number of edges of G in a given set $S \subset E$ with $|S| = \Omega(n^2)$ is very close to the sum of the entries of the maximum entropy matrix indexed by the elements of S.

Theorem 1.6. *Let us fix numbers $\kappa > 0$ and $0 < \delta \leq 1/2$. Then there exists a number $\gamma(\kappa,\delta) > 0$ such that the following holds.*

Suppose that $n \geq \gamma(\kappa,\delta)$ and that $D = (d_1, \ldots, d_n)$ is a δ-tame degree sequence such that $d_1 + \cdots + d_n \equiv 0 \mod 2$. For a set $S \subset E$, let $\sigma_S(G)$ be the number of edges of graph $G \in G(D)$ that belong to set S and let

$$\sigma_S(z) = \sum_{\{j,k\}\in S} \zeta_{\{j,k\}},$$

where $z = (\zeta_{\{j,k\}})$ is the maximum entropy matrix. Suppose that $|S| \geq \delta n^2$ and let

$$\epsilon = \delta \frac{\ln n}{\sqrt{n}}.$$

If $\epsilon \leq 1$ then for a uniformly chosen random graph $G \in G(D)$, we have

$$\mathbf{P}\{G \in G(D) : (1-\epsilon)\sigma_S(z) \leq \sigma_S(G) \leq (1+\epsilon)\sigma_S(z)\} \geq 1 - 2n^{-\kappa n}.$$

The idea of the proof is as follows. For $1 \leq j \neq k \leq n$, let $x_{\{j,k\}}$ be independent Bernoulli random variables such that

$$\mathbf{P}\{x_{\{j,k\}} = 1\} = \zeta_{\{j,k\}} \quad \text{and} \quad \mathbf{P}\{x_{\{j,k\}} = 0\} = 1 - \zeta_{\{j,k\}}.$$

As is shown in [3], the probability mass function of the random vector $X = (x_{\{j,k\}})$ is constant on the integer points of $\mathcal{P}(D)$ and is equal to $e^{-H(z)}$ at each $G \in G(D)$, so that the vector X conditioned on $G(D)$ is uniform. Theorem 1.4 then implies that the probability that $X \in G(D)$ is not too small. On the other hand, standard large deviation inequalities imply that the sum $\sum_{\{j,k\} \in S} x_{\{j,k\}}$ concentrates about the value of $\sigma_S(z) = \sum_{\{j,k\}} \zeta_{\{j,k\}}$. We supply the details of the proof in Section 10.

In many respects random graphs $G \in G(D)$ behave like random graphs on the set $\{1, \ldots, n\}$ of vertices, with pairs $\{j, k\}$ chosen as the edges of G independently with probabilities $\zeta_{\{j,k\}}$, where $z = (\zeta_{\{j,k\}})$ is the maximum entropy matrix. As is discussed in [3], the distribution of the multivariate Bernoulli random vector $X = (x_{\{j,k\}})$ is the distribution of the largest entropy among all multivariate Bernoulli random vectors constrained by

$$\mathbf{E} y_k = d_k \quad \text{for } k = 1, \ldots, n$$

where

$$y_k = \sum_{j:\ j \neq k} x_{\{j,k\}}.$$

We remark that we obtain the "Gaussian approximation" term (1.4.2) if we assume that the vector of random variables $Y = (y_1, \ldots, y_n)$ is asymptotically Gaussian around its expectation $(d_1, \ldots, d_n)$. As it turns out, Y is not exactly Gaussian but is not very far from it.

It looks plausible that both Theorem 1.4 and Theorem 1.6 can be extended to degree sequences D allowing a moderate number of entries $\zeta_{\{j,k\}}$ of the maximum entropy matrix to be arbitrarily close to 1 or 0. Our proofs, however, do not seem to allow such an extension with Theorem 4.1 being the main obstacle. Some of our proofs (mostly in Sections 5 and 6) are similar to those of [4], where we applied the maximum entropy approach of [3] to count non-negative integer matrices with prescribed row and column sums.

The paper is organized as follows.

In Section 2, we give several examples and extensions concerning our main result, Theorem 1.4 and also discuss related work in the literature.

In Section 3, we present an integral representation for the number $|G(D)|$ of graphs and also describe the plan of the proof of Theorem 1.4.

The rest of the paper deals with the proofs.

2. EXAMPLES AND EXTENSIONS

Sometimes one can tell that a degree sequence is tame without computing the maximum entropy matrix.

Theorem 2.1. *Let us fix real numbers* $0 < \alpha < \beta < 1$ *such that*

$$\beta < 2\sqrt{\alpha} - \alpha, \quad \textit{or, equivalently,} \quad (\alpha+\beta)^2 < 4\alpha.$$

Then there exists a real number $\delta = \delta(\alpha, \beta) > 0$ *and a positive integer* $n_0 = n_0(\alpha, \beta)$ *such that any degree sequence* $D = (d_1, \ldots, d_n)$ *satisfying*

$$\alpha < \frac{d_i}{n-1} < \beta \quad \textit{for } i = 1, \ldots, n$$

is δ*-tame provided* $n > n_0$.

One can choose

$$n_0 = \max\left\{\frac{\beta}{\alpha(1-\beta)}, \quad \frac{4(\beta-\alpha)}{4\alpha-(\alpha+\beta)^2}\right\} + 1 \quad \textit{and}$$

$$\delta = \frac{\epsilon^6}{1+\epsilon^6} \quad \textit{where } \epsilon = \min\left\{\alpha,\ \alpha - \frac{(\alpha+\beta)^2}{4}\right\}.$$

For example, degree sequences $D = (d_1, \ldots, d_n)$ satisfying

$$0.25 < \frac{d_i}{n-1} < 0.74 \quad \text{for } i = 1, \ldots, n$$

or

$$0.01 < \frac{d_i}{n-1} < 0.18 \quad \text{for } i = 1, \ldots, n$$

or

$$0.81 < \frac{d_i}{n-1} < 0.89 \quad \text{for } i = 1, \ldots, n$$

are δ-tame for some $\delta > 0$ and all sufficiently large n.

We prove Theorem 2.1 in Section 12.

2.2. On the Boundary of δ-Tameness

Let us choose rational $0 < \alpha < \beta < 1$ such that

$$\beta = 2\sqrt{\alpha} - \alpha. \tag{2.2.1}$$

Clearly, $\beta > \alpha$. Let us choose a positive integer n such that αn and βn are even integers and let us consider the degree sequence

$$d_1 = \ldots = d_k = \beta n \quad \text{and} \quad d_{k+1} = \ldots = d_n = \alpha n \quad \text{for } k = n\sqrt{\alpha}$$

(note that k is necessarily integral). The Erdős-Gallai condition (1.1.2) for $k = n\sqrt{\alpha}$, reduces to

$$\beta \leq 2\sqrt{\alpha} - \alpha - \frac{1}{n}. \tag{2.2.2}$$

In particular, (2.2.1) does not even guarantee that the polytope $\mathcal{P}(D)$ is non-empty.

In [12] Jerrum, Sinclair and McKay discuss under what conditions an approximation formula for $|G(D)|$ which depends "smoothly" on D may exist. They describe the phenomenon of the number of graphs $|G(D)|$ changing sharply when the degree sequence D is varying only slightly around some special values of D. This phenomenon is apparently explained by the fact that the dimension of the polytope $\mathcal{P}(D)$ may change abruptly or the polytope may disappear altogether when D lies on the boundary of the Erdős-Gallai conditions (1.1.2). Theorem 8.1 of [12] states that for

$$d_+ = \max\{d_i,\ i = 1, \ldots, n\} \quad \text{and} \quad d_- = \min\{d_i,\ i = 1, \ldots, n\},$$

as long as

$$(d_+ - d_- + 1)^2 \leq 4d_-(n - d_+ - 1), \tag{2.2.3}$$

the degree sequence D is *P-stable*, meaning that increasing one of the degrees d_i and decreasing another by 1 does not change $|G(D)|$ by more than a factor of n^{10} (this, in turn, implies that there are polynomial time randomized approximation algorithms for computing $|G(D)|$ and sampling a random graph $G \in G(D)$). The condition of our Theorem 2.1 is only marginally stronger than (2.2.3).

As Sourav Chatterjee pointed out to us, Lemma 4.1 of recent [9] shows that a sequence D is δ-tame provided it lies sufficiently deep inside the polyhedron defined by the Erdős-Gallai conditions (1.1.2).

Our example shows that the bounds of Theorem 2.1 are essentially the best possible if we take into account only the largest and the smallest degree of a vertex of the graph.

2.3. Regular Graphs

In [18] McKay and Wormald compute the asymptotic of $|G(D)|$ for regular graphs, where

$$d_1 = \ldots = d_n = d,$$

and almost regular graphs, where

$$|d_i - d| < n^{\frac{1}{2}+\epsilon} \quad \text{for } i = 1, \ldots, n$$

for a sufficiently small $\epsilon > 0$; see also [17] for recent developments and [16] for a survey. One can show that the formula of Theorem 1.4 is equivalent to the asymptotic formula of [18] for regular or almost regular graphs.

In the case of regular graphs, symmetry requires that

$$\zeta_{\{j,k\}} = \frac{d}{n-1} \quad \text{for all } 1 \leq j \neq k \leq n$$

for the maximum entropy matrix $z = (\zeta_{\{j,k\}})$.

2.4. Approximations in the Cut Norm

The *cut norm* (sometimes called the *normalized* cut norm) of a real $m \times n$ matrix $A = (a_{jk})$ is defined by

$$\|A\|_{\text{cut}} = \frac{1}{mn} \max_{J,K} \left| \sum_{j \in J, k \in K} a_{jk} \right|,$$

where the maximum is taken over all non-empty subsets $J \subset \{1, \ldots, m\}$ and $K \subset \{1, \ldots, n\}$. Let us choose set S in Theorem 1.6 of the form

$$S = \{\{j,k\} : j \in J, k \in K, j \neq k\} \qquad \text{for some } J, K \subset \{1, \ldots, n\}.$$

We note that there are not more than 2^{2n} distinct sets S of this form. Theorem 1.6 implies that as n grows, the maximum entropy matrix $z(D)$ approximates the adjacency matrix of the overwhelming majority of graphs $G \in G(D)$ within an error of $O(n^{-1/2} \ln n)$ in the cut norm.

Shortly after the first version of this paper appeared, using a different approach, Chatterjee, Diaconis and Sly [9] described *graph limits* of graphs from $G(D)$ as n grows. A graph limit is a certain function on $[0, 1] \times [0, 1]$, viewed as an "infinite matrix", which naturally arises as a limit object for a Cauchy sequence in the cut norm of adjacency matrices of graphs [14]. Graph limits constructed in [9] can indeed be viewed as "infinite maximum entropy matrices".

2.5. Enumeration of Bipartite Graphs

A natural version of the problem concerns enumeration of labeled *bipartite* graphs with a given degree sequence or, equivalently, $m \times n$ matrices with 0-1 entries and prescribed row sums $R = (r_1, \ldots, r_m)$ and column sums $C = (c_1, \ldots, c_n)$. We assume that

$$r_1 + \cdots + r_m = c_1 + \cdots + c_n.$$

A simple necessary and sufficient condition for a 0-1 matrix with prescribed row and column sums to exist is given by the Gale-Ryser Theorem, see, for example, Corollary 6.2.5 of [6].

Let us consider the polytope $\mathcal{P}(R, C)$ of $m \times n$ matrices $x = (\xi_{jk})$ defined by the equations

$$\sum_{k=1}^{n} \xi_{jk} = r_j \quad \text{for } j = 1, \ldots, m \qquad \text{and} \qquad \sum_{j=1}^{m} \xi_{jk} = c_k \quad \text{for } k = 1, \ldots, n$$

and inequalities

$$0 \leq \xi_{jk} \leq 1 \quad \text{for all } j, k.$$

Let us compute the *maximum entropy* matrix $z = (\zeta_{jk})$ as the necessarily unique matrix $z \in \mathcal{P}(R, C)$ that maximizes

$$H(x) = \sum_{jk} \left(\xi_{jk} \ln \frac{1}{\xi_{jk}} + (1 - \xi_{jk}) \ln \frac{1}{1 - \xi_{jk}} \right) \quad \text{for } x = (\xi_{jk})$$

on $\mathcal{P}(R, C)$. For $0 < \delta \leq 1/2$, we say that the margins (R, C) are δ-*tame* if

$$\delta m \leq n \quad \text{and} \quad \delta n \leq m$$

and

$$\delta \leq \xi_{jk} \leq 1 - \delta \quad \text{for all } j, k.$$

Suppose that the margins (R, C) are indeed δ-tame for some $\delta > 0$. Let us define a quadratic form $q : \mathbb{R}^{m+n} \longrightarrow \mathbb{R}$ by

$$q(s, t) = \frac{1}{2} \sum_{j,k} \left(\zeta_{jk} - \zeta_{jk}^2\right)(\sigma_j + \tau_k)^2 \quad \text{for } (s, t) = (\sigma_1, \ldots, \sigma_m; \tau_1, \ldots, \tau_n). \tag{2.5.1}$$

Let

$$u = \Bigg(\underbrace{1, \ldots, 1}_{m \text{ times}}; \underbrace{-1, \ldots, -1}_{n \text{ times}}\Bigg) \tag{2.5.2}$$

and let $L = u^{\perp}$ be the orthogonal complement to u in $\mathbb{R}^{m+n}$. Then the restriction $q|L$ of q onto L is strictly positive definite and we define $\det q|L$ as the product of the non-zero eigenvalues of q. We consider the Gaussian probability measure on L with density proportional to e^{-q} and define random variables $f, g : L \longrightarrow \mathbb{R}$ by

$$\begin{aligned} f(s, t) &= \frac{1}{6} \sum_{j,k} \zeta_{jk}(1 - \zeta_{jk})(2\zeta_{jk} - 1)(\sigma_j + \tau_k)^3 \quad \text{and} \\ h(s, t) &= \frac{1}{24} \sum_{j,k} \zeta_{jk}(1 - \zeta_{jk})\left(6\zeta_{jk}^2 - 6\zeta_{jk} + 1\right)(\sigma_j + \tau_k)^4 \\ &\quad \text{for } (s, t) = (\sigma_1, \ldots, \sigma_m; \tau_1, \ldots, \tau_n). \end{aligned} \tag{2.5.3}$$

We define

$$\mu = \mathbf{E} f^2 \quad \text{and} \quad \nu = \mathbf{E} h.$$

Then the number $|R, C|$ of 0-1 matrices with row sums R and column sums C is

$$|R, C| = \frac{e^{H(z)}\sqrt{m + n}}{(4\pi)^{(m+n-1)/2}\sqrt{\det q|L}} \exp\left\{-\frac{\mu}{2} + \nu\right\} (1 + o(1)) \tag{2.5.4}$$

provided $m, n \longrightarrow +\infty$ in such a way that the margins (R, C) remain δ-tame for some $\delta > 0$. We sketch the proof of (2.5.4) in Section 11.

Canfield and McKay [8] obtained an asymptotic formula of $|R, C|$ when all row sums are equal, $r_1 = \cdots = r_m$ and all column sums are equal, $c_1 = \cdots = c_n$, which was later extended to the case of "almost equal" row sums and "almost equal" column sums [7], see also [11]. The maximum entropy matrix z was introduced in [2] where a cruder asymptotic formula

$$\ln |R, C| \approx H(z)$$

was established without the δ-tameness assumption and for a wider class of enumeration problems, including enumeration of 0-1 matrices with prescribed row and column sums *and* zeros in prescribed position. It was also shown in [2] that a random matrix 0-1 with prescribed row and column sums concentrates about the maximum entropy matrix z.

3. AN INTEGRAL REPRESENTATION FOR THE NUMBER OF GRAPHS

In [3] we proved the following general result; see Theorem 5, Lemma 11 and formula (16) there.

Theorem 3.1. *Let $P \subset \mathbb{R}^p$ be a polyhedron defined by the system of linear equations $Ax = b$, where A is a $n \times p$ matrix with columns $a_1, \dots, a_p \in \mathbb{Z}^n$ and $b \in \mathbb{Z}^n$ is an integer vector, and inequalities $0 \leq x \leq 1$ (the inequalities are understood coordinate-wise). Suppose that P has a non-empty interior, that is, contains a point $x = (\xi_1, \dots, \xi_p)$ such that $0 < \xi_j < 1$ for $j = 1, \dots, p$.*

Then the function

$$H(x) = \sum_{j=1}^{p} \left(\xi_j \ln \frac{1}{\xi_j} + (1 - \xi_j) \ln \frac{1}{1 - \xi_j} \right) \quad \text{for } x = (\xi_1, \dots, \xi_p)$$

attains its maximum on P at a unique point $z = (\zeta_1, \dots, \zeta_p)$ such that $0 < \zeta_j < 1$ for $j = 1, \dots, p$.

