Limiting Spectral Measures for Random Matrix Ensembles with a Polynomial Link Function

by

Kirk Swanson

Steven J. Miller, Advisor

A thesis submitted in partial fulfillment
of the requirements for the
Degree of Bachelor of Arts with Honors
in Mathematics

Williams College
Williamstown, Massachusetts

May 14, 2014

ABSTRACT

Given an ensemble of $N \times N$ random matrices, an interesting question to ask is: Do the empirical spectral measures of typical matrices converge to some limiting measure as $N \to \infty$? The limiting measures of several canonical matrix ensembles, such as the symmetric Wigner, Toeplitz, and Hankel matrices, have been well studied. It is known that in the limit, the Wigner matrices have a semicircular distribution, the Toeplitz have a near-Gaussian distribution, and the Hankel have a non-unimodal distribution. Although it is not fully understood why, these ensembles exhibit the remarkable property that as more constraints are introduced to the structure of a random matrix ensemble in the form of a pattern on the matrix entries, new limiting distributions other than the semicircle can arise. It is natural, then, to explore the question: To what extent will a patterned random matrix continue to have a semicircular limiting eigenvalue distribution? In the following, we explore this question by generalizing the Toeplitz and Hankel ensembles. The resulting matrix ensembles with bivariate polynomial link functions have unique limiting spectral distributions. In specific cases, we establish that when the variables in the polynomial are raised to the same power, the limiting distribution becomes non-semicircular, but when the variables are raised to different powers, the limiting distribution remains semicircular.

ACKNOWLEDGEMENTS

I would like to thank Professor Steven J. Miller for being an excellent advisor by providing continual support and encouragement and for introducing me to the fascinating subject of random matrix theory. I am indebted to the faculty of the Department of Mathematics & Statistics at Williams College for providing an atmosphere that encourages all students to learn to enjoy mathematics. I am grateful for the numerous math and physics students at Williams who have inspired me by their enthusiasm for these subjects. I am thankful for the insights of my brother, Kyle Swanson, and Jared Hallet '14. I would not have been able to write this thesis without relying on the many papers on random matrix theory completed by Professor Miller's previous students.

I welcome any comments on this thesis, and I am responsible for all errors.

Contact: ks8@williams.edu or swansonkirk80@gmail.com

CONTENTS

1. Introduction

This section contains the history and background of the field of random matrix theory as well as important technical definitions and methods.

1.1. **History and Background.**

Although random matrices were first used by John Wishart as a tool in the study of population statistics in the 1920s [Wis], the field really flourished in the 1950s when Eugene Wigner conjectured that random matrices could be used to approximate the spacing between adjacent energy levels in heavy nuclei [Wig1, Wig2, Wig3, Wig4, Wig5]. His work was supported several years later by contributions from Freeman Dyson [Dy1, Dy2], and then in the 1970s, Hugh Montgomery discovered that random matrices could also predict answers to problems in number theory, including the distribution of the zeroes of the Riemann zeta function [Mon]. Since then, random matrix theory has had significant application not only to nuclear physics and number theory, but also to engineering, data analysis, multivariate statistics, operator algebra, wireless communications, dynamical systems, finance, and diffusion processes, as described in Bose [B] and Firk and Miller [FM] and exemplified in Miller, Novikoff, and Sabelli [MNS], Baik, Borodin, Deift, and Suidan [BBDS], and Krbalek and Seba [KrSe].

The original physics problem that motivated Wigner was how to describe the energy levels of large atoms, as explained in Firk and Miller. In quantum mechanics, it is well known that particles or systems can occupy different energy levels. The nature of the energy levels for hydrogen, an atom with just one electron and one proton, is completely understood through the Schrödinger equation. However, complicated atoms with more than two subatomic particles are not fully understood; uranium, for example, has over two hundred protons and neutrons in the nucleus. In operator form, the Schrödinger equation is described by an energy operator H, a wavefunction Ψ, and energy levels E:

$$H\Psi = E\Psi. \tag{1.1}$$

In this formulation of quantum mechanics, the energy operator takes the form of an infinite-dimensional matrix so that the wavefunctions can be thought of as eigenfunctions and the energy levels can be thought of as eigenvalues. Understanding the energy levels of an atom is then equivalent to understanding the eigenvalues of a matrix operator. Wigner's fascinating discovery was that the energy operator H can be modeled by a sequence of matrices. He considered collections of $N \times N$ matrices in which the entries were independently chosen from a fixed probability distribution p. By taking an average over all such random matrices in the limit as $N \to \infty$, the eigenvalues in this averaging and limiting process become excellent predictors for the energy levels of heavy nuclei. The image in Figure 1 is

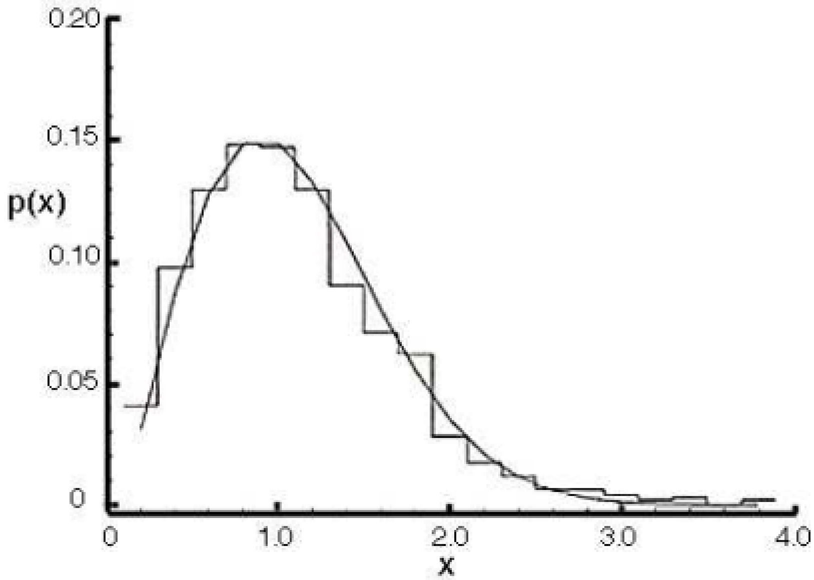

FIGURE 1. A Wigner distribution of adjacent eigenvalues fitted to the spacing distribution of 932 s-wave resonances in the interaction ^{238}U + n at energies up to 20 keV.

of sample data fitted to a random matrix eigenvalue distribution, taken from Firk and Miller with permission.

Discoveries such as this launched random matrix theory as a field of mathematical research. The research typically investigates the properties of random matrix eigenvalues as the dimensions of these matrices become very large. The main attributes studied are the spacing distribution between adjacent eigenvalues, the limiting spectral distribution, and the spectral width, or range from minimum to maximum eigenvalue.[1] In this paper, we investigate the limiting spectral distribution.

The limiting spectral distributions of several canonical matrix ensembles, such as the symmetric Wigner, Toeplitz, and Hankel matrices, have been well studied. The Wigner matrices are simply real symmetric matrices:

[1]"Eigenvalue distribution" is synonymous with "spectral distribution", and often the words "distribution" and "measure" are used interchangeably.

$$W_N = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1N} \\ a_{12} & a_{22} & a_{23} & \dots & a_{2N} \\ a_{13} & a_{23} & a_{33} & \dots & a_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1N} & a_{2N} & a_{3N} & \dots & a_{NN} \end{pmatrix}. \tag{1.2}$$

The entries in the upper triangle are all independent, identically distributed random variables, while those in the lower triangle are fixed by the symmetry constraint. It has been proved that Wigner matrices have a semicircular limiting spectral distribution for the normalized eigenvalues [Wig6], given by

$$f_{\text{Wigner}}(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2} & |x| \leq 2 \\ \\ 0 & \text{otherwise} \end{cases}. \tag{1.3}$$

The Wigner matrices have $\frac{N(N+1)}{2}$ independent parameters. For matrix ensembles with fewer independent parameters, or "degrees of freedom," different limiting spectral distributions other than the semicircle can arise. The real symmetric Toeplitz and Hankel matrices, for example, each have N degrees of freedom, and they do not have semicircular limiting spectral distributions. These matrices have appeared in time-series analysis and combinatorics. The Toeplitz matrices are fixed along diagonals

$$T_N = \begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{N-1} \\ a_1 & a_0 & a_1 & \dots & a_{N-2} \\ a_2 & a_1 & a_0 & \dots & a_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N-1} & a_{N-2} & a_{N-3} & \dots & a_0 \end{pmatrix} \tag{1.4}$$

while the Hankel matrices are fixed along skew diagonals

$$H_N = \begin{pmatrix} a_2 & a_3 & a_4 & \dots & a_{N+1} \\ a_3 & a_4 & a_5 & \dots & a_{N+2} \\ a_4 & a_5 & a_6 & \dots & a_{N+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N+1} & a_{N+2} & a_{N+3} & \dots & a_{2N} \end{pmatrix}. \tag{1.5}$$

Although it is not fully understood why, these ensembles exhibit the intriguing property that as more constraints are introduced to a patterned random matrix, new limiting measures other than the semicircle can arise. It is known that the Toeplitz matrices have a near-Gaussian distribution, as shown in Bose, Chatterjee, and Gangopadhyay [BCG], Bryc, Dembo, and Jiang [BDJ], and Hammond and Miller. It is also known that the Hankel matrices have a non-unimodal distribution, as shown by Bryc, Dembo, and Jiang. It is natural, then, to explore the following questions: To what extent will a patterned random matrix continue to have a semicircular limiting eigenvalue distribution? Which classes, or ensembles, of matrices will have a non-semicircular distribution?

We explore these questions by generalizing the Toeplitz and Hankel ensembles. These matrices have already been generalized in several ways. For example, it has been proved by Massey, Miller, and Sinsheimer [MMS] that Toeplitz matrices with palindromic rows have a Gaussian limiting spectral distribution, and it has been demonstrated by Jackson, Miller, and Pham [JMP] that Toeplitz matrices whose rows contain more than one palindrome have a limiting spectral distribution with very fat tails. In our generalization, the Toeplitz matrices can be thought of as having fixed, or constant, entries along lines of slope -1, and the Hankel matrices can be thought of as having fixed entries along lines of slope 1. What if we changed that slope to $\frac{1}{2}$, or some other slope? What if we required that all entries lying on the same parabola were constant, or along the same curve of some other polynomial? A 5×5 matrix with slope -$\frac{1}{2}$ might be like this:

$$A_5 = \begin{pmatrix} a_3 & a_4 & a_5 & a_6 & a_7 \\ a_4 & a_6 & a_7 & a_8 & a_9 \\ a_5 & a_7 & a_9 & a_{10} & a_{11} \\ a_6 & a_8 & a_{10} & a_{12} & a_{13} \\ a_7 & a_9 & a_{11} & a_{13} & a_{15} \end{pmatrix}. \tag{1.6}$$

In this paper, we explore matrices in which the pattern structure is described by some polynomial. In specific cases, we establish that when the variables in the polynomial are raised to the same power, the limiting measure is non-semicircular, but when the variables are raised to different powers, the limiting measure is semicircular.

1.2. **Definitions and Methodology.**

1.2.1. *Random Matrices and Link Functions.* A *random matrix* A_N is an $N \times N$ matrix whose entries are random variables drawn from a fixed probability distribution $p(x)$. The notion of randomness for a particular matrix depends on the distribution from which its entries are drawn. For our purposes, we will be considering probability distributions with all moments finite[2]:

(1) $p(x) \geq 0$ for all x
(2) $\int_{-\infty}^{\infty} p(x)dx = 1$
(3) $\int_{-\infty}^{\infty} |x|^k p(x)dx < \infty$ for all $k \geq 0$.

Since any such probability distribution can be scaled to have mean 0 and variance 1, we will also assume that for any random variable X with distribution $p(x)$,

$$\mathbb{E}[X] = 0 \text{ and } \mathbb{E}[X^2] = 1. \tag{1.7}$$

A particular random matrix is constructed from a sequence of independent, identically distributed random variables with distribution $p(x)$, called the *input sequence*: $\{a_i : i \in \mathbb{Z}\}$ or $\{a_{ij} : i, j \in \mathbb{Z}\}$. The way in which the input sequence gives a pattern to a random matrix is dictated by the *link function*, $L(i, j)$, as it tells us which entries (i, j) are constructed from which random variables. [3] It is therefore a function that maps the entries of a matrix to the input sequence of random variables:

$$L(i, j) : \{1, 2, ..., N\}^2 \to \mathbb{Z} \text{ for all } 1 \leq i, j \leq N. \tag{1.8}$$

A random matrix constructed from an *input sequence* and *link function* is of the form

$$A_N = [[a_{L(i,j)}]]. \tag{1.9}$$

Since we will be working with symmetric matrices, we will also impose the symmetry condition $L(i, j) = L(j, i)$. The aforementioned Wigner, Toeplitz, and Hankel matrices are listed below with their particular link functions.

---

[2]For several convergence results, we will also assume that all the moments are uniformly bounded.

[3]We borrow the excellent notation in Bose, Hammond and Miller, and Xiong [X] in the remaining introductory sections.

Wigner:

$$L(i,j) = (\text{min}[i,j], \text{max}[i,j]) \tag{1.10}$$

$$W_N = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1N} \\ a_{12} & a_{22} & a_{23} & \dots & a_{2N} \\ a_{13} & a_{23} & a_{33} & \dots & a_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1N} & a_{2N} & a_{3N} & \dots & a_{NN} \end{pmatrix} \tag{1.11}$$

Toeplitz:

$$L(i,j) = |i-j| \tag{1.12}$$

$$T_N = \begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{N-1} \\ a_1 & a_0 & a_1 & \dots & a_{N-2} \\ a_2 & a_1 & a_0 & \dots & a_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N-1} & a_{N-2} & a_{N-3} & \dots & a_0 \end{pmatrix} \tag{1.13}$$

Hankel:

$$L(i,j) = i+j \tag{1.14}$$

$$H_N = \begin{pmatrix} a_2 & a_3 & a_4 & \dots & a_{N+1} \\ a_3 & a_4 & a_5 & \dots & a_{N+2} \\ a_4 & a_5 & a_6 & \dots & a_{N+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N+1} & a_{N+2} & a_{N+3} & \dots & a_{2N} \end{pmatrix}. \tag{1.15}$$

Since these matrices are governed by random variables, we can express the probability of observing a particular matrix, or more exactly, the probability that the entry (i,j) lies in the interval $[\alpha_{ij}, \beta_{ij}]$ for a matrix A_N contained in the *outcome space* Ω_N:

$$\text{Prob}\left(A_N \in \Omega_N : a_{L(i,j)} \in [\alpha_{ij}, \beta_{ij}]\right) = \Pi_{1 \le i \le j \le N} \int_{\alpha_{ij}}^{\beta_{ij}} p(x)dx. \tag{1.16}$$

We can also define a probability measure for our matrices, which holds the information about its eigenvalues. For real symmetric matrices of size N, there are N real eigenvalues (including multiplicity) that can be ordered as $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N$. The probability measure for a matrix of size N is called the *empirical spectral measure* for normalized eigenvalues. It is denoted by

$$\mu_{A_N}(x)dx = \frac{1}{N}\sum_{i=1}^{N}\delta\left(x - \frac{\lambda_i(A_N)}{\sqrt{N}}\right)dx, \tag{1.17}$$

where $\delta(x)$ is the Dirac-delta functional.[4] This measure is "empirical" because it depends on the unknown eigenvalues of the matrix. Integrating the empirical measure gives the fraction of eigenvalues less than or equal to x. We can use this to define the *empirical spectral distribution*, a cumulative distribution function, for a matrix of size N:

$$F_{A_N}(x) = \int_{-\infty}^{x}\mu_{A_N}(x)dx = \frac{\#\left\{i \le N : \frac{\lambda_i}{\sqrt{N}} \le x\right\}}{N}. \tag{1.18}$$

Since the entries of any matrix A_N are random, the empirical measure and spectral distribution are both random.

1.2.2. *The Method of Moments.* We would like to understand the distribution of eigenvalues in the limit as the size of the matrices grows to infinity. In order to do so, it is important to connect the behavior of the empirical distributions to the moments that characterize them, because it will be the moments that we are able to compute by hand. The critical connection is established in the following moment convergence theorem from Bose and Hammond and Miller.

Theorem 1.1 (The Method of Moments.) *Let $\{A_N\}_{N=1}^{\infty}$ be a sequence of random variables and $\{F_N\}_{N=1}^{\infty}$ be the corresponding sequence of cumulative distribution functions such that their moments, $M_k(N) = \int_{-\infty}^{\infty} x^k dF_N(x)$, exist for all k. Let $\{M_k\}_{k=1}^{\infty}$ be a sequence of moments that uniquely determine a probability distribution whose cumulative distribution function is denoted by F. If $\lim_{N\to\infty} M_k(N) = M_k$ for each $k \ge 1$, then the sequence of cumulative distribution functions for the random variables converges weakly to the limiting distribution: $\lim_{N\to\infty} F_N = F$.*[5]

We will use the moments of the empirical spectral distributions to investigate the limiting spectral distribution. The k^{th} moment for the empirical spectral distribution of a random

[4]See section 7.1 for a heuristic on the size of the normalization factor.

[5]See section 7.2 for a description of Riesz's condition, which determines whether a sequence of moments uniquely determines a probability distribution.

matrix A of size N is given by:

$$M_k(A_N) = \int_{-\infty}^{\infty} x^k \mu_{A_N}(x) dx \tag{1.19}$$

$$= \int_{-\infty}^{\infty} x^k \frac{1}{N} \sum_{i=1}^{N} \delta\left(x - \frac{\lambda_i(A_N)}{\sqrt{N}}\right) dx = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\lambda_i(A_N)}{\sqrt{N}}\right)^k = \frac{1}{N^{\frac{k}{2}+1}} \sum_{i=1}^{N} \lambda_i^k(A_N). \tag{1.20}$$

Our goal is to explore the behavior of the limiting spectral distribution for a typical sequence of random matrices. So instead of considering a particular sequence of matrices and their moments, we compute the average moment values over all such matrices where, for a given N, the average moment is computed by weighting each matrix of size N by the probability of observing that matrix (Eq. 1.16). The average k^{th} moment for matrices of size N is given by

$$M_k(N) = \mathbb{E}\left[M_k(A_N)\right]. \tag{1.21}$$

The moments of the typical, or expected, limiting spectral distribution are given by

$$M_k = \lim_{N\to\infty} M_k(N). \tag{1.22}$$

With these tools, we can formally state the Moment Convergence Theorem for Random Matrices, which follows directly from the Method of Moments.

Theorem 1.2 (Moment Convergence Theorem for Random Matrices.) *Suppose* $\{A_N\}_{N=1}^{\infty}$ *is an arbitrary sequence of random matrices with distributions* $\{F_{A_N}\}_{N=1}^{\infty}$*. Suppose there exists some sequence of moments* $\{M_k\}_{k=1}^{\infty}$ *such that they uniquely determine a probability distribution whose cumulative distribution function is denoted by F. If* $\lim_{N\to\infty} M_k(N) = M_k$ and $\lim_{N\to\infty} \mathrm{Var}[M_k(A_N)] = 0$ *for every positive integer k, then the sequence* $\{F_{A_N}\}_{N=1}^{\infty}$ *converges in probability to the limiting spectral distribution of the ensemble,* F.

This theorem contains the same ideas in the Method of Moments, except instead of a sequence of fixed moments, it considers a sequence of average moments over the ensemble with diminishing variance. This condition ensures that the limiting moments and limiting distribution will hold for most sequences of random matrices that one might construct.

1.2.3. *Eigenvalue-Trace Lemma and Circuits.* It would be impossible to compute the moments of an empirical spectral distribution directly from the eigenvalues, since we do not yet know the eigenvalues. Instead, we use the Eigenvalue-Trace Lemma to rewrite the moments in terms of the matrix trace, which is what we do know.

