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 12, 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], Kargin [Kar] and Liu and Wang [LW] 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 is 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:*

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

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

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

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

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

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 $\lambda_i(A_N)$:

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

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 weak convergence and almost sure convergence.

Specifically, our main result is the following.

Theorem 1.2. (Weak and Strong Convergence) *Let n be a fixed positive integer, and for each N a multiple of 2^n consider the ensemble of real symmetric $N \times N$ palindromic Toeplitz matrices whose first row is 2^n copies of a fixed palindrome (see Definition 1.1), where the independent entries are independent, identically distributed random variables arising from a probability distribution p with mean 0, variance 1 and finite higher moments. Then as $N \to \infty$ the measures μ_{A_N} (see Definition 1.5) converge weakly to a limiting spectral measure with unbounded support. If additionally p is even, then the measures converge strongly to the limiting spectral measure.*

As in other related ensembles, it is very difficult to obtain closed form expressions for the general moments of the limiting spectral measure **MAYBE ADD SOMETHING ABOUT THE DOUBLY PALINDROMIC CASE?**. We can, however, analyze the moments well enough to determine the limiting distribution has unbounded support; in fact, as the following theorem shows it has fatter tails than previously studied ensembles.

Theorem 1.3. (Fat Tails) *Consider the ensemble from Theorem 1.2. For any fixed $n \geq 1$, the moments grow faster than the corresponding moments of the standard normal; specifically, if $M_{2m,n}$ denotes the $2m^{th}$ moment of the limiting spectral measure of our ensemble for a given n, then*

$$M_{2m,n} \gg \frac{2^{mn}}{m} \cdot (2m-1)!!. \tag{1.6}$$

The limiting spectral measure thus has unbounded support, and fatter tails than the standard normal (or in fact any of the known limiting spectral measures arising from an ensemble where the independent entries are chosen from a density whose moment generating function converges in a neighborhood of the origin).

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 $\lambda_i(A_N)$:

$$\mu_{A_N}(x)dx \ := \ \frac{1}{N}\sum_{i=1}^{N}\delta\left(x - \frac{\lambda_i(A_N)}{\sqrt{N}}\right)dx, \tag{2.7}$$

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

$$M_{k,n;N}(A_N) \ := \ \int_{-\infty}^{\infty} x^k \mu_{A_N}(x)dx \ = \ \frac{1}{N^{k/2+1}}\sum_{i=1}^{N}\lambda_i^k(A_N). \tag{2.8}$$

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

$$\begin{aligned} M_{k,n;N} \ := \ \mathbb{E}[M_{k,n;N}(A_N)] \quad &= \quad \frac{1}{N^{k/2+1}}\sum_{1\le i_1,\dots,i_k\le N}\mathbb{E}[a_{i_1i_2}a_{i_2i_3}\cdots a_{i_ki_1}] \\ &= \quad \frac{1}{N^{k/2+1}}\sum_{1\le i_1,\dots,i_k\le N}\mathbb{E}[b_{\psi(i_1,i_2)}b_{\psi(i_2,i_3)}\cdots b_{\psi(i_k,i_1)}], \end{aligned} \tag{2.9}$$

where from (1.4) the expectation equals

$$\mathbb{E}[b_{\psi(i_1,i_2)}b_{\psi(i_2,i_3)}\cdots b_{\psi(i_k,i_1)}] \ := \ \int\cdots\int b_{\psi(i_1,i_2)}b_{\psi(i_2,i_3)}\cdots b_{\psi(i_k,i_1)}\prod_{i=0}^{\frac{N}{2^n}-1}p(b_i)db_i. \tag{2.10}$$

We let $M_{k,n}$ be the limit of the average moments; thus

$$M_{k,n} \ := \ \lim_{N\to\infty} M_{k,n;N}; \tag{2.11}$$

we will prove later that these limits exist.

Our goal is to understand the $M_{k,n}$, 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 $M_{m,n} = \lim_{N\to\infty} M_{m,n;N} = \lim_{N\to\infty}\mathbb{E}[M_{m,n;N}(A_N)]$ exists for m a positive integer, with the $M_{m,n}$'s satisfying Carleman's condition: $\sum_{m=1}^{\infty} M_{2m}^{-1/2m} = \infty$. As these are the moments of the empirical distribution measures, this implies that the $M_{m,n}$'s are the moments of a distribution.
- Weak convergence follows from analyzing the second moment, namely showing $\mathrm{Var}(M_{m,n;N}(A_N) - M_{m,n})$ 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.* $\lim_{N\to\infty} M_{2m+1,n;N} = 0$

Proof. For the $2m+1^{\text{th}}$ moment, we consider $\mathbb{E}[b_{\psi(i_1,i_2)} b_{\psi(i_2,i_3)} \cdots b_{\psi(i_{2m+1},i_1)}]$; 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 $2m+1$ 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 m 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 $m+1$ degrees of freedom. Thus the number of indices $i_1, \dots, i_k \in \{1, \dots, N\}$ that can contribute to the moment in (2.9) for a given matrix is $O_n(N^{m+1})$ (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 $N^{m+3/2}$ in (2.9), the odd moments are $O_n(N^{-1/2})$, and thus vanish in the limit as $N \to \infty$. □

Corollary 2.2. *For fixed n, as $N \to \infty$ there is no contribution to the average $2m^{th}$ moment from any tuple where the b_ℓ's are not matched in pairs.*

Proof. The corollary follows from a similar analysis as in Lemma 2.1. □

From the above corollary, we see that in order to study the eigenvalues of our matrices we need to know how many different ways the $k = 2m$ entries (the $a_{i_j i_{j+1}}$'s) in our tuples can be matched into $k/2 = m$ pairs. Letting $r!! = r(r-2)(r-4)\cdots$, where the product stops at 1 if r is odd and 2 if r is even, we see there are at most $(2m-1)!!$ ways to match in pairs.[3] Note $(2m-1)!!$ is the $2m^{\text{th}}$ moment of the standard normal, and has the combinatorial interpretation of being the number of ways of matching $2m$ objects in m pairs where order does not matter. For each legitimate matching we obtain a system of m equations, one

[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.

[3]There are $\binom{2m}{2}$ ways to choose the first pair, $\binom{2m-2}{2}$ ways to choose the second and so on; we must divide by $m!$ as it does not matter which pair we call the first. The claim follows by elementary algebra. Alternatively we can prove this by induction. Assume there are $(2m-3)!!$ ways to match $2m-2$ objects in pairs. If we have $2m$ objects, there are $2m-1$ choices of an element to pair with the first element in our list, and then by induction there are $(2m-3)!!$ ways of pairing the remaining $2m-2$ elements.

for each pair of entries, for which the number of solutions is the contribution of the matching to the $2m^{\text{th}}$ moment.

