ASYMPTOTIC ENUMERATION OF SYMMETRIC INTEGER MATRICES WITH UNIFORM ROW SUMS

BRENDAN D. MCKAY and JEANETTE C. MCLEOD $^{\boxtimes}$

Abstract

We investigate the number of symmetric matrices of nonnegative integers with zero diagonal such that each row sum is the same. Equivalently, these are zero-diagonal symmetric contingency tables with uniform margins, or loop-free regular multigraphs. We determine the asymptotic value of this number as the size of the matrix tends to infinity, provided the row sum is large enough. We conjecture that one form of our answer is valid for all row sums. An example appears in Figure 1.

1. Introduction

Let $M(n, \ell)$ be the number of $n \times n$ symmetric matrices over $\{0, 1, 2, \dots\}$ with zeros on the main diagonal and each row summing to ℓ. Our interest is in the asymptotic value of $M(n, \ell)$ as $n \to \infty$ with ℓ being a function of n. Alternative descriptions of the class $M(n, \ell)$ are: adjacency matrices of loop-free regular multigraphs of order n and degree ℓ, and zero-diagonal symmetric contingency tables of dimension n with uniform margins equal to ℓ.

Very little seems to be known about this problem. The asymptotic value of $M(n, 3)$ was determined by Read in 1958 [12]. According to Bender and Canfield [3], de Bruijn extended this to $M(n, \ell)$ for fixed ℓ but failed to publish it. In any case, [3] generalized the result to bounded but possibly nonequal row sums. By the method of switchings, Greenhill and McKay [7] found the asymptotic number of matrices with given small row sums over a range that includes $M(n, \ell)$ for $\ell = o(n^{1/2})$.

In this paper we treat the case of large ℓ and manage to find the asymptotics whenever $\ell > Cn/\log n$ for any $C > \frac{1}{6}$. We will use the multi-dimensional saddle-point method, which was previously applied successfully to the corresponding $\{0, 1\}$

$$
\begin{array}{ccccccccc}
0 & 4 & 1 & 3 & 2 & 2 & 3 & 1 & 4 \\
4 & 0 & 4 & 1 & 3 & 2 & 2 & 3 & 1 \\
1 & 4 & 0 & 7 & 1 & 0 & 2 & 2 & 3 \\
3 & 1 & 7 & 0 & 1 & 1 & 3 & 2 & 2 \\
2 & 3 & 1 & 1 & 0 & 7 & 1 & 3 & 2 \\
2 & 2 & 0 & 1 & 7 & 0 & 4 & 1 & 3 \\
3 & 2 & 2 & 3 & 1 & 4 & 0 & 4 & 1 \\
1 & 3 & 2 & 2 & 3 & 1 & 4 & 0 & 4 \\
4 & 1 & 3 & 2 & 2 & 3 & 1 & 4 & 0
\end{array}
$$

$n = 9$

$\ell = 20$

$\lambda = \frac{\ell}{n-1} = 2.5$

FIGURE 1. An example of a matrix counted by $M(9, 20) = 1955487489759152410696$.

problem by McKay and Wormald [11] and to the corresponding nonsymmetric problem by Canfield and McKay [5]. For the nonsymmetric problem with mixed row and column sums, see Barvinok and Hartigan [1].

Our theorem is as follows.

THEOREM 1.1. *Let a and b be positive real numbers such that $a + b < \frac{1}{2}$. Let $\ell = \ell(n)$ be such that ℓn is even and $\lambda = \ell/(n-1)$ satisfies*

$$
\lambda \geq \frac{1}{3a \log n}. \tag{1.1}
$$

Then as $n \to \infty$,

$$
\begin{aligned}
M(n, \ell) &= \sqrt{2} \, (2\pi n(1 + \lambda)^{-\ell-n+2} \lambda^{\ell+1})^{-n/2} \exp\left(\frac{14\lambda^2 + 14\lambda - 1}{12\lambda(1 + \lambda)} + O(n^{-b}) \right) \\
&= \left(\frac{\lambda^\lambda}{(1 + \lambda)^{1+\lambda}} \right)^{\binom{n}{2}} \binom{n + \ell - 2}{\ell}^n \sqrt{2} \, e^{3/4} \, (1 + O(n^{-b})).
\end{aligned} \tag{1.2}
$$

In Section 2 we express $M(n, \ell)$ as an integral in n-dimensional complex space and divide the domain of integration into three parts, then in Section 3 we estimate the integral in two of the parts. In Section 4 we show that the third part is negligible in comparison provided ℓ is bounded by a polynomial in n. We complete the proof for large ℓ in Section 5 using the theory of Ehrhart quasipolynomials.

In Section 6 we show that the form of expression (1.2) is motivated by a naïve probabilistic model. We also note that (1.2) agrees with [7], apart from the error term, when $1 \leq \ell = o(n^{1/2})$, and closely matches many exact values computed as described in Section 7. This leads us to suspect that (1.2) is true whenever $\ell > 0$, and we conjecture explicit bounds for $M(n, \ell)$ in Conjecture 7.1.

Throughout the paper, asymptotic notation like $O(f(n))$ refers to the passage of n to ∞. We will also use a modified notation $\widetilde{O}(f(n))$. A function $g(n)$ belongs to this class provided that

$$
g(n) = O(f(n)n^{a\varepsilon}),
$$

for some numerical constant a that might be different at each use of the notation.

2. An integral for $M(n, \ell)$

We now express $M(n, \ell)$ as an integral in n-dimensional complex space and outline a plan for estimating it.

We begin with a generating function in n variables $x_1, \ldots, x_n$,

$$\prod_{1 \leq j < k \leq n} (1 - x_j x_k)^{-1},$$

for which the coefficient of $x_1^{\ell_1} \cdots x_n^{\ell_n}$ is the number of $n \times n$ symmetric matrices over $\{0, 1, 2, \ldots\}$ with zeros on the main diagonal and row sums $\ell_1, \ldots, \ell_n$. In particular, $M(n, \ell)$ is the coefficient of $x_1^\ell \cdots x_n^\ell$.

Applying Cauchy's integral formula, we have

$$M(n, \ell) = \frac{1}{(2\pi i)^n} \oint \cdots \oint \frac{\prod_{1 \leq j < k \leq n} (1 - x_j x_k)^{-1}}{x_1^{\ell+1} \cdots x_n^{\ell+1}} \, dx_1 \cdots dx_n,$$

where each variable is integrated along a contour circling the origin once in the anticlockwise direction. It will suffice to take the contours to be circles; specifically, we will put $x_j = re^{i\theta_j}$ for each j, where, for reasons that will become clear in Section 3, we choose

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

This gives

$$M(n, \ell) = \frac{1}{(2\pi)^n} (\lambda^{-\lambda} (1 + \lambda)^{1+\lambda})^{\binom{n}{2}} I(n),$$

where

$$I(n) = \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \frac{\prod_{1 \leq j < k \leq n} (1 - \lambda(e^{i(\theta_j + \theta_k)} - 1))^{-1}}{e^{i\ell \sum_{j=1}^{n} \theta_j}} \, d\theta. \tag{2.1}$$

Let $F(\boldsymbol{\theta})$ be the integrand in (2.1).