Let us consider the parallelepiped $\Pi = [-\pi, \pi]^n$, $\Pi \subset \mathbb{R}^n$. Then the number $|P \cap \{0, 1\}^p|$ of 0-1 points in P can be written as

$$|P \cap \{0, 1\}^p| = \frac{e^{H(z)}}{(2\pi)^n} \int_{\Pi} e^{-i\langle t, b\rangle} \prod_{j=1}^{p} \left(1 - \zeta_j + \zeta_j e^{i\langle a_j, t\rangle}\right) dt,$$

where $\langle \cdot, \cdot \rangle$ is the standard scalar product in $\mathbb{R}^n$, dt is the standard Lebesgue measure in $\mathbb{R}^n$ and $i = \sqrt{-1}$. ∎

The idea of the proof is as follows. Let $X = (x_1, \dots, x_p)$ be a random vector of independent Bernoulli random variables such that $\mathbf{P}\{x_j = 1\} = \zeta_j$ and $\mathbf{P}\{x_j = 0\} = 1 - \zeta_j$ for $j = 1, \dots, p$. It turns out that the probability mass function of X is constant on the set $P \cap \{0, 1\}^p$ and equals $e^{-H(z)}$ for every 0-1 point in P. Letting $Y = AX$, we obtain

$$|P \cap \{0, 1\}^p| = e^{H(z)} \mathbf{P}\{X \in P\} = e^{H(z)} \mathbf{P}\{Y = b\}$$

and the probability in question is written as the integral of the characteristic function of Y.

Since

$$\sum_{j=1}^{p} \zeta_j a_j = b,$$

in a neighborhood of the origin $t = 0$ the integrand can be written as

$$e^{-i\langle t,b\rangle} \prod_{j=1}^{p} \left(1 - \zeta_j + \zeta_j e^{i\langle a_j,t\rangle}\right) = \exp\left\{ -\frac{1}{2} \sum_{j=1}^{p} \zeta_j(1-\zeta_j)\langle a_j, t\rangle^2 \right.$$
$$+ \frac{i}{6} \sum_{j=1}^{p} \zeta_j(1-\zeta_j)(2\zeta_j - 1)\langle a_j, t\rangle^3$$
$$+ \frac{1}{24} \sum_{j=1}^{p} \zeta_j(1-\zeta_j)(6\zeta_j^2 - 6\zeta_j + 1)\langle a_j, t\rangle^4$$
$$\left. + O\left(\sum_{j=1}^{p} (\zeta_j + 1)^5 \langle a_j, t\rangle^5 \right) \right\}. \tag{3.2}$$

Note that the linear term is absent in the expansion.

We obtain the following corollary.

Corollary 3.3. *Let $D = (d_1, \ldots, d_n)$ be a degree sequence such that the polytope $\mathcal{P}(D)$ defined in Section 1.2 has a non-empty interior and let $z = (\zeta_{\{j,k\}})$ be the maximum entropy matrix. Let*

$$F(t) = \exp\left\{ -i \sum_{m=1}^{n} d_m \tau_m \right\} \prod_{\{j,k\}} (1 - \zeta_{\{j,k\}} + \zeta_{\{j,k\}} e^{i(\tau_j + \tau_k)}) \quad \text{for } t = (\tau_1, \ldots, \tau_n).$$

Then for the parallelepiped $\Pi = [-\pi, \pi]^n$, we have

$$|G(D)| = \frac{e^{H(z)}}{(2\pi)^n} \int_{\Pi} F(t)\,dt.$$

Proof. Follows by Theorem 3.1. ∎

We note that in the case of regular and almost regular graphs (see Section 2.3) the integral of Corollary 3.3 is the same as the one evaluated by McKay and Wormald [18].

3.4. Plan of the Proof of Theorem 1.4

We use the integral representation of Corollary 3.3. Let us define subsets $\mathcal{U}, \mathcal{W} \subset \Pi$ by

$$\mathcal{U} = \left\{ (\tau_1, \ldots, \tau_n) : \ |\tau_j| \leq \frac{\ln n}{\sqrt{n}} \text{ for } j = 1, \ldots, n \right\}$$

and

$$\mathcal{W} = \left\{ (\tau_1, \ldots, \tau_n) : \ \left|\tau_j - \sigma_j \pi\right| \leq \frac{\ln n}{\sqrt{n}} \text{ for some } \sigma_j = \pm 1 \text{ and } j = 1, \ldots, n \right\}.$$

We show that the integral of $F(t)$ over $\Pi \setminus (\mathcal{U} \cup \mathcal{W})$ is asymptotically negligible. Namely, in Section 8 we prove that the integral

$$\int_{\Pi \setminus (\mathcal{U} \cup \mathcal{W})} |F(t)|\,dt$$

is asymptotically negligible compared to the integral

$$\int_{\mathcal{U}} |F(t)|dt. \tag{3.4.1}$$

It is easy to show that

$$\int_{\mathcal{U}} F(t)dt = \int_{\mathcal{W}} F(t)dt,$$

provided $d_1 + \cdots + d_n$ is even.

In Section 7, we evaluate

$$\int_{\mathcal{U}} F(t)dt. \tag{3.4.2}$$

In particular, we show that the integrals (3.4.1) and (3.4.2) have the same order of magnitude and so the integral of $F(t)$ outside of $\mathcal{U} \cup \mathcal{W}$ is indeed asymptotically irrelevant.

From (3.2) one can deduce that asymptotically as $n \longrightarrow +\infty$,

$$F(t) \approx \exp\{-q(t) + if(t) + h(t)\} \quad \text{for } t \in \mathcal{U},$$

where q is defined by (1.3.1) and f and h are defined by (1.3.2).

Let us consider the Gaussian probability measure in $\mathbb{R}^n$ with density proportional to e^{-q}. In Section 6, we prove that with respect to that measure

$$h(t) \approx \mathbf{E}h = \nu \quad \text{almost everywhere in } \mathcal{U}. \tag{3.4.3}$$

This allows us to conclude that

$$\int_{\mathcal{U}} \exp\{-q(t) + if(t) + h(t)\}dt \approx e^{\nu} \int_{\mathcal{U}} \exp\{-q(t) + if(t)\}dt.$$

In Section 5, we prove that asymptotically, as $n \longrightarrow +\infty$, function f is a Gaussian random variable, so

$$\begin{aligned} \int_{\mathcal{U}} \exp\{-q(t) + if(t)\}dt &\approx \int_{\mathbb{R}^n} \exp\{-q(t) + if(t)\}dt \\ &\approx \exp\left\{-\frac{1}{2}\mathbf{E}f^2\right\} \int_{\mathbb{R}^n} e^{-q(t)}dt, \end{aligned} \tag{3.4.4}$$

which concludes the evaluation of (3.4.2).

The crucial consideration used in proving (3.4.3) and (3.4.4) is that with respect to the Gaussian probability measure in $\mathbb{R}^n$ with density proportional to e^{-q}, the coordinate functions $\tau_1, \ldots, \tau_n$ are weakly correlated, that is,

$$\begin{aligned} |\mathbf{E}\tau_j\tau_k| &= O\left(\frac{1}{n^2}\right) \quad \text{for } j \neq k \quad \text{and} \\ \mathbf{E}\tau_j^2 &= O\left(\frac{1}{n}\right) \quad \text{for } j = 1, \ldots, n. \end{aligned} \tag{3.4.5}$$

We prove (3.4.5) in Section 4, where we essentially use the δ-tameness assumption.

3.5. Notation

By γ, sometimes with an index or a list of parameters, we denote a positive constant depending only on the listed parameters. The most common appearance will be $\gamma(\delta)$, a positive constant depending only on the δ-tameness constant δ.

As usual, for two functions g_1 and g_2, where g_2 is non-negative, we write $g_1 = O(g_2)$ if $|g_1| \leq \gamma g_2$ and $g_1 = \Omega(g_2)$ if $g_1 \geq \gamma g_2$ for some $\gamma > 0$.

4. CORRELATIONS

Let $z = (\zeta_{\{j,k\}})$ be the maximum entropy matrix as defined in Section 1.2. We assume that

$$0 < \zeta_{\{j,k\}} < 1 \quad \text{for all } j \neq k.$$

We define the quadratic form $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ by

$$q(t) = \frac{1}{2} \sum_{\{j,k\}} \left(\zeta_{\{j,k\}} - \zeta^2_{\{j,k\}} \right) (\tau_j + \tau_k)^2 \quad \text{for } t = (\tau_1, \ldots, \tau_n).$$

For $n > 2$ the quadratic form q is strictly positive definite. We consider the Gaussian probability measure on $\mathbb{R}^n$ with density proportional to e^{-q}. We consider a point $t = (\tau_1, \ldots, \tau_n)$ as a random vector and $\tau_1, \ldots, \tau_n$ as random variables.

The main result of this section is as follows.

Theorem 4.1. *For any* $0 < \delta \leq 1/2$ *there exists* $\gamma(\delta) > 0$ *such that the following holds. Suppose that*

$$\delta \leq \zeta_{\{j,k\}} \leq 1 - \delta \quad \textit{for all } j \neq k.$$

Then

$$\left| \mathbf{E} \tau_j \tau_k \right| \leq \frac{\gamma(\delta)}{n^2} \quad \textit{provided } j \neq k \quad \textit{and}$$
$$\mathbf{E} \tau_j^2 \leq \frac{\gamma(\delta)}{n} \quad \textit{for } j = 1, \ldots, n.$$

We will often consider the following situation. Let $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a positive definite quadratic form. We consider the Gaussian probability measure in $\mathbb{R}^n$ with density proportional to $e^{-\psi}$. For a polynomial (random variable) $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ we denote by $\mathbf{E}(f; \psi)$ its expectation with respect to the measure. For a subspace $L \subset \mathbb{R}^n$, we consider the restriction $\psi|L$ of ψ onto L and the Gaussian probability measure on L with density proportional to $e^{-\psi|L}$. For a polynomial $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, we denote by $\mathbf{E}(f;\ \psi|L)$ the expectation of the restriction $f : L \longrightarrow \mathbb{R}$ with respect to that Gaussian probability measure on L. We will use the following standard fact: suppose that $\mathbb{R}^n = L_1 \oplus L_2$ is a decomposition of $\mathbb{R}^n$ into the direct sum of orthogonal subspaces such that

$$\psi(t_1 + t_2) = \psi(t_1) + \psi(t_2) \quad \text{for all } t_1 \in L_1 \text{ and } t_2 \in L_2,$$

so that the coordinates $t_1 \in L_1$ and $t_2 \in L_2$ of the point $t = t_1 + t_2$, $t \in \mathbb{R}^n$ are independent. Let $\ell_1, \ell_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ be linear functions. Then

$$\mathbf{E}(\ell_1\ell_2;\ \psi) = \mathbf{E}(\ell_1\ell_2;\ \psi|L_1) + \mathbf{E}(\ell_1\ell_2;\ \psi|L_2).$$

Indeed, writing $t = t_1+t_2$ with $t_1 \in L_1$ and $t_2 \in L_2$ and noting that $\ell_{1,2}(t) = \ell_{1,2}(t_1)+\ell_{1,2}(t_2)$, we obtain

$$\begin{aligned}
\mathbf{E}(\ell_1(t)\ell_2(t);\ \psi) &= \mathbf{E}(\ell_1(t_1)\ell_2(t_1);\ \psi) + \mathbf{E}(\ell_1(t_1)\ell_2(t_2);\ \psi)\\
&\quad + \mathbf{E}(\ell_1(t_2)\ell_2(t_1);\ \psi) + \mathbf{E}(\ell_1(t_2)\ell_2(t_2);\ \psi)\\
&= \mathbf{E}(\ell_1\ell_2;\ \psi|L_1) + 2\mathbf{E}(\ell_1;\ \psi)\mathbf{E}(\ell_2;\ \psi) + \mathbf{E}(\ell_1\ell_2;\ \psi|L_2)\\
&= \mathbf{E}(\ell_1\ell_2;\ \psi|L_1) + \mathbf{E}(\ell_1\ell_2;\ \psi|L_2).
\end{aligned}$$

We deduce Theorem 4.1 from the following result.

Proposition 4.2. *Let $n > 2$ and let $\xi_{\{j,k\}}$, $1 \le j \ne k \le n$ be a set of numbers such that*

$$\alpha \le \xi_{\{j,k\}} \le \beta \quad \textit{for all } j,k$$

and some $\beta > \alpha > 0$.

Let

$$\sigma_k = \sum_{j:j\ne k} \xi_{\{j,k\}} \quad \textit{for } k = 1,\dots,n.$$

Let us consider the quadratic form $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by

$$\psi(t) = \frac{1}{2}\sum_{\{j,k\}} \xi_{\{j,k\}} \left(\frac{\tau_j}{\sqrt{\sigma_j}} + \frac{\tau_k}{\sqrt{\sigma_k}}\right)^2 \quad \textit{for } t = (\tau_1,\dots,\tau_n),$$

where the sum is taken over all unordered pairs of indices $1 \le j \ne k \le n$. Then ψ is a positive definite quadratic form and we consider the Gaussian probability measure in $\mathbb{R}^n$ with density proportional to $e^{-\psi}$.

Let

$$\epsilon = \frac{\alpha}{\beta}.$$

Then for $n > 2/\epsilon$ we have

$$\begin{aligned}
|\mathbf{E}\tau_j\tau_k| &\le \frac{n^2}{\epsilon^{5/2}(n-\epsilon)(n\epsilon-2)(n-1)} + \frac{3}{2\epsilon n} \quad \textit{provided } j \ne k \textit{ and}\\
\left|\mathbf{E}\tau_j^2 - 1\right| &\le \frac{n^2}{\epsilon^{5/2}(n-\epsilon)(n\epsilon-2)(n-1)} + \frac{3}{2\epsilon n} \quad \textit{for } j = 1,\dots,n.
\end{aligned}$$

Proof. Clearly, ψ is positive definite. Let

$$v = (\sqrt{\sigma_1},\dots,\sqrt{\sigma_n}).$$

Then v is an eigenvector of ψ with eigenvalue 1. Indeed, the gradient of ψ at $t = v$ is $2v$:

$$\frac{\partial}{\partial \tau_j} \psi(t)\bigg|_{t=v} = \frac{1}{\sqrt{\sigma_j}} \sum_{k:\ k \neq j} 2\xi_{\{j,k\}} = 2\sqrt{\sigma_j} \quad \text{for } j = 1, \ldots, n.$$

Let

$$L = v^{\perp} \subset \mathbb{R}^n$$

be the orthogonal complement to v. Hence L is defined in $\mathbb{R}^n = \{(\tau_1, \ldots, \tau_n)\}$ by the equation

$$\sum_{j=1}^{n} \tau_j \sqrt{\sigma_j} = 0.$$

We write

$$\psi(t) = \frac{1}{2} \sum_{j=1}^{n} \tau_j^2 + \sum_{\{j,k\}} \frac{\xi_{\{j,k\}}}{\sqrt{\sigma_j \sigma_k}} \tau_j \tau_k.$$

We remark that

$$\alpha(n-1) \leq \sigma_j \leq \beta(n-1) \quad \text{for } j = 1, \ldots, n. \tag{4.2.1}$$

Let

$$\omega = \sum_{j=1}^{n} \sigma_j \geq \alpha n(n-1)$$

and let us define the quadratic form $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$ by

$$\phi(t) = \frac{1}{\omega} \left(\sum_{j=1}^{n} \tau_j \sqrt{\sigma_j} \right)^2 \quad \text{for } t = (\tau_1, \ldots, \tau_n).$$

Hence $\phi(t)$ is a form of rank 1 and v is an eigenvector of ϕ with eigenvalue 1.