In order to understand the even moments, we need to know more about the permissible matchings, 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

$$|i_m - i_{m+1}| \;=\; |i_l - i_{l+1}|. \tag{2.12}$$

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
$$|i_m - i_{m+1}| \;=\; |i_l - i_{l+1}| + C_1; \tag{2.13}$$

(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
$$|i_m - i_{m+1}| \;=\; -|i_l - i_{l+1}| + C_2; \tag{2.14}$$

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

As a consequence of (2.13) and (2.14), for the matchings above there is some C such that

$$i_n - i_{n+1} \;=\; \pm(i_l - i_{l+1}) + C. \tag{2.15}$$

As there are two choices for sign and m matchings, there are potentially 2^m 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 (2.15)) 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 2.3. *Consider the contribution to the $2m^{th}$ moment from all tuples $(i_1, \dots, i_{2m})$ in which the corresponding b_ℓ's are matched in pairs. If an $a_{i_n i_{n+1}}$ is matched with an $a_{i_l i_{l+1}}$ with a positive sign (which means*

$$i_n - i_{n+1} \;=\; +(i_l - i_{l+1}) + C$$

for some C as defined in (2.13) or (2.14)), then it contributes $O_m(1/N)$ to $M_{2m,n;N}$ and therefore the contribution of all but one of the 2^m choices for the m 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 $(i_1, \dots, i_{2m})$ in which the corresponding b_ℓ's are matched in pairs, there exist k equations, one for each pairing, of the form

$$i_n - i_{n+1} \;=\; \epsilon_m(i_l - i_{l+1}) + C_l \text{ where } \epsilon_l \;=\; 1 \text{ or } -1. \tag{2.16}$$

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

$$\sum_{i=1}^{2k} \tilde{x}_i \;=\; 0. \tag{2.17}$$

Each $\tilde{x}_m$ can be expressed in two ways. By breaking the absolute value sign in (2.13) or (2.14), 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 (2.17) such that $\tilde{x}_m = \epsilon_m \tilde{x}_l + C_m$ for some l. Thus

$$\tilde{x}_m \;=\; \eta_j x_j = \epsilon_m \tilde{x}_l + C_l. \tag{2.18}$$

Then since $\epsilon_m^2 = 1$,

$$\tilde{x}_l \;=\; \epsilon_n \eta_j x_j - \epsilon_n C_l. \tag{2.19}$$

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

$$\sum_{i=1}^{2m} \tilde{x}_i \;=\; \sum_{j=1}^{m} [\eta_j x_j + (\epsilon_n \eta_j x_j - \epsilon_n C_n)] \;=\; \sum_{j=1}^{m} (n_j x_j (1 + \epsilon_n) - \epsilon_n C_j) \;=\; 0. \tag{2.20}$$

If any $\epsilon_m = 1$, then the x_j's are not linearly independent and we would have less than $m+1$ degree of freedom.[4] The contribution from such tuples to the moment in (2.9) for a given matrix is therefore $O(1/N)$ (as we divide by N^{m+1}), which vanishes in the limit as $N \to \infty$ and can thus be safely ignored. □

Lemma 2.3 immediately implies

Lemma 2.4. *If the indices of* $a_{i_l i_{l+1}}$ *and* $a_{i_n i_{n+1}}$ *satisfy* (2.13) *for some* C_1*, then* $|i_n - i_{n+1}| = |i_l - i_{l+1}| + C_1$ *implies*

$$\begin{cases} i_n - i_{n+1} \;=\; -(i_l - i_{l+1}) + C_1 \\ i_n \;>\; \max\{i_{n+1}, i_{m+1} + C_1\} \end{cases} \quad \textit{or} \quad \begin{cases} i_n - i_{n+1} \;=\; -(i_l - i_{l+1}) - C_1 \\ i_n \;<\; \min\{i_{n+1}, i_{n+1} - C_1\}. \end{cases} \tag{2.21}$$

Similarly, if the indices satisfy (2.14) *for some* C_2*, then* $|i_n - i_{n+1}| = -|i_l - i_{l+1}| + C_2$ *implies*

$$\begin{cases} i_n - i_{n+1} \;=\; -(i_l - i_{l+1}) + C_2 \\ i_{n+1} \;<\; i_n \;<\; i_{n+1} + C_2, \end{cases} \quad \textit{or} \quad \begin{cases} i_n - i_{n+1} \;=\; -(i_l - i_{l+1}) - C_2 \\ i_{n+1} - C_2 \;<\; i_n \;<\; i_{n+1}. \end{cases} \tag{2.22}$$

[4]As in the proof of Lemma 2.3, 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 $m-1$ independent x_j's, and thus at most m degrees of freedom.

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 removes some of the Diophantine obstructions that arise when evaluating (2.21) or (2.22) 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:

$$\begin{cases} i_m \in \{a, \dots, b\} \\ i'_m \in \{0, \dots, a\} \cup \{b, \dots, N\} \end{cases} \tag{2.23}$$

3. Properties of the Even Moments

In Lemma 2.1 we showed that the average odd moments vanish in the limit. In this section we analyze the even moments. While the low moments may be computed by brute force, similar to other ensembles we are unable to obtain nice closed form expressions for the higher moments. We have the same difficulties seen in [BB, BDJ, HM]; however, as we shall see in §4, the situation is different for the case of Doubly Palindromic Toeplitz (DPT) Matrices. There, similar to [MMS], we are able to obtain closed form expressions in the general case.

3.1. General Properties.

We first handle the zeroth and second moments.

Lemma 3.1. *Assume that p has mean 0, variance 1 and finite higher moments, and fix the degree of palindromicity n. Notation as above, for all A_N we have $M_{0,n;N}(A_N) = 1$ and $M_{2,n;N}(A_N) = 1$, which implies the average moments in the limit are both 1 (explicitly, $M_{0,n} = 1$ and $M_{2,n} = 1$).*

Proof. From (2.8), we see $M_{0,n,N}(A_N) = 1$. For the second moment, we have

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

Since we choose the b's from a distribution with mean zero and variance 1, the expected value above is just the variance (which is 1), and hence $M_{2,n;N} = 1$, which implies $M_{2,n} = \lim_{N\to\infty} M_{2,n;N} = 1$. □

We need to consider the general even moment. By Corollary 2.2, the only contributions to the moments $M_{2m,n;N}$ (see (2.9)) is when the $a_{i_j i_{j+1}}$'s are matched in pairs. There are $(2m-1)!!$ such matchings; we need to determine the contribution of each matching to $M_{2m,n;N}$.

Each of the $(2m-1)!!$ matchings, hereafter referred to as a **configuration**, leads to a system of m equations of the form (2.21) or (2.22) (with the C's coming from (2.13) and (2.14)), for which each distinct solution gives us one possible choice for the tuples $(i_1, \dots, i_{2m})$ and contributes one to the sum. The analysis is completed by counting how many valid configurations there are (or at least determining the main term).

Determining the exact value is complicated by the fact that there are many ways for an $a_{i_j i_{j+1}}$ and an $a_{i_v i_{v+1}}$ to be paired; they must correspond to the same b_ℓ, but there are many diagonals each can lie on (with the number of diagonals growing with n). Fortunately, we can obtain a weak bound depending on n that nevertheless suffices to prove the existence of a limiting spectral measure. By standard arguments, it suffices to show the average even moments converge as $N \to \infty$ to a sequence satisfying Carleman's condition, and then perform a similar analysis on the variance (for weak convergence) or the fourth moment (for almost sure convergence). We leave the convergence issues to §5, and instead prove the existence of the limits.

Lemma 3.2. *For fixed n, $M_{2m,n}$ exists and*

$$M_{2m,n} \;=\; \lim_{N\to\infty} M_{2m,n;N} \leq (2\cdot 2^n)^m (2m-1)!!, \tag{3.25}$$

which implies the $M_{2m,n}$ satisfy Carleman's condition.

Proof. Fix n and m. Consider one of the $(2m-1)!!$ pairings. We have m equations, and thus we must choose m values for the C's. For each equation there are at most $2 \cdot 2^n$ possible choices for a C; the largest the moment can be is if each possible choice of the C's lead to valid configurations. We therefore assume that happens. The contribution of each configuration to the moment is at most 1. To see this, note that as in all arguments in the subject, the number of tuples that contribute in a given configuration is at most N^{m+1}, which is precisely what we divide by in (2.9). Thus $M_{2m,n;N} \leq (2\cdot 2^n)^m (2m-1)!!$, which implies

$$(2\cdot 2^n)^{-1/2}((2m-1)!!)^{-1/2m} \;\leq\; M_{2m,n;N}^{-1/2m}. \tag{3.26}$$

The existence of the limit is proved analogously to [BB, BDJ, HM, MMS]; now that we know $M_{2m,n}$ is bounded, it is easy to see that the main term of the contribution from each possible configuration is independent of N.