The quantity $(1 - \lambda(e^{i(\theta_j + \theta_k)} - 1))^{-1}$, and thus $F(\boldsymbol{\theta})$, has greatest magnitude when $\theta_j + \theta_k \in \{0, 2\pi\}$ for each distinct pair j, k. It is easy to see that these constraints have only two solutions: $\theta_j = 0$ for all j, and $\theta_j = \pi$ for all j. We will show that the value of $I(n)$ comes mostly from the neighbourhoods of these two points; specifically, it comes from two boxes $\mathcal{R}_0, \mathcal{R}_\pi \subseteq [-\pi, \pi]^n$ defined, for sufficiently small ε, as

$$\mathcal{R}_0 = \{\boldsymbol{\theta} : |\theta_j| \leq n^{-1/2+\varepsilon}(1 + \lambda)^{-1} \text{ for all } j\} \quad \text{and}$$
$$\mathcal{R}_\pi = \{\boldsymbol{\theta} : |\theta_j + \pi| \leq n^{-1/2+\varepsilon}(1 + \lambda)^{-1} \text{ for all } j\},$$

where $\theta_j + \pi$ is taken mod 2π. Note that the operation $\theta_j \mapsto \theta_j + \pi$ for all j, which maps $\mathcal{R}_0$ and $\mathcal{R}_\pi$ onto each other, preserves $F(\boldsymbol{\theta})$ since $n\ell$ is even. Also note that $\mathcal{R}_0 \cap \mathcal{R}_\pi = \emptyset$. We denote the region outside of the boxes as

$$\mathcal{R}^c = [-\pi, \pi]^n \setminus (\mathcal{R}_0 \cup \mathcal{R}_\pi). \tag{2.2}$$

If $X \subseteq [-\pi, \pi]^n$, then we let $I_X(n) = \int_X F(\theta) \, d\theta$. For $\lambda = O(n^5)$ we will evaluate the integral $I(n)$ defined in (2.1) in the following way:

$$
\begin{aligned}
I(n) &= I_{\mathcal{R}_0}(n) + I_{\mathcal{R}_\pi}(n) + I_{\mathcal{R}^c}(n) \\
&= 2I_{\mathcal{R}_0}(n) + O(1) \int_{\mathcal{R}^c} |F(\theta)| \, d\theta \\
&= 2I_{\mathcal{R}'}(n) + O(1) \int_{\mathcal{R}^c} |F(\theta)| \, d\theta
\end{aligned}
$$

for any $\mathcal{R}'$ with $\mathcal{R}_0 \subseteq \mathcal{R}' \subseteq [-\pi, \pi]^n \setminus \mathcal{R}_\pi$.

3. The main part of the integral

In this section we estimate the value of the integral $I(n)$ in a convenient region $\mathcal{R}'$ that contains $\mathcal{R}_0$. We begin by quoting several results required for the calculation.

The following theorem, simplified from [9], estimates the value of a certain multi-dimensional integral.

THEOREM 3.1. *Let ε', ε'', ε''', $\breve{\varepsilon}$ be constants such that $0 < \varepsilon' < \varepsilon'' < \varepsilon'''$, and $\breve{\varepsilon} > 0$. The following is true if ε''' is sufficiently small.*

Let $\hat{A} = \hat{A}(n)$ be a real-valued function such that $\hat{A}(n) = \Omega(n^{-\varepsilon'})$. Let $\hat{B} = \hat{B}(n)$, $\hat{C} = \hat{C}(n)$, $\hat{E} = \hat{E}(n)$, $\hat{F} = \hat{F}(n)$, $\hat{G} = \hat{G}(n)$, $\hat{H} = \hat{H}(n)$, and $\hat{I} = \hat{I}(n)$ be complex-valued functions of n such that $\hat{B}, \hat{C}, \hat{E}, \hat{F}, \hat{G}, \hat{H}, \hat{I} = O(1)$. Suppose that $\hat{\varepsilon}(n)$ satisfies $\varepsilon'' \leq 2\hat{\varepsilon}(n) \leq \varepsilon'''$ for all n and define

$$
U_n = \{z \subseteq \mathbb{R}^n : |z_j| \leq n^{-1/2+\hat{\varepsilon}(n)} \text{ for } 1 \leq j \leq n\}.
$$

Suppose that, for $z = (z_1, z_2, \ldots, z_n) \in U_n$,

$$
\begin{aligned}
f(z) = \exp\Bigl(&-\hat{A}n \sum_{j=1}^{n} z_j^2 + \hat{B}n \sum_{j=1}^{n} z_j^3 + \hat{C} \sum_{j,k=1}^{n} z_j z_k^2 + \hat{D}n^{-1} \sum_{j,k,p=1}^{n} z_j z_k z_p \\
&+ \hat{E}n \sum_{j=1}^{n} z_j^4 + \hat{F} \sum_{j,k=1}^{n} z_j^2 z_k^2 + \hat{G}n^{1/2} \sum_{j,k=1}^{n} z_j z_k^3 \\
&+ \hat{H}n^{-1/2} \sum_{j,k,p=1}^{n} z_j z_k z_p^2 + \hat{I}n^{-3/2} \sum_{j,k,p,q=1}^{n} z_j z_k z_p z_q + \delta(z)\Bigr),
\end{aligned}
$$

where $\delta(z)$ is continuous and $\delta(n) = \max_{z \in U_n} |\delta(z)| = o(1)$. Then, provided the $O(\cdot)$ term in the following converges to zero,

$$
\int_{U_n} f(z) \, dz = \left(\frac{\pi}{\hat{A}n}\right)^{n/2} \exp\Bigl(\Theta_1 + O(n^{-1/2+\breve{\varepsilon}} + (n^{-3/4} + \delta(n))\hat{Z})\Bigr),
$$

where

$$\Theta_1 = \frac{15\hat{B}^2}{16\hat{A}^3} + \frac{3\hat{B}\hat{C}}{8\hat{A}^3} + \frac{\hat{C}^2}{16\hat{A}^3} + \frac{3\hat{E}}{4\hat{A}^2} + \frac{\hat{F}}{4\hat{A}^2} \quad and$$

$$\hat{Z} = \exp\left(\frac{15\,\mathrm{Im}(\hat{B})^2 + 6\,\mathrm{Im}(\hat{B})\,\mathrm{Im}(\hat{C}) + \mathrm{Im}(\hat{C})^2}{16\hat{A}^3}\right).$$

The following lemma defines a linear transformation, adapted from [11].

LEMMA 3.2. *Define c and* $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ *by*

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

$$(1 + \lambda)\,\theta_j = z_j - \frac{c}{n}\sum_{k=1}^{n} z_k \qquad (1 \le j \le n). \tag{3.1}$$

The transformation $\boldsymbol{\theta} = T(\mathbf{z})$ *defined by (3.1) has determinant* $(1-c)/(1+\lambda)^n$. *For* $m \ge 1$, *define* $\mu_m = \sum_{j=1}^{n} z_j^m$. *Then we have the following translations.*

$$(1 + \lambda)\sum_{j=1}^{n}\theta_j = (1-c)\mu_1,$$

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

$$(1+\lambda)^3\sum_{1\le j<k\le n}(\theta_j + \theta_k)^3 = (n-4)\mu_3 + (3(1-2c) + 12c/n)\mu_1\mu_2$$
$$+ ((-6c + 12c^2 - 4c^3)/n - 4c^2(3-c)/n^2)\mu_1^3,$$

$$(1+\lambda)^4\sum_{1\le j<k\le n}(\theta_j + \theta_k)^4 = (n-8)\mu_4 + 3\mu_2^2 + (4(1-2c) + 32c/n)\mu_1\mu_3$$
$$- (24c(1-c)/n + 48c^2/n^2)\mu_1^2\mu_2$$
$$+ (8c^2(1-c)(3-c)/n^2 + 8c^3(4-c)/n^3)\mu_1^4.$$

From Taylor's theorem with remainder we have the following lemma.

