THE LIMITING SPECTRAL DISTRIBUTION OF VARIOUS MATRIX ENSEMBLES UNDER THE ANTICOMMUTATOR OPERATION

GLENN BRUDA, BRUCE FANG, RAUL MARQUEZ, STEVEN J. MILLER, BENI PRAPASHTICA, VISMAY SHARAN, DANIEL SON, SAAD WAHEED, AND JANINE WANG

ABSTRACT. We introduce the anticommutator operation $\{\cdot,\cdot\}$, where $\{A, B\} = AB + BA$, to real symmetric random matrix ensembles including Gaussian orthogonal ensemble (GOE), real symmetric palindromic Toeplitz ensemble (PTE), k-checkerboard ensemble, and real symmetric block k-circulant ensemble (k-BCE). Using combinatorial and topological techniques related to non-crossing and free matching properties of GOE and PTE, we obtain closed form formulae for the moments of the limiting spectral distribution of $\{$GOE, GOE$\}$, $\{$PTE, PTE$\}$, $\{$GOE, PTE$\}$ and corresponding convergence results. In particular, we employ novel integration techniques to show that $\{$PTE, PTE$\}$ converges almost surely to the difference between two i.i.d. χ^2_1 distributions. On the other hand, $\{k$-checkerboard, j-checkerboard$\}$ exhibits entirely different spectral behavior from the other anticommutator ensembles: its spectrum consists of 1 bulk regime of $\Theta(N)$, 4 intermediary blip regimes of $\Theta(N^{3/2})$, and 1 largest blip regime of $\Theta(N^2)$. Using appropriate weight function, we are able to isolate each regime separately and analyze the moments and convergence results via combinatorics. We end with some numerical computation of moments of $\{$GOE, k-BCE$\}$ and $\{k$-BCE, k-BCE$\}$ based on genus expansion.

CONTENTS

Date: December 1, 2024.

This research was supported by the National Science Foundation grant DMS-2241623, Churchill College Cambridge, the Finnerty fund, and Williams College.

1. Introduction

1.1. **Background.** Random matrices were first introduced by Wishart [Wis] in the 1920s to estimate the sample covariance matrices of large samples. Since then, random matrices and statistics of their eigenvalues have been widely studied, with numerous important applications to number theory, quantum mechanics, condensed matter physics, wireless communications [Mez, Meh, For, Cou], etc. Discovered by Wigner while he was investigating nuclear resonance levels [Wig1, Wig2], Wigner's semi-circle law is a central result in random matrix theory and states that under certain conditions, the normalized spectral distribution of a symmetric random matrix with i.i.d. entries converges almost surely to a semi-circle as the dimension of the matrix tends to infinity. The limiting normalized spectral distribution of different types of random matrix ensembles has now been extensively researched with numerous surveys relating to the topic [BasBo1, BasBo2, BLMST, BHS1, BHS2, FM, GKMN, HM, McK, Me].

Even though the spectral distribution of matrices from various random matrix ensembles converges to the semicircle distribution, additional symmetries on the entries of a matrix ensemble often allow us to observe new behavior. Examples of such ensembles include Toeplitz [HM], palindromic Toeplitz [MMS], k-checkerboard [BCDHMSTPY], adjacency matrices of d-regular graphs [GKMN], and block circulant [KKMSX] ensembles. In general, it is rare to find a well-known, closed-form expression of the limiting spectral distribution of a random matrix ensemble.

The new construction in this paper was motivated by the deep connections between random matrix theory and number theory, which originated from a chance meeting between Montgometry and Dyson at the Institute for Advanced Study at Princeton. They observed that the pair correlation of the nontrivial zeros of the Riemann zeta function, which encodes information about the spacing between adjacent zeros, matches with that of the eigenvalues of random Hermitian matrices in the Gaussian Unitary Ensemble (GUE) [BFMT]. This observation started a long and fruitful relationship between these two areas. Subsequent work by Rudnick and Sarnak [RS] generalized this connection to zeros of automorphic L-functions, which are generalization of Riemann zeta function in number theory. In the theory of integral representation of L-functions, Rankin-Selberg convolution is an important method to obtain a new family of L-function from the given L-function families. Considering fruitful connections between random matrix theory and number theory, it is a natural question to ask if there is an analogue of such convolution in the context of random matrix theory.

Previous research on analogues of Rankin Selberg convolution has considered the "swirl" operation and the "disco" operation. Both operations take as input different types of random matrices and concatenate them together to form a larger block matrix [BBDLMSWX, DFJKRSSW]. Inspired by the fact that the distribution of the sum of two independent random variables is the convolution of the two probability distributions, we consider the convolution realized through anticommutating of two random matrices, defined as

$$\{A, B\} \;:=\; AB + BA,$$

where A, B are real symmetric random matrices of the same dimension. We call $\{A, B\}$ the **anticommutator** of A and B. We choose to sample A and B from random matrix ensembles known either for lack of symmetry or for additional symmetrical structure, such as GOE (lack of symmetry), palindromic Toeplitz (additional symmetry), block circulant (additional symmetry), and k-checkerboard (additional symmetry), in hope to see how the additional symmetry affects the spectral distribution of the anticommutator. Moreover, we choose the anticommutator operation in our definition of the convolution two random matrices over the normal sum operation, as the former operation seems to yield a more intricate matrix structure and spectral distribution.

The anticommutator of random matrices has been studied by Nica and Speicher in [NR] using the machinery of free probability. They provided an analytic formula for the moment series of the spectral distribution of the anticommutator of random matrices in terms of expressions involving compositional inversion and R-transform. The analytic formula has successfully yielded the moment series, and thus the spectral distribution of the anticommutator of random matrices whose spectral distributions are semicircle, free Poisson, arcsine, or Bernoulli, etc. However, this method would generally fail in the case of matrix ensembles with additional symmetries, due to the intractability of the R-transform and the compositional inversion computation. In our paper, we demonstrate that we can study the anticommutator of these matrix ensembles using combinatorial, topological, and moment methods. We define these matrix ensembles and summarize the results in the next section.

1.2. **Preliminaries.** We study the anticommutators of the matrices from real symmetric matrix ensembles as defined in this section. All these ensembles have been studied separately using the moment method, but not all of them have closed form expression for their moments.

Definition 1.1. *The $N \times N$ **Gaussian Orthogonal Ensemble (GOE)** is a random matrix ensemble whose matrices $A_N = (a_{ij})$ are given by*

$$a_{ij} \ = \ a_{ji} \ \sim \begin{cases} b_{ij} & \text{if } i \neq j \\ c_{ii} & \text{if } i = j. \end{cases} \tag{1.1}$$

where $b_{ij} \sim N(0,1)$ and $c_{ii} \in N(0,2)$ for all $1 \leq i, j \leq N$.

Definition 1.2. *The $N \times N$ **real symmetric palindromic Toeplitz ensemble (PTE)** (where N is assumed to be even for simplicity) is a random matrix ensemble whose matrices M_N have entries paramatrized by $b_0, b_1, \dots, b_{N/2-1}$, where the b_i's are i.i.d. random variables with mean 0, variance 1, and finite higher moments:*

$$a_{ij} = \begin{cases} b_{|i-j|} & \textit{if } 0 \leq |i-j| \leq \frac{N}{2} - 1 \\ b_{N-1-|i-j|} & \textit{if } \frac{N}{2} \leq |i-j| \leq N-1. \end{cases} \tag{1.2}$$

This matrix is therefore of the form

$$\begin{pmatrix} b_0 & b_1 & b_2 & \cdots & b_2 & b_1 & b_0 \\ b_1 & b_0 & b_1 & \cdots & b_3 & b_2 & b_1 \\ b_2 & b_1 & b_0 & \cdots & b_4 & b_3 & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ b_2 & b_3 & b_4 & \cdots & b_0 & b_1 & b_2 \\ b_1 & b_2 & b_3 & \cdots & b_1 & b_0 & b_1 \\ b_0 & b_1 & b_2 & \cdots & b_2 & b_1 & b_0 \end{pmatrix}. \tag{1.3}$$

Definition 1.3. *Let $k|N$. The $N \times N$ **real symmetric k-block circulant ensemble** is a random matrix ensemble whose matrices are block matrices of the form*

$$\begin{pmatrix} B_0 & B_1 & B_2 & \cdots & B_{\frac{N}{k}-1} \\ B_{-1} & B_0 & B_1 & \cdots & B_{\frac{N}{k}-2} \\ B_{-2} & B_{-1} & B_0 & \cdots & B_{\frac{N}{k}-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_{1-\frac{N}{k}} & B_{2-\frac{N}{k}} & B_{3-\frac{N}{k}} & \cdots & B_0 \end{pmatrix}, \tag{1.4}$$

where each B_i is an $k \times k$ real matrix, each $B_{-i} = B_i^T$, and specifically B_0 is symmetric. Unless otherwise specified (i.e., the restriction that $B_{-i} = B_i^T$), the entries of these block matrices are i.i.d. with mean 0, variance 1, and finite higher moments.

Definition 1.4. *For $k \in \mathbb{Z}_{>0}$ and $w \in \mathbb{R}$, the $N \times N$ (k, w)**-checkerboard ensemble** is a random matrix ensemble whose matrices $M = (m_{ij})$ are given by*

$$m_{ij} = \begin{cases} a_{ij}, & \text{if } i \not\equiv j \mod k \\ w, & \text{if } i \equiv j \mod k, \end{cases} \tag{1.5}$$

where $a_{ij} = a_{ji}$ and all of the distinct a_{ij} terms are sampled from a distribution with mean 0, variance 1, and finite higher moments. We refer to the $(k, 1)$-checkerboard ensemble as the k-checkerboard ensemble. Unless specified otherwise, we always assume that the weight of the checkerboard be 1.

Remark 1.5. *We say that $f(n) = O(g(n))$ if there exist a positive real number C and a real number M such that $f(n) \leq Cg(n)$ for all $n \geq M$. If $f(n) = O(g(n))$ and $g(n) = O(f(n))$, then we say that $f(n) = \Theta(g(n))$. Moreover, we say that $f(n) = o(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.*

Definition 1.6. *As we see later, the eigenvalues of the anticommutator of two random matrices from the above ensembles (except for the k-checkerboard ensemble) are typically of $\Theta(N)$. Hence, we define the **empirical spectral measure** associated to such an $N \times N$ random matrix C_N as*

$$\mu_{C_N,N}(x)dx \ = \ \frac{1}{N}\sum_{i=1}^{N} \delta\left(x - \frac{\lambda_i}{N}\right) dx, \tag{1.6}$$

*where $\{\lambda_i\}_{i=1}^N$ are the eigenvalues of C_N and δ is the Dira-delta functional. The **empirical spectral distribution** $F^{C_N/N}$ of A_N/N is defined as*

$$F^{C_N/N}(x) \ := \ \int_{-\infty}^{x} \mu_{C_N,N}(y)dy \ = \ \frac{\#\{i \leq N : \lambda_i/N \leq x\}}{N}. \tag{1.7}$$

*If as $N \to \infty$ we have $F^{C_N/N}$ converges in some sense (in probability or almost surely) to a distribution F, then we say that F is the **limiting spectral distribution** of the matrix ensemble. The **mth moment of the empirical spectral distribution** $F^{C_N/N}$ **of** C_N/N, denoted by $M_m(C_N)$, is naturally defined as*

$$M_m(C_N) \ := \ \int_{-\infty}^{\infty} x^m dF^{A_N/N}(x) \ = \ \int_{-\infty}^{\infty} x^m \mu_{A_N,N}(x)dx \ = \ \frac{\sum_{i=1}^N \lambda_i^m}{N^{k+1}}. \tag{1.8}$$

*Moreover, we define the **expected m^{th} moment of the empirical spectral distribution** of C_N/N, denoted by $M_m(N)$, as the average of $M_m(C_N)$ over all the C_N in our chosen anticommutator matrix ensemble, i.e. $M_m(N) = \mathbb{E}[M_m(C_N)]$. Finally, we let M_m to be the limit of $M_m(N)$ as $N \to \infty$.*

To establish the limiting spectral distribution of a random matrix ensemble, it is crucial to understand the limiting expected moments of the empirical spectral distribution. The formula below, known as the Eigenvalue Trace Formula, allows us to reduce the moment calculation into a combinatorial problem:

Proposition 1.7. *Let A_N be an $N \times N$ real symmetric matrix, then*

$$M_m(A_N) \ = \ \frac{1}{N^{m+1}}\mathrm{Tr}(A_N^m) \ = \frac{1}{N^{m+1}} \sum_{1\leq i_1,\cdots,i_m\leq N} a_{i_1i_2}a_{i_2i_3}\cdots a_{i_mi_1}, \tag{1.9}$$

where $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix and a_{ij}'s are the entries of A_N indexed by ij. Similarly, if A_N is a random matrix drawn from an $N \times N$ real symmetric matrix ensemble, then

$$M_m(N) \ = \ \frac{1}{N^{m+1}}\mathbb{E}\left[\mathrm{Tr}(A_N^m)\right] \ = \frac{1}{N^{m+1}} \sum_{1\leq i_1,\cdots,i_m\leq N} \mathbb{E}\left[a_{i_1i_2}a_{i_2i_3}\cdots a_{i_mi_1}\right], \tag{1.10}$$

*where by $\mathbb{E}[\mathrm{Tr}(A_N^m)]$ we mean averaging over the $N \times N$ random matrix ensemble with each matrix A_N weighted by its probability of occurring. We refer to each $a_{i_1i_2}a_{i_2i_3}\cdots a_{i_mi_1}$ as a **cyclic product** in the expected m^{th} moment of the random matrix ensemble.*

1.3. **Results.** In this section, we summarize our findings about the anticommutator of random matrices sampled from ensembles including GOE, real symmetric PTE, real symmetric k-BCE, and k-checkerboard. Using the method of genus expansion as we shall develop in Section 2, the non-crossing property of cyclic products of GOE, and the free matching property of the cyclic products of PTE, we are able to obtain closed-form expressions for the limiting expected moments of the empirical spectral distributions of {GOE, GOE}, {PTE, PTE}, {GOE, PTE}. Specifically, the closed-form expressions for {GOE, GOE} and {GOE, PTE} are given by recurrence relations, while that for {PTE, PTE} is given by an explicit, simple expression. We only concern ourselves with the limiting expected even moments, since as we shall see in Section 2, the limiting expected odd moments vanish for all three ensembles. Note that from now on, unless otherwise specified, we refer to the limiting expected moments of the empirical spectral distribution of a matrix ensemble simply as the limiting expected moments of the matrix ensemble.

Lemma 1.8. *The limiting expected $2m^{th}$ moment M_{2m} of* {GOE, GOE} *is given by $M_{2m} = 2f(m)$, where $f(0) = f(1) = 1$, $g(1) = 1$, and*

$$f(m) \;=\; 2\sum_{j=1}^{m-1} g(j)f(m-j) + g(m), \tag{1.11}$$

$$g(m) \;=\; 2f(m-1) + \sum_{\substack{0\le x_1,x_2\le m-2\\ x_1+x_2\le m-2}} (1+\mathbb{1}_{x_1>0})(1+\mathbb{1}_{x_2>0})f(x_1)f(x_2)g(m-1-x_1-x_2). \tag{1.12}$$

Theorem 1.9. *The limiting expected $2m^{th}$ moment of* {GOE, PTE} *is given by $\sigma_{2m,0,m}$, where $\sigma_{n,s,k}$ is given by the conditions:*

(1) $\sigma_{n,s,k} = 0$ *if* $k < 0$,
(2) $\sigma_{n,s,k} = 0$ *if* $s + k > n$,
(3) $\sigma_{n,s,2k+1} = 0$,
(4) $\sigma_{n,s,0} = (2n-1)!!$,

and the recurrence relation

$$\sigma_{n,s,2k} \;=\; \sum_{p=s+1}^{n}\sum_{q=p+1}^{n}\sum_{r=0}^{2k}\left[\sigma_{n-q+p,p,r}\cdot\sigma_{q-p-1,0,2k-2-r} + \sigma_{n-q+p-1,p-1,r}\cdot\sigma_{q-p,1,2k-2-r}\right]. \tag{1.13}$$

Theorem 1.10. *The limiting expected $2m^{th}$ moment M_{2m} of* {PTE, PTE} *is* $4^m((2m-1)!!)^2$.