It remains to show that the $M_{2m,n}$ satisfy Carleman's condition by showing the sum of the reciprocals of their $2m^{\text{th}}$ roots diverge. Trivial estimation suffices. As $(2m-1)!! < (2m)^{2m}$, we have $(2m-1)!!^{-1/2m} > 1/2m$, and thus

$$\sum_m M_{2m,n}^{-1/2m} \;>\; \sum_m (2\cdot 2^n)^{-1/2} \cdot \frac{1}{2m}. \tag{3.27}$$

The latter sum is the harmonic sum and diverges, completing the proof. □

In the arguments above we did not attempt to find optimal or even good bounds, as these are not needed for convergence; however, if we want to understand the properties of the limiting spectral measure whose moments are the $M_{2m,n}$'s, then we need a more careful analysis (which we perform in §4).

In the case of real symmetric Toeplitz matrices, some configurations lead to a system of equations with significantly fewer solutions than others. The first instance of this was for the non-adjacent matchings in the fourth moment, which contributed 2/3 and not 1 (see [HM]). In the case of real symmetric palindromic Toeplitz matrices, though, all matchings contributed equally (see [MMS]). Thus the presence of a palindrome leads to very different behavior. This leads to the natural question of what is the effect of increasing the palindromicity. There are now more possible configurations as a given b value is on more diagonals; do all configurations contribute equally? The answer is yes. We prove this in stages.

3.2. **The Fourth Moment.** We analyze the fourth moment in detail below, proving in particular that the adjacent and non-adjacent matchings contribute equally (which we compute in §4). As this analysis generalizes readily to higher moments, we provide complete details. Further, as any distribution with finite mean and variance can be normalized to have mean 0 and variance 1, if the distribution is even then the fourth moment is the first moment to show the 'shape' of the distribution, and thus merits special consideration.

Lemma 3.3 (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:

$$\begin{cases} a_{i_1 i_2} &= a_{j_1 j_2} \\ a_{l_1 l_2} &= a_{m_1 m_2}, \end{cases} \tag{3.28}$$

where in the adjacent configuration case $i_2 = j_1, j_2 = m_1, m_2 = l_1$ and $l_2 = i_1$, while in the non-adjacent configuration case $i_2 = m_1, m_2 = j_1, j_2 = l_1$ and $l_2 = i_1$. We have here a system of equation with 4 unknown variables, since two adjacent entries in the tuples share one common index, and two equations. At least two of those four unknown variables can be free indices. Our goal is to show that the contribution from this general configuration depends on only at most two nontrivially unequal 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:

$$\begin{cases} |i_1 - i_2| &= \pm|j_1 - j_2| + A \\ |l_1 - l_2| &= \pm|m_1 - m_2| + B. \end{cases} \tag{3.29}$$

According the Lemma 2.3, in order for a tuple to contribute to the fourth moment in the limit, we must have

$$\begin{cases} i_1 - i_2 \;=\; -(j_1 - j_2) + A' \\ l_1 - l_2 \;=\; -(m_1 - m_2) + B'. \end{cases} \tag{3.30}$$

Where either $A' = A$ or $A' = -A$ (similarly for B') depending on how we expand the absolute value equations. Also, independent of whether this is the adjacent or non-adjacent configuration, it must be true that:

$$A' + B' \;=\; i_1 - i_2 + j_1 - j_2 + l_1 - l_2 + m_1 - m_2 \;=\; 0. \tag{3.31}$$

This implies that A and B must be of the same form, either in (2.13) or (2.14). If A is of the form C_2 in (2.14), then it follows immediately that $A = B$, while if A is of the form C_1 in (2.13), then it can be either be the case that $A = B$ or it could happen that $A = -B$. For each A, we have a system of two equations with four unknown variables so we can always pick at least two free indices among them. For convenience, we specify i_1, i_2 as free indices by choosing the first entry $a_{i_1 i_2}$ at random. 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, since A can only be of the form C_2 in (2.14) or C_1 in (2.13), we first consider the case where A is of the form C_2 in (2.14), and thus $A = B = C_2$ for some C_2. We therefore find

$$\begin{cases} |j_1 - j_2| \;=\; -|i_1 - i_2| + C_2 \\ |l_1 - l_2| \;=\; -|m_1 - m_2| + C_2 \end{cases} \Longrightarrow \begin{cases} j_1 - j_2 \;=\; -(i_1 - i_2) + C_2 \\ j_1 > j_2 \\ l_1 - l_2 \;=\; -(m_1 - m_2) - C_2 \\ m_1 < m_2 < m_1 + A. \end{cases} \tag{3.32}$$

We now consider A of the form C_1 in (2.13); then $A = \pm B = C_1$ for some C_1 such that $C_1 + C_2 = N - 1$. The value C_1 is unique for each choice of C_2 and the contribution from the pair (C_1, C_2) complements nicely one another as we will show below. We have

$$\begin{cases} |j_1 - j_2| \;=\; |i_1 - i_2| + C_1 \\ |l_1 - l_2| \;=\; |m_1 - m_2| \pm C_1 \end{cases} \Longrightarrow \begin{cases} j_1 - j_2 \;=\; -(i_1 - i_2) - C_1 \\ j_1 < j_2 \\ l_1 - l_2 \;=\; -(m_1 - m_2) + C_1 \\ m_1 < m_2 \textbf{ or } m_2 < m_1 - C_1. \end{cases} \tag{3.33}$$

Since we have already picked the first entry $a_{i_1 i_2}$, choosing the entry $a_{m_1 m_2}$ suffices to specify at least three out of the four unknown variables, and once we know three of the variables then the last variable is determined. Our choice of C_1 (or C_2 since the pair is unique) indicates the diagonals that $a_{m_1 m_2}$ lies on. Finally, as only one of the indices m_1 or m_2 need to be specified (since the other is restricted by the diagonal), without loss of generality we choose m_2. We now use our previous analysis from (2.19) and Lemma 2.4 to analyze the diagonals associated to $A = C_1$ and $A = C_2$.

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

$$m_2 \in \{m_1, \dots, m_1 + C_1\} \cap \{a, \dots, b\} \tag{3.34}$$

(2) On the diagonal associated with $A = C_2$:

$$m_2 \in (\{0, \dots, m_1 - C_2\} \cup \{m_1, \dots, N\}) \cap (\{0, \dots, a\} \cup \{b, \dots, N\}). \tag{3.35}$$

Therefore, there are exactly C_1 out of $N+1$ values of m_2 we can pick (or exactly C_2 out of $N+1$ value of m_2 we cannot pick). Since we have N^2 choices for picking the initial entry $a_{i_1 i_2}$, the contribution to the fourth moment from the pair (C_1, C_2) is given by

$$N^2 \cdot C_1 \;=\; \left(\frac{N^3}{2^n}\right)\left(-\left\lfloor \frac{|i_1 - i_2|}{2^n} \right\rfloor + k - 1\right). \tag{3.36}$$

This contribution only depends on the initial choice of $a_{i_1 i_2}$ and the choice of A. Summing over all possible choices of A of the form C_1, we obtain the same contribution to the fourth moment from either configuration. □

Corollary 3.4. *Given any configuration at the fourth moment, the set of possible values for each of the indices i, j, k and l is same.*

The corollary follows immediately from our formula in **??**. Given an initial entry a_{ij}, we have $\sum_{C_i} N^2 C_i$ satisfying tuples. Since the value of C_i depends only on $\lfloor \frac{|i_1 - i_2|}{2^n} \rfloor$, the number of satisfying tuples with initial entry a_{ji} is also $\sum_{C_i} N^2 C_i$. Therefore, over all satisfying tuples, if i takes on values $\{a, b, c, ...\}$ then j must also takes on the same value $\{a, b, c, ...\}$. Furthermore, should we choose the initial entry at a different vertex of the configuration, we would obtain the same result for the other index k and l.