Lemma 1.1 (Eigenvalue-Trace Lemma). *For a square matrix of size N denoted by A_N and with eigenvalues $\lambda_i(A_N)$,*

$$\text{Trace}(A_N^k) = \sum_{i=1}^{N} \lambda_i^k(A_N). \tag{1.23}$$

Using the Eigenvalue-Trace Lemma, the empirical moments are written as

$$M_k(A_N) = \frac{1}{N^{\frac{k}{2}+1}} \sum_{i=1}^{N} \lambda_i^k(A_N) = \frac{1}{N^{\frac{k}{2}+1}} \text{Tr}(A_N^k). \tag{1.24}$$

Expanding $\text{Tr}(A_N^k)$, we have

$$M_k(A_N) = \frac{1}{N^{\frac{k}{2}+1}} \sum_{1 \le i_1, \ldots, i_k \le N} a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_k i_1}, \tag{1.25}$$

where $a_{i_1 i_2}$ denotes the value of the entry of A_N with indices (i_1, i_2). Using linearity of expectation, we can write the k^{th} expected moment as

$$M_k(N) = \frac{1}{N^{\frac{k}{2}+1}} \sum_{1 \le i_1, \ldots, i_k \le N} \mathbb{E}\left[a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_k i_1}\right]. \tag{1.26}$$

Using the link function, we can write this expansion in terms of the input sequence elements[6]:

$$M_k(N) = \frac{1}{N^{\frac{k}{2}+1}} \sum_{1 \le i_1, \ldots, i_k \le N} \mathbb{E}[a_{L(i_1,i_2)} a_{L(i_2,i_3)} \cdots a_{L(i_k,i_1)}]. \tag{1.27}$$

The above sum is taken over all combinations of positive integers $\{i_1, ..., i_k\}$ less than N. Each distinct combination of index values is a *circuit*. We will therefore define a circuit π as a function from the entry indices to their integer values:[7]

$$\pi : \{0, 1, 2, ..., k\} \to \{1, 2, ..., N\} \text{ such that } \pi(0) = \pi(k). \tag{1.28}$$

The k^{th} average moment is then written succinctly as

$$M_k(N) = \frac{1}{N^{\frac{k}{2}+1}} \sum_{\pi : \pi \text{circuit}} \mathbb{E}[X_\pi], \tag{1.29}$$

where

$$X_\pi = a_{L(\pi(0),\pi(1))} a_{L(\pi(1),\pi(2))} \cdots a_{L(\pi(k-1),\pi(k))}. \tag{1.30}$$

[6]Although technically we might say that a_{ij} refers to a matrix entry and $a_{L(i,j)}$ refers to the input sequence variable whose value describes that entry, we use both notations interchangeably.

[7]Here the index values have been shifted down by one, so that the first index i_1 is mapped by $\pi(0)$, and the last index i_1 is relabeled $\pi(k)$ under the constraint that $\pi(0) = \pi(k)$. This seemingly confusing switch of notation will actually help us in later proofs.

We call an input variable index $L(\pi(i-1), \pi(i))$ an *L-value*. If an L-value is repeated exactly *e* times in a circuit, then the circuit has an edge of order *e*. If a circuit has all edges $e \geq 2$, then the circuit is *L-matched*. Circuits that are not L-matched do not contribute to the moment, since the expected value of a product of independent random variables is the product of the expected values and since the distribution $p(x)$ is assumed to have mean zero.

It is possible for two different circuits to yield the same set of L-values; these are *equivalent circuits*. Two circuits, π_1 and π_2, are equivalent if and only if their L-values match at the same locations for all $1 \leq i, j \leq N$:

$$L(\pi_1(i-1), \pi_1(i)) = L(\pi_1(j-1), \pi_1(j)) \Longleftrightarrow L(\pi_2(i-1), \pi_2(i)) = L(\pi_2(j-1), \pi_2(j)). \tag{1.31}$$

An equivalence class of circuits is a partition of the set $\{1, 2, ..., k\}$. We can label an equivalence class by a *word* of length k, where the first occurrence of each letter in the word is in alphabetical order. If we let $w[i]$ denote the i^{th} entry of word w, the equivalence class of circuits corresponding to w will be given by

$$\Pi(w) = \{\pi : w[i] = w[j] \Longleftrightarrow L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j))\}. \tag{1.32}$$

For example, if $k = 6$, the partition $\{\{1, 3, 5, 6\}, \{2, 4\}\}$ is represented by the word $ababaa$. This identifies all circuits π for which $L(\pi(0), \pi(1)) = L(\pi(2), \pi(3)) = L(\pi(4), \pi(5)) = L(\pi(5), \pi(6))$ and $L(\pi(1), \pi(2)) = L(\pi(3), \pi(4))$.

The *size* of w, or number of distinct letters, is denoted by $|w|$:

$$|w| = \#\{L(\pi(i-1), \pi(i)) : 1 \leq i \leq k\}. \tag{1.33}$$

The positions of the letters in a word i, for $1 \leq i \leq k$, along with the additional value $i = 0$, are called *vertices*.[8] A vertex is *generating* if either $i = 0$ or $w[i]$ is the first occurrence of a letter in the word. Otherwise, the vertex is *non-generating*. For example, if $w = ababcb$, then the generating vertices are $\{0, 1, 2, 5\}$ and the nongenerating vertices are $\{3, 4\}$.

The number of generating vertices in an L-matched word is equivalent to the maximum number of degrees of freedom one has in choosing a circuit that corresponds to that word, because once the generating vertices are chosen, the non-generating vertices are fixed by the fact that they have to satisfy matched L-values. For example, consider the word $abab$ for the Toeplitz link function $L(i, j) = |i - j|$. The word dictates the following system of

[8]"Vertex" and "index" are basically synonymous in this paper.

equations:

$$|\pi(0) - \pi(1)| = |\pi(2) - \pi(3)| \text{ and } |\pi(1) - \pi(2)| = |\pi(3) - \pi(4)|. \tag{1.34}$$

We can choose $\pi(0),\ \pi(1),$ and $\pi(2)$ freely, but then $\pi(3)$ is fixed by the matching constraints and $\pi(4)$ is defined to be equal to $\pi(0)$.

Since there are at most N choices for each generating vertex, and since there are $|w|+1$ generating vertices, the size of the equivalence class for word w is at most

$$\#\Pi(w) = O(N^{|w|+1}). \tag{1.35}$$

In fact, we can rewrite the average k^{th} moment using words and their equivalence classes:

$$M_k(N) = \sum_{w:w \text{ is L-matched and of length } k} \frac{1}{N^{\frac{k}{2}+1}} \sum_{\pi:\pi\in\Pi(w)} \mathbb{E}[X_\pi]. \tag{1.36}$$

We are interested in computing M_k to determine the limiting spectral distribution, but not all circuits will contribute in the limit as $N \to \infty$. We have already seen that only L-matched words contribute to the moment. Moreover, it turns out that if a matrix ensemble satisfies certain properties, then only *pair-matched* words contribute, words in which every letter appears exactly twice. Bose calls the sufficient property for this condition Property B:

$$\Delta(L) = \sup_N \sup_{t\in\mathbb{Z}^+} \sup_{1\le k\le N} \#\{m : 1 \le m \le n, L(k,m) = t\} < \infty. \tag{1.37}$$

For a matrix satisfying Property B, its $\Delta(L)$ value, the maximum number of repetitions of the same random variable in any row or column, is finite. For the Wigner, Toeplitz, and Hankel matrices, their $\Delta(L)$ values are 1, 2, and 2, respectively. The following lemma is from Bose, the excellent proof of which we follow closely:

Lemma 1.2: (Pair-Matched Words.) *Only pair-matched words contribute and odd moments are zero.*

Proof. Let w be a word with at least one edge of order greater than or equal to three. Since contributing words must be L-matched, there are at most $k+1$ generating vertices. For a word of length $2k$, constructing one edge of order three requires eliminating one generating vertex, leaving k generating vertices. For a word of length $2k+1$, because there is an odd number of vertices, there must already be at least one edge of order three, which again allows for at most k generating vertices. So there are k degrees of freedom from the generating vertices in both cases. For each non-generating vertex, there are at most $\Delta(L)$ choices. Once all the generating vertices are chosen, each non-generating vertex must satisfy a particular L-value in a particular row or column of the matrix, and by Property B, there are at most $\Delta(L)$ choices for a particular random variable, or L-value, in any row or column. Since there are at most k non-generating vertices, we then have

$$\#\Pi(w) \leq \Delta(L)^k N^k = O(N^k). \tag{1.38}$$

From (Eq. 1.36), there must be $k+1$ degrees of freedom for the contribution to M_k to be nonzero. Terms of order $O(N^k)$ will not contribute. Since odd moments sum over words with at least one edge greater than or equal to three, odd moments are zero, and it suffices to check even moments. For the even moments, only words that are pair-matched contribute. □

Since only even moments with pair-matched words of length $2k$ contribute, the moments of the limiting spectral distribution can finally be written as

$$M_{2k} = \lim_{N\to\infty} M_{2k}(N) = \sum_{w:w \text{ is pair-matched of length 2k}} \lim_{N\to\infty} \frac{1}{N^{k+1}} \#\Pi(w). \tag{1.39}$$

In other words, computing the limiting moments reduces to checking all possible pair-matched words, and for each word, finding the number circuits corresponding to that word. Counting the number of circuits for a given word becomes equivalent to counting the number of integer solutions to a set of Diophantine, or integer-valued, equations.[9]

If the moments M_{2k} can be computed or shown to exist, it is possible to prove convergence to a limiting spectral distribution via Theorem 1.2. The following sections, essentially paraphrasing Bose and Hammond and Miller, describe general proofs of convergence for broad classes of random matrix ensembles. These convergence results apply to all of the matrices considered in this paper.

1.2.4. *Existence and Uniqueness of the Limiting Spectral Distribution.* Bose proves that if a random matrix ensemble has a link function that satisfies Property B and the limiting moments exist, then the limiting spectral distribution exists. Moreover, the limiting spectral distribution is uniquely specified by its moments. We sketch the proof given by Bose. By Property B, odd moments are zero. The average $2k^{th}$ moment is given by

$$M_{2k} = \sum_{w:w \text{ is pair-matched of length 2k}} \lim_{N\to\infty} \frac{1}{N^{k+1}} \#\Pi(w). \tag{1.40}$$

It can be shown that there are $(2k-1)!!$ ways of grouping $2k$ objects in pairs, so for words of length $2k$, the maximum number of words w we sum over is $(2k-1)!! = \frac{(2k)!}{2^k k!}$. For each word, there are at most $k+1$ degrees of freedom for the generating vertices,

[9]In this paper we only consider matrix ensembles that satisfy Property B.

leaving $\Delta(L)^k$ choices for the non-generating vertices so that each word contributes at most $\Delta(L)^k N^{k+1}$. We then have that $M_{2k}(N) \leq \frac{(2k)!}{2^k k!}\Delta(L)^k + O\left(\frac{1}{N}\right)$. Also by Property B under the assumption of the existence of the limiting moments, the expected moments converge to the limiting moments almost surely, so that $M_{2k} \leq \frac{(2k)!}{2^k k!}\Delta(L)^k$. These moments satisfy Riesz's condition.[10] By Theorem 1.2, the limiting spectral distribution of the ensemble exists and is uniquely determined.

1.2.5. *Convergence in Probability.* Assume that all moments M_k exist, are finite, and uniquely determine a probability distribution. By Theorem 1.2, it suffices to show that Var$[M_k(A_N)] \to 0$ to prove convergence in probability. Here is a sketch of the proof from Hammond and Miller. Although it was specifically designed for real symmetric Toeplitz matrices, it is general enough to apply to any matrix ensemble considered in this paper.

The empirical spectral distributions converge in probability to the limiting spectral distribution if the empirical moments converge in probability to the limiting moments. The moments converge in probability if $\forall \epsilon > 0$,

$$\lim_{N\to\infty} \text{Prob}\left(\{A_N \in \Omega_N : |M_{2k}(A_N) - M_{2k}| > \epsilon\}\right) = 0. \tag{1.41}$$

By the triangle inequality,

$$|M_{2k}(A_N) - M_{2k}| \leq |M_{2k}(A_N) - M_{2k}(N)| + |M_{2k}(N) - M_{2k}|. \tag{1.42}$$

By Chebyshev's inequality,

$$\begin{aligned} \text{Prob}\left(\{A_N \in \Omega_N : |M_{2k}(A_N) - M_{2k}(N)| > \epsilon\}\right) &\leq \frac{\text{Var}[M_{2k}(A_N)]}{\epsilon^2} \\ &\leq \frac{\mathbb{E}[M_{2k}(A_N)^2] - \mathbb{E}[M_{2k}(A_N)]^2}{\epsilon^2}. \end{aligned} \tag{1.43}$$

Since all higher moments exist and are finite by assumption, $|M_{2k}(N) - M_{2k}| \to 0$ as $N \to \infty$. It suffices to show that for all $2k$,

$$\lim_{N\to\infty}\left(\mathbb{E}\left[M_{2k}(A_N)^2\right] - \mathbb{E}\left[M_{2k}(A_N)\right]^2\right) = 0, \tag{1.44}$$

By (Eq. 1.26), we have

$$\mathbb{E}\left[M_{2k}(A_N)^2\right] = \frac{1}{N^{2k+2}} \sum_{1\leq i_1,\dots,i_{2k}\leq N} \sum_{1\leq j_1,\dots,j_{2k}\leq N} \mathbb{E}\left[a_{i_1 i_2}\cdots a_{i_{2k} i_1} a_{j_1 j_2}\cdots a_{j_{2k} j_1}\right], \tag{1.45}$$

[10]See section 7.2.

$$\mathbb{E}\left[M_{2k}\left(A_N\right)\right]^2 = \frac{1}{N^{2k+2}} \sum_{1\le i_1,\dots,i_{2k}\le N} \mathbb{E}\left[a_{i_1 i_2}\cdots a_{i_{2k} i_1}\right] \sum_{1\le j_1,\dots,j_{2k}\le N} \mathbb{E}\left[a_{j_1 j_2}\cdots a_{j_{2k} j_1}\right]. \tag{1.46}$$

There are two possibilities for the contribution from the i configurations, $a_{i_1 i_2}\cdots a_{i_{2k} i_1}$, and the j configurations, $a_{j_1 j_2}\cdots a_{j_{2k} j_1}$,. If in an i configuration, any entry $a_{i_s i_{s+1}}$ is not equal to any entry $a_{j_t j_{t+1}}$ in a j configuration, then together, the these two configurations contribute equally to $\mathbb{E}\left[M_{2k}\left(A_N\right)^2\right]$ and $\mathbb{E}\left[M_{2k}\left(A_N\right)\right]^2$. It suffices to estimate the difference for the crossover cases, where we have at least one pair of entries from the i and j configurations matched, $a_{i_s i_{s+1}} = a_{j_t j_{t+1}}$. These cases contribute unequally to the two expected values above. We adopt the standard method of counting degrees of freedom in Hammond and Miller and show that the contribution from crossover cases is $O_{2k}\left(\frac{1}{N}\right)$ to both $\mathbb{E}\left[M_{2k}\left(A_N\right)^2\right]$ and $\mathbb{E}\left[M_{2k}\left(A_N\right)\right]^2$. Essentially, it can be shown that one crossover is associated with at least one loss of degrees of freedom. As in the full proof in Hammond and Miller, we will show that for all of the matrix ensembles in this paper, only entries paired in opposite triangles of a matrix can be matched. For our matrix ensembles, all remaining steps of the proof follow trivially except changes in the constants $O_k\left(\frac{1}{N}\right)$, which do not alter the result that the contribution from crossover cases diminishes as $N\to\infty$.

1.2.6. *Almost Sure Convergence.* Assume that all moments M_k exist, are finite, and uniquely determine a probability distribution. Then, if the empirical moments converge almost surely to the limiting moments, the empirical spectral distributions converge almost surely to the limiting spectral distribution. We appeal to the excellent proofs of almost sure convergence in Bose and Hammond and Miller. The arguments in Hammond and Miller were designed for Toeplitz matrices, but they can be generalized easily. The arguments in Bose apply to any ensemble of matrices whose link function satisfies Property B. To prove almost sure convergence of moments, it suffices to show that

$$\sum_{n=1}^{\infty} \mathbb{E}\left[M_k\left(A_N\right) - \mathbb{E}\left[M_k\left(A_N\right)\right]\right]^4 < \infty \text{ for every } k\ge 1. \tag{1.47}$$

If we assume that the input distribution $p(x)$ has mean zero, variance one, and uniformly bounded moments of all order, it is possible to show

$$\mathbb{E}\left[\frac{1}{N} Tr\left(A_N\right)^k - \mathbb{E}\left[\frac{1}{N} Tr\left(A_N\right)^k\right]\right]^4 = O\left(N^{-2}\right) \tag{1.48}$$

Almost sure convergence then follows.

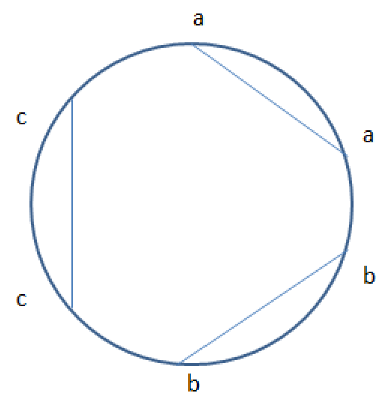

FIGURE 2. *aabbcc* is a Catalan word.

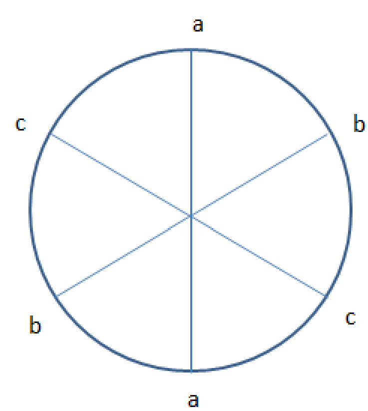

FIGURE 3. *abcabc* is not a Catalan word.

1.2.7. *Non-Crossing Pair Partitions, Catalan Words, and the Semicircle Measure.* Pair-matched words of length $2k$ can be classified as *non-crossing partitions* or *crossing partitions*. Consider the set $\{1, 2, ..., 2k\}$. Arrange the elements on a circle sequentially. Consider any pair partition of this set and draw an edge between two points of each partition. The partition is said to be non-crossing if none of the edges crosses another, and crossing otherwise.

It can be shown that non-crossing partitions are in bijection with Catalan words. A pair-matched word is called a *Catalan word* if (1) there is at least one double letter, (2) if any double letter is deleted, the remaining word of length $2k-2$ is either empty or has a double letter, and (3) repeating the process in the previous step ultimately leads to an empty word. For example, $aabbcc$ is a Catalan word, while $abcabc$ is not a Catalan word, as shown in Figures 2 and 3.

The number of Catalan words of length $2k$ is given by the $2k^{th}$ *Catalan number*

$$C_{2k} := \frac{1}{k+1}\binom{2k}{k}. \tag{1.49}$$

It can also be shown that the $2k^{th}$ moment of the semicircle measure is exactly C_{2k}, with odd moments zero.[11] This is critical, because then one can prove that an ensemble of matrices has a semicircular limiting spectral distribution by showing that all Catalan words contribute one to the $2k^{th}$ moment and all non-Catalan words contribute zero.

2. Polynomials of Order One

In this section we will generalize the real symmetric Toeplitz and Hankel link functions to a class of bivariate polynomial link functions in which the polynomials are order one. This lets us investigate all matrices in which random variables are fixed along lines with rational slopes.

2.1. **Generalized Toeplitz Matrices.** Real symmetric Toeplitz matrices are matrices that are constant along the diagonals. In other words, the entries in the upper triangle of the matrix are described by the same random variable if they lie along the same line of slope -1, as though the columns j were the x-axis and the rows i were the y-axis. What if we change the slope of these lines to $-\frac{1}{2}$, or $-\frac{3}{2}$, or some other negative rational number?

The link function for Toeplitz matrices was written as $L(i,j) = |i-j|$, but we can also write the link function as:

$$L(i,j) = \begin{cases} i-j & i \leq j \\ \\ -i+j & i > j. \end{cases} \tag{2.1}$$

This splits any matrix into two zones, where Zone 1 describes the upper triangle $i \leq j$, and Zone 2 describes the lower triangle $i > j$, as in Figure 4.[12] For example, if two matched entries $a_{i_1 i_2} = a_{i_3 i_4}$ are such that $a_{i_1 i_2},\ a_{i_3 i_4} \in$ Zone 1, their L-values must satisfy

$$\pi(0) - \pi(1) = \pi(2) - \pi(3), \tag{2.2}$$

[11]See section 7.4.

[12]We could have defined the zones to exclude the main diagonal, as the values of the main diagonal do not affect the limiting distribution of the eigenvalues. Likewise, we could have set the main diagonal to be zero by definition.