The limiting expected moments of {PTE, PTE} are precisely the moments of the difference between two i.i.d. χ_1^2 distribution, or equivalently, the product of two i.i.d. $N(0,2)$. As a result, the limiting spectral distribution of {PTE, PTE} is given by the convolution of the distributions of χ_1^2 and $-\chi_1^2$, as shown in Figure 1. We render a proof of this fact in Section 2 using novel integration techniques. On the other hand, as it is very hard to find explicit expressions for the recurrence relations in Theorem 1.8 and 1.9, we are unable to pinpoint the limiting spectral distribution of {GOE, GOE} and {GOE, PTE}, which are shown in Figure 2 and Figure 3. Nevertheless, a proof that the empirical spectral distributions of {GOE, GOE} and {GOE, PTE} converges to some distribution shall be given in Appendix C.

For {GOE, BCE} and {BCE, BCE}, even recurrence relations for the limiting expected moments are hard to come by due to the complexity of the matrix structure and matching rules. Instead, we provide genus expansion formulae as shown in Table 1 for these moments and numerical computation for lower even moments in Appendix F. We need figures of these two ensembles.

So far, all the random matrix ensembles we have considered have eigenvalues of magnitude $\Theta(N)$. However, we begin to observe new behaviors as we consider anticommutator of random matrices sampled from k-checkerboard. In particular, if we take $\gcd(k,j) = 1$ and $k, j \mid N$, a condition crucial to our computation, then the empirical spectral distribution of an $N \times N$ {k-checkerboard, j-checkerboard} consists of 1 bulk regime of $\Theta(N)$ that contains the vast majority of eigenvalues, 1 largest blip regime that contains the largest

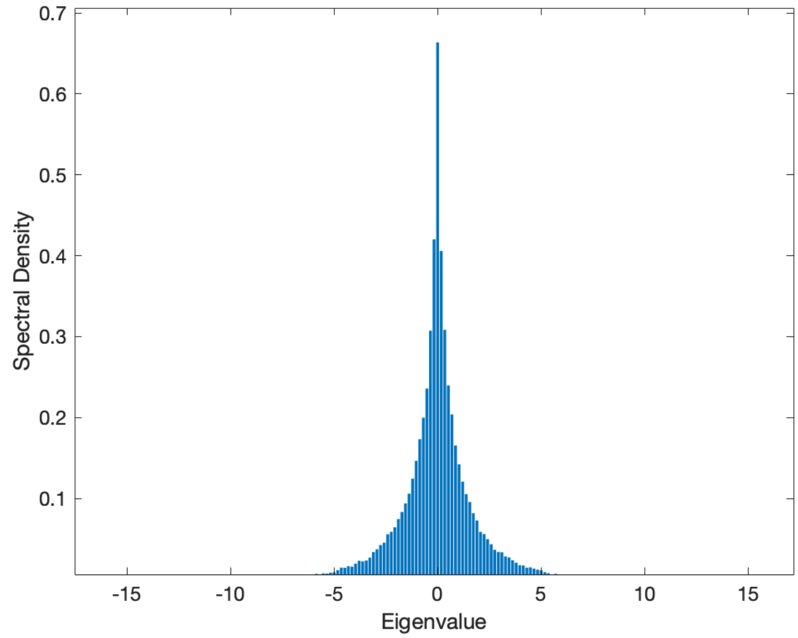

FIGURE 1. Normalized empirical spectral distribution for one hundred 1000×1000 matrices from $\{$PTE, PTE$\}$

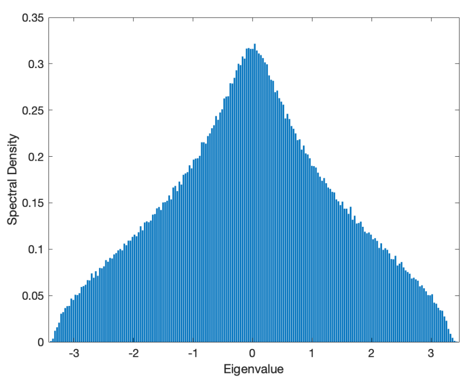

FIGURE 2. Normalized empirical spectral distribution for one hundred 1000×1000 matrices from $\{$GOE, GOE$\}$

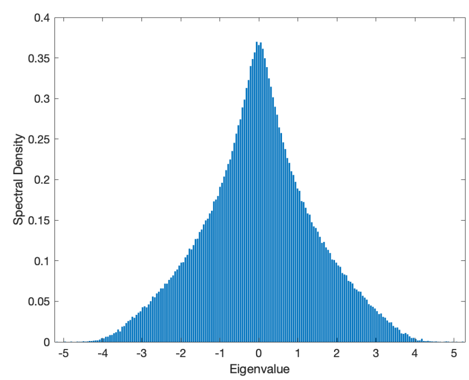

FIGURE 3. Normalized empirical spectral distribution for one hundred 1000×1000 matrices from $\{$GOE, PTE$\}$

eigenvalue (which is positive), and 4 regimes whose magnitudes are in between the bulk regime and the largest blip regime that the number of eigenvalues they contain depends on k and j. We call these regimes **intermediary blip regimes** and their eigenvalues **intermediary blip eigenvalues**. Specifically, two of the intermediary blip regimes each contains $k-1$ eigenvalues and are at $\pm 1/k\sqrt{1-1/j}N^{3/2}+O(N)$, the other two intermediary blip regimes each contains $j-1$ eigenvalues and are at $\pm 1/j\sqrt{1-1/k}N^{3/2}+O(N)$.

	k-Block Circulant
GOE	$\sum_{C\in\mathcal{C}_{2,4m}}\sum_{\pi_C\in NCF_{2,C}(4m)} k^{\#(\gamma_{4m}\pi)-(2m+1)}$
k-Block Circulant	$\sum_{C\in\mathcal{C}_{2,4m}}\sum_{\pi_C\in\mathcal{P}_{2,C}(4m)} k^{\#(\gamma_{4m}\pi)-(2m+1)}$

TABLE 1. Moments of $\{$GOE, k-BCE$\}$ and $\{k$-BCE, k-BCE$\}$.

Moreover, the largest blip regime is at $2/jkN^2+O(N)$. These are shown in Figure **??**(Need a figure here). The existence of these blip regimes are established in Appendix B through Weyl's inequality.

The presence of multiple blip regimes poses a challenge to analyzing the blip regimes one at a time. To this end, we introduce suitable weight functions and their corresponding empirical blip spectral measures that allow us to localize at each blip. Then, we can exploit combinatorics and cancellation techniques to obtain the expected moments of the empirical blip spectral measures. A proof for the convergence of empirical blip spectral measures will be given in Appendix C Which appendix?

Definition 1.11. *Suppose that $n=\log\log(N)$. The* ***empirical largest blip spectral measure*** *associated to the anticommutator of an $N\times N$ k-checkerboard and j-checkerboard $\{A_N,B_N\}$ is*

$$\mu_{\{A_N,B_N\}}(x)dx=\sum_{\lambda\text{ eigenvalues}} g\left(\frac{jk\lambda}{2N^2}\right)\delta\left(x-\left(\frac{\lambda-\frac{2}{jk}N^2}{N}\right)\right)dx, \tag{1.14}$$

where the weight function is given by $g(x)=x^{2n}(1-x)^{2n}$. Let $w_1=1/k\sqrt{1-1/j}$ and $w_2=1/j\sqrt{1-1/k}$. The ***empirical intermediary blip spectral measure*** *associated to $\{A_N,B_N\}$ around $\pm w_sN^{3/2}$ is*

$$\mu_{\{A_N,B_N\},s}(x)\ =\ \frac{1}{2h_s}\sum_{\lambda\text{ eigenvalues}} f_s\left(\frac{\lambda}{w_sN^{3/2}}\right)\delta\left(x-\left(\frac{\lambda^2-w_s^2N^3}{N^{5/2}}\right)\right), \tag{1.15}$$

where $h_s=\begin{cases}k-1 & \text{if } s\ =\ 1\\ j-1 & \text{if } s\ =\ 2\end{cases}$ *and* $f_s(x)\ =\ \dfrac{x^{2n}\prod_{i=1;i\neq s}^{2}\left(x^2-\frac{w_i^2}{w_s^2}\right)^{2n}\left(x^2-\frac{w_5^2N}{w_s^2}\right)^{10n}}{\prod_{i=1;i\neq s}^{2}\left(1-\frac{w_i^2}{w_s^2}\right)^{2n}\left(1-\frac{w_5^2N}{w_s^2}\right)^{10n}}.$

Theorem 1.12. *The limiting expected m^{th} moment of the empirical largest blip spectral distribution is*

$$\mathbb{E}\left[\mu_{\{A_N,B_N\}}^{(m)}\right]\ =\ \sum_{\substack{m_{1a}+m_{1b}+m_{2a}+m_{2b}=m;\\ m_{1a},m_{1b}\text{ even}}} C(m,m_{1a},m_{2a},m_{1b},m_{2b}),$$

where

$$C(m,m_{1a},m_{2a},m_{1b},m_{2b})\ =\ m!\left(\frac{2}{jk}\right)^m\frac{2^{\frac{m_{1a}+m_{1b}}{2}-2(m_{2a}+m_{2b})}m_{1a}!!m_{1b}!!}{m_{1a}!m_{1b}!m_{2a}!m_{2b}!}\left(k\sqrt{1-\frac{1}{k}}\right)^{m_{1a}+2m_{2a}}\left(j\sqrt{1-\frac{1}{j}}\right)^{m_{1b}+2m_{2b}}. \tag{1.16}$$

Theorem 1.13. *Suppose that $j=\log(N)$. Then the limiting expected m^{th} moment of the empirical intermediary blip spectral distribution around $\pm\frac{1}{k}\sqrt{1-\frac{1}{j}}=\pm\frac{1}{k}+o(1)$ is*

$$\mathbb{E}\left[\mu_{\{A_N,B_N,1\}}^{(m)}\right]\ =\ \frac{1}{2k-2}\left(\frac{1}{2k}\right)^m m!. \tag{1.17}$$

Similarly, suppose that $k = \log(N)$, then the limiting expected m^{th} moment of the empirical intermediary blip spectral distribution around $\pm\frac{1}{j}\sqrt{1-\frac{1}{k}} = \pm\frac{1}{j} + o(1)$ is

$$\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N,2\}}\right] = \frac{1}{2j-2}\left(\frac{1}{2j}\right)^m m!. \tag{1.18}$$

2. Limiting Expected Moments of Some Anticommutator Ensembles

In this section, we provide formulae for the limiting expected moments of the following anticommutator ensembles: $\{$GOE,GOE$\}$, $\{$PTE, PTE$\}$, $\{$GOE, PTE$\}$, $\{$GOE, k-BCE$\}$, and $\{k$-BCE, k-BCE$\}$. Along the way, we develop genus expansion formulae for these matrices based on the matching properties of their cyclic products. We first start with some definitions and results to facilitate our discussion.

Definition 2.1. *For a positive integer n, let $[2n] := \{1, 2, \dots, 2n\}$ and $C = c_1c_2\cdots c_{2n}$ be a configuration satisfying $(c_{2s-1}, c_{2s}) \in \{(a_{2s-1}, b_{2s}), (b_{2s-1}, a_{2s})\}$ for all $1 \le s \le n$. Let $\mathcal{C}_{2,2n}$ be the set of all such configurations. Then, a **partition of** $[2n]$ **with respect to** C, $\pi_C = (V_1, \dots, V_t)$, is a tuple of subsets of $[2n]$ such that the following holds:*

(1) $V_i \neq \emptyset$ *for all* $1 \le i \le t$,
(2) $V_1 \cup \cdots \cup V_t = [2n]$,
(3) $V_i \cap V_j = \emptyset$ *for* $i \neq j$,
(4) *For all* $1 \le i \le t$ *and* $i_1, i_2 \in V_i$, $\{c_{i_1}, c_{i_2}\} \in \{\{a_{i_1}, a_{i_2}\}, \{b_{i_1}, b_{i_2}\}\}$.

*We call $V_1, V_2, \dots, V_t$ **blocks** of π_C. Let $\mathcal{P}_C(2n)$ denote the set of all partitions with respect to C of $[2n]$. A partition is called a **pairing** (or matchings) if each block is of size 2. We denote all the pairings with respect to C of $[2n]$ as $\mathcal{P}_{2,C}(2n)$.*

Since there is a bijective correspondence between a configuration $C = c_1c_2\cdots c_{2n}$ and the set $[2n]$, namely $c_i \leftrightarrow i$, then we can often identify a partition of $[2n]$ with respect to C as a partition of the configuration C itself. Note that this definition can be easily extended to any arbitrary subset $S \subseteq [n]$ and configuration C_S (indexed by S), where S is not necessarily equal to $[k]$ for any $k \in \mathbb{N}$.

Definition 2.2. *A partition with respect to C, $\pi_C = (V_1, \dots, V_t)$, of $[2n]$ is **crossing** if there exists blocks V and W with $i, k \in V$ and $j, l \in W$ such that $i < j < k < l$. We denote the set of non-crossing partitions with respect to C of $[2n]$ by $NC_C(2n)$ and the set of non-crossing pairings with respect to C of $[2n]$ by $NC_{2,C}(2n)$.*

We can represent a pairing of the set $[2n]$ with respect to C by drawing lines that connect pairs of numbers from $[2n]$. Then a non-crossing pairing matches a's and b's within themselves and no two lines cross in the diagram. For example, suppose that $C = a_1b_2a_3b_4b_5a_6b_7a_8$. Then $(\{1,8\},\{2,7\},\{3,6\},\{4,5\})$ is a non-crossing pairing with respect to C and $(\{1,3\},\{2,4\},\{5,7\},\{6,8\})$ is a crossing pairing with respect to C, as shown in Figure 4 and 5:

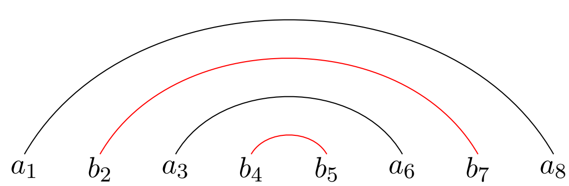

Figure 4. $(\{1,8\},\{2,7\},\{3,6\},\{4,5\})$

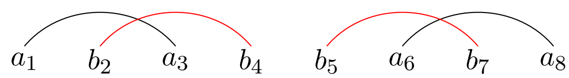

FIGURE 5. $(\{1,3\},\{2,4\},\{5,7\},\{6,8\})$

The following formula, known as Wick's formula, provides a connection between expected moments and non-crossing partitions useful to us in the next section:

Proposition 2.3 (Wick's formula)**.** *For each $\pi \in \mathcal{P}_2(2m)$, let $\mathbb{E}_\pi(X_1,\ldots,X_{2m}) = \prod_{(r,s)\in\pi}\mathbb{E}(X_rX_s)$. Suppose that $(X_1,\ldots,X_n)$ be a real Gaussian random vector. Then*

$$\mathbb{E}[X_{i_1},\ldots,X_{i_{2m}}] \;=\; \sum_{\pi\in\mathcal{P}_2(2m)} \mathbb{E}_\pi[X_{i_1},\ldots,X_{i_{2m}}], \tag{2.1}$$

for any $i_1,\ldots,i_{2m}\in[n]$.