LEMMA 3.3. *For all real X,*

$$(1 - \lambda(e^{iX} - 1))^{-1} = \exp(\lambda iX - \tfrac{1}{2}\lambda(1+\lambda)X^2 - \tfrac{1}{6}i\lambda(1+\lambda)(1+2\lambda)X^3$$
$$+ \tfrac{1}{24}\lambda(1+\lambda)(1+6\lambda+6\lambda^2)X^4 + O((\lambda+\lambda^5)X^5)).$$

We now present the main result of this section.

THEOREM 3.4. *Under the conditions of Theorem 1.1, there is a region* $\mathcal{R}'$ *such that* $\mathcal{R}_0 \subseteq \mathcal{R}' \subseteq 3\mathcal{R}_0 \subseteq [-\pi, \pi]^n \setminus \mathcal{R}_\pi$ *and*

$$I_{\mathcal{R}'}(n) = \frac{1}{\sqrt{2}}\left(\frac{2\pi}{\lambda(1+\lambda)n}\right)^{n/2}\exp\left(\frac{14\lambda^2 + 14\lambda - 1}{12\lambda(\lambda+1)} + O(n^{-b})\right).$$

PROOF. Consider the transformation $\theta = T(z)$ defined by (3.1). Define

$$\mathcal{R}_z = \{z : |z_j| \leq 2n^{-1/2+\varepsilon}\} \quad \text{and} \quad \mathcal{R}' = T(\mathcal{R}_z).$$

From (3.1),

$$|\theta_j| \leq y \quad \text{for all } j \implies |z_j| \leq (1 + \lambda)(1 - c)^{-1}y \quad \text{for all } j,$$
$$|z_j| \leq y \quad \text{for all } j \implies |\theta_j| \leq (1 + \lambda)^{-1}(1 + c)y \quad \text{for all } j.$$

These imply, for $n \geq 2$, that $T^{-1}\mathcal{R}_0 \subseteq \mathcal{R}_z$ and

$$\mathcal{R}_0 \subseteq \mathcal{R}' \subseteq 3\mathcal{R}_0.$$

From Lemma 3.3 we have, for $\theta \in \mathcal{R}'$,

$$F(\theta) = \exp\Big(-A_2 \sum_{1 \leq j < k \leq n} (\theta_j + \theta_k)^2 - iA_3 \sum_{1 \leq j < k \leq n} (\theta_j + \theta_k)^3$$
$$+ A_4 \sum_{1 \leq j < k \leq n} (\theta_j + \theta_k)^4 + \widetilde{O}(n^{-1/2})\Big),$$

where

$$A_2 = \tfrac{1}{2}\lambda(1 + \lambda), \quad A_3 = \tfrac{1}{6}\lambda(1 + \lambda)(1 + 2\lambda), \quad A_4 = \tfrac{1}{24}\lambda(1 + \lambda)(1 + 6\lambda + 6\lambda^2).$$

The absence of a linear term is due to our particular choice of r in Section 2.

Using Lemma 3.2, we perform the transformation $\theta = T(z)$. This diagonalizes the quadratic form in $F(\theta)$, and $I_{\mathcal{R}'}$ becomes:

$$\frac{1}{\sqrt{2}}\Big(\frac{2\pi}{\lambda(1 + \lambda)n}\Big)^{n/2} \int_{\mathcal{R}_z} F(T(z)) \, dz,$$

where

$$F(T(z)) = \exp(-A_2 B_2 (1 + \lambda)^{-2} \mu_2 - iA_3 B_3 (1 + \lambda)^{-3} \mu_3$$
$$- iA_3 B_{1,2} (1 + \lambda)^{-3} \mu_1 \mu_2 - iA_3 B_{1,1,1} (1 + \lambda)^{-3} \mu_1^3$$
$$+ A_4 B_4 (1 + \lambda)^{-4} \mu_4 + A_4 B_{2,2} (1 + \lambda)^{-4} \mu_2^2$$
$$+ A_4 B_{1,3} (1 + \lambda)^{-4} \mu_1 \mu_3 - A_4 B_{1,1,2} (1 + \lambda)^{-4} \mu_1^2 \mu_2$$
$$+ A_4 B_{1,1,1,1} (1 + \lambda)^{-4} \mu_1^4 + \widetilde{O}(n^{-1/2})),$$

in which

$$B_2 = n - 2,$$
$$B_3 = n - 4,$$
$$B_4 = n - 8,$$
$$B_{1,2} = 3(1 - 2c) + \frac{12c}{n} = -3 + 3\sqrt{2} + O(n^{-1}),$$
$$B_{1,3} = 4(1 - 2c) + \frac{32c}{n} = -4 + 4\sqrt{2} + O(n^{-1}),$$

$$B_{2,2} = 3,$$

$$B_{1,1,1} = \frac{-6c + 12c^2 - 4c^3}{n} - \frac{4c^2(3-c)}{n^2} = \frac{2 - 2\sqrt{2}}{n} + O(n^{-2}),$$

$$B_{1,1,2} = -\frac{24c(1-c)}{n} + \frac{48c^2}{n^2} = O(n^{-1}),$$

$$B_{1,1,1,1} = \frac{8c^2(1-c)(3-c)}{n^2} + \frac{8c^3(4-c)}{n^3} = O(n^{-2}).$$

In order to apply Theorem 3.1 we choose $\hat{\varepsilon}(n) = \varepsilon + \log 2/\log n$, $\varepsilon' = \frac{1}{2}\varepsilon$, $\varepsilon'' = \varepsilon$, $\varepsilon''' = 3\varepsilon$, $\check{\varepsilon} = \varepsilon$, $\delta(n) = \widetilde{O}(n^{-1/2})$ and

$$\hat{A} = \frac{A_2 B_2}{(1+\lambda)^2 n} = -\frac{\lambda}{2(1+\lambda)}\left(1 - \frac{2}{n}\right),$$

$$\hat{B} = -i\frac{A_3 B_3}{(1+\lambda)^3 n} = -i\frac{\lambda(1+2\lambda)}{6(1+\lambda)^2} + O(n^{-1}),$$

$$\hat{C} = -i\frac{A_3 B_{1,2}}{(1+\lambda)^3} = i\frac{\lambda(1+2\lambda)(1-\sqrt{2})}{2(1+\lambda)^2} + O(n^{-1}),$$

$$\hat{D} = -i\frac{A_3 B_{1,1,1} n}{(1+\lambda)^3} = -i\frac{\lambda(1+2\lambda)(1-\sqrt{2})}{3(1+\lambda)^2} + O(n^{-1}),$$

$$\hat{E} = \frac{A_4 B_4}{(1+\lambda)^4 n} = \frac{\lambda(1+6\lambda+6\lambda^2)}{24(1+\lambda)^3} + O(n^{-1}),$$