We define a perturbation

$$\tilde{\psi} = \psi - \frac{\epsilon^2}{2} \phi \quad \text{for } \epsilon = \frac{\alpha}{\beta}.$$

Hence $\tilde{\psi}$ is a positive definite quadratic form such that

$$\tilde{\psi}(t) = \psi(t) \quad \text{for all } t \in L,$$

and v is an eigenvector of $\tilde{\psi}$ with eigenvalue $1 - \epsilon^2/2$.

Let us consider the Gaussian probability measure on $\mathbb{R}^n$ with density proportional to $e^{-\tilde{\psi}}$. Our immediate goal is to estimate the covariances $\mathbf{E}\tau_j\tau_k$ with respect to that measure.

Denoting by $\langle \cdot, \cdot \rangle$ the standard scalar product in $\mathbb{R}^n$, we can write

$$\tilde{\psi}(t) = \frac{1}{2} \langle (I + Q)t,\ t \rangle,$$

where I is the $n \times n$ identity matrix and $Q = (q_{jk})$ is an $n \times n$ symmetric matrix such that v is an eigenvector of Q with eigenvalue $1 - \epsilon^2$. We have

$$q_{jk} = \frac{\xi_{\{j,k\}}}{\sqrt{\sigma_j \sigma_k}} - \epsilon^2 \frac{\sqrt{\sigma_j \sigma_k}}{\omega} \quad \text{for } j \neq k \quad \text{and}$$
$$q_{jj} = -\frac{\epsilon^2 \sigma_j}{\omega} \quad \text{for } j = 1, \ldots, n.$$

It follows by (4.2.1) that

$$\frac{1}{\epsilon(n-1)} \geq q_{jk} \geq 0 \quad \text{for } j \neq k \quad \text{and}$$
$$0 \geq q_{jj} \geq -\frac{\epsilon}{n} \quad \text{for } j = 1, \ldots, n.$$

The covariance matrix $R = (\mathbf{E}\tau_j\tau_k;\ \tilde{\psi})$ of the Gaussian measure with density proportional to $e^{-\tilde{\psi}}$ is

$$(I+Q)^{-1} = \left(\left(1 - \frac{\epsilon}{n}\right) I + \left(\frac{\epsilon}{n} I + Q\right)\right)^{-1} = \left(1 - \frac{\epsilon}{n}\right)^{-1} (I+P)^{-1},$$
$$\text{where} \quad P = \left(1 - \frac{\epsilon}{n}\right)^{-1} \left(\frac{\epsilon}{n} I + Q\right).$$

Hence $P = (p_{jk})$ is a symmetric matrix such that

$$0 \leq p_{jk} \leq \left(1 - \frac{\epsilon}{n}\right)^{-1} \frac{1}{\epsilon(n-1)} \quad \text{for all } j, k. \tag{4.2.2}$$

Furthermore, v is an eigenvector of P with eigenvalue $(1 - \epsilon^2 + \epsilon/n)/(1 - \epsilon/n)$, so

$$Pv = \lambda v \quad \text{for } \lambda = \left(1 - \frac{\epsilon}{n}\right)^{-1} \left(1 - \epsilon^2 + \frac{\epsilon}{n}\right). \tag{4.2.3}$$

Let us bound the entries of a positive integer power $P^d = (p_{jk}^{(d)})$ of P. Let

$$\kappa = \left(1 - \frac{\alpha}{\beta n}\right)^{-1} \frac{\beta}{\alpha^{3/2}(n-1)^{3/2}} \quad \text{and let } y = \kappa v, \quad y = (\eta_1, \ldots, \eta_n).$$

By (4.2.1) and (4.2.2), we have

$$p_{jk} \leq \eta_j \quad \text{for all } j, k. \tag{4.2.4}$$

Also, by (4.2.1) we have

$$\eta_j \leq \left(1 - \frac{\epsilon}{n}\right)^{-1} \frac{1}{\epsilon^{3/2}(n-1)} \quad \text{for all } j. \tag{4.2.5}$$

Furthermore, y is an eigenvector of P with eigenvalue λ defined by (4.2.3), and hence y is an eigenvector of the d-th power $P^d = (p_{jk}^{(d)})$ with eigenvalue λ^d. Combining this with (4.2.4) and (4.2.5), for $d \geq 0$ we obtain

$$p_{jk}^{(d+1)} = \sum_{m=1}^{n} p_{jm}^{(d)} p_{mk} \leq \sum_{m=1}^{n} p_{jm}^{(d)} \eta_m = \lambda^d \eta_j \leq \lambda^d \left(1 - \frac{\epsilon}{n}\right)^{-1} \frac{1}{\epsilon^{3/2}(n-1)}.$$

We note that for $n > 2/\epsilon$ we have $0 < \lambda < 1$. Consequently, the series

$$(I+P)^{-1} = I + \sum_{d=1}^{+\infty} (-1)^d P^d$$

converges absolutely and we can bound the entries of the matrix

$$R = (I+Q)^{-1} = \left(1 - \frac{\epsilon}{n}\right)^{-1} (I+P)^{-1},$$

$R = (r_{jk})$ by

$$\begin{aligned} |r_{jk}| &\leq \left(1 - \frac{\epsilon}{n}\right)^{-2} \frac{1}{\epsilon^{3/2}(n-1)} \frac{1}{1-\lambda} \quad \text{if } j \neq k \quad \text{and} \\ |r_{jj} - 1| &\leq \left(1 - \frac{\epsilon}{n}\right)^{-2} \frac{1}{\epsilon^{3/2}(n-1)} \frac{1}{1-\lambda} \quad \text{for } j = 1, \ldots, n. \end{aligned}$$

We have

$$\frac{1}{1-\lambda} = \left(1 - \frac{\epsilon}{n}\right)\left(\epsilon^2 - \frac{2\epsilon}{n}\right)^{-1}.$$

Since R is the covariance matrix of the Gaussian probability measure with density proportional to $e^{-\tilde{\psi}}$, we obtain

$$\begin{aligned} |\mathbf{E}(\tau_j \tau_k;\ \tilde{\psi})| &\leq \frac{n^2}{\epsilon^{5/2}(n-\epsilon)(n\epsilon - 2)(n-1)} \quad \text{provided } j \neq k \quad \text{and} \\ \left|\mathbf{E}\left(\tau_j^2;\ \tilde{\psi}\right) - 1\right| &\leq \frac{n^2}{\epsilon^{5/2}(n-\epsilon)(n\epsilon - 2)(n-1)} \quad \text{for } j = 1, \ldots, n. \end{aligned} \tag{4.2.6}$$

Now we go back to the form ψ and the Gaussian probability measure with density proportional to $e^{-\psi}$. Since v is an eigenvector of both ψ and $\tilde{\psi}$, since $L = u^{\perp}$ and since ψ and $\tilde{\psi}$ coincide on L, for any linear functions $\ell_1, \ell_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$, we have

$$\begin{aligned} \mathbf{E}(\ell_1\ell_2;\ \psi) &= \mathbf{E}(\ell_1\ell_2;\ \psi|L) + \mathbf{E}(\ell_1\ell_2;\ \psi|\text{span}(v)) \\ &= \mathbf{E}(\ell_1\ell_2;\ \tilde{\psi}|L) + \mathbf{E}(\ell_1\ell_2;\ \psi|\text{span}(v)) \\ &= \mathbf{E}(\ell_1\ell_2;\ \tilde{\psi}) - \mathbf{E}(\ell_1\ell_2;\ \tilde{\psi}|\text{span}(v)) + \mathbf{E}(\ell_1\ell_2;\ \psi|\text{span}(v)). \end{aligned} \tag{4.2.7}$$

We note that the gradient of the coordinate function τ_j restricted to span(v) is $\sqrt{\sigma_j/\omega}$. Since v is an eigenvector of ψ with eigenvalue 1 and an eigenvector of $\tilde{\psi}$ with eigenvalue $1-\epsilon^2/2$, we have

$$\begin{aligned} \mathbf{E}(\tau_j\tau_k; \psi|\text{span}(v)) &= \frac{\sqrt{\sigma_j\sigma_k}}{2\omega} \quad \text{and} \\ \mathbf{E}(\tau_j\tau_k; \tilde{\psi}|\text{span}(v)) &= \frac{\sqrt{\sigma_j\sigma_k}}{(2-\epsilon^2)\omega} \quad \text{for all } j, k. \end{aligned}$$

By (4.2.1) we have

$$|\mathbf{E}(\tau_j\tau_k;\ \psi|\mathrm{span}(v))| \le \frac{1}{2\epsilon n} \quad \text{and} \quad |\mathbf{E}(\tau_j\tau_k;\ \tilde{\psi}|\mathrm{span}(v))| \le \frac{1}{\epsilon n}.$$

The proof now follows by (4.2.6) and (4.2.7). ∎

Now we are ready to prove Theorem 4.1.

Proof of Theorem 4.1. Let us define

$$\xi_{\{j,k\}} = \zeta_{\{j,k\}} - \zeta^2_{\{j,k\}} \quad \text{for all } j \ne k$$

and let us choose $\alpha = \delta - \delta^2$ and $\beta = 1/4$ in Proposition 4.2. We define σ_j and ψ as in Proposition 4.2 and consider a linear transformation

$$(\tau_1, \dots, \tau_n) \longmapsto (\tau_1\sqrt{\sigma_1}, \dots, \tau_n\sqrt{\sigma_n}).$$

Then the push-forward of the Gaussian probability measure with density proportional to e^{-q} is the Gaussian probability measure with density proportional to $e^{-\psi}$. Therefore,

$$\mathbf{E}(\tau_j\tau_k;\ q) = \frac{1}{\sqrt{\sigma_j\sigma_k}}\mathbf{E}(\tau_j\tau_k;\ \psi).$$

Since

$$\sigma_j \ge \delta\alpha(n-1) \quad \text{for } j = 1, \dots, n,$$

The proof follows by Proposition 4.2. ∎

We will need the following lemma.

Lemma 4.3. *Let $q_0 : \mathbb{R}^n \longrightarrow \mathbb{R}$, $n \ge 2$, be the quadratic form defined by the formula*

$$q_0(t) = \frac{1}{2}\sum_{\{j,k\}}(\tau_j + \tau_k)^2 \quad \text{for } t = (\tau_1, \dots, \tau_n).$$

Then the eigenspaces of q_0 are as follows: the 1-dimensional eigenspace E_1 with eigenvalue $n-1$ spanned by the vector $u = (1, \dots, 1)$ and the $(n-1)$-dimensional eigenspace $E_2 = u^{\perp}$ with eigenvalue $(n-2)/2$.

Proof. We have

$$\frac{\partial}{\partial \tau_k} q(t)\bigg|_{t=u} = 2n - 2.$$

Hence the gradient of $q_0(t)$ at $t = u$ is $(2n-2)u$, so u is an eigenvector with eigenvalue $(n-1)$. For $t \in u^{\perp}$ we have $\tau_1 + \cdots + \tau_n = 0$ and hence

$$\frac{\partial}{\partial \tau_k} q(t) = \sum_{j:\ j \ne k}(\tau_j + \tau_k) = (n-2)\tau_k.$$

Therefore, the gradient of $q_0(t)$ at $t \in u^{\perp}$ is $(n-2)t$, and so t is an eigenvector with eigenvalue $(n-2)/2$. ∎

5. THE THIRD DEGREE TERM

The main result of this section is the following theorem.

Theorem 5.1. *For unordered pairs $\{j,k\}$, $1 \leq j \neq k \leq n$, let $u_{\{j,k\}}$ be Gaussian random variables such that*

$$\mathbf{E}u_{\{j,k\}} = 0 \quad \textit{for all } j,k.$$

Suppose further that for some $\theta > 0$ we have

$$\mathbf{E}u^2_{\{j,k\}} \leq \frac{\theta}{n} \quad \textit{for all } j,k$$

and that

$$|\mathbf{E}u_{\{j_1,k_1\}}u_{\{j_2,k_2\}}| \leq \frac{\theta}{n^2} \quad \textit{provided } \{j_1,k_1\} \cap \{j_2,k_2\} = \emptyset.$$

Let

$$U = \sum_{\{j,k\}} u^3_{\{j,k\}}.$$

Then for some constant $\gamma(\theta) > 0$ and any $0 < \epsilon < 1/2$ we have

$$\left|\mathbf{E}\exp\{iU\} - \exp\left\{-\frac{1}{2}\mathbf{E}U^2\right\}\right| \leq \epsilon$$

provided

$$n \geq \left(\frac{1}{\epsilon}\right)^{\gamma(\theta)}.$$

Furthermore,

$$\mathbf{E}U^2 \leq \gamma(\theta)$$

for some $\gamma(\theta) > 0$. Here $i = \sqrt{-1}$.

We apply Theorem 5.1 in the following situation. Let $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ be the quadratic form defined by (1.3.1).

Let us consider the Gaussian probability measure on $\mathbb{R}^n$ with density proportional to e^{-q}. We define random variables $u_{\{j,k\}}$ by

$$u_{\{j,k\}}(t) = \sqrt[3]{\frac{1}{6}\zeta_{\{j,k\}}(1-\zeta_{\{j,k\}})(2\zeta_{\{j,k\}}-1)(\tau_j+\tau_k)} \quad \text{for } t = (\tau_1,\ldots,\tau_n).$$

Then for the function $f(t)$ defined by (1.3.2) we have

$$f = \sum_{\{j,k\}} u^3_{\{j,k\}}.$$

In this section, all implied constants in the "O" notation are absolute.

Our main tool is Wick's formula for the expectation of the product of random Gaussian variables.

5.2. Wick's Formula

Let $w_1, \dots, w_l$ be Gaussian random variables such that

$$\mathbf{E}w_1 = \ldots = \mathbf{E}w_l = 0.$$

Then

$$\mathbf{E}(w_1 \cdots w_l) = 0 \quad \text{if } l \text{ is odd} \quad \text{and}$$
$$\mathbf{E}(w_1 \ldots w_l) = \sum (\mathbf{E}w_{i_1}w_{i_2}) \cdots (\mathbf{E}w_{i_{l-1}}w_{i_l}) \quad \text{if } l = 2r \text{ is even,}$$

where the sum is taken over all $(2r)!/r!2^r$ unordered partitions of the set of indices $\{1, \dots, l\}$ into $r = l/2$ pairwise disjoint unordered pairs $\{i_1, i_2\}, \dots, \{i_{l-1}, i_l\}$, see for example, [20]. Such a partition is called a *matching* of the random variables $w_1, \dots, w_l$ and we say that w_i and w_j are *matched* if they form a pair in the matching.

In particular,

$$\mathbf{E}w^{2r} = \frac{(2r)!}{r!2^r}(\mathbf{E}w^2)^r \tag{5.2.1}$$

for a centered Gaussian random variable w. We will also use that

$$\mathbf{E}w_1^3w_2^3 = 9\left(\mathbf{E}w_1^2\right)\left(\mathbf{E}w_2^2\right)(\mathbf{E}w_1w_2) + 6(\mathbf{E}w_1w_2)^3 \tag{5.2.2}$$

and later in Section 6 that

$$\begin{aligned}\mathbf{cov}\left(w_1^4, w_2^4\right) &= \mathbf{E}\left(w_1^4w_2^4\right) - \left(\mathbf{E}w_1^4\right)\left(\mathbf{E}w_2^4\right)\\ &= 72(\mathbf{E}w_1w_2)^2\left(\mathbf{E}w_1^2\right)\left(\mathbf{E}w_2^2\right) + 24(\mathbf{E}w_1w_2)^4.\end{aligned} \tag{5.2.3}$$

5.3. Representing Monomials by Graphs

Let $x_{\{j,k\}} : 1 \le j \ne k \le n$ be formal commuting variables. We interpret a monomial in $x_{\{j,k\}}$ as a weighted graph as follows. Let K_n be the complete graph with vertices $1, \dots, n$ and edges $\{j, k\}$ for $1 \le j \ne k \le n$. A *weighted graph* G is a set of edges $\{j, k\}$ of K_n with positive integer weights $\alpha_{\{j,k\}}$ on them. The set of vertices of G consists of all vertices of the edges of G. With G, we associate a monomial

$$m_G(x) = \prod_{\{j,k\} \in G} x_{\{j,k\}}^{\alpha_{\{j,k\}}}.$$

The *weight* of G is the degree of $m_G(x)$, that is, $\sum_{\{j,k\} \in G} \alpha_{\{j,k\}}$. We observe that for any p there are not more than $r^{O(r)}n^p$ distinct weighted graphs G of weight $2r$ on p vertices.