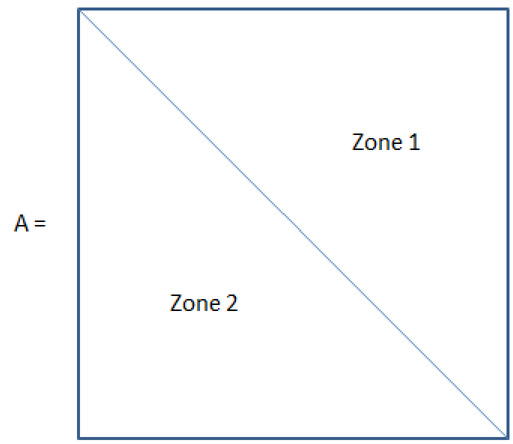

FIGURE 4. A matrix A is split into two zones, where Zone 1 is the upper triangle of the matrix including the main diagonal, and Zone 2 is the lower triangle of the matrix.

while if one entry $a_{i_1 i_2} \in$ Zone 1 and the other entry $a_{i_3 i_4} \in$ Zone 2, their L-values must satisfy

$$\pi(0) - \pi(1) = -\pi(2) + \pi(3). \tag{2.3}$$

We can change the link function slope by introducing parameters α and β to define a *generalized Toeplitz link function*.[13] This is, for fixed $\alpha,\ \beta \in \mathbb{Q}^+$

$$L_{\alpha,\beta}(i,j) = \begin{cases} \alpha i - \beta j & i \leq j \\ \\ -\beta i + \alpha j & i > j. \end{cases} \tag{2.4}$$

A matrix with $\alpha = \beta$ reduces to the original Toeplitz, while a 5×5 matrix with $\alpha = 2$ and $\beta = 1$ would have the structure

$$A = \begin{pmatrix} a_1 & a_0 & a_{-1} & a_{-2} & a_{-3} \\ a_0 & a_2 & a_1 & a_0 & a_{-1} \\ a_{-1} & a_1 & a_3 & a_2 & a_1 \\ a_{-2} & a_0 & a_2 & a_4 & a_3 \\ a_{-3} & a_{-1} & a_1 & a_3 & a_5 \end{pmatrix}. \tag{2.5}$$

It is known that the moments of the limiting spectral distribution of large Toeplitz matrices are bounded above by the moments of the Gaussian distribution, where the $2k^{th}$

[13]We disregard irrational parameters, since this would simply give us Wigner matrices. Moreover, for other general link functions, we will suppress the subscripts α and β on $L_{\alpha,\beta}(i,j)$.

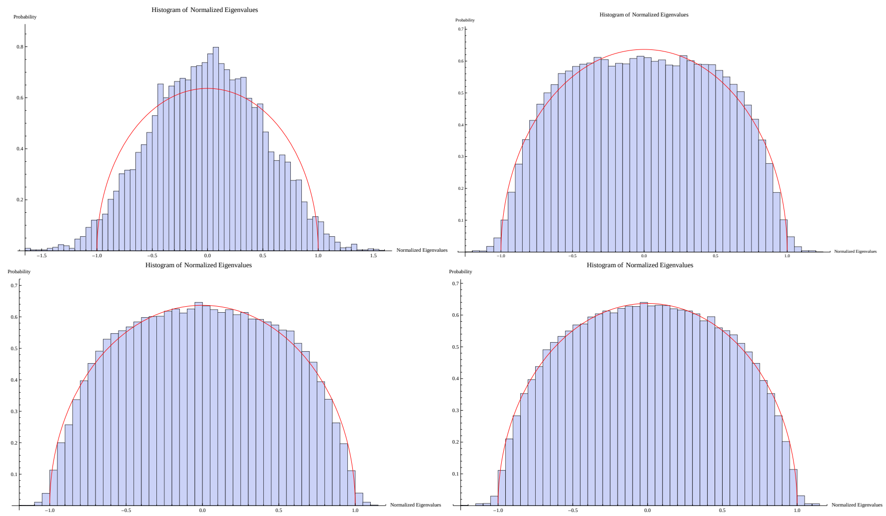

FIGURE 5. Histograms of numerical eigenvalues of 100 generalized Toeplitz matrices of size 1200×1200. Each has $\alpha = 1$. Clockwise from the upper left, β is equal to 1, 2, 3, and 4. The red curve is the semicircle distribution of Eq. 1.3 for an eigenvalue normalization of $2\sqrt{N}$. See section 7.1 for details on normalizations for the eigenvalues.

Gaussian moment is given by $(2k-1)!!$ The following table compares low moments for the Gaussian, Toeplitz, and semicircular distributions:

Moment	Gaussian	Toeplitz	Semicircle
M_4	3	$2\frac{2}{3}$	2
M_6	15	11	5
M_8	105	$64\frac{4}{16}$	14

(2.6)

Numerics suggest that the generalized Toeplitz matrices have a near-semicircular limiting spectral distribution.[14] Figure 5 shows simlutions of the limiting spectral distribution for generalized Toeplitz matrices for $\alpha = 1$ and several values of β. The first histogram clearly shows the near-Gaussian behavior for $\beta = 1$, which corresponds to original Toeplitz matrices. Although the other distributions appear semicircular, note the slight tails, ears, and dip in the histogram for $\beta = 2$ that show deviation from the semicircle, shown in red. As β increases, the histograms show that the limiting distribution more closely resembles the semicircle distribution. We can compute low moments for the generalized Toeplitz ensemble and show that they deviate from the Catalan numbers by a factor that depends on α and β. We can also prove that as either α or β tends to infinity, the limiting spectral distribution converges to the semicircle.

[14]Explanation of numerical calculations is provided in section 6.

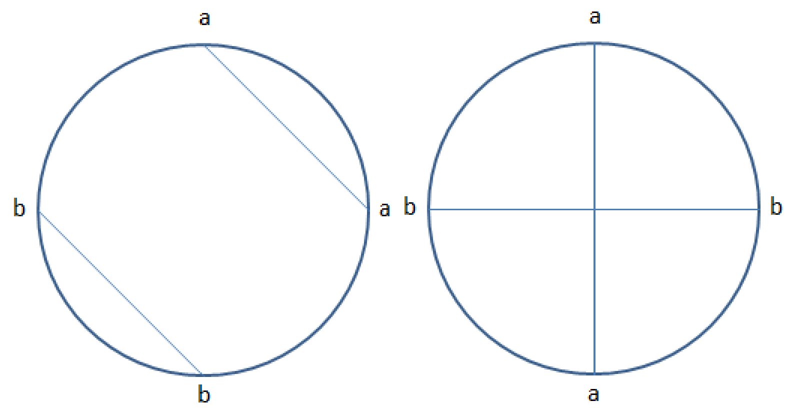

FIGURE 6. Partitions for the Catalan word *aabb* and non-Catalan word *abab*.

2.1.1. *Odd Moments.* Because the slopes of lines connecting random variables via this link function can never be zero, $\Delta(L)$ is at most one. Odd moments are then zero because these matrices satisfy Property B.

2.1.2. *Zeroth and Second Moments.* Calculating the zeroth and second moments is simple:

$$\begin{aligned} M_0 &= \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\left[\lambda_i^0(A_N)\right] = 1, \\ M_2 &= \lim_{N\to\infty} \frac{1}{N^2} \sum_{1\le i_1,\ i_2\le N} \mathbb{E}\left[a_{i_1 i_2} a_{i_2 i_1}\right] = \lim_{N\to\infty} \frac{1}{N^2} \sum_{1\le i_1,\ i_2\le N} \mathbb{E}\left[a^2_{L(i_1 i_2)}\right] = 1. \end{aligned} \tag{2.7}$$

The expected value above is 1, since we are drawing from a variance one distribution.

2.1.3. *Fourth Moment.* To calculate the fourth moment, we compute

$$\begin{aligned} M_4 &= \lim_{N\to\infty} \frac{1}{N^3} \sum_{1\le i_1,i_2,i_3,i_4\le N} \mathbb{E}\left[a_{i_1 i_2} a_{i_2 i_3} a_{i_3 i_4} a_{i_4 i_1}\right] \\ &= \sum_{w:w \text{ is pair-matched of length 4}} \lim_{N\to\infty} \frac{1}{N^3} \#\Pi(w). \end{aligned} \tag{2.8}$$

The pair-matched words of length four are *aabb*, *abba*, and *abab*, which give us the non-isomorphic configurations in Figure 6. By relabeling indices, it is easy to see that the matchings for *aabb* and *abba*, respectively, give equivalent systems of equations:

$$a_{L(i_1,i_2)} = a_{L(i_2,i_3)} \text{ and } a_{L(i_3,i_4)} = a_{L(i_4,i_1)} \tag{2.9}$$

$$a_{L(i_1,i_2)} = a_{L(i_4,i_1)} \text{ and } a_{L(i_2,i_3)} = a_{L(i_3,i_4)}. \tag{2.10}$$

The system of equations for the word *abab* is

$$a_{L(i_1,i_2)} = a_{L(i_3,i_4)} \text{ and } a_{L(i_2,i_3)} = a_{L(i_4,i_1)}. \tag{2.11}$$

Since any entry can be located in one of two zones, there are $2^4 = 16$ possible diophantine L-value equations for each of the above sets of entry matchings.

We first count the number of circuits for the word *aabb*, which will be equivalent to the number of cuircuits for *abba*. We define an *adjacent pair* as a matching between matrix entries that share one index, as in $a_{L(i_1,i_2)} = a_{L(i_2,i_3)}$. We reduce the number of relevant cases with the following.

Lemma 2.1: (Fourth Moment Adjacent Pairs.) *Fourth moment adjacent pairs must be in opposite zones when $\alpha \neq \beta$ to yield a nonzero contribution.*

Proof. Assume that for an adjacent pair, one entry $a_{i_1 i_2} \in$ Zone 1 and the other entry $a_{i_2 i_3} \in$ Zone 1. From the link function, we have

$$\alpha\pi(0) - \beta\pi(1) = \alpha\pi(1) - \beta\pi(2) \Longrightarrow \pi(1) = \frac{\alpha\pi(0) + \beta\pi(2)}{\alpha + \beta}. \tag{2.12}$$

Start by choosing values for the four variables $\pi(0),\ \pi(1),\ \pi(2),$ and $\pi(3)$[15]. Once $\pi(0)$ and $\pi(2)$ are chosen freely,[16] $\pi(1)$ is fixed as above and $\pi(3)$ is fixed by the second L-value equation. There are at most two degrees of freedom here, but there is only a contribution to the fourth moment if there are more than two degrees of freedom, since we divide by N^3.

Assume that for an adjacent pair, one entry $a_{i_1 i_2} \in$ Zone 2 and the other entry $a_{i_2 i_3} \in$ Zone 2. From the link function, we have

$$-\beta\pi(0) + \alpha\pi(1) = -\beta\pi(1) + \alpha\pi(2) \Longrightarrow \pi(1) = \frac{\alpha\pi(2) + \beta\pi(0)}{\alpha + \beta}. \tag{2.13}$$

By the same reasoning, this case does not contribute. Therefore, adjacent matchings must be in opposite zones. □

Since adjacent pairs must be in opposite zones, we calculate four possibilities:

(1) $a_{i_1 i_2} \in$ Zone 1, $a_{i_2 i_3} \in$ Zone 2, $a_{i_3 i_4} \in$ Zone 1, and $a_{i_4 i_1} \in$ Zone 2:
$\alpha\pi(0) - \beta\pi(1) = -\beta\pi(1) + \alpha\pi(2)$ and $\alpha\pi(2) - \beta\pi(3) = -\beta\pi(3) + \alpha\pi(0) \longrightarrow$
$\pi(0) = \pi(2),\ \pi(1) > \pi(0)$ and $\pi(3) > \pi(0)$

[15] It is not always necessary to distinguish which vertices are generating and which are non-generating in some counting arguments.

[16] A "free" index/vertex is one with N possible values.

(2) $a_{i_1i_2} \in$ Zone 1, $a_{i_2i_3} \in$ Zone 2, $a_{i_3i_4} \in$ Zone 2, and $a_{i_4i_1} \in$ Zone 1:
$\alpha\pi(0) - \beta\pi(1) = -\beta\pi(1) + \alpha\pi(2)$ and $-\beta\pi(2) + \alpha\pi(3) = \alpha\pi(3) - \beta\pi(0) \longrightarrow$
$\pi(0) = \pi(2),\ \pi(1) > \pi(0)$ and $\pi(3) < \pi(0)$

(3) $a_{i_1i_2} \in$ Zone 2, $a_{i_2i_3} \in$ Zone 1, $a_{i_3i_4} \in$ Zone 1, and $a_{i_4i_1} \in$ Zone 2:
$-\beta\pi(0) + \alpha\pi(1) = \alpha\pi(1) - \beta\pi(2)$ and $\alpha\pi(2) - \beta\pi(3) = -\beta\pi(3) + \alpha\pi(0) \longrightarrow$
$\pi(0) = \pi(2),\ \pi(1) < \pi(0)$ and $\pi(3) > \pi(0)$

(4) $a_{i_1i_2} \in$ Zone 2, $a_{i_2i_3} \in$ Zone 1, $a_{i_3i_4} \in$ Zone 2, and $a_{i_4i_1} \in$ Zone 1:
$-\beta\pi(0) + \alpha\pi(1) = \alpha\pi(1) - \beta\pi(2)$ and $-\beta\pi(2) + \alpha\pi(3) = \alpha\pi(3) - \beta\pi(0) \longrightarrow$
$\pi(0) = \pi(2),\ \pi(1) < \pi(0)$ and $\pi(3) < \pi(0)$.

(2.14)

Solutions to the four sets of inequalities above are valid as long as $\pi(0),\ \pi(1),$ and $\pi(3) \in \{1, 2, ..., N\}$. We define a new function that incorporates this restriction: $v_x = \frac{\pi(x)}{N},$ where for each $x \in \{0, 1, 3\},\ v_x \in \{\frac{1}{N}, \frac{2}{N}, ..., \frac{N}{N}\}$. Now, the four sets of inequalities are given by:

$$\begin{aligned} &v_1 > v_0 \text{ and } v_3 > v_0 \\ &v_1 > v_0 \text{ and } v_3 < v_0 \\ &v_1 < v_0 \text{ and } v_3 > v_0 \\ &v_1 < v_0 \text{ and } v_3 < v_0. \end{aligned} \tag{2.15}$$

Now we count the contribution. By transforming to v_x, we have already divided by N^3. In the limit of large N, then, a sum over the appropriate region gives us the moment contribution. For $\mathbb{I}(G)$, the indicator function on the region G, it is

$$\sum_{v_0, v_1, v_3 \in \{\frac{1}{N}, \frac{2}{N}, ..., \frac{N}{N}\}} \mathbb{I}(v_1 > v_0 \text{ and } v_3 > v_0, \text{ or } v_1 > v_0 \text{ and } v_3 < v_0, \text{ or } v_1 < v_0 \text{ and } v_3 > v_0,$$
$$\text{or } v_1 < v_0 \text{ and } v_3 < v_0). \tag{2.16}$$

In the limit of large N, this just becomes a triple integral:

$$\int_0^1 \int_0^1 \int_0^1 dv_0 dv_1 dv_3 = 1. \tag{2.17}$$

Together, *aabb* and *abba* contribute two to the fourth moment.

Now we count the contribution for the word *abab* and show that it only contributes to the moment when $\alpha = \beta$. Choose a zone for the first entry, say Zone 1. Then, the L-value equations have the form

$$\begin{aligned} \alpha\pi(0) - \beta\pi(1) &= a_{L(i_3,i_4)} \\ a_{L(i_2,i_3)} &=_{L(i_4,i_1)} . \end{aligned} \tag{2.18}$$

Assume that $\alpha \neq \beta$. We immediately see that in order to avoid introducing an extra linear constraint, and therefore a loss of degrees of freedom, we must choose $a_{i_2 i_3} \in$ Zone 2. If we were to sum the two L-value equations after choosing $a_{i_2 i_3} \in$ Zone 1, for example, we could derive an equation for $\pi(1)$ in terms of the other indices and apply the arguments from Lemma 2.1. Following similar arguments from Bose, there are at most $2k+1$ degrees of freedom to start with: $2k$ L-values and the value of the first index, $\pi(0)$. The L-value equations introduce k constraints, leaving us with at most $k+1$ degrees of freedom. Introducing another constraint leaves only k degrees of freedom, but more than k degrees of freedom are needed for a nonzero contribution.

By similar reasoning, we would have to choose $a_{i_3 i_4} \in$ Zone 1 and $a_{i_4 i_1} \in$ Zone 2, which gives

$$\begin{aligned} \alpha\pi(0) - \beta\pi(1) &= \alpha\pi(2) - \beta\pi(3) \\ -\beta\pi(1) + \alpha\pi(2) &= -\beta\pi(3) + \alpha\pi(4). \end{aligned} \tag{2.19}$$

Nevertheless, by summing these two equations, we can still introduce a new linear constraint, $\pi(1) = \pi(3)$. If we choose the first entry to be in Zone 2, we run into the same problem. Therefore, the word *abab* does not contribute when $\alpha \neq \beta$.

For $\alpha = \beta$, we are reduced to original Toeplitz matrices, and Bose and Hammond and Miller show that the contribution is $\frac{2}{3}$. Thus, we have[17]

$$M_4(\alpha,\beta) = \begin{cases} 2 & \alpha \neq \beta \\ \\ 2\frac{2}{3} & \alpha = \beta. \end{cases} \tag{2.20}$$

2.1.4. *Sixth Moment.* The formula for the sixth moment is

$$\begin{aligned} M_6 &= \lim_{N\to\infty} \frac{1}{N^4} \sum_{1 \leq i_1,i_2,i_3,i_4,i_5,i_6 \leq N} \mathbb{E}\left[a_{i_1 i_2} a_{i_2 i_3} a_{i_3 i_4} a_{i_4 i_5} a_{i_5 i_6} a_{i_6 i_1}\right] \\ &= \sum_{w: w \text{ is pair-matched of length 6}} \lim_{N\to\infty} \frac{1}{N^4} \#\Pi(w). \end{aligned} \tag{2.21}$$

[17] $M_k(\alpha,\beta)$ is the k^{th} limiting moment as a function of α and β.

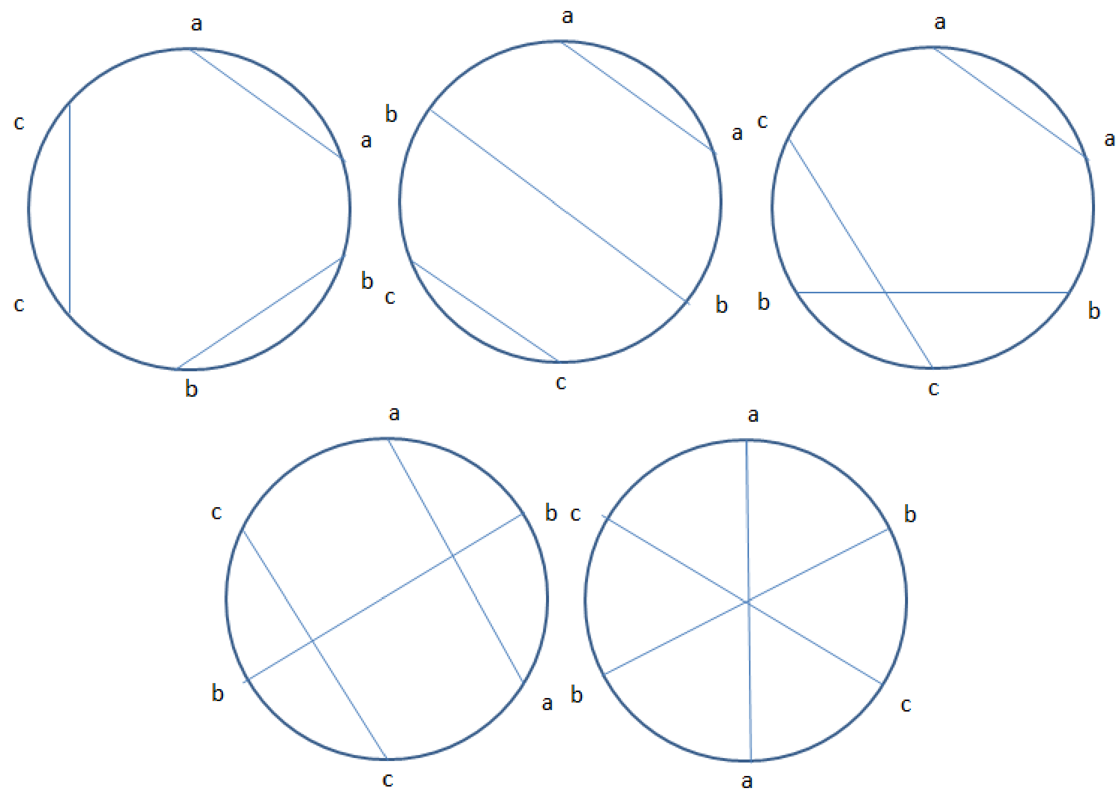

FIGURE 7. Words for the sixth moment.

The pair-matched words of length six are *aabbcc*, *aabccb*, *aabcbc*, *abacbc*, and *abcabc*, along with other words that are isomorphic to these. The five non-isomorphic configurations are shown in Figure 7. Respectively, there are 2, 3, 6, 3, and 1 versions for the above configurations.[18] We first count the number of circuits for the Catalan words, the first two words. We begin by showing that, for any moment, adjacent pairs must be located in opposite zones for a general class of link functions.

Lemma 2.2 (Adjacent Pairs.) *Let L(i,j) be a bivariate polynomial link function, where* $p_1(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0$ *for some* $m \in \mathbb{R}^+$, $p_2(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0$ *for some* $n \in \mathbb{R}^+$, *all the coefficients in both polynomials are non-negative, and the leading coefficients are also non-zero:*

$$L(i,j) = \begin{cases} p_1(i) \pm p_2(j) & i \leq j \\ \pm p_2(i) + p_1(j) & i > j. \end{cases} \tag{2.22}$$

[18] A configuration has x versions if there are x words that yield that configuration. For example, the configuration corresponding to the word *aabb* has two versions, since it also corresponds to the word *abba*.