$$\hat{F} = \frac{A_4 B_{2,2}}{(1+\lambda)^4} = \frac{\lambda(1+6\lambda+6\lambda^2)}{8(1+\lambda)^3},$$

$$\hat{G} = \frac{A_4 B_{1,3}}{(1+\lambda)^4 n^{1/2}} = O(n^{-1/2}),$$

$$\hat{H} = \frac{A_4 B_{1,1,2} n^{1/2}}{(1+\lambda)^4} = O(n^{-1/2}),$$

$$\hat{I} = \frac{A_4 B_{1,1,1,1} n^{3/2}}{(1+\lambda)^4} = O(n^{-1/2}),$$

$$\hat{Z} = \exp\left(\frac{nA_3^2(15B_3^2 + 6B_3 B_{1,2} n + B_{1,2}^2 n^2)}{16 B_2^3 A_2^3}\right)$$

$$= \exp\left(\frac{(1+2\lambda)^2}{3\lambda(1+\lambda)} + \widetilde{O}(n^{-1})\right),$$

$$\Theta_1 = \frac{2\lambda^2 + 2\lambda - 1}{12\lambda(1+\lambda)} + \widetilde{O}(n^{-1}).$$

Theorem 3.4 now follows from Theorem 3.1. □

4. Concentration of the integral

In the previous section we proved that the contribution to $I(n)$ from the box $\mathcal{R}'$ is

$$I_{\mathcal{R}'}(n) = \left(\frac{\pi}{A_2 n}\right)^n \exp(O(1 + \lambda^{-1})).$$

We now consider the contribution to $I(n)$ from the region $\mathcal{R}^c$ (defined in (2.2)) and show, provided λ is not too large, that it is negligible compared to $I_{\mathcal{R}'}(n)$.

First we import from [5] some useful lemmas.

LEMMA 4.1. *The absolute value of the integrand $F(\boldsymbol{\theta})$ of (2.1) is*

$$|F(\boldsymbol{\theta})| = \prod_{1 \le j < k \le n} f(\theta_j + \theta_k),$$

where

$$f(z) = (1 + 4A_2(1 - \cos z))^{-1/2}.$$

Moreover, for all real z with $|z| \le \frac{1}{10}(1 + \lambda)^{-1}$,

$$0 \le f(z) \le \exp(-A_2 z^2 + (\tfrac{1}{12} A_2 + A_2^2) z^4). \tag{4.1}$$

LEMMA 4.2. *Define $t = \frac{1}{60}(1 + \lambda)^{-1}$ and $g(x) = -A_2 x^2 + (\frac{3}{4} A_2 + 9 A_2^2) x^4$. Then, uniformly for $\lambda > 0$ and $K \ge 1$,*

$$\int_{-2t}^{2t} \exp(K g(x)) \, dx \le \sqrt{\pi/(A_2 K)} \, \exp(O(K^{-1} + (A_2 K)^{-1})).$$

THEOREM 4.3. *Suppose that the conditions of Theorem 1.1 hold, and in addition that $\lambda = n^{O(1)}$. Then*

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

PROOF. The proof follows a similar pattern to that of [11, Theorem 1]. Define t and $g(z)$ as in Lemma 4.2.

Define n_0, n_1, n_2, n_3, functions of $\boldsymbol{\theta}$, to be the number of indices j such that θ_j lies in $[-t, t]$, $(t, \pi - t)$, $[\pi - t, \pi + t]$, and $(-\pi + t, -t)$, respectively. Let $\mathcal{R}''$ be the set of all $\boldsymbol{\theta}$ such that

$$\max\left\{ n_0 n_2, \binom{n_1}{2}, \binom{n_3}{2} \right\} \ge n^{1+\varepsilon}.$$

Any $\boldsymbol{\theta} \in \mathcal{R}''$ has the property that $f(\theta_j + \theta_k) \le f(2t)$ for at least $n^{1+\varepsilon}$ pairs j, k. Since $f(z) \le 1$ for all z, and the volume of $\mathcal{R}''$ is less than $(2\pi)^n$,

$$\int_{\mathcal{R}''} |F(\boldsymbol{\theta})| \, d\boldsymbol{\theta} \le (2\pi)^n f(2t)^{n^{1+\varepsilon}}.$$

Applying (4.1) and the assumption that $\lambda = O(n^{O(1)})$,

$$\int_{\mathcal{R}''} |F(\boldsymbol{\theta})| \, d\boldsymbol{\theta} = O(e^{-c_1 n^{1+\varepsilon/2}}) I_{\mathcal{R}'}(n) \tag{4.2}$$

for some $c_1 > 0$.

For $\boldsymbol{\theta} \in \mathcal{R}^c \setminus \mathcal{R}''$ we must have $n_1, n_3 = O(n^{1/2+\varepsilon})$ and either $n_0 = O(n^{1/2+\varepsilon})$ or $n_2 = O(n^{1/2+\varepsilon})$. The latter two cases are equivalent, so we will assume that $n_2 = O(n^{1/2+\varepsilon})$, which implies that $n_0 = n - O(n^{1/2+\varepsilon})$.

Define S_0, S_1, S_2, functions of $\boldsymbol{\theta}$, as follows.

$$S_0 = \{j : |\theta_j| \le t\},$$
$$S_1 = \{j : t < |\theta_j| \le 2t\},$$
$$S_2 = \{j : |\theta_j| > 2t\}.$$

Define $s_i = |S_i|$ for each i. Since $s_0 = n_0$, we know that $s_1 + s_2 = O(n^{1/2+\varepsilon})$. Now we bound $|F(\boldsymbol{\theta})|$ in $\mathcal{R}^c \setminus \mathcal{R}''$ using

$$f(\theta_j + \theta_k) \le \begin{cases} f(t) \le \exp\left(-\dfrac{\lambda}{14400(1+\lambda)}\right) & \text{if } j \in S_0, k \in S_2, \\ \exp(-A_2(\theta_j+\theta_k)^2 + (\tfrac{1}{12}A_2 + A_2^2)(\theta_j+\theta_k)^4) & \text{if } j, k \in S_0, \\ 1 & \text{otherwise.} \end{cases}$$

Let $I_2(s_2)$ be the contribution to $I(n)$ from those $\boldsymbol{\theta} \in \mathcal{R}^c \setminus \mathcal{R}''$ with the given value of s_2, and let $\boldsymbol{\theta}'$ denote the vector $(\theta_j)_{j \in S_0}$. The set S_2 can be chosen in at most n^{s_2} ways. Applying the bounds above, and allowing $(2\pi)^{s_1+s_2}$ for integration over $\theta_j \in S_1 \cup S_2$,

$$I(s_2) \le n^{s_2}(2\pi)^{s_1+s_2} \exp\left(-\frac{s_0 s_2 \lambda}{14400(1+\lambda)}\right) I'(s_0), \tag{4.3}$$