In what follows, given a set of random variables, we construct auxiliary Gaussian random variables with the same matrix of covariances. This is always possible since the matrix of covariances is positive semi-definite.

Our proof of Theorem 5.1 is based on the following combinatorial lemma.

Lemma 5.4. *For the Gaussian random variables $u_{\{j,k\}}$ of Theorem 5.1, let us introduce auxiliary Gaussian random variables $v_{\{j,k\}}$ such that*

$$\mathbf{E}v_{\{j,k\}} = 0 \quad \textit{for all } 1 \le j \ne k \le n \qquad \textit{and}$$
$$\mathbf{E}v_{\{j_1,k_1\}}v_{\{j_2,k_2\}} = \mathbf{E}u^3_{\{j_1,k_1\}}u^3_{\{j_2,k_2\}} \quad \textit{for all } 1 \le j_1 \ne k_1,\ j_2 \ne k_2 \le n.$$

Given a weighted graph G of weight $2r$, $r > 1$, let us represent it as a vertex-disjoint union

$$G = G_0 \cup G_1,$$

where G_0 consists of s isolated edges of weight 1 each and G_1 is a graph with no isolated edges of weight 1 (we may have $s = 0$ and G_0 empty).

Then

1. *We have*

$$\left|\mathbf{E}m_G\left(u^3_{\{j,k\}}\ :\ 1 \le j \ne k \le n\right)\right|,\ |\mathbf{E}m_G(v_{\{j,k\}}\ :\ 1 \le j \ne k \le n)| \le \frac{r^{O(r)}\theta^{3r}}{n^{3r+s/2}}$$

if s is even and

$$\left|\mathbf{E}m_G\left(u^3_{\{j,k\}}\ :\ 1 \le j \ne k \le n\right)\right|,\ |\mathbf{E}m_G(v_{\{j,k\}}\ :\ 1 \le j \ne k \le n)| \le \frac{r^{O(r)}\theta^{3r}}{n^{3r+(s+1)/2}}$$

if s is odd.

2. *If s is even and G_1 is a vertex-disjoint union of $r - s/2$ connected components, each consisting of a pair of edges of weight 1 sharing exactly one common vertex, then*

$$\left|\mathbf{E}m_G\left(u^3_{\{j,k\}}\ :\ 1 \le j \ne k \le n\right) - \mathbf{E}m_G(v_{\{j,k\}}\ :\ 1 \le j \ne k \le n)\right| \le \frac{r^{O(r)}\theta^{3r}}{n^{3r+s/2+1}}.$$

3. *If s is even and G_1 is a vertex-disjoint union of $r - s/2$ connected components, each consisting of a pair of edges of weight 1 sharing exactly one common vertex, then G has precisely $3r + s/2$ vertices. In all other cases, G has strictly fewer than $3r + s/2$ vertices.*

Proof. If $\{j_1, k_1\} \cap \{j_2, k_2\} = \emptyset$ we say that the pair of variables $u_{\{j_1,k_1\}}, u_{\{j_2,k_2\}}$ and the pair of variables $v_{\{j_1,k_1\}}, v_{\{j_2,k_2\}}$ are *weakly correlated*. If $\{j_1, k_1\} \cap \{j_2, k_2\} \ne \emptyset$ we say that the pairs of variables are *strongly correlated*. Pairs of variables indexed by edges in different connected components of G are necessarily weakly correlated.

To prove Part (1) we use Wick's formula of Section 5.2. By (5.2.2), we obtain

$$\begin{aligned} \mathbf{E}(v_{\{j_1,k_1\}}v_{\{j_2,k_2\}}) &= O\left(\frac{\theta^3}{n^4}\right) \quad \text{if the pair } v_{\{j_1,k_1\}}, v_{\{j_2,k_2\}} \text{ is weakly correlated,} \\ \mathbf{E}(v_{\{j_1,k_1\}}v_{\{j_2,k_2\}}) &= O\left(\frac{\theta^3}{n^3}\right) \quad \text{if the pair } v_{\{j_1,k_1\}}, v_{\{j_2,k_2\}} \text{ is strongly correlated.} \end{aligned} \tag{5.4.1}$$

Since for each isolated edge $\{j_1, k_1\} \in G_0$ variable $v_{\{j_1,k_1\}}$ has to be matched with variable $v_{\{j_2,k_2\}}$ indexed by an edge $\{j_2, k_2\}$ in a different connected component, we conclude that every matching of the set

$$\{v_{\{j,k\}} : \{j, k\} \in G\} \tag{5.4.2}$$

contains at least $s/2$ weakly correlated pairs and hence

$$|\mathbf{E}m_G(v_{\{j,k\}}\ :\ 1 \le j \ne k \le n)| \le r^{O(r)}\left(\frac{\theta^3}{n^4}\right)^{s/2}\left(\frac{\theta^3}{n^3}\right)^{r-s/2}.$$

Moreover, if s is odd, then the number of weakly correlated pairs is at least $(s+1)/2$ and hence

$$|\mathbf{E}m_G(v_{\{j,k\}} : 1 \leq j \neq k \leq n)| \leq r^{O(r)} \left(\frac{\theta^3}{n^4}\right)^{(s+1)/2} \left(\frac{\theta^3}{n^3}\right)^{r-(s+1)/2}.$$

Similarly, since for each isolated edge $\{j_1, k_1\} \in G_0$ at least one copy of the variable $u_{\{j_1,k_1\}}$ has to be matched with a copy of variable $u_{\{j_2,k_2\}}$ indexed by an edge in a different connected component, we conclude that every matching of the multiset

$$\{u_{\{j,k\}}, u_{\{j,k\}}, u_{\{j,k\}} : \{j,k\} \in G\} \tag{5.4.3}$$

contains at least $s/2$ weakly correlated pairs, and hence

$$\left|\mathbf{E}m_G\left(u^3_{\{j,k\}} : 1 \leq j \neq k \leq n\right)\right| \leq r^{O(r)} \left(\frac{\theta}{n^2}\right)^{s/2} \left(\frac{\theta}{n}\right)^{3r-s/2}.$$

Moreover, if s is odd, then the number of weakly correlated pairs is at least $(s+1)/2$ and hence

$$\left|\mathbf{E}m_G\left(u^3_{\{j,k\}} :\ 1 \leq j \neq k \leq n\right)\right| \leq r^{O(r)} \left(\frac{\theta}{n^2}\right)^{(s+1)/2} \left(\frac{\theta}{n}\right)^{3r-(s+1)/2}.$$

This concludes the proof of Part (1).

To prove Part (2), let us define $\Sigma_v(G)$ as the sum in the Wick's formula for $\mathbf{E}m_G(v_{\{j,k\}} : 1 \leq j \neq k \leq n)$ taken over all matchings of the set (5.4.2) of the following structure: we split the edges of G into r pairs, pairing each isolated edge with another isolated edge and pairing each edge in a connected component of G consisting of two edges with the remaining edge in the same connected component. Then we match every variable $v_{\{j_1,k_1\}}$ with the variable $v_{\{j_2,k_2\}}$ such that $\{j_1, k_1\}$ is paired with $\{j_2, k_2\}$. Reasoning as in the proof of Part (1), we conclude that

$$|\mathbf{E}m_G(v_{\{j,k\}} : 1 \leq j \neq k \leq n) - \Sigma_v(G)| \leq \frac{r^{O(r)}\theta^{3r}}{n^{3r+s/2+1}},$$

since every matching of the set (5.4.2) which is not included in $\Sigma_v(G)$ contains at least $s/2 + 1$ weakly correlated pairs.

Similarly, let us define $\Sigma_u(G)$ as the sum in the Wick's formula for $\mathbf{E}m_G(u^3_{\{j,k\}} :\ 1 \leq j \neq k \leq n)$ taken over all matchings of the multiset (5.4.3) of the following structure: we split the edges of G into r pairs as above and match every copy of variable $u_{\{j_1,k_1\}}$ with a copy of variable $u_{\{j_2,k_2\}}$ indexed by an edge in the same pair (in particular, we may match copies of the same variable). Reasoning as in the proof of Part (1), we conclude that

$$\left|\mathbf{E}m_G\left(u^3_{\{j,k\}} :\ 1 \leq j \neq k \leq n\right) - \Sigma_u(G)\right| \leq \frac{r^{O(r)}\theta^{3r}}{n^{3r+s/2+1}},$$

since every matching of the multiset (5.4.3) which is not included in $\Sigma_u(G)$ contains at least $s/2 + 1$ weakly correlated pairs.

The proof of Part (2) follows since

$$\Sigma_u(G) = \Sigma_v(G)$$

by Wick's formula.

To prove Part (3), we note that a connected weighted graph G of weight e contains a spanning tree with at most e edges and hence has at most $e+1$ vertices. In particular, a connected graph G of weight e contains fewer than $3e/2$ vertices unless G is an isolated edge of weight 1 or a pair of edges of weight 1 each, sharing one common vertex. Therefore, G has at most

$$2s + \frac{3}{2}(2r - s) = 3r + \frac{s}{2}$$

vertices and strictly fewer vertices, unless s is even and the connected components of G_1 are pairs of edges of weight 1 each sharing one common vertex. ∎

5.5. Proof of Theorem 5.1

Let $v_{\{j,k\}}$ be Gaussian random variables defined in Lemma 5.4 and let

$$V = \sum_{\{j,k\}} v_{\{j,k\}}.$$

Since there are $O(n^3)$ strongly correlated pairs v_{j_1,k_1}, v_{j_2,k_2} and there are $O(n^4)$ weakly correlated pairs, by (5.4.1) we have

$$\mathbf{E}V^2 = \mathbf{E}U^2 = O(\theta^3). \tag{5.5.1}$$

Since V is a Gaussian random variable, we have

$$\mathbf{E}e^{iV} = \exp\left\{-\frac{1}{2}\mathbf{E}V^2\right\} = \exp\left\{-\frac{1}{2}\mathbf{E}U^2\right\}. \tag{5.5.2}$$

Our goal is to show that $\mathbf{E}e^{iV}$ and $\mathbf{E}e^{iU}$ are asymptotically equal as $n \longrightarrow +\infty$.

By symmetry, the odd moments of U and V are 0:

$$\mathbf{E}U^k = \mathbf{E}V^k = 0 \quad \text{if } k > 0 \text{ is odd.} \tag{5.5.3}$$

The even moments of U and V can be expressed as

$$\mathbf{E}U^{2r} = \sum_G a_G \mathbf{E}m_G\left(u^3_{\{j,k\}} : 1 \le j \ne k \le n\right)$$
$$\mathbf{E}V^{2r} = \sum_G a_G \mathbf{E}m_G(v_{\{j,k\}} : \ 1 \le j \ne k \le n),$$

where the sum is taken over all weighted graphs G of weight $2r$ and

$$1 \le a_G \le (2r)!.$$

Let $\mathcal{G}_{2r}$ be the set of weighted graphs G of weight $2r$ whose connected components are an even number s of isolated edges of weight 1 and $r - s/2$ pairs of edges of weight 1 sharing one common vertex. Since there are no more than $r^{O(r)}n^p$ distinct weighted graphs of weight $2r$ with p vertices, by Parts (1) and (3) of Lemma 5.4, we have

$$\left|\mathbf{E}U^{2r} - \sum_{G\in\mathcal{G}_{2r}} a_G \mathbf{E} m_G\big(u^3_{\{j,k\}}:\ 1 \le j \ne k \le n\big)\right| \le \frac{r^{O(r)}\theta^{3r}}{n} \quad \text{and}$$
$$\left|\mathbf{E}V^{2r} - \sum_{G\in\mathcal{G}_{2r}} a_G \mathbf{E} m_G(v_{\{j,k\}}:\ 1 \le j \ne k \le n)\right| \le \frac{r^{O(r)}\theta^{3r}}{n}.$$

Therefore, by Part (2) of Lemma 5.4,

$$|\mathbf{E}U^{2r} - \mathbf{E}V^{2r}| \le \frac{r^{O(r)}\theta^{3r}}{n}. \tag{5.5.4}$$

From Taylor's Theorem

$$\left|e^{ix} - \sum_{s=0}^{2r-1} i^s \frac{x^s}{s!}\right| \le \frac{x^{2r}}{(2r)!} \quad \text{for } x \in \mathbb{R},$$

it follows that

$$|\mathbf{E}e^{iU} - \mathbf{E}e^{iV}| \le \frac{\mathbf{E}U^{2r}}{(2r)!} + \frac{\mathbf{E}V^{2r}}{(2r)!} + \sum_{s=0}^{2r-1} \frac{|\mathbf{E}U^s - \mathbf{E}V^s|}{s!}.$$

From (5.5.1) and (5.2.1) we deduce that for a positive integer r we have

$$\mathbf{E}V^{2r} \le \frac{(2r)!2^{O(r)}\theta^{3r}}{r!}.$$

Therefore, by (5.5.4)

$$|\mathbf{E}e^{iU} - \mathbf{E}e^{iV}| \le \frac{2^{O(r)}\theta^{3r}}{r!} + \frac{r^{O(r)}\theta^{3r}}{n}. \tag{5.5.5}$$

Given $0 \le \epsilon \le 1/2$, one can choose a positive integer r such that

$$r \ln r = O\left(\theta^2 \ln \frac{1}{\epsilon}\right)$$

so that the first term in the right hand side of (5.5.5) does not exceed $\epsilon/2$. It follows then that for all

$$n \ge \left(\frac{1}{\epsilon}\right)^{\gamma(\theta)}$$

we have

$$|\mathbf{E}e^{iU} - \mathbf{E}e^{iV}| \le \epsilon,$$

and the proof follows by (5.5.2).

6. THE FOURTH DEGREE TERM

The main result of this section is the following theorem.

Theorem 6.1. *For unordered pairs $\{j,k\}$, $1 \leq j \neq k \leq n$, let $w_{\{j,k\}}$ be Gaussian random variables such that*

$$\mathbf{E}w_{\{j,k\}} = 0 \quad \text{for all } j,k,$$

and let $\sigma_{\{j,k\}} \in \{-1,1\}$ be numbers.

Suppose further that for some $\theta > 0$ we have

$$\mathbf{E}w_{\{j,k\}}^2 \leq \frac{\theta}{n} \quad \text{for all } j,k,$$

and that

$$|\mathbf{E}w_{\{j_1,k_1\}}w_{\{j_2,k_2\}}| \leq \frac{\theta}{n^2} \quad \text{provided } \{j_1,k_1\} \cap \{j_2,k_2\} = \emptyset.$$

Let

$$W = \sum_{\{j,k\}} \sigma_{\{j,k\}} w_{\{j,k\}}^4.$$

Then for some constant $\gamma(\theta) > 0$ we have

(1) $\mathbf{E}|W| \leq \gamma(\theta)$;

(2) $\mathbf{var}W \leq \dfrac{\gamma(\theta)}{n}$;

(3) $\mathbf{P}\{|W| > \gamma(\theta)\} \leq \exp\{-n^{1/5}\}$

provided $n \geq \gamma_1(\theta)$ for some constant $\gamma_1(\theta) > 0$.

We apply Theorem 6.1 in the following situation. Let $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ be the quadratic form defined by (1.3.1). Let us consider the Gaussian probability measure on $\mathbb{R}^n$ with density proportional to e^{-q}. We define random variables $w_{\{j,k\}}$ by

$$w_{\{j,k\}}(t) = \sqrt[4]{\frac{1}{24}\zeta_{\{j,k\}}(1-\zeta_{\{j,k\}})\left|6\zeta_{\{j,k\}}^2 - 6\zeta_{\{j,k\}} + 1\right|}(\tau_j + \tau_k) \quad \text{for } t = (\tau_1, \ldots, \tau_n)$$

and let

$$\sigma_{\{j,k\}} = \operatorname{sign}\left(6\zeta_{\{j,k\}}^2 - 6\zeta_{\{j,k\}} + 1\right).$$

Then for the function h defined by (1.3.2), we have

$$h = \sum_{\{j,k\}} \sigma_{\{j,k\}} w_{\{j,k\}}^4.$$

While the proof of Parts (1)–(2) is done by a straightforward computation, to prove Part (3) we need reverse Hölder inequalities for polynomials with respect to the Gaussian measure.

Lemma 6.2. *Let p be a polynomial of degree d in random Gaussian variables $w_1, \ldots, w_l$. Then for $r > 2$ we have*

$$(\mathbf{E}|p|^r)^{1/r} \leq r^{d/2}(\mathbf{E}p^2)^{1/2}.$$

Proof. This is Corollary 5 of [10]. ∎

6.3. Proof of Theorem 6.1

All implied constants in the "O" notation below are absolute.