Then adjacent pairs must be in opposite zones when $p_1(x) \neq p_2(x)$.

Proof. Any adjacent pair will have the form $a_{L(i_1,i_2)} = a_{L(i_2,i_3)}$. First, assume that $a_{i_1 i_2} \in$ Zone 1, $a_{i_2 i_3} \in$ Zone 1, and the coefficient of $p_2(x)$ in the link function is positive. The corresponding L-value equation will have the form

$$p_1(\pi(0)) + p_2(\pi(1)) = p_1(\pi(1)) + p_2(\pi(2)). \tag{2.23}$$

From the choice of zones, $\pi(0) \leq \pi(1)$ and $\pi(1) \leq \pi(2)$. The only way to satisfy the L-value equation is to have $\pi(0) = \pi(1) = \pi(2)$. If we choose $a_{i_1 i_2}$ to correspond to the first letter in a word,[19] then by definition, $\pi(0)$ and $\pi(1)$ are generating vertices. As in the proof of Lemma 1.2, the degrees of freedom originate from the $k+1$ generating vertices. Once we choose $\pi(0)$, $\pi(1)$ is fixed, we lose a generating vertex, and there are at most $O(N^k)$ solutions where we need at least $k+1$ degrees of freedom for a contribution. If we assume that the coefficient of $p_2(x)$ in the link function is negative, then the L-value equation will have the form

$$p_1(\pi(0)) - p_2(\pi(1)) = p_1(\pi(1)) - p_2(\pi(2)). \tag{2.24}$$

Rearranging, we have

$$p_1(\pi(1)) + p_2(\pi(1)) = p_1(\pi(0)) + p_2(\pi(2)). \tag{2.25}$$

Choose all k generating vertices except $\pi(1)$. This occurs with at most k degrees of freedom. Then, the non-generating vertices will be fixed by the other L-value equations, with a total of $\Delta(L)^k$ choices. Now both $\pi(0)$ and $\pi(2)$ are chosen and $\pi(1)$ is fixed by the above equation. This loss of the generating vertex $\pi(1)$ means there are at most $O(N^k)$ solutions and therefore a contribution of zero to the moment.

Now, assume that $a_{i_1 i_2} \in$ Zone 2 and $a_{i_2 i_3} \in$ Zone 2, and the coefficient of $p_2(x)$ in the link function is positive. The corresponding L-value equation will again have the form

$$p_1(\pi(0)) + p_2(\pi(1)) = p_1(\pi(1)) + p_2(\pi(2)). \tag{2.26}$$

The zonewise conditions give $\pi(0) < \pi(1) < \pi(2)$, so the L-value equation cannot even be satisfied. If we assume that the coefficient of $p_2(x)$ in the link function is negative, then the L-value equation will again have the form

$$p_1(\pi(1)) + p_2(\pi(1)) = p_1(\pi(0)) + p_2(\pi(2)) \tag{2.27}$$

and the previous argument applies. □

[19]For any word, we are free to cycle the letters in the word without changing the underlying configuration.

Instead of counting the contribution for each sixth moment word, we show that it suffices to know the contribution of the word *aabb* to the fourth moment. Consider one version of the word *aabbcc*:

$$\begin{aligned} a_{L(i_1,i_2)} &= a_{L(i_2,i_3)} \\ a_{L(i_3,i_4)} &= a_{L(i_4,i_5)} \\ a_{L(i_5,i_6)} &= a_{L(i_6,i_1)}. \end{aligned} \tag{2.28}$$

No matter what contributing zones $a_{i_5 i_6}$ and $a_{i_6 i_1}$ are located in, $i_5 = i_1$, since adjacent pairs must be in opposite zones by Lemma 2.2. Substituting for i_5, we can rewrite the first two equations of the matching as:

$$\begin{aligned} a_{L(i_1,i_2)} &= a_{L(i_2,i_3)} \\ a_{L(i_3,i_4)} &= a_{L(i_4,i_1)}. \end{aligned} \tag{2.29}$$

This has the same structure as the fourth moment adjacent matching. It is as though we had "lifted" the adjacent pair, $a_{L(i_5,i_6)} = a_{L(i_6,i_1)}$. For each of the four contributing zonewise cases in the fourth moment structure, there are now two additional possibilities for the third pair, $a_{i_5 i_6} \in$ Zone 1 and $a_{i_6 i_1} \in$ Zone 2, or $a_{i_5 i_6} \in$ Zone 2 and $a_{i_6 i_1} \in$ Zone 1. These two possibilities give $\pi(5) \leq \pi(0)$ or $\pi(5) > \pi(0)$, which leaves $\pi(5)$ as a free index. To compute the contribution to the sixth moment, then, we just integrate over the same region as in the fourth moment case, with an additional variable:

$$\int_0^1 \int_0^1 \int_0^1 \int_0^1 dv_0 dv_1 dv_3 dv_5 = 1. \tag{2.30}$$

The word *aabccb* is handled similarly. The system of equations is:

$$\begin{aligned} a_{L(i_1,i_2)} &= a_{L(i_4,i_5)} \\ a_{L(i_2,i_3)} &= a_{L(i_3,i_4)} \\ a_{L(i_5,i_6)} &= a_{L(i_6,i_1)}. \end{aligned} \tag{2.31}$$

We "lift" the second matched pair by noticing that for any set of zones, $\pi(1) = \pi(3)$. Relabeling $\pi(4)$ and $\pi(5)$, we are left with the fourth moment structure again. As above, the case contributes one to the moment. We can now prove the following general lemma:

Lemma 2.3 (Adjacent Lifting.) *Let L(i,j) be a bivariate polynomial link function, where* $p_1(x) = a_m x^m + a_{m-1} x^{m-1} + \ldots + a_0$ *for some* $m \in \mathbb{R}^+$ *and* $p_2(x) = b_n x^n + b_{n-1} x^{n-1} + \ldots +$

b_0 for some $n \in \mathbb{R}^+$, and where all the coefficients in both polynomials are non-negative:

$$L(i,j) = \begin{cases} p_1(i) \pm p_2(j) & i \leq j \\ \pm p_2(i) + p_1(j) & i > j. \end{cases} \tag{2.32}$$

Then, any Catalan word of length $2k$ for matrices with such a link function contributes one to the $2k^{th}$ moment.

Proof. We begin by counting the number of circuits for the word *aabb*. Adjacent pairs must be in opposite zones, by Lemma 2.2. So we calculate four possibilities:

(1) $a_{i_1,i_2} \in$ Zone 1, $a_{i_2,i_3} \in$ Zone 2, $a_{i_3,i_4} \in$ Zone 1, and $a_{i_4,i_1} \in$ Zone 2: $p_1(\pi(0)) \pm p_2(\pi(1)) = \pm p_2(\pi(1)) + p_1(\pi(2))$ and $p_1(\pi(2)) \pm p_2(\pi(3)) = \pm p_2(\pi(3)) + p_1(\pi(0)) \longrightarrow \pi(0) = \pi(2),\ \pi(1) > \pi(0)$ and $\pi(3) > \pi(0)$

(2) $a_{i_1,i_2} \in$ Zone 1, $a_{i_2,i_3} \in$ Zone 2, $a_{i_3,i_4} \in$ Zone 2, and $a_{i_4,i_1} \in$ Zone 1: $p_1(\pi(0)) \pm p_2(\pi(1)) = \pm p_2(\pi(1)) + p_1(\pi(2))$ and $\pm p_2(\pi(2)) + p_1(\pi(3)) = p_1(\pi(3)) \pm p_2(\pi(0)) \longrightarrow \pi(0) = \pi(2),\ \pi(1) > \pi(0)$ and $\pi(3) < \pi(0)$

(3) $a_{i_1,i_2} \in$ Zone 2, $a_{i_2,i_3} \in$ Zone 1, $a_{i_3,i_4} \in$ Zone 1, and $a_{i_4,i_1} \in$ Zone 2: $\pm p_2(\pi(0)) + p_1(\pi(1)) = p_1(\pi(1)) \pm p_2(\pi(2))$ and $p_1(\pi(2)) \pm p_2(\pi(3)) = \pm p_2(\pi(3)) + p_1(\pi(0)) \longrightarrow \pi(0) = \pi(2),\ \pi(1) < \pi(0)$ and $\pi(3) > \pi(0)$

(4) $a_{i_1,i_2} \in$ Zone 2, $a_{i_2,i_3} \in$ Zone 1, $a_{i_3,i_4} \in$ Zone 2, and $a_{i_4,i_1} \in$ Zone 1: $\pm p_2(\pi(0)) + p_1(\pi(1)) = p_1(\pi(1)) \pm p_2(\pi(2))$ and $\pm p_2(\pi(2)) + p_2(\pi(3)) = p_1(\pi(3)) \pm p_2(\pi(0)) \longrightarrow \pi(0) = \pi(2),\ \pi(1) < \pi(0)$ and $\pi(3) < \pi(0)$.

We can follow the same calculation as in the Toeplitz case to show that the word *aabb* contributes one to the fourth moment. For higher moments, any non-crossing pair partition must have at least one adjacent pair of the form $a_{i_1 i_2} = a_{i_2 i_3}$, by the definition of a Catalan word. Since adjacent pairs must be located in opposite zones, any such adjacent pair must require $\pi(0) = \pi(2)$. Since there are two sets of zones for the pair, the remaining index is bound either by $\pi(1) \leq \pi(0)$ or $\pi(1) > \pi(0)$, leaving $\pi(1)$ as a free index. "Lift" this pair by setting $\pi(0) = \pi(2)$ and relabeling the remaining indices appropriately. Since there are now $2k - 2$ indices left, the remaining structure is a Catalan word for the $(2k-2)^{th}$ moment, and the contribution can be computed with $\pi(1)$ as an extra degree of freedom. Using this process, any adjacent matching can be reduced to the fourth moment structure, and since that structure contributes one, any other Catalan word contributes one. In other words, the contribution of a configuration remains unchanged if only adjacent pairs are added to the structure, since the extra degree of freedom from an adjacent pair balances out the extra factor of $\frac{1}{N}$ in the moment formula. □

Using Lemma 2.3 the words *aabbcc*, *aabccb*, and other words isomorphic to them contribute one to the sixth moment.

For the word *aabcbc*, we can also "lift" the adjacent pair and relabel the remaining indices so that the fourth moment structure for the word *abab* remains. Therefore, versions of this word contribute 0 when $\alpha \neq \beta$ and $\frac{2}{3}$ when $\alpha = \beta$. The process of "lifting" is useful for both Catalan and non-Catalan words.

For the word *abacbc*, we use an argument similar to that for the fourth moment. Assume $\alpha \neq \beta$. We first reason that by checking for extra linear constraints, the only possible zonewise cases that might contribute are

$$\begin{aligned} \alpha\pi(0) - \beta\pi(1) &= \alpha\pi(4) - \beta\pi(5) \\ -\beta\pi(1) + \alpha\pi(2) &= -\beta\pi(3) + \alpha\pi(4) \\ \alpha\pi(2) - \beta\pi(3) &= -\beta\pi(5) + \alpha\pi(0) \end{aligned} \tag{2.33}$$

and

$$\begin{aligned} -\beta\pi(0) + \alpha\pi(1) &= -\beta\pi(4) + \alpha\pi(5) \\ \alpha\pi(1) - \beta\pi(2) &= \alpha\pi(3) - \beta\pi(4) \\ -\beta\pi(2) + \alpha\pi(3) &= \alpha\pi(5) - \beta\pi(0). \end{aligned} \tag{2.34}$$

Nevertheless, summing each of these equations still produces an extra linear constraint; hence, there is no contribution to the moment.

For $\alpha = \beta$, the original Toeplitz matrices, Bose and Hammond and Miller compute the contribution to be $\frac{1}{2}$.

For the word *abcabc*, there are two sets of zones that do not produce extra linear constraints. The L-value equations for these cases are:

$$\begin{aligned} \alpha\pi(0) - \beta\pi(1) &= -\beta\pi(3) + \alpha\pi(4) \\ -\beta\pi(1) + \alpha\pi(2) &= \alpha\pi(4) - \beta\pi(5) \\ \alpha\pi(2) - \beta\pi(3) &= -\beta\pi(5) + \alpha\pi(0) \end{aligned} \tag{2.35}$$

and

$$\begin{aligned} -\beta\pi(0) + \alpha\pi(1) &= \alpha\pi(3) - \beta\pi(4) \\ \alpha\pi(1) - \beta\pi(2) &= -\beta\pi(4) + \alpha\pi(5) \\ -\beta\pi(2) + \alpha\pi(3) &= \alpha\pi(5) - \beta\pi(0). \end{aligned} \tag{2.36}$$

Let's begin with the first case. We can choose $\pi(0)$, $\pi(1)$, $\pi(2)$, and $\pi(3)$ freely. Then $\pi(4)$ and $\pi(5)$ are fixed:

$$\begin{aligned}\pi(4) &= \pi(0) - \frac{\alpha}{\beta}\pi(1) + \frac{\alpha}{\beta}\pi(3) \\ \pi(5) &= \pi(3) - \frac{\beta}{\alpha}\pi(2) + \frac{\beta}{\alpha}\pi(0).\end{aligned} \tag{2.37}$$

Using the same variable transformation as in the fourth moment, we can rewrite these constraints as

$$\begin{aligned}v_4 &= v_0 - \frac{\alpha}{\beta}v_1 + \frac{\alpha}{\beta}v_3 \\ v_5 &= v_3 - \frac{\beta}{\alpha}v_2 + \frac{\beta}{\alpha}v_0.\end{aligned} \tag{2.38}$$

To count the contribution, we just integrate the indicator function acting on the regions defined above along with the choices of zones. In the limit of large N, this becomes a quadruple integral.

For $a = \frac{\beta}{\alpha}$ and $\beta > \alpha$, or $a > 1$, the contribution is

$$\begin{aligned}&\int_0^1\int_0^1\int_0^1\int_0^1 \mathbb{I}(0 \leq v_0 - \frac{v_1}{a} + \frac{v_3}{a} \leq 1 \text{ and } 0 \leq av_0 - av_2 + v_3 \leq 1 \text{ and } v_0 > v_1 \text{ and} \\ &v_1 < v_2 \text{ and } v_2 > v_3 \text{ and } v_3 < v_0 - \frac{v_1}{a} + \frac{v_3}{a} \text{ and } v_0 - \frac{v_1}{a} + \frac{v_3}{a} > av_0 - av_2 + v_3 \\ &\text{and } av_0 - av_2 + v_3 \leq v_0) dv_1 dv_2 dv_3 dv_0.\end{aligned} \tag{2.39}$$

The contribution from this case is[20]

$$\frac{1}{4(1+a)} = \frac{\alpha}{4(\alpha+\beta)}. \tag{2.40}$$

Now consider the case $\alpha > \beta$. The set of L-value equtions we have been using is

$$\begin{aligned}-\beta\pi(0) + \alpha\pi(1) &= \alpha\pi(3) - \beta\pi(4) \\ \alpha\pi(1) - \beta\pi(2) &= -\beta\pi(4) + \alpha\pi(5) \\ -\beta\pi(2) + \alpha\pi(3) &= \alpha\pi(5) - \beta\pi(6).\end{aligned} \tag{2.41}$$

[20] See section 7.3 for the full calculation of the integral.

Perform a symmetric change of the indices, which will not affect the number of solutions:

$$\begin{aligned}\pi(0) &\longleftrightarrow \pi(3)\\ \pi(1) &\longleftrightarrow \pi(4)\\ \pi(2) &\longleftrightarrow \pi(5).\end{aligned} \tag{2.42}$$

Under this transformation, we have the rewritten configuration:

$$\begin{aligned}\alpha\pi(0)-\beta\pi(1) &= -\beta\pi(3)+\alpha\pi(4)\\ -\beta\pi(1)+\alpha\pi(2) &= \alpha\pi(4)-\beta\pi(5)\\ \alpha\pi(2)-\beta\pi(3) &= -\beta\pi(5)+\alpha\pi(6).\end{aligned} \tag{2.43}$$

Following the same steps above, and switching the relevant inequalities to conform to the transformation, these equations lead to another quadruple integral:

$$\begin{aligned}&\int_0^1\int_0^1\int_0^1\int_0^1 \mathbb{I}(0 \le v_0 - \frac{\beta}{\alpha}v_1 + \frac{\beta}{\alpha}v_3 \le 1 \text{ and } 0 \le \frac{-\alpha}{\beta}v_2 + v_3 + \frac{\alpha}{\beta}v_0 \le 1 \text{ and } v_0 \le v_1\\ &\text{and } v_1 > v_2 \text{ and } v_2 \le v_3 \text{ and } v_3 > v_0 - \frac{\beta}{\alpha}v_1 + \frac{\beta}{\alpha}v_3 \text{ and } v_0 - \frac{\beta}{\alpha}v_1 + \frac{\beta}{\alpha}v_3 \le \frac{-\alpha}{\beta}v_2 + v_3\\ &+ \frac{\alpha}{\beta}v_0 \text{ and } \frac{-\alpha}{\beta}v_2 + v_3 + \frac{\alpha}{\beta}v_0 > v_0)dv_1dv_2dv_3dv_0.\end{aligned} \tag{2.44}$$

Let $a = \frac{\alpha}{\beta}$. By a similar calculation, we get a symmetric result:

$$\frac{1}{4(1+a)} = \frac{1}{4(1+\frac{\alpha}{\beta})} = \frac{\beta}{4(\alpha+\beta)}. \tag{2.45}$$

It is simple to calculate the contribution from Eq. 2.36. Notice that we just calculated the contribution for the set of L-value equations in Eq. 2.43. This is exactly the same set of equations in Eq. 2.36. Therefore, both cases contribute equally.

Since there are two versions of the configuration corresponding to the word *aabbcc* and three versions of the configuration corresponding to the word *aabccb*, the Catalan words contribute a total of 5 to the moment. Including the extra factor from the word *abcabc*, we have

$$M_6(\alpha,\beta) = \begin{cases} 5 + \frac{\alpha}{2}\frac{1}{\alpha+\beta} & \alpha < \beta \\ \\ 5 + \frac{\beta}{2}\frac{1}{\alpha+\beta} & \alpha > \beta \\ \\ 11 & \alpha = \beta. \end{cases} \tag{2.46}$$

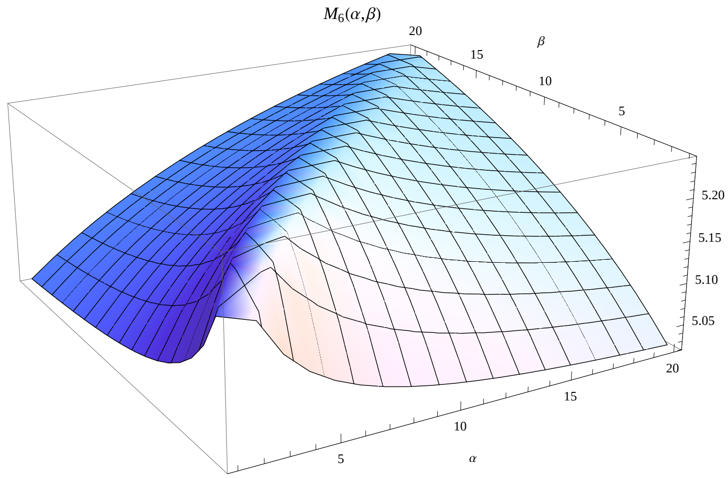

FIGURE 8. A 3D Mathematica listplot of the generalized Toeplitz sixth moment, $M_6(\alpha,\beta)$, for integer values of α and β with $\alpha \neq \beta$ up to 20.

Using Mathematica we computed the relevant integrals for low fixed values of α and β and had agreement with the above formula:

α	β	$M_6(\alpha,\beta)$
1	2	$\frac{31}{6} = 5 + \frac{1}{6}$
1	3	$\frac{41}{8} = 5 + \frac{1}{8}$
2	1	$\frac{31}{6} = 5 + \frac{1}{6}$
2	3	$\frac{41}{8} = 5 + \frac{1}{8}$

(2.47)

We collect the full results on the moments:

Moment	Gaussian	Toeplitz	Generalized Toeplitz	Semicircle
M_4	3	$2\frac{2}{3}$	$\begin{cases} 2 & \alpha \neq \beta \\ 2\frac{2}{3} & \alpha = \beta \end{cases}$	2
M_6	15	11	$\begin{cases} 5 + \frac{\alpha}{2}\frac{1}{\alpha+\beta} & \beta > \alpha \\ 5 + \frac{\beta}{2}\frac{1}{\alpha+\beta} & \alpha > \beta \end{cases}$	5
M_8	105	$64\frac{4}{16}$		14

(2.48)

2.1.5. *Existence of Higher Moments.* Unfortunately, all higher moments for this link function become increasingly computationally intensive. Although we cannot find a closed-form expression for higher moments, we can show that all higher moments exist and are finite.

