DISTRIBUTION OF EIGENVALUES OF HIGHLY PALINDROMIC TOEPLITZ MATRICES

STEVEN JACKSON, STEVEN J. MILLER, AND THUY PHAM

ABSTRACT. Consider the ensemble of real symmetric Toeplitz matrices whose entries are i.i.d random variables chosen from a fixed probability distribution p of mean 0, variance 1 and finite higher moments. Previous works showed that the limiting spectral measures (the density of normalized eigenvalues) converge weakly and almost surely to a universal distribution almost that of the Gaussian, independent of p. The deficit from the Gaussian distribution is due to obstructions to solutions of Diophantine equations and can be removed by making the first row palindromic. In this paper, we study the case where there is more than one palindrome in the first row of a real symmetric Toeplitz matrix. Using the method of moments and an analysis of the resulting Diophantine equations, we show that the moments of this ensemble converge to an universal distribution with very fat tails.

CONTENTS

Date: January 13, 2010.

1991 *Mathematics Subject Classification.* 15B52, 60F05, 11D45 (primary), 60F15, 60G57, 62E20 (secondary).

Key words and phrases. limiting spectral measure, Toeplitz matrices, random matrix theory, Diophantine equations, convergence, method of moments.

This work was done at the 2009 SMALL Undergraduate Research Project at Williams College, funded by NSF GRANT DMS-0850577 and Williams College; it is a pleasure to thank them and the other participants. The second named author was also partly supported by NSF grant DMS0600848.

1. Introduction

1.1. **Background.** Since its inception, Random Matrix Theory has been a powerful tool in modeling highly complicated systems, with applications in statistics [Wis], nuclear physics [Wig1, Wig2, Wig3, Wig4, Wig5] and number theory [KS1, KS2, KeSn]; see [FM] for a history of the development of some of these connections. An interesting problem in Random Matrix Theory is to study sub-ensembles of real symmetric matrices by introducing additional structure. One of those sub-ensembles is the family of real symmetric Toeplitz matrices; these matrices are constant along the diagonals:

$$A_N = \begin{pmatrix} b_0 & b_1 & b_2 & \cdots & b_{N-1} \\ b_{-1} & b_0 & b_1 & \cdots & b_{N-2} \\ b_{-2} & b_{-1} & b_0 & \cdots & b_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{1-N} & b_{2-N} & b_{3-N} & \cdots & b_0 \end{pmatrix}, \quad a_{ij} = b_{j-i}. \tag{1.1}$$

Initially numerical investigations suggested that the density of the normalized eigenvalues was given by the standard normal; however, Bose, Chatterjee, Gangopadhyay [BCG], Bryc, Dembo and Jiang [BDJ] and Hammond and Miller [HM] showed that this is not the case (in particular, the fourth moment is 2 2/3 and not 3). The analysis in [HM] shows that although the moments grow more slowly than the Gaussian's, they grow sufficiently fast to determine a universal distribution with unbounded support. The deficit from the standard Gaussian's moments is due to obstructions to Diophantine equations.

In [MMS], Massey, Miller and Sinsheimer found that, by imposing additional structure on the Toeplitz matrices by making the first row a palindrome, the Diophantine obstructions vanish and the limiting spectral measure converges weakly and almost surely to the standard Gaussian. A fascinating question to ask here is how the behavior of the normalized eigenvalues changes if we impose other constraints. Basak and Bose [BB] obtain results for ensembles of Toeplitz (and other) matrices that are also band matrices, with the results depending on the relative size of the band length to the dimension of the matrices. In this paper we explore the effect of increasing the palindromicity on the distribution of the eigenvalues. Before stating our results, we first list our notation.

1.2. **Notation.**

Definition 1.1. *For fixed n, we consider $N \times N$ real symmetric Toeplitz matrices in which the first row has 2^n copies of a palindrome. We always assume N to be a multiple of 2^n so that each element occurs exactly 2^{n+1} times in the first row. For instance, a doubly palindromic Toeplitz matrix (henceforth referred to as a DPT* *matrix) is of the form:*

$$A_N = \begin{pmatrix} b_0 & b_1 & \cdots & b_1 & b_0 & b_0 & b_1 & \cdots & b_1 & b_0 \\ b_1 & b_0 & \cdots & b_2 & b_1 & b_0 & b_0 & \cdots & b_2 & b_1 \\ b_2 & b_1 & \cdots & b_3 & b_2 & b_1 & b_0 & \cdots & b_3 & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b_2 & b_3 & \cdots & b_0 & b_1 & b_2 & b_3 & \cdots & b_1 & b_2 \\ b_1 & b_2 & \cdots & b_0 & b_0 & b_1 & b_2 & \cdots & b_0 & b_1 \\ b_0 & b_1 & \cdots & b_1 & b_0 & b_0 & b_1 & \cdots & b_1 & b_0 \end{pmatrix} \quad (1.2)$$

We always assume the entries of our matrices are i.i.d.r.v. chosen from some distribution p with mean 0, variance 1 and finite higher moments. The entries of the matrices are constant along diagonals. Furthermore, entries on two diagonals that are $N/2^n$ diagonals apart from each other are also equal. Finally, entries on two diagonals symmetric within a palindrome are also equal.

To succinctly keep track of which elements are equal, we may introduce a *link function* $\psi : \{1, \dots, N\}^2 \to \{1, \dots, N\}$ and new parameters b_ℓ such that $a_{ij} = b_{\psi(i,j)}$, where

$$\psi(i,j) = \begin{cases} |i-j| \mod 2^n & \text{if} \quad |i-j| \mod 2^n < N/2^{n+1} \\ -|i-j| \mod 2^n & \text{if} \quad |i-j| \mod 2^n > N/2^{n+1}. \end{cases} \quad (1.3)$$

Each $N \times N$ matrix A_N in this ensemble can be identified with a vector in $\mathbb{R}^{N/2^n}$ by $A_N \leftrightarrow (b_0(A_N), b_1(A_N), \dots, b_{N/2^n}(A_N))$. We denote the set of $N \times N$ real symmetric Toeplitz matrices with 2^n palindromes by $\Omega_{N,n}$ and subsequently construct a probability space $(\Omega_{N,n}, \mathcal{F}_N, \mathbb{P}_N)$ by

$$\begin{aligned} &\mathbb{P}_N(\{A_n \in \Omega_{N,n} : b_i(A_N) \in [\alpha_i, \beta_i] \text{ for } i \in \{0, 1, \dots, N/2^n - 1\}) \\ &= \prod_{i=0}^{\frac{N}{2^n}-1} \int_{\alpha_i}^{\beta_i} p(x_i) dx_i, \end{aligned} \quad (1.4)$$

where each dx_i is Lebesgue measure. For each matrix $A_N \in \Omega_{N,n}$ we associate a probability measure by placing a point mass of size $1/N$ at each of its normalized eigenvalues $\lambda_i(A_N)$:

$$\mu_{A_N}(x)dx := \frac{1}{N} \sum_{i=1}^{N} \delta\left(x - \frac{\lambda_i(A_N)}{\sqrt{N}}\right) dx, \quad (1.5)$$

where $\delta(x)$ is the Dirac delta function.

1.3. **Results.** Our main result concerns the limiting behavior (as a function of the palindromicity n) of the μ_{A_N} for generic A_N as $N \to \infty$. We analyze these limits using the method of moments. Specifically, for each A_N we calculate the moments of μ_{A_N} by using the Eigenvalue Trace Lemma to relate the k^{th} moment to the trace of A_N^k. We show the average k^{th} moment tends to the k^{th} moment of a distribution with unbounded support. By analyzing the rate of convergence, we obtain results on convergence in probability and almost sure convergence.

Specifically, we prove the following.

GIVE PRECISE STATEMENTS

The rest of the paper is organized as follows. We first establish some basic results about our ensembles and the associated measures in §2. We then analyze the even moments in detail in §3. We give the proof on the vanishing Diophantine obstructions for highly palindromic Toeplitz matrices and show that all the configurations of highly palindromic Toeplitz matrices contribute equally at any general even moment. While it is difficult to isolate the exact value of these moments, we are able to analyze these moments well enough to prove our convergence claims and to have some understanding of the limiting spectral measure. The situation is different for both the fourth moment for any palindromicity and all even moments for the doubly palindromic Toeplitz matrices, and we determine the exact values in §4. We conclude in §5 by proving the convergence claims.

2. Diophantine Formulation

In this section we begin our analysis of the moments. We prove some combinatorial results which restrict the number of configurations which can contribute a main term; we then analyze the potential main terms in the following section.