By formula (5.2.1),

$$\mathbf{E}w_{\{j,k\}}^4 = 3\big(\mathbf{E}w_{\{j,k\}}^2\big)^2 = O\left(\frac{\theta^2}{n^2}\right)$$

and hence

$$\mathbf{E}|W| = O(\theta^2)$$

and Part (1) follows. Furthermore,

$$\mathbf{var}\,W = \sum_{\substack{\{j_1,k_1\}\\ \{j_2,k_2\}}} \sigma_{\{j_1,k_1\}}\sigma_{\{j_2,k_2\}}\mathbf{cov}\big(w_{\{j_1,k_1\}}^4, w_{\{j_2,k_2\}}^4\big).$$

By (5.2.3) we have

$$\mathbf{cov}\big(w_{\{j_1,k_1\}}^4, w_{\{j_2,k_2\}}^4\big) = O\left(\frac{\theta^4}{n^4}\right)$$

and, additionally,

$$\mathbf{cov}\big(w_{\{j_1,k_1\}}^4, w_{\{j_2,k_2\}}^4\big) = O\left(\frac{\theta^4}{n^6}\right) \quad \text{provided } \{j_1, k_1\} \cap \{j_2, k_2\} = \emptyset.$$

Therefore,

$$\mathbf{var}\,W = O\left(\frac{\theta^4}{n}\right), \tag{6.3.1}$$

which proves Part (2).

Finally, applying Lemma 6.2 with $d = 4$ we deduce from (6.3.1) that for any $r > 2$

$$\mathbf{E}|W - \mathbf{E}W|^r \leq r^{2r}n^{-r/2}2^{O(r)}\theta^{2r}.$$

Choosing $r = n^{1/5}$, we conclude that for all sufficiently large $n \geq n_0(\theta)$ we have

$$\mathbf{E}|W - \mathbf{E}W|^r \leq \exp\{-n^{1/5}\}.$$

By Markov's inequality, we obtain

$$\mathbf{P}\{|W - \mathbf{E}W| > 1\} \leq \exp\{-n^{1/5}\}$$

for all sufficiently large n and the proof follows from Part (1). ∎

7. COMPUTING THE INTEGRAL OVER A NEIGHBORHOOD OF THE ORIGIN

We consider the integral

$$\int_{\Pi} F(t)dt$$

of Corollary 3.3. Hence

$$F(t) = \exp\left\{-i\sum_{m=1}^{n} d_m\tau_m\right\}\prod_{\{j,k\}}(1-\zeta_{\{j,k\}}+\zeta_{\{j,k\}}e^{i(\tau_j+\tau_k)}) \text{ for } t=(\tau_1,\ldots,\tau_n),$$

where $D=(d_1,\ldots,d_n)$ is a given degree sequence, $z=(\zeta_{\{j,k\}})$ is the maximum entropy matrix and Π is the parallelepiped $[-\pi,\pi]^n$. We recall that the quadratic form $q:\mathbb{R}^n\longrightarrow\mathbb{R}$ is defined by

$$q(t)=\frac{1}{2}\sum_{\{j,k\}}\left(\zeta_{\{j,k\}}-\zeta^2_{\{j,k\}}\right)(\tau_j+\tau_k)^2 \quad \text{for } t=(\tau_1,\ldots,\tau_n).$$

In this section, we prove the following main result.

Theorem 7.1. *Let us fix a number* $0<\delta\le 1/2$ *and suppose that*

$$\delta\le\zeta_{\{j,k\}}\le 1-\delta \quad \textit{for all } j\ne k.$$

Let $f,h:\mathbb{R}^n\longrightarrow\mathbb{R}$ *be polynomials defined by (1.3.2). Let us define a neighborhood of the origin* $\mathcal{U}\subset\Pi$ *by*

$$\mathcal{U}=\left\{(\tau_1,\ldots,\tau_n):|\tau_k|\le\frac{\ln n}{\sqrt{n}} \textit{ for } k=1,\ldots,n\right\}.$$

Let

$$\Xi=\int_{\mathbb{R}^n}e^{-q(t)}dt$$

and let us consider the Gaussian probability measure in $\mathbb{R}^n$ *with density* $\Xi^{-1}e^{-q}$*. Let*

$$\mu=\mathbf{E}f^2 \quad \textit{and} \quad \nu=\mathbf{E}h.$$

Then

(1) $\Xi\ge\left(\frac{4\pi}{n}\right)^{n/2}$;

(2) $\mu,\ |\nu|,\ \mathbf{E}|h|\le\gamma(\delta)$

for some constant $\gamma(\delta)>0$*;*

(3) *For any* $0<\epsilon\le 1/2$

$$\left|\int_{\mathcal{U}}|F(t)|dt-\exp\{\nu\}\Xi\right|\le\epsilon\,\Xi$$

provided

$$n \geq \left(\frac{1}{\epsilon}\right)^{\gamma(\delta)}$$

for some constant $\gamma(\delta) > 0$*;*

(4) *For any* $0 < \epsilon \leq 1/2$

$$\left| \int_{\mathcal{U}} F(t)dt - \exp\left\{-\frac{\mu}{2} + \nu\right\} \Xi \right| \leq \epsilon \Xi$$

provided

$$n \geq \left(\frac{1}{\epsilon}\right)^{\gamma(\delta)}$$

for some $\gamma(\delta) > 0$.

Proof. In what follows, all constants implied in the "O" and "Ω" notation depend only on the parameter δ.

Let

$$q_0(t) = \frac{1}{2} \sum_{\{j,k\}} (\tau_j + \tau_k)^2 \text{ for } t = (\tau_1, \ldots, \tau_n)$$

as in Lemma 4.3. Then $q(t) \leq \frac{1}{4} q_0(t)$ and hence

$$\int_{\mathbb{R}^n} e^{-q} dt \geq \int_{\mathbb{R}^n} e^{-\frac{1}{4} q_0(t)} dt = \pi^{n/2} \sqrt{\frac{4}{n-1}} \left(\frac{8}{n-2}\right)^{\frac{n-1}{2}} \geq \left(\frac{4\pi}{n}\right)^{n/2},$$

which proves Part (1).

Let us think of the coordinate functions τ_j as random variables with respect to the Gaussian probability measure with density proportional to e^{-q}.

By Theorem 4.1, we have

$$\begin{aligned} |\mathbf{E}\tau_j \tau_k| &= O\left(\frac{1}{n^2}\right) \quad \text{provided } j \neq k \quad \text{and} \\ \mathbf{E}\tau_j^2 &= O\left(\frac{1}{n}\right) \quad \text{for } j = 1, \ldots, n. \end{aligned} \tag{7.1.1}$$

For an unordered pair $1 \leq j \neq k \leq n$, let us define

$$u_{\{j,k\}} = \sqrt[3]{\frac{1}{6} \zeta_{\{j,k\}} (1 - \zeta_{\{j,k\}})(2\zeta_{\{j,k\}} - 1)} (\tau_j + \tau_k).$$

Then

$$f = \sum_{\{j,k\}} u_{\{j,k\}}^3.$$

Similarly, let us define

$$w_{\{j,k\}} = \sqrt[4]{\frac{1}{24}\zeta_{\{j,k\}}(1-\zeta_{\{j,k\}})\left|6\zeta_{\{j,k\}}^2 - 6\zeta_{\{j,k\}} + 1\right|(\tau_j + \tau_k)} \quad \text{and}$$
$$\sigma_{\{j,k\}} = \operatorname{sign}\left(6\zeta_{\{j,k\}}^2 - 6\zeta_{\{j,k\}} + 1\right).$$

Then

$$h = \sum_{\{j,k\}} \sigma_{\{j,k\}} w_{\{j,k\}}^4.$$

By (7.1.1) the random variables $u_{\{j,k\}}$ satisfy the conditions of Theorem 5.1 and hence the upper bound for $\mu = \mathbf{E}f^2$ follows by Theorem 5.1. Similarly, by (7.1.1) the random variables $w_{\{j,k\}}$ satisfy the conditions of Theorem 6.1 and hence the upper bound for $|\nu| = \mathbf{E}h$ and $\mathbf{E}|h|$ follows by Part (1) of Theorem 6.1. This concludes the proof of Part (2) of the theorem.

By Lemma 4.3, eigenvalues of q are $\Omega(n)$ from which it follows that

$$\left|\int_{\mathbb{R}^n\setminus\mathcal{U}} e^{-q(t)}\,dt\right| \leq \exp\{-\Omega(\ln^2 n)\}\Xi. \tag{7.1.2}$$

From Theorem 5.1, we have

$$\left|\int_{\mathbb{R}^n} e^{-q(t)+if(t)}\,dt - \exp\left\{-\frac{\mu}{2}\right\}\Xi\right| \leq \epsilon\Xi \quad \text{provided } n \geq \left(\frac{1}{\epsilon}\right)^{O(1)},$$

which, combined with (7.1.2), results in

$$\left|\int_{\mathcal{U}} e^{-q(t)+if(t)}\,dt - \exp\left\{-\frac{\mu}{2}\right\}\Xi\right| \leq \epsilon\Xi \quad \text{provided } n \geq \left(\frac{1}{\epsilon}\right)^{O(1)}. \tag{7.1.3}$$

By Part (2) of Theorem 6.1 and Chebyshev's inequality, we have

$$\mathbf{P}\{|h-\nu| > \epsilon\} = O\left(\frac{1}{\epsilon^2 n}\right), \tag{7.1.4}$$

while by Part (3) of Theorem 6.1, we have

$$\mathbf{P}\{h > \gamma(\delta)\} = O(\exp\{-n^{1/5}\}) \tag{7.1.5}$$

for some constant $\gamma(\delta) > 0$. In addition,

$$|h(t)| = O(\ln^4 n) \quad \text{for } t \in \mathcal{U}. \tag{7.1.6}$$

Combining (7.1.4)–(7.1.6) and Part (2) of the theorem, we deduce from (7.1.2) and (7.1.3) that

$$\begin{gathered}\left|\int_{\mathcal{U}} e^{-q(t)+h(t)}\,dt - \exp\{\nu\}\Xi\right| \leq \epsilon\Xi \\ \text{and} \\ \left|\int_{\mathcal{U}} e^{-q(t)+if(t)+h(t)}\,dt - \exp\left\{-\frac{\mu}{2}+\nu\right\}\Xi\right| \leq \epsilon\Xi \quad \text{provided } n \geq \left(\frac{1}{\epsilon}\right)^{O(1)}.\end{gathered} \tag{7.1.7}$$

From the Taylor series expansion, cf. (3.2), we obtain

$$F(t) = \exp\{-q(t) + if(t) + h(t) + \rho(t)\}, \quad \text{where}$$
$$|\rho(t)| = O\left(\frac{\ln^5 n}{\sqrt{n}}\right) \quad \text{for } t \in \mathcal{U}.$$

Therefore, for any $\epsilon > 0$ we have

$$||F(t)| - e^{-q(t)+h(t)}| \leq \epsilon e^{-q(t)+h(t)} \quad \text{and}$$
$$|F(t) - e^{-q(t)+if(t)+h(t)}| \leq \epsilon e^{-q(t)+h(t)} \quad \text{for all } t \in \mathcal{U} \quad \text{provided} \quad n \geq \left(\frac{1}{\epsilon}\right)^{O(1)}.$$

The proof of Parts (3) and (4) now follows from (7.1.7) and Part (2). ∎

8. BOUNDING THE INTEGRAL OUTSIDE OF THE SPECIAL POINTS

We consider the integral representation of Corollary 3.3. Our goal is to show that the integral of $F(t)$ for t outside of the neighborhood of the special points

$$(0, \ldots, 0) \quad \text{and} \quad (\pm\pi, \ldots, \pm\pi)$$

is asymptotically negligible.

In this section, we prove the following main result.

Theorem 8.1. *Let us fix a number $0 < \delta \leq 1/2$ and let $D = (d_1, \ldots, d_n)$ be a δ-tame degree sequence. Let us define subsets $\mathcal{U}, \mathcal{W} \subset \Pi$ by*

$$\mathcal{U} = \left\{(\tau_1, \ldots, \tau_n) : |\tau_j| \leq \frac{\ln n}{\sqrt{n}} \text{ for } j = 1, \ldots, n\right\} \quad \text{and}$$
$$\mathcal{W} = \left\{(\tau_1, \ldots, \tau_n) : |\tau_j - \sigma_j\pi| \leq \frac{\ln n}{\sqrt{n}} \quad \text{for some } \sigma_j = \pm 1 \quad \text{and all } j = 1, \ldots, n\right\}.$$

Then for any $\kappa > 0$

$$\int_{\Pi \setminus (\mathcal{U} \cup \mathcal{W})} |F(t)| dt \leq n^{-\kappa} \int_{\mathcal{U}} |F(t)| dt$$

provided $n > \gamma(\delta, \kappa)$.

The plan of the proof of Theorem 8.1 is as follows: first, using some combinatorial arguments we show that for any positive constant $\epsilon > 0$ the integral is asymptotically negligible outside of the areas where $|\tau_j| \leq \epsilon$ for all j or where $|\tau_j - \sigma_j\pi| \leq \epsilon$ for some $\sigma_j = \pm 1$ and all j. Then we note that $|F(t)|$ is log-concave for t in a neighborhood of the origin and use a concentration inequality for log-concave measures.

We introduce the following metric ρ.

8.2. Metric ρ

Let us define a function $\rho : \mathbb{R} \longrightarrow [0, \pi]$ as follows:

$$\rho(x) = \min_{k \in \mathbb{Z}} |x - 2\pi k|.$$

In words: $\rho(x)$ is the distance from x to the nearest integer multiple of 2π. Clearly,

$$\rho(-x) = \rho(x) \quad \text{and} \quad \rho(x+y) \leq \rho(x) + \rho(y)$$

for all $x, y \in \mathbb{R}$.

We will use that

$$1 - \frac{1}{2}\rho^2(x) \leq \cos x \leq 1 - \frac{1}{5}\rho^2(x). \tag{8.2.1}$$

8.3. The Absolute Value of $F(t)$

Let

$$\alpha_{\{j,k\}} = 2\zeta_{\{j,k\}}(1 - \zeta_{\{j,k\}}) \quad \text{for all } j \neq k.$$

If D is δ-tame, we have

$$2\delta^2 \leq \alpha_{\{j,k\}} \leq \frac{1}{2} \quad \text{for all } j \neq k. \tag{8.3.1}$$

We have

$$|F(t)| = \left(\prod_{\{j,k\}} (1 - \alpha_{\{j,k\}} + \alpha_{\{j,k\}} \cos(\tau_j + \tau_k)) \right)^{1/2}.$$

For $1 \leq j \neq k \leq n$ let us define a function of $\tau \in \mathbb{R}$,

$$f_{\{j,k\}}(\tau) = \sqrt{1 - \alpha_{\{j,k\}} + \alpha_{\{j,k\}} \cos \tau},$$

so

$$|F(t)| = \prod_{\{j,k\}} f_{\{j,k\}}(\tau_j + \tau_k). \tag{8.3.2}$$

We note that

$$f_{\{j,k\}}(0) = 1.$$

It follows by (8.2.1) and (8.3.1) that for $\epsilon > 0$

$$f_{\{j,k\}}(x) \leq \exp\{-\gamma(\delta, \epsilon)\} f_{\{j,k\}}(y) \quad \text{provided } \rho(x) \geq 2\epsilon \text{ and } \rho(y) \leq \epsilon \tag{8.3.3}$$

for some $\gamma(\epsilon, \delta) > 0$. Furthermore,

$$\frac{d^2}{d\tau^2} \ln f_{\{j,k\}}(\tau) = -\frac{\alpha_{\{j,k\}}(\alpha_{\{j,k\}} + \cos \tau - \alpha_{\{j,k\}} \cos \tau)}{2(1 - \alpha_{\{j,k\}} + \alpha_{\{j,k\}} \cos \tau)^2}.$$

In particular, by (8.3.1)

$$\frac{d^2}{d\tau^2} \ln f_{\{j,k\}}(\tau) \le -\frac{\delta^2}{2} \quad \text{for } -\frac{\pi}{3} \le \tau \le \frac{\pi}{3}$$

and hence $\ln f_{\{j,k\}}$ is strictly concave on the interval $[-\pi/3, \pi/3]$:

$$\ln f_{\{j,k\}}(x) + \ln f_{\{j,k\}}(y) - 2\ln f_{\{j,k\}}\left(\frac{x+y}{2}\right) \le -\frac{\delta^2}{8}|x-y|^2 \quad \text{for all } x, y \in [-\pi/3, \pi/3]. \tag{8.3.4}$$

In what follows, we fix a particular parameter $\epsilon > 0$. All implied constants in the "O" and "Ω" notation below may depend only on the parameters δ and ϵ. We say that n is *sufficiently large* if $n \ge \gamma(\delta, \epsilon)$ for some constant $\gamma(\delta, \epsilon) > 0$.