Lemma 2.4 (Existence of Higher Moments.) *If the probability distribution $p(x)$ has mean* *zero and variance one, then for all non-negative integers k, $M_k = lim_{N\to\infty} M_k(N)$ exists and is finite.*

Proof. Since odd moments are zero, it suffices to check the limiting behavior of the even moments. As exemplified in the fourth and sixth moment calculations, for any word w of length $2k$, we obtain a system of linear equations relating the variables $\pi(0), ..., \pi(2k-1) \in \{1, ..., N\}$. These variables together have at most $k+1$ degrees of freedom. Letting $v_x = \frac{\pi(x)}{N} \in \{\frac{1}{N}, \frac{2}{N}, ..., \frac{1}{N}\}$, the system of equations then determines a nice region in the $(k+1)$-dimensional unit cube. As $N \to \infty$, we obtain the finite volume of this region. This volume, which we denote $M_{2k}(w)$, is the coefficient of the leading order term in the number of solutions to the original system of equations. Transforming back to the variables $\pi(x)$, then, we obtain the contribution of this word to the $2k^{th}$ moment, before dividing by N^{k+1}, to be $M_{2k}(w)N^{k+1} + O_k\left(N^k\right)$. Summing over all pair-matched words of length $2k$ gives $M_{2k}N^{k+1} + O_k\left(N^k\right)$. We extract the finite limiting moment, M_{2k}, by dividing by N^{k+1} for N large. □

2.1.6. *Bounds on the Moments.* It is easy to argue the following bounds for the moments.

Lemma 2.5 (Bounds on the Moments.) *Let C_{2k} be the $2k^{th}$ moment of the semicircle distribution, $M_{2k}(T)$ the $2k^{th}$ moment of the Toeplitz ensemble limiting distribution, and $M_{2k}\left(\alpha,\beta\right)$ the $2k^{th}$ moment of the generalized Toeplitz ensemble limiting distribution. Then, $C_{2k} \leq M_{2k}\left(\alpha,\beta\right) < M_{2k}(T)$ for all non-negative integers k.*

Proof. Since each Catalan word contributes one, the moments are at least as large as the semicircle moments, and the lower bound holds. For a non-Catalan word, either the α and β in the link function prevent contributions that otherwise occur when $\alpha = \beta$, or they decrease these contributions. Let $g = gcd(\alpha, \beta)$. Then for a given matrix entry, there are at most $\left\lceil \frac{N}{\max(\alpha,\beta)/g} \right\rceil$ L-matches in the upper triangle of the matrix. There are always fewer matchings than the Toeplitz case, for which $\alpha = \beta$, which means there are always fewer solutions to the relevant Diophantine equations, and hence the upper bound holds. □

2.1.7. *Convergence.* Now that we have calculated low moments of the limiting spectral distribution and proved that all higher moments exist and are finite, we can show that the empirical measures for generalized Toeplitz matrices converge in probability and almost surely to a unique limiting spectral distribution that is *universal*, or independent of $p(x)$. By the arguments in section 1.2.4 and the fact that these matrices satisfy Property B, the limiting moments determine a unique limiting spectral distribution. By the arguments in section 1.2.5, Var$[M_k(A_N)] \to 0$ and the empirical spectral distributions converge in probability to the limiting spectral distribution. By the arguments in section 1.2.6, the empirical distributions converge almost surely to the limiting distribution. All of the above arguments only depend on $p(x)$ having mean zero, variance one, and uniformly bounded moments of all order. Hence, the convergence is universal. [21]

2.1.8. *Limiting Behavior.* We can also show that in the limit as either α or β becomes very large, the moments of the limiting distribution for the generalized Toeplitz ensemble approach those of the semicircle measure. This is clear for the sixth moment, for example, which approaches the sixth Catalan number in either limit:

$$\begin{aligned} &\text{For fixed } \alpha,\ \lim_{\beta\to\infty}\ M_6(\alpha,\beta) = \lim_{\beta\to\infty}\ 5 + \frac{\alpha}{2}\frac{1}{\alpha+\beta} = 5. \\ &\text{For fixed } \beta,\ \lim_{\alpha\to\infty}\ M_6(\alpha,\beta) = \lim_{\alpha\to\infty}\ 5 + \frac{\beta}{2}\frac{1}{\alpha+\beta} = 5. \end{aligned} \tag{2.49}$$

In general, we assert the following:

Lemma 2.6 (Limiting Behavior for Generalized Toeplitz Matrices.) *For fixed α, lim$_{\beta\to\infty}$ $M_{2k}(\alpha,\beta) = C_{2k}$, and for fixed β, lim$_{\alpha\to\infty}$ $M_{2k}(\alpha,\beta) = C_{2k}$ for generalized Toeplitz matrices when the limits are taken appropriately.*

Proof. Consider the limit as $\beta \to \infty$. To prove that the limiting spectral measure is a semicircle, it suffices to show that all non-Catalan words contribute zero. Specifically, we will show that the contribution of a word that is fully crossed, or a word that has no adjacent pairs, is zero. This shows that all non-Catalan words contribute zero, since the contribution of any non-Catalan word with adjacent pairs is calculated by "lifting" the adjacent pairs, and any loss in degrees of freedom in lower moments will propagate through these adjacent pairs. We will let β grow to infinity as a function of N, such that $\lim_{N\to\infty} f(N) = \infty$:

$$L(i,j) = \begin{cases} \alpha i - f(N)j & i \le j \\ -f(N)i + \alpha j & i > j. \end{cases} \tag{2.50}$$

Any fully crossed matching will have non-adjacent pairs of the form $a_{L(i_1,i_2)} = a_{L(i_3,i_4)}$. For any such pair there are four possible sets of zones for the matching, each of which will result in one index with coefficient α on both sides of the equation and one index with coefficient β on both sides of the L-value equation. Without loss of generality, then, we

[21] See Bose for proof of the fact that the limiting distribution is universal for all matrix ensembles that satisfy Property B.

will assume that $a_{i_1i_2} \in$ Zone 1 and $a_{i_3i_4} \in$ Zone 1. The corresponding L-value equation has the form

$$\alpha\pi(0) - f(N)\pi(1) = \alpha\pi(2) - f(N)\pi(3). \tag{2.51}$$

In the following argument, we also want to choose entries $a_{i_1i_2}$ and $a_{i_3i_4}$ to be specific letters in the fully crossed word. We will pick this matched pair such that all letters between the letters that correspond to these entries are distinct. Pick a matched pair, then check if there are any duplicated letters between these matched letters. If there are duplicated letters, then choose the entries corresponding to those letters to be the matched pair. Continue this process until all the intermediate letters are distinct. There must be at least one distinct letter between the matched letters remaining; otherwise, the word would have an adjacent pair. For the resulting pair, choose the first vertex to be the generating vertex $\pi(0)$. Then the next vertex, $\pi(1)$, is also a generating vertex, as it now corresponds to the first letter in the word. Although we cannot know in general the location of the other vertices, we will just name them $\pi(2)$ and $\pi(3)$. Since intermediate letters are distinct and the word began with generating vertex $\pi(0)$, $\pi(2)$ must correspond to the first occurrence of a letter in the word, making $\pi(2)$ another generating vertex. Therefore, there are at least three generating vertices in this pair. Let's say that there are x degrees of freedom in choosing all four vertices in this matching. There were initially $k+1$ generating vertices, but now there are at most $k-2$ generating vertices remaining and at most $x+k-2$ degrees of freedom in total. If we can show that $x < 3$, then any non-Catalan word of length $2k$ contributes zero to the $2k^{th}$ moment, since there must be $k+1$ degrees of freedom for a contribution.

Here there are at most N choices for $\pi(0)$ and N choices for $\pi(1)$. Then the number of choices for $\pi(2)$ and $\pi(3)$ is equivalent to the number of matchings in the matrix for entry $a_{i_1i_2}$, given fixed values for $\pi(0)$ and $\pi(1)$. In the upper triangle of the matrix, when N is large enough, there will be at most $\left\lceil \frac{N}{f(N)/g} \right\rceil$ matchings where $g = gcd(\alpha, f(N))$. So in totall, there are at most $2\left\lceil \frac{N}{f(N)/g} \right\rceil$ matchings. To see if there are fewer than three degrees of freedom, we check if $\lim_{N\to\infty} \frac{\#\text{solutions to Eq. 2.51}}{N^3} = 0$. Dropping the ceiling notation, we have

$$\lim_{N\to\infty} \frac{\#\text{solutions to Eq. 2.51}}{N^3} \propto \lim_{N\to\infty} \frac{2N^3}{N^3 f(N)/g} = \lim_{N\to\infty} \frac{2g}{f(N)} = 0. \tag{2.52}$$

Since the same argument holds for every set of zones, $x < 3$. A similar proof holds for $\alpha \to \infty$. □

2.2. **Generalized Hankel Matrices.** Real symmetric Hankel matrices are matrices that are constant along the skew diagonals, or lines with slope 1. The link function for Hankel

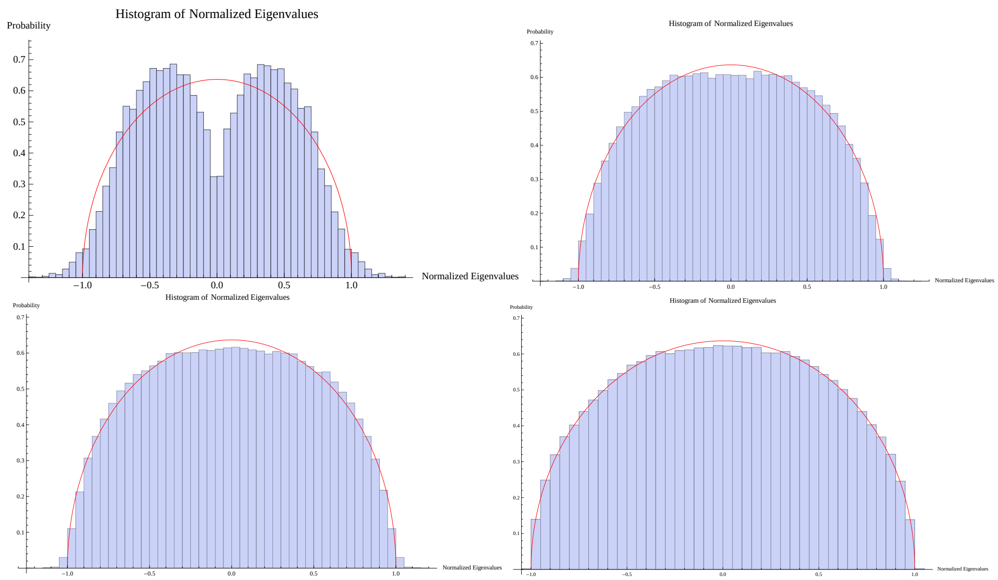

FIGURE 9. Histograms of numerical eigenvalues of 100 generalized Hankel matrices of size 1200×1200. Each has $\alpha = 1$. Clockwise from the upper left, β is equal to 1, 2, 3, and 4. The red curve is the semicircle distribution of Eq. 1.3 for an eigenvalue normalization of $2\sqrt{N}$. See section 7.1 for details on normalizations for the eigenvalues.

matrices can be written as $L(i, j) = i+j$. We will generalize the link function to all positive rational slopes in the upper triangle by introducing parameters α and β for a *generalized Hankel link function*.[22] For fixed $\alpha, \beta \in \mathbb{Q}^+$,

$$L_{\alpha,\beta}(i,j) = \begin{cases} \alpha i + \beta j & i \leq j \\ \\ \beta i + \alpha j & i > j. \end{cases} \tag{2.53}$$

A matrix with $\alpha = \beta$ reduces to the original Hankel matrix, while a 5×5 matrix with $\alpha = 2$ and $\beta = 1$ would have the structure

$$A = \begin{pmatrix} a_3 & a_4 & a_5 & a_6 & a_7 \\ a_4 & a_6 & a_7 & a_8 & a_9 \\ a_5 & a_7 & a_9 & a_{10} & a_{11} \\ a_6 & a_8 & a_{10} & a_{12} & a_{13} \\ a_7 & a_9 & a_{11} & a_{13} & a_{15} \end{pmatrix}. \tag{2.54}$$

[22] As in the generalized Toeplitz case, we restrict the parameters to rational numbers.

Lower moments of the Hankel matrices are known, and it has been proved that the Hankel limiting spectral distribution is not unimodal. The following table compares low moments for the Gaussian, Toeplitz, Hankel, and semicircular distributions:

Moment	Gaussian	Toeplitz	Hankel	Semicircle
M_4	3	$2\frac{2}{3}$	2	2
M_6	15	11	$5\frac{1}{2}$	5
M_8	105	$64\frac{4}{16}$	$18\frac{11}{15}$	14

(2.55)

As in the generalized Toeplitz case, numerics suggest that the generalized Hankel matrices have a near-semicircular limiting spectral distribution. Figure 9 shows simulations of the limiting spectral distribution for generalized Hankel matrices for $\alpha = 1$ and several values of β. The histograms clearly show the bimodal behavior for $\beta = 1$, which corresponds to original Hankel matrices. For larger values of β, however, the distribution looks increasingly semicircular. We now compute low moments for the generalized Hankel ensemble and prove that low even moments deviate from the Catalan numbers by a factor that depends on α and β, which is smaller than that of the generalized Toeplitz matrices. We again prove that as either α or β tends to infinity, the limiting spectral distribution converges to the semicircle.

2.2.1. *Odd Moments.* Because the slopes in the upper triangle for matrices with the generalized Hankel link function can never be zero, $\Delta(L)$ is at most one. Odd moments are then zero, because these matrices satisfy Property B.

2.2.2. *Zeroth and Second Moments.* By the same calculation in the generalized Toeplitz case, $M_0 = M_2 = 1$.

2.2.3. *Fourth Moment.* Again following the same procedure as in the generalized Toeplitz case, we first count the number of circuits for the word *aabb*. Since the generalized Hankel link function satisfies the link function properties required for Lemma 2.3, all Catalan words contribute one. Now consider contributions from the word *abab*. By the same argument that applied in the generalized Toeplitz case, there will be an extra constraint in the L-value equations if $\alpha \neq \beta$. For $\alpha = \beta$, we are reduced to original Hankel matrices, and Bose shows that the contribution for this word is zero. Thus, we have

$$M_4(\alpha, \beta) = 2. \tag{2.56}$$

2.2.4. *Sixth Moment.* By Lemma 2.3, the words *aabbcc*, *aabccb* and other words isomorphic to them contribute one to the sixth moment. Using the process of "lifting", the word *aabcbc* contributes zero, since the non-contributing structure *bcbc* is embedded within that word. Counting linear constraints, the word *abacbc* contributes zero when $\alpha \neq \beta$. It also contributes zero when $\alpha = \beta$, according to calculations in Bose.

For the word *abcabc* and others isomorphic to it, we follow the generalized Toeplitz calculation. The two sets of contributing L-value equations are

$$\begin{aligned} \beta\pi(0) + \alpha\pi(1) &= \alpha\pi(3) + \beta\pi(4) \\ \alpha\pi(1) + \beta\pi(2) &= \beta\pi(4) + \alpha\pi(5) \\ \beta\pi(2) + \alpha\pi(3) &= \alpha\pi(5) + \beta\pi(6) \end{aligned} \tag{2.57}$$

and

$$\begin{aligned} \alpha\pi(0) + \beta\pi(1) &= \beta\pi(3) + \alpha\pi(4) \\ \beta\pi(1) + \alpha\pi(2) &= \alpha\pi(4) + \beta\pi(5) \\ \alpha\pi(2) + \beta\pi(3) &= \beta\pi(5) + \alpha\pi(6). \end{aligned} \tag{2.58}$$

In addition to the zonewise constraints, we have the following constraints for the transformed variables:

$$\begin{aligned} v_4 &= v_0 + \frac{\beta}{\alpha}v_1 - \frac{\beta}{\alpha}v_3 \\ v_5 &= \frac{\alpha}{\beta}v_2 + v_3 - \frac{\alpha}{\beta}v_0. \end{aligned} \tag{2.59}$$

For $a = \frac{\alpha}{\beta}$ and $\beta > \alpha$, we integrate

$$\int_0^1\int_0^1\int_0^1\int_0^1 \mathbb{I}(0 \leq v_0 + \frac{v_1}{a} - \frac{v_3}{a} \leq 1 \text{ and } 0 \leq av_2 + v_3 - av_0 \leq 1 \text{ and } v_0 > v_1 \text{ and } v_1 < v_2 \text{ and } v_2 > v_3 \text{ and } v_3 < v_0 + \frac{v_1}{a} - \frac{v_3}{a} \text{ and } v_0 + \frac{v_1}{a} - \frac{v_3}{a} > av_2 + v_3 - av_0 \text{ and } av_2 + v_3 - av_0 < v_0) dv_1 dv_2 dv_3 dv_0. \tag{2.60}$$

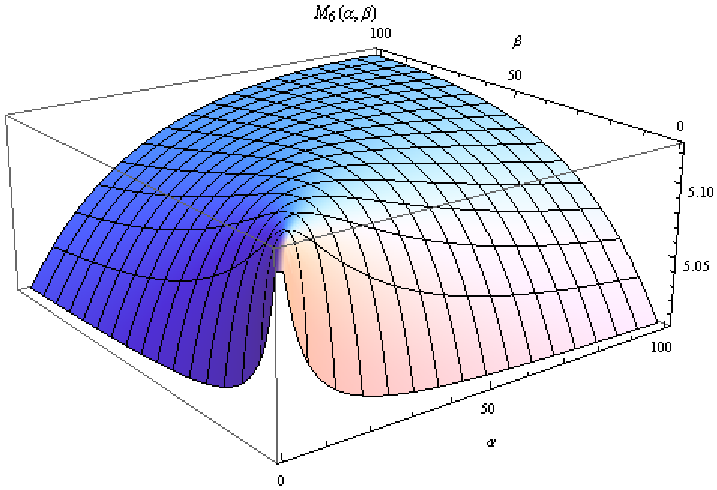

FIGURE 10. A 3D Mathematica listplot of the sixth moment for generalized Hankel matrices, $M_6(\alpha,\beta)$, for integer values of α and β with $\alpha \neq \beta$ up to 20.

This integral can be calculated in a similar manner as the generalized Toeplitz case, and the same transformation of indices applies for computing the integral for $\beta < \alpha$. The result is:

$$M_6(\alpha,\beta) = \begin{cases} 5 + \frac{\alpha}{2}\frac{\beta}{(\alpha+\beta)^2} & \alpha < \beta \\ \\ 5 + \frac{\alpha}{2}\frac{\beta}{(\alpha+\beta)^2} & \alpha > \beta \\ \\ 5\frac{1}{2} & \alpha = \beta. \end{cases} \tag{2.61}$$

Using Mathematica we computed the relevant integrals for low fixed values of α and β and had agreement with the above formula:

α	β	$M_6(\alpha,\beta)$
1	2	$\frac{46}{9} = 5 + \frac{1}{9}$
1	3	$\frac{163}{32} = 5 + \frac{3}{32}$
2	1	$\frac{46}{9} = 5 + \frac{1}{9}$
2	3	$\frac{163}{32} = 5 + \frac{3}{32}$

(2.62)

We collect the calculations for the moments of generalized Toeplitz and generalized Hankel matrices:

$$\begin{array}{|l|l|l|}
\hline
\text{Moment} & \text{Generalized Toeplitz} & \text{Generalized Hankel} \\
\hline
M_4 & \begin{cases} 2 & \alpha \neq \beta \\ 2\frac{2}{3} & \alpha = \beta \end{cases} & 2 \\
\hline
M_6 & \begin{cases} 5 + \frac{\alpha}{2}\frac{1}{\alpha+\beta} & \alpha \leq \beta \\ 5 + \frac{\beta}{2}\frac{1}{\alpha+\beta} & \alpha > \beta \\ 11 & \alpha = \beta \end{cases} & \begin{cases} 5 + \frac{\alpha}{2}\frac{\beta}{(\alpha+\beta)^2} & \alpha \leq \beta \\ 5 + \frac{\alpha}{2}\frac{\beta}{(\alpha+\beta)^2} & \alpha > \beta \\ 5\frac{1}{2} & \alpha = \beta \end{cases} \\
\hline
\end{array} \tag{2.63}$$

2.2.5. *Existence of Higher Moments.* As in the generalized Toeplitz case, although we cannot find a closed-form expression for all higher moments, we can show that higher moments exist and are finite.

Lemma 2.7 (Existence of Higher Moments.) *If the probability distribution $p(x)$ has mean zero and variance one, then for all k, $M_k = lim_{N\to\infty} M_k(N)$ exists and is finite.*

Proof. The proof follows as in Lemma 2.4. □

2.2.6. *Bounds on the Moments.* It is easy to argue the following bounds for the moments.