For the remainder of the section, we investigate the limiting expected moments of $\{$GOE,GOE$\}$, $\{$PTE, PTE$\}$, $\{$GOE, PTE$\}$, $\{$GOE, k-BCE$\}$, $\{k$-BCE, k-BCE$\}$ one by one. Note that in any expected moment calculation, it is helpful to characterize the pairings that do contribute in the limit, as they have proven to considerably simplify the calculation. This requires us to extend the method of genus expansion originally used in the moment calculation of the GUE (see [MS] section 1.7 for more details) to our ensembles.

2.1. **Limiting Expected Moments of $\{$GOE, GOE$\}$.** Let $A_N = (a_{ij})$ and $B_N = (b_{ij})$ be $N\times N$ matrices sampled independent from GOE. We consider the m^{th} moment of $\{A_N, B_N\} := A_NB_N + B_NA_N$:

$$M_m(N) \;=\; \frac{1}{N^{m+1}}\mathbb{E}[\mathrm{Tr}(A_NB_N+B_NA_N)^m] \;=\; \frac{1}{N^{m+1}}\sum_{C\in\mathcal{C}_{2,2m}}\;\sum_{1\le i_1,\cdots,i_{2m}\le N}\mathbb{E}[c_{i_1i_2}c_{i_2i_3}\ldots c_{i_{2m}i_1}], \tag{2.2}$$

where for the second equality 2.2, we first expand $\mathrm{Tr}(A_NB_N+B_NA_N)^m$ using additivity of trace and identify each summand as a configuration in $\mathcal{C}_{2,2m}$ (where we identify A_N as a and B_N as b). Then we apply eigenvalue trace lemma to each summand, which gives us the RHS. Now, we apply genus expansion to each $\mathbb{E}[c_{i_1i_2}c_{i_2i_3}\ldots c_{i_{2m}i_1}]$. The argument that follows is essentially the same argument as the genus expansion of the $2m^{\text{th}}$ moment of the GUE, since treating a's and b's both as c's while ensuring that they are matched within themselves preserves the "non-crossing" property of pairings that contribute in the limit.

Now, as $N\to\infty$, we have $M_m(N)\to 0$ when m is odd, since by standard argument the contribution from each type of configuration is $O(N^m)$, but the number of types of configurations depends only on m. When m is even, by Wick's formula

$$\mathbb{E}[c_{i_1i_2}c_{i_2i_3}\cdots c_{i_{2m}i_1}] \;=\; \sum_{\pi_C\in\mathcal{P}_{2,C}(2m)}\mathbb{E}_{\pi_C}[c_{i_1i_2},c_{i_2i_3},\ldots,c_{i_{2m}i_1}]. \tag{2.3}$$

Since $\mathbb{E}[c_{i_ri_{r+1}}c_{i_si_{s+1}}] = 1$ when $i_r = i_{s+1}$ and $i_{r+1} = i_s$ and is 0 otherwise given that $i_r\ne i_s$ (as we shall see, pairings with $i_r = i_s$ vanish in the limit because they are crossing pairings), then $\mathbb{E}[c_{i_1i_2}c_{i_2i_3}\cdots c_{i_{2m}i_1}]$ is the number of pairings π_C with respect to C of $[2m]$ such that $i_r = i_{s+1}$, $i_{r+1} = i_s$, and a's and b's are matched within themselves (i.e., an a is not matched with a b). Now, we think of a tuple of indices $(i_1,\ldots,i_{2m})$ as a function $i:[2m]\to[N]$ and write the pairing $\pi_C = \{(r_1,s_1),(r_2,s_2),\cdots,(r_k,s_m)\}$, as the product of transpositions $(r_1,s_1)(r_2,s_2)\ldots(r_k,s_m)$. We also take γ_{2m} to be the cycle $(1,2,3,\ldots,2m)$. If π_C is a pairing of $[2m]$ and (r,s) is a pair of π_C, then we express our conditions $i_r = i_{s+1}$ and $i_s = i_{r+1}$ as $i(r) = i(\gamma_{2m}(\pi_C(r)))$ and $i(s) = i(\gamma_{2k}(\pi_C(s)))$ respectively. Hence, $\mathbb{E}_{\pi_C}[c_{i_1i_2}c_{i_2i_3}\cdots,c_{i_{2m}i_1}] = 1$ if i is constant on the orbits of $\gamma_{2m}\pi_C$ (e.g. $i(r) = i(s+1)$) and 0 otherwise. Let $\#(\sigma)$ denote the number of cycles of a permutation σ, then

$$M_m(N) \;=\; \frac{1}{N^{m+1}}\mathbb{E}[\mathrm{Tr}(A_N B_N + B_N A_N)^m] \;=\; \frac{1}{N^{m+1}} \sum_{C\in\mathcal{C}_{2,2m}} \sum_{\pi_C\in\mathcal{P}_{2,C}(2m)} N^{\#(\gamma_{2m}\pi_C)}. \tag{2.4}$$

The following proposition provides a powerful characterization of $\#(\gamma_{2m}\pi)$ based on whether π is non-crossing or not. A proof of this proposition can be found in section 1.8 of [MS].

Proposition 2.4. *If π is a pairing of $[2m]$ then $\#(\gamma_{2m}\pi) \leq m-1$ unless π is non-crossing in which case $\#(\gamma_{2m}\pi) = m+1$.*

Combined with Equation 2.4, Proposition 2.4 intuitively says that for large N, $M_m(N)$ is the number of non-crossing pairings with respect to each configuration C summed over all configurations in $\mathcal{C}_{2,2m}$. This leads us to the following lemma:

Lemma 2.5. *As $N \to \infty$,*

$$M_m \;=\; \lim_{N\to\infty} M_m(N) \;=\; \sum_{C\in\mathcal{C}_{2,2m}} \sum_{\pi_C\in NC_{2,C}(2m)} 1. \tag{2.5}$$

Given genus expansion formula 2.5, we are then able to obtain a recurrence relation for the even moment of {GOE, GOE}:

Theorem 2.6. *The limiting expected $2m^{th}$ moment M_{2m} of* {GOE, GOE} *is given by $M_{2m} = 2f(m)$, where $f(0) = f(1) = 1$, $g(1) = 1$, and*

$$f(m) \;=\; 2\sum_{j=1}^{m-1} g(j)f(m-j) + g(m), \tag{2.6}$$

$$g(m) \;=\; 2f(m-1) + \sum_{\substack{0\leq x_1,x_2\leq m-2\\ x_1+x_2\leq m-2}} (1+\mathbb{1}_{x_1>0})(1+\mathbb{1}_{x_2>0})f(x_1)f(x_2)g(m-1-x_1-x_2). \tag{2.7}$$

Proof. Let $f(m)$ be the number of non-crossing pairings with respect to all configurations in $\mathcal{C}_{2,4m}$ starting with an a, and $g(m)$ be the number of non-crossing pairings with respect to all configurations in $\mathcal{C}_{2,4m}$ starting and ending with an a such that these two a's are matched together (i.e., a configuration $a_{i_1i_2}b_{i_2i_3}\cdots b_{i_{4m-1}i_{4m}}a_{i_{4m}i_1}$ with $i_{4m} = i_2$).

We first find the recurrence relation for $f(m)$. We know that $a_{i_1i_2}$ is matched with some $a_{i_{4j}i_{4j+1}}$ with $j \leq m$ (in case when $j = m$, we identify $4m+1$ as 1) since there should be an even number of both a and b terms between $a_{i_1i_2}$ and $a_{i_{4j}i_{4j+1}}$ to ensure non-crossing pairings. When $j = m$, the number of non-crossing pairings is just $g(m)$ by definition. When $j < m$, the number of non-crossing pairings within $a_{i_1,i_2}b_{i_2,i_3}\cdots b_{i_{4j-1}i_{4j}}a_{i_{4j}i_{4j+1}}$ is $g(j)$. We multiply this by the number of non-crossing pairings within the rest of the cyclic product which have no restrictions and is therefore simply $2f(m-j)$, with the 2 accounting for starting with either an a or b. Thus, summing over all possible j's, we have

$$f(m) \;=\; 2\sum_{j=1}^{m-1} g(j)f(m-j) + g(m). \tag{2.8}$$

Similarly, we know that either $b_{i_2i_3}$ is matched with $b_{i_{4m-1}i_{4m}}$, or $b_{i_2i_3}$ is matched with $b_{i_{4x_1+3}i_{4x_1+4}}$ and $b_{i_{4m-1}i_{4m}}$ is matched with $b_{i_{4m-4x_2-2}i_{4m-4x_2-1}}$, with $4x_1+4 < 4m-4x_2-2$, or $x_1+x_2 \leq m-2$. In the first case, since there are no restrictions on the $4k-4$ terms between $b_{i_2i_3}$ and $b_{i_{4m-1}i_{4m}}$, the number of non-crossing pairings is $2f(m-1)$. In the second case, the number of non-crossing pairings of terms between $b_{i_2i_3}$ and $b_{i_{4x_1+3}i_{4x_1+4}}$ is $(1+\mathbb{1}_{x_1>0})f(x_1)$, the number of non-crossing pairings of terms between $b_{i_{4m-4x_2-2}i_{4m-4x_2-1}}$ and $b_{i_{4m-1}i_{4m}}$ is $(1+\mathbb{1}_{x_2>0})f(x_2)$, with the indicator functions accounting for the intermediary terms starting with either an a or a b, and lastly the number of non-crossing pairings of terms between $b_{i_{4x_1+3}i_{4x_1+4}}$ and $b_{i_{4m-4x_2-2}i_{4m-4x_2-1}}$ is $g(m-1-x_1-x_2)$. Thus,

$$g(m) \;=\; 2f(m-1) + \sum_{\substack{0\le x_1,x_2\le m-2\\ x_1+x_2\le m-2}} (1+\mathbb{1}_{x_1>0})(1+\mathbb{1}_{x_2>0})f(x_1)f(x_2)g(m-1-x_1-x_2). \tag{2.9}$$

We have now defined our recurrence for $f(m)$, which counts the number of non-crossing pairings with respect to configurations in $\mathcal{C}_{2,4m}$ starting with an a. Since general configurations in $\mathcal{C}_{2,4m}$ can start with either an a or a b, we multiply $f(m)$ by 2 to get all possible non-crossing pairings of configurations in $\mathcal{C}_{2,4m}$, and we arrive at the even moments being $M_{2m} = 2f(m)$. □

A natural extension of the anticommutator $\{A_N, B_N\}$ is the **ℓ-anticommutator** of ℓ matrix ensembles $A_N^{(1)}, A_N^{(2)}, \dots, A_N^{(\ell)}$, defined as

$$\{A_N^{(1)}, A_N^{(2)}, \dots, A_N^{(\ell)}\} \;:=\; \sum_{\sigma\in S_\ell} A_N^{(\sigma(1))}A_N^{(\sigma(2))}\cdots A_N^{(\sigma(\ell))}, \tag{2.10}$$

where S_ℓ is the symmetric group of order ℓ. By employing the same method as in the proof of Lemma 2.6, we are able to obtain a recurrence relation for the limiting expected moments of the ℓ-anticommutator of GOE. Now, however, instead of two interdependent recurrence relations, we have ℓ interdependent recurrence relations. We leave the result and proof in Appendix A.

2.2. **Limiting Expected Moments of {PTE, PTE}.** The palindromic Toeplitz ensemble (PTE) was introduced by Massey-Miller-Sinsheimer in [MMS] to remove the Diophantine obstruction encountered in the moment calculation of Toeplitz ensemble in [HM]. Essentially, the additional symmetry in the structure of PTE allows almost all pairings to have consistent choices of indexing and contribute in the limit (the pairings that don't have consistent choice of indexing are negligible in the limit). Through this fact, they showed that the limiting expected $2m^{\text{th}}$ moment of the PTE is $(2m-1)!!$, which is exactly the $2m^{\text{th}}$ moment of standard Gaussian, and hence the spectral distribution of PTE converges almost surely to Gaussian. The limiting expected moment calculation of PTE can be naturally extended to that of {PTE, PTE}. By standard argument, the expected odd moments of {PTE, PTE} vanish in the limit. For expected even moments, we can view each cyclic product in the expected $2m^{\text{th}}$ moment of {PTE, PTE} as a cyclic product in the expected $4m^{\text{th}}$ moment of PTE. Thus, the matching in each cyclic product is again free, that is, almost all pairings have consistent choices of indexing and those pairings that don't have consistent choices of indexing are negligible in the limit. Thus, this gives us the genus expansion formula:

$$M_m \;=\; \lim_{N\to\infty} M_m(N) \;=\; \sum_{C\in\mathcal{C}_{2,2m}} \sum_{\pi_C\in P_{2,C}(2m)} 1. \tag{2.11}$$

Intuitively, this is the statement that M_m is equal to the number of pairings with respect to configuration C summed over all configurations in $\mathcal{C}_{2,2m}$.

Theorem 2.7. *The limiting expected $2m^{th}$ moment M_{2m} of* {PTE, PTE} *is* $4^m((2m-1)!!)^2$.

Proof. Recall that a valid configuration $C = c_1c_2\cdots c_{4m}$ satisfy $(c_{2s-1}, c_{2s}) \in \{(a_{2s-1}, b_{2s}), (b_{2s-1}, a_{2s})\}$, then the number of valid configurations is $2^{2m} = 4^m$. Moreover, each configuration has $(2m-1)!!$ ways of matching up the a's and $(2m-1)!!$ ways of matching up the b's, then by Equation 2.11 we have

$$M_{2m} \;=\; 4^m((2m-1)!!)^2. \tag{2.12}$$

□

2.3. **Limiting Expected Moments of {GOE, PTE}.** So far, we've only been looking at **homogeneous** anticommutator ensembles $\{A_N, B_N\}$, i.e., A_N and B_N are the same ensembles. Genus expansions of {GOE, GOE} and {PTE, PTE} suggest that in general, genus expansion of a homogeneous anticommutator ensemble $\{A_N, B_N\}$ is a straightforward generalization of the genus expansion of A_N (or B_N). A natural question to ask is: what does genus expansion of an **inhomogeneous** anticommutator ensembles $\{A_N, B_N\}$ (i.e., A_N and B_N are different ensembles) look like? In this section, we turn our attention to an inhomogeneous anticommutator ensemble, namely {GOE, PTE}. Interestingly, we see that the matching properties of GOE and PTE are well preserved under the anticommutator operator, that is, the contributions to the expected moments of {GOE, PTE} in the limit come solely from non-crossing matchings of the GOE terms and free matchings of the PTE terms that don't cross the matchings of the GOE terms.