Recall that for each matrix $A_N \in \Omega_{N,n}$ we associate a probability measure by placing a point mass of size $1/N$ at each of its normalized eigenvalues $\lambda_i(A_N)$:

$$\mu_{A_N}(x)dx \ := \ \frac{1}{N}\sum_{i=1}^{N}\delta\left(x-\frac{\lambda_i(A_N)}{\sqrt{N}}\right)dx, \tag{2.6}$$

where $\delta(x)$ is the Dirac delta function. Thus the k^{th} moment of $\mu_{A_N}(x)$ is

$$M_{k,N}(A_N) \ := \ \int_{-\infty}^{\infty} x^k \mu_{A_N}(x)dx \ = \ \frac{1}{N^{k/2+1}}\sum_{i=1}^{N}\lambda_i^k(A_N). \tag{2.7}$$

The expected value of the k^{th} moment of the $N\times N$ matrices in our ensemble, found by averaging over the ensemble with each A_N weighted by (1.4) and using the Eigenvalue Trace Lemma, is

$$\begin{aligned} M_{k,N} \ := \ \mathbb{E}[M_{k,N}(A_N)] \quad &= \quad \frac{1}{N^{k/2+1}}\sum_{1\le i_1,\dots,i_k\le N}\mathbb{E}[a_{i_1i_2}a_{i_2i_3}\cdots a_{i_ki_1}] \\ &= \quad \frac{1}{N^{k/2+1}}\sum_{1\le i_1,\dots,i_k\le N}\mathbb{E}[b_{\psi(i_1,i_2)}b_{\psi(i_2,i_3)}\cdots b_{\psi(i_k,i_1)}], \end{aligned} \tag{2.8}$$

where from (1.4) the expectation equals

$$\mathbb{E}[b_{\psi(i_1,i_2)}b_{\psi(i_2,i_3)}\cdots b_{\psi(i_k,i_1)}] \ := \ \int\cdots\int b_{\psi(i_1,i_2)}b_{\psi(i_2,i_3)}\cdots b_{\psi(i_k,i_1)}\prod_{i=0}^{\frac{N}{2^n}-1}p(b_i)db_i. \tag{2.9}$$

We let M_k be the limit of the average moments; thus

$$M_k \ := \ \lim_{N\to\infty} M_{k,N}; \tag{2.10}$$

we will prove later that these limits exist.

Our goal is to understand the M_k, i.e., the limiting behavior of the moments in these ensembles. We use Markov's method of moments, which we summarize below. This is a standard method for proving results in the subject; a nice explicit summary of this method begins Section 3 of [BB].

- We first show $M_k = \lim_{N\to\infty} M_{k,N} = \lim_{N\to\infty} \mathbb{E}[M_{k,N}(A_N)]$ exists for k a positive integer, with the M_k's satisfying Carleman's condition: $\sum_{k=1}^{\infty} M_{2k}^{-1/2k} = \infty$. As these are the moments of the empirical distribution measures, this implies that the M_k's are the moments of a distribution.
- Convergence in probability follows from analyzing the second moment, namely showing $\mathrm{Var}(M_{k,N}(A_N) - M_k)$ tends to zero as $N \to \infty$.
- Almost sure convergence follows from showing the fourth moment tends to zero and then applying the Borel-Cantelli lemma.

We do the convergence calculations in §5; in this and the next few sections we determine the limiting behavior of the ensemble averages.

The odd moments are readily determined, as counting the degrees of freedom show the average odd moments vanish in the limit as $N \to \infty$.

Lemma 2.1. *All the average odd moments vanish in the limit; i.e.,* $\lim_{N\to\infty} M_{2m+1,N} = 0$.

Proof. For odd k, we consider $\mathbb{E}[b_{\psi(i_1,i_2)} b_{\psi(i_2,i_3)} \cdots b_{\psi(i_k,i_1)}]$; we may write this as $\mathbb{E}[b_{\ell_1}^{r_1} \cdots b_{\ell_j}^{r_j}]$ with $r_1 + \cdots + r_j = k$ and the b_ℓ's distinct. As k is odd, at least one b_ℓ is raised to an odd power. If any of these occur to just the first power, then the expectation is zero as the b's are drawn from a mean zero distribution.[1] Thus at least one of the b_ℓ's above occurs at least three times, and every b_ℓ occurs at least twice. The maximum number of distinct b_ℓ's occurs when everything is matched in pairs except for one triple matching. Thus there are at most $\frac{k-1}{2}$ different b_ℓ's in our tuple, and the number of tuples is bounded independent of N. We have two degrees of freedom from the first matching of the b_ℓ's and one degree of freedom for each other matching,[2] for a total of at most $\frac{k-1}{2} + 1 = \frac{k+1}{2}$ degrees of freedom. Thus the number of indices $i_1, \dots, i_k \in \{1, \dots, N\}$ that can contribute to the moment in (2.8) for a given matrix is $O_n(N^{(k+1)/2})$ (where the big-Oh constant may depend on n, as the larger n is the more choices we have for diagonals). As we divide by $N^{(k/2)+1}$ in (2.8), the odd moments are $O_n(N^{-1/2})$, and thus vanish in the limit as $N \to \infty$. □

[1]If we assume our distribution p is even, then a similar argument immediately implies all the odd moments vanish.

[2]For example, say $b_{\psi(i_1,i_2)} = b_{\psi(i_v,i_{v+1})}$, with i_1 our first index. Both i_1 and i_2 are free variables and we have N choices for each; however, i_v is not (it will have occurred in a matching before this point), and i_{v+1} is determined by requiring the two b_ℓ's under consideration to be equal. The number of choices for i_{v+1} depends on n (the larger n is, the more diagonals work); what matters is that the number of choices for i_{v+1} is independent of N. Whenever we have a new pair, we have a new choice for the value of the link function, and thus gain a degree of freedom.

In order to understand the even moments, we need to know more about which matchings are permissible, and how many choices of the indices lead to valid configurations. In the original case of the ensemble of real symmetric Toeplitz matrices [HM], the only way any two entries b_ℓ could match was for them to lie on the same diagonal or on the reflection of that diagonal over the main diagonal. That is, they matched if and only if

$$|i_m - i_{m+1}| \ = \ |i_l - i_{l+1}|. \tag{2.11}$$

For highly palindromic Toeplitz matrices, more relations give matchings (as seen in the investigation of palindromic matrices in [MMS]). An entry for which the absolute value of the difference between its indices is in a given congruence class modulo 2^n can match with another entry if and only if it is in the same congruence class or its negative. That is, two entries $a_{i_m i_{m+1}}$ and $a_{i_l i_{l+1}}$ can be matched in a pair if and only if their indices satisfy one of the following relations:

(1) there is a $C_1 \in \{(-\lfloor\frac{|i_l - i_{l+1}|}{2^n}\rfloor + k - 1)\frac{N}{2^n} \mid k \in \{1, \dots, 2^n\}\}$ such that
$$|i_m - i_{m+1}| \ = \ |i_l - i_{l+1}| + C_1; \tag{2.12}$$

(2) there is a $C_2 \in \{(\lfloor\frac{|i_l - i_{l+1}|}{2^n}\rfloor + k)\frac{N}{2^n} \mid k \in \{1, \dots, 2^n\}\}$ such that
$$|i_m - i_{m+1}| \ = \ -|i_l - i_{l+1}| + C_2; \tag{2.13}$$

as is standard, $\lfloor x \rfloor$ represents the largest integer at most x.

As a consequence of (2.12) and (2.13), for the matchings above there is some C such that

$$i_m - i_{m+1} \ = \ \pm(i_l - i_{l+1}) + C. \tag{2.14}$$

As there are two choices for sign and k matchings, there are potentially 2^k cases that can contribute. We now prune down the number of possibilities greatly by showing only one case contributes in the limit, namely the case when all the signs are negative.

In the Toeplitz ensembles studied in [HM] and [MMS], it was shown that any matching with a positive sign (i.e., as in (2.14)) in any pair contributes a lower order term to the moments, and thus it sufficed to consider the case where only negative signs occurred. A similar result holds here, which greatly prunes the number of cases we need to investigate. Note by Lemma 2.1 we need only investigate the even moments.

Lemma 2.2. *Consider the contribution to the $2k^{th}$ moment from all tuples $(i_1, \dots, i_{2k})$ in which the corresponding b_ℓ's are matched in pairs. If an $a_{i_m i_{m+1}}$ is matched with an $a_{i_l i_{l+1}}$ with a positive sign (which means*

$$i_m - i_{m+1} \ = \ +(i_l - i_{l+1}) + C$$

for some C as defined in (2.12) or (2.13)), then it contributes $O_k(1/N)$ to $M_{2k}(N)$ and therefore the contribution of all but one of the 2^k choices for the k signs vanishes in the limit, with only the choice of all negative signs being able to contribute in the limit.

