The expected characteristic and permanental polynomials of the random Gram matrix

Jacob G. Martin

Department of Neuroscience
Georgetown University Medical Center
Washington, DC 20007, U.S.A.

jm733@georgetown.edu

E. Rodney Canfield*

Department of Computer Science
University of Georgia
Athens, GA 30602, U.S.A.

erc@cs.uga.edu

Abstract

A $t \times n$ random matrix A can be formed by sampling n independent random column vectors, each containing t components. The *random Gram matrix* of size n, $G_n = A^T A$, contains the dot products between all pairs of column vectors in the randomly generated matrix A, and has characteristic roots coinciding with the singular values of A. Furthermore, the sequences $\det(G_i)$ and $\text{perm}(G_i)$ (for $i = 0, 1, \ldots, n$) are factors that comprise the expected coefficients of the characteristic and permanental polynomials of G_n. We prove theorems that relate the generating functions and recursions for the traces of matrix powers, expected characteristic coefficients, expected determinants $E(\det(G_n))$, and expected permanents $E(\text{perm}(G_n))$ in terms of each other. Using the derived recursions, we exhibit the efficient computation of the expected determinant and expected permanent of a random Gram matrix G_n, formed according to any underlying distribution. These theoretical results may be used both to speed up numerical algorithms and to investigate the numerical properties of the expected characteristic and permanental coefficients of any matrix comprised of independently sampled columns.

1 Introduction

Let w be a t-tall vector whose components w_i are random variables (not necessarily independent)

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_t \end{bmatrix}.$$

Next, sample n independent vectors $w^{(1)}, \ldots, w^{(n)}$; creating a $t \times n$ matrix A. Then, the *random Gram matrix* of size n,

$$G_n = A^T A \,,$$

has a distribution that depends on the underlying distribution of the random vector w. (The symbol T as a superscript is used to denote transpose.)

Some general features and convergence properties of the eigenvalues of certain random Gram matrices were derived by Fannes and Spincemaille [9]. Fyodorov formulated correlation functions for permanental polynomials of certain random matrices and noted some similarities and differences between their characteristic and permanental polynomials [10]. The paper's main result, which treats the particular case of the Gaussian unitary ensemble, is derived from an expression of an arbitrary $n \times n$ permanent as a $2n$-dimensional contour integral (see Lemma 2.1). The integrand in the latter formula is the exponential of the trace of a certain tensor product. Likewise the exponential of the trace appears also in our formulas. We present combinatorial theory and an efficient algorithm for calculating $E(\det(G_n))$ and $E(\text{perm}(G_n))$, which are factors comprising the coefficients of the expected characteristic and expected permanental polynomials of G_n.

The computation of the determinant is equivalent to matrix multiplication and is therefore contained in the complexity class P (see Chapter 16 of [5]). Currently, the fastest asymptotic algorithm for matrix multiplication is $O(n^{2.376})$ [8], with a recent unpublished work [23] claiming an improvement to $O(n^{2.3727})$. Some researchers have suggested that group theoretic observations imply that $O(n^2)$ algorithms also exist [20].

At the other complexity extreme, even though the sign is the only difference between the formula for the determinant,

$$\det(A) = \sum_{\sigma \in S_n} (-1)^{\text{sgn}(\sigma)} \prod_{i=1}^{n} A_{i,\sigma(i)}, \tag{1}$$

and the formula for the permanent,

$$\text{perm}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} A_{i,\sigma(i)}, \tag{2}$$

the computation of the permanent is #P-Complete [21, 3]. The standard reference for properties of permanents is Minc [16]. The most efficient algorithm currently known for

calculating the exact permanent has complexity $O(2^n n^2)$, due to Ryser [18]. Jerrum, Sinclair and Vigoda provided a fully-polynomial randomized approximation scheme for approximating permanents of nonnegative matrices [13, 14]. Matrix permanents have found applications in physics for calculating Bose-Einstein corrections [24] and in quantum computing for encoding quantum circuit amplitudes [17, 4]. Permanental polynomials have been used as invariants for chemical structures [6, 7, 15].

The rest of the paper is organized as follows: Section 2 is a statement of results, Section 3 contains all proofs, Section 4 reports on some numerical experiments, Section 5 points out a connection to prior work involving the cycle index polynomial of the symmetric group, and Section 6 presents summary and conclusions.

2   Statement of Results

Before stating our results, we explain all notation. Let w be a t-tall vector whose components w_i are random variables (not necessarily independent). Let A be a $t \times n$ matrix whose columns are a random sample of n vectors $w^{(1)}, \ldots, w^{(n)}$. Let $G_n = A^T A$; we call G_n the random Gram matrix of size n, it being understood that the exact distribution of G_n depends on the underlying distribution on t-dimensional vectors $w = (w_1, \ldots, w_t)^T$. Although G_n is an $n \times n$ matrix, its rank is at most t, and generally speaking we take the viewpoint henceforth that n is much larger than t. One may even regard t as fixed, and $n \to \infty$, as we study the effect of taking larger and larger samples. We are especially interested in two expected values, the determinant and the permanent of G_n; these are denoted a_n, p_n respectively:

$$
\begin{aligned}
a_n &= E(\det(G_n)) \\
p_n &= E(\operatorname{perm}(G_n)).
\end{aligned}
$$