Similar to the previous examples, we have that the m^{th} moment of $\{A_N, B_N\}$ is given by

$$M_m(N) \;=\; \frac{1}{N^{m+1}} \sum_{C \in \mathcal{C}_{2,2m}} \sum_{1 \leq i_1, \cdots, i_m \leq N} \mathbb{E}[c_{i_1 i_2} c_{i_2 i_3} \cdots c_{i_{2m} i_1}]. \tag{2.13}$$

Definition 2.8. *For positive integers n and m, let $C = c_1 c_2 \cdots c_{2m}$ be a $(2, m)$-configuration, where $c_i \in \{a_i, b_i\}$ for all $1 \leq i \leq 2m$ under the restriction that $(c_{2s-1}, c_{2s}) \in \{(a_{2s-1}, b_{2s}), (b_{2s-1}, a_{2s})\}$ and $S \subseteq [2n]$ be the set of all the indices of the a's. Let π_S be a partition of S. Then a **layer of** $[2n]$ **with respect to** S is a maximal subset $B_S^{(i)} \subseteq [2n] \setminus S$ such that for any $j, k \in B_S^{(i)}$, there doesn't exist $(p, q) \in \pi_S$ such that $j < p < k < q$ or $p < j < q < k$. It's clear from definition that distinct layers must be disjoint. Then we denote the union of all the layers with respect to S by $B_S := \cup_{i=1}^{t} B_S^{(i)}$, where t is the total number of layers.*

Lemma 2.9. *Consider a cyclic product in* (2.13)*. Let S be the set of all the indices of the a's, π_S be a matching of the a's and $\pi_{[2n]\setminus S}$ be a matching of the b's. If π_S is non-crossing, then the matching $\pi_S \circ \pi_{[2n]\setminus S}$ contribute to* (2.13) *in the limit if and only if there exists i such that $j, k \in B_S^{(i)}$ for each $(j, k) \in \pi_{[2n]\setminus S}$, i.e. every layer is matched within itself. For the $2m^{th}$ moment, the number of ways to assign indices for all the t layers is $N^{m+t} + O(N^{m+t-1})$.*

Proof. First, observe that each layer $B_S^{(i)}$ can be thought of as a cyclic product. For example, consider the following layer $B_S^{(i)}$ consisting of 2ℓ b's. For clarity, we include some of the a's to highlight how the matching the a's give rise to the cyclic product:

$$\begin{aligned} &a_{i_{j_1-1} i_{j_1}} b_{i_{j_1} i_{j_1+1}} a_{i_{j_1+1} i_{j_1+2}} \cdots a_{i_{j_2-1} i_{j_2}} b_{i_{j_2} i_{j_2+1}} a_{i_{j_2+1} i_{j_2+2}} \cdots \\ &a_{i_{j_3-1} i_{j_3}} b_{i_{j_3} i_{j_3+1}} a_{i_{j_3+1} i_{j_3+2}} \cdots b_{i_{j_{2\ell}} i_{j_{2\ell}+1}} a_{i_{j_{2\ell}+1} i_{j_{2\ell}+2}} \cdots . \end{aligned} \tag{2.14}$$

Since the b's form a layer, then for every neighboring two b's, the inner adjacent two a's must be paired together. For example, $a_{i_{j_1+1} i_{j_1+2}}$ and $a_{i_{j_2-1} i_{j_2}}$, which are adjacent $b_{i_{j_1} i_{j_1+1}}$ and $b_{i_{j_2} i_{j_2+1}}$, must be paired together to ensure that all the b's form a layer. Hence, the indices must satisfy the relations $i_{j_1} = i_{j_{2\ell+1}}$, $i_{j_1+1} = i_{j_2}$, $i_{j_2+1} = i_{j_3}$, $\ldots$, $i_{j_{2\ell-1}+1} = i_{j_{2\ell}}$, which allows us to think of $B_S^{(i)}$ as $b_{i_1 i_2} b_{i_2 i_3} \cdots b_{i_{2\ell} i_1}$. Let $\#(B_S^{(i)})$ be the number of b's in the layer $B^{(i)}$. For each cyclic product, if the matching of all the b's is within each layer, then the number of ways to choose indices for all the b's is $\prod_{i=1}^{t} (N^{\#(B_S^{(i)})/2+1} + O(N^{\#(B_S^{(i)})/2})) = N^{m+t} + O(N^{m+t-1})$ by [MMS].

We move on to the case where the matchings of the b's are across different layers. Now, for two arbitrary layers $B_S^{(i_1)}$ and $B_S^{(i_2)}$, suppose that all the b's are paired within these two layers except for at least two b's that are paired across these two layers. Due to the special structure of palindromic Toeplitz, if $b_{i_j i_{j+1}}$ and $b_{i_k i_{k+1}}$ are paired together, then the indices must satisfy the equation $i_{j+1} - i_j + i_{k+1} - i_k = C_j$ for some $C_j \in \{0, \pm(N-1)\}$. Hence, similar to [MMS], we can think of the matching of all the indices as a system of $M := (m(B_S^{(i_1)}) + m(B_S^{(i_2)})/2$ equations, where each index appears exactly twice. After labeling these equations, we pick any equation as eq(M), and choose an index that has occured only once. Then, we select the equation in which this index first appeared and label this equation as eq$(M-1)$. This index is one of our dependent indices and guarantees consistency choice of indices for the other indices in eq$(M-1)$. We can continue this process, and at stage s, if at least one index of eq$(M-s)$ has occured only once among eq$(M-s)$, eq$(M-s+1), \dots,$ eq(M), then we can choose such an index as one of our dependent indices and continue this process. The only way to terminate this process at stage $s < M-1$ is for all the indices among eq$(M-s)$, eq$(M-s+1), \dots,$ eq(M) to occur twice, which implies that each layer is paired within itself, a contradiction. Hence, if at least two b's are paired across these two layers, then the number of dependent indices is $M/2-1$ and the number of ways to choose indices for all the b's is $N^{M/2+1}$. This is a lower order term compared to the case where each layer is paired within itself, which gives that the number of ways to choose indices for all the b's is $N^{M/2+2}$.

Finally, we consider the case where the matchings of the a's cross each other. If a matching of two a's cross another matching of two a's, then we automatically have three layers $B_S^{(i_1)}, B_S^{(i_2)}$, and $B_S^{(i_3)}$. Due to the mismatch, we can no longer view different layers as independent cyclic products, but all three layers as a single cyclic product. The total number of ways to assign the indices for the three layers is $N^{(m(B_S^{(i_1)})+m(B_S^{(i_2)}+m(B^{(i_3)}))/2)+1}$, which is a lower order term compared to the case where each of the three layers is matched within itself. Thus, the number of ways to assign indices for all the layers is $O(N^{m+t-1})$, which is again a lower order term. □

Lemma 2.10. *With the same notation as in Lemma 2.9, regardless of whether π_S is non-crossing or not, the number of ways to assign the remaining indices for the $2m^{th}$ moment is $N^{m+1-t} + O(N^{m-t})$.*

Proof. For a fixed m, when a cyclic product has only one layer, the only possible configurations for the cyclic product are $abba \cdots abba$ or $baab \cdots baab$; moreover, all the a's must be matched in adjacent pairs. Since there is one free index for each adjacent pair of a, then the number of ways to assign the remaining indices is $N^m = N^{(m+1)-1}$. This proves the base case.

When a cyclic product C has two layers $B_S^{(1)}$ and $B_S^{(2)}$, suppose that layer $B_S^{(1)}$ is contained in the cyclic product C_1 with $2k_1$ total b's and layer $B_S^{(2)}$ is contained in the cyclic product C_2 with $2k_2$ total b's, where $C_1 \cap C_2 = \emptyset$. We can think of C as inserting C_2 into C_1. From the base case, C_1 and C_2 are either $abba \cdots abba$ or $baab \cdots baab$. Then, without loss of generalilty, suppose that C_1 has the configuration $abba \cdots abba$. If C_2 is inserted between two a's in C_1, then it must have the configuration $abba \cdots abba$, otherwise C has only one layer instead of two layers. Let C_2 be $a_{i'_1 i'_2} b_{i'_2 i'_3} \cdots b_{i'_{4k-1} i'_{4k}} a_{i'_{4k} i'_1}$ and surrounded by $a_{i_\ell i_{\ell+1}}$ and $a_{i_{\ell+1} i_{\ell+2}}$ in C_1. Since $a_{i'_1 i'_2}$ and $a_{i'_{4k} i'_1}$ as well as $a_{i_\ell i_{\ell+1}}$ and $a_{i_{\ell+1} i_{\ell+2}}$ are no longer adjacent, then we lose one additional degrees of freedom and the number of ways to assign the remaining indices is $k_1 + k_2 - 1$. If C_2 is inserted between two b's in C_1, then it can either be $abba \cdots abba$ or $baab \cdots baab$. Similarly, we can see that the number of ways to assign the remaining indices is $k_1 + k_2 - 1$. Similar constructions follow when we have an arbitrary number of layers in the cyclic product, and whenever we get another layer we lose one extra degree of freedom, giving us $N^{(m+1)-t} + O(N^{m-t})$ ways of assigning the remaining indices for t layers.

□

By (2.9) and (2.10), for the $2m^{\text{th}}$ moment, the number of ways to assign all the indices is $N^{2m+1} + O(N^{2m})$ when π_S is non-crossing and every layer is matched within itself, and $O(N^{2m})$ otherwise. In other words, in the limit, the only contributions come from non-crossing matchings of the a's and matchings of the b's within the same layer. This leads us to Theorem 2.11, as follows.

Theorem 2.11. *The $2m^{th}$ moment of* {GOE, PTE} *is given by $\sigma_{2m,0,m}$, where $\sigma_{n,s,k}$ is given by the conditions*

(1) $\sigma_{n,s,k} = 0$ *if* $k < 0$,
(2) $\sigma_{n,s,k} = 0$ *if* $s + k > n$,

(3) $\sigma_{n,s,2k+1} = 0$,
(4) $\sigma_{n,s,0} = (2n-1)!!$,

and the recurrence relation

$$\sigma_{n,s,2k} \;=\; \sum_{p=s+1}^{n} \sum_{q=p+1}^{n} \sum_{r=0}^{2k} \left[\sigma_{n-q+p,p,r} \cdot \sigma_{q-p-1,0,2k-2-r} + \sigma_{n-q+p-1,p-1,r} \cdot \sigma_{q-p,1,2k-2-r}\right]. \tag{2.15}$$

Proof. Let $\sigma_{n,s,k}$ be the total number of matchings of any cyclic products of a's and b's of length $2n$ that starts with at least s adjacent pairs of bb and has k adjacent pairs of ab and ba in total. It's clear that the $2m^{\text{th}}$ moment of $\{\text{GOE, PTE}\}$ is given by $\sigma_{2m,0,m}$ and conditions (1), (2), (3) trivially follows from the definition. Moreover, since $\sigma_{n,s,0}$ is the number of matchings of cyclic products of b's of length $2n$, where the matchings of b's are free. Then, $\sigma_{n,s,0} = (2n-1)!!$.

Now, we move on to prove the recurrence relation for $\sigma_{n,s,2k}$. Suppose that the p^{th} adjacent pair is the first occurrence of ab or ba pair and that the a is paired with another a in the q^{th} block. Since no matchings can cross the matching of two a's, then if the p^{th} block is ab, the q^{th} block must be ba, and vice versa. In both cases, the matching of the two a's split the cyclic product into two smaller cyclic product, as illustrated in the following example:

Example 2.12. *If $(n,s,k) = (5,1,2)$, then an example of a cyclic product with $p = 3$ and $q = 5$ is $b_{i_1i_2}b_{i_2i_3}b_{i_3i_4}b_{i_4i_5}b_{i_5i_6}a_{i_6i_7}b_{i_7i_8}b_{i_8i_9}a_{i_9i_{10}}b_{i_{10}i_1}$. The matching of the a's partitions the cyclic product into two smaller cyclic products $b_{i_7i_8}b_{i_8i_9}$ (inner cyclic product) and $b_{i_{10}i_1}b_{i_1i_2}b_{i_2i_3}b_{i_3i_4}b_{i_4i_5}b_{i_5i_6}$ (outer cyclic product), where a term from either smaller cyclic product is paired with another term in the same smaller cyclic product.*

If the p^{th} and the q^{th} adjacent pairs are both ba, then the matching of the a's partitions the cyclic product into two smaller cyclic products C_1' (inner cyclic product) and C_2' (outer cyclic product), where C_1' is of length $2(q-p-1)$ and has no restrictions on the number of starting adjacent pairs of bb and C_2' is of length $2(n-(q-p))$ and starts with at least p adjacent pairs of bb. Then the total number of matchings is $\sigma_{q-p-1,0,r}$ for C_1' and $\sigma_{n-q+p,p,2k-2-r}$ for some r. Similarly, if p^{th} and the q^{th} adjacent pairs are both ab, then the total number of matchings is $\sigma_{q-p,1,r}$ for C_1' and $\sigma_{n-q+p-1,p-1,2k-2-r}$ for C_2'.

Summing over all possible p's, q's and r's, we have

$$\sigma_{n,s,2k} \;=\; \sum_{p=s+1}^{n} \sum_{q=p+1}^{n} \sum_{r=0}^{2k} \left[\sigma_{n-q+p,p,r} \cdot \sigma_{q-p-1,0,2k-2-r} + \sigma_{n-q+p-1,p-1,r} \cdot \sigma_{q-p,1,2k-2-r}\right]. \tag{2.16}$$

□

2.4. **Moments of $\{$GOE, k-BCE$\}$ and $\{k$-BCE, k-BCE$\}$.** The real symmetric k-block circulant ensemble is introduced by Koloğlu-Kopp-Miller in [KKMSX] and possesses an even more complicated symmetry structure than the palindromic Toeplitz ensemble: not only do entries on different diagonals satisfy relations analogous to the palindromic Toeplitz ensemble, but the entries on the same diagonal also appear periodically due to the k-block structure. Because of its complicated structure, the $2m^{\text{th}}$ moments of the spectral distribution are not given explicitly, but expressed in terms of the number of pairings of the edges of a $2m$-gon which give rise to a genus g surface. Similar to the palindromic case, we can extend the moment calculation of k-block circulant ensemble to that of $\{k\text{-BCE}, k\text{-BCE}\}$ and $\{\text{GOE}, k\text{-BCE}\}$.

Suppose that $b_{i_si_{s+1}}$ and $b_{i_ti_{t+1}}$ are entries from an $N \times N$ real symmetric k-block circulant matrix, then $b_{i_si_{s+1}}$ and $b_{i_ti_{t+1}}$ are matched iff either of the following relations hold:

(1) $i_{s+1} - i_s = i_{t+1} - i_t + C_s$ and $i_s \equiv i_t \pmod k$, or
(2) $i_{s+1} - i_s = -(i_{t+1} - i_t) + C_s$ and $i_s \equiv i_{t+1} \pmod k$,

where $C_s \in \{0, \pm N\}$. The difference in sign in the two relations above allows us to think of the matching of $(s, s+1)$ and $(t, t+1)$ as having the same or different orientations. For both $\{k\text{-BCE}, k\text{-BCE}\}$ and

{GOE, k-BCE}, we can apply the same argument from [HM], [MMS], and [KKMSX] to show that the total contribution of all the pairings with at least one matching of the same orientation is $O(1/N)$. Hence, it suffices to consider those pairings with matchings of the same orientation, i.e. $i_{s+1}-i_s=-(i_{t+1}-i_t)+C_s$ and $i_s \equiv i_t$ (mod k). By assumption $k=o(N)$, then the modular restrictions do not reduce the total degrees of freedom. Hence, analogous to the palindromic Toeplitz case, we can think of the pairing of terms of in the $2m^{\text{th}}$ moment of an $N \times N$ real symmetric k-block circulant matrix as a system of m linear equations each of the form $i_{s+1}-i_s=-(i_{t+1}-i_t)+C_s$. This gives us $m+1$ free indices with $m-1$ dependent indices and constants $C_s \in \{0, \pm N\}$ uniquely determined, except for a lower order term of choices of free indices.