where

$$\begin{aligned} I'(s_0) &= \int_{-t}^{t} \cdots \int_{-t}^{t} \prod_{j,k \in S_0, j<k} f(\theta_j + \theta_k)\, d\boldsymbol{\theta}' \\ &\le \int_{-t}^{t} \cdots \int_{-t}^{t} \exp\Big(-A_2 \sum_{j,k \in S_0, j<k} (\theta_j+\theta_k)^2 \\ &\qquad + (\tfrac{1}{12}A_2 + A_2^2) \sum_{j,k \in S_0, j<k} (\theta_j+\theta_k)^4\Big)\, d\boldsymbol{\theta}' \\ &\le \int_{-t}^{t} \cdots \int_{-t}^{t} \exp\Big(-A_2(s_0-2)\sum_{j \in S_0} \theta_j^2 \\ &\qquad + 8(s_0-1)(\tfrac{1}{12}A_2 + A_2^2)\sum_{j \in S_0} \theta_j^4\Big)\, d\boldsymbol{\theta}' \\ &\le \Big(\int_{-t}^{t} \exp(-(s_0-2)g(z))\, dz\Big)^{s_0} \quad \text{for } s_0 \ge 10, \\ &\le \Big(\sqrt{\frac{\pi}{A_2(s_0-2)}}\, \exp(O(1+\lambda^{-1})n^{-1})\Big)^{s_0} \\ &\le \Big(\frac{\pi}{A_2 n}\Big)^{n/2} \exp(O(n^{1/2+2\varepsilon})). \end{aligned}$$

The third line of the above follows from the bounds

$$\sum_{1 \le j < k \le p} (x_j + x_k)^2 \ge (p-2) \sum_{j=1}^{p} x_j^2 \quad \text{and} \quad \sum_{1 \le j < k \le p} (x_j + x_k)^4 \le 8(p-1) \sum_{j=1}^{p} x_j^4$$

valid for all $x_1, x_2, \ldots, x_p$. The fifth line follows from Lemma 4.1, and the last line follows from $s_0 = n - O(n^{1/2+\varepsilon})$. Substituting this bound into (4.3),

$$\sum_{s_2 \ge 1} I_2(s_2) = O(\exp(-c_2 n / \log n)) I_{\mathcal{R}'}(n) \tag{4.4}$$

for some $c_2 > 0$.

With the cases of (4.2) and (4.4) excluded, we are left with the problem of bounding the contribution of $\boldsymbol{\theta} \in [-2t, 2t]^n \setminus \mathcal{R}'$. Let $u = n^{-1/2+\varepsilon}/(1 + \lambda)$. First note that, by arguing as above,

$$|F(\boldsymbol{\theta})| \le \prod_{j=1}^{n} \exp((n-2) g(\theta_j))$$

for all $\boldsymbol{\theta} \in [-2t, 2t]^n$. Also note that, by Lemma 4.2,

$$\int_{-2t}^{2t} \exp((n-2) g(z)) \, dz \le \sqrt{\pi/(A_2 n)} \exp(O(1 + \lambda^{-1}) n^{-1}). \tag{4.5}$$

The function $g(z)$ has at most one minimum in $[u, 2t]$, and $g(2t) < g(u)$ for sufficiently large n, so

$$\left(\int_{-2t}^{-u} + \int_{u}^{2t} \right) \exp((n-2) g(z)) \, dz \le 4t \exp((n-2) g(u))$$

$$\le \exp\left(-\frac{\lambda}{4(1+\lambda)} n^{\varepsilon} \right). \tag{4.6}$$

Let J_1, J_2 be the right-hand sides of (4.5) and (4.6), respectively. Then

$$\int_{[-2t,2t]^n \setminus \mathcal{R}} |F(\boldsymbol{\theta})| \, d\boldsymbol{\theta} \le \sum_{q=1}^{n} \binom{n}{q} J_2^q J_1^{n-q}$$

$$= J_1^n ((1 + J_2/J_1)^n - 1)$$

$$= O(e^{-c_3 n^{\varepsilon}}) I_{\mathcal{R}'}(n) \tag{4.7}$$

for some $c_3 > 0$.

The lemma now follows from (4.2), (4.4) and (4.7). □

5. Proof of Theorem 1.1

In the case where $\lambda = n^{O(1)}$, Theorem 1.1 follows from Theorems 3.4 and 4.3. For larger λ, the method used in the proof of Theorem 4.3 is insufficient so we need a new approach.

Let us assume that we have already proved Theorem 1.1 for $\lambda = O(n^5)$. Now we want to show that it must be true for larger λ as well. First note that for such large λ (indeed for $\lambda/n \to \infty$), Theorem 1.1 is equivalent to

$$M(n, \ell) = \sqrt{2}\, (\lambda + \tfrac{1}{2})^{n(n-3)/2}\, \frac{e^{\binom{n}{2}+7/6}}{(2\pi n)^{n/2}} (1 + O(n^2/\lambda^2 + n^{-b})). \tag{5.1}$$

Let $\mathcal{P}_n$ be the polytope of symmetric $n \times n$ real nonnegative matrices with zero diagonal whose rows sum to 1. Then $M(n, \ell)$ is the number of integer points in $\ell\mathcal{P}_n$. That is, $M(n, \ell)$ is the *Ehrhart quasipolynomial* of $\mathcal{P}_n$.

According to [4, Theorem 8.2.6], the vertices of $\mathcal{P}_n$ are the adjacency matrices of graphs whose components are either isolated edges of weight 1 or odd cycles with edges of weight $\frac{1}{2}$. That is, the coordinates of the vertices are multiples of $\frac{1}{2}$. By a result of Ehrhart (see [2, Example 3.25]), there is a polynomial $f_n(z)$ with nonnegative integer coefficients such that

$$\sum_{\ell \leq 0} M(n, \ell) z^\ell = \frac{f_n(z)}{(1 - z^2)^{d+1}}, \tag{5.2}$$

where d is the dimension of $\mathcal{P}_n$. By [13], $d = n(n-3)/2$.

By applying the binomial expansion to $(1 - z^2)^{-d-1}$ in (5.2), we find that $M(n, \ell)$ is a polynomial in ℓ for even ℓ and a possibly different polynomial in ℓ for odd ℓ. Explicitly, there are nonnegative integers $h_0, \dots, h_d$ (dependent on n and the parity of ℓ) such that

$$M(n, \ell) = \sum_{i=0}^{d} h_{d-i} \binom{\ell + i}{d}.$$

Arguing as in [5], we infer that there is a function $\alpha(n, \ell)$ such that

$$M(n, \ell) = \binom{\ell}{d}\left(\sum_{i=0}^{d} h_i\right)(1 + \alpha(n, \ell)/\ell) \tag{5.3}$$

$$\alpha(n, \ell) \geq 0 \quad \text{for } \ell \geq d \tag{5.4}$$

$$\alpha(n, \ell + 2) \leq \alpha(n, \ell) \quad \text{for } \ell \geq d. \tag{5.5}$$

Equations (5.1) and (5.3) both apply for $\ell = \Theta(n^5)$, so for $m \in \{n^5, n^5 + 1\}$,

$$\frac{M(n, 3m)}{M(n, m)} = 3^d \frac{1 + \tfrac{1}{3}\alpha(n, 3m)/m^5}{1 + \alpha(n, m)/m^5} (1 + O(n^{-1})) = 3^d(1 + O(n^{-b})),$$

where the first estimate comes from (5.3) and the second comes from (5.1). Comparing these two estimates, and noting from (5.4) and (5.5) that $0 \leq \alpha(n, 3m) \leq \alpha(n, m)$, we conclude that $\alpha(n, m) = O(n^{5-b})$. By (5.5), this implies that $\alpha(n, \ell) = O(n^{5-b})$ for all $\ell \geq n^5$. Now we can see from (5.3) that

$$M(n, \ell) = M(n, m)\frac{M(n, \ell)}{M(n, m)} = M(n, m)\left(\frac{\ell}{m}\right)^d (1 + O(n^{-b}))$$

and apply (5.1) to $M(n, m)$. This shows that (5.1) holds for all $\ell \geq n^5$. The proof of Theorem 1.1 is now complete.