We define M to be the $t \times t$ matrix of underlying second moments,

$$
M_{ij} = E(w_i w_j), \quad 1 \leqslant i, j \leqslant t,
$$

and define the infinite sequence t_n as the traces of the powers of M:

$$
t_n = \operatorname{trace}(M^n).
$$

Finally, we define c_i, $0 \leqslant i \leqslant t$, to be the sign-adjusted coefficients of the characteristic polynomial of M, with the familiar indexing:

$$
\det(\lambda I - M) = c_0 \lambda^t - c_1 \lambda^{t-1} + \cdots + (-1)^t c_t.
$$

Theorem 1 *Let a_n, p_n, t_n denote $E(\det(G_n))$, $E(\operatorname{perm}(G_n))$, $\operatorname{trace}(M^n)$, respectively, as given above. Then,*

$$
\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} = \exp\left\{ \frac{t_1 x}{1} - \frac{t_2 x^2}{2} + \frac{t_3 x^3}{3} - \ldots \right\}. \tag{3}
$$

and

$$\sum_{n=0}^{\infty} p_n \frac{x^n}{n!} = \exp\left\{\frac{t_1 x}{1} + \frac{t_2 x^2}{2} + \frac{t_3 x^3}{3} + \dots\right\}. \tag{4}$$

The generating function identities in the previous theorem lead immediately to recursions for the sequences a_n, p_n as given in the corollary:

Corollary 2 *Let a_n, p_n, t_n denote $E(\det(G_n))$, $E(\text{perm}(G_n))$, $\text{trace}(M^n)$, respectively, as given above. Then, we have the recursions*

$$a_0 = 1$$
$$a_{n+1} = \sum_j \binom{n}{j}(-1)^j j! a_{n-j} t_{j+1} \tag{5}$$

and

$$p_0 = 1$$
$$p_{n+1} = \sum_j \binom{n}{j} j! p_{n-j} t_{j+1} \tag{6}$$

The next theorem relates the expected values $E(\det(G_n))$, $E(\text{perm}(G_n))$ to the coefficients c_i of the characteristic polynomial for M.

Theorem 3 *Let G_n, M, c_n be respectively the random Gram matrix of size n, the underlying $t \times t$ matrix of second moments, and the sign-adjusted coefficients of the characteristic polynomial,*

$$\det(\lambda I - M) = c_0 \lambda^t - c_1 \lambda^{t-1} + \dots + (-1)^t c_t.$$

Then

$$E(\det(G_n)) = n! c_n.$$

and

$$E(\text{perm}(G_n)) = n! \times [x^n] \left(1 - c_1 x + c_2 x^2 - \dots\right)^{-1}.$$

The last theorem concerns the expected values of the coefficients of the characteristic and permanental polynomials of G_n.

Theorem 4 *Let a_n, p_n denote $E(\det(G_n))$, $E(\text{perm}(G_n))$, respectively, as given above. Let b_i, d_i be the sign-adjusted coefficients of, respectively, the characteristic and permanental polynomials G_n:*

$$\begin{aligned}
\det(\lambda I - G_n) &= b_0 \lambda^n - b_1 \lambda^{n-1} + \dots + (-1)^n b_n \\
\text{perm}(\lambda I - G_n) &= d_0 \lambda^n - d_1 \lambda^{n-1} + \dots + (-1)^n d_n.
\end{aligned}$$

Then,

$$E(b_i) = \binom{n}{i} a_i \tag{7}$$

and

$$E(d_i) = \binom{n}{i} p_i. \tag{8}$$

Remark. The characteristic polynomials $\det(\lambda I - AA^T)$ and $\det(\lambda I - A^T A)$ have exactly (including multiplicity) the same nonzero roots. With A a $t \times n$ matrix, and assuming $n \geqslant t$, then, the latter characteristic polynomial has a factor of λ^{n-t}, and so $b_i = 0$ for $i > t$. This is consistent with the fact that the a_i are nonzero for at most $0 \leqslant i \leqslant t$.

3 Proofs

Proof of Theorem 1: The Leibniz formula for the determinant is

$$\det(G_n) = \sum_{\sigma \in S_n} (-1)^{\mathrm{sgn}(\sigma)} \mathrm{term}_\sigma , \tag{9}$$

with

$$\mathrm{term}_\sigma = \prod_{i=1}^{n} (G_n)_{i,\sigma(i)} ,$$

where S_n is the *symmetric group* on $\{1, 2, \ldots, n\}$, and $\mathrm{sgn}(\sigma)$ signifies the *sign* of the permutation σ. Similarly, for the permanent,

$$\mathrm{perm}(G_n) = \sum_{\sigma \in S_n} \mathrm{term}_\sigma . \tag{10}$$

Since expectation is a linear operator, $a_n = E(\det(G_n))$ may be obtained by the following strategy

$$(*) \begin{cases} 1. \text{ Determine } E(\mathrm{term}_\sigma) \\ 2. \text{ Multiply by } (-1)^{\mathrm{sgn}(\sigma)} \\ 3. \text{ Sum over } \sigma \in S_n . \end{cases}$$

Furthermore, $p_n = E(\mathrm{perm}(G_n))$ can be obtained in the same manner but omitting step 2.