Our first goal is to show that only the points $t \in \Pi$ for which the inequality $\rho(\tau_j + \tau_k) \le \epsilon$ holds for an overwhelming majority of pairs $\{j, k\}$ contribute significantly to the integral of $|F(t)|$ on Π.

Lemma 8.4. *For $t \in \Pi$, $t = (\tau_1, \ldots, \tau_n)$, and $\epsilon > 0$ let us define a set $K(t, \epsilon) \subset \{1, \ldots, n\}$ consisting of the indices k such that*

$$\rho(\tau_j + \tau_k) \le \epsilon$$

for more than $n/2$ distinct indices j. Let $\overline{K(t, \epsilon)} = \{1, \ldots, n\} \setminus K(t, \epsilon)$. Then

1. $|F(t)| \le \exp\{-\gamma(\delta)\epsilon^2 n |\overline{K(t, \epsilon)}|\}$ *for some* $\gamma(\delta) > 0$;
2. $\rho(\tau_{k_1} - \tau_{k_2}) \le 2\epsilon$ *for all* $k_1, k_2 \in K(t, \epsilon)$;
3. *Suppose that $|K(t, \epsilon)| > n/2$. Then*

$$\rho(\tau_{k_1} + \tau_{k_2}) \le 3\epsilon \quad \textit{for all } k_1, k_2 \in K(t, \epsilon).$$

Proof. For every $k \in \overline{K(t, \epsilon)}$ there are at least $(n-2)/2$ distinct $j \ne k$ for which

$$\rho(\tau_j + \tau_k) > \epsilon \tag{8.4.1}$$

and so by (8.2.1) we have

$$\cos(\tau_j + \tau_k) \le 1 - \frac{1}{5}\epsilon^2.$$

Since there are at least $|\overline{K(t, \epsilon)}|(n-2)/4$ pairs $\{j, k\}$ for which (8.4.1) holds, the proof of Part (1) follows from (8.3.2) and (8.3.3).

For any $k_1, k_2 \in K$ there is $j \in \{1, \ldots, n\}$ such that

$$\rho(\tau_j + \tau_{k_1}), \rho(\tau_j + \tau_{k_2}) \le \epsilon.$$

Therefore,

$$\begin{aligned} \rho(\tau_{k_1} - \tau_{k_2}) &= \rho(\tau_j + \tau_{k_1} - \tau_{k_2} - \tau_j) \le \rho(\tau_j + \tau_{k_1}) + \rho(-\tau_j - \tau_{k_2}) \\ &= \rho(\tau_j + \tau_{k_1}) + \rho(\tau_j + \tau_{k_2}) \le 2\epsilon \end{aligned}$$

and Part (2) follows.

Let us choose a $k_1 \in K(t,\epsilon)$. Since $|K(t,\epsilon)| > n/2$, there is a $j \in K(t,\epsilon)$ such that

$$\rho(\tau_j + \tau_{k_1}) \leq \epsilon.$$

By Part (2), for any $k_2 \in K(t,\epsilon)$ we have

$$\begin{aligned}\rho(\tau_{k_1} + \tau_{k_2}) &= \rho(-\tau_j + \tau_{k_1} + \tau_j + \tau_{k_2}) \leq \rho(\tau_j + \tau_{k_1}) + \rho(-\tau_j + \tau_{k_2}) \\ &\leq 3\epsilon\end{aligned}$$

and Part (3) follows. ∎

Corollary 8.5. *For an $\epsilon > 0$ let us define a set $V(\epsilon) \subset \Pi$ consisting of the points $t \in \Pi$ such that*

$$\overline{K(t,\epsilon)} \leq \ln^2 n,$$

where $\overline{K(t,\epsilon)}$ is defined in Lemma 8.4. Then

$$\int_{\Pi \setminus V(\epsilon)} |F(t)| dt \leq n^{-n} \int_{\Pi} |F(t)| dt$$

provided $n \geq \gamma(\delta,\epsilon)$ for some constant $\gamma(\delta,\epsilon) > 0$.

Proof. By Parts (1)–(3) of Theorem 7.1 we have

$$\int_{\Pi} |F(t)| dt \geq \Omega(n^{-n/2}).$$

The proof now follows by Part (1) of Lemma 8.4. ∎

Next, we show that only the points $t \in \Pi$ such that $\rho(\tau_j + \tau_k) \leq \epsilon$ for *all* $1 \leq j,k \leq n$ contribute substantially to the integral of $|F(t)|$ on Π.

Lemma 8.6. *For $\epsilon > 0$ let us define a set $X(\epsilon) \subset \Pi$,*

$$X(\epsilon) = \{t \in \Pi, t = (\tau_1, \ldots, \tau_n) : \rho(\tau_j + \tau_k) \leq \epsilon \text{ for all } j,k\}.$$

Then

$$\int_{\Pi \setminus X(\epsilon)} |F(t)| dt \leq \exp\{-\gamma_1(\delta,\epsilon)n\} \int_{\Pi} |F(t)| dt$$

for all $n \geq \gamma_2(\delta,\epsilon)$ for some constants $\gamma_1(\delta,\epsilon), \gamma_2(\delta,\epsilon) > 0$.

Proof. Let us consider the set $V(\epsilon/60) \subset \Pi$ and n large enough so that the conclusion of Corollary 8.5 holds, that is, the integral of $|F(t)|$ over $\Pi \setminus V(\epsilon/60)$ is asymptotically negligible. For a set $A \subset \{1, \ldots, n\}$ such that

$$|\overline{A}| \leq \ln^2 n,$$

let us define a set $P_A \subset \Pi$ (we call it a *piece*) such that

$$\rho(\tau_j - \tau_k) \leq \epsilon/30 \quad \text{and} \quad \rho(\tau_j + \tau_k) \leq \epsilon/20 \quad \text{for all } j,k \in A.$$

If n is large enough, by Lemma 8.4 for every $t \in V(\epsilon/60)$ we can choose $A = K(t, \epsilon)$, so we have

$$V(\epsilon/60) \subset \bigcup_{A : |\overline{A}| \leq \ln^2 n} P_A. \tag{8.6.1}$$

Our next goal is to show that the integral of $|F(t)|$ over $P_A \setminus X(\epsilon)$ is negligible compared to the integral of $|F(t)|$ over P_A.

Let us choose a point $t \in P_A \setminus X(\epsilon)$. Thus we have

$$\rho(\tau_{i_0} + \tau_{j_0}) > \epsilon \quad \text{for some } i_0, j_0.$$

Let us choose any $k_0 \in A$. Then

$$\rho(\tau_{i_0} + \tau_{j_0}) = \rho(\tau_{i_0} + \tau_{j_0} + \tau_{k_0} - \tau_{k_0}) \leq \rho(\tau_{i_0} + \tau_{k_0}) + \rho(\tau_{j_0} - \tau_{k_0}).$$

Hence we have either

$$\rho(\tau_{i_0} + \tau_{k_0}) > \epsilon/2 \quad \text{or } \rho(\tau_{j_0} - \tau_{k_0}) > \epsilon/2.$$

In the first case, for every $k \in A$ we have

$$\begin{aligned}\rho(\tau_{i_0} + \tau_{k_0}) &= \rho(\tau_{i_0} + \tau_{k_0} + \tau_k - \tau_k) \leq \rho(\tau_{i_0} + \tau_k) + \rho(\tau_{k_0} - \tau_k) \\ &\leq \rho(\tau_{i_0} + \tau_k) + \epsilon/30,\end{aligned}$$

from which

$$\rho(\tau_{i_0} + \tau_k) \geq \epsilon/2 - \epsilon/30 = 7\epsilon/15.$$

In the second case, for every $k \in A$, we have

$$\begin{aligned}\rho(\tau_{j_0} - \tau_{k_0}) &= \rho(\tau_{j_0} - \tau_{k_0} + \tau_k - \tau_k) \leq \rho(\tau_{j_0} + \tau_k) + \rho(-\tau_{k_0} - \tau_k) \\ &= \rho(\tau_{j_0} + \tau_k) + \rho(\tau_{k_0} + \tau_k) \leq \rho(\tau_{j_0} + \tau_k) + \epsilon/20,\end{aligned}$$

from which

$$\rho(\tau_{j_0} + \tau_k) > \epsilon/2 - \epsilon/20 = 9\epsilon/20.$$

In either case, for any $t \in P_A \setminus X(\epsilon)$, $t = (\tau_1, \ldots, \tau_n)$, there exists an index $i \notin A$ such that

$$\rho(\tau_i + \tau_k) > 0.45\epsilon \quad \text{for all } k \in A.$$

For $i \notin A$, we define

$$Q_{A,i} = \{t \in P_A : \rho(\tau_i + \tau_k) > 0.45\epsilon \text{ for all } k \in A\}. \tag{8.6.2}$$

Hence

$$P_A \setminus X(\epsilon) \subset \bigcup_{i \in \overline{A}} Q_{A,i}.$$

Given a point $t \in P_A$, we obtain another point in P_A if we arbitrarily change the coordinate $\tau_i \in [-\pi, \pi]$ for $i \in \overline{A}$. We obtain a *fiber* $E \subset P_A$ if we fix all other coordinates and let $\tau_i \in [-\pi, \pi]$ vary. Geometrically, each fiber E is an interval of length 2π. Let us construct a set $I \subset E$ as follows. We choose an arbitrary $k_0 \in A$ and let τ_i vary in such a way that $\rho(\tau_{k_0} + \tau_i) \leq 0.05\epsilon$. Geometrically, I is an interval of length 0.1ϵ or a union of two non-overlapping intervals of total length 0.1ϵ. Moreover,

$$\rho(\tau_k + \tau_i) \leq \rho(\tau_{k_0} + \tau_i) + \rho(\tau_k - \tau_{k_0}) \leq 0.05\epsilon + 0.05\epsilon = 0.1\epsilon$$

for all $k \in A$ and all $\tau \in I$.

Using (8.3.2) and (8.3.3), we conclude from (8.6.2) that for any $t \in Q_{A,i} \cap E$ and for any $s \in P_A \cap I$ we have

$$|F(t)| \leq \exp\{-\Omega(n)\}|F(s)|$$

provided n is large enough. Therefore,

$$\int_{E \cap Q_{A,i}} |F(t)|dt \leq \exp\{-\Omega(n)\} \int_E |F(t)|dt$$

for all sufficiently large n.

Integrating over all fibers E, we establish that

$$\int_{Q_{A,i}} |F(t)|dt \leq \exp\{-\Omega(n)\} \int_{P_A} |F(t)|dt$$

for all sufficiently large n. Since the number of different subsets $A \subset \{1, \ldots, n\}$ with $|\overline{A}| \leq \ln^2 n$ in (8.6.1) does not exceed $\exp\{O(\ln^3 n)\}$, the proof follows. ∎

Next, we prove that only the points in the neighborhood of the origin or the corners of Π contribute significantly to the integral of $|F(t)|$ over Π.

Lemma 8.7. *For $0 < \epsilon < 1$, let $X(\epsilon)$ be the set defined in Lemma 8.6. Let us define $Y(\epsilon), Z(\epsilon) \subset \Pi$ by*

$$\begin{aligned} Y(\epsilon) &= \{t \in \Pi : \ t = (\tau_1, \ldots, \tau_n), |\tau_i| \leq \epsilon/2 \text{ for } i = 1, \ldots, n\} \quad \text{and} \\ Z(\epsilon) &= \{t \in \Pi : \ t = (\tau_1, \ldots, \tau_n), |\tau_i - \sigma_i \pi| \leq \epsilon/2 \text{ for some } \sigma_i = \pm 1 \\ &\qquad \text{and all } i = 1, \ldots, n\}. \end{aligned}$$

Then

$$X(\epsilon) = Y(\epsilon) \cup Z(\epsilon) \quad \text{and} \quad Y(\epsilon) \cap Z(\epsilon) = \emptyset.$$

Moreover,

$$\int_{Y(\epsilon)} |F(t)|dt = \int_{Z(\epsilon)} |F(t)|dt.$$

Proof. Let us pick a point $t \in X(\epsilon)$. Then for each k we have $\rho(2\tau_k) \leq \epsilon$ and hence either

$$|\tau_k| \leq \epsilon/2 \quad \text{or} \quad |\tau_k - \pi| \leq \epsilon/2 \quad \text{or} \quad |\tau_k + \pi| \leq \epsilon/2.$$

Since $\rho(\tau_k + \tau_j) \leq \epsilon$ for all k, j, we conclude that if $|\tau_k| \leq \epsilon/2$ for some k then $|\tau_k| \leq \epsilon/2$ for all k. Hence $X(\epsilon) \subset (Y(\epsilon) \cup Z(\epsilon))$. The inclusion $(Z(\epsilon) \cup Y(\epsilon)) \subset X(\epsilon)$ is obvious. Since $\epsilon < 1$, we have $Y(\epsilon) \cap Z(\epsilon) = \emptyset$.

The set $Z(\epsilon)$ is a union of 2^n pairwise disjoint *corners*, where each corner is determined by a choice of the interval $[-\pi, -\pi + \epsilon/2]$ or $[\pi - \epsilon/2, \pi]$ for each coordinate τ_i. The transformation

$$\tau_k \longmapsto \begin{cases} \tau_k + \pi & \text{if } \tau_k \in [-\pi, -\pi + \epsilon/2] \\ \tau_k - \pi & \text{if } \tau_k \in [\pi - \epsilon/2, \pi] \end{cases}$$

is a volume-preserving transformation which maps $Z(\epsilon)$ onto $X(\epsilon)$ and does not change the value of $|F(t)|$. ∎

Finally, we will use that $|F(t)|$ is strictly log-concave on the set $Y(1)$. For Euclidean space V with the norm $\|\cdot\|$, a point $x \in V$ and a closed set $A \subset V$ we define the distance

$$\text{dist}(x, A) = \min_{y \in A} \|x - y\|.$$

We will need the following concentration inequality for strictly log-concave measures.

Theorem 8.8. *Let V be Euclidean space with the norm $\|\cdot\|$, let $B \subset V$ be a convex body and let us consider a probability measure supported on B with density e^{-u}, where $u : B \longrightarrow \mathbb{R}$ is a function satisfying*

$$u(x) + u(y) - 2u\left(\frac{x+y}{2}\right) \geq c\|x - y\|^2 \quad \textit{for all} \quad x, y \in B$$

and some constant $c > 0$.

Let $A \subset B$ be a closed subset such that $\mathbf{P}(A) \geq 1/2$. Then, for $r \geq 0$, we have

$$\mathbf{P}\{x \in B : \text{dist}(x, A) \geq r\} \leq 2e^{-cr^2}.$$

Proof. See Section 2.2 of [13] and Theorem 8.1 and its proof in [1]. ∎

Lemma 8.9. *Let $Y(1) \subset \Pi$ be the set defined by*

$$Y(1) = \left\{ t \in \Pi,\ t = (\tau_1, \ldots, \tau_n) : |\tau_i| \leq \frac{1}{2} \textit{ for } i = 1, \ldots, n \right\}.$$

Let $\mathcal{U} \subset \Pi$ be the set

$$\mathcal{U} = \left\{ t \in \Pi,\ t = (\tau_1, \ldots, \tau_n) : |\tau_i| \leq \frac{\ln n}{\sqrt{n}} \textit{ for } \quad i = 1, \ldots, n \right\}.$$

Then for any $\kappa > 0$ we have

$$\int_{Y(1) \setminus \mathcal{U}} |F(t)| dt \leq n^{-\kappa} \int_{Y(1)} |F(t)| dt,$$

provided $n \geq \gamma(\delta, \kappa)$ is big enough. ∎

Proof. Let us consider the probability measure on $Y(1)$ with density proportional to $|F(t)|$. Let us consider a map $M : \mathbb{R}^n \longrightarrow \mathbb{R}^{\binom{n}{2}}$, where the coordinates of $\mathbb{R}^{\binom{n}{2}}$ are indexed by unordered pairs $\{j,k\}$ and

$$M_{\{j,k\}}(\tau_1,\ldots,\tau_n) = \tau_j + \tau_k.$$

By Lemma 4.3,

$$\|M(t)\|^2 \geq (n-2)\|t\|^2 \quad \text{for all } t \in \mathbb{R}^n.$$

Since from Section 8.3,

$$\ln|F(t)| = \frac{1}{2}\sum_{\{j,k\}} \ln f_{\{j,k\}}(\tau_j + \tau_k),$$

it follows by (8.3.4) that for any constant a and

$$u(t) = -\ln|F(t)| + a,$$

we have

$$u(t_1) + u(t_2) - 2u\left(\frac{t_1+t_2}{2}\right) \geq \gamma(\delta)n\|t_1 - t_2\|^2 \text{ for all } t_1, t_2 \in Y(1)$$

and some constant $\gamma(\delta) > 0$. We choose a so that e^{-u} is a probability density on $Y(1)$.