Proof. The argument is essentially the same as in [MMS]. For any tuple $(i_1, \dots, i_{2k})$ in which the corresponding b_ℓ's are matched in pairs, there exist k equations, one for each pairing, of the form

$$i_m - i_{m+1} \ = \ \epsilon_m(i_l - i_{l+1}) + C_l \text{ where } \epsilon_l \ = \ 1 \text{ or } -1. \tag{2.15}$$

Let $x_1, x_2, \dots, x_{2k}$ denote the absolute value of the difference between two indices of each entry (so for $a_{i_l,i_{l+1}}$ it would be $x_j = |i_1 - i_{l+1}|$), and let $\tilde{x}_1 = i_1 - i_2, \tilde{x}_2 = i_2 - i_3, \dots$ and $\tilde{x}_{2k} = i_{2k} - i_1$ (i.e., the unsigned differences). It follows immediately that

$$\sum_{i=1}^{2k} \tilde{x}_i \ = \ 0. \tag{2.16}$$

Each $\tilde{x}_m$ can be expressed in two ways. By breaking the absolute value sign in (2.12) or (2.13), we have $\tilde{x}_m = \eta_j x_j$ for some j with $\eta_j = 1$ or -1. We can also express it through an equation like the one in (2.16) such that $\tilde{x}_m = \epsilon_m \tilde{x}_l + C_m$ for some l. Thus

$$\tilde{x}_m \ = \ \eta_j x_j = \epsilon_m \tilde{x}_l + C_l. \tag{2.17}$$

Then since $\epsilon_m^2 = 1$,

$$\tilde{x}_l \ = \ \epsilon_m \eta_j x_j - \epsilon_m C_l. \tag{2.18}$$

Note each absolute value of a difference occurs twice, as everything is matched in pairs. We therefore have

$$\sum_{i=1}^{2k} \tilde{x}_i \ = \ \sum_{j=1}^{k} [\eta_j x_j + (\epsilon_m \eta_j x_j - \epsilon_m C_m)] \ = \ \sum_{j=1}^{k} (n_j x_j (1 + \epsilon_m) - \epsilon_m C_j) \ = \ 0. \tag{2.19}$$

If any $\epsilon_m = 1$, then the x_j's are not linearly independent and we would have less than $k+1$ degree of freedom.[3] The contribution from such tuples to the moment in (2.8) for a given matrix is therefore $O(1/N)$ (as we divide by $N^{k/2+1}$), which vanishes in the limit as $N \to \infty$ and can thus be safely ignored. □

If the indices of $a_{i_l i_{l+1}}$ and $a_{i_m i_{m+1}}$ satisfy (2.12) for some C_1, then it follows immediately from Lemma 2.2 that $|i_m - i_{m+1}| = |i_l - i_{l+1}| + C_1$ implies

$$\begin{cases} i_m - i_{m+1} \ = \ -(i_l - i_{l+1}) + C_1 \\ i_m \ > \ \max\{i_{m+1}, i_{m+1} + C_1\} \end{cases} \quad \text{or} \quad \begin{cases} i_m - i_{m+1} \ = \ -(i_l - i_{l+1}) - C_1 \\ i_m \ < \ \min\{i_{m+1}, i_{m+1} - C_1\}. \end{cases} \tag{2.20}$$

Similarly, if they satisfy (2.12) for some C_2, then it follows from Lemma 2.2 that $|i_m - i_{m+1}| = -|i_l - i_{l+1}| + C_2$ implies

$$\begin{cases} i_m - i_{m+1} \ = \ -(i_l - i_{l+1}) + C_2 \\ i_{m+1} \ < \ i_m \ < \ i_{m+1} + C_2, \end{cases} \quad \text{or} \quad \begin{cases} i_m - i_{m+1} \ = \ -(i_l - i_{l+1}) - C_2 \\ i_{m+1} - C_2 \ < \ i_m \ < \ i_{m+1}. \end{cases} \tag{2.21}$$

Instead of considering each value of C (either C_1 or C_2) individually, we will consider a pair of constants C_1, C_2 such that $C_1 + C_2 = N - 1$. We claim that this

[3]As in the proof of Lemma 2.2, the first pair gives us two degrees of freedom and each subsequent pair gives at most one degree of freedom. If the x_j's are not linearly independent, there can be at most $k/2 - 1$ independent x_j's, and thus at most $k/2$ degrees of freedom.

removes some of the Diophantine obstructions that arise when evaluating (2.20) or (2.21) individually. Given an entry $a_{i_l i_{l+1}}$, we can associate each value of C with one diagonal whose entries, generally denoted by $a_{i_m i_{m+1}}$, all equal $a_{i_l i_{l+1}}$. Except for the main diagonal, every other diagonal has fewer than N entries and therefore the index $i_m \in \{a, \dots, b\}$ where $1 \le a < b \le N$ rather than $i_m \in \{1, \dots, N\}$. Here we only need to restrict one of the two indices of $a_{i_m i_{m+1}}$ and the other one will automatically be determined. However, by considering $a_{i_m i_{m+1}}$ on a pair of diagonals associated with C_1, C_2, we can take the index i_m (or i_{m+1}) to be any value between 1 and N. Furthermore, except for $O(1)$ values, the first index of entries from the pair of diagonals associated with C_1, C_2 are distinct, and similarly for the second index. Therefore, if $a_{i_m i_{m+1}}$ is on the diagonal associated with C_1 and $a_{i'_m i'_{m+1}}$ is on the diagonal associated with C_2, then for some $a, b \in \{1, \dots, N\}$, we have:

$$\begin{cases} i_m \in \{a, \dots, b\} \\ i'_m \in \{0, \dots, a\} \cup \{b, \dots, N\} \end{cases} \tag{2.22}$$

3. Spectral Characteristic of Highly Palindromic Toeplitz Matrices

3.1. **The Moment Problems.** To study the distribution of eigenvalues of highly palindromic Toeplitz matrices, we rely on Markov's method of moments.

Lemma 3.1. *Assume that p has mean 0, variance 1 and finite higher moments. Then $M_0 = 1$, $M_1 = 0$ and $M_2 = 1$ and all odd moments vanish.*

Proof. For all N, $M_0(A_N) = M_0(N) = 1$. For the first moment, we have:

$$M_1(N) = \frac{1}{N^{3/2}} \sum_{1 \le i_1 \le N} \mathbb{E}(a_{i_1 i_1}) = 0 \tag{3.23}$$

And for the second moment, we have:

$$\begin{aligned} M_2(N) &= \frac{1}{N^2} \sum_{1 \le i_1, i_2 \le N} \mathbb{E}(a_{i_1 i_2} \cdot a_{i_2 i_1}) \\ &= \frac{1}{N^2} \sum_{1 \le i_1, i_2 \le N} \mathbb{E}(a_{i_1 i_2}^2) = \frac{1}{N^2} \sum_{1 \le i_1, i_2 \le N} \mathbb{E}(b_{\psi(i_1, i_2)}^2). \end{aligned} \tag{3.24}$$

Since we choose b's from a normalized distribution with variance 1 distribution, the expected value above is 1. Thus, $M_2 = \lim_{N\to\infty} M_2(N) = \lim_{N\to\infty} \frac{N^2}{N^2} = 1$. We have shown that all the odd moments vanish in the limit in Lemma (2.1). □

Thus, we only need to consider the general even moment. For the $2k^{\text{th}}$ moment, we have a tuples with $2k$ indices and therefore $(2k-1)!!$ ways to match them in pairs. Each way of matching, hereafter referred to as configuration, leads to a system of k equations of the form (2.20) or (2.21), for which each distinct solution gives us one possible choice for the tuples $(i_1, \dots, i_{2k})$ and contribute one to the sum. In the case of real symmetric Toeplitz matrices, some configurations lead to a system of equations with more solutions than others, whereas in the case of real symmetric palindromic Toeplitz matrices, all configurations lead to a system of equations with the same number of solutions, agreeing up to $O(\frac{1}{N})$. Fortunately, this behavior is preserved in the case of highly palindromic Toeplitz matrices and greatly simplifies the calculation of the moments.

3.2. **The Matching Lemmas.**

Lemma 3.2 (Equal Contribution - Fourth Moment)**.** *The non-adjacent configuration and the adjacent configuration contribute equally to the fourth moment.*

FIGURE 1. The adjacent and the non-adjacent configurations of the fourth moment.