Using the idea of layers developed in subsection 2.3, we see that if π is a pairing of $\{\text{GOE}, k\text{-BCE}\}$, then π contributes to the moment of $\{\text{GOE}, k\text{-BCE}\}$ iff the GOE terms are matched non-crossing and the real symmetric k-block circulant terms are matched within each layer. Now, consider a pairing π in the $2m^{\text{th}}$ moment of $\{\text{GOE}, k\text{-BCE}\}$. We identify the matched indices in the same congruence class modulo k by the equivalence relation $\sim$. For example, if $a_{i_s i_{s+1}}$ and $a_{i_t i_{t+1}}$ are matched, i.e. $i_s = i_{t+1}$ and $i_{s+1}=i_t$, then $i_s \sim i_{t+1}$ and $i_{s+1} \sim i_t$. If $b_{i_s i_{s+1}}$ and $b_{i_t i_{t+1}}$ are matched, i.e. $i_{s+1}-i_s=-(i_{t+1}-i_t)+C_s$ and $i_s \equiv i_{t+1}$ (mod k), then we also have $i_{s+1} \equiv i_t$ (mod k). Hence, we still have $i_s \sim i_{t+1}$ and $i_{s+1} \sim i_t$. We see that the number of equivalence classes of indices of the pairing π is $\#(\gamma_{4m}\pi)$. For each equivalence class, there are k ways to choose congruence classes. So the number of ways to choose congruence class for all the indices is $k^{\#(\gamma_{4m}\pi)}$. Since there are N/k choices for indices for each congruence class, then $2m^{\text{th}}$ moment of $\{\text{GOE}, k\text{-BCE}\}$ is given by

$$M_{2m} \;=\; \sum_{C \in \mathcal{C}_{2,4m}} \sum_{\pi_C \in NCF_{2,C}(4m)} k^{\#(\gamma_{4m}\pi)-(2m+1)},$$

where $NCF_{2,C}(4m)$ denotes the set of all the pairings with respect to C of $[4m]$ where the GOE terms are matched non-crossingly and the palindromic Toeplitz terms are matched freely without crossing the matchings of the GOE terms. Similarly, the $2m^{\text{th}}$ moment of $\{k\text{-BCE}, k\text{-BCE}\}$ is given by

$$M_{2m} \;=\; \sum_{C \in \mathcal{C}_{2,4m}} \sum_{\pi_C \in \mathcal{P}_{2,C}(4m)} k^{\#(\gamma_{4m}\pi)-(2m+1)}.$$

3. The Blip Spectral Measure of anticommutators

In this section, we consider the blip spectral measures of two ensembles: $\{\text{GOE}, k\text{-checkerboard}\}$ and $\{k\text{-checkerboard}, j\text{-checkerboard}\}$. We always assume that $\gcd(k,j)=1$ and $jk \mid N$, which as we will see later on is crucial to the structure of the anticommutator. Even though the bulk eigenvalues of these two ensembles are both of order $O(N)$ (with largest and smallest eigenvalues $\Theta(N)$), we observe drastically different splitting behaviors: while $\{\text{GOE}, k\text{-checkerboard}\}$ has blip eigenvalues only of order $\Theta(N^{3/2})$, $\{k\text{-checkerboard}, j\text{-checkerboard}\}$ has blip eigenvalues of order $\Theta(N^2)$ and $\Theta(N^{3/2})$. Specifically, $\{\text{GOE}, k\text{-checkerboard}\}$ has $2k$ eigenvalues of order $\Theta(N^{3/2})$ (called the **blip** eigenvalues), of which k are $\frac{N^{3/2}}{k}+O(N)$ and k are $-\frac{N^{3/2}}{k}+O(N)$; by contrast $\{k\text{-checkerboard}, j\text{-checkerboard}\}$ has 1 eigenvalue of order $\Theta(N^2)$ (called the **largest blip** eigenvalue) at $\frac{2}{jk}N^2+O(N)$ and $2k+2j-4$ eigenvalues of order $\Theta(N^{3/2})$ (called the **intermediary blip** eigenvalues), of which two intermediary blips each containing $k-1$ eigenvalues are $\pm\frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}+O(N)$ and two intermediary blips each containing $j-1$ eigenvalues are $\pm\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N)$. For proofs of different regimes, see Appendix B.

3.1. Structural Preliminaries.

We first define the empirical blip spectral measure using appropriate weight functions and reduce the blip moment calculation to combinatorics. Next, using the language developed in [BCDHMSTPY], we identify the types of cyclic products that contribute to the expected m^{th} moments of the blip spectral measures of both ensembles. Then, we explicitly obtain the expected m^{th} moments of the blip spectral measures of {GOE, k-checkerboard} and of {k-checkerboard, j-checkerboard} around the far away blip. Finally, we highlight the combinatorial challenge that we encounter in the calculation of the expected m^{th} moments of the blip spectral measures of {k-checkerboard, j-checkerboard} around the intermediary blips.

Definition 3.1. *Let $w_s = \frac{(-1)^{s+1}}{k}$ for $s \in \{1,2\}$. Then the **empirical blip spectral measure** associated to the anticommutator of an $N \times N$ GOE and k-checkerboard $\{A_N, B_N\}$ around $w_s N^{3/2}$ is*

$$\mu_{\{A_N,B_N\},s}(x) \;=\; \frac{1}{k} \sum_{\lambda \text{ eigenvalues}} f_s^{2n}\left(\frac{\lambda}{w_s N^{3/2}}\right) \delta\left(\frac{x - \left(\lambda - w_s N^{3/2}\right)}{N}\right), \tag{3.1}$$

where $n(N)$ is a function satisfying $n(N) = \log\log(N)$ (note that when we use n in this section we are referring to $n(N)$) and

$$f_s^{2n}(x) \;:=\; \left(\frac{x(2-x)(x+1)(3-x)}{4}\right)^{2n}. \tag{3.2}$$

Definition 3.2. *The **empirical largest blip spectral measure** associated to the anticommutator of an $N \times N$ k-checkerboard and j-checkerboard $\{A_N, B_N\}$, where $\gcd(k,j) = 1$ and $jk \mid N$, is*

$$\mu_{\{A_N,B_N\}}(x) = \sum_{\lambda \text{ eigenvalues}} g_0^{2n}\left(\frac{jk\lambda}{2N^2}\right) \delta\left(x - \left(\frac{\lambda - \frac{2}{jk}N^2}{N}\right)\right), \tag{3.3}$$

where $g_0^{2n}(x) := x^{2n}(2-x)^{2n}$. Let $w_s = \frac{(-1)^{s+1}}{k}\sqrt{1-\frac{1}{j}}$ and $h_s = k$ for $s \in \{1,2\}$ and $w_s = \frac{(-1)^{s+1}}{j}\sqrt{1-\frac{1}{k}}$ and $h_s = j$ for $s \in \{3,4\}$ and

$$g(x) = Something$$

*as introduced in Definition 1.11. Then the **empirical intermediary blip spectral measure** associated to $\{A_N, B_N\}$ around $w_s N^{3/2}$ is*

$$\mu_{\{A_N,B_N\},s}(x) \;=\; \frac{1}{h_s} \sum_{\lambda \text{ eigenvalues}} g\left(\frac{\lambda}{w_s N^{3/2}}\right) \delta\left(\frac{x - \left(\lambda - w_s N^{3/2}\right)}{N}\right). \tag{3.4}$$

We again require that $n(N)$ is a function satisfying $\lim_{N\to\infty} n(N) = \infty$ and $n(N) = \log\log(N)$.

We first consider empirical blip spectral measure associated to $\{A_N, B_N\}$ around $\frac{N^{3/2}}{k}$. As we see later in this section and by symmetry and from 3.5 with 3.22, the empirical blip spectral measure associated to $\{A_N, B_N\}$ around $-\frac{N^{3/2}}{k}$ is the same as that around $\frac{N^{3/2}}{k}$. Since the weight polynomial $f_1^{2n}(x)$ can be written as $\sum_{\alpha=2n}^{8n} c_\alpha x^\alpha$, where $c_\alpha \in \mathbb{R}$, then by eigenvalue trace lemma, the expected m^{th} moment of the empirical blip spectral measure associated to $\{A_N, B_N\}$ around $\frac{N^{3/2}}{k}$ is

$$\begin{aligned}
\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\},1}\right] \;&=\; \mathbb{E}\left[\frac{1}{k} \sum_{\lambda \text{ eigenvalues}} \sum_{\alpha=2n}^{8n} c_\alpha \left(\frac{k\lambda}{N^{3/2}}\right)^{\alpha} \left(\frac{\lambda - w_1 N^{3/2}}{N}\right)^{m}\right] \\
&=\; \mathbb{E}\left[\frac{1}{k} \sum_{\alpha=2n}^{8n} c_\alpha \left(\frac{k}{N^{3/2}}\right)^{\alpha} \frac{1}{N^m} \sum_{i=0}^{m} \binom{m}{i} \left(-\frac{N^{3/2}}{k}\right)^{m-i} \mathrm{Tr}(\{A_N, B_N\}^{\alpha+i})\right] \\
&=\; \frac{1}{k} \sum_{\alpha=2n}^{8n} c_\alpha \left(\frac{k}{N^{3/2}}\right)^{\alpha} \frac{1}{N^m} \sum_{i=0}^{m} \binom{m}{i} \left(-\frac{N^{3/2}}{k}\right)^{m-i} \mathbb{E}[\mathrm{Tr}(\{A_N, B_N\}^{\alpha+i})]. \qquad (3.5)
\end{aligned}$$

Let $g_0^{2n}(x) = \sum_{\beta=2n}^{4ln} d_\beta x^\beta$, then similarly by eigenvalue trace lemma, the expected m^{th} moment of the empirical largest blip spectral measure associated to $\{A_N, B_N\}$ is

$$\begin{aligned}
\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\}}\right] &= \mathbb{E}\left[\sum_{\lambda \text{ eigenvalues}} \sum_{\beta=2n}^{4nl} d_\beta \left(\frac{jk\lambda}{2N^2}\right)^\beta \left(\frac{\lambda - \frac{2}{jk}N^2}{N}\right)^m\right] \\
&= \mathbb{E}\left[\sum_{\beta=2n}^{4nl} d_\beta \left(\frac{jk}{2N^2}\right)^\beta \frac{1}{N^m} \sum_{i=0}^{m} \binom{m}{i} \left(-\frac{2}{jk}N^2\right)^{m-i} \operatorname{Tr}(\{C_N, D_N\}^{\beta+i})\right] \\
&= \sum_{\beta=2n}^{4nl} d_\beta \left(\frac{jk}{2N^2}\right)^\beta \frac{1}{N^m} \sum_{i=0}^{m} \binom{m}{i} \left(-\frac{2}{jk}N^2\right)^{m-i} \mathbb{E}[\operatorname{Tr}\{C_N, D_N\}^{\beta+i}]. \qquad (3.6)
\end{aligned}$$

For the intermediary blip spectral measure we have $g(x) = x^{2n} e^{-\sqrt{N}(x-1)^2}$ which is not a polynomial, but we can use the Taylor expansion of $e^{-\sqrt{N}(x-1)^2}$ which is an infinite polynomial. From properties of the Taylor expansion this infinite polynomial agrees with $e^{-\sqrt{N}(x-1)^2}$ on all finite values and any finite derivatives also agree. This means that we can write $g(x) = \sum_{\alpha=2n}^{\infty} c_\alpha x^\alpha$. Also note that our normalization factor here is $w_1 = \frac{1}{k}\sqrt{1-\frac{1}{j}}$. So we get that the expected m^{th} moment of the empirical intermediate blip spectral measure associated to $\{A_N, B_N\}$ is

$$\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\},1}\right] = \mathbb{E}\left[\frac{1}{2k-2} \sum_{\lambda \text{ eigenvalues}} \sum_{\alpha=2n}^{20n} c_\alpha \left(\frac{\lambda}{w_1 N^{3/2}}\right)^\alpha \left(\frac{\lambda^2 - w_1^2 N^3}{N^{5/2}}\right)^m\right] \qquad (3.7)$$

$$= \frac{(w_1^2 N)^{\frac{1}{2}m}}{2k-2} \sum_{\alpha=2n}^{20n} c_\alpha \left(\frac{1}{w_1 N^{3/2}}\right)^\alpha \sum_{i=0}^{m} \binom{m}{i} (-1)^i \left(\frac{1}{w_1 N^{3/2}}\right)^{2i} \mathbb{E}\left[\mu^{(\alpha+2i)}_{A_N,B_N}\right]. \qquad (3.8)$$

We know that for an $N \times N$ anticommutator ensemble $\{X_N, Y_N\}$, the $(\alpha+i)^{\text{th}}$ expected moment is

$$\mathbb{E}[\operatorname{Tr}(\{X_N, Y_N\}^{\alpha+i})] = \sum_{1 \le i_1, \cdots, i_2 m \le N} \mathbb{E}[c_{i_1 i_2} \cdots c_{i_{2m} i_1}]. \qquad (3.9)$$

Hence, the calculation of the blip moment has now been transformed into a combinatorics problem of counting different types of products of entries. For the rest of this section, we use a to denote a non-weight term of A_N, w a weight term of A_N, b a non-weight term of B_N, v a weight term of B_N, and c any term of A_N or B_N.

Definition 3.3. *A **block** is a set of adjacent a's and b's surrounded by w's and v's in a cyclic product, where the last term of a cyclic product is considered to be adjacent to the first. We refer to a block of length ℓ as an ℓ block or sometimes a block of size ℓ.*

Definition 3.4. *A **weight block** is a set of adjacent w's and v's surrounded by a's and b's in a cyclic product. We similarly refer to a weight block of length ℓ as an ℓ weight block or sometimes a weight block of size ℓ.*

Definition 3.5. *An **adjacent pair** is a pair of adjacent entries of the form $c_{i_{2\ell-1} i_{2\ell}}$, where the first term starts with an odd index.*

Definition 3.6. *A **weight pair** is a pair of adjacent weight terms $c_{i_{2\ell-1} i_{2\ell}} c_{i_{2\ell} i_{2\ell+1}}$. Due to the structure of anticommutator, $\{c_{i_{2\ell-1} i_{2\ell}}, c_{i_{2\ell} i_{2\ell+1}}\} \in \{\{w_{i_{2\ell-1} i_{2\ell}}, v_{i_{2\ell} i_{2\ell+1}}\}, \{v_{i_{2\ell-1} i_{2\ell}}, w_{i_{2\ell} i_{2\ell+1}}\}\}$.*

Definition 3.7. *An **mixed pair** is a pair of adjacent weight and non-weight terms $c_{i_{2\ell-1} i_{2\ell}} c_{i_{2\ell} i_{2\ell+1}}$. Due to the structure of anticommutator,*

$$\{c_{i_{2\ell-1} i_{2\ell}}, c_{i_{2\ell} i_{2\ell+1}}\} \in \{\{a_{i_{2\ell-1} i_{2\ell}}, v_{i_{2\ell} i_{2\ell+1}}\}, \{v_{i_{2\ell-1} i_{2\ell}}, a_{i_{2\ell} i_{2\ell+1}}\}, \{b_{i_{2\ell-1} i_{2\ell}}, w_{i_{2\ell} i_{2\ell+1}}\}, \{w_{i_{2\ell-1} i_{2\ell}}, b_{i_{2\ell} i_{2\ell+1}}\}\}. \qquad (3.10)$$

Definition 3.8. *A* ***configuration*** *is a set of all cyclic products for which it is specified (a) how many blocks there are and what each of them compose of (e.g., a block of $abba$); and (b) in what order these blocks appear (up to cyclic permutation); However, it is not specified how many w's and v's there are between each block.*

Definition 3.9. *A* ***congruence configuration*** *is a configuration together with a choice of the congruence class modulo k every index.*

Definition 3.10. *Given a configuration, a* ***matching*** *is an equivalence relation $\sim$ on the a's and b's in the cyclic product which constrains the way of indexing: for any $c_{i_\ell i_{\ell+1}}$ and $c_{i_t i_{t+1}}$, if $\{c_{i_\ell i_{\ell+1}}, c_{i_t i_{t+1}}\} \in \{\{a_{i_\ell i_{\ell+1}}, a_{i_t i_{t+1}}\}, \{b_{i_\ell i_{\ell+1}}, b_{i_t i_{t+1}}\}\}$, then $\{i_\ell, i_{\ell+1}\} = \{i_t, i_{t+1}\}$ if and only if $c_{i_\ell i_{\ell+1}} \sim c_{i_t i_{t+1}}$.*

Definition 3.11. *Given a configuration, matching, and length of the cyclic product, then an* ***indexing*** *is a choice of*

(1) *the (positive) number of w's and v's between each pair of adjacent blocks (in the cyclic sense), and*
(2) *the integer indices of each a, b, w, v in the cyclic product.*

Definition 3.12. *A* ***configuration equivalence*** *$\sim_C$ is an equivalence relation on the set of all configurations such that for any configurations C_1, C_2, $C_1 \sim_C C_2$ if and only they have the same blocks (but they may have different number and arrangement of w's and v's between their blocks). Every equivalence class under $\sim_C$ is called an S-**class**, specified by the blocks in all the configurations in the equivalence class.*

The following lemma characterizes the S-class with the highest degree of freedom and boils down the blip moment calculation for both $\{$GOE, k-checkerboard$\}$ and $\{k$-checkerboard, j-checkerboard$\}$ to consideration of some nice configurations.