6. Naïve thinking

In this section we consider a 'naïve' random matrix model and show how it motivates our estimate for $M(n, \lambda)$.

Define $\mathcal{G}_\lambda$ to be the geometric distribution with mean λ. That is, for a random variable X distributed according to $\mathcal{G}_\lambda$,

$$\text{Prob}(X = j) = \frac{1}{1 + \lambda}\left(\frac{\lambda}{1 + \lambda}\right)^j \tag{6.1}$$

for $j \geq 0$.

Define $\mathcal{S} = \mathcal{S}(n, \ell)$ to be the probability space of $n \times n$ nonnegative symmetric integer matrices with zero diagonal, where each element of the upper triangle is independently chosen from $\mathcal{G}_\lambda$. Define events on $\mathcal{S}$:

$$E_j : \text{row } j \text{ has sum } \ell,$$

$$E_{\text{all}} : \bigcap_{j=1}^{n} E_j,$$

$$E_0 : \text{the whole matrix has sum } n\ell.$$

Note that $E_{\text{all}} \subseteq E_0$. Also note that each matrix in E_0 has the same probability, namely

$$P_0 = \left(\frac{1}{1 + \lambda}\right)^{\binom{n}{2}}\left(\frac{\lambda}{1 + \lambda}\right)^{n\ell/2}.$$

(Proof: Apply (6.1) to each entry in the upper triangle and use the assumed independence of the entries there. The result is independent of the actual matrix entries.) Therefore,

$$M(n, \ell) = \frac{\text{Prob}(E_{\text{all}})}{P_0}.$$

Now make a *naïve assumption* that the events E_j are independent. By symmetry, $\text{Prob}(E_j)$ is independent of j, so we get a naïve estimate of $M(n, \ell)$:

$$M_{\text{naïve}}(n, \ell) = \frac{\text{Prob}(E_1)^n}{P_0}. \tag{6.2}$$

Now consider $\text{Prob}(E_1)$. The number of possible first rows is

$$\binom{n + \ell - 2}{\ell}.$$

(This is the number of ways of writing ℓ as the sum of $n-1$ nonnegative integers.) In space $\mathcal{S}$, each such first row has probability

$$\left(\frac{1}{1+\lambda}\right)^{n-1}\left(\frac{\lambda}{1+\lambda}\right)^{\ell}.$$

Therefore,

$$\text{Prob}(E_1) = \binom{n+\ell-2}{\ell}\left(\frac{1}{1+\lambda}\right)^{n-1}\left(\frac{\lambda}{1+\lambda}\right)^{\ell}.$$

Substituting this value into (6.2),

$$M_{\text{naïve}}(n,\ell) = \left(\frac{\lambda^{\lambda}}{(1+\lambda)^{1+\lambda}}\right)^{\binom{n}{2}}\binom{n+\ell-2}{\ell}^n.$$

Therefore, formula (1.2) in Theorem 1.1 can be written

$$M(n,\ell) = M_{\text{naïve}}(n,\ell)\,\sqrt{2}\,\exp(\tfrac{3}{4}+O(n^{-b})).$$

Note that $\sqrt{2}\,e^{3/4} \approx 2.9939$.

7. Exact values

As noted in Section 5, $M(n,\ell)$ is the number of integer points in $\ell\mathcal{P}_n$, where $\mathcal{P}_n$ is the polytope defined in that section. Lattice point enumeration techniques such as the algorithm in [6] therefore allow the exact computation of $M(n,\ell)$ for small n. In practice this is feasible for $n \le 9$ or with difficulty $n \le 10$, almost irrespective of ℓ.

By interpolating the computed values, we obtain the Ehrhart quasipolynomial for small n. Recall that $M(n,\ell)$ is a polynomial $M_e(n,\ell)$ for even ℓ and a polynomial $M_o(n,\ell)$ for odd ℓ. We have $M_o(n,\ell) = 0$ if n is odd, and the following.

$$M_e(3,\ell) = 1$$
$$M_e(4,\ell) = M_o(4,\ell) = \tfrac{1}{2}\,\ell^2 + \tfrac{3}{2}\,\ell + 1$$
$$M_e(5,\ell) = \tfrac{5}{256}\,\ell^5 + \tfrac{25}{128}\,\ell^4 + \tfrac{155}{192}\,\ell^3 + \tfrac{55}{32}\,\ell^2 + \tfrac{47}{24}\,\ell + 1$$
$$M_e(6,\ell) = \tfrac{19}{120960}\,\ell^9 + \tfrac{19}{5376}\,\ell^8 + \tfrac{143}{4032}\,\ell^7 + \tfrac{5}{24}\,\ell^6 + \tfrac{4567}{5760}\,\ell^5 + \tfrac{785}{384}\,\ell^4 + \tfrac{10919}{3024}\,\ell^3$$
$$\qquad + \tfrac{955}{224}\,\ell^2 + \tfrac{857}{280}\,\ell + 1$$
$$M_o(6,\ell) = M_e(6,\ell) - \tfrac{5}{256}$$
$$M_e(7,\ell) = \tfrac{533}{3633315840}\,\ell^{14} + \tfrac{533}{86507520}\,\ell^{13} + \tfrac{279413}{2335703040}\,\ell^{12} + \tfrac{9233}{6488064}\,\ell^{11}$$
$$\qquad + \tfrac{3076459}{265420800}\,\ell^{10} + \tfrac{151339}{2211840}\,\ell^9 + \tfrac{4679131}{15482880}\,\ell^8 + \tfrac{9367}{9216}\,\ell^7 + \tfrac{43502617}{16588800}\,\ell^6$$
$$\qquad + \tfrac{478009}{92160}\,\ell^5 + \tfrac{71076539}{9123840}\,\ell^4 + \tfrac{661673}{76032}\,\ell^3 + \tfrac{1712147}{246400}\,\ell^2 + \tfrac{9649}{2640}\,\ell + 1$$

$$M_e(8, \ell) = \frac{70241}{5088422500761600}\, \ell^{20} + \frac{70241}{72691750010880}\, \ell^{19} + \frac{18703309}{585359881666560}\, \ell^{18}$$

$$+ \frac{12330581}{18582853386240}\, \ell^{17} + \frac{428460619}{44144787456000}\, \ell^{16} + \frac{33009749}{310542336000}\, \ell^{15}$$

$$+ \frac{90842880341}{100429391462400}\, \ell^{14} + \frac{5580172163}{910924185600}\, \ell^{13} + \frac{1110632463421}{33108590592000}\, \ell^{12}$$

$$+ \frac{4381892419}{29196288000}\, \ell^{11} + \frac{304644862903}{551809843200}\, \ell^{10} + \frac{22001378209}{13138329600}\, \ell^{9}$$

$$+ \frac{262880239845943}{62768369664000}\, \ell^{8} + \frac{12867890603299}{1494484992000}\, \ell^{7} + \frac{3890196991231}{269007298560}\, \ell^{6}$$