Suppose the permutation σ contains k_i cycles of size i, where $k_i \geqslant 0$ and $n = k_1 + 2k_2 + \cdots$. The cycle structure of σ alone is enough to determine its sign by the relation: $\mathrm{sgn}(\sigma) = k_2 + k_4 + \cdots$. What can be said about the expected value $E(\mathrm{term}_\sigma)$, given only the cycle structure of σ ? We claim that, like the sign, the latter expected value is determined completely by the cycle structure, as given in the relation

$$E(\mathrm{term}_\sigma) = \prod_{i \geqslant 1} (t_i)^{k_i}. \tag{11}$$

Indeed, if $(i_1, i_2, \ldots, i_l)$ is a cycle of σ having length l, then the quantity product$_C$, defined by

$$\text{product}_C = (G_n)_{i_1,i_2}(G_n)_{i_2,i_3} \ldots (G_n)_{i_l,i_1} \ ,$$

is a *subproduct* of term$_\sigma$. Moreover, it is seen that the various subproducts associated with the different cycles comprising σ have no rows or columns of the matrix G_n in common. These subproducts are consequently independent, and we have

$$E(\text{term}_\sigma) = \prod_{C:C \text{ is a cycle of } \sigma} E(\text{product}_C) \ .$$

The entry $(G_n)_{ij}$ of the Gram matrix is the dot product $w^{(i)} \cdot w^{(j)}$ of columns in the sample matrix A, and so

$$\text{product}_C = (w^{(i_1)} \cdot w^{(i_2)}) \times (w^{(i_2)} \cdot w^{(i_3)}) \times \cdots \times (w^{(i_l)} \cdot w^{(i_1)}) \ .$$

From this we observe that the expectation $E(\text{product}_C)$ depends only on the length of the cycle C, and not on the particular columns of A which are involved. That the common value of $E(\text{product}_C)$ over all cycles C of length l is equal to t_l, the trace of the power M^ℓ is seen as follows

$$E\left[\left(w^{(1)} \cdot w^{(2)}\right)\left(w^{(2)} \cdot w^{(3)}\right) \ldots \left(w^{(n)} \cdot w^{(1)}\right)\right]$$

$$= E\left[\left(w_1^{(1)}w_1^{(2)} + \cdots + w_t^{(1)}w_t^{(2)}\right) \times \left(w_1^{(2)}w_1^{(3)} + \cdots + w_t^{(2)}w_t^{(3)}\right) \times \cdots \right.$$

$$\left. \cdots \times \left(w_1^{(n)}w_1^{(1)} + \cdots + w_t^{(n)}w_t^{(1)}\right)\right]$$

$$= E\left[\sum_{(i_1,\ldots,i_n)\in\{1,\ldots,t\}^n} w_{i_1}^{(1)}w_{i_1}^{(2)}w_{i_2}^{(2)}w_{i_2}^{(3)} \ldots w_{i_n}^{(n)}w_{i_n}^{(1)}\right]$$

$$= \sum_{(i_1,\ldots,i_n)\in\{1,\ldots,t\}^n} E\left(w_{i_1}^{(1)}w_{i_n}^{(1)}\right) E\left(w_{i_1}^{(2)}w_{i_2}^{(2)}\right) E\left(w_{i_2}^{(3)}w_{i_3}^{(3)}\right) \ldots E\left(w_{i_{n-1}}^{(n)}w_{i_n}^{(n)}\right)$$

$$= \sum_{(i_1,\ldots,i_n)\in\{1,\ldots,t\}^n} M_{i_1 i_n} M_{i_1 i_2} M_{i_2 i_3} \ldots M_{i_{n-1} i_n}$$

$$= \text{trace}(M^n)$$

$$= t_n \ .$$

Thus the claim (11) is justified.

Continuing the proof, we introduce the sign of the permutation to obtain

$$(-1)^{\text{sgn}(\sigma)} E(\text{term}_\sigma) = (-1)^{k_2+k_4+\cdots} \prod_{i \geqslant 1}(t_i)^{k_i}. \tag{12}$$

We are now in position to carry out the three-step strategy (*) proposed above. The number of permutations σ which have a given cycle structure $(k_1, k_2, \ldots)$ is

$$\frac{n!}{1^{k_1}2^{k_2}\ldots k_1! k_2! \ldots} \ , \qquad n = k_1 + 2k_2 + \cdots \ .$$