3.3. **Lifting Configurations and Contributions.** Before extending the Lemma 3.3 to the general even moment, we introduce some notation for a "lift map", which is a way of relating one configuration of an even moment (say one of the $(2m-1)!!$ configurations of the $2m^{\text{th}}$ moment) to one configuration of the next higher even moment (to one of the $(2m+1)!!$ configurations of the $(2m+2)^{\text{nd}}$ moment. If we add a pair of entries to a configuration, this moves us from our initial 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 3.5 (Configuration Lifting - Adjacent Case)**.** *Consider a configuration of the $2m^{th}$ moment. All configurations of the $(2m+2)^{th}$ moment obtained by adding a pair of adjacent entries to this configuration contribute equally to the $(2m+2)^{th}$ 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 from one of the $(2k-1)!!$ configurations of the $2k^{\text{th}}$ moment; for brevity's sake we call this configuration (1). We let

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

be the new tuple obtained by adding the pair of adjacent, matched entries $a_{jo} = a_{os}$; we denote this by configuration (2). Let Ω_{2m} be the set of all tuples that work for configuration (1) and Ω_{2m+2} be the set of all tuples that work for configuration (2). We define a "lift map" $F : \Omega_{2m} \to \Omega_{2m+2}$ by

$$F\left((\dots, a_{pq}, \dots, a_{ij}, a_{jk}, \dots, a_{lm}, \dots)\right) \;=\; (\dots, a_{pq}, \dots, a_{ij}, a_{jo}, a_{os}, a_{sk}, \dots, a_{lm}, \dots). \tag{3.37}$$

Note F maps each index in $(\dots, a_{pq}, \dots, a_{ij}, a_{jk}, \dots, a_{lm}, \dots)$ to itself and inserts a new index $s = j - B + B'$ where B is the value of C corresponding to the pair of entries $(a_{jk} = a_{pq})$, and B' is any value of C such that $s \in \{1, \dots, N\}$ and $(B - B')$ is a valid value 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:

$$\begin{cases} l - m \;=\; -(i-j) + A \\ p - q \;=\; -(j-k) + B \\ \dots \\ \dots . \end{cases} \tag{3.38}$$

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

$$\begin{cases} l - m \;=\; -(i-j) + A \\ p - q \;=\; -(s-k) + B' \\ j - o \;=\; -(o-s) + (B - B') \\ \dots . \end{cases} \tag{3.39}$$

Except for the two equations $p-q = -(s-k)+B'$ and $j-o = -(o-s)+(B-B')$, every other equation of configuration (1) is preserved under F and therefore still holds in configuration (2). Furthermore, since both B' and $(B - B')$ are valid choices of the C value by the construction of F, the two equations $p - q = -(s - k) + B'$ and $j - o = -(o - s) + (B - B')$ are also valid. Thus

$$(\dots, a_{pq}, \dots, a_{ij}, a_{jo}, a_{os}, a_{sk}, \dots, a_{lm}, \dots) \in \Omega_{2m+2}. \tag{3.40}$$

Next, we need to worry about the issue of whether the lift map F is well-defined and exhausting. However, we fortunately only need to show that F is onto to claim equal contribution. This is due to the fact that the map F only depends on the choice of one index j and one C value B from the original configuration (1). First, we can take j to be any value in $\{1, \dots, N\}$. For any configuration in the $2m^{\text{th}}$ moment, we have $2m$ indices (unknown variables) and m equations with the last equation linearly dependent on the rest. Thus we must have at least two completely free indices that can take on any value between 1 and N. We specify the two completely free indices by choosing the very first entry of the tuple, which obviously can be any entry in the matrix. Furthermore, we can also choose the first entry to be any vertex 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. Second, 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 combinations of the C's that work for the system of equations corresponding to configuration (1).

Hence, starting from a configuration of the $2m^{\text{th}}$ 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 a consequence, the number of tuples resulting from the map F is the same regardless of where we add the adjacent pair *if* the map F can reach every possible tuple of configuration (2).

We complete the proof by showing F is onto. **(IS THIS MORE INVOLVED THAN NEEDED? DO WE HAVE F ONTO BY THE DEFINITION OF THE TWO SETS Ω_{2k} AND Ω_{2k+2}? OR IS IT THAT WE HAVE A VALID C CHOICE AS WELL?)** Assume not; then F is not onto and hence there exists some tuple

$$b \;=\; (\dots, a_{pq}, \dots, a_{ij}, a_{jo}, a_{os}, a_{sk}, \dots, a_{lm}, \dots) \in \Omega_{2k+2}$$

such that for all tuples $a \in \Omega_{2m}$ we have $F(a) \neq b$. Since $a_{jo} = a_{os}$ and $a_{sk} = a_{pq}$ then for some C-value B and C

$$\begin{cases} p - q \;=\; -(s-k) + A \\ j - o \;=\; -(o-s) + B \end{cases} \Longrightarrow \begin{cases} p - q \;=\; -(j-k) + (A+B) \\ s \;=\; j - B. \end{cases} \tag{3.41}$$

Now we consider the tuple $a = (\dots, a_{pq}, \dots, a_{ij}, a_{jk}, \dots, a_{lm}, \dots)$ where $j = s + B$. Everything entry except a_{jk} and a_{pq} are matched in pair. Furthermore, a_{jk} and a_{pq} would also be matched in pair, and therefore $a \in \Omega_{2m}$, if for some C-value A and B then $A + B$ is a valid C-value. There are $k+1$ equations corresponding to the configuration Ω_{2m+2}. Let $C_1, \dots, C_{k+1}$ be the C-value corresponding to each of those equations. Lemma 2.3 implies that

$$C_1 + C_2 + \cdots + C_{m+1} \;=\; 0. \tag{3.42}$$

We first analyze the low moments to build intuition. For the forth moment ($k = 1$ so $2k + 2 = 4$) then $0 = C_1 + C_2$ is obviously a valid C-value. For the sixth moment ($k = 2$) then $-C_3 = C_1 + C_2$ is also a valid C-value. More generally, for any even moment, if there does not exist a valid C-value for which $A + B$ is also a valid C-value then $p - q = -(j-k) + (A+B)$ would imply that given the entry a_{pq} there exists no entry on the j^{th} column of our matrix that can be matched in pair with a_{pq}. This is a contradiction since just the single palindromic condition guarantees that for any entry on the matrix, there exists at least one entry on each column (or row) that can be matched in pair with it. **(DOES SOMETHING LIKE THIS MAKE ALL THIS TRIVIAL? ARE WE JUST DOING EXISTENCE OR NUMBER OF MATCHINGS?)** Finally, it is clear by constructioin that $F(a) = b$, which completes the proof that F is onto. $\square$