Lemma 2.8 (Bounds on the Moments.) *Let C_{2k} be the $2k^{th}$ moment of the semicircle distribution, $M_{2k}(H)$ the $2k^{th}$ moment of the Hankel ensemble limiting distribution, and $M_{2k}\left(\alpha, \beta\right)$ the $2k^{th}$ moment of the generalized Hankel ensemble limiting distribution. Then, $C_{2k} \leq M_{2k}\left(\alpha, \beta\right) < M_{2k}(H)$ for all non-negative integers k.*

Proof. The proof follows as in Lemma 2.5 □

2.2.7. *Convergence.* Now that we have calculated low moments of the limiting spectral distribution and proved that all higher moments exist and are finite, we can show that the empirical measures for generalized Hankel matrices converge in probability and almost surely to a unique limiting spectral distribution that is universal. By the arguments in section 1.2.4 and the fact that these matrices satisfy Property B, the limiting moments determine a unique limiting spectral distribution. By the arguments in section 1.2.5, Var$[M_k(A_N)] \to 0$ and the empirical spectral distributions converge in probability to the limiting spectral distribution. By the arguments in section 1.2.6, the empirical distributions converge almost surely to the limiting distribution. All of the above arguments only depend on $p(x)$ having mean zero, variance one, and uniformly bounded moments of all order. Hence, the convergence is universal.

2.2.8. *Limiting Behavior.* We can also show that in the limit as either α or β becomes very large, the moments of the generalized Hankel ensemble approach those of the semicircle measure. This is clear for the sixth moment, for example, which approaches the sixth Catalan number in either limit:

$$\begin{aligned}\text{For fixed } \alpha,\ \lim_{\beta\to\infty}\ M_6(\alpha,\beta) &= \lim_{\beta\to\infty}\ 5+\frac{\alpha}{2}\frac{\beta}{(\alpha+\beta)^2}=5.\\ \text{For fixed } \beta,\ \lim_{\alpha\to\infty}\ M_6(\alpha,\beta) &= \lim_{\alpha\to\infty}\ 5+\frac{\alpha}{2}\frac{\beta}{(\alpha+\beta)^2}=5.\end{aligned} \tag{2.64}$$

In general, we assert the following:

Lemma 2.8 (Limiting Behavior for Generalized Hankel Matrices.) *For fixed α, $\lim_{\beta\to\infty}\ M_{2k}(\alpha,\beta) = C_{2k}$ and for fixed β, $\lim_{\alpha\to\infty}\ M_{2k}(\alpha,\beta) = C_{2k}$ for generalized Hankel matrices when the limits are taken appropriately.*

Proof. The proof follows as in Lemma 2.6. □

3. Higher Order Polynomials with Equal Powers

3.1. **Hyperbolic Matrices.** We can further generalize the Toeplitz matrices by raising the variables in the link function to powers higher than one. Here we will raise both indices to the same power by introducing the parameter, n, giving us a link function for the *hyperbolic matrices*. For $\alpha,\ \beta \in \mathbb{Q}^+$ and $n > 1$, we will define

$$L_{\alpha,\beta}(i,j) = \begin{cases} \alpha i^n - \beta j^n & i \le j \\ \\ -\beta i^n + \alpha j^n & i > j. \end{cases} \tag{3.1}$$

We call this the hyperbolic ensemble, since for $n = 2$, entries share the same L-value in the upper triangle of the matrix if they lie along the same hyperbola. Although higher moment calculations become very difficult, we are able to calculate the fourth moment for $n = 2$ and $\alpha = \beta$ and show that it deviates from the corresponding semicircle moment.

3.1.1. *Odd Moments.* Because entries that share L-values lie along hyperbolas, $\Delta(L)$ is at most one. Odd moments are then zero, because these matrices satisfy Property B.

3.1.2. *Zeroth and Second Moments.* By the same calculation that applied in the generalized Toeplitz case, $M_0 = M_2 = 1$.

3.1.3. *Fourth Moment.* Following the same procedure that applied in generalized Toeplitz case, we first count the number of circuits for the word *aabb*. Since the link function satisfies the required conditions in Lemma 2.3, all Catalan words contribute one.

Now consider the word *abab*. When $\alpha \neq \beta$, obstructions arise that lead to a loss in degrees of freedom, by the same reasoning as in the Toeplitz case. However, when $\alpha = \beta$, there is the possibility of a nonzero contribution. We calculate this for $n = 2$. Interestingly, the contribution is less than the corresponding Toeplitz contribution, $\frac{2}{3}$. Since pairs must be in opposite zones, we are left with four possibilities for matchings in the word *abab*, which we integrate using similar methods as in the generalized Toeplitz case:

(1) $a_{i_1i_2} \in$ Zone 1, $a_{i_3i_4} \in$ Zone 2, $a_{i_2i_3} \in$ Zone 1, and $a_{i_4i_1} \in$ Zone 2:
$\pi(0)^2 - \pi(1)^2 = -\pi(2)^2 + \pi(3)^2$ and $\pi(1)^2 - \pi(2)^2 = -\pi(3)^2 + \pi(0)^2 \longrightarrow$
$\int_0^1 \int_0^1 \int_0^1 \mathbb{I}(0 \leq \sqrt{v_0^2 - v_1^2 + v_2^2} \leq 1$ and $v_0 < v_1$ and $v_1 < v_2$ and
$v_0 < \sqrt{v_0^2 - v_1^2 + v_2^2}$ and $\sqrt{v_0^2 - v_1^2 + v_2^2} < v_2) dv_0 dv_1 dv_2 = \frac{1}{6}$

(2) $a_{i_1i_2} \in$ Zone 1, $a_{i_3i_4} \in$ Zone 2, $a_{i_2i_3} \in$ Zone 2, and $a_{i_4i_1} \in$ Zone 1:
$\pi(0)^2 - \pi(1)^2 = -\pi(2)^2 + \pi(3)^2$ and $-\pi(1)^2 + \pi(2)^2 = \pi(3)^2 - \pi(0)^2 \longrightarrow$
$\int_0^1 \int_0^1 \int_0^1 \mathbb{I}(0 \leq \sqrt{v_0^2 - v_1^2 + v_2^2} \leq 1$ and $v_0 < v_1$ and $v_1 > v_2$ and
$v_0 > \sqrt{v_0^2 - v_1^2 + v_2^2}$ and $\sqrt{v_0^2 - v_1^2 + v_2^2} < v_2) dv_0 dv_1 dv_2 = \frac{4-\pi}{12}$

(3) $a_{i_1i_2} \in$ Zone 2, $a_{i_3i_4} \in$ Zone 1, $a_{i_2i_3} \in$ Zone 1, and $a_{i_4i_1} \in$ Zone 2:
$-\pi(0)^2 + \pi(1)^2 = \pi(2)^2 - \pi(3)^2$ and $\pi(1)^2 - \pi(2)^2 = -\pi(3)^2 + \pi(0)^2 \longrightarrow$
$\int_0^1 \int_0^1 \int_0^1 \mathbb{I}(0 \leq \sqrt{v_0^2 - v_1^2 + v_2^2} \leq 1$ and $v_0 > v_1$ and $v_1 < v_2$ and
$v_0 < \sqrt{v_0^2 - v_1^2 + v_2^2}$ and $\sqrt{v_0^2 - v_1^2 + v_2^2} > v_2) dv_0 dv_1 dv_2 = \frac{\mathrm{Log}(2)}{3}$

(4) $a_{i_1i_2} \in$ Zone 2, $a_{i_3i_4} \in$ Zone 1, $a_{i_2i_3} \in$ Zone 2, and $a_{i_4i_1} \in$ Zone 1:
$-\pi(0)^2 + \pi(1)^2 = \pi(2)^2 - \pi(3)^2$ and $-\pi(1)^2 + \pi(2)^2 = \pi(3)^2 - \pi(0)^2 \longrightarrow$
$\int_0^1 \int_0^1 \int_0^1 \mathbb{I}(0 \leq \sqrt{v_0^2 - v_1^2 + v_2^2} \leq 1$ and $v_0 > v_1$ and $v_1 > v_2$ and
$v_0 > \sqrt{v_0^2 - v_1^2 + v_2^2}$ and $\sqrt{v_0^2 - v_1^2 + v_2^2} > v_2) dv_0 dv_1 dv_2 = \frac{1}{6}$.

Therefore, for $\alpha = \beta$ and $n = 2$,

$$M_4 = 2 + \frac{8 - \pi + 2\mathrm{Log}(4)}{12} < \frac{8}{3} \tag{3.2}$$

and

$$M_4 \neq C_4. \tag{3.3}$$

3.2. **Elliptical Matrices.** We can further generalize the Hankel matrices by raising the variables in the link function to powers higher than one. Here we will raise both indices to the same power by introducing the parameter, n, giving us a link function for the *elliptic matrices*. For $\alpha,\ \beta \in \mathbb{Q}^+$ and $n > 1$,

$$L_{\alpha,\beta}(i,j) = \begin{cases} \alpha i^n + \beta j^n & i \leq j \\ \\ \beta i^n + \alpha j^n & i > j. \end{cases} \tag{3.4}$$

We call this the elliptical ensemble, since for $n = 2$, entries share the same L-value in the upper triangle of the matrix if they lie along the same ellipse. Although higher moment calculations become very difficult, we are able to calculate the fourth moment for $n = 2$ and any α and β and show that it is equal to the corresponding semicircle moment.

3.2.1. *Odd Moments.* Because entries that share L-values lie along ellipses, $\Delta(L)$ is at most one. Odd moments are then zero, because these matrices satisfy Property B.

3.2.2. *Zeroth and Second Moments.* $M_0 = M_2 = 1$ by the same calculation that applied generalized Toeplitz case.

3.2.3. *Fourth Moment.* Following the same procedure that applied in the generalized Toeplitz case, we first count the number of circuits for the word *aabb*. The link function satisfies the requirements of Lemma 2.3, so all Catalan words contribute one. Now we consider versions of the word *abab*. When $\alpha \neq \beta$, obstructions arise that lead to a loss in degrees of freedom. When $\alpha = \beta$, the contribution is zero via previous integration techniques. Therefore,

$$M_4(\alpha, \beta) = 2. \tag{3.5}$$

It would require calculating higher moments to determine if the limiting spectral distribution is semicircular or non-semicircular.

4. Polynomials with Different Powers

4.1. **Parabolic Toeplitz Matrices.** We will further generalize the Toeplitz matrices by expanding the variables in their link function using two polynomials of different order. Let $p_1(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0$ for some $m \in \mathbb{R}^+$ and $p_2(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0$ for some $n \in \mathbb{R}+$, where all the coefficients in both polynomials are non-negative, the leading order coefficients are both non-zero, and $m \neq n$. The link function for the *parabolic Toeplitz matrices* is:

$$L(i,j) = \begin{cases} p_1(i) - p_2(j) & i \leq j \\ -p_2(i) + p_1(j) & i > j. \end{cases} \tag{4.1}$$

We call this the parabolic Toeplitz ensemble since, for $m = 2$ and $n = 1$, entries share the same L-value in the upper triangle of the matrix if they lie along the same parabola. In the following sections, we will prove that matrices with this link function have a semicircular limiting spectral distribution.

Proposition 4.1 (Limiting Spectral Distribution of Parabolic Toeplitz Matrices.) *For matrices with the parabolic Toeplitz link function, $M_{2k} = C_{2k}$ for every positive integer k, and the limiting spectral distribution of the ensemble is the semicircle.*

Before proving most general case, we will first assume that only the coefficients of the leading order term in each polynomial are non-zero:

$$L_{\alpha,\beta}(i,j) = \begin{cases} \alpha i^m - \beta j^n & i \leq j \\ \\ -\beta i^n + \alpha j^m & i > j. \end{cases} \tag{4.2}$$

A 5×5 matrix with $\alpha = \beta$, $m = 2$, and $n = 1$ would have the structure

$$A = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & a_{-3} & a_{-4} \\ a_{-1} & a_2 & a_1 & a_0 & a_{-1} \\ a_{-2} & a_1 & a_6 & a_5 & a_4 \\ a_{-3} & a_0 & a_5 & a_{12} & a_{11} \\ a_{-4} & a_{-1} & a_4 & a_{11} & a_{20} \end{pmatrix}. \tag{4.3}$$

We now show that this link function yields a semicircular limiting spectral distribution.

4.1.1. *Odd Moments.* Fix a row and column in the upper triangle of the matrix to the get L-value $\alpha i^m - \beta j^n$. Any other column in that row will produce a change in j and thus a change in the L-value. $\Delta(L)$ is at most one, then, and odd moments are zero because these matrices satisfy Property B.

4.1.2. *Zeroth and Second Moments.* By the same calculation that applied in Toeplitz case, $M_0 = M_2 = 1$.

4.1.3. *Catalan Words.* Since the link function satisfies the conditions in Lemma 2.3, every Catalan word of length $2k$ contributes one to the $2k^{th}$ moment.

4.1.4. *Crossed Words.* To prove that the limiting spectral measure is a semicircle, it suffices to show that all non-Catalan words contribute zero. Specifically, we will show that the contribution of a word that is fully crossed, or a word that has no adjacent pairs, is zero. This shows that all non-Catalan words contribute zero, since the contribution of any non-Catalan word with adjacent pairs is calculated by adjacent "lifting", and any loss in degrees of freedom in lower moments will propagate through these adjacent pairs. Again, we can use adjacent "lifting" because all adjacent pairs must be in opposite zones by Lemma 2.2.

Any fully crossed matching will have non-adjacent pairs of the form $a_{i_1 i_2} = a_{i_3 i_4}$. For any such pair, there are four possible sets of zones for the matching, each of which will result in one index with coefficient α and power m on each side of the L-value equation and one index with coefficient $-\beta$ and power n on each side of the equation. Without loss of generality, then, we will assume that $a_{i_1 i_2} \in$ Zone 1 and $a_{i_3 i_4} \in$ Zone 1. We will also assume, without loss of generality, that $m > n$. The corresponding L-value equation has the form:

$$\alpha\pi(0)^m - \beta\pi(1)^n = \alpha\pi(2)^m - \beta\pi(3)^n. \tag{4.4}$$

In the following argument, we also want to choose the entries $a_{i_1 i_2}$ and $a_{i_3 i_4}$ to be specific letters in the fully crossed word. This argument follows as in Lemma 2.6. We will pick this matched pair such that all letters between the letters that correspond to these entries are distinct. Pick a matched pair, then check if there are any duplicated letters between these matched letters. If there are duplicated letters, then choose the entries corresponding to those letters to be the matched pair. Continue this process until all the intermediate letters are distinct. There must be at least one distinct letter between the matched letters remaining; otherwise, the word would have an adjacent pair. For the resulting pair, choose the first index to be the generating vertex $\pi(0)$. Then the next vertex, $\pi(1)$, is also a generating vertex, as it now corresponds to the first letter in the word. Although we cannot know in general the location of the other vertices, we will just name them $\pi(2)$ and $\pi(3)$. Since intermediate letters are distinct and the word began with generating vertex $\pi(0)$, $\pi(2)$ must correspond to the first occurrence of a letter in the word, making $\pi(2)$ another generating vertex. Therefore, there are at least three generating vertices in this pair. Let's say that there are x degrees of freedom in choosing all four vertices in this matching. There were initially $k+1$ generating vertices, but now there are at most $k-2$ generating vertices remaining and at most $x+k-2$ degrees of freedom in total. If we can show that $x < 3$, then any non-Catalan word contributes zero to its moment, since there must be $k+1$ degrees of freedom for a contribution.

First, rearrange the L-value equation to get:

$$\alpha\left(\pi(0)^m - \pi(2)^m\right) = \beta\left(\pi(1)^n - \pi(3)^n\right). \tag{4.5}$$

Although we have chosen the entries to be in specific zones, we will relax these restrictions and assume that the only constraint is from the L-value equation. First, assume that $\pi(1) \geq \pi(3)$. Since vertices are at least one and at most N, $\pi(0)$ and $\pi(2)$ must be chosen so that

$$0 \leq \pi(0)^m - \pi(2)^m \leq \frac{\beta}{\alpha}\left(N^n - 1\right). \tag{4.6}$$

Otherwise, there will be no valid choices for $\pi(1)$ and $\pi(3)$. Again, we will relax the upper bound and simply impose

$$0 \leq \pi(0)^m - \pi(2)^m \leq \frac{\beta}{\alpha}N^n. \tag{4.7}$$

Choose $\pi(1)$, which must be between 1 and N. Then $\pi(2)$ is constrained by

$$\begin{cases} \left(\pi(0)^m - \frac{\alpha}{\beta}N^n\right)^{\frac{1}{m}} \leq \pi(2) \leq \pi(0) & \text{if } \left(\pi(0)^m - \frac{\alpha}{\beta}N^n\right)^{\frac{1}{m}} > 1 \\ 1 \leq \pi(2) \leq \pi(0) & \text{if } \left(\pi(0)^m - \frac{\alpha}{\beta}N^n\right)^{\frac{1}{m}} \leq 1. \end{cases} \tag{4.8}$$

Again, we will relax the constraint without artificially restricting degrees of freedom by choosing the second lower bound above to be zero:

$$\begin{cases} \left(\pi(0)^m - \frac{\alpha}{\beta}N^n\right)^{\frac{1}{m}} \leq \pi(2) \leq \pi(0) & \text{if } \left(\pi(0)^m - \frac{\alpha}{\beta}N^n\right)^{\frac{1}{m}} > 1 \\ 0 \leq \pi(2) \leq \pi(0) & \text{if } \left(\pi(0)^m - \frac{\alpha}{\beta}N^n\right)^{\frac{1}{m}} \leq 1. \end{cases} \tag{4.9}$$

The number of valid choices for $\pi(0)$ and $\pi(2)$ is then given by:

$$\sum_{\pi(0)=1}^{N} \sum_{\pi(2)=1}^{N} \mathbb{I}\left(\pi(2) \leq \pi(0) \text{ and } \pi(2) \geq \left(\pi(0)^m - \frac{\alpha}{\beta}N^n\right)^{\frac{1}{m}}\right). \tag{4.10}$$

For convenience we will split the outer sum at $\left(\frac{\alpha}{\beta}\right)^{\frac{1}{m}} N^{\frac{n}{m}}$:

$$\sum_{\pi(0)=1}^{\left(\frac{\alpha}{\beta}\right)^{\frac{1}{m}} N^{\frac{n}{m}}} \sum_{\pi(2)=0}^{\pi(0)} 1 + \sum_{\pi(0)=\left(\frac{\alpha}{\beta}\right)^{\frac{1}{m}} N^{\frac{n}{m}}}^{N} \sum_{\pi(2)=\left(\pi(0)^m - \frac{\alpha}{\beta}N^n\right)^{\frac{1}{m}}}^{\pi(0)} 1. \tag{4.11}$$

Since we only care about very large matrices, we will let $N \to \infty$. In this limit, the sums become integrals:

$$\int_{\pi(0)=1}^{\left(\frac{\alpha}{\beta}\right)^{\frac{1}{m}} N^{\frac{n}{m}}} \int_{\pi(2)=0}^{\pi(0)} d_{\pi(0)} d_{\pi(2)} + \int_{\pi(0)=\left(\frac{\alpha}{\beta}\right)^{\frac{1}{m}} N^{\frac{n}{m}}}^{N} \int_{\pi(2)=\left(\pi(0)^m - \frac{\alpha}{\beta} N^n\right)^{\frac{1}{m}}}^{\pi(0)} d_{\pi(0)} d_{\pi(2)}. \tag{4.12}$$

From Mathematica, the count is[23]:

$$\begin{aligned} &\frac{\left(\frac{\alpha}{\beta}\right)^{\frac{2}{m}} N^{\frac{2n}{m}}}{2} - \frac{1}{2} + \frac{N^2}{2} - \frac{\left(\frac{\alpha}{\beta}\right)^{\frac{2}{m}} N^{\frac{2n}{m}}}{2} + \frac{2^{\frac{-2}{m}} N^{\frac{2n}{m}} \Gamma\left(\frac{1}{2} - \frac{1}{m}\right) \Gamma\left(1 + \frac{1}{m}\right)}{2\sqrt{\pi}} \\ &- \frac{N^2}{2} {}_2F_1\left[-\frac{2}{m}, -\frac{1}{m}, \frac{m-2}{m}, \frac{\frac{\alpha}{\beta}}{N^{m-n}}\right], \end{aligned} \tag{4.13}$$

where ${}_2F_1[a,b,c,z] = 1$ is the *hypergeometric function*, defined for $|z| < 1$ on Wolfram MathWorld [WMW] by the power series

$${}_2F_1[a,b,c,z] = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n z^n}{(c)_n n!}, \tag{4.14}$$

with

$$(q)_n = \begin{cases} 1 & \text{if } n = 0 \\ q(q+1)\cdots(q+n-1) & \text{if } n > 0. \end{cases} \tag{4.15}$$

To check if there are at least two degrees of freedom, we divide by N^2 and take the limit $N \to \infty$. Since $\frac{2n}{m} < 2$, the first, second, fourth, and fifth terms vanish in the limit. From the power series expansion, we see that for any $a \in \mathbb{C},\ b \in \mathbb{C},\ \text{ and }\ c \in \{\mathbb{C} \backslash (\mathbb{Z}^- \cup \{0\})\}$, ${}_2F_1[a,b,c,0] = 1$. Since $\frac{1}{N^{m-n}} \to 0$, we are left with $\frac{1}{2} - \frac{1}{2}$ = 0.