If we multiply the right side of (12) by the latter multiplicity and by $x^n/n!$ – note the resulting cancellation of $n!$ – and then sum over all sequences $(k_1, k_2, \dots)$ of nonnegative integers which are zero from some point on, we obtain the desired exponential generating function. Hence,

$$
\begin{aligned}
\sum_{n=0}^{\infty} a_n \frac{x^n}{n!}
&= \sum_{(k_1, k_2, \dots)} (-1)^{k_2 + k_4 + \cdots} x^{k_1 + 2k_2 + \cdots} \prod_{i \geqslant 1} \frac{t_i^{k_i}}{i^{k_i} k_i!} \\
&= \prod_{i \geqslant 1} \sum_{k_i=0}^{\infty} \frac{((-1)^{i-1} t_i x^i / i)^{k_i}}{k_i!} \\
&= \exp\left(\frac{t_1 x}{1} - \frac{t_2 x^2}{2} + \cdots \right),
\end{aligned}
$$

as was to be shown. This completes the proof of the first part of Theorem 1. The second part of the theorem, equation (4) giving the exponential generating function of the sequence of permanents p_n, is proven in a similar manner. ∎

Proof of Corollary 2: These are proven in the standard manner by comparing the coefficients of x^n on both sides of the identities obtained from (3) and (4) by differentiating with respect to x. ∎

In the next proof of Theorem 3 we use the identity

$$
\det(\exp(B)) = \exp(\text{trace}(B)), \tag{13}
$$

valid for any complex square matrix B. See, for example, Section 1.1.10, item 7, page 11 of [[12]], where the identity is attributed to Jacobi.

Proof of Theorem 3: We start with the expected determinant,

$$
\begin{aligned}
E(\det(G_n)) &= \left[\frac{x^n}{n!} \right] \exp\left\{ \frac{t_1 x}{1} - \frac{t_2 x^2}{2} + \frac{t_3 x^3}{3} - \cdots \right\}. \\
&= \left[\frac{x^n}{n!} \right] \exp\left\{ \text{trace}\left(\frac{M^1 x}{1} - \frac{M^2 x^2}{2} + \frac{M^3 x^3}{3} - \cdots \right) \right\} \\
&= \left[\frac{x^n}{n!} \right] \det\left(\exp\left\{ \frac{M^1 x}{1} - \frac{M^2 x^2}{2} + \frac{M^3 x^3}{3} - \cdots \right\} \right) \\
&= \left[\frac{x^n}{n!} \right] \det\left(\exp\left\{ \log\left(I + xM \right) \right\} \right) \\
&= \left[\frac{x^n}{n!} \right] \det\left(I + xM \right) \\
&= \left[\frac{x^n}{n!} \right] (-x)^t \cdot \det\left(\lambda I - M \right)\big|_{\lambda = -\frac{1}{x}}
\end{aligned}
$$

$$= \left[\frac{x^n}{n!} \right] (-x)^t \left(\lambda^t - c_1 \lambda^{t-1} + \cdots \right) |_{\lambda = -\frac{1}{x}}$$

$$= \left[\frac{x^n}{n!} \right] (-x)^t \left(\left(-\frac{1}{x} \right)^t - c_1 \left(-\frac{1}{x} \right)^{t-1} + \cdots \right)$$

$$= \left[\frac{x^n}{n!} \right] \left(1 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots \right)$$

$$= n! c_n \ .$$

The first equality comes from (3), the third from (13), and the rest are straightforward manipulations. The proof for the expected permanent is similar:

$$E(\mathrm{perm}(G_n)) = \left[\frac{x^n}{n!} \right] \exp \left\{ \frac{t_1 x}{1} + \frac{t_2 x^2}{2} + \frac{t_3 x^3}{3} + \ldots \right\}$$

$$= \left[\frac{x^n}{n!} \right] \exp \left\{ \mathrm{trace} \left(\frac{M^1 x}{1} + \frac{M^2 x^2}{2} + \frac{M^3 x^3}{3} + \ldots \right) \right\}$$

$$= \left[\frac{x^n}{n!} \right] \det \left(\exp \left\{ \frac{M^1 x}{1} + \frac{M^2 x^2}{2} + \frac{M^3 x^3}{3} + \ldots \right\} \right)$$

$$= \left[\frac{x^n}{n!} \right] \det \left(\exp \left\{ \log \left(I - xM \right)^{-1} \right\} \right)$$

$$= \left[\frac{x^n}{n!} \right] \det \left(I - xM \right)^{-1}$$

$$= \left[\frac{x^n}{n!} \right] \frac{1}{\det \left(I - xM \right)}$$

$$= \left[\frac{x^n}{n!} \right] \frac{1}{(x)^t \cdot \det \left(\lambda I - M \right) |_{\lambda = \frac{1}{x}}}$$

$$= \left[\frac{x^n}{n!} \right] \frac{1}{(x)^t \left(\left(\frac{1}{x} \right)^t - c_1 \left(\frac{1}{x} \right)^{t-1} + \cdots \right)}$$

$$= \left[\frac{x^n}{n!} \right] \frac{1}{1 - c_1 x + c_2 x^2 - c_3 x^3 + \cdots}$$

This time the first equality follows from (4), the third again is from (13), and, as was the case with the determinant, the rest are straightforward manipulations. ∎

Remark. The function

$$\zeta_G(u) = \frac{1}{\det(I - Tu)},$$

where T is Hashimoto's edge adjacency operator, is called the *Ihara zeta-function of the graph G*, see [22], [19].