Lemma 3.6 (Configuration Lifting - Nonadjacent Case)**.** *Consider a configuration of matchings for the $2m^{th}$ moment. All configurations at the $(2m+2)^{th}$ moment obtained by adding a pair of non-adjacent entries contribute equally to the $(2m+2)^{nd}$ 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 a configuration for the $2k^{\text{th}}$ moment; denote this by configuration (1). We let $(\dots, a_{ij}, a_{jo}, a_{om}, \dots, a_{lp}, a_{ps}, a_{sq}, \dots)$ 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 configuration (1) and $\Omega^{'}_{2k+2}$ be the set of all tuples that work for the configuration (2). We define the "lift map" $FF^{'} : \Omega^{'}_{2k} \to \Omega^{'}_{2k+2}$ in this case by

$$F^{'}\left((\dots, a_{ij}, a_{jm}, \dots, a_{lp}, a_{pq}, \dots)\right) \;=\; (\dots, a_{ij}, a_{jo}, a_{om}, \dots a_{lp}, a_{ps}, a_{sq}, \dots) \tag{3.43}$$

such that $F^{'}$ maps every index in $(\dots, a_{ij}, a_{jm}, \dots, a_{lp}, a_{pq}, \dots)$ to itself and adds 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 value of C. For the tuple $(\dots, a_{ij}, a_{jm}, \dots, a_{lp}, a_{pq}, \dots)$ we have the following system of equations:

$$\begin{cases} i_0 - j_0 \;=\; -(i-j) + A \\ j_0 - m_0 \;=\; -(j-m) + B \\ l_0 - p_0 \;=\; -(l-p) + C \\ p_0 - q_0 \;=\; -(p-q) + D \\ \dots \end{cases} \tag{3.44}$$

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:

$$\begin{cases} i_0 - j_0 \;=\; -(i-j) + A \\ j_0 - m_0 \;=\; -(o-m) + B^{'} \\ l_0 - p_0 \;=\; -(l-p) + C \\ p_0 - q_0 \;=\; -(s-q) + D^{'} \\ j - o \;=\; -(p-s) + (D^{'} + B^{'} - D - B) \\ \dots \end{cases} \tag{3.45}$$

Similar to the analysis in Lemma 3.5, all equations except for $j_0 - m_0 = -(o - m) + B^{'}$ and $p_0 - q_0 = -(s-q) + D^{'}$ and $j - o = -(p-s) + (D^{'} + B^{'} - D - B)$ are preserved under the map $F^{'}$ so they 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 3.5, 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. □

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

(Vincent) Please ignore this theorem for now. I'm making new pictures to reflect the proof better

Proof. Given any configuration at the $(2k+2)^{\mathrm{nd}}$ 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

2

For any configuration at the $(2k+2)^{\mathrm{th}}$ 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 3.3 still holds true for this case if we choose the two completely free indices in Lemma 3.3 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 $(2k+2)^{\mathrm{th}}$ moment and therefore all configurations at the $(2k+2)^{\mathrm{th}}$ 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 3.3 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

We determine below closed form expressions for the moments where possible. Similar to other ensembles, there appear to be no nice closed form expressions for general moments for arbitrary n; however, we are able to calculate the fourth moment for any n, and in the special case of doubly palindromic matrices we calculate *all* the even moments. Since each matching configuration contributes equally, we need only determine the contribution from an adjacent matching in the doubly palindromic case and then multiply by the number of configurations to obtain the value for the even moments. We end this section with a discussion on why this approach is not readily generalizable to matrices with a greater number of palindromes.

4.1. The Fourth Moment.

For the fourth moment, we have four indices i, j, k, and l, and we consider an adjacent matching where

$$a_{ij} = a_{jk}, \qquad a_{kl} = a_{li}.$$

We think about this as follows. A pair i and j gives us a matrix element a_{ij}; we want to find all pairs j and k such that $a_{ij} = a_{jk}$. This could happen by having the two on the same diagonal, or it could happen that a_{jk} is on a palindromically equivalent diagonal. As the formula for the fourth moment of our matrix A_N involves division by N^3, we need only worry about situations where we have on the order of N^3 tuples. Clearly we may choose i and j freely. The matching then forces there to be on the order of 1 choice for k (the exact answer depends on n, the degree of palindromicity; what matters is that the answer is independent of N in the limit), and on the order of N choices for l. The last is important, as unless the number of choices of l is proportional to N, we will obtain a negligible contribution from the matching $a_{ij} = a_{jk}$ and $a_{kl} = a_{li}$. Exploiting the symmetry of the matrix, this reduces to choosing k so that $a_{ij} = a_{kj}$ and $a_{kl} = a_{il}$. That is, we want row i and row k to match well.

Figure 6. An example highlighting matchings for l in medium shading, with mismatching in dark shading. Note that any anomalous matchings won't contribute in the limit.

We isolate some of the most useful features of our matrices in the following lemma. The proof follows immediately from the previous discussions and the structure of the matrices in our ensemble.

Lemma 4.1. *Fix n and consider the ensemble of $N \times N$ real symmetric palindromic Toeplitz matrices with 2^n palindromes in the first row. The main diagonal is the only place (excluding the border of the matrix) where b_0 occurs once rather than twice. This implies the following useful properties.*
Moving to the corresponding point in the next palindrome can require either moving $N2^{-n} - 1$ elements when crossing the main diagonal or $N2^{-n}$ elements otherwise. As pictured in Figure 6, a given row and the row $bN2^{-n}$ rows down from that given row do not match perfectly, but rather become unaligned when one row has reached the main diagonal but the other row has not. Moreover, the row $bN2^{-n} - 1$ 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.

Proof. The first item follows directly from the observation that the main diagonal is the only place where b_0 appears once rather than consecutively. We also see that, neglecting the first row which starts on the main diagonal, that the first elements of a row and one $bN2^{-n}$ rows away match initially. Moreover, they evolve the same way when moving from left to right, except when the first one hits the main diagonal, in which case it skips forward one place in the palindrome, in which case they do not match except possibly for repeating elements at the beginning/end or middle of palindromes, like b_0. However, once the second row hits the main diagonal, it also skips forward, and they become realigned. The case for rows $bN2^{-n} - 1$ rows away from each other is argued similarly.

To prove that no other rows match sufficiently well we need to show that there are only $O_n(1)$ matchings in any of the other rows. Suppose we do have a matching in one of the other rows. Since we can't be at the corresponding point in the palindrome, we must be at the other end of the palindrome. Thus, the two rows will evolve in opposite directions, so although there may be additional anomalous matchups, there will certainly not be more than four per palindrome, giving us the desired maximum of $O_n(1)$ possible matchings. □

Lemma 4.2. *Let $b \in \{0, 1, \ldots, 2^{n-1}\}$ and $k = i + bN2^{-n}$. There are then*

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

good matchings, whereas if $k = i + bN2^{-n} - 1$, then there are

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

good matchings, where the big-Oh constants depend on n (which is fixed).[5]

[5]The constants may be taken to depend on b as well; however, as n is fixed and $b \in \{0, \ldots, 2^{n-1}\}$, we may take the maximum of all the constants and may replace b dependence with n dependence.

Proof. We begin by noting that by Lemma 4.1 above, choosing k so that a_{ij} and a_{kj} are at corresponding points in a palindrome guarantees that $a_{ij} = a_{kj}$ and that there are $O_n(N)$ choices of l satisfying $a_{kl} = a_{il}$, as desired. Moreover, if a_{ij} and a_{kj} aren't at corresponding points in the palindrome, then there are only $O_n(1)$ good choices of l, and since there are at most n such possible cases, this contribution can be ignored. Thus, we only consider the cases where a_{ij} and a_{kj} are at corresponding places in a palindrome.