Therefore, there are fewer than two degrees of freedom when choosing $\pi(0)$ and $\pi(2)$. Choose $\pi(1)$, then $\pi(3)$ is fixed by the L-value equation. There are fewer than three degrees of freedom here. Assuming $\pi(1) < \pi(3)$ gives the same number of solutions, as do all other sets of zones for the pairs. Letting $n > m$ would lead to a similar calculation and the same results. Therefore, $x < 3$ and the limiting spectral distribution is semicircular.

The general proof is motivated by the idea that the leading order terms in the polynomials dominate. We proceed in a similar manner.

[23] We could apply a binomial approximation here and prove the result by hand. We will do this for the general case.

4.1.5. *Odd Moments.* Fix a row and column in the upper triangle of the matrix to get the L-value $p_1(i) - p_2(j)$. Any other column in that row will produce a change in the L-value, because $p_2(j)$ is a monotonically increasing function for non-negative arguments. $\Delta(L)$ is at most one, then, and odd moments are zero because these matrices satisfy Property B.

4.1.6. *Zeroth and Second Moments.* By the same calculation that applied in the generalized Toeplitz case, $M_0 = M_2 = 1$.

4.1.7. *Catalan Words.* Since the link function satisfies conditions in Lemma 2.3, every Catalan word of length $2k$ contributes one to the $2k^{th}$ moment.

4.1.8. *Crossed Words.* To prove that the limiting spectral measure is a semicircle, it suffices to show that all non-Catalan words contribute zero. Following the argument above, we pick matched entries appropriately and count the number of solutions to show that $x < 3$. Without loss of generality, we assume that $a_{i_1 i_2} \in$ Zone 1, $a_{i_3 i_4} \in$ Zone 1, and $m > n$. The relevant L-value equation is then

$$p_1(\pi(0)) - p_2(\pi(1)) = p_1(\pi(2)) - p_2(\pi(3)). \tag{4.16}$$

Rearranging the equation, we have

$$p_1(\pi(0)) - p_1(\pi(2)) = p_2(\pi(1)) - p_2(\pi(3)). \tag{4.17}$$

Although we have chosen the entries to be in specific zones, we will relax these restrictions and assume that the only constraint is from the L-value equation. First, assume that $\pi(1) \geq \pi(3)$, and let $b = \max(b_i)$ for $i \in \{0, n\}$. Since indices are at least one and at most N, $\pi(0)$ and $\pi(2)$ must be chosen so that

$$0 \leq p_1(\pi(0)) - p_1(\pi(2)) \leq b(n+1)N^n - \sum_{i=0}^{n} b_i. \tag{4.18}$$

Otherwise, there will be no valid choices for $\pi(1)$ and $\pi(3)$. Again, we relax the upper bound and simply impose

$$0 \leq p_1(\pi(0)) - p_1(\pi(2)) \leq b(n+1)N^n. \tag{4.19}$$

Choose $\pi(1)$, which must be between 1 and N. Then $\pi(2)$ is constrained by

$$\begin{cases} p_1(\pi(0)) - b(n+1)N^n \leq p_1(\pi(2)) \leq \pi(0) & \text{if } p_1(\pi(0)) - b(n+1)N^n > 1 \\ 1 \leq p_1(\pi(2)) \leq \pi(0) & \text{if } p_1(\pi(0)) - b(n+1)N^n \leq 1. \end{cases} \tag{4.20}$$

Once again, we will relax the constraints without artificially restricting degrees of freedom by choosing the second lower bound to be zero:

$$\begin{cases} p_1(\pi(0)) - b(n+1)N^n \leq p_1(\pi(2)) \leq \pi(0) & \text{if } p_1(\pi(0)) - b(n+1)N^n > 1 \\ 0 \leq p_1(\pi(2)) \leq \pi(0) & \text{if } p_1(\pi(0)) - b(n+1)N^n \leq 1. \end{cases} \tag{4.21}$$

For large N, the number of valid choices for $\pi(0)$ and $\pi(2)$ becomes an integral:

$$\int_1^N \int_1^N \mathbb{I}\left(\pi(2) \leq \pi(0) \text{ and } p_1(\pi(2)) \geq p_1(\pi(0)) - b(n+1)N^n\right) d_{\pi(2)} d_{\pi(0)}. \tag{4.22}$$

Let Π be a placeholder for the integrand and let C be a constant independent of N. We then have

$$\begin{aligned} \int_1^N \int_1^N \Pi d_{\pi(2)} d_{\pi(0)} &= \int_1^C \int_1^N \Pi d_{\pi(2)} d_{\pi(0)} + \int_C^N \int_1^N \Pi d_{\pi(2)} d_{\pi(0)} \\ &\leq \int_1^C \int_0^N \Pi d_{\pi(2)} d_{\pi(0)} + \int_C^N \int_0^C \Pi d_{\pi(2)} d_{\pi(0)} + \int_C^N \int_C^N \Pi d_{\pi(2)} d_{\pi(0)} \\ &\leq N(C-1) + C(N-C) + \int_C^N \int_C^N \Pi d_{\pi(2)} d_{\pi(0)}. \end{aligned} \tag{4.23}$$

Since we are dividing by N^2 and taking the limit of large N, we can ignore the first two terms. For the third term, if $p_1'(x) = \frac{p_1(x)}{a_m}$

$$\begin{aligned} &\int_1^N \int_1^N \Pi d_{\pi(2)} d_{\pi(0)} \\ &\approx \int_C^N \int_C^N \mathbb{I}\left(\pi(2) \leq \pi(0) \text{ and } p_1(\pi(2)) \geq p_1(\pi(0)) - b(n+1)N^n\right) d_{\pi(2)} d_{\pi(0)} \\ &\approx \int_C^N \int_C^{\pi(0)} \mathbb{I}\left(p_1(\pi(2)) \geq p_1(\pi(0)) - b(n+1)N^n\right) d_{\pi(2)} d_{\pi(0)} \\ &\approx \int_C^N \int_C^{\pi(0)} \mathbb{I}\left(p_1'(\pi(2)) \geq p_1'(\pi(0)) - \frac{b(n+1)}{a_m} N^n\right) d_{\pi(2)} d_{\pi(0)}. \end{aligned} \tag{4.24}$$

We prove the following lemma.

Lemma 4.1 (Polynomials Dominated by Leading Term.) *Let* $p(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_0$. *We can choose* x *large enough such that* $(1-\epsilon)x^k < p(x) < (1+\epsilon)x^k$.

Proof. Let $q(x) = a_{k-1}x^{k-1} + \cdots + a_0$. Then $\frac{p(x)-x^k}{x^k} = \frac{q(x)}{x^k}$. Since x^k dominates the leading order terms of $q(x)$ in the limit of large x, we can choose x large enough such that for any $\epsilon' > 0$, $\frac{p(x)-x^k}{x^k} = \epsilon'$. This gives us $p(x) = (1+\epsilon')x^k$. If we let $\epsilon = 2\epsilon'$, we have $(1-\epsilon)x^k < p(x) < (1+\epsilon)x^k$. □

We now expand the region of integration; by Lemma 4.1, for any ϵ there exists a C large enough such that

$$\int_1^N \int_1^N \Pi d_{\pi(2)} d_{\pi(0)} \leq \int_C^N \int_C^{\pi(0)} \mathbb{I}\left((1+\epsilon)\pi(2)^m \geq (1-\epsilon)\pi(0)^m - \frac{b(n+1)}{a_m} N^n\right) d_{\pi(2)} d_{\pi(0)}$$
$$\leq \int_C^N \int_C^{\pi(0)} \mathbb{I}\left(\pi(2) \geq \left(\frac{(1-\epsilon)}{(1+\epsilon)}\pi(0)^m - \frac{b(n+1)}{a_m(1+\epsilon)} N^n\right)^{\frac{1}{m}}\right) d_{\pi(2)} d_{\pi(0)}$$
$$\leq \int_C^N \int_C^{\pi(0)} \mathbb{I}\left(\pi(2) \geq \left((1-2\epsilon+O(\epsilon^2))\pi(0)^m - \frac{b(n+1)}{a_m(1+\epsilon)} N^n\right)^{\frac{1}{m}}\right) d_{\pi(2)} d_{\pi(0)}$$
$$\leq \int_C^{\left(\frac{3b(n+1)N^n}{a_m(1+\epsilon)}\right)^{\frac{1}{m}}} \int_C^{\pi(0)} d_{\pi(2)} d_{\pi(0)} +$$
$$\int_{\left(\frac{3b(n+1)N^n}{a_m(1+\epsilon)}\right)^{\frac{1}{m}}}^N \int_C^{\pi(0)} \mathbb{I}\left(\pi(2) \geq \left((1-2\epsilon+O\left(\epsilon^2\right))\pi(0)^m - \frac{b(n+1)}{a_m(1+\epsilon)} N^n\right)^{\frac{1}{m}}\right)$$
$$d_{\pi(2)} d_{\pi(0)}. \tag{4.25}$$

The first term becomes negligible when we divide by N^2 and take the limit

$$\int_C^{\left(\frac{3b(n+1)N^n}{a_m(1+\epsilon)}\right)^{\frac{1}{m}}} \int_C^{\pi(0)} d_{\pi(2)} d_{\pi(0)}$$
$$= \frac{(3b)^{\frac{2}{m}}(n+1)^{\frac{2}{m}}}{2(a_m)^{\frac{2}{m}}(1+\epsilon)^{\frac{1}{m}}} N^{\frac{2n}{m}} - \frac{C\left(3b(n+1)\right)^{\frac{1}{m}}}{\left(a_m(1+\epsilon)\right)^{\frac{1}{m}}} N^{\frac{n}{m}} - \frac{3C^2}{2}, \tag{4.26}$$

since $\frac{2n}{m} < 2$. For the second term, we use the indicator function to define a lower bound for the inner integral, integrate that integral, and pull out a factor of $\pi(0)^m$ from the resulting second term:

$$\int_{\left(\frac{3b(n+1)N^n}{a_m(1+\epsilon)}\right)^{\frac{1}{m}}}^N \pi(0) - \pi(0)\left((1-2\epsilon+O\left(\epsilon^2\right)) - \frac{b(n+1)}{a_m(1+\epsilon)\pi(0)^m} N^n\right)^{\frac{1}{m}} d_{\pi(2)} d_{\pi(0)}. \tag{4.27}$$

From the lower bound of $\pi(0)$, the term raised to power $\frac{1}{m}$ is at most $1-\frac{1}{3}-2\epsilon+O\left(\epsilon^2\right)$. We can therefore apply a binomial expansion and just keep the leading order term:

$$\int_{\left(\frac{3b(n+1)N^n}{a_m(1+\epsilon)}\right)^{\frac{1}{m}}}^{N} \pi(0)-\pi(0)\left(1-\frac{2\epsilon}{m}+\frac{O\left(\epsilon^2\right)}{m}-\frac{b(n+1)}{ma_m(1+\epsilon)\pi(0)^m}N^n\right)d_{\pi(2)}d_{\pi(0)}. \tag{4.28}$$

Since $m>n$ the only resulting terms that appear to depend on N^2 are

$$\frac{N^2\epsilon}{m}-\frac{N^2O\left(\epsilon^2\right)}{2m}, \tag{4.29}$$

but we can choose C large enough that $\epsilon\to 0$. So,

$$\lim_{N\to\infty}\frac{1}{N^2}\int_1^N\int_1^N \Pi d_{\pi(2)}d_{\pi(0)}=0. \tag{4.30}$$

Choose $\pi(0)$ and $\pi(2)$ with fewer than two degrees of freedom. Choose $\pi(1)$, then $\pi(3)$ is fixed by the L-value equation, and there are fewer than three degrees of freedom. Letting $\pi(1)<\pi(3)$ yields the same results. Choosing any other set of zones for the matched pair will also yield the same results, as will allowing $n>m$. Therefore $x<3$, and the limiting spectral distribution is semicircular.

4.1.9. *Convergence.* By the arguments in section 1.2.5, Var$[M_k(A_N)]\to 0$ and the empirical spectral distributions converge in probability to the semicircle. By the arguments in section 1.2.6, the empirical distributions converge almost surely to the semicircle. All of the above arguments only depend on $p(x)$ having mean zero, variance one, and uniformly bounded moments of all order. Hence, the convergence is universal.

4.2. **Parabolic Hankel Matrices.** We can likewise generalize the Hankel matrices by expanding the variables in their link function with polynomials of different order. Let $p_1(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_0$ for some $m\in\mathbb{R}^+$ and $p_2(x)=b_nx^n+b_{n-1}x^{n-1}+\cdots+b_0$ for some $n\in\mathbb{R}+$, where all the coefficients in both polynomials are non-negative, the leading order coefficients are both non-zero, and $m\neq n$. The link function for the *parabolic Hankel matrices* is given by

$$L(i,j)=\begin{cases} p_1(i)+p_2(j) & i\leq j\\ p_2(i)+p_1(j) & i>j.\end{cases} \tag{4.31}$$

We call this the parabolic Hankel ensemble since, for $m=2$ and $n=1$, entries share the same L-value in the upper triangle of the matrix if they lie along the same parabola. In the following sections, we prove that matrices with this link function have a semicircular limiting spectral distribution.

Proposition 4.2(Limiting Spectral Distribution of Parabolic Hankel Matrices.) *For matrices with the parabolic Hankel link function, $M_{2k} = C_{2k}$ for every positive integer k, and the limiting spectral distribution of the ensemble is the semicircle.*

4.2.1. *Odd Moments.* Fix a row and column in the upper triangle of the matrix to get the L-value $p_1(i) + p_2(j)$. Any other column in that row will produce a change in the L-value because $p_2(j)$ is a monotonically increasing function for non-negative arguments. $\Delta(L)$ is at most one, then, and odd moments are zero because these matrices satisfy Property B.

4.2.2. *Zeroth and Second Moments.* By the same calculation in the Toeplitz case, $M_0 = M_2 = 1$.

4.2.3. *Catalan Words.* Since the link function satisfies the conditions in Lemma 2.3, every Catalan word contributes one to its corresponding moment.

4.2.4. *Crossed Words.* To prove that the limiting spectral measure is a semicircle, it suffices to show that all non-Catalan words contribute zero. Following the same argument as above, we pick matched entries appropriately and count the number of solutions to show that $x < 3$. Without loss of generality, we assume that $a_{i_1 i_2} \in$ Zone 1, $a_{i_3 i_4} \in$ Zone 1, and $m > n$. The relevant L-value equation is then

$$p_1\left(\pi(0)\right) + p_2\left(\pi(1)\right) = p_1\left(\pi(2)\right) + p_2\left(\pi(3)\right). \tag{4.32}$$

Rearranging the equation, we have

$$p_1\left(\pi(0)\right) - p_1\left(\pi(2)\right) = p_2\left(\pi(3)\right) - p_2\left(\pi(1)\right). \tag{4.33}$$

At this point, the methods from the previous section apply directly and the proof follows.

4.2.5. *Convergence.* By the arguments in section 1.2.5, Var$[M_k(A_N)] \to 0$ and the empirical spectral distributions converge in probability to the semicircle. By the arguments in section 1.2.6, the empirical distributions converge almost surely to the semicircle. All of the above arguments only depend on $p(x)$ having mean zero, variance one, and uniformly bounded moments of all order. Hence, the convergence is universal.

5. Future Research

We have shown that several bivariate polynomial link functions in which the variables in the link function are raised to the same power have a non-semicircular limiting spectral distribution. Although computing higher moments became intractable, we were able to show the dependence of the moments on the link function parameters, α and β. We also proved that for any bivariate polynomial link function in which the polynomials in the two variables are raised to different powers and all coefficients are non-negative, the limiting spectral distribution is semicircular.

For the generalized Toeplitz, generalized Hankel, and hyperbolic matrices, future work includes developing a closed-form expression for all higher moments, analyzing rates of convergence, and developing sharper moment bounds. For the elliptic ensemble, it would be interesting to directly compute the sixth moment and higher moments to check whether the limiting distribution is non-semicircular. Since integration over the appropriate regions becomes difficult, it might be worthwhile to recast the problem as counting solutions to Diophantine equations or summing over lattice points inside a specified region, instead of integration. In addition, it would be worthwhile to make Proposition 4.1 as general as possible. What can be said about the limiting spectral distribution for a link function composed of different-order polynomials that do not necessarily have all non-negative coefficients? What can be said in general about link functions composed of polynomials of the same order? Researching these questions would not only help describe why only certain polynomial link functions have a semicircular limiting spectral distribution, but also illuminate the general problem of determining to what extent patterned random matrices maintain the semicircular distribution of the original Wigner matrices.

Many of the desired properties for the link functions studied in this thesis, such as odd moments being zero and almost sure convergence to the limiting spectral distribution, follow from Property B. An interesting question to ask, then, is what happens to link functions that do not satisfy Property B? This property requires that as a matrix grows to infinite size, the maximum number of repetitions of the same random variable in any row or column is finite. We might consider breaking this requirement in the weakest way possible. For example, the repetitions of a random variable might grow to infinity much slower than N. Consider, for example, the first row of a random matrix to be described by the sequence of random variables,$\{a_1 a_2 a_1 a_2 a_3 a_1 a_2 a_3 a_4 \cdots\}$. Pushing the bounds of Property B would similarly illuminate the general problem of describing bulk eigenvalue distributions for patterned random matrices.

6. Numerical Simulations

We performed numerical simulations of limiting spectral distributions in Mathematica. Random numbers were generated from a standard normal distribution. The code allows for computing the scaled eigenvalues of any number of matrices of arbitrary size, printing a historam of these normalized eigenvalues, and printing the average k^{th} moment for the

```
create[size_, α_, β_] := Do[b[i] = RandomVariate[NormalDistribution[0, 1]], {i, α * size + β * size}];
entry[i1_, j1_, α_, β_] := b[α * i1 + β * j1];
matrix[size_, α_, β_] := Table[If[i1 < j1, entry[i1, j1, α, β], entry[j1, i1, α, β]], {i1, size}, {j1, size}];
computation[size_, numberiterations_, α_, β_] :=
  Do[m1 = {}; m2 = {}; Do[create[size, α, β]; symmetricmatrix = matrix[size, α, β]; For[k = 1, k < size + 1, k++, AppendTo[m1, Eigenvalues[symmetricmatrix / (2 Sqrt[size])][[k]]]];
    AppendTo[m2, 1/(2^4 size^(4/2 + 1)) Tr[MatrixPower[symmetricmatrix, 4]]], {numberiterations}]; Print["Average Fourth Moment" Total[m2]/Length[m2]]; Print[α, β];
   Print[Histogram[m1, 50, "Probability", PlotLabel → "Histogram of Normalized Eigenvalues", AxesLabel → {"Normalized Eigenvalues", "Probability"}]], {1}];
fullcomputation[size_, numberiterations_, numberparameters_] :=
 Do[computation[size, numberiterations, 1, 0]; computation[size, numberiterations, 0, 1];
  Do[computation[size, numberiterations, α, β], {α, 1, numberparameters}, {β, 1, numberparameters}], {1}]
```

FIGURE 11. Sample Mathematica code for limiting spectral distribution and average k^{th} moment simulations.

matrices. Sample code is shown in Figure 11. The link function appears as $b[\alpha \times i_1 + \beta \times j_1]$.

7. APPENDIX

7.1. Scaling of Random Matrix Eigenvalues. We give a heuristic for the eigenvalues of our $N \times N$ matrix ensembles being roughly of size $\sqrt{N}$. Let a matrix be denoted A_N whose entries a_{ij} are randomly and independently chosen from a fixed probability distribution $p(x)$ with mean 0 and variance 1. For real symmetric matrices, $A_N = A_N^T$, and

$$\text{Trace}(A_N^2) = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} a_{ji} = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}^2. \tag{7.1}$$

From our assumptions on $p(x)$, we expect each a_{ij}^2 to be of size 1. Therefore, we expect that

$$\sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}^2 \approx N^2. \tag{7.2}$$

Thus by Lemma 1.1, we have

$$\sum_{i=1}^{N} \lambda_i(A)^2 \approx N^2. \tag{7.3}$$

Taking the expected value, we get

$$N\text{Ave}(\lambda_i^2(A_N)) \approx N^2 \tag{7.4}$$

For heuristic purposes we pass the square root through the average to get

$$|\text{Ave}(\lambda_i(A_N))| \approx \sqrt{N} \tag{7.5}$$

A more precise argument would show that the scaling factor to normalize the eigenvalues to size one is $2\sqrt{N}$. Although we have chosen to keep $\sqrt{N}$ as the normalization, it would be just as effective to scale them by $2\sqrt{N}$. For a scaling of $\sqrt{N}$, real symmetric matrices have a semicircular limiting spectral distribution given by

$$f_{\text{Wigner}}(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2} & |x| \leq 2 \\ \\ 0 & \text{otherwise} \end{cases},$$

whose moments are exactly the Catalan numbers. For a scaling of $2\sqrt{N}$, real symmetric matrices have a semicircular limiting spectral distribution given by

$$f_{\text{Wigner}}(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2} & |x| \leq 1 \\ \\ 0 & \text{otherwise} \end{cases},$$

whose moments are proportional to the Catalan numbers. Either choice of scaling is fine.