Proof. The general configuration of the fourth moment of highly palindromic Toeplitz matrices satisfies the following equations:

$$\begin{cases} a_{i_1 i_2} = a_{j_1 j_2} \\ a_{l_1 l_2} = a_{m_1 m_2}. \end{cases} \tag{3.25}$$

While we denote the indices by 8 different variables for full generality, there are actually only four different indices here (two adjacent entries in the tuples always share one index). Depending on which two variables refer to which index, we obtain different configurations for the fourth moment. Our goal is to show that the contribution from this general configuration is independent of the indices and therefore all configurations contribute equally. From the above system of equations relating the matchings, we obtain the corresponding system of equations for the indices:

$$\begin{cases} |i_1 - i_2| = \pm |j_1 - j_2| + A \\ |l_1 - l_2| = \pm |m_1 - m_2| + B \end{cases} \tag{3.26}$$

According the Lemma 2.2, in order for a tuples to contribute to the fourth moment in the limit, we must have

$$\begin{cases} i_1 - i_2 = -(j_1 - j_2) + A' \\ l_1 - l_2 = -(m_1 - m_2) + B'. \end{cases} \tag{3.27}$$

Since all equations occur with a negative sign before the indices on the left-hand side, we have $A' + B' = 0$ **I KNOW THIS IS TRUE, BUT IT'S HARD FOR ME TO SEE IN THIS FULLY GENERAL CASE WHY IT MUST BE TRUE, AND I CAN'T QUITE SEE WHAT THE PREVIOUS COMMENT HAS TO DO WITH IT AT THE MOMENT.**. If A is of the form C_2 in (2.13), then it follows immediately that $A = B$ since C_2 is always positive. But if A is of the form C_1 in (2.12), then it can be either that $A = B$ or $A = -B$. To find the contribution of this configuration to the fourth moment, we need to find the number of solutions of (3.22) **WHY ARE WE REFERENCING A LATER EQUATION??** for each A then sum over all possible values of A. For each A, we have a system of two equations with four variables so we can always at least pick two free indices among them. For convenience, we specify i_1, i_2 by choosing the first entry $a_{i_1 i_2}$. Moreover, we assume that we only pick $a_{i_1 i_2}$ in the lower diagonal half of the matrix so that $i_1 > i_2$. By the symmetry of the matrix, picking $a_{i_1 i_2}$ in the upper diagonal half would follow the same procedure. Finally, without loss of generality, we assume that A is of the form C_2 in (2.13). Thus

$$\begin{cases} |j_1 - j_2| = -|i_1 - i_2| + A \\ |l_1 - l_2| = -|m_1 - m_2| + A \end{cases} \Rightarrow \begin{cases} j_1 - j_2 = -(i_1 - i_2) + A \\ j_1 > j_2 \\ l_1 - l_2 = -(m_1 - m_2) - A \\ m_1 < m_2 < m_1 + A. \end{cases} \tag{3.28}$$

We now consider $A^{'}$ of the form C_1 in (2.12) such that $A + A^{'} = N - 1$. This contribution nicely complements the previous one:

$$\begin{cases} |j_1 - j_2| = |i_1 - i_2| + A^{'} \\ |l_1 - l_2| = |m_1 - m_2| \pm A^{'} \end{cases} \Rightarrow \begin{cases} j_1 - j_2 = -(i_1 - i_2) - A^{'} \\ j_1 < j_2 \\ l_1 - l_2 = -(m_1 - m_2) + A^{'} \\ m_1 < m_2 \textbf{ or } m_2 < m_1 - A^{'}. \end{cases} \tag{3.29}$$

Now consider the entry $a_{m_1 m_2}$. One of its indices, either m_1 or m_2, is determined by the choice of $a_{i_1 i_2}$. **IT'S NOT EXACTLY CLEAR TO ME WHAT IS MEANT BY THE FOLLOWING.** By swapping the position of $a_{m_1 m_2}$ and $a_{l_1 l_2}$, we see that we can safely assume that m_1 is the predetermined index. The other index m_2 can take on any value:

(1) On the diagonal associated with $A = C_1$:

$$m_2 \in \{m_1, \ldots, m_1 + A\} \cap \{a, \ldots, b\} \tag{3.30}$$

(2) On the diagonal associated with $A^{'} = C_2$:

$$m_2 \in (\{0, \ldots, m_1 - A^{'}\} \cup \{m_1, \ldots, N\}) \cap (\{0, \ldots, a\} \cup \{b, \ldots, N\}) \tag{3.31}$$

IN THESE EQUATIONS WE HAVE INCLUSIVE, YET THE ABOVE RELATIONS ARE INEQUALITIES. EITHER FIX THIS OR PUT A DISCLAIMER SAYING WE'LL BE LOOSE WITH INEQUALITIES SINCE SINGLE VALUES DON'T MATTER IN THE LIMIT. THIS MAY ALSO AFFECT SOME OF THE "EXACTLY" COMMENTS BELOW.

Therefore, there are exactly A out of $N+1$ values of m_2 we can pick (or exactly $A^{'}$ out of $N+1$ value of m_2 we cannot pick). Furthermore, since i_1, i_2 and m_2 already determined the final index is determined. Finally, since we have N^2 choices for picking the initial entry $a_{i_1 i_2}$, the contribution to the fourth moment from A and $A^{'}$ is given by:

$$N^2 \cdot A = \left(\frac{N^3}{2^n}\right) \left(-\left\lfloor \frac{|i_1 - i_2|}{2^n} \right\rfloor + k - 1\right) \tag{3.32}$$

The above formula only depends on the initial choice of $a_{i_1 i_2}$ and the choice of A, and not on how we organize the four entries to form each configuration. Summing over all possible choices of C_1, we obtain the same contribution to the fourth moment from either configuration. □

Before extending the above lemma to the general even moment, we discuss briefly some notation of a "lift map", a way of relating one configuration of an even moment to one configuration of the next higher even moment. Adding a pair of entries to a configuration will move us from that configuration to some configuration of the next even moment. There are only two ways to add these entries: adding a pair of adjacent entries or adding a pair of non-adjacent entries.

Lemma 3.3 (Configuration Lifting - Adjacent Case). *Conider a configuration of the $2k^{th}$ moment. All configurations of the $(2k+2)^{th}$ moment obtained by adding a pair of adjacent entries to this configuration contribute equally to the $(2k+2)^{th}$ moment.*

FIGURE 2. Moment Lifting by adding a pair of adjacent entries

Proof. Let

$$(\dots, a_{pq}, \dots, a_{ij}, a_{jk}, \dots a_{lm}, \dots)$$

be a tuple of configuration (1) of the moment M_{2k} and

$$(\dots, a_{pq}, \dots, a_{ij}, a_{jo}, a_{os}, a_{sk}, \dots, a_{lm}, \dots)$$

be the new tuple obtained by adding the pair of entries $a_{jo} = a_{os}$. Let Ω_{2k} be the set of all tuples that work for configuration (1) and Ω_{2k+2} be the set of all tuples that work for the configuration (2) **WHAT ARE CONFIGURATION (1) AND CONFIGURATION (2)? I DON'T RECALL SEEING THESE DEFINED IN THE FIGURES EITHER.**. We define a "lift map" $F : \Omega_{2k} \to \Omega_{2k+2}$ such that:

$$F[(\dots, a_{pq}, \dots, a_{ij}, a_{jk}, \dots, a_{lm}, \dots)] = (\dots, a_{pq}, \dots, a_{ij}, a_{jo}, a_{os}, a_{sk}, \dots, a_{lm}, \dots)$$

We want the map F to map each index in $(\dots, a_{pq}, \dots, a_{ij}, a_{jk}, \dots, a_{lm}, \dots)$ to itself and add the new index $s = j - B + B'$ where B is the value of C corresponding to the pair of entries containing a_{jk}, and B' is any value of C such that $s \in \{1, \dots, N\}$ and $(B - B')$ is a valid choice of C. The system of equations corresponding to the tuples $(\dots, a_{pq}, \dots, a_{ij}, a_{jk}, \dots, a_{lm}, \dots)$ is given as follows:

$$\begin{cases} l - m = -(i-j) + A \\ p - q = -(j-k) + B \\ \dots \end{cases} \tag{3.33}$$

Under the map F, we obtain a new tuple $(\dots, a_{pq}, \dots, a_{ij}, a_{jo}, a_{os}, a_{sk}, \dots, a_{lm}, \dots)$ satisfying the system of equations:

$$\begin{cases} l - m = -(i-j) + A \\ p - q = -(s-k) + B' \\ j - o = -(o-s) + (B - B') \\ \dots \end{cases} \tag{3.34}$$

Except for the entry a_{sk} immediately following the added pair of adjacent entries $(a_{jo},\ a_{os})$, every other entry is preserved under F. Therefore, the equations and C values that held for the original configuration still hold true. Since both B' and $(B - B')$ are valid choices for the C value, the two equations associated with a_{sk} and the added pair of adjacent entries $(a_{jo}a_{os})$ are also valid. Therefore

$$(\ldots, a_{pq}, \ldots, a_{ij}, a_{jo}, a_{os}, a_{sk}, \ldots, a_{lm}, \ldots) \in \Omega_{2k+2}. \tag{3.35}$$

THIS PARAGRAPH IS PRETTY AWKWARD. Thus, the map F lifts any tuple of the first configuration to a tuple of the configuration of the next even moment. It is worth noticing that the map F is not one-to-one. In fact, how many tuples of configuration (2) would result from the map F depending on the choice of j, B and B'. Nor did we consider the Diophantine obstructions or other conditions arising from solving the absolute value equations to obtain (3.35). One reason is that we are only concerned with the number of tuples that work for each configuration, rather than the specific value of each tuples. In doing so, we would need to apply the map F to every possible tuples of configuration (1), as well as considering every legitimate choice of B and B'. Fortunately, the map F only depends on the choice of one index j and one C value B from the configuration (1).