We now consider the case when $k = i + bN2^{-n}$, hence a_{ij} and a_{kj} must be on the same side of the main diagonal in order to match. Moreover, to have $k \in \{1, 2, \dots, N\}$ we must have $i \in \{1, 2, \dots, N - bN2^{-n}\}$. Another restriction arises from the fact that they are on the same side of the main diagonal. We note that we won't cross the main diagonal when moving down from any a_{ij} below the main diagonal to a_{kj}. There will similarly be no crossing if a_{ij} lies more than $bN2^{-n}$ elements above the main diagonal. This defines two right-triangular regions of height $N - bN2^{-n} + O_n(1)$, which in total gives a square of area

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

from which to choose a_{ij}, thus giving that many valid choices of a_{ij}. We also have the restriction on the values of l as explained in Lemma 4.1, leaving $N - bN2^{-n} + O_n(1)$ good values of l for each of these a_{ij}. In total $bN2^{-n}$ contributes

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

matchings to the fourth moment.

Next we consider the case where we cross the main diagonal when moving from a_{ij} to a_{kj}, so that $k = i + bN2^{-n} - 1$. In this case, the area of values of a_{ij} from which we will cross the diagonal to give a matching will be mostly defined by the parallelogram bordered by the triangles from the previous constant. However, there may also be additional strips as depicted in light shading in Figure 7, but these will only be of width 1, so the area is essentially that of the parallelogram of height $N - bN2^{-n} + O_n(1)$ and width $bN2^{-n} + O_n(1)$. There will also be $bN2^{-n} + O_n(1)$ good values of l, so in all this constant contributes

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

matchings. □

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

We now comment on what happens for the negative constants $\{-bN2^{-n}, -(bN2^{-n}-1)\}$ for $b \in \{1, 2, \dots, 2^n - 1\}$, in which case we are now moving up b palindromes, and either crossing or not crossing the mian diagonal, respectively. We easily see that this is essentially switching the roles of a_{ij} and a_{jk}, so the contributions will be the same. If we repeat the analysis above we find regions of identical size that thus give identical contributions to the fourth moment. Pictorially, what happens for a negative constant is that of the positive one rotated 180 degrees. Thus, the contribution to the fourth moment will be given by the contributions from the positive constants $(\{bN2^{-n}, bN2^{-n} - 1\}$ for $b \in \{1, 2, \dots, 2^n - 1\}$ multiplied by a factor of 2 to account for the negative constants.

Lemma 4.3. *Fix n and consider the limit as $N \to \infty$ of the average fourth moment of our ensemble. The contribution from one of the adjacent matching configurations (i.e., $a_{ij} = a_{jk}$ and $a_{kl} = a_{li}$) to this limit is*

$$\frac{2}{3}2^n + \frac{1}{3}2^{-n}. \tag{4.46}$$

Proof. For each value of b, we note that the contribution to $M_{4,n}(N)$ is

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

Taking the limit as $N \to \infty$ yields the contribution to $M_{4,n}$ 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 2 to account for the negative constants, and include the contribution from $C = 0$, known to be 1 from [MMS] to obtain the contribution to the fourth moment from the adjacent matching case:

$$M_{4,n}^{\text{adj}} \;=\; 1 + \frac{2}{2^{3n}} \sum_{b=1}^{2^n} \left((2^n - b)^3 + 2^n b^2 - b^3\right). \tag{4.47}$$

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:

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

completing the proof. □

4.2. **The General Even Moments of DPT Matrix.** Using Theorem 3.8, we can determine any moment by calculating the contribution from one of the adjacent configurations. we generalize the pictorial method of the previous subsection to higher moments. For the $2m^{\text{th}}$ moment, we find that our final system of equations for an adjacent configuration becomes

$$\begin{aligned} i_3 &= i_1 + C_1 \\ i_5 = i_3 + C_2 &= i_1 + C_1 + C_2 \\ &\vdots \\ i_1 = i_{2m-1} + C_m &= i_1 + \sum_{k=1}^{m} C_k. \end{aligned}$$

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} analagous to the restrictions on l for the fourth moment. Moreover, as each i_{2n+1} is related back to i_1, the difference between the maximum and minimum partial sums must be strictly less than $N + O(1)$ or we lose a degree of freedom.*

These observations allow us to calculate the value of every moment for the doubly palindromic case. First, we give some useful definitions.

Definition 4.5. *A **C-vector** is the ordered collection of constants relating the odd indices to each other. In the example at the beginning of this subsection, the C-vector would be $(C_1, C_2, \ldots, C_m)$.*

*A **core** of a C-vector is the ordered collection of nonzero constants in the C-vector. That is, we collapse down the C-vector to its core by removing all of the zero constants from it. We can also think of the C-vector as being built up from the core by adding back the zeros in the correct places.*

Theorem 4.6. *The $2m^{th}$ moment averaged over the ensemble of doubly palindromic Toeplitz matrices is given by:*

$$M_{2m,1} \;=\; (2m-1)!!\left(-2 + 2^{-m}\left(\sum_{b=1}^{3} b^m\right)\right). \tag{4.48}$$

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

- If the constants $\pm\frac{N}{2}$ appear in the C-vector $(C_1, \ldots, C_n)$, then $\pm\frac{N}{2} - 1$ cannot occur as a main term. If it did, we would lose a degree of freedom in i_2, as $a_{i_1 i_2}$ would need to lie on a very specific set of diagonals.
- If some C_j is non-zero, then the next non-zero C chosen must be $-C_j$, as we would otherwise lose a complete degree of freedom in i_1.

Now consider the $2m^{\text{th}}$ moment, which will have a C-vector of length m. We can then consider a subset of length k (k even) of $(\frac{N}{2}, -\frac{N}{2}, \ldots, \frac{N}{2}, -\frac{N}{2})$ that forms the core of the C-vector, with the remaining entries being zero. There are then

$\binom{m}{k}$ distinct ways to insert the remaining zeros, and thus $\binom{m}{k}$ ways to build a C-vector around this core.

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

$$\sum_{\substack{k \text{ even} \\ k=2}}^{m} \binom{m}{k} \left(\frac{1}{2}\right)^{k+1}. \tag{4.49}$$

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

$$\frac{1}{4}\left(\left(1+\frac{1}{2}\right)^m + \left(1-\frac{1}{2}\right)^m\right) - \frac{1}{2}. \tag{4.50}$$

The contribution from a core of $(-\frac{N}{2}, \frac{N}{2}, \dots, -\frac{N}{2}, \frac{N}{2})$ will be the same as that above. The cores of $(\pm(\frac{N}{2}-1), \mp(\frac{N}{2}-1), \dots, \pm(\frac{N}{2}-1), \mp(\frac{N}{2}-1))$ 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 (4.50) by 4. We also include the contribution from the 0-vector, which is 1 for the adjacent case. Thus, the contribution from each configuration is

$$-2 + \left(1+\frac{1}{2}\right)^m + 1^m + \left(1-\frac{1}{2}\right)^m \;=\; -2 + 2^{-m}\left(\sum_{b=1}^{3} b^m\right). \tag{4.51}$$

Appealing to Theorem 3.8 and multiplying by $(2m-1)!!$ (the number of configurations, i.e., the number of ways of matching $2m$ objects in m pairs with order not mattering), we have

$$M_{2m,2} \;=\; (2m-1)!!\left(-2 + 2^{-m}\left(\sum_{b=1}^{3} b^m\right)\right), \tag{4.52}$$

completing the proof. $\square$

Remark 4.7. *Unfortunately, this method does not readily generalize to matrices with a greater number of palindromes. The fundamental reason is that the observations made at the beginning of Theorem 4.6 no longer hold for these matrices, which then makes it tremendously more difficult to generate all valid C-vectors.*

To demonstrate these difficulties, we investigate the 6^{th} moment of a matrix with four palindromes. While we can construct C-vectors in the same manner as for the doubly palindromic ensemble, we clearly will be missing a substantial number of possible C-vectors. For instance, we miss the vector $\left(\frac{N}{2}, \frac{N}{4}, -\frac{3N}{4}\right)$.