Proof of Theorem 4: For simplicity, let us assume the probability distribution on vectors w is discrete; say, v_1, v_2, ... with probabilities $p(v_1)$, $p(v_2)$, ... In order to

obtain a term λ^{n-t} in the expansion of $\det(\lambda I - A^T A)$, we choose the $n - t$ λ's from the main diagonal, and then expand the remaining principal submatrix of size t. Since the remaining submatrix is that of $-A^T A$, we obtain the sign-adjustment $(-1)^t$ and find

$$b_i = \sum_{\substack{i \times i \\ \text{principal} \\ \text{submatrices } \alpha}} \det(\alpha) \ .$$

Since expectation is a linear operator, the expected value of the i'th coefficient of the characteristic polynomial of $A^T A$ is

$$
\begin{aligned}
E(b_i) &= E\left(\sum_{\substack{i \times i \\ \text{principle} \\ \text{submatrices } \alpha}} \det(\alpha) \right) \\
&= \binom{n}{i} \sum_{(v_1, \ldots, v_i)} \det \begin{pmatrix} (v_1 \cdot v_1) & (v_1 \cdot v_2) & \cdots & (v_1 \cdot v_i) \\ (v_2 \cdot v_1) & (v_2 \cdot v_2) & \cdots & (v_2 \cdot v_i) \\ \vdots & \vdots & \vdots & \vdots \\ (v_i \cdot v_1) & (v_i \cdot v_2) & \cdots & (v_i \cdot v_i) \end{pmatrix} p(v_1) \ldots p(v_i) \\
&= \binom{n}{i} E(\det(G_i)) \ .
\end{aligned}
$$

This proves the first assertion of the theorem, and the second assertion regarding the permanental polynomial is demonstrated in a similar manner. ∎

4 Experimental results

We wrote a Matlab program to compare the expected characteristic and permanental polynomials given by Theorem 4 to those of randomly sampled matrices of various sizes. We computed all permanents using a programmatic link with Maple via the Maple Toolbox for Matlab, and all characteristic polynomials using the Matlab command *poly*.

There are, of course, infinitely many different distributions which might underly the vectors w. We chose to use one based on *sample counts*. This is an easily understood distribution, and of interest for possible applications. The idea is to assume a set X of size t, $\{x_1, \ldots, x_t\}$, with probabilities $p_1, \ldots, p_t$, where $\sum_{i=1}^{t} p_i = 1$. We take a sample, with replacement, from X of size ℓ, and record w_i as the number of times that the element x_i is chosen. Then, each of our t-tall vectors w is integral, satisfies $\sum_{i=1}^{t} w_i = \ell$, and the distribution on these vectors is the familiar multinomial distribution:

$$\text{Prob} \left\{ \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_t \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_t \end{bmatrix} \right\} = \binom{l}{b_1 \ldots b_t} p_1^{b_1} \cdots p_t^{b_t} \ .$$

The corresponding matrix M of second moments is found to be

$$M_{ij} = E(w_i w_j) = \begin{cases} \ell(\ell-1)p_i^2 + \ell p_i & \text{if } i = j \\ \ell(\ell-1)p_i p_j & \text{if } i \neq j \end{cases}.$$

We have derived the matrix M for several other scenarios which seem natural for applications, but do not report any of these results in the present paper, with one exception. Namely, suppose that the random vectors w are generated as counts, much as above, except the sample size ℓ is also a random quantity. That is, the w come about by a compound process. If we assume the sample size ℓ to be given by a distribution $\text{prob}(\ell)$, then

$$M_{ij} = \begin{cases} \sum_\ell \left(\ell(\ell-1)p_i^2 + \ell p_i \right) \text{prob}(\ell), & i = j \\ \sum_\ell \left(\ell(\ell-1)p_i p_j \right) \text{prob}(\ell), & i \neq j \end{cases}.$$

In our experiments, we used the above multinomial distribution to generate random vectors with $\ell = 10$, $t = 4$ and $p = [3/8, 1/4, 1/4, 1/8]$.

We know, theoretically, that $a_i = 0$ for $i > t$. However, we computed these for confirmation. Thus, we computed t_i for i up to 7.

i	t_i
1	565/16
2	210825/256
3	93917125/4096
4	42581180625/65536
5	19338382478125/1048576
6	8784040432265625/16777216
7	3990026079685703125/268435456

Then, from recursion (5), we compute a_i. We also generated 1000 matrices $G_{4000} = A^T A$ at random (by sampling 4000 random vectors from the counting distribution described above to form A) and computed the mean and standard deviation of the appropriate coefficient in $\det(\lambda I - G_{4000})$, divided by a binomial coefficient as given in equation (7). The exact values and the sample values were then compared. Here are the results:

i	a_i, recursion	a_i, sample	std dev
1	35.31	35.32	0.12
2	423.44	423.59	5.91
3	2648.44	2649.99	67.54
4	7031.25	7035.47	263.93
5	0.0	0.0	0.0
6	0.0	0.0	0.0