We apply Theorem 8.8 with $c = \gamma(\delta)n$. For $k = 1,\ldots,n$, let $A_k^- \subset Y(1)$ be the set of points with $\tau_k \leq 0$ and let $A_k^+ \subset Y(1)$ be the set of points with $\tau_k \geq 0$. Since both $Y(1)$ and the probability measure are invariant under the symmetry $t \longmapsto -t$, we have

$$\mathbf{P}(A_k^-) = \mathbf{P}(A_k^+) = \frac{1}{2} \quad \text{for } k = 1,\ldots,n.$$

Therefore, by Theorem 8.8, all but a $n^{-\kappa}$ fraction of all points in $Y(1)$ lie within a distance of $\ln n/\sqrt{n}$ from each of the sets A_k^- and A_k^+, provided n is large enough. ∎

8.10. Proof of Theorem 8.1

For $0 < \epsilon < 1$ let us define the set $X(\epsilon)$ as in Lemma 8.6 and the sets $Y(\epsilon)$ and $Z(\epsilon)$ as in Lemma 8.7. In particular,

$$\mathcal{U} = Y\left(\frac{2\ln n}{\sqrt{n}}\right) \quad \text{and} \quad \mathcal{W} = Z\left(\frac{2\ln n}{\sqrt{n}}\right).$$

By Lemma 8.6, for any $\kappa > 0$ we have

$$\int_{\Pi\setminus X(1)} |F(t)|dt \leq n^{-\kappa}\int_{\Pi} |F(t)|dt$$

for all sufficiently large n, so that the integral outside of $X(1)$ is asymptotically negligible.

By Lemma 8.7, $X(1) = Y(1) \cup Z(1)$ with $Y(1) \cap Z(1) = \emptyset$ and

$$\int_{Y(1)} |F(t)|dt = \int_{Z(1)} |F(t)|dt \quad \text{and} \quad \int_{\mathcal{U}} |F(t)|dt = \int_{\mathcal{W}} |F(t)|dt. \tag{8.10.1}$$

By Lemma 8.9,

$$\int_{Y(1)\setminus\mathcal{U}} |F(t)|dt \leq n^{-\kappa} \int_{Y(1)} |F(t)|dt$$

for all sufficiently large n, so that the integral over $Y(1)\setminus\mathcal{U}$ is asymptotically negligible. By (8.10.1), the integral over $Z(1)\setminus\mathcal{W}$ is asymptotically negligible. The proof now follows. ∎

9. PROOF OF THEOREM 1.4

By Corollary 3.3, we have the integral representation for the number $|G(D)|$ of graphs:

$$|G(D)| = \frac{e^{H(z)}}{(2\pi)^n} \int_{\Pi} F(t)dt.$$

Let us define subsets $\mathcal{U}, \mathcal{W} \subset \Pi$ as in Theorem 8.1. Let us consider the transformation $\mathcal{W} \longrightarrow \mathcal{U}$,

$$\tau_k \longmapsto \begin{cases} \tau_k + \pi & \text{if } -\pi \leq \tau_k \leq -\pi + \dfrac{\ln n}{\sqrt{n}} \\ \tau_k - \pi & \text{if } \pi - \dfrac{\ln n}{\sqrt{n}} \leq \tau_k \leq \pi \end{cases} \qquad \text{for } k = 1, \dots, n.$$

As in the proof of Lemma 8.7, this is a measure-preserving transformation which maps $\mathcal{W}$ onto $\mathcal{U}$. Since $d_1 + \cdots + d_n$ is even, the transformation does not change the value of $F(t)$ (if $d_1 + \cdots + d_n$ is odd, the transformation changes the sign of $F(t)$). Hence

$$\int_{\mathcal{U}} F(t)dt = \int_{\mathcal{W}} F(t)dt. \tag{9.1}$$

By Theorem 7.1, the integrals of $F(t)$ and $|F(t)|$ over $\mathcal{U}$ have the same order of magnitude, that is,

$$\int_{\mathcal{U}} |F(t)|dt \leq \gamma(\delta) \left| \int_{\mathcal{U}} F(t)dt \right|$$

for some constant $\gamma(\delta) > 1$. Therefore, from Theorem 8.1, the integral outside of $\mathcal{U} \cup \mathcal{W}$ is asymptotically negligible, so that for any $\kappa > 0$ and all sufficiently large $n \geq \gamma(\delta, \kappa)$, we have

$$\int_{\Pi\setminus(\mathcal{U}\cup\mathcal{W})} |F(t)|dt \leq n^{-\kappa} \left| \int_{\mathcal{U}} F(t)dt \right|.$$

The proof now follows by Parts (2) and (4) of Theorem 7.1, identity (9.1) and the formula

$$\Xi = \int_{\mathbb{R}^n} e^{-q(t)}dt = \frac{(2\pi)^{n/2}}{\sqrt{\det Q}}.$$

∎

10. PROOF OF THEOREM 1.6

The proof is very similar to that of Theorem 3 of [2], which deals with a similar situation in the case of bipartite graphs. All implicit constant in the "O" and "Ω"-notation below may depend only on the parameter $\delta > 0$.

For pairs $1 \leq j \neq k \leq n$, let $x_{\{j,k\}}$ be independent Bernoulli random variables such that

$$\mathbf{P}\{x_{\{j,k\}} = 1\} = \zeta_{\{j,k\}} \quad \text{and} \quad \mathbf{P}\{x_{\{j,k\}} = 0\} = 1 - \zeta_{\{j,k\}}.$$

As is implied by Theorem 5 of [3], the probability mass function of the random vector $X = (x_{\{j,k\}})$ is constant on the integer points of $\mathcal{P}(D)$ and is equal to $e^{-H(z)}$ at each $G \in G(D)$.

Let us define

$$\sigma_S(X) = \sum_{\{j,k\} \in S} x_{\{j,k\}}.$$

Then

$$\begin{aligned}\mathbf{P}&\{G \in G(D) : \sigma_S(G) \leq (1-\epsilon)\sigma_S(z)\} \\ &= \frac{\mathbf{P}\{X : \sigma_S(X) \leq (1-\epsilon)\sigma_S(z) \text{ and } X \in G(D)\}}{\mathbf{P}\{X : X \in G(D)\}}\end{aligned}$$

and, similarly,

$$\begin{aligned}\mathbf{P}&\{G \in G(D) : \quad \sigma_S(G) \geq (1+\epsilon)\sigma_S(z)\} \\ &= \frac{\mathbf{P}\Big\{X : \quad \sigma_S(X) \geq (1+\epsilon)\sigma_S(z) \quad \text{and} \quad X \in G(D)\}}{\mathbf{P}\{X : \quad X \in G(D)\}}\end{aligned}$$

Applying Theorem 1.4 and Parts (1) and (2) of Theorem 7.1, we get

$$\mathbf{P}\{X : X \in G(D)\} = e^{-H(z)}|G(D)| \geq n^{-O(n)}.$$

On the other hand, standard large deviation inequalities for sums of bounded independent random variables (see, for example, Corollary 5.3 of [15]) imply that

$$\mathbf{P}\{X : \sigma_S(X) \geq (1+\epsilon)\sigma_S(z)\} \leq \exp\{-\Omega(n \ln^2 n)\}$$

and, similarly,

$$\mathbf{P}\{X : \sigma_S(X) \leq (1-\epsilon)\sigma_S(z)\} \leq \exp\{-\Omega(n \ln^2 n)\}$$

and the proof follows. ∎

11. COUNTING BIPARTITE GRAPHS

Here we list some modifications needed to establish the asymptotic formula (2.5.4). We adhere to the notation of Section 2.5.

As in Corollary 3.3, we represent the number 0-1 matrices with row sums $R = (r_1, \ldots, r_m)$ and column sums $C = (c_1, \ldots, c_n)$ as an integral. Let us define

$$F(s,t) = \exp\left\{-i\sum_{j=1}^{m} r_j\sigma_j - i\sum_{k=1}^{n} c_k\tau_k\right\} \prod_{\substack{1\le j\le m\\ 1\le k\le n}} (1 - \zeta_{jk} + \zeta_{jk}e^{i(\sigma_j+\tau_k)})$$
$$\text{for } (s,t) = (\sigma_1, \ldots, \sigma_m; \tau_1, \ldots, \tau_n)$$

and let $\Pi \subset \mathbb{R}^{m+n}$ be the parallelepiped

$$\Pi = \{(\sigma_1, \ldots, \sigma_m; \tau_1, \ldots \tau_n) : -\pi \le \sigma_j, \tau_k \le \pi \quad \text{for } j = 1, \ldots, m;\ k = 1, \ldots, n\}.$$

Let $\Pi_0 \subset \Pi$ be the facet of Π defined by the equation $\tau_n = 0$.

Since the constraints are not independent (the sum of all row sums is equal to the sum of all column sums), we can drop one of the constraints and represent the desired number $|R, C|$ as an integral over Π_0,

$$|R, C| = \frac{e^{H(z)}}{(2\pi)^{m+n-1}} \int_{\Pi_0} F(s,t)\,ds\,dt,$$

cf. Section 2 of [4].

Let $\mathcal{U} \subset \Pi$ be the neighborhood of the origin,

$$\mathcal{U} = \left\{(\sigma_1, \ldots, \sigma_m; \tau_1, \ldots, \tau_n) : |\sigma_j|,\ |\tau_k| \le \frac{\ln n}{\sqrt{n}} \quad \text{for } j = 1, \ldots, m;\ k = 1, \ldots, n\right\}$$

and let $\mathcal{U}_0$ be the intersection of $\mathcal{U}$ with the hyperplane $\tau_n = 0$. We prove that the integral

$$\int_{\Pi_0 \setminus \mathcal{U}_0} |F(s,t)|\,ds\,dt$$

is asymptotically negligible relative to the integral

$$\int_{\mathcal{U}_0} |F(s,t)|\,ds\,dt. \tag{11.1}$$

The proof is a modification of that of Theorem 8.1 (note that here we don't have another set $\mathcal{W} \subset \Pi$ contributing large values of $|F(s,t)|$) and very similar to that of Theorem 7.1 of [4]. A different line of proof can be inferred from [5].

Our next goal is to evaluate asymptotically as $m, n \longrightarrow +\infty$ the integral

$$\int_{\mathcal{U}_0} F(s,t)\,ds\,dt. \tag{11.2}$$

In particular, we need to show that (11.2) and (11.1) have about the same order of magnitude. From (3.2), we can write the expansion

$$F(s,t) = \exp\{-q(s,t) + if(s,t) + h(s,t)\}(1 + o(1)) \quad \text{for } (s,t) \in \mathcal{U},$$

as $m, n \longrightarrow +\infty$, where q, f and h are as defined by formulas (2.5.1) and (2.5.3) respectively.

We note that $q(s,t)$ is *not* strictly positive definite, since its kernel is spanned by the vector $u \in \mathbb{R}^{m+n}$ defined by (2.5.2). However, the restriction of q onto any hyperplane $L \subset \mathbb{R}^{m+n}$ which does not contain u is strictly positive definite and allows us to define the Gaussian probability measure in L with density proportional to e^{-q}. It is easy to prove (see Lemma 3.1 of [4]) that the expectation of any polynomial in the sums $\sigma_j + \tau_k$ does not depend on the choice of L. To evaluate (11.2), we need to show that asymptotically

$$\mathbf{E} \exp\{if + h\} = \exp\left\{-\frac{1}{2}\mathbf{E}f^2 + \mathbf{E}h\right\}(1 + o(1)),$$

if we choose the hyperplane L defined by the equation $\tau_n = 0$. However, since the expectation on the left hand side does not depend on the choice of the hyperplane L, we can choose L in such a way that

$$\begin{aligned} |\mathbf{E}\tau_j\tau_k|,\ |\mathbf{E}\sigma_j\sigma_k| &= O\left(\frac{1}{mn}\right) \quad \text{provided } j \neq k \\ |\mathbf{E}\sigma_j\tau_k| &= O\left(\frac{1}{mn}\right) \quad \text{for all } j, k \text{ and} \\ \mathbf{E}\sigma_j^2, \mathbf{E}\tau_k^2 &= O\left(\frac{1}{m+n}\right) \quad \text{for all } j, k. \end{aligned} \tag{11.3}$$

As is shown in [4] (see Theorem 3.2 there), to ensure (11.3), one has to choose L defined by the equation

$$\sum_{j=1}^{m} \alpha_j \sigma_j = \sum_{k=1}^{n} \beta_k \tau_k, \quad \text{where}$$

$$\alpha_j = \sum_{k=1}^{n} \left(\zeta_{jk} - \zeta_{jk}^2\right) \quad \text{and} \quad \beta_k = \sum_{j=1}^{m} \left(\zeta_{jk} - \zeta_{jk}^2\right).$$

The proof then proceeds as in Theorem 1.4.

12. PROOF OF THEOREM 2.1

In what follows, it is convenient to define the polytope $\mathcal{P}(D) \subset \mathbb{R}^{\binom{n}{2}}$ for positive, not necessarily integer, sequences $D = (d_1, \ldots, d_n)$. Recall that $\mathcal{P}(D)$ consists of the vectors $(\xi_{\{j,k\}})$ for $1 \leq j \neq k \leq n$ such that

$$\sum_{j:\ j \neq k} \xi_{\{j,k\}} = d_k \quad \text{for } k = 1, \ldots, n$$

and

$$0 \leq \xi_{\{j,k\}} \leq 1 \quad \text{for } 1 \leq j \neq k \leq n.$$

We say that $\mathcal{P}(D)$ has a *non-empty interior* if there is a point $y \in \mathcal{P}(D)$, $y = (\eta_{\{j,k\}})$, such that

$$0 < \eta_{\{j,k\}} < 1 \quad \text{for all } 1 \leq j \neq k \leq n.$$

The following two lemmas are probably known in greater generality, but since we are unable to provide a precise reference, we prove only the parts we need to obtain Theorem 2.1.

Lemma 12.1. *Let $D = (d_1, \ldots, d_n)$ be a sequence of positive rational numbers such that*

$$d_1 \geq \ldots \geq d_n$$

and the Erdős-Gallai conditions

$$\sum_{i=1}^{k} d_i \leq k(k-1) + \sum_{i=k+1}^{n} \min\{k, d_i\} \quad \text{for } k = 1, \ldots, n$$

are satisfied. Then the polytope $\mathcal{P}(D)$ is non-empty.

Proof. Let q be a positive integer such that qd_i are even integer for $i = 1, \ldots, n$. Clearly, $\mathcal{P}(D)$ is non-empty if and only if the dilated polytope $q\mathcal{P}(D)$ is non-empty. By Theorem 6.3.5 of [6] there exists an $n \times n$ symmetric non-negative integer matrix with zero trace, row/column sums $qd_1, \ldots, qd_n$ and the entries not exceeding q if and only if

$$\sum_{i=1}^{k} qd_i \leq qk(k-1) + \sum_{i=k+1}^{n} \min\{qk, qd_i\} \quad k = 1, \ldots, n.$$

Hence if the Erdős-Gallai conditions are satisfied, the polytope $q\mathcal{P}(D)$ is non-empty, and hence the polytope $\mathcal{P}(D)$ is non-empty. ∎

Next, we prove a sufficient condition for the polytope $\mathcal{P}(D)$ to have a non-empty interior.