Lemma 3.13. *Fix $m \geq 1$, consider all the S-classes with $|S| = m$. Then a S-class with a matching $\sim$ yields the highest degrees of freedom iff it satisfies the following conditions.*

(1) *It consists only of the following blocks: (i) 1-block of a; (ii) 1-block of b; (iii) 2-block of aa, (iv) 2-block of bb.*
(2) *Each 1-block is paired up to another 1-block and the two terms in each 2-block are paired up with each other.*

Proof. Similar to Lemma 3.14 of [BCDHMSTPY], we see that when a 1-block of a (resp. b) is paired up with another 1-block of a (resp. b) or when the letters in a 2-block of a's (resp. b's) are paired up with each other, there is one degree of freedom lost per block. Now, fix a configuration $\mathcal{C}$ with α from the a's and β from the b's and a matching $\sim$. Suppose that $\sim$ partitions all the a's into equivalence classes $\mathcal{E}_1, \cdots, \mathcal{E}_{s_a}$ and $\mathcal{E}'_1, \cdots, \mathcal{E}'_{s_b}$. Then, without any matching restrictions, the degrees of freedom of $\mathcal{C}$ is

$$\tilde{\mathcal{F}}_{\mathcal{C}} \;=\; \sum_{\text{blocks } \mathcal{B}} (\text{len}(\mathcal{B}) + 1) \;=\; \alpha + \beta + m. \tag{3.11}$$

To find the actual degree of freedom $\mathcal{F}_{\mathcal{C}}$ of $\mathcal{C}$, we can choose two indices from each equivalence class. However, adjacent a's and b's from different equivalence classes (which we call cross-overs) place restrictions on the indices and cause additional loss of degrees of freedom. Let the loss of degrees of freedom due to cross-overs be γ, then $\mathcal{F}_{\mathcal{C}} = 2s_a + 2s_b - \gamma$. Thus, the degree of freedom lost per block is

$$\overline{\mathcal{L}}_{\mathcal{C}} \;=\; \frac{\tilde{\mathcal{F}}_{\mathcal{C}} - \mathcal{F}_{\mathcal{C}}}{m} \;=\; 1 + \frac{\alpha + \beta + \gamma - 2s_a - 2s_b}{m}. \tag{3.12}$$

Since $|\mathcal{E}_i|, |\mathcal{E}'_j| \geq 2$ for $1 \leq i \leq s_a$ and $1 \leq j \leq s_b$, then $s_a \leq \frac{\alpha}{2}$ and $s_b \leq \frac{\beta}{2}$, and so $\overline{\mathcal{L}} \geq 1$. We've shown that if $\mathcal{C}$ satisfies the conditions (1) and (2), then $\overline{\mathcal{L}}_{\mathcal{C}} = 1$. Hence, it suffices to show that if $\mathcal{C}$ with a matching $\sim$ loses one degree of freedom per block (or equivalently, satisfies $\frac{\alpha+\beta+\gamma}{s_a+s_b} = 2$), then it must satisfy the conditions (1) and (2). Since $|\mathcal{E}_i|, |\mathcal{E}'_j| \geq 2$, we get $\alpha \geq 2s_a$ and $\beta \geq 2a_b$. Hence, if some $|\mathcal{E}_i| > 2$

or $|\mathcal{E}_j'| > 2$, then $\frac{\alpha+\beta+\gamma}{s_a+s_b} > 2$. Moreover, if $\gamma > 0$, then $\frac{\alpha+\beta+\gamma}{s_a+s_b} > 2$. Therefore, if $\mathcal{C}$ with a matching $\sim$ loses one degree of freedom per block, then all the blocks are paired up and there can be no cross-overs from different equivalence classes. Thus, the only possible S-classes and matchings are those satisfying conditions (1) and (2). $\square$

Due to the structure of anticommutator, there are eight possibilities for each adjacent pair: ab, ba, av, va, bw, wb, vw, wv. Hence, after specifying the 1-block, we know the other term in the mixed pair, i.e. if the 1-block is of the form $c_{i_{2\ell-1}i_{2\ell}}$ and $c_{i_{2\ell-1}i_{2\ell}} = a_{i_{2\ell-1}i_{2\ell}}$ (resp. $c_{i_{2\ell-1}i_{2\ell}} = b_{i_{2\ell-1}i_{2\ell}}$), then $c_{i_{2\ell}i_{2\ell+1}} = v_{i_{2\ell}i_{2\ell+1}}$ (resp. $c_{i_{2\ell}i_{2\ell+1}} = w_{i_{2\ell}i_{2\ell+1}}$); similar conclusion holds for 1-block of the form $c_{i_{2\ell}i_{2\ell+1}} = a_{i_{2\ell}i_{2\ell+1}}$ (resp. $c_{i_{2\ell}i_{2\ell+1}} = b_{i_{2\ell}i_{2\ell+1}}$). Moreover, after specifying the 2-block, we know both of its adjacent terms (or the two mixed pairs the 2-block belongs to), i.e. every 2-block is of the form $c_{i_{2\ell}i_{2\ell+1}}c_{i_{2\ell+1}i_{2\ell+2}}$, and if $\{c_{i_{2\ell}i_{2\ell+1}}, c_{i_{2\ell+1}i_{2\ell+2}}\} = \{a_{i_{2\ell}i_{2\ell+1}}, a_{i_{2\ell+1}i_{2\ell+2}}\}$, then $\{c_{i_{2\ell-1}i_{2\ell}}, c_{i_{2\ell+2}i_{2\ell+3}}\} = \{v_{i_{2\ell-1}i_{2\ell}}, v_{i_{2\ell+2}i_{2\ell+3}}\}$; if $\{c_{i_{2\ell}i_{2\ell+1}}, c_{i_{2\ell+1}i_{2\ell+2}}\} = \{b_{i_{2\ell}i_{2\ell+1}}, b_{i_{2\ell+1}i_{2\ell+2}}\}$, then $\{c_{i_{2\ell-1}i_{2\ell}}, c_{i_{2\ell+2}i_{2\ell+3}}\} = \{w_{i_{2\ell-1}i_{2\ell}}, w_{i_{2\ell+2}i_{2\ell+3}}\}$. We also know that a weight pair $\{c_{i_{2\ell-1}i_{2\ell}}, c_{i_{2\ell}i_{2\ell+1}}\} \in \{\{w_{i_{2\ell-1}i_{2\ell}}, v_{i_{2\ell}i_{2\ell+1}}\}, \{v_{i_{2\ell-1}i_{2\ell}}, w_{i_{2\ell}i_{2\ell+1}}\}\}$. Thus, we can view specifying a cyclic product of length 2η as only specifying η terms, where specifying one term of the pair $c_{i_{2\ell-1}i_{2\ell}}c_{i_{2\ell}i_{2\ell+1}}$ and whether the pair is weight/non-weight or not uniquely determines the other term of the pair.

3.2. $\{$**GOE,** k**-checkerboard**$\}$.

Lemma 3.14. *For $m_1, m_2 \in \mathbb{Z}_{\geq 0}$, the total contribution to $\mathbb{E}[\mathrm{Tr}\{A_N, B_N\}^\eta]$ of an S-class with m_1 1-blocks of a's and m_2 2-blocks of a is*

$$\frac{p(\eta)N^{\frac{3}{2}\eta-\frac{1}{2}m_1}}{k^\eta}\mathbb{E}_k[\mathrm{Tr}\, C^{m_1}] + O\left(N^{\frac{3}{2}\eta-\frac{1}{2}m_1-1}\right), \tag{3.13}$$

where $p(\eta) = \frac{2\eta^{m_1}}{m_1!} + O(\eta^{m_1-1})$ and C is a $k \times k$ Gaussian Wigner matrix.

Proof. We begin by noting that by 3.13, an S-class containing any b would have fewer degrees of freedom and hence would contribute at most $O(N^{\frac{3}{2}\eta-\frac{m_1}{2}-1})$. Thus, it suffices to consider the case when $m_2 = \frac{\eta-m_1}{2}$ and there are no b's. The rest of the proof is divided into two parts: we first count the number of ways to arrange a prescribed number of blocks into a cyclic product of length 2η; we then count the number of ways to pair together 1-blocks and assign indices that are consistent throughout the cyclic product.

Given $m_1 = o(\eta)$, we claim that the number of ways $q(\eta)$ of arranging m_1 1-blocks and $\frac{\eta-m_1}{2}$ 2-blocks of a's into a cyclic product of length 2η is $\frac{2\eta^{m_1}}{m_1!} + O(\eta^{m_1-1})$.

Indeed, there are two ways to choose the $\frac{\eta-m_1}{2}$ 2-blocks since we can either start with aw or wa, and there are $2^{m_1}\binom{\frac{\eta-m_1}{2}}{m_1} = \frac{\eta^{m_1}}{m_1!} + O(\eta^{m_1-1})$ ways to choose the m_1 1-blocks between adjacent 2-blocks assuming that $m_1 = o(\eta)$. Moreover, as we see later in the proof, we require that mixed pairs containing the 1-blocks are not placed adjacent to each other, which is possible since the number of ways of having at least one 2-block formed from adjacent mixed pairs is $2\binom{\eta-m_1}{2}\binom{\frac{\eta-m_1}{2}-2}{m_1-2} = O(\eta^{m_1-1})$ and thus a lower order term. Hence, overall we get $\frac{2\eta^{m_1}}{m_1!} + O(\eta^{m_1-1})$ ways to arrange the prescribed blocks.

Now, we observe that the second and the first index respectively of two adjacent 1-blocks are congruent mod k, as illustrated in the example below.

Example 3.15. *Consider the configuration*

$$\cdots v_{i_1i_2}a_{i_2i_3}v_{i_3i_4}v_{i_4i_5}a_{i_5i_6}a_{i_6i_7}v_{i_7i_8}a_{i_8i_9}\cdots.$$

Since a's within a 2-block are matched together, we have $i_5 = i_7$ with i_6 being free, and that

$$i_3 \equiv i_4 \equiv i_5 \equiv i_7 \equiv i_8 \pmod k.$$

Thus, all the indices of terms between a pair of 1-blocks, except for those within 2-blocks, share a congruence class. The number of ways to specify the congruence classes of 1-blocks and to pair the 1-blocks up is

$$\sum_{1\le i_1,i_2,\dots,i_{m_1}\le k} \mathbb{E}[a_{i_1i_2}a_{i_2i_3}\dots a_{i_{m_1}i_1}], \tag{3.14}$$

where each $c_{ij}\sim\mathcal{N}(0,1)$. The above expression is simply the ${m_1}^{\text{th}}$ expected moment of $k\times k$ GOE.

Since the $2m_1$ 1-blocks are paired together, then there are only m_1 free indices. As the congruence class of these indices are fixed, the number of choices of these indices is $(\frac{N}{k})^{m_1}$. Similarly, the number of choices of indices for all the 2-blocks is $(\frac{N^2}{k})^{\frac{\eta-m_1}{2}}$, since the indices of each 2-block $a_{i_\ell i_{\ell+1}}a_{i_{\ell+1}i_{\ell+2}}$ must satisfy $i_\ell=i_{\ell+2}$, and there are $\frac{N}{k}$ choices for $i_\ell=i_{\ell+2}$ whose congruence class is fixed and N choices for i_2 that is free. The remaining indices are those of the weight blocks, which must satisfy congruence mod k and hence are each restricted to $\frac{N}{k}$ choices. By the structure imposed by the anticommutator, the total number of indices of all weight blocks is $\eta-\left(\left(\frac{\eta-m_1}{2}\right)+m_1\right)=\frac{\eta-m_1}{2}$. Thus, the total number of ways to assign indices is $\left(\frac{N}{k}\right)^{m_1}\left(\frac{N^2}{k}\right)^{\frac{\eta-m_1}{2}}\left(\frac{N}{k}\right)^{\frac{\eta-m_1}{2}}=\frac{N^{\frac{3}{2}\eta-\frac{1}{2}m_1}}{k^\eta}$. After combining all these pieces, we arrive at the desired result for the contribution of a fixed S-class. □

In the expected m^{th} moment calculation, the following two combinatorial equalities from [BCDHMSTPY] are extremely useful for cancelling the contribution of S-classes with fewer than m blocks.

Lemma 3.16. *For any* $0\le p<m$*,*

$$\sum_{i=0}^{m}(-1)^i\binom{m}{i}i^p \;=\; 0, \tag{3.15}$$

$$\sum_{i=0}^{m}(-1)^{m-i}\binom{m}{i}i^m \;=\; m!. \tag{3.16}$$

Observe that if $m_1>m$, then by Lemma 3.14 the contribution of an S class with m_1 1-block is

$$\begin{aligned}
&\frac{1}{k}\sum_{\alpha=2n}^{8n}c_\alpha\left(\frac{k}{N^{3/2}}\right)^{\alpha}\frac{1}{N^m}\sum_{i=0}^{m}\binom{m}{i}\left(-\frac{N^{3/2}}{k}\right)^{m-i}p(\alpha+i)\frac{N^{\frac{3}{2}(\alpha+i)-\frac{1}{2}m_1}}{k^{\alpha+i}}\\
&=\;\frac{C_{k,m}}{N^{\frac{1}{2}(m_1-m)}}\sum_{\alpha=2n}^{8n}c_\alpha\sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}p(\alpha+i)\\
&=\;\frac{C_{k,m}}{N^{\frac{1}{2}(m_1-m)}}\sum_{\alpha=2n}^{8n}c_\alpha\sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}\left(\frac{2(\alpha+i)^{m_1}}{m_1}+O\left((\alpha+i)^{m_1-1}\right)\right)\\
&=\;\frac{C_{k,m,m_1}}{N^{\frac{1}{2}(m_1-m)}}\sum_{\alpha=2n}^{8n}c_\alpha\alpha^{m_1}.
\end{aligned} \tag{3.17}$$

Since $f_1^{2n}(x)=\left(\frac{x(2-x)(x+1)(3-x)}{4}\right)^{2n}$, then $|c_\alpha|\ll C_0^{2n}$ for some $C_0>0$. Moreover, $\alpha\ll\log\log(N)$, then for some $\epsilon>0$

$$\sum_{\alpha=2n}^{8n}c_\alpha\alpha^{m_1}\ll n^{m_1+1}C_0^{2n}\ll(\log\log(N))^{m_1+1}\log(N)\ll N^{1/2(m_1-m)-\epsilon} \tag{3.18}$$

Hence, as $N \to 0$, the contribution of an S-class with $m_1 > m$ total a blocks and m_2 total aa blocks is negligible. Moreover, if $m_1 < m$, then the contribution of an S-class with m_1 total a blocks is

$$\begin{aligned}
&\frac{1}{k}\sum_{\alpha=2n}^{8n} c_\alpha \left(\frac{k}{N^{3/2}}\right)^\alpha \left(\frac{1}{N^m}\sum_{i=0}^{m}\binom{m}{i}\left(-\frac{N^{3/2}}{k}\right)^{m-i} p(\alpha+i)\left(\frac{N^{\frac{3}{2}(\alpha+i)-\frac{1}{2}m_1}}{k^{\alpha+i}}\right)\right)\\
&= \frac{C_{k,m}}{N^{\frac{1}{2}(m_1-m)}}\sum_{\alpha=2n}^{8n} c_\alpha \sum_{i=0}^{m}\binom{m}{i}(-1)^i p(\alpha+i)\\
&= \frac{C_{k,m}}{N^{\frac{1}{2}(m_1-m)}}\sum_{\alpha=2n}^{8n} c_\alpha \sum_{q=0}^{m_1} c_q \alpha^{m_1-q}\sum_{i=0}^{m}(-1)^i\binom{m}{i} i^q \;=\; 0. \qquad (3.19)
\end{aligned}$$

Thus, we must have $m_1 = m$.