7.2. **Riesz's Condition.** The moments to which the expected values converge determine a unique distribution if these limiting moments satisfy a certain conditions. Let $\{M_k\}_{k=1}^{\infty}$ be the sequence of moments for the limiting spectral distribution F. Then, F is the unique distribution with these moments if the following holds:

$$\lim_{k\to\infty}\inf\frac{1}{k}M_{2k}^{\frac{1}{2k}} < \infty. \tag{7.6}$$

See Bose for a detailed proof.

7.3. **Full Calculation of Eq. 2.39.**

$$\int_0^1\int_0^1\int_0^1\int_0^1 \mathbb{I}(0 \leq v_0 - \frac{v_1}{a} + \frac{v_3}{a} \leq 1 \text{ and } 0 \leq av_0 - av_2 + v_3 \leq 1 \text{ and } v_0 > v_1 \text{ and } v_1 < v_2 \text{ and } v_2 > v_3 \text{ and } v_3 < v_0 - \frac{v_1}{a} + \frac{v_3}{a} \text{ and } v_0 - \frac{a_1}{a} + \frac{v_3}{a} > av_0 - av_2 + v_3 \text{ and } av_0 - av_2 + v_3 \leq v_0)dv_1dv_2dv_3dv_0. \tag{7.7}$$

First, choose v_0, v_2, v_3. Then, the above integral reduces to:

$$\int_0^1 \int_0^1 \int_0^1 \mathbb{I}(0 \le av_0 - av_2 + v_3 \le 1 \text{ and } av_0 - av_2 + v_3 \le v_0) dv_0 dv_2 dv_3$$
$$\int_0^1 \mathbb{I}(v_1 \le av_0 + v_3 \text{ and } v_1 \ge v_3 + av_0 - a \text{ and } v_1 < v_0 \text{ and } v_1 < v_2 \text{ and}$$
$$v_1 < av_0 + v_3(1-a) \text{ and } v_1 < v_0 a(1-a) + v_3(1-a) + a^2 v_2 \tag{7.8}$$

$$= \int_0^1 \int_0^1 \int_0^1 \mathbb{I}(0 \le av_0 - av_2 + v_3 \le 1 \text{ and } av_0 - av_2 + v_3 \le v_0)$$
$$\int_{\max(v_3+av_0-a,0)}^{\min(av_0+v_3, v_0, v_2, av_0+v_3(1-a), v_0 a(1-a)+v_3(1-a)+a^2 v_2)} dv_1 dv_2 dv_3 dv_0.$$

By brute force, this integral can be calculated by splitting it into six regions:

1. Let $v_0 < v_2$ and $v_0 < v_3$ and $v_3 + av_0 - a < 0$:

$$\int_0^1 \int_0^1 \int_0^1 \mathbb{I}(0 \leq av_0 - av_2 + v_3 \leq 1 \text{ and } av_0 - av_2 + v_3 \leq v_0)$$
$$\int_0^{av_0+(1-a)v_3} dv_1 dv_2 dv_3 dv_0$$
$$= \int_0^1 \int_0^1 \int_0^1 \mathbb{I}(v2 \leq v_0 + \frac{v_3}{a} \text{ and } v_2 \geq v_0 + \frac{v_3}{a} - \frac{1}{a} \text{ and } v_2 > v_3 \text{ and } v_2 > \frac{v_0(1-a)}{a}$$
$$+ \frac{v_3}{a} \text{ and } av_0 + (1-a)v_3 \geq 0 \text{ and } v_2 \geq \frac{v_0(1-a)}{a} + \frac{v_3(a-1)}{a^2} \text{ and } v_0 < v_2 \text{ and } v_0 < v_3$$
$$\text{and } v_3 + av_0 - a < 0, av_0 + (1-a)v_3, 0) dv_2 dv_3 dv_0$$
$$= \int_0^1 \int_0^1 \mathbb{I}(0 \leq av_0 - av_2 + v_3 \leq 1 \text{ and } av_0 - av_2 + v_3 \leq v_0)$$
$$\int_{\max(v_0+\frac{v_3}{a}-\frac{1}{a}, v_3, \frac{(a-1)v_0}{a}+\frac{v_3}{a}, \frac{(a-1)v_0}{a}+\frac{(a-1)v_3}{a^2}, v_0, 0)}^{\min(v_0+\frac{v_3}{a}, 1)} av_0 + (1-a)v_3 dv_2 dv_3 dv_0$$
$$= \int_0^1 \int_0^1 \mathbb{I}(0 \leq av_0 - av_2 + v_3 \leq 1 \text{ and } av_0 - av_2 + v_3 \leq v_0)$$
$$\int_{v_3}^{v_0+\frac{v_3}{a}} av_0 + (1-a)v_3 dv_2 dv_3 dv_0$$
$$= \int_0^1 \int_0^1 \mathbb{I}(v_3 \leq \frac{av_0}{a-1} \text{ and } v_3 > v_0 \text{ and } v_3 < a - av_0,\ (av_0 + (1-a)v_3)(v_0 + \frac{v_3}{a})$$
$$- (av_0 + (1-a)v_3)v_3, 0) dv_3 dv_0$$
$$= \int_0^1 \int_{\max(v_0,0)}^{\min(\frac{av_0}{a-1}, a-av_0, 1)} \frac{(a(v_0 - v_3) + v_3)^2}{a}$$
$$= \int_0^{\frac{a-1}{a}} \int_{v_0}^{\frac{av_0}{a-1}} \frac{(a(v_0 - v_3) + v_3)^2}{a} + \int_{\frac{a-1}{a}}^{\frac{a}{a+1}} \int_{v_0}^{a-av_0} \frac{(a(v_0 - v_3) + v_3)^2}{a}. \tag{7.9}$$

This reduces to

$$= \frac{a}{12(1+a)^3}. \tag{7.10}$$

2. Let $v_0 < v_2$ and $v_0 < v_3$ and $v_3 + av_0 - a > 0$:

$$\int_0^1 \int_0^1 \int_0^1 \mathbb{I}(0 \le av_0 - av_2 + v_3 \le 1 \text{ and } av_0 - av_2 + v_3 \le v_0) \int_{v_3+av_0-a}^{av_0+(1-a)v_3} dv_1 dv_2 dv_3 dv_0$$
$$= \int_0^1 \int_0^1 \int_0^1 If(v_2 \le v_0 + \frac{v_3}{a} \text{ and } v_2 \ge v_0 + \frac{v_3}{a} - \frac{1}{a} \text{ and } v_2 > v_3 \text{ and } v_2 > \frac{(a-1)v_0}{a} +$$
$$\frac{v_3}{a} \text{ and } av_0 + (1-a)v_3 \ge 0 \text{ and } v_2 \ge \frac{(a-1)v_0}{a} + \frac{(a-1)v_3}{a^2} \text{ and } v_0 < v_2 \text{ and } v_0 < v_3 \text{ and }$$
$$v_3 + av_0 - a > 0, av_0 + (1-a)v_3 - (v_3 + av_0 - a), 0) dv_2 dv_3 dv_0$$
$$= \int_0^1 \int_0^1 \mathbb{I}(0 \le av_0 - av_2 + v_3 \le 1 \text{ and } av_0 - av_2 + v_3 \le v_0)$$
$$\int_{\max(v_0+\frac{v_3}{a}-\frac{1}{a}, v_3, \frac{(a-1)v_0}{a}+\frac{v_3}{a}, \frac{(a-1)v_0}{a}+\frac{(a-1)v_3}{a^2}, v_0, 0)}^{\min(v_0+\frac{v_3}{a}, 1)} a - av_3 dv_3 dv_0$$
$$= \int_0^1 \int_0^1 \mathbb{I}(v_0 < v_3 \text{ and } v_0 \ge \frac{(a-1)v_3}{a} \text{ and } v_0 > 1 - \frac{v_3}{a}, (a - av_3) - (a - av_3)v_3, 0) dv_3 dv_0$$
$$= \int_0^1 \int_{max(v_0, a-av_0, 0)}^{min(\frac{av_0}{a-1}, 1)} a(v_3 - 1)^2 dv_3 dv_0$$
$$= \int_{\frac{a-1}{a}}^{\frac{a}{a+1}}] \int_{a-av_0}^1 a(v_3 - a)^2 dv_3 dv_0 + \int_{\frac{a}{a+1}}^1 \int_{v_0}^1 a(v_3 - a)^2 dv_3 dv_0. \tag{7.11}$$

This reduces to

$$= \frac{1}{12(1+a)^3}. \tag{7.12}$$

3. Let $v_0 < v_2$ and $v_0 > v_3$ and $v_3 = av_0 - a < 0$:

$$
\begin{aligned}
&\int_0^1 \int_0^1 \int_0^1 \mathbb{I}(0 \leq av_0 - av_2 + v_3 \leq 1 \text{ and } av_0 - av_2 + v_3 \leq v_0) \int_0^{v_0} dv_1 dv_2 dv_3 dv_0 \\
&\int_0^1 \int_0^1 \int_0^1 If(v_2 \leq v_0 + \frac{v_3}{a} \text{ and } v_2 \geq v_0 + \frac{v_3}{a} - \frac{1}{a} \text{ and } v_2 > v_3 \text{ and } v_2 > \frac{(a-1)v_0}{a} \\
&+ \frac{v_3}{a} \text{ and } av_0 + (1-a)v_3 \geq 0 \text{ and } v_2 \geq \frac{(a-1)v_0}{a} + \frac{(a-1)v_3}{a^2} \text{ and } v_0 < v_2 \text{ and } \\
&v_0 > v_3 \text{ and } v_3 + av_0 - a < 0, v_0, 0) dv_2 dv_3 dv_0 \\
&\int_0^1 \int_0^1 \mathbb{I}(0 \leq av_0 - av_2 + v_3 \leq 1 \text{ and } av_0 - av_2 + v_3 \leq v_0) \\
&\int_{\max(v_0 + \frac{v_3}{a} - \frac{1}{a}, v_3, \frac{(1-a)v_0}{a} + \frac{v_3}{a}, \frac{(a-1)v_0}{a} + \frac{(a-1)v_3}{a^2}, v_0, 0)}^{\min(v-0 + \frac{v_3}{a}, 1)} v_0 dv_2 dv_3 dv_0 \\
&= \int_0^1 \int_0^1 \mathbb{I}(v_3 < v_0 \text{ and } v_3 \leq \frac{av_0}{a-1} \text{ and } v_3 < a - av_0, \frac{v_0 v_3}{a}, 0) dv_3 dv_0 \\
&= \int_0^1 \int_0^{min(v_0, \frac{av_0}{a-1}, a - av_0, 1)} \frac{v_0 v_3}{a} dv_3 dv_0 \\
&= \int_0^{\frac{a}{a+1}} \int_0^{v_0} \frac{v_0 v_3}{a} + \int_{\frac{a}{a+1}}^1 \int_0^{a - av_0} \frac{v_0 v_3}{a}
\end{aligned}
\tag{7.13}
$$

This reduces to

$$
= \frac{a(3a+1)}{24(a+1)^3}. \tag{7.14}
$$

4. Let $v_0 < v_2$ and $v_0 > v_3$ and $v_3 + av_0 - a > 0$:

$$\int_0^1\int_0^1\int_0^1 \mathbb{I}(0 \le av_0 - av_2 + v_3 \le 1 \text{ and } av_0 - av_2 + v_3 \le v_0)\int_{v_3+av_0-a}^{v_0} dv_1 dv_2 dv_3 dv_0$$
$$= \int_0^1\int_0^1\int_0^1 If(v_2 \le v_0 + \frac{v_3}{a} \text{ and } v_2 \ge v_0 + \frac{v_3}{a} - \frac{1}{a} \text{ and } v_2 > v_3 \text{ and } v_2 > \frac{(a-1)v_0}{a}$$
$$+ \frac{v_3}{a} \text{ and } av_0 + (1-a)v_3 \ge 0 \text{ and } v_2 \ge \frac{(1-a)v_0}{a} + \frac{(a-1)v_3}{a^2} \text{ and } v_0 < v_2 \text{ and } v_0 > v_3$$
$$\text{and } v_3 + av_0 - a > 0, v_0 - (v_3 + av_0 - a), 0) dv_2 dv_3 dv_0$$
$$\int_0^1\int_0^1\int_{\max(v_0+\frac{v_3}{a}-\frac{1}{a}, v_3, \frac{(a-1)v_0}{a}+\frac{v_3}{a}, \frac{(a-1)v_0}{a}+\frac{(a-1)v_3}{a^2}, v_0, 0)}^{\min(v_0+\frac{v_3}{a}, 1)} (1-a)v_0 - v_3 + adv_3 dv_0$$
$$= \int_0^1\int_0^1 \text{If}(v_3 < v_0 \text{ and } v_3 \le \frac{av_0}{a-1} \text{ and } v_3 > a - av_0,$$
$$(1-a)v_0 - v_3 + a - v_0((1-a)v_0 - v_3 + a), 0) dv_3 dv_0$$
$$\int_0^1\int_{\max(a-av_0, 0)}^{\min(v_0, \frac{av_0}{a-1}, 1)} (1-a)v_0 - v_3 + a - v_0((1-a)v_0 - v_3 + a). \tag{7.15}$$

This reduces to

$$= \frac{3a+1}{24(a+1)^3}. \tag{7.16}$$

5. Let $v_0 > v_2$ and $v_0 > v_3$ and $v_3 + av_0 - a < 0$:

$$\int_0^1\int_0^1\int_0^1 \text{If}(v_2 \le v_0 + \frac{v_3}{a} \text{ and } v_2 \ge v_0 + \frac{v_3}{a} - \frac{1}{a} \text{ and } v_2 > v_3 \text{ and } v_3 > \frac{(a-1)v_0}{a} + \frac{v_3}{a}$$
$$\text{and } av_0 + (1-a)v_3 \ge 0 \text{ and } v_2 \ge \frac{(a-1)v_0}{a} + \frac{(a-1)v_3}{a^2} \text{ and } v_0 > v2 \text{ and } v_0 > v_3$$
$$\text{and } v_3 + av_0 - a < 0, \min(v_2, a(1-a)v_0 + (1-a)v_3 + a^2 v_2), 0) dv_2 dv_3 dv_0. \tag{7.17}$$

There are two subcases here.
a) Let $v_2 < a(1-a)v_0 + (1-a)v_3 + a^2 v_2$:

$$\int_0^1\int_0^1\int_{\max(v_0+\frac{v_3}{a}-\frac{1}{a}, v_3, \frac{(a-1)v_0}{a}+\frac{v_3}{a}, \frac{(a-1)v_0}{a}+\frac{(a-1)v_3}{a^2}, \frac{a(1-a)v_0}{1-a^2}+\frac{(1-a)v_3}{1-a^2}, 0)}^{\min(v_0+\frac{v_3}{a}, v_0, 1)} v_2 dv_2 dv_3 dv_0$$
$$= \int_0^1\int_0^1 \text{If}(v_3 < v_0 \text{ and } v_3 \le \frac{av_0}{a-1} \text{ and } v_3 < a - av_0, \frac{v_0^2}{2} - \frac{(\frac{a(1-a)v_0}{1-a^2} + \frac{(1-a)v_3}{1-a^2})^2}{2}, 0). \tag{7.18}$$

This reduces to

$$= \frac{a(3a+1)}{24(a+1)^3}. \tag{7.19}$$

b) Let $v_2 > a(1-a)v_0 + (1-a)v_3 + a^2 v_2$:

$$\int_0^1 \int_0^1 If(v_3 < v_0 \text{ and } v_3 \leq \frac{av_0}{a-1} \text{ and } v_3 < a - av_0, \frac{(v_0 - v_3)(2v_3 + a(v_0 + v_3))}{2a(1+a)^2}, 0) dv_3 dv_0. \tag{7.20}$$

This reduces to

$$= \frac{a}{12(1+a)^3}. \tag{7.21}$$

6. Let $v_0 > v_2$ and $v_0 > v_3$ and $v_3 + av_0 - a > 0$:

$$\int_0^1 \int_0^1 \int_0^1 If(v_2 \leq v_0 + \frac{v_3}{a} \text{ and } v_2 \geq v_0 + \frac{v_3}{a} - \frac{1}{a} \frac{and}{v}_2 > v_3 \text{ and } v_2 > \frac{(a-1)v_0}{a} + \frac{v_3}{a}$$
$$\text{and } av_0 + (1-a)v_3 \geq 0 \text{ and } v_2 \geq \frac{(a-1)v_0}{a} + \frac{(a-1)v_3}{a^2} \text{ and } v_0 > v_2 \text{ and } v_0 > v_3 \text{ and}$$
$$v_3 + av_0 - a > 0, \min(v_2, a(1-a)v_0 + (1-a)v_3 + a^2 v_2 - (v_3 + av_0 - a)), 0) dv_2 dv_3 dv_0. \tag{7.22}$$

This reduces to

$$= \frac{1}{8(a+1)^2}. \tag{7.23}$$

Collecting all the terms, for $\beta > \alpha$ the contribution from this case is

$$\frac{1}{4(1+a)} = \frac{\alpha}{4(\alpha+\beta)}. \tag{7.24}$$

7.4. **Semicircle Moments.** For positive integers k, the semicircle moments are:

$$\begin{aligned} M_{2k+1} &= 0 \\ M_{2k} &= \frac{1}{k+1}\binom{2k}{k} \end{aligned} \tag{7.25}$$

Proof. Since the semicircular distribution is symmetric about zero, integration for the odd moments gives zero, since an odd function integrated over a symmetric region is zero. Also,

$$
\begin{aligned}
M_{2k} &= \frac{1}{2\pi}\int_{-2}^{2} x^{2k}\sqrt{4-x^2}dx \\
&= \frac{1}{\pi}\int_{0}^{2} x^{2k}\sqrt{4-x^2}dx \text{ and set } x = 2\sqrt{y} \\
&= \frac{2^{2k+1}}{\pi}\int_{0}^{1} y^{k=\frac{1}{2}}(1-y)^{\frac{1}{2}}dy \\
&= \frac{2^{2k+1}}{\pi}\frac{\Gamma(k+\frac{1}{2})\Gamma(\frac{3}{2})}{\Gamma(k+2)} = \frac{1}{k+1}\binom{2k}{k}
\end{aligned}
\tag{7.26}
$$

□

REFERENCES

[B] A. Bose, *Introduction to Patterned Random Matrices*, draft edition of forthcoming book (2013).

[BBDS] J. Baik, A. Borodin, P. Deift, and T. Suidan, *A Model for the Bus System in Cuernevaca (Mexico)*, Math. Phys. 2005, 1-9. Online available at http://arxiv.org/abs/math/0510414.

[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, and T. Jiang, *Spectral Measure of Large Random Hankel, Markov, and Toeplitz Matrices*, Annals of Probability **34** (2006), no. 1, 1-38.

[Dy1] F. Dyson, *Statistical theory of the energy levels of complex systems: I, II, III*, J. Mathematical Phys. **3** (1962) 140-156, 157-165, 166-175.

[Dy2] F. Dyson, *The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics*, J. Mathematical Phys., **3** (1962) 1199-1215.

[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.

[JMP] S. Jackson, S. J. Miller, and V. Pham, *Distribution of Eigenvalues of Highly Palindromic Toeplitz Matrices*, http://arxiv.org/abs/1003.2010.

[KrSe] M. Krbalek and P. Seba, *The statistical properties of the city transport in Cuernavaca (Mexico) and Random matrix ensembles*, J. Phys. A: Math. Gen 2000, 33, L229L234.

[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

[MNS] S. J. Miller, T. Novikoff and, A. Sabelli, *The distribution of the second largest eigenvalue in families of random regular graphs*, Experimental Mathematics **17** (2008), no. 2, 231-244.

[Mon] H. Montgomery, *The pair correlation of zeros of the zeta function*. Pages 181-193 in *Analytic Number Theory*, Proceedings of Symposia in Pure Mathematics, vol. 24, AMS, Providence, RI, 1973.

[Q] R. E. Quandt, *Some Basic Matrix Theorems*, www.quandt.com.

[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.

[Wig6] E. Wigner, *Random Matrices in Physics*, SIAM Review, Vol. 9, No. 1 (Jan., 1967), pp. 1-23.

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

[WMW] "Hypergeometric Function." Wolfram MathWorld. <http://mathworld.wolfram.com/HypergeometricFunction.html>. Viewed May 11, 2014.

[X] W. Xiong, *The Limiting Spectral Measure for the Ensemble of Generalized Real Symmetric Block m-Circulant Matrices*, thesis draft no. 3 (2011).