Furthermore, boxplots of the sample coefficient distributions as the sample size increases are depicted in Figure 1. These plots show that, for this distribution, the standard deviation decreases according to a power law.
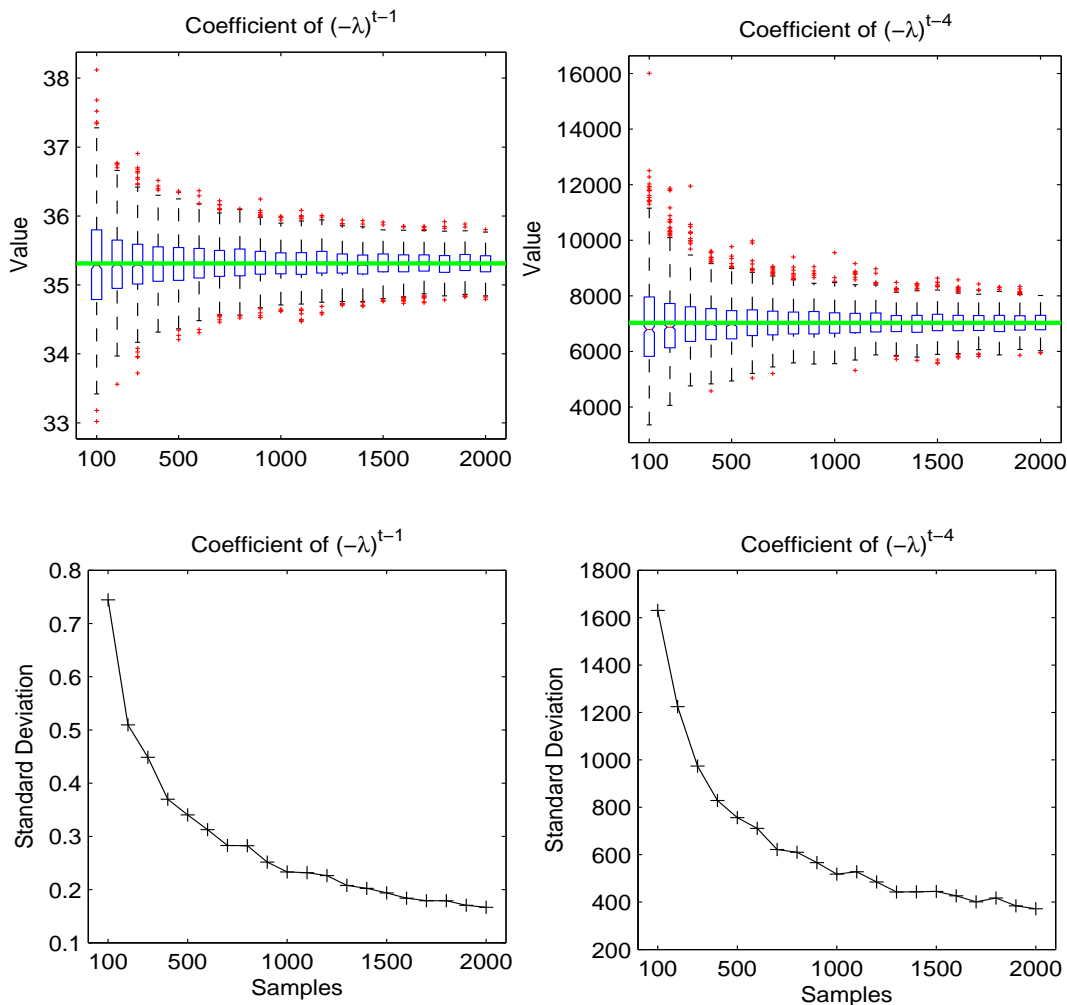
Figure 1: Experimental results of the coefficients of **characteristic** polynomials of 1000 matrices for particular sample sizes (columns) with the model $l = 10$, $t = 4$ and $p = [3/8, 1/4, 1/4, 1/8]$. To normalize the results, each coefficient of $(-\lambda)^{t-i}$ was divided by $\binom{n}{i}$ before plotting. Green line represents $E(\det(G_i))$ as computed by the recursion (5).

Using the same t_n as above, we used the permanental coefficient recursion (6) to compute the exact values of $p_n = E(\text{perm}(G_n))$. We also computed the coefficients of permanental polynomials of 1000 random gram matrices $G_{18} = A^T A$ (created by sampling 18 random vectors from the distribution) and subsequently compared the sampled results with those provided by the recursion. We computed the mean and standard deviation after division by the binomial coefficient as given in (8). The exact and sampled values were then compared. Due to the intractable computational complexity of computing the exact permanent, we were computationally limited to computing only matrices with 18 randomly sampled columns. Here are the results (see Figure 2 for boxplots):