Theorem 3.17. *The expected m^{th} moment associated to the empirical blip spectral measure is*

$$\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\},1}\right] \;=\; 2\left(\frac{1}{k}\right)^{m+1}\mathbb{E}_k[\text{Tr }C^m]. \qquad (3.20)$$

Proof. By the discussion above, we know that $m_1 = m$. Then

$$\begin{aligned}
&\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\},1}\right] \qquad (3.21)\\
&= \frac{1}{k}\sum_{\alpha=2n}^{8n} c_\alpha\left(\frac{k}{N^{3/2}}\right)^\alpha \frac{1}{N^{m+\frac{1}{2}m}}\sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}\left(\frac{N^{3/2}}{k}\right)^{m+\alpha}\frac{2(\alpha+i)^m}{m!}\mathbb{E}_k[\text{Tr}C^m]\\
&= \frac{2}{m!}\left(\frac{1}{k}\right)^{m+1}\mathbb{E}_k[\text{Tr}C^m]\sum_{\alpha=2n}^{8n} c_\alpha\sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}(\alpha+i)^m\\
&= \frac{2}{m!}\left(\frac{1}{k}\right)^{m+1}\mathbb{E}_k[\text{Tr}C^m]\sum_{\alpha=2n}^{8n} c_\alpha\sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}\sum_{p=0}^{m}\binom{m}{p}\alpha^p i^{m-p}\\
&= \frac{2}{m!}\left(\frac{1}{k}\right)^{m+1}\mathbb{E}_k[\text{Tr}C]\sum_{\alpha=2n}^{8n}\sum_{p=0}^{m}\binom{m}{p}c_\alpha\alpha^p\sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i}i^{m-p}. \qquad (3.22)
\end{aligned}$$

Since the inner sum is 0 if $p > 0$ and $m!$ if $p = 0$ by Lemma 3.16 and $f_1^{(2n)}(1) = \sum_{\alpha=2n}^{8n} c_\alpha = 1$, then

$$\begin{aligned}
\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\},1}\right] &= \frac{2}{m!}\left(\frac{1}{k}\right)^{m+1}\mathbb{E}_k[\text{Tr}C^m]\sum_{\alpha=2n}^{8n} c_\alpha m!\\
&= 2\left(\frac{1}{k}\right)^{m+1}\mathbb{E}_k[\text{Tr}C^m]. \qquad (3.23)
\end{aligned}$$

□

3.3. Moments of $\{k$-checkerboard, j-checkerboard$\}$.

Proposition 3.18. *For $m_{1a}, m_{2a}, m_{1b}, m_{2b} \in \mathbb{Z}_{\geq 0}$, define $m_1 := m_{1a} + m_{1b}$ and $m_2 := m_{2a} + m_{2b}$. If $m_1 + m_2 = o(\eta)$, then the total contribution to $\mathbb{E}[\text{Tr}\{A_N, B_N\}^\eta]$ of an S-class with m_{1a} 1-blocks of a, m_{1b} 1-blocks of b, m_{2a} 2-blocks of a, and m_{2b} 2-blocks of b is*

$$\frac{2^{\eta-2m_2}\eta^{m_1+m_2}}{m_{1a}!m_{1b}!m_{2a}!m_{2b}!}2^{\frac{m_1}{2}}(m_{1a})!!(m_{1b})!!\left(\tfrac{1}{k}\right)^{\eta-m_{1a}-2m_{2a}}\left(\tfrac{1}{j}\right)^{\eta-m_{1b}-2m_{2b}}\left(1-\tfrac{1}{k}\right)^{\frac{m_{1a}}{2}+m_{2a}}\left(1-\tfrac{1}{j}\right)^{\frac{m_{1b}}{2}+m_{2b}}N^{2\eta-(m_1+m_2)}.$$

Proof. The proof is divided into two parts. First, we count the number of ways to arrange the prescribed blocks into the cyclic product of length 2η and assign the weight pairs; second, we count the number of ways to pair up all the 1-blocks and assign indices that ensures consistent indexing throughout the cyclic product.

Given $m_1 + m_2 = o(\eta)$, we claim that the number of ways $q(\eta)$ of arranging the prescribed blocks into the cyclic product of length 2η and assign the adjacent weight pairs is $\frac{2^{\eta-2m_2}\eta^{m_1+m_2}}{m_{1a}!m_{1b}!m_{2a}!m_{2b}!} + O(2^\eta \eta^{m_1+m_2-1})$. Naively, $q(\eta)$ is simply

$$\binom{\eta - m_2}{m_1 + m_2}\binom{m_1 + m_2}{m_2}2^{\eta-2m_2}\binom{m_1}{m_{1a}}\binom{m_2}{m_{2a}} \;=\; \frac{2^{\eta-2m_2}\eta^{m_1+m_2}}{m_{1a}!m_{1b}!m_{2a}!m_{2b}!} + O(2^\eta \eta^{m_1+m_2-1}). \quad (3.24)$$

That is, we first choose all the $m_1 + m_2$ blocks, viewing each of the m_2 2-block as a 1-block (where $\eta - m_2$ comes from), which can be done in $\binom{\eta-m_2}{m_1+m_2}$ ways. Then, we choose the m_2 2-blocks from all the $m_1 + m_2$ blocks, m_{1a} 1-blocks of a from all the m_1 1-blocks, m_{2a} 2-blocks of a from all the m_2 2-blocks, and finally specifying the mixed pairs the 1-blocks belong to and the weight pairs, all of which can be done in $\binom{m_1+m_2}{m_2}2^{\eta-2m_2}\binom{m_1}{m_{1a}}\binom{m_2}{m_{2a}}$ ways.

However, this naive counting method fails to account for the restriction that different mixed pairs of a and v and of b and w cannot be placed adjacent to each other to form a 2-block (e.g. if two mixed pairs va and av are adjacent to each other, then we have a 2-block of a). The number of ways of having at least one 2-block formed from different mixed pairs of a and v and of b and w is $2^{\eta-2m_2-1}(\eta - m_2)\binom{\eta-m_2-2}{m_1+m_2-2}\binom{m_1+m_2}{m_2}\binom{m_1}{m_{1a}}\binom{m_2}{m_{2a}} = O(2^\eta \eta^{m_1+m_2-1})$ if $m_1 + m_2 = o(\eta)$. Thus,

$$q(\eta) \;=\; \frac{2^{\eta-2m_2}\eta^{m_1+m_2}}{m_{1a}!m_{1b}!m_{2a}!m_{2b}!} + O(2^\eta \eta^{m_1+m_2-1}). \quad (3.25)$$

Similarly, we can also guarantee that none of the 1-blocks are adjacent to each other, since the number of ways of having at least two adjacent 1-blocks is $2^{\eta-2m_2}(\eta - m_2)\binom{\eta-m_2-2}{m_1+m_2-2}\binom{m_1+m_2}{m_2}\binom{m_1}{m_{1a}}\binom{m_2}{m_{2a}} = O(2^\eta \eta^{m_1+m_2-1})$ if $m_1 + m_2 = o(\eta)$.

Now, we count the number of ways to assign the indices for the S-class. In contrast to [BCDHMSTPY] where there are restrictions on the indices of the 1-blocks, we demonstrate in the example below that we can remove such restrictions.

Example 3.19. *Consider a weight block surrounded by two non-weight terms $c_{i_{\ell-1}i_\ell}c_{i_\ell i_{\ell+1}}\cdots c_{i_t i_{t+1}}c_{i_{t+1}i_{t+2}}$, where $t - \ell$ is sufficiently large. We assume without loss of generality that $c_{i_{\ell-1}i_\ell} = a_{i_{\ell-1}i_\ell}$ and $c_{i_{t+1}i_{t+2}} = a_{i_{t+1}i_{t+2}}$. After specifying i_ℓ, if $c_{i_\ell i_{\ell+1}} = w_{i_\ell i_{\ell+1}}$, then $i_\ell \equiv i_{\ell+1} \pmod k$ and there are $\frac{N}{k}$ choices of indices for $i_{\ell+1}$; if $c_{i_\ell i_{\ell+1}} = v_{i_\ell i_{\ell+1}}$, then $i_\ell \equiv i_{\ell+1} \pmod j$ and there are $\frac{N}{j}$ choices of indices for $i_{\ell+1}$. After specifying i_{t+1}, there are similar number of choice of indices for i_t. Since $t - \ell$ is sufficiently large, then with high probability there exists $\ell + 1 \le s \le t - 2$ such that $\{c_{i_s i_{s+1}}, c_{i_{s+1}i_{s+2}}\} = \{w_{i_s i_{s+1}}, v_{i_{s+1}i_{s+2}}\}$. We can specify the indices $i_{\ell+2}, \dots, i_s$ and $i_{s+2}, \dots, i_{t-1}$ the same way as before. Then we have $i_{s+1} \equiv i_s \pmod k$ and $i_{s+1} \equiv i_{s+2} \pmod j$. Since $\gcd(k, j) = 1$ and $jk \mid N$, then by Chinese remainder theorem, there are $\frac{N}{kj}$ choices of indices i_{s+1}. If the number of w's and v's in the weight block is r and $t - \ell + 1 - r$, respectively, then there are $\left(\frac{1}{k}\right)^r \left(\frac{1}{j}\right)^{t-\ell+1-r} N^{t-\ell}$ ways of specifying the $i_{\ell+1}, \dots, i_t$. Thus, regardless of the indices we specify for the two non-weight terms surrounding a weight block, we can guarantee with high probability consistency of indexing throughout the weight block.*

We know that the total number of w's and v's in a cyclic product of length 2η is $\eta - m_{1a} - 2m_{2a}$ and $\eta - m_{1b} - 2m_{2b}$, respectively. Then the number of choices of congruence classes of indices for all the w's and v's has the corresponding factors $\left(\frac{1}{k}\right)^{\eta-m_{1a}-2m_{2a}}$ and $\left(\frac{1}{j}\right)^{\eta-m_{1b}-2m_{2b}}$. Now, by Lemma 3.13, each 1-block is paired up with another 1-block and the two terms of each 2-block are paired up with each other. Moreover, the indices of any non-weight terms $a_{i_\ell i_{\ell+1}}$ and $b_{i_t i_{t+1}}$ must satisfy the modular restrictions

$i_\ell \not\equiv i_{\ell+1} \pmod k$ and $i_t \not\equiv i_{t+1} \pmod j$. Then similarly, the number of choices of congruence classes of indices for all the a's and b's has the corresponding factors $\left(1-\frac{1}{k}\right)^{\frac{m_{1a}}{2}+m_{2a}}$ and $\left(1-\frac{1}{j}\right)^{\frac{m_{1b}}{2}+m_{2b}}$. Since the loss of degrees of freedom per block in a contributing configuration is 1, then the contribution of actually specifying all the indices is $N^{2\eta-(m_1+m_2)}$. Thus, the number of ways of assigning the indices that guarantees consistent indexing is

$$\left(\frac{1}{k}\right)^{\eta-m_{1a}-2m_{2a}}\left(\frac{1}{j}\right)^{\eta-m_{1b}-2m_{2b}}\left(1-\frac{1}{k}\right)^{\frac{m_{1a}}{2}+m_{2a}}\left(1-\frac{1}{j}\right)^{\frac{m_{1b}}{2}+m_{2b}}N^{2\eta-(m_1+m_2)}. \tag{3.26}$$

Finally, since there are m_{1a} 1-block of a and m_{1b} 1-block of b, and there are no restrictions on their indices, then the number of ways of matching up all the 1-blocks is $2^{\frac{m_1}{2}}(m_{1a})!!(m_{1b})!!$. Note that the $2^{\frac{m_1}{2}}$ factor is due to the fact that for any two paired 1-blocks $c_{i_\ell i_{\ell+1}}, c_{i_t, i_{t+1}}$, we either have $i_\ell = i_{t+1}$ and $i_{\ell+1} = i_t$ or $i_\ell = i_t$ and $i_{\ell+1} = i_{t+1}$. This completes the proof. □

Theorem 3.20. *The m^{th} moment of the largest blip spectral measure is*

$$\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\}}\right] = \sum_{\substack{m_{1a}+m_{1b}+m_{2a}+m_{2b}=m;\\ m_{1a},m_{1b}\text{ even}}} C(m,m_{1a},m_{2a},m_{1b},m_{2b})\left(k\sqrt{1-\frac{1}{k}}\right)^{m_{1a}+2m_{2a}}\left(j\sqrt{1-\frac{1}{j}}\right)^{m_{1b}+2m_{2b}},$$

where $C(m,m_{1a},m_{2a},m_{1b},m_{2b}) := m!\left(\frac{2}{jk}\right)^m \frac{2^{\frac{m_{1a}+m_{1b}}{2}-2(m_{2a}+m_{2b})}m_{1a}!!m_{1b}!!}{m_{1a}!m_{1b}!m_{2a}!m_{2b}!}$.

Proof. By Lemma 3.13, it suffices to consider the contributions from S-classes with 1-blocks and 2-blocks. Applying Proposition 3.18 to Equation (3.7), we have

$$\begin{aligned}\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\}}\right] &= \sum_{\beta=2n}^{4nl} d_\beta\left(\frac{jk}{2N^2}\right)^\beta \frac{1}{N^m}\sum_{i=0}^m \binom{m}{i}\left(-\frac{2}{jk}N^2\right)^{m-i}\left(\frac{2}{kj}\right)^{\beta+i}N^{2(\beta+i)-(m_1+m_2)}\\ &\sum_{\substack{m_{1a},m_{1b},m_{2a},m_{2b}\\ m_{1a},m_{1b}\text{ even}}} \frac{2^{-2m_2+m_1/2}(m_{1a})!!(m_{1b})!!}{m_{1a}!m_{1b}!m_{2a}!m_{2b}!}\left(k\sqrt{1-\frac{1}{k}}\right)^{m_{1a}+2m_{2a}}\left(j\sqrt{1-\frac{1}{j}}\right)^{m_{1b}+2m_{2b}}(\beta+i)^{m_1+m_2}.\end{aligned} \tag{3.27}$$

Similar to the blip moment calculation of {GOE, k-checkerbaord}, we require $m_1+m_2 \leq m$, since otherwise the contribution from $m_1+m_2 > m$ vanishes in the limit. Moreover, by Lemma 3.16, the sum $\sum_{i=0}^m \binom{m}{i}(-1)^{m-i}(\beta+i)^{m_1+m_2}$ vanishes except for the m^{th} power of i. Hence, the contribution to the moment in the limit only comes from $m_1+m_2=m$. After combining terms and canceling out the dependency on i, and noting that $g_0^{2n}(1) = \sum_{\beta=2n}^{4ln} d_\beta x^\beta$, we have

$$\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\}}\right] = \sum_{\substack{m_{1a}+m_{1b}+m_{2a}+m_{2b}=m;\\ m_{1a},m_{1b}\text{ even}}} C(m,m_{1a},m_{2a},m_{1b},m_{2b})\left(k\sqrt{1-\frac{1}{k}}\right)^{m_{1a}+2m_{2a}}\left(j\sqrt{1-\frac{1}{j}}\right)^{m_{1b}+2m_{2b}}. \tag{3.28}$$

□

It is very difficult to differentiate the intermediary blips when they are of the same power of N. Thus, we create a difference in power of N by taking either k or j to infinite. Suppose that here we take j to infinite, say that $j = j(N) = \Theta(\log N)$. We let $w_1 := \frac{1}{k}\sqrt{1-\frac{1}{j}}$, $w_2 := \frac{1}{j}\sqrt{1-\frac{1}{k}}$, $w_3 := \frac{1}{kj}$. Notice that in order for these different regimes to not cluster together, we need $j \ll \sqrt{N}$. Since we set $j(N) = \log N$ we see that $\frac{1}{k}\sqrt{1-\frac{1}{j}} \to \frac{1}{k}$ and $w_2 \to \frac{1}{\log N}\sqrt{1-\frac{1}{k}}$. Note that the regimes of eigenvalues are located at the bulk at $O(N)$, the intermediate blips at $\pm w_1 N^{3/2} + O(N)$, $\pm w_2 N^{3/2} + O(N)$, and the far away blip at $w_3 N^2 + O(N)$. So in order to isolate the intermediate blip at $\pm w_1 N^{3/2} + O(N)$ we define the weight function below.