In addition to these vectors, we have even more problematic vectors such as $\left((\frac{N}{2}, \frac{N}{4} - 1, -\left(\frac{3N}{4} - 1\right)\right)$, which turns out to be valid for a_{ij} chosen within a certain band. These new vectors, in which "mixing" is important, make it fiendishly difficult to determine precisely which C-vectors to include and which to exclude.

While we could in principle calculate any given moment for any number of palindromes, there is no apparent method that will work for all of these simultaneously. Similar to other investigations on related ensembles, we are left with general existence proofs of the moments, as well as estimates on their rate of growth.

Whenever one has involved combinatorial arguments such as the ones above, it is worthwhile to numerically test our theory. We looked at 1000 real symmetric doubly palindromic 2048×2048 Toeplitz matrices, and compared the even moments to our predicted values (as expected, the odd moments were small). Unfortunately the rate of convergence is slow in N due to the presence of large big-Oh constants.

moment	predicted	observed	observed/predicted
2	1.000	1.001	1.001
4	4.500	4.521	1.005
6	37.500	37.887	1.010
8	433.125	468.53	1.082
10	6260.63	107717.3	17.206

It is worth noting how slow the convergence is. For example, when we considered 1000 real symmetric doubly palindromic 96×96 Toeplitz matrices, the observed second moment was 0.990765, the fourth moment was 4.75209, the sixth was 45.7965 (for a ratio of 1.22) and the eight was 737.71 (for a ratio of 1.70).

5. Convergence

In §2 (specifically Lemma 2.1 and Corollary 2.2) we showed that the limit of the average moments exist as $N \to \infty$, and in Lemma 3.2 we proved that the moments grow slowly enough to uniquely determine a probability distribution. We now show convergence in probability, and if $p(x)$ is even, we prove almost sure convergence. As these arguments closely follow those in [HM, MMS], we concentrate on the novelties introduced by the higher palindromicity. We conclude by obtaining lower bounds for the moments. These bounds imply that our limiting distributions have unbounded support and fatter tails than the standard normal (possibly the fattest tails observed from a random matrix ensemble arising from entries chosen independently from a distribution whose moment generating function converges in a neighborhood of the origin).

5.1. **Weak Convergence.** Recall

Definition 5.1 (Weak Convergence)**.** *A family of probability distributions μ_n weakly converges to μ if and only if for any bounded, continuous function f we have*

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(x)\mu_n(dx) \ = \ \int_{-\infty}^{\infty} f(x)\mu(dx). \tag{5.53}$$

We will prove our probability measures converge weakly. We follow the arguments in [HM]. We begin by defining our random variables. Let A be a sequence of real numbers to which we associate an $N \times N$ real symmetric Toeplitz matrix with 2^n palindromes, which we denote by A_N. Thus we may view A as $(b_0, b_1, b_2, \dots)$, and we form A_N by considering the initial segment of length $N/2^{n+1}$, taking that as the first half of our palindrome, and then building the matrix by having 2^n palindromes in the first row.

Let $X_{m,n;N}(A)$ be a random variable that equals the m^{th} moment of A_N (so $X_{m,n;N}(A) = M_{m,n;N}(A_N)$), and set $M_{m,n;N}$ to the m^{th} moment averaged over the ensemble as above (so $M_{m,n;N} = \mathbb{E}[X_{m,n;N}]$).

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

$$\lim_{N\to\infty} \mathbb{P}_{\mathbb{N}}(\{A \in \Omega_{\mathbb{N}} : |X_{m,n;N} - X_{m,n}| > \epsilon\}) \ = \ 0. \tag{5.54}$$

Using Chebyshev's inequality and the fact that $\mathrm{Var}(Y) = \mathbb{E}[(Y - \mathbb{E}[Y])^2] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$, we have

$$\begin{aligned} \mathbb{P}_{\mathbb{N}}(\{A \in \Omega_{\mathbb{N}} : |X_{m,n;N} - \mathbb{E}[X_{m,n;N}]| > \epsilon\}) \\ \leq\ & \frac{\mathbb{E}[(X_{m,n;N} - M_{m,n;N})^2]}{\epsilon^2}. \\ =\ & \frac{\mathbb{E}[X_{m,n;N}^2] - M_{m,n;N}^2}{\epsilon^2}. \end{aligned} \tag{5.55}$$

Thus, it suffices to show

$$\lim_{N\to\infty} (\mathbb{E}[X_{m,n;N}^2] - M_{m,n;N}^2) \ = \ 0 \tag{5.56}$$

to prove convergence in probability.

We have

$$\begin{aligned} \mathbb{E}[X_{m,n;N}^2] \ &= \ \frac{1}{N^{m+2}} \sum_{1 \leq i_1, \dots, i_m \leq N} \\ &\quad \times \sum_{1 \leq j_1, \dots, j_m \leq N} \mathbb{E}[b_{|i_1 - i_2|} \cdots b_{|i_m - i_1|} b_{|j_1 - j_2|} \cdots b_{|j_m - j_1|}], \\ M_{m,n;N}^2 \ &= \ \frac{1}{N^{m+2}} \sum_{1 \leq i_1, \dots, i_m \leq N} \mathbb{E}[b_{|i_1 - i_2|} \cdots b_{|i_m - i_1|}] \\ &\quad \times \sum_{1 \leq j_1, \dots, j_m \leq N} \mathbb{E}[b_{|j_1 - j_2|} \cdots b_{|j_m - j_1|}]. \end{aligned} \tag{5.57}$$

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 $\mathbb{E}[X_{m,n;N}^2]$ and to $M_{m,n;N}^2$ 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 weak convergence follow.

5.2. **Almost Sure Convergence.** We assume that $p(x)$ is even for convenience, and use the same notation as above; in particular,

$$M_{m,n} \;=\; \lim_{N\to\infty} M_{m,n;N} \;=\; \lim_{N\to\infty} \mathbb{E}[X_{m,n;N}(A)]. \tag{5.58}$$

Almost sure convergence follows from showing that as $N \to \infty$ the event

$$\{A \in \Omega_{\mathbb{N}} \;:\; \lim_{N\to\infty} X_{m,n;N}(A) \to M_{m,n}\}$$

occurs with probability one for all non-negative integers m.

By the triangle inequality, we have that

$$|X_{m,n;N}(A) - M_{m,n}| \;\le\; |X_{m,n;N}(A) - M_{m,n;N}| + |M_{m,n;N} - M_{m,n}|. \tag{5.59}$$

We have already shown that $\lim_{N\to\infty} |M_{m,n;N} - M_{m,n}| = 0$, so we must show that $|X_{m,n;N}(A) - M_{m,n;N}|$ almost surely tends to zero. Clearly $\mathbb{E}[X_{m,n;N}(A) - M_{m,n;N}] = 0$, and we can modify the arguments in [HM] to show that the fourth moment of $X_{m,n;N}(A) - M_{m,n;N}$ 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) in [HM] hold for any of our C-vectors, so a similar result holds in this case. We then apply Chebyshev's inequality to find

$$\mathbb{P}_{\mathbb{N}}(|X_{m,n;N}(A) - M_{m,n;N}| \ge \epsilon) \;\le\; \frac{\mathbb{E}[|X_{m,n;N}(A) - M_{m,n;N}|^4]}{\epsilon^4} \;\le\; \frac{C_{m,2^n}}{N^2\epsilon^4}. \tag{5.60}$$

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

5.3. **Moment bounds and fat tails.** We now extend Theorem 4.6 to matrices with greater palindromicity. In doing so, we miss many of the C-vectors that contribute to these moments, but exact calculations for even a quadruply palindromic matrix have proven difficult. The goal is to obtain good enough bounds on the moments to deduce properties of the limiting spectral measures. We begin with the following lemma.