Lemma 12.2. *Let $D = (d_1, \ldots, d_n)$ be a sequence of positive integers such that*

$$d_1 \geq \ldots \geq d_n$$

and the strict Erdős-Gallai conditions

$$\sum_{i=1}^{k} d_i < k(k-1) + \sum_{i=k+1}^{n} \min\{k, d_i\} \quad \text{for } k = 1, \ldots, n$$

are satisfied. Then $\mathcal{P}(D)$ has a non-empty interior.

Proof. For a sufficiently small rational $\epsilon \geq 0$, let us define

$$d_i(\epsilon) = \frac{d_i - (n-1)\epsilon}{1 - 2\epsilon} \quad \text{for } i = 1, \ldots, n.$$

Clearly, $d_i(0) = d_i$ and

$$d_1(\epsilon) \geq \ldots \geq d_n(\epsilon).$$

For all sufficiently small $\epsilon > 0$ we have $d_i(\epsilon) > 0$ for $i = 1, \ldots, n$ and the Erdős-Gallai conditions of Lemma 12.1 are satisfied for $d_i(\epsilon)$. Therefore, by Lemma 12.1, the polytope

$\mathcal{P}_\epsilon = \mathcal{P}(D_\epsilon)$ for $D = \big(d_1(\epsilon), \ldots, d_n(\epsilon)\big)$ is non-empty. Let $x = (\xi_{\{j,k\}})$, $x \in \mathcal{P}_\epsilon$, be a point. Then the point $y = (\eta_{\{j,k\}})$ defined by

$$\eta_{\{j,k\}} = (1-2\epsilon)\xi_{\{j,k\}} + \epsilon \quad \text{for all } 1 \leq j \neq k \leq n$$

is the desired interior point in $\mathcal{P}(D)$. ■

Next, we prove that our conditions on the minimum and maximum degree ensure that $\mathcal{P}(D)$ has a non-empty interior.

Lemma 12.3. *Let us fix real numbers $0 < \alpha < \beta < 1$ such that*

$$\beta < 2\sqrt{\alpha} - \alpha, \quad \textit{or, equivalently, } (\alpha+\beta)^2 < 4\alpha.$$

and let $D = (d_1, \ldots, d_n)$ be an integer sequence such that

$$\alpha < \frac{d_i}{n-1} < \beta \quad \textit{for } i = 1, \ldots, n.$$

Then for

$$n > \max\left\{\frac{\beta}{\alpha(1-\beta)}, \frac{4(\beta-\alpha)}{4\alpha-(\alpha+\beta)^2}\right\} + 1$$

the polytope $\mathcal{P}(D)$ has a non-empty interior.

Proof. Without loss of generality, we assume that

$$d_1 \geq \ldots \geq d_n.$$

Let us show that the strict Erdős-Gallai conditions

$$\sum_{i=1}^k d_i < k(k-1) + \sum_{i=k+1}^n \min\{k, d_i\} \quad \text{for } k = 1, \ldots, n$$

are satisfied.

We consider three different cases for k.

Suppose that $k \leq \alpha(n-1)$. Then

$$\sum_{i=1}^k d_i < k\beta(n-1) \quad \text{and} \quad \min\{k, d_i\} = k \quad \text{for all } i.$$

Therefore,

$$k(k-1) + \sum_{i=k+1}^n \min\{k, d_i\} = k(k-1) + k(n-k) = k(n-1)$$

and the strict Erdős-Gallai conditions are satisfied.

Suppose that $k \geq \beta(n-1)$. Then

$$\sum_{i=1}^k d_i < k\beta(n-1) \quad \text{and} \quad \min\{k, d_i\} = d_i > \alpha(n-1).$$

If $k \geq \beta(n-1)+1$ then $k(k-1) \geq k\beta(n-1)$ and the strict Erdős-Gallai conditions are satisfied. If $\beta(n-1) \leq k \leq \beta(n-1)+1$ then

$$\begin{aligned} k(k-1) + \sum_{i=k+1}^{n} \min\{k, d_i\} &\geq \beta(n-1)(k-1) + \alpha(n-1)(n-k) \\ &= \beta k(n-1) + (n-1)(\alpha(n-k) - \beta) \\ &\geq \beta k(n-1) + (n-1)(\alpha(n-1)(1-\beta) - \beta) \end{aligned}$$

and the strict Erdős-Gallai conditions are satisfied provided

$$n \geq \frac{\beta}{\alpha(1-\beta)} + 1.$$

Finally, suppose that $k = \gamma(n-1)$ for some $\alpha < \gamma < \beta$. Then

$$\sum_{i=1}^{k} d_i < k\beta(n-1) = \gamma\beta(n-1)^2 \quad \text{and} \quad \min\{k, d_i\} > \alpha(n-1).$$

Therefore,

$$k(k-1) + \sum_{i=k+1}^{n} \min\{k, d_i\} > \gamma^2(n-1)^2 - (\gamma - \alpha)(n-1) + (1-\gamma)\alpha(n-1)^2.$$

The minimum value of the function

$$\gamma \longmapsto \gamma^2 + (1-\gamma)\alpha - \gamma\beta$$

is attained at $\gamma = (\alpha+\beta)/2$ and equal to

$$\alpha - \frac{(\alpha+\beta)^2}{4} > 0$$

Therefore, the strict Erdős-Gallai conditions are satisfied, provided

$$n > \frac{4(\beta - \alpha)}{4\alpha - (\alpha+\beta)^2} + 1.$$

The proof now follows by Lemma 12.2. ∎

12.4. Proof of Theorem 2.1

By Lemma 12.3, the polytope $\mathcal{P}(D)$ contains a point $y = (\eta_{\{j,k\}})$ such that $0 < \eta_{\{j,k\}} < 1$ for all j, k. First, we show that the maximum entropy matrix z lies in the interior of $\mathcal{P}(D)$, that is, $0 < \zeta_{\{j,k\}} < 1$ for all j, k.

We have

$$\frac{\partial}{\partial \xi_{\{j,k\}}} H(x) = \ln \frac{1 - \xi_{\{j,k\}}}{\xi_{\{j,k\}}}.$$

We note that the value of the derivative is $+\infty$ at $\xi_{\{j,k\}} = 0$ (we consider the right derivative there), is $-\infty$ at $\xi_{\{j,k\}} = 1$ (we consider the left derivative there) and is finite for $0 < \xi_{\{j,k\}} < 1$.

Therefore, if for the maximum point z and some $j \neq k$ we have $\zeta_{\{j,k\}} \in \{0,1\}$ then for $\tilde{z} = (1-\epsilon)z + \epsilon y$ for a sufficiently small $\epsilon > 0$, we have $\tilde{z} \in \mathcal{P}(D)$ and $H(\tilde{z}) > H(z)$, which is a contradiction.

Since the maximum value of H is attained at an interior point of $\mathcal{P}(D)$, the gradient of H at the maximum point is orthogonal to the affine span of $\mathcal{P}(D)$, that is,

$$\ln \frac{1-\zeta_{\{j,k\}}}{\zeta_{\{j,k\}}} = \lambda_j + \lambda_k,$$

or, equivalently,

$$\zeta_{\{j,k\}} = \frac{1}{1+e^{\lambda_j+\lambda_k}} \quad \text{for all } 1 \leq j \neq k \leq n \tag{12.4.1}$$

for some real $\lambda_1, \ldots, \lambda_n$. Without loss of generality, we assume that

$$\lambda_1 \leq \lambda_j \leq \lambda_n \quad \text{for all } j. \tag{12.4.2}$$

From the choice of ϵ in Theorem 2.1, it follows that

$$\epsilon \leq \alpha \quad \text{and} \quad \beta \leq 2\sqrt{\alpha-\epsilon} - \alpha. \tag{12.4.3}$$

Our next goal is to show that

$$\lambda_n \leq 2 \ln \frac{1}{\epsilon}. \tag{12.4.4}$$

Aiming for a contradiction, suppose that

$$\lambda_n > 2 \ln \frac{1}{\epsilon}.$$

Then, necessarily,

$$\lambda_1 < \ln \epsilon$$

since otherwise by (12.4.1) and (12.4.2) we have

$$\zeta_{\{j,n\}} = \frac{1}{1+e^{\lambda_j+\lambda_n}} \leq \frac{1}{1+e^{\lambda_1+\lambda_n}} < \epsilon$$

and

$$d_n = \sum_{j:\, j \neq n} \zeta_{\{j,n\}} < \epsilon(n-1),$$

which by (12.4.3) contradicts the lower bound for d_i.

Since $\lambda_1 < \ln \epsilon$ and $\lambda_n > -2\ln \epsilon$, we deduce from (12.4.1) that for $1 < j < n$ we have

$$\begin{gathered} \zeta_{\{j,n\}} = \frac{1}{1+e^{\lambda_j+\lambda_n}} < \frac{1}{1+e^{\lambda_n}} < \epsilon \quad \text{provided } \lambda_j \geq 0 \\ \text{and} \\ \zeta_{\{1,j\}} = \frac{1}{1+e^{\lambda_1+\lambda_j}} > \frac{1}{1+e^{\lambda_1}} > 1-\epsilon \quad \text{provided } \lambda_j \leq 0. \end{gathered} \tag{12.4.5}$$

Denoting

$$\tau = \zeta_{\{1,n\}} = \frac{1}{1 + e^{\lambda_1 + \lambda_n}} < 1,$$

by (12.4.1) and (12.4.2) we obtain that for $1 < j < n$ we have

$$\begin{gathered} \zeta_{\{1,j\}} = \frac{1}{1 + e^{\lambda_1 + \lambda_j}} \geq \frac{1}{1 + e^{\lambda_1 + \lambda_n}} = \tau \\ \text{and} \\ \zeta_{\{j,n\}} = \frac{1}{1 + e^{\lambda_j + \lambda_n}} \leq \frac{1}{1 + e^{\lambda_1 + \lambda_n}} = \tau. \end{gathered} \tag{12.4.6}$$

Let

$$|\{1 \leq j < n : \quad \lambda_j \leq 0\}| = \gamma(n-1) \quad \text{for some } 0 \leq \gamma \leq 1.$$

Combining (12.4.5) and (12.4.6), we obtain

$$\begin{aligned} \beta(n-1) > d_1 = \sum_{j:\ j \neq 1} \zeta_{\{1,j\}} &= \sum_{j \neq 1:\ \lambda_j \leq 0} \zeta_{\{1,j\}} + \sum_{j \neq 1:\ \lambda_j > 0} \zeta_{\{1,j\}} \\ &> (1-\epsilon)\gamma(n-1) + (n-1)(1-\gamma)\tau \end{aligned}$$

and

$$\begin{aligned} \alpha(n-1) < d_n = \sum_{j:\ j \neq n} \zeta_{\{j,n\}} &= \sum_{j \neq n:\ \lambda_j > 0} \zeta_{\{j,n\}} + \sum_{j \neq n:\ \lambda_j \leq 0} \zeta_{\{j,n\}} \\ &\leq \epsilon(1-\gamma)(n-1) + (n-1)\gamma\tau. \end{aligned}$$

Consequently,

$$\beta > (1-\epsilon)\gamma + (1-\gamma)\tau \quad \text{and} \quad \alpha < \epsilon(1-\gamma) + \gamma\tau.$$

Therefore,

$$\beta + \epsilon > \gamma + (1-\gamma)\tau \quad \text{and} \quad \alpha - \epsilon < \gamma\tau.$$

Since the function $2\sqrt{x} - x$ is increasing for $0 < x < 1$, from (12.4.3) it follows that

$$\gamma + (1-\gamma)\tau < 2\sqrt{\gamma\tau} - \gamma\tau,$$

or, equivalently,

$$\frac{\gamma + \tau}{2} < \sqrt{\gamma\tau},$$

which is a contradiction.

The contradiction shows that (12.4.4) indeed holds. Then, by (12.4.2), we have

$$\lambda_j \leq 2 \ln \frac{1}{\epsilon} \quad \text{for } j = 1, \ldots, n.$$

We claim now that

$$\lambda_1 \geq 3 \ln \epsilon. \tag{12.4.7}$$

Indeed, if $\lambda_1 < 3 \ln \epsilon$ then by (12.4.1)

$$\zeta_{\{1,j\}} = \frac{1}{1 + e^{\lambda_1 + \lambda_j}} > \frac{1}{1 + \epsilon} > 1 - \epsilon \quad \text{for } j = 1, \dots, n-1$$

and

$$\beta(n-1) > d_1 = \sum_{j:\ j \neq n} \zeta_{\{1,j\}} > (n-1)(1-\epsilon),$$

which contradicts (12.4.3).

Summarizing, from (12.4.4) and (12.4.7), we obtain

$$\frac{\epsilon^4}{1+\epsilon^4} \leq \frac{1}{1+e^{2\lambda_n}} \leq \zeta_{\{j,k\}} \leq \frac{1}{1+e^{2\lambda_1}} \leq \frac{1}{1+\epsilon^6} \quad \text{for all } j \neq k,$$

which completes the proof.

ACKNOWLEDGMENTS

The authors are grateful to Brendan McKay and Sourav Chatterjee for useful comments.

REFERENCES

[1] K. Ball, An elementary introduction to modern convex geometry, Flavors of Geometry, Mathematical Sciences Research Institute Publications 31, Cambridge Univeristy Press, Cambridge, 1997, pp. 1–58.

[2] A. Barvinok, On the number of matrices and a random matrix with prescribed row and column sums and 0-1 entries, Adv Math 224 (2010), 316–339.

[3] A. Barvinok and J. A. Hartigan, Maximum entropy Gaussian approximations for the number of integer points and volumes of polytopes, Adv Appl Math 45 (2010), 252–289.

[4] A. Barvinok and J. A. Hartigan, An asymptotic formula for the number of non-negative integer matrices with prescribed row and column sums. Preprint `arXiv:0910.2477`, to appear in Trans Am Math Soc, 2009.

[5] A. Barvinok and J. A. Hartigan, Maximum entropy Edgeworth estimates of volumes of polytopes. preprint `arXiv:0910.2497`, 2009.

[6] R. A. Brualdi and H. J. Ryser, Combinatorial Matrix Theory, Encyclopedia of Mathematics and its Applications Vol. 39, Cambridge University Press, Cambridge, 1991.

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

[8] E. R. Canfield and B. D. McKay, Asymptotic enumeration of dense 0-1 matrices with equal row sums and equal column sums, Electron J Combin 12 (2005), 31.

[9] S. Chatterjee, P. Diaconis, and A. Sly, Random graphs with a given degree sequence, Ann Appl Probab 21 (2011), 1400–1435.

[10] J. Duoandikoetxea, Reverse Hölder inequalities for spherical harmonics, Proc Am Math Soc 101 (1987), 487–491.

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

[12] M. Jerrum, A. Sinclair, and B. McKay, When is a graphical sequence stable?, Random Graphs, Vol. 2 (Poznań, 1989), Wiley, New York, 1992, pp. 101–115.

[13] M. Ledoux, The Concentration of Measure Phenomenon, Mathematical Surveys and Monographs, 89, American Mathematical Society, Providence, RI.

[14] L. Lovász and B. Szegedy, Limits of dense graph sequences, J Combin Theory Ser B 96 (2006), 933–957.

[15] C. McDiarmid, On the method of bounded differences, Surveys in combinatorics (Norwich, 1989), London Mathematical Society of Lecture Note Series, 141, Cambridge Univeristy Press, Cambridge, 1989, pp. 148–188.

[16] B. D. McKay, Subgraphs of random graphs with specified degrees, In Proceedings of the International Congress of Mathematicians 2010 (ICM 2010), Hyderabad, 2010, pp. 2489–2501.

[17] B. D. McKay, Subgraphs of dense random graphs with specified degrees, Combinator Probab Comput 20 (2011), 413–433.

[18] B. D. McKay and N. C. Wormald, Asymptotic enumeration by degree sequence of graphs of high degree, Eur J Combin 11 (1990), 565–580.

[19] Y. Nesterov and A. Nemirovskii, Interior-Point Polynomial Algorithms in Convex Programming, SIAM Studies in Applied Mathematics, 13, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994.

[20] A. Zvonkin, Matrix integrals and map enumeration: An accessible introduction, Combinatorics and physics (Marseilles, 1995), Math Comput Model 26 (1997), 281–304.