HAVING REMARKS WITHIN A PROOF STRIKES ME AS SOMEWHAT WEIRD. THIS DOESN'T NECESSARILY MEAN THEY SHOULD BE REMOVED OR MOVED ELSEWHERE.

Remark 3.4. *We can take B to be any possible value of our C values since in order to obtain all the tuples of configuration (1), we need to sum over all possible combination of C that work for the system of equations corresponding to configuration (1).*

Remark 3.5. *We can also take j to be any value in $\{1, \ldots, N\}$. For any general $2k^{th}$ even moment, we have $2k$ indices (unknown variables) and k equations with the last equation not linearly independent of the rest. Therefore, we must have at least two completely free indices that can take on any value between 1 and N. Customarily, we choose the two completely free indices by choosing the very first entry of the tuple, which obviously can be any entry on the matrix. Furthermore, we can also choose the first entry to be any vertice of the configuration. So if we choose a_{ij} to be our first entry, we can pick j to be any value between 1 and N.*

Hence, starting from a configuration at the $2k^{th}$ moment and adding in a pair of adjacent entries to move up to a configuration at the next even moment, we can always pick the entry preceding the location where we would add the new adjacent pair to be the first entry in the tuples. As the consequence, the number of tuples resulted from the map F would be the same regardless of where we add the adjacent pair. □

Lemma 3.6 (Configuration Lifting - Nonadjacent Case)**.** *Consider a configuration of matchings for the $2k^{th}$ moment. All configurations at the $2k^{th}$ moment obtained by adding a pair of non-adjacent entries contribute equally to the $(2k + 2)^{th}$ moment.*

FIGURE 3. Moment Lifting by adding a pair of non-adjacent entries

Proof. Let $(\dots, a_{ij}, a_{jm}, \dots, a_{lp}, a_{pq}, \dots)$ be a tuple of the configuration (1) for the $2k^{\text{th}}$ moment and $(\dots, a_{ij}, a_{jo}, a_{om}, \dots, a_{lp}, a_{ps}, a_{sq}, \dots)$ be the new tuple obtained by adding the pair of entries $a_{jo} = a_{ps}$. As before, let $\Omega_{2k}^{'}$ be the set of all tuples that work for the configuration $(1)^{'}$ and $\Omega_{2k+2}^{'}$ be the set of all tuples that work for the configuration $(2)^{'}$ **WHERE DID THESE PRIMES COME FROM? IF THEY ARE HERE TO DISTINGUISH FROM THE PREVIOUS LEMMA, WHY WERE THEY NOT THERE FROM THE VERY BEGINNING?**. We define a "lift map" $F^{'} : \Omega_{2k}^{'} \to \Omega_{2k+2}^{'}$ such that:

$$F^{'}[(\dots, a_{ij}, a_{jm}, \dots, a_{lp}, a_{pq}, \dots)] = (\dots, a_{ij}, a_{jo}, a_{om}, \dots a_{lp}, a_{ps}, a_{sq}, \dots) \quad (3.36)$$

We want the map $F^{'}$ to map every index in $(\dots, a_{ij}, a_{jm}, \dots, a_{lp}, a_{pq}, \dots)$ to itself and add two new indices $o = j + B - B^{'}$ and $s = p + D - D^{'}$ where B and D are the C values associated with the pairs containing a_{jm} and a_{pq} respectively. Also, $B^{'}$ and $D^{'}$ are any value of C such that $o, s \in \{1, \dots, N\}$ and $(D^{'} + B^{'} - D - B)$ is a valie choice of C. For the tuple $(\dots, a_{ij}, a_{jm}, \dots, a_{lp}, a_{pq}, \dots)$ we have the following system of equations:

$$\begin{cases} i_0 - j_0 = -(i - j) + A \\ j_0 - m_0 = -(j - m) + B \\ l_0 - p_0 = -(l - p) + C \\ p_0 - q_0 = -(p - q) + D \\ \dots \end{cases} \quad (3.37)$$

WHY ARE THINGS SUBSCRIPTED WITH ZEROS NOW?

Under the map $F^{'}$, we obtain a new tuples $(\dots, a_{ij}, a_{jo}, a_{om}, \dots, a_{lp}, a_{ps}, a_{sq}, \dots)$ satisfying the system of equations:

$$\begin{cases} i_0 - j_0 = -(i - j) + A \\ j_0 - m_0 = -(o - m) + B^{'} \\ l_0 - p_0 = -(l - p) + C \\ p_0 - q_0 = -(s - q) + D^{'} \\ j - o = -(p - s) + (D^{'} + B^{'} - D - B) \\ \dots \end{cases} \quad (3.38)$$

Like before, all entries except for the two immediately following the added entries are preserved under the map $F^{'}$ so their associated equations and C value still hold. Furthermore, since $B^{'}$, $D^{'}$ and $(D^{'} + B^{'} - D - B)$ are all valid choices of C, the other three equations also hold true. Lastly, the existence of at least two completely free indices allow us to choose them to be the first index in a_{jm} and the second index in a_{lp}. Thus, following the same line of argument in Lemma 3.3, we can always choose the two free indices such that the number of tuples resulting from the map $F'(\Omega'_{2k})$ are the same regardless of where we add the non-adjacent pair. ☐

I'M NOT SURE THAT I'M 100% SATISFIED BY THE LIFTING THEOREMS, BUT THE ONLY PART THAT SEEMS WEEK IS THE APPARENT ASSUMPTION THAT THE VALUES OF B' AND D' SUCH THAT $D' + B' - D - B$ IS A VALID C VALUE ARE THE ONLY VALUES THAT CAN WORK FOR THE $(2k+2)^{\mathbf{th}}$ MOMENT CONFIGURATIONS. I'LL HAVE TO THINK ABOUT IT MORE.

Corollary 3.7. *Given any configuration, we can replace one of its adjacent pairs by another adjacent pair, and similarly for non-adjacent pairs, without changing its contribution to the corresponding moment.*

Theorem 3.8. *If all configurations at the $2k^{th}$ moment contribute equally, then all configurations at the $(2k+2)^{th}$ moment also contribute equally.*

Proof. Given any configuration at the $(2k+2)^{th}$ moment, Corollary 3.7 allows us to repeatedly replace adjacent pairs with other adjacent pairs, and similarly for non-adjacent pairs. By iterating this process, we can move any configuration down to the following two configurations:

Figure 4. Two possible final configurations

For any configuration at the $(2k+2)^{th}$ moment with $k+1$ pairs, we first move all adjacent pairs to the left-hand side until there are only non-adjacent pairs left at the right-hand side. We can replace those non-adjacent pairs to form the following structure, which is possible since it contains only non-adjacent pairs:

Figure 5. Completely non-adjacent configuration

Consider the structure on the right-hand side. Since they **WHAT ARE THEY?** are the same regardless of what initial configuration we start with, we expect the same number of choices for the entries $(a_{jk}, \dots, a_{hl})$. For each choice of entries $(a_{jk}, \dots, a_{hl})$ we need to find the number of choices for $(a_{ni}, a_{ij}, a_{lm}, a_{mn})$ that work, then sum over all possible choices for $(a_{jk}, \dots, a_{hl})$ to find the contribution to the moment.