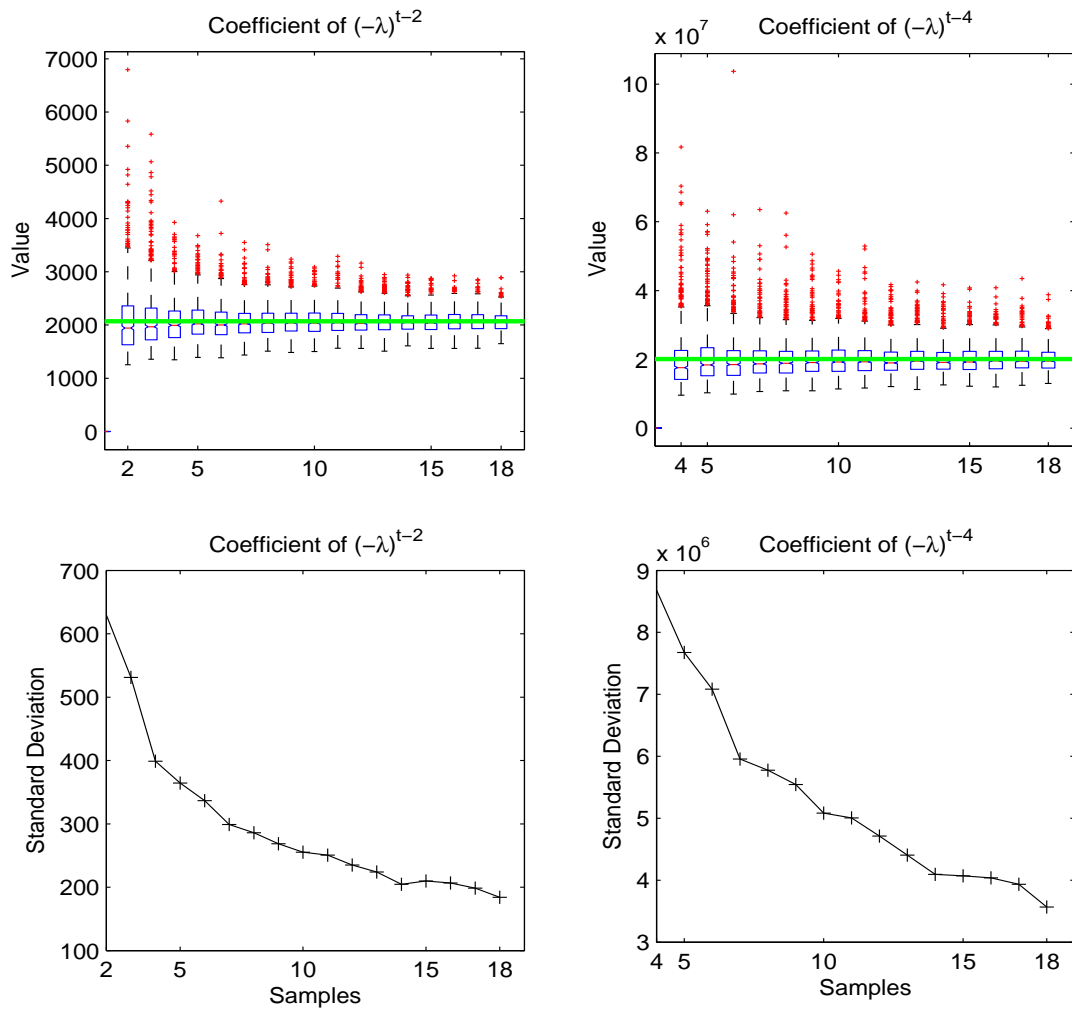| i	p_i, recursion	p_i, sample	std dev
1	35.31	35.27	1.921
2	2070.51	2071.54	206.63
3	177134.95	177679.22	26299.36
4	20126988.14	20245985.44	4037955.48
5	2857210195.90	2882490271.36	729446452.08
6	486697830067.95	492457249647.11	151888811726.52

Figure 2: Experimental results of the **permanental** coefficients from 1000 matrices for particular sample sizes (columns) with the model $l = 10$, $t = 4$ and $p = [3/8, 1/4, 1/4, 1/8]$. To normalize the results, each coefficient of $(-\lambda)^{t-i}$ was divided by $\binom{n}{i}$ before plotting. Green line represents $E(\text{perm}(G_i))$ as computed by the recursion (6).

5   Connection to the cycle index polynomial

For a permutation $\sigma \in S_n$ belonging to the symmetric group of order n, let $N_i(\sigma)$ denote the number of cycles in σ of size i. Let $X_1, X_2, \ldots$ be a countably infinite sequence of variables, and define the polynomial $P_n(X_1, \ldots, X_n)$ by

$$P_n(X_1, \ldots, X_n) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} X_i^{N_i(\sigma)}.$$

Then the quotient $P_n(X_1, \ldots, X_n)/n!$ is called the *cycle index polynomial of the symmetric group*. The generating function identity

$$\sum_{n=0}^{\infty} P_n(X_1, \ldots, X_n)\frac{u^n}{n!} \quad = \quad \exp\left(\sum_{i=1}^{\infty} \frac{X_i u^i}{i} \right)$$

was observed in [2]. The latter paper was devoted to proving that assigning nonnegative real values to the variables X_i subject to certain inequalities would result in the real values $P_n(X_1, \ldots, X_n)$ satisfying similar inequalities. Coincidentally, the pairs of quantities $(-1)^{n-1}t_n, a_n$ and t_n, p_n studied in this paper satisfy identical generating function identities. In particular, the sequence of expected permanents $p_n = E(\mathrm{perm}(A^T A))$ are hereby identified as evaluations of the cycle index polynomials at certain weights t_i.