$$+ \frac{9530810537}{485222400}\, \ell^{5} + \frac{76295531167}{3592512000}\, \ell^{4} + \frac{1100694281}{61776000}\, \ell^{3}$$

$$+ \frac{50787821048}{4583103525}\, \ell^{2} + \frac{135038369}{29099070}\, \ell + 1$$

$$M_o(8, \ell) = M_e(8, \ell) - \frac{35}{1048576}\, \ell^{5} - \frac{1225}{2097152}\, \ell^{4} - \frac{13685}{3145728}\, \ell^{3} - \frac{17885}{1048576}\, \ell^{2}$$

$$- \frac{233261}{6291456}\, \ell - \frac{78057}{2097152}$$

$$M_e(9, \ell) = \frac{863924282670630091}{773269480488761839429720473600000000}\, \ell^{27}$$

$$+ \frac{863924282670630091}{715990259711816517990481920000000}\, \ell^{26}$$

$$+ \frac{10311705659720524879}{1652285214719576579978035200000000}\, \ell^{25}$$

$$+ \frac{44159888290330963}{2145824954181268285685760000}\, \ell^{24} + \frac{44603828594214317123}{9179362303997647666544640000}\, \ell^{23}$$

$$+ \frac{4134171051301720697}{47262162892436401029120000}\, \ell^{22} + \frac{2139768518991928638127}{1714327544916556728238080000}\, \ell^{21}$$

$$+ \frac{2365877475528196499}{16326928999205302173696000}\, \ell^{20} + \frac{16736477703747399001}{12005094852356839833600000}\, \ell^{19}$$

$$+ \frac{4321022180965196023}{3836214520489967616000000}\, \ell^{18} + \frac{7598908879241416557943}{98462839359242502144400000}\, \ell^{17}$$

$$+ \frac{784730469955519797477}{1737579518104279449600000}\, \ell^{16} + \frac{269041737824782010522933}{118589802110617072435200000}\, \ell^{15}$$

$$+ \frac{12598164604216578106061}{1284394167815700048000000}\, \ell^{14} + \frac{39802237244716247322233}{1085987198815174656000000}\, \ell^{13}$$

$$+ \frac{18331564888365520768}{155141028402167808000}\, \ell^{12} + \frac{11492891877126624163867}{34965496931549184004000}\, \ell^{11}$$

$$+ \frac{20646561932994651460327}{26224122269866188800000}\, \ell^{10} + \frac{1269904196062353431485303}{784756871757456998400000}\, \ell^{9}$$

$$+ \frac{35369366351570841001}{1245645828186439680000}\, \ell^{8} + \frac{60277662215801786423929}{14273025114636288000000}\, \ell^{7}$$

$$+ \frac{8959111748174759872739}{169916965650432000000}\, \ell^{6} + \frac{6214960986028675406647}{113985964457164800000}\, \ell^{5}$$

$$+ \frac{416558573311485749}{9089789829120000}\, \ell^{4} + \frac{7739053944610908107}{254233401117696000}\, \ell^{3}$$

$$+ \frac{1309315468639693}{85753329742080}\, \ell^{2} + \frac{94565099767}{17847429600}\, \ell + 1.$$

The same method would yield $M(10, \ell)$ with a plausible but large amount of computation. For completeness, we also give the Ehrhart series $L_n(x) = \sum_{\ell \geq 0} M(n, \ell) x^{\ell}$ for $n \leq 9$.

$$(1 - x^2) L_3(x) = 1$$

$$(1 - x)^3 L_4(x) = 1$$

$$(1 - x^2)^6 L_5(x) = (x^8 + 1) + 16\,(x^6 + x^2) + 41\,x^4$$

$$(1 - x)^{10}(1 + x) L_6(x) = (x^6 + 1) + 6\,(x^5 + x) + 30\,(x^4 + x^2) + 40\,x^3$$

$$(1 - x^2)^{15} L_7(x) = (x^{24} + 1) + 807\,(x^{22} + x^2) + 81483\,(x^{20} + x^4)$$
$$+ 1906342\,(x^{18} + x^6) + 15277449\,(x^{16} + x^8)$$
$$+ 50349627\,(x^{14} + x^{10}) + 74301542\,x^{12}$$

$$(1 - x)^{21}(1 + x)^6 L_8(x) = (x^{20} + 1) + 90\,(x^{19} + x) + 4726\,(x^{18} + x^2)$$
$$+ 107050\,(x^{17} + x^3) + 1261121\,(x^{16} + x^4)$$
$$+ 8761248\,(x^{15} + x^5) + 39187016\,(x^{14} + x^6)$$
$$+ 119662536\,(x^{13} + x^7) + 259344246\,(x^{12} + x^8)$$
$$+ 408811676\,(x^{11} + x^9) + 475095180\,x^{10}$$

$$(1 - x^2)^{28} L_9(x) = (x^{48} + 1) + 52524\,(x^{46} + x^2)$$
$$+ 169345602\,(x^{44} + x^4)$$
$$+ 78276428212\,(x^{42} + x^6)$$
$$+ 10217460516057\,(x^{40} + x^8)$$
$$+ 527531262668208\,(x^{38} + x^{10})$$
$$+ 13016462628712186\,(x^{36} + x^{12})$$
$$+ 172410423955058664\,(x^{34} + x^{14})$$
$$+ 1322251960254170931\,(x^{32} + x^{16})$$
$$+ 6176715510750440488\,(x^{30} + x^{18})$$
$$+ 18182086106689738044\,(x^{28} + x^{20})$$
$$+ 34470475812807166836\,(x^{26} + x^{22})$$
$$+ 42606701216240491693\,x^{24}.$$

For larger n, $\mathcal{P}_n$ has too many vertices for this method to be useful, but we can use the technique of [5] and [8]. Define $f(z) = 1 + z + z^2 + \cdots + z^\ell$. Then $M(n, \ell)$ is the coefficient of $x_1^\ell x_2^\ell \cdots x_n^\ell y^{n\ell/2}$ in $\prod_{1 \le j < k \le n} f(x_j x_k y)$.

If q is any integer greater than $\max\{n\ell/2, n^2(n-1)/2 - n\ell/2\}$, then $M(n, \ell)$ is the coefficient of the only term in

$$y^{-n\ell/2} x_1 \cdots x_n \prod_{1 \le j < k \le n} f(x_j x_k y)$$

in which each x_j appears with a power that is a multiple of $\ell + 1$ and y appears with a power that is a multiple of q. Now let p be a prime number for which $p - 1$ is a multiple of both $\ell + 1$ and q. Let α and β be a primitive $(\ell + 1)$th root and a primitive qth root of unity in $\mathbb{Z}_p$, respectively. Then, modulo p,

$$M(n, \ell) = \frac{n!}{q(\ell + 1)^n}$$
$$\times \sum_{r_0 + \cdots + r_d = n} \prod_{i=0}^{d} \frac{\alpha^{ir_i}}{r_i!} \sum_{k=0}^{q-1} \beta^{-kn\ell/2} \prod_{i=0}^{d} f(\alpha^{2i}\beta^k)^{\binom{r_i}{2}} \prod_{0 \le i < j \le d} f(\alpha^{i+j}\beta^k)^{r_i r_j},$$

where the first summation is over all nonnegative integers $r_0, r_1, \ldots, r_d$ which sum to n. Using sufficiently many primes p, we can extract the exact value of $M(n, \ell)$ using the Chinese Remainder Theorem. As an example of a value computed using this method,

$$M(19, 10) = 613329062511931789477677176839174642138032757885191693120,$$

which is about 2% higher than the estimate of Theorem 1.1.