Consider the structure on the left-hand side $(a_{ni}, a_{ij}, a_{lm}, a_{mn})$. They are only slightly different from the adjacent and non-adjacent matching of the fourth moment. The only different is the the index j of a_{ij} and the index l of a_{lm} are not required to be the same like before. Nonetheless, Lemma 3.2 still holds true for this case if we choose the two completely free indices in Lemma 3.2 to be j and l instead of choosing the first entries at random. So the two configurations in Figures ? **ADD REFERENCES** and ? contribute equally to the $(2k+2)^{th}$ moment and therefore all configurations at the $(2k+2)^{th}$ moment contribute equally. ☐

I'M NOT SURE THAT I FULLY FOLLOW THIS AS IT IS IT ISN'T ALWAYS CLEAR WHAT IS BEING REFERRED TO DUE TO LACK OF NOTATION IN THE PICTURES AND SOME LACK OF SPECIFIC LANGUAGE, BUT I THINK THE ARGUMENT IS OK.

It follows immediately by induction from Lemma 3.2 and Theorem 3.8 that every configuration at any even moment contributes equally. The fact greatly reduces the complexity of our moment problem as we only need to calculate the contribution of the completely adjacent matching, and immediately get the same contribution from the other $(2k-1)!!-1$ configurations.

4. Calculating the Moments

4.1. The Fourth Moment.

We consider the two cases of adjacent matching, where either

$$\begin{cases} a_{ij} = a_{jk} \\ a_{kl} = a_{li}, \end{cases} \qquad \text{or} \qquad \begin{cases} a_{ij} = a_{li} \\ a_{jk} = a_{kl}. \end{cases}$$

We note that these are equivalent by relabeling, so we will focus on the first case and multiply by two to account for the two cases.

Given a_{ij} we want to find a matching a_{jk}, but we also need a third degree of freedom proportional to N to have a nonzero contribution. Thus, we choose a_{jk} so that it matches and many choices of l satisfy $a_{kl} = a_{li}$. Exploiting the matrix symmetry, this reduces to choosing k so that $a_{ij} = a_{kj}$ and $a_{kl} = a_{il}$. That is, row i and row k should match up well.

Figure 6. An example highlighting matchings for l in green. Note that any anomalous matchings won't contribute in the limit.

Remark 4.1. *We now note the most useful features of our matrices. Foremost among these is the special feature of the main diagonal: It is the only place (excluding the border of the matrix) where b_0 occurs once rather than twice, leading to a few interesting properties. Firstly, moving to the corresponding point in the next palindrome can require either moving $\frac{N}{2^n}-1$ elements when crossing the main diagonal or $\frac{N}{2^n}$ elements otherwise. Secondly, as pictured in Figure 6, it means that a row and the row $\frac{bN}{2^n}$ rows down don't match perfectly, but rather become unaligned when one has reached the main diagonal but the other hasn't. Moreover, the row $\frac{bN}{2^n}-1$ rows down starts out unaligned, but then becomes aligned in this same region. Furthermore, only rows of this form match up well with the original row.*

Thus, due to the way the rows match, choosing k so that a_{ij} and a_{kj} are at corresponding points in a palindrome guarantees $a_{ij} = a_{kj}$ and that there are many choices of l satisfying $a_{kl} = a_{il}$, as desired.

More specifically, consider when a_{kj} is b palindromes away from a_{ij}, with b positive for simplicity. Then $k = i + \frac{bN}{2^n}$ if a_{ij} and a_{kj} are on the same side of the main diagonal, and $k = i + \frac{bN}{2^n} - 1$ if crossing the main diagonal. In order to have $k \in \{1, 2, \dots, N\}$, we must have

$$i \in \{1, 2, \dots, N - \frac{bN}{2^n} + O(1)\}$$

as our first restriction.

FIGURE 7. Regions where $k = i + \frac{N}{2}$ gives a matching are indicated in green, whereas those where $k = i + \frac{N}{2} - 1$ are indicated in red. Regions where both are satisfied are indicated in yellow: These are 1-dimensional, and thus won't contribute in the limit.

Another restriction comes about due to where the constant can be $\frac{bN}{2^n}$ and where it will instead be $\frac{bN}{2^n} - 1$. For $\frac{bN}{2^n}$, clearly any a_{ij} below the main diagonal won't cross the main diagonal when moving down to a_{kj}. Similarly, any a_{ij} that lies more than $\frac{bN}{2^n}$ elements above the main diagonal won't cross the main diagonal when moving to a_{kj}. Thus, we have two triangular regions of height $N - \frac{bN}{2^n} + O_b(1)$ defined by the main diagonal, which sum to give a square of area $\left(\frac{2^n - b}{2^n}\right)^2 N^2 + O_b(N)$. As explained above, the only values of l that won't work occur when the rows are unaligned, leaving $\frac{2^n - b}{2^n} N + O_b(1)$ good values of l. Thus, these areas contribute a total of

$$\left(\frac{2^n - b}{2^n}\right)^3 N^3 + O_b(N^2)$$

matchings to the fourth moment.

We now consider the situation where $\frac{bN}{2^n} - 1$ is the constant. The area of possible a_{ij} is the parallelogram in bordered by the triangles defined above, and thus it is of height $\frac{2^n - b}{2^n} N + O_b(1)$ and width $\frac{b}{2^n} N + O_b(1)$. In this case, the contributing values of ℓ will be those when one row has reached the diagonal but the other hasn't yet, and will thus be $\frac{b}{2^n} N + O_b(1)$. This constant will therefore contribute

$$\frac{2^n b^2 - b^3}{2^{3n}} N^3 + O_b(N^2)$$

matchings.

Remark 4.2. *We have not yet considered what happens for the negatives of these constants. However, repeating the same analysis gives identical regions and thus identical contributions to the fourth moment. Pictorially, what happens for a negative constant is that of the positive one rotated* $180°$. *Thus, the contribution to the fourth moment will be given by the contributions for positive constants multiplied first by a factor of* 2 *to account for the negatives, and a further factor of* 2 *to account for the two adjacent configurations.*

Theorem 4.3. *The adjacent contribution to the* 4^{th} *moment averaged over the ensemble of real symmetric Toeplitz matrices with* 2^n *palindromes is*

$$\frac{4}{3} 2^n + \frac{2}{3} 2^{-n}. \tag{4.39}$$

Proof. For each value of b, we note that the contribution to $M_4(N)$ is

$$\frac{1}{N^3}\left(\left(\frac{2^n-b}{2^n}\right)^3 N^3 + \frac{2^n b^2 - b^3}{2^{3n}} N^3 + O_b(N^2)\right) = \left(\frac{2^n-b}{2^n}\right)^3 + \frac{2^n b^2 - b^3}{2^{3n}} + O_b\left(\frac{1}{N}\right).$$

Thus, the contribution to M_4 is

$$\left(\frac{2^n-b}{2^n}\right)^3 + \frac{2^n b^2 - b^3}{2^{3n}}.$$

We sum over each value of b, multiply by 4 to account for the negative constants and the two adjacent configurations, and include the contribution from $C = 0$, known to be 2 [MMS] to obtain the adjacent contribution to the fourth moment:

$$M_4(\text{adj}) \; = = 2 + \frac{4}{2^{3n}} \sum_{b=1}^{2^n} \left((2^n - b)^3 + 2^n b^2 - b^3\right). \tag{4.40}$$

Extending the sum to include $b = 0$ cancels the first and last terms of the sum, but we must subtract 4 to compensate. This then leaves a sum of squares which is easily evaluated:

$$\begin{aligned}
M_4(\text{adj}) &= -2 + \frac{4}{2^{3n}} \sum_{b=0}^{2^n} 2^n b^2 \\
&= -2 + \frac{4}{2^{2n}} \frac{2^n(2^n+1)(2\cdot 2^n + 1)}{6} \\
&= -2 + 2\frac{(1+2^{-n})(2\cdot 2^n+1)}{3} \\
&= -2 + \frac{2}{3}(2\cdot 2^n + 2 + 1 + 2^{-n}) \\
&= \frac{4}{3}2^n + \frac{2}{3}2^{-n},
\end{aligned}$$

completing the proof. ☐

4.2. **The General Even Moments of DPT Matrix.** Theorem 3.8 **ADD REFERENCE** allows us to solve the higher moments and higher palindromicities provided that we can solve a single matching case in general. Thus, we generalize the above pictorial method for higher moments of the adjacent case. For the $2k^{\text{th}}$ moment, we find that our final system of equations becomes

$$\begin{gathered}
i_3 = i_1 + C_1 \\
i_5 = i_3 + C_2 = i_1 + C_1 + C_2 \\
\vdots \\
i_1 = i_{2k-1} + C_k = i_1 + \sum_{n=1}^{k} C_n.
\end{gathered}$$

Remark 4.4. *The even indices don't appear because the n^{th} matching gives the equation $i_{2n} - i_{2n-1} \;=\; -(i_{2n+1} - i_{2n}) + C_n$, and the i_{2n} terms cancel. However, for each non-zero constant C_l, we will have a picture similar to Figure 6, which limits the number of good values of the even indices i_{2l}. Moreover, as every i_{2n+1} is related back to i_1, the difference between the maximum and minimum partial sums must be stricly less than $N + O(1)$ or we lose a degree of freedom.*

These observations allow us to solve the doubly palindromic case.