6   Summary and conclusion

We have introduced the notion of a random Gram matrix, and provided theory enabling the efficient computation of the expected determinant and expected permanent of it. The random Gram matrix consists of dot products of vectors taken from various distributions. We further proved generating function identities and recursions relating these expectations to the traces of powers of a second moment matrix. The expected coefficients of the characteristic and permanental polynomials have also been studied, with some numerical experiments checking on the theory. Some of the formulas found are the same as those studied in earlier work in an entirely different context [2].

We have observed empirically that as the number of columns in the sample matrix A increases, the standard deviation of the normalized expected coefficients of the determinantal and permanental polynomials decreases according to a power law. Although the empirical data presented in this paper was limited to the multinomial counting model, the theoretical relationships between the different quantities remain no matter which representation is used. In future work, the theoretical rate of convergence should be formulated according to the representation and probability model used to generate the matrix A (e.g. trivially, when A=0, the truth converges immediately to the expected value).

Can the probabilistic results presented in this work be of any help in managing the complexity of computing the permanent? Already, [1], there is a polynomial time algorithm for computing the permanent of an $n \times n$ matrix of rank t, t being fixed. One way

for our probabilistic methods to impact complexity considerations would be via finding a distribution on t-vectors (t small) such that a given $n \times n$ permanent per(H) is equal to or well approximated by the expected value of per($A^T A$). We have no ideas in this direction.

It is hoped that the theoretical observations we have made will prove useful in processing and comparing large amounts of numerical data, such as those algorithms that use permanental polynomials of large chemical graphs [6, 7]. Moreover, the combinatorial relationships between traces of matrix powers, characteristic coefficients, expected permanents, and expected determinants will help us better understand how to use these quantities, create bounds for them, and illuminate what has made them so especially useful in applied numerical science.

References

[1] Alexander I. Barvinok, Two algorithmic results for the traveling salesman problem, *Mathematics of Operations Research* **21** (1996) 65–84.

[2] Edward A. Bender and E. Rodney Canfield, Log-concavity and related properties of the cycle index polynomials, *J. Combinatorial Theory* **74** (1996) 57–70.

[3] Amir Ben-Dor and Shai Halevi, Zero-One Permanent is #P-Complete, A Simpler Proof, in *Proceedings of the 2nd Israel Symposium on Theory and Computing Systems* (1993) 108–117, `doi:10.1109/ISTCS.1993.253457`.

[4] M. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. Ralph, A. White, *Photonic Boson Sampling in a Tunable Circuit*, Science, 2012.

[5] Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi, *Algebraic Complexity Theory*, Springer-Verlag, 1997.

[6] Gordon G. Cash, A fast computer algorithm for finding the permanent of adjacency matrices, *J. Mathematical Chemistry* **18** 115–119.

[7] Gordon G. Cash, The permanental polynomial, *J. Chem. Inf. Comput. Sci.* **40** (2000) 1203–1206.

[8] D. Coppersmith and S. Winograd, Matrix multiplication via arithmetic progressions, in *STOC '87: Proceedings of 19th Annual ACM Symposium on Theory of computing*, (1987) 1–6.

[9] M. Fannes and P. Spincemaille, The mutual affinity of random measures, *Periodica Mathematica Hungarica* **47** (2003) 51–71.

[10] Yan V. Fyodorov, On permanental polynomials of certain random matrices, *Int. Math. Res. Not. 2006* Art. ID 61570, 37 pages.

[11] Gene H. Golub and Henk A. van der Vorst, Eigenvalue computation in the 20th century, *J. Comput. Appl. Math.* **123** (2000) 35–65.

[12] I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, Wiley, 1983.

[13] Mark Jerrum and Alistair Sinclair, Approximating the permanent, *SIAM J. Comput.* **18** (1989) 1149–1178.

[14] Mark Jerrum, Alistair Sinclair, and Eric Vigoda, A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries, *J. ACM* **51** (2004) 671–697.

[15] R. Merris, K.R. Rebman, and W. Watkins, Permanental polynomials of graphs, *Linear Alg. Appl.* **38** (1981) 273–278.

[16] H. Minc, *Permanents*, Reading, 1978.

[17] Terry Rudolph, Simple encoding of a quantum circuit amplitude as a matrix permanent, *Physical Review A* **80** 054302 (2009).

[18] H. J. Ryser, *Combinatorical Mathematics*, Carus Monograph, Mathematical Association of America, 1963.

[19] Audrey Terras *Zeta Functions of Graphs: A Stroll through the Garden*, Cambridge University Press, 2010.

[20] Christopher Umans, Group-theoretic algorithms for matrix multiplication, in *ISSAC '06: Proceedings of the 2006 International Symposium on Symbolic and Algebraic Computation* (2006) 5–5.

[21] L. G. Valiant, The complexity of computing the permanent, *Theoretical Computer Science* **8** (1979) 189–211.

[22] Ihara zeta function, http://en.wikipedia.org/wiki/Ihara_zeta_function

[23] Virginia Vassilevska Williams, Multiplying matrices faster than Coppersmith-Winograd, unpublished paper available at http://www.cs.stanford.edu/~virgi/matrixmult-f.pdf

[24] J. Wosiek, A simple formula for Bose–Einstein corrections, *Physics Letters B* **399** (1997) 1–2.