Machine-readable versions of these exact formulas, along with many other exact values of $M(n, \ell)$, can be found in [10].

After observing a large number of exact values, we have noted that (1.2) appears to have an accuracy much wider than we can prove. We can even guess extra terms. We express our observations in the following conjecture.

CONJECTURE 7.1. *For even $n\ell$, define $\Delta(n, \ell)$ by*

$$M(n, \ell) = M_{na\ddot{i}ve}(n, \ell)\, \sqrt{2}\, \exp\left(\frac{3}{4} + \frac{3\ell + 1}{12\ell(n-1)} + \frac{\Delta(n, \ell)}{n(n-1)}\right).$$

Then $|\Delta(n, \ell)| < 1$ for $n \geq 5$, $\ell \geq 1$.

8. The minimum entry

In this section we note a simple corollary of Theorem 1.1. Choose X uniformly at random from the set $\mathcal{M}(n, \ell)$ of zero-diagonal symmetric nonnegative integer matrices of order n and row sums ℓ. Let $X_{\min}$ be the least off-diagonal entry of X. If $X_{\min} \geq k$ for integer $k \geq 0$, we can subtract k from each entry to make a matrix of row sums $\ell - (n-1)k$. This elementary observation shows that

$$\text{Prob}(X_{\min} \geq k) = \frac{M(n, \ell - (n-1)k)}{M(n, \ell)}.$$

Theorem 1.1 can thus be used to estimate this probability whenever it applies to the quantities on the right. We can provide some information even in other cases; note that (1.1) is not required for the following.

THEOREM 8.1. *Let $k = k(n) \geq 0$ and $\ell = \ell(n) \geq 0$ with $n\ell$ even. Define $a = kn^3/\ell$. Then, as $n \to \infty$,*

$$\text{Prob}(X_{\min} \geq k) \begin{cases} \to 0 & \text{if } a \to \infty, \\ \sim e^{-a/2} & \text{if } a = O(1). \end{cases}$$

PROOF. We begin with a case incompletely covered by Theorem 1.1, namely $\ell = o(n^3)$. Define $\mathcal{M}_0, \mathcal{M}_1$ to be the sets of those matrices in $\mathcal{M}(n, \ell)$ with no off-diagonal zeros, and exactly two or four off-diagonal zeros, respectively. Given $X \in \mathcal{M}_0$, choose distinct q, r, s, t and replace $a_{qr}, a_{rs}, a_{st}, a_{tq}$ (and $a_{rq}, a_{sr}, a_{ts}, a_{qt}$ consistently) by $a_{qr} - \delta, a_{rs} + \delta, a_{st} - \delta, a_{tq} + \delta$, where $\delta = \min\{a_{qr}, a_{st}\}$. This can be done in $\Theta(n^4)$ ways and creates an element of $\mathcal{M}_1$. Alternatively, if $X \in \mathcal{M}_1$, choose distinct q, r, s, t such that either a_{qr}

or a_{st} or both are 0. Then replace $a_{qr}, a_{rs}, a_{st}, a_{tq}$ (and $a_{rq}, a_{sr}, a_{ts}, a_{qt}$ consistently) by $a_{qr} + \delta, a_{rs} - \delta, a_{st} + \delta, a_{tq} - \delta$, where $1 \le \delta \le \min\{a_{rs}, a_{tq}\} - 1$. If this produces an element of $\mathcal{M}_0$, it is the inverse of the previous operation. Given a choice of $a_{qr} = 0$, s and δ can be chosen in at most ℓ ways since $\sum_s a_{rs} = \ell$, then t can be chosen in at most n ways. Similarly for $q_{st} = 0$. Therefore, this operation can be done in at most $O(\ell n)$ ways. It follows that either $|\mathcal{M}_0| = 0$ or $|\mathcal{M}_0| = o(|\mathcal{M}_1|)$, which completes this case since $\mathrm{Prob}(X_{\min} \ge k) \le \mathrm{Prob}(X_{\min} \ge 1)$ for $k \ge 1$.

In case $\ell = \Theta(n^3)$, define $k' = \min\{k, \lfloor \ell/(2n) \rfloor\}$ and estimate the value of $\mathrm{Prob}(X_{\min} \ge k')$ using (5.1). This gives the desired result when $k = k'$. For $k > k'$ the value obtained tends to 0, so again the desired result follows by monotonicity with respect to k.   □

9. Concluding remarks

In this paper we have begun the asymptotic enumeration of dense symmetric nonnegative integer matrices with given row sums, by considering the special case of uniform row sums and zero diagonal. Further cases, which can be approached by the same method, are to allow the row sums to vary, and to allow diagonals other than zero. The structure of random matrices in the class can also be investigated by specifying some forced matrix entries. We hope to return to these problems in the future.

References

[1]   A. Barvinok and J. A. Hartigan, 'An asymptotic formula for the number of nonnegative integer matrices with prescribed row and column sums' (2010), arXiv:0910.2477v2.

[2]   M. Beck and S. Robins, *Computing the Continuous Discretely* (Springer, New York, 2006).

[3]   E. A. Bender and E. R. Canfield, 'The asymptotic number of labeled graphs with given degree sequences', *J. Combin. Theory, Ser. A* **24** (1978), 296–307.

[4]   R. A. Brualdi, *Combinatorial Matrix Classes* (Cambridge University Press, Cambridge, 2006).

[5]   E. R. Canfield and B. D. McKay, 'Asymptotic enumeration of integer matrices with large equal row and column sums', *Combinatorica* **30** (2010), 655–680.

[6]   J. A. De Loera, D. Haws, R. Hemmecke, P. Huggins, J. Tauzer and R. Yoshida, *A User's Guide for LattE v1.1*, 2003 and software package LattE, available online at http://www.math.ucdavis.edu/~latte/.

[7]   C. S. Greenhill and B. D. McKay, 'Asymptotic enumeration of sparse nonnegative integer matrices with specified row and column sums', *Adv. Appl. Math.* **41** (2008), 459–481.

[8]   B. D. McKay, 'Applications of a technique for labelled enumeration', *Congr. Numer.* **40** (1983), 207–221.

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

[10]   B. D. McKay *et al.*, Tables of counts of integer matrices, available online at http://cs.anu.edu.au/~bdm/data/intmat.html.

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

[12]   R. C. Read, 'Some enumeration problems in graph theory', Doctoral Thesis, University of London, 1958.

[13]   R. P. Stanley, 'Linear homogeneous Diophantine equations and magic labelings of graphs', *Duke Math. J.* **40** (1973), 607–632.

BRENDAN D. MCKAY, Research School of Computer Science,
Australian National University, Canberra ACT 0200, Australia
e-mail: bdm@cs.anu.edu.au

JEANETTE C. MCLEOD, Department of Mathematics and Statistics,
University of Canterbury, Christchurch 8140, New Zealand
e-mail: jeanette.mcleod@canterbury.ac.nz