Theorem 4.5. *The $2k^{th}$ moment averaged over the ensemble of doubly palindromic Toeplitz matrices is given by:*

$$M_{2k} \;=\; (2k-1)!!\left(-2 + 2^{-k}\left(\sum_{b=1}^{3} b^k\right)\right). \tag{4.41}$$

Proof. The following observations greatly simplify the analysis for this case:

- If the constants $\pm\frac{N}{2}$ appear in the C-vector, then $\pm\frac{N}{2} - 1$ can't occur as we would lose a degree of freedom in i_2, as $a_{i_1 i_2}$ would need to lie on a certain diagonal.
- If some C_j is non-zero, then the next non-zero C chosen must be $-C_j$, as we would otherwise lose a degree of freedom in i_1.

Now consider the $2k^{\text{th}}$ moment, which will have a C-vector of length k. We can then consider a subset of length m (m even) of $(\frac{N}{2}, -\frac{N}{2}, \frac{N}{2}, -\frac{N}{2}, \dots)$ that forms the core of the C-vector, with the remaining entries being zero. There are then $\binom{k}{m}$ ways to insert the zeros, and thus $\binom{k}{m}$ ways to build a C-vector around this core.

We now consider the contribution from each of these C-vectors. By Remark 4.4, we see that we will have $\frac{N}{2}$ values of i_1 to choose from, and there will be m other i_{2l} (corresponding to the m nonzero C_l) that will have $(N - \frac{N}{2}) + O(1)$ good values. Thus, the contribution for each of these cases will be $(\frac{1}{2})^{m+1}$. Therefore, the total contribution to the $2k^{\text{th}}$ moment from this configuration, summing over all possible C-vectors, will be

$$\sum_{\substack{m \text{ even} \\ m=2}}^{k} \binom{k}{m}\left(\frac{1}{2}\right)^{m+1}.$$

If we pull out a factor of $\frac{1}{2}$ and include $m = 0$ in the sum, we can use the binomial theorem to express this as

$$\frac{1}{4}\left(\left(1+\frac{1}{2}\right)^k + \left(1-\frac{1}{2}\right)^k\right) - \frac{1}{2}. \tag{4.42}$$

The contribution from a core of $(-\frac{N}{2}, \frac{N}{2}, -\frac{N}{2}, \frac{N}{2}, \dots)$ will be the same. The cores of $(\pm(\frac{N}{2} - 1), \mp(\frac{N}{2} - 1), \dots)$ can be similarly analyzed, and they will also have the same contribution since $N - \frac{N}{2} + O(1) = \frac{N}{2} + O(1)$, so we multiply (4.42) by 4. We must also include the contribution from the 0-vector, which is always 1 for the adjacent case. Thus, the contribution from each configuration is

$$-2+\left(1+\frac{1}{2}\right)^k+1^k+\left(1-\frac{1}{2}\right)^k \;=\; -2+2^{-k}\left(\sum_{b=1}^{3}b^k\right). \tag{4.43}$$

Appealing to Theorem 3.8 **ADD REFERENCE** and multiplying by the number of configurations, we have

$$M_{2k} \;=\; (2k-1)!!\left(-2+2^{-k}\left(\sum_{b=1}^{3}b^k\right)\right), \tag{4.44}$$

completing the proof. □

5. Convergence

In this section, we show that the limit of the average moments exists, and that the moments grow slowly enough to determine a unique probability distribution. With this result, we then show convergence in probability. Finally, assuming that $p(x)$ is even, we prove almost sure convergence. These arguments closely follow those in [HM] and [MMS], with modifications where necessary.

5.1. A Non-trivial Lower Bound for Higher Moments.

We now extend Theorem 4.5 to matrices with greater palindromicity. In doing so, we will miss many of the C-vectors that contribute to these moments, but exact calculations for even a quadruply palindromic matrix have proven difficult. We begin with the following lemma:

Lemma 5.1. *For a constant $C=\frac{bN}{2^n}$ for $b\in\{1,\dots,2^n-1\}$, the total contribution from the cores of $(\pm C,\mp C,\pm C,\mp C,\dots)$, and its complement $(\pm(N-1-C),\mp(N-1-C),\dots)$ is*

$$-2+\left(2-\frac{b}{2^n}\right)^k+\left(\frac{b}{2^n}\right)^k. \tag{5.45}$$

Proof. This goes back to the observation in Figure 6 that for $C=\frac{bN}{2^n}$, the number of free l values is $N-\frac{bN}{2^n}+O(1)$, whereas if $C=\frac{bN}{2^n}-1$, then there are $\frac{bN}{2^n}+O(1)$ good l values. Thus, the complementary C will give the same restrictions on the number of l values.

Moreover, the restrictions on i_1 from $\frac{bN}{2^n}$ and $N-1-\frac{bN}{2^n}$ sum to 1. Thus, as there are the two cases (plus first or minus first) for each, when we sum them and extend the sums back to 0, we have

$$-2+2\cdot\sum_{m\ \text{even}}^{k}\binom{k}{m}\left(\frac{2^n-b}{2^n}\right)^{m+1}=-2+\left(\frac{2\cdot 2^n-b}{2^n}\right)^k+\left(\frac{b}{2^n}\right)^k, \tag{5.46}$$

completing the proof. □

In order to get our lower bound for the higher moments, we then repeat this for every value of $b\in\{1,2,\dots,2^n-1\}$. Adding in the contribution from the zero vector, we obtain

Lemma 5.2. *The lower bound for the $2k^{th}$ moment averaged over the ensemble of Toeplitz matrices with 2^n palindromes is given by*

$$M_{2k}(2^n) \geq -2\cdot(2^n-1) + 2^{-kn}\left(\sum_{b=1}^{2\cdot 2^n-1} b^k\right), \tag{5.47}$$

which is easily summed for any value of k.

5.2. **Upper Bound for the Higher Moments.** We first show that the higher moments have a limit regardless of the palindromicity, even though we may not be able to calculate it in closed form, and that these limits grow slowly enough to ensure convergence to a unique probability distribution (see [Fe]).

For notational convenience, in this section we use $C \pm |i_n - i_{n+1}|$ when we really mean either only $C + |i_n - i_{n+1}|$ or $C - |i_n - i_{n+1}|$, where the choice of sign depends on the value of C. Of course, the total number of possible values of C depends on the palindromicity. If we have 2^n palindromes, then we will have $4\cdot 2^n - 1$ possible values of C_j. Thus, for the $2k^{th}$ moment, we will have at most $(4\cdot 2^n-1)^{k-1}$ possible C-vectors (as C_k will be chosen to return to 0, if possible).

By matching $2k$ entries from the matrix in pairs, we obtain k equations of the form

$$|i_n - i_{n+1}| = C_j \pm |i_m - i_{m+1}|. \tag{5.48}$$

From these equations, we then move to a system of equations of the form

$$i_n - i_{n+1} = C_l - (i_m - i_{m+1}), \tag{5.49}$$

where C_l can be either C_j or $-C_j$, and $\sum_{l=1}^{k} C_l = 0$. In moving from (5.48) to (5.49), we also obtain a system of inequalities that the i values must satisfy.

As was argued in Lemma 2.9 of [MMS] the k equations corresponding to (5.49) leave $k+1$ free indices. Of course, further degrees of freedom can be lost from the inequalities, but none can be gained. If we divide by N^{k+1} and take the limit $N\to\infty$, we can interpret these as uniform variables as in [BDJ] and [HM], and thus interpret the equations as determining a region of a $(k+1)$-dimensional unit cube. Thus, the contribution must converge to some $c \in [0,1]$.

Finally, we show that these moments grow sufficiently slowly to give a probability distribution. For each configuration, each C-vector can contribute at most 1 (if every free index ranges freely over $\{1,2,\dots,N\}$), giving an upper bound for the $2k^{th}$ moment of $(2k-1)!!\cdot(4\cdot 2^n-1)^{k-1}$.