Lemma 5.2. *For any fixed n we have*

$$M_{2m,n} \;\ge\; \left(-2\cdot(2^n-1) + 2^{-mn}\left(\sum_{b=1}^{2\cdot 2^n-1} b^m\right)\right)\cdot(2m-1)!!. \tag{5.61}$$

As $m \to \infty$, we have

$$M_{2m,n} \;\gg\; \frac{2^{mn}}{m}\cdot(2m-1)!!. \tag{5.62}$$

Proof. Let $C_b = \frac{bN}{2^n}$ for $b \in \{1,\dots,2^n-1\}$. We determine the contribution to the average $2m^{\text{th}}$ moment as $N \to \infty$ from one of the two adjacent matchings. That is, consider the core (i.e., the non-zero part) of the corresponding C-vectors is $(\pm C_b, \mp C_b, \ldots, \pm C_b, \mp C_b)$ and its complement $(\pm(N-1-C_b), \mp(N-1-C_b), \ldots, \pm(N-1-C_b), \mp(N-1-C_b))$. They contribute

$$-2 + \left(2 - \frac{b}{2^n}\right)^m + \left(\frac{b}{2^n}\right)^m. \tag{5.63}$$

The proof of this claim goes back to the observation in Figure 6 that for $C_b = \frac{bN}{2^n}$, the number of free l values is $N - \frac{bN}{2^n} + O(1)$, whereas if $C_b = \frac{bN}{2^n} - 1$, then there are $\frac{bN}{2^n} + O(1)$ good l values. Thus the complementary C_b 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

$$-2 + 2 \cdot \sum_{j \text{ even}}^{m} \binom{m}{j} \left(\frac{2^n - b}{2^n}\right)^{j+1} = -2 + \left(\frac{2 \cdot 2^n - b}{2^n}\right)^m + \left(\frac{b}{2^n}\right)^m. \tag{5.64}$$

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

$$-2 \cdot (2^n - 1) + 2^{-mn} \left(\sum_{b=1}^{2 \cdot 2^n - 1} b^m\right), \tag{5.65}$$

which is easily summed for any value of k. From Theorem 3.8, each of the $(2m-1)!!$ matchings contribute equally, and hence

$$M_{2m,n} \ \geq \ \left(-2 \cdot (2^n - 1) + 2^{-mn} \left(\sum_{b=1}^{2 \cdot 2^n - 1} b^m\right)\right) \cdot (2m-1)!!. \tag{5.66}$$

The behavior for large m follows by approximating the sum with an integral. $\square$

We can now turn to an analysis of the properties of the limiting spectral measures. Note $n = 0$ corresponds to the real symmetric palindromic Toeplitz matrices studied in [MMS], and $n = 1$ corresponds to the doubly palindromic Toeplitz matrices. We now prove Theorem 1.3, which we restate below for the reader's convenience.

Theorem 1.3 (Fat Tails). *Consider the ensemble from Theorem 1.2. For any fixed $n \geq 1$, the moments grow faster than the corresponding moments of the standard normal; specifically, if $M_{2m,n}$ denotes the $2m^{th}$ moment of the limiting spectral measure of our ensemble for a given n, then*

$$M_{2m,n} \ \gg \ \frac{2^{mn}}{m} \cdot (2m-1)!!. \tag{5.67}$$

The limiting spectral measure thus has unbounded support, and fatter tails than the standard normal (or in fact any of the known limiting spectral measures arising from an ensemble where the independent entries are chosen from a density whose moment generating function converges in a neighborhood of the origin).

Proof. From Lemma 5.2 we know

$$M_{2m,n} \gg \frac{2^{mn}}{m} \cdot (2m-1)!!. \tag{5.68}$$

As $n \geq 1$, for m large this is greater than the m^{th} moment of the standard normal, which is $(2m-1)!!$. Thus our limiting spectral measure has unbounded support, and more mass in the tails than the standard normal, or in fact, any normal if $n \geq 2$. To see the last claim, note that if $X \sim N(0, \sigma^2)$ then the $2m^{\text{th}}$ moment of X is $\sigma^{2m} \cdot (2m-1)!!$, and thus when $n \geq 2$ eventually the moment of our ensemble is greater than the moment of this normal. □

I DON'T KNOW HOW TO START AN APPENDIX, SO I'M JUST PUTTING IN A NEW SECTION FOR NOW. I'M ALSO NOT SURE EXACTLY HOW FORMAL AND FLESHED OUT THIS SECTION SHOULD BE.

6. Numerical Methods

While they can never be accepted as proof, numerical simulations did much to guide our efforts in attacking this problem, and we would not have been successful without it, as our naive adaptations of previous works on this subject failed to give even remotely accurate predictions. Therefore we give a brief outline of our use of these simulations below.

At first, we primarily used a basic, direct method to approximate the momemts of the eigenvalue distribution. We first set up a matrix with 2^n palindromes and choose N so that the matrix has the desired form (every element appears exactly 2^{n+1} times in the first row). For each moment we use the eigenvalue trace lemma to calculate the moment of the eigenvalue distribution for this particular matrix, then we average over a large number of such random matrices to get an approximation for that moment averaged over the ensemble of Toeplitz matrices with 2^n palindromes. To get increased accuracy, we would simply increase N.

While this method was quite useful and accurate for lower moments or for a small number of palindromes, for larger values of these quantities the big-Oh constants grew quite large, making it computationally prohibitive to simulate a representative sample of sufficiently large matrices, and thus leaving us with a rather poor estimate of the moments and providing no guide to whether or not our formulas were accurate.

To avoid simulating ever-larger matrices, we instead realized that the $2m^{\text{th}}$ moment of an N by N matrix will satisfy

$$M_{2m,n;N} = M_{2m,n} + \frac{C_1}{N} + \frac{C_2}{N^2} + \cdots + \frac{C_m}{N^m}.$$

Thus, rather than simulating prohibitively large matrices, we could instead simulate large numbers (to increase the likelihood of a representative sample) of several sizes of smaller matrices then perform a regression to estimate the value of $M_{2m,n}$.

In performing these regressions, we sometimes found big-Oh constants so large that it would have been impossible to sufficiently large matrices to get an accuracy of within a few percent for the moments. For example, for the fourth moment of a Toeplitz matrix with 64 palindromes we found the big-Oh constant to be above $30,000$ so that averages of quite large matrices would give an approximation for the fourth moment that would be off by 10, compared to a true value of about 128.

While this again cannot replace a proof, it was successful in verifying our predictions for the higher moments of the doubly palindromic Toeplitz matrix and for the fourth moment of the 64-palindrome Toeplitz matrix.

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.

[Kar] V. Kargin, *Spectrum of random Toeplitz matrices with band structure*, Elect. Comm. in Probab. **14** (2009), 412–421.

[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.

[LW] D.-Z. Liu and Z.-D. Wang, *Limit Distribution of Eigenvalues for Random Hankel and Toeplitz Band Matrices*, preprint, 2009.

[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