Carleman's condition ([Fe]) states that if a distribution μ has finite moments and the moments satisfy

$$\sum_{k=1}^{\infty} M_{2k}^{-1/(2k)} = +\infty$$

then the moments uniquely determine μ. For each moment, we have

$$\begin{aligned} M_{2k}^{-1/(2k)} &\geq \left((2k-1)!!\cdot(4\cdot 2^n-1)^{k-1}\right)^{-1/(2k)} && (5.50)\\ &= \left((2k-1)!!\right)^{-1/(2k)}\cdot(4\cdot 2^n-1)^{\frac{1-k}{2k}}. && (5.51)\end{aligned}$$

In the limit $k \to \infty$, the latter factor approaches $(4 \cdot 2^n - 1)^{-1/2}$, so as the sum of first factor diverges to infinity[4], the product must also diverge to infinity, and Carleman's condition is satisfied.

5.3. **Convergence in Probability.** We begin by defining our random variables. Let A be a sequence of real numbers to which we associate A_N, an $N \times N$ real symmetric Toeplitz matrix with 2^n palindromes. Let $X_{m;N}(A)$ be a random variable that equals the m^{th} moment of A_N, and set $X_m(A)$ to the m^{th} moment averaged over the ensemble as above.

Thus, we have convergence in probability if for all $\epsilon > 0$

$$\lim_{N\to\infty} \mathbb{P}_{\mathbb{N}}(\{A \in \Omega_{\mathbb{N}} : |X_{m;N} - X_m| > \epsilon\}) \ = \ 0. \tag{5.52}$$

Chebychev's inequality states that

$$\begin{aligned} \mathbb{P}_{\mathbb{N}}(\{A \in \Omega_{\mathbb{N}} : |X_{m;N} - X_m| > \epsilon\}) \quad &\leq \quad \frac{\mathbb{E}[(X_{m;N} - X_m)^2]}{\epsilon^2}. \\ &= \quad \frac{\mathbb{E}[M_m(A_N)^2] - \mathbb{E}[M_m(A_N)]^2}{\epsilon^2}. \end{aligned} \tag{5.53}$$

Thus, it suffices to show

$$\lim_{N\to\infty} (\mathbb{E}[M_m(A_N)^2] - \mathbb{E}[M_m(A_N)]^2) \ = \ 0 \tag{5.54}$$

to prove convergence in probability.

We have

$$\begin{aligned} \mathbb{E}[M_m(A_N)^2] \quad &= \quad \frac{1}{N^{m+2}} \sum_{1\leq i_1,\dots,i_m\leq N} \\ &\qquad \times \sum_{1\leq j_1,\dots,j_m\leq N} \mathbb{E}[b_{|i_1-i_2|}\cdots b_{|i_m-i_1|} b_{|j_1-j_2|}\cdots b_{|j_m-j_1|}], \\ \mathbb{E}[M_m(A_N)]^2 \quad &= \quad \frac{1}{N^{m+2}} \sum_{1\leq i_1,\dots,i_m\leq N} \mathbb{E}[b_{|i_1-i_2|}\cdots b_{|i_m-i_1|}] \\ &\qquad \times \sum_{1\leq j_1,\dots,j_m\leq N} \mathbb{E}[b_{|j_1-j_2|}\cdots b_{|j_m-j_1|}]. \end{aligned} \tag{5.55}$$

We can break this up into two cases. If the entries of the i diagonals are entirely distinct from those of the j diagonals, then the contribution to $\mathbb{E}[M_m(A_N)^2]$ and to $\mathbb{E}[M_m(A_N)]^2$ will clearly be the same. Thus, we need to approximate the contribution from the cases where there are one or more shared diagonals. The degree of freedom arguments of [HM] immediately apply here, though our big-Oh constants will now depend on the value of 2^n as we now have many more C-vectors to which we apply these arguments. Thus, as $N \to \infty$ these two quantities will converge, and convergence in probability and thus weak convergence follow.

---

[4] $(2n-1)!! = \frac{(2n)!}{2^n n!}$, then apply Stirling's approximation: $((2k-1)!!)^{-1/(2k)} \approx \left(\frac{e}{2k}\right)^{1/2}$. Thus, the sum diverges.

5.4. **Almost Sure Convergence.** We assume that $p(x)$ is even for convenience. Almost sure convergence follows from showing that as as $N \to \infty$

$$\{A_N \in \Omega_{\mathbb{N}} \ : \ \lim_{N\to\infty} X_{m;N}(A_N) \to X_m(A_N)\}$$

is an event with probability one for all non-negative integers m.

Let $M_m(N) = \mathbb{E}(M_m(A_N)$. By the triangle inequality, we have that

$$|M_m(A_N) - M_m| \ \leq \ |M_m(A_N) - M_m(N)| - |M_m(N) - M_m|. \tag{5.56}$$

We have already shown that $\lim_{N\to\infty} |M_m(N) - M_m| = 0$, so we must show that $|M_m(A_N) - M_m(N)|$ almost surely tends to zero. Clearly, $\mathbb{E}[|M_m(A_N) - M_m(N)|] = 0$, and we can modify the arguments in [HM] to show that the fourth moment of $M_m(A_N) - M_m(N)$ is $O_{m,2^n}(\frac{1}{N^2})$. All of the degree of freedom arguments can be applied directly for each C-vector.

However, Theorems 6.15 and 6.16 of [HM] require greater care as these use more than degree of freedom arguments. Fortunately, equations (50) and (51) will hold for any of our C-vectors, so a similar result will hold in this case. We then apply Chebyshev's inequality to find

$$\mathbb{P}_{\mathbb{N}}(|X_{m;N}(A) - X_m(A)| \geq \epsilon) \ \leq \ \frac{\mathbb{E}[|M_m(A_N) - M_m(N)|^4]}{\epsilon^4} \ \leq \ \frac{C_{m,2^n}}{N^2\epsilon^4}. \tag{5.57}$$

Finally, applying the Borel-Cantelli Lemma shows that we have convergence everywhere except for a set of zero probability, thus proving almost sure convergence.

References

[BB] A. Basak and A. Bose, *Limiting spectral distribution of some band matrices*, preprint 2009.

[BCG] A. Bose, S. Chatterjee and S. Gangopadhyay, *Limiting spectral distributions of large dimensional random matrices*, J. Indian Statist. Assoc. (2003), **41**, 221–259.

[BDJ] W. Bryc, A. Dembo, T. Jiang, *Spectral Measure of Large Random Hankel, Markov, and Toeplitz Matrices*, Annals of Probability **34** (2006), no. 1, 1–38.

[Fe] W. Feller, *Introduction to Probability Theory and its Applications, Volume 2*, first edition, Wiley, New York, 1966.

[FM] F. W. K. Firk and S. J. Miller, *Nuclei, Primes and the Random Matrix Connection*, Symmetry **1** (2009), 64–105; doi:10.3390/sym1010064.

[GS] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, third edition, Oxford University Press, 2005.

[HM] C. Hammond and S. J. Miller, *Eigenvalue spacing distribution for the ensemble of real symmetric Toeplitz matrices*, Journal of Theoretical Probability **18** (2005), no. 3, 537–566.

[KS1] N. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues and Monodromy*, AMS Colloquium Publications, Vol. 45, AMS, Providence, RI, 1999.

[KS2] N. Katz and P. Sarnak, *Zeros of zeta functions and symmetries*, Bull. AMS **36** (1999), 1–26.

[KeSn] J. P. Keating and N. C. Snaith, *Random matrices and L-functions.* In *Random Matrix Theory*, J. Phys. A **36** (2003), no. 12, 2859–2881.

[MMS] A. Massey, S. J. Miller, J. Sinsheimer, *Distribution of eigenvalues of real symmetric palindromic Toeplitz matrices and circulant matrices*, Journal of Theoretical Probability **20** (2007), no. 3, 637-662.

[Wig1] E. Wigner, *On the statistical distribution of the widths and spacings of nuclear resonance levels*, Proc. Cambridge Philo. Soc. **47** (1951), 790–798.

[Wig2] E. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Ann. of Math. **2** (1955), no. 62, 548–564.

[Wig3] E. Wigner, *Statistical Properties of real symmetric matrices.* Pages 174–184 in *Canadian Mathematical Congress Proceedings*, University of Toronto Press, Toronto, 1957.

[Wig4] E. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions. II*, Ann. of Math. Ser. 2 **65** (1957), 203–207.

[Wig5] E. Wigner, *On the distribution of the roots of certain symmetric matrices*, Ann. of Math. Ser. 2 **67** (1958), 325–327.

[Wis] J. Wishart, *The generalized product moment distribution in samples from a normal multivariate population*, Biometrika **20 A** (1928), 32–52.

E-mail address: `Steven.R.Jackson@williams.edu`

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267

E-mail address: `Steven.J.Miller@williams.edu`

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267

E-mail address: `tvp1@williams.edu`

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267