Definition 3.21 (Weight Function)**.** *For example, we can take it to be $j(N) := \log(N)$. We define our weight function to be*

$$f(x) \;=\; \frac{x^{2n}\left(x^2 - \frac{w_2^2}{w_1^2}\right)^{2n}\left(x^2 - \frac{w_3^2}{w_1^2}N\right)^{6n}}{\left(1 - \frac{w_2^2}{w_1^2}\right)^{2n}\left(1 - \frac{w_3^2}{w_1^2}N\right)^{6n}},$$

since this weight function evaluates to 1 *at the intermediate blip and* 0 *at the locations of all other regimes of eigenvalues.*

Definition 3.22 (Empirical Intermediate Blip Measure)**.** *The centered mth moment for the intermediary regimes at $\pm w_1$ is*

$$\begin{aligned}\mathbb{E}\left[\mu^{(m)}_{\{A_N,B_N\},1}\right] \;&=\; \mathbb{E}\left[\frac{1}{2k-2}\sum_{\lambda \text{ eigenvalues}}\sum_{\alpha=2n}^{20n} c_\alpha \left(\frac{\lambda}{w_1 N^{3/2}}\right)^{\alpha}\left(\frac{\lambda^2 - w_1^2 N^3}{N^{5/2}}\right)^{m}\right]\\ &=\; \frac{(w_1^2 N)^{\frac{1}{2}m}}{2k-2}\sum_{\alpha=2n}^{20n} c_\alpha \left(\frac{1}{w_1 N^{3/2}}\right)^{\alpha}\sum_{i=0}^{m}\binom{m}{i}(-1)^i\left(\frac{1}{w_1 N^{3/2}}\right)^{2i}\mathbb{E}\left[\mu^{(\alpha+2i)}_{A_N,B_N}\right].\end{aligned}$$

Remark 3.23. *The reason the moment is defined this way is because empirically there are two separate blips located at $\pm w_1 N^{3/2}$. We consider opposite blips as part of a single similar structure since the previous literature does not give any explanation for the difference between these opposite sides. For this reason we consider the quantity $\lambda^2 - w_1^2 N^3$ since this function sends both empirical blips to the same location and measures the deviation from the center to the power m, essentially capturing the m*th *centered moment of this function of the distribution.*

We also see that empirically the eigenvalues in the blips are located at $w_1 N^{3/2} + O(N)$ which makes that when we take λ^2 all of the values are located at $w_1^2 N^3 + O(N^{5/2})$ so in order to make the error term $O(1)$ we must normalize by $N^{5/2}$ in order to get finite moments.

Lemma 3.24. *For the weight function associated to the blip and the empirical blip spectral measure as defined above. For each of the other blips, which are empirically located at $O(N)$, $\pm w_2 N^{3/2} + O(N)$, and $w_3 N^2 + O(N)$, their contribution to the m^{th} moment is* 0*. So in particular*

$$f\left(\frac{\lambda}{w_1 N^{3/2}}\right)\left(\frac{\lambda^2 - w_1^2 N^3}{N^{5/2}}\right)^m \to 0, \tag{3.29}$$

for $\lambda = O(N), \pm w_2 N^{3/2} + O(N), w_3 N^2 + O(N)$.

Proof. Now we can show that at all of the other regimes the weight function cancels out the contribution from the other terms. The other regimes are located at $O(N)$, $\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2} + O(N) = O(\frac{N^{3/2}}{\log N})$ since we set j to $O(\log N)$, and $\frac{N^2}{\log N} + O(N)$.

Then we see that when we plug in these eigenvalues into our weight function we get that since we divide by $w_1 N^{3/2}$ so we see that the first regime at $O(N)$ maps to $O(1/\sqrt{N})$. So then we can plug this in and we get $\lim_{N\to\infty} \frac{\left(\frac{1}{N}-\frac{w_3^2}{w_1^2}N\right)^{2n}}{\left(1-\frac{w_3^2}{w_1^2}N\right)^{2n}} = 1$ since the N terms dominate the remainder. Similarly we get that $\lim_{N\to\infty} \frac{\left(\frac{1}{N}-\frac{w_2^2}{w_1^2}\right)^{2n}}{\left(1-\frac{w_2^2}{w_1^2}\right)^{2n}} = 1$ since the constant terms dominate the $1/N$. So this leaves the final limit $\frac{N^{m/2}}{N^n} \to 0$ since for the mth moment contributes a factor of $N^{m/2}$ at most and m is a constant while $n = O(\log\log N)$ so the N^n in the dominator dominates.

Then we can consider the middle regime located at $\pm w_2 N^{3/2} + O(N)$. This is equivalent to evaluating $f(\pm w_2/w_1 + O(1/\sqrt{N}))$. Note that even though w_2 goes to 0, this intermediate blip asymptotically moves to 0 more slowly than the bulk so still need to evaluate this limit. Similarly to the previous paragraph the terms corresponding to N cancel to 1 since the N terms dominate. Then we see $\lim_{N\to\infty}\left(\left(\pm\frac{w_2}{w_1} + O(1/\sqrt{N})\right)^2 - \frac{w_2^2}{w_1^2}\right)^{2n} = O(1/N^n)$ since the w_2/w_1 terms cancel. Then we see that $\lim_{N\to\infty} \frac{\left(\frac{w_2^2}{w_1^2}\right)^{2n}}{\left(1-\frac{w_2^2}{w_1^2}\right)^{2n}} = \lim_{N\to\infty}\left(\frac{w_2^2}{w_1^2-w_2^2}\right)^{2n} = O(1/\log N)^{2n}$because $w_2 = O(1/\log N)$ while w_1 is a constant. So combining these terms we get $O(1/N^n)\cdot O(1/\log N)^{2n}\cdot O(N^{m/2})$ which clearly goes to 0 since $n\to\infty$ while m is just a constant. So the second blip also goes to 0. Finally we consider the far away blip located at $w_3N^2 + O(N)$. So this reduces to evaluating $f((w_3\sqrt{N})/w_1 + O(1/\sqrt{N}))$. So then we see that $\lim_{N\to\infty} \frac{\left((w_3\sqrt{N})/w_1+O(1/\sqrt{N})^2-(w_3^2N/w_1^2)\right)^{6n}}{(1-w_2^2/w_1^2)^{2n}} = C^{2n}$ for some constant C. Then we see that since the $\sqrt{N}$ term dominates we get that

$$\lim_{N\to\infty} \frac{((w_3\sqrt{N})/w_1 + O(1/\sqrt{N}))^{2n}((w_3\sqrt{N})/w_1 + O(1/\sqrt{N})^2) - w_2^2/w_1^2}{(1-(w_3^2N)/w_1^2)^{6n}} = O\left(\frac{1}{N}\right)^n. \tag{3.30}$$

So then we see that $C^{2n}\cdot O(\frac{1}{N})^n$ goes to 0 since C is a constant while N and n go to infinity. So this proves that this weight function gives negligible contribution for all other eigenvalues for sufficiently large N.

□

Lemma 3.25. *For the eigenvalues at the intermediate blip located at $\pm w_1N^{3/2}+O(N)$ the weight function equals* 1*, so the contribution to the moment calculation is a constant $O(1)$. In particular*

$$f\left(\frac{\lambda}{w_1N^{3/2}}\right)\left(\frac{\lambda^2 - w_1^2N^3}{N^{5/2}}\right)^m \to O(1), \tag{3.31}$$

for $\lambda = \pm w_1N^{3/2} + O(N)$.

Proof. First we see that $(\pm w_1N^{3/2} + O(N))^2 = w_1^2N^3 + O(N^{5/2})$ by basic binomial expansion, so when we set $\lambda = \pm w_1N^{3/2} + O(N)$ we get $\left(\frac{\lambda^2-w_1^2N^3}{N^{5/2}}\right) = O(1)$. So since m is also a constant we get that $\left(\frac{\lambda^2-w_1^2N^3}{N^{5/2}}\right)^m = O(1)$ also.

Then we see that since we divide out $w_1N^{3/2}$ within the function we must evaluate $f\left(\pm1 + O(1/\sqrt{N})\right)$ then we see that similarly to the previous lemma both of the terms with a factor of N cancel since this factor dominates. Also both of the $1-w_2^2/w_1^2$ and $1+O(1/\sqrt{N})-w_2^2/w_1^2$ terms also cancel since the constant terms dominate because $1/\sqrt{N}\to 0$. So this leaves us with $\left(1+O(1/\sqrt{N})\right)^{2n}$. Since $\lim_{y\to\infty}(1+1/y)^y = e$ and $n \ll \log\log(N)$, then $1 \le \lim_{N\to\infty}(1+O(1/\sqrt{N}))^{2n} \le \lim_{N\to\infty}\left((1+O(1/\sqrt{N}))^{O(\sqrt{N})}\right)^{2n/O(\sqrt{N})}$

$= \lim_{N\to\infty} e^{2n/O(\sqrt{N})} = 1$. Hence, $\lim_{N\to\infty}(1+O(1/\sqrt{N}))^{2n} = 1$. So since we are multiplying quantities whose limits are 1 and $O(1)$ the product is a constant $O(1)$. This shows that the contribution of the desired intermediate blip is factored into the calculation. $\square$

Since weight function can reduce the contribution from the other regimes, we can choose to only consider the intermediary regime of our interest. Since the eigenvalues of our interest are of order $N^{3/2}$, then in the expansion of $\mathbb{E}\left[\mu^{(\alpha+2i)}_{A_N,B_N}\right]$ we are only expecting contribution of order $N^{3/2(\alpha+2i)}$. However, we can show that the only configurations we need to consider are those consisting of wb, bw blocks (could be 1 block or 2 block).

Lemma 3.26. *The only configurations that contribute the the empirical intermediate blip spectral measure, $\mathbb{E}\left[\mu^{(\alpha+2i)}_{A_N,B_N}\right]$ are those configurations that only consist of wb or bw blocks. In particular any configuration that contains any block of the form wv, vw, av, or va does not contribute to the intermediate blip spectral measure.*

Proof. Suppose that we have other blocks, va, va, wv, vw. Suppose that number of $1a$ blocks is m_{1a}, the number of $2a$ blocks is m_{2a}, the number of wv or vw is $m_{vw,vw}$, then the power of N of such a configuration is $\frac{N^{(m_{1a}+\frac{3}{2}m_{2a}+m_{wv,vw})+(m_{1b}+\frac{3}{2}m_{2b})}}{(\log(N))^{\frac{m_{1a}}{2}+m_{2a}+m_{wv,vw}}}$. Note that we can take n to be a small power of N, say $\log\log(N)$. Let's compare the denominator with N. Taking $\log$ for both, we have

$$\log\left((\log(N))^{\frac{m_{1a}}{2}+m_{2a}+m_{wv,vw}}\right) = \log\log(N)\left(\frac{m_{1a}}{2}+m_{2a}+m_{wv,vw}\right) = O(\log\log(N))$$

where $\log$ of N is $\log(N)$. Thus, we see that $(\log(N))^{\frac{m_{1a}}{2}+m_{2a}+m_{wv,vw}} \ll N$, and is not an integer power of N. Thus, if we have any $1a$ blocks, $2a$ blocks, wv or vw blocks, then either the power of N is too big or too small compared to $N^{3/2(\alpha+2i)}$. Thus, we can safely ignore any contributions from these configurations. $\square$

Theorem 3.27. *The m^{th} moment of the empirical intermediate blip measure is*

$$\frac{1}{2k-2}\left(\frac{1}{2k}\right)^m m!. \tag{3.32}$$

We apply the same idea from the calculation of the largest blip. So the contribution of a configuration with m_{1b} $1b$ blocks and m_{2b} $2b$ blocks is

$$\frac{(w_1^2N)^{\frac{1}{2}m}}{2k-2}\sum_{\alpha=2n}^{(6+2\ell)n} c_\alpha\left(\frac{1}{w_1N^{3/2}}\right)^{\alpha}\sum_{i=0}^{m}\binom{m}{i}(-1)^i\left(\frac{1}{w_1N^{3/2}}\right)^{2i}\binom{\alpha+2i-m_{2b}}{m_{1b}}$$
$$\left(\frac{1}{k}\sqrt{1-\frac{1}{j}}\right)^{m_{1b}+2m_{2b}} N^{2(\alpha+2i)-(m_{1b}+m_{2b})}$$

Note that since $w_1 = \frac{1}{2}\sqrt{1-\frac{1}{j}}$ and $m_{1b}+2m_{2b} = \alpha+2i$, then $\left(\frac{1}{k}\sqrt{1-\frac{1}{j}}\right)^{m_{1b}+2m_{2b}} = w_1^{\alpha+2i}$, so the power of w_1 involving α and $2i$ in the above expression all cancel out. Now, we analyze the power of N in the expression above. The power of N is $N^{2(\alpha+2i)-(m_{2b}+m_{1b})-\frac{3}{2}(\alpha+2i)+\frac{1}{2}m} = N^{\frac{1}{2}(\alpha+2i)+\frac{1}{2}m-(m_{1b}+m_{2b})}$. To have a contribution, we need $m_{1b}+m_{2b} \le \frac{1}{2}(\alpha+2i)+\frac{1}{2}m$. Moreover, we have that $2m_{2b}+m_{1b} = \alpha+2i$, substituting $m_{2b} = \frac{1}{2}(\alpha+2i)-\frac{1}{2}m_{1b}$, we get that $m_{1b} \le m$. Thus, $m_{2b} \ge \frac{\alpha+2i}{2}-\frac{1}{2}m$. So $\binom{\alpha+2i-m_{2b}}{m_{1b}} = \binom{\frac{\alpha+2i+m_{1b}}{2}}{m_{1b}} = (\frac{\alpha}{2}+i)^{m_{1b}}+(\alpha+i)^{m_{1b}-1}$. Recall that $\sum_{i=0}^m\binom{m}{i}i^p = 0$ for all $0 \le p \le m-1$, then the only contribution to the above expression is the case when $p = m$. This occurs only when $m_{1b} = m$. Hence, we only need to consider $m_{1b} = m$, in which case the above expression is $\frac{w_1^m}{2k-2}\sum_{i=0}^m\binom{m}{i}(-1)^i i^m\left(\frac{1}{2}\right)^m = \frac{1}{2k-2}\left(\frac{w_1}{2}\right)^m m!$. Then since $w_1 \to \frac{1}{k}$ since j grows with N we see that the moments become $\frac{1}{2k-2}\left(\frac{1}{2k}\right)^m m!$.

APPENDIX A. MOMENTS OF ANTICOMMUTATORS

A.1. Moments of ℓ anticommutators.

Proposition A.1. *For sequences $f_0, f_1, \dots, f_\ell$ defined such that $f_0(0) = 1$, and $f_0(1) = f_1(1) = f_2(1) = \cdots = f_\ell(1) = 1$, and with recurrence relations for $m > 1$ given by*

$$f_0(m) = f_1(m) + \ell! \sum_{j=1}^{m-1} f_1(j) f_0(m-j),$$

$$f_k(m) = f_{k+1}(m) + \sum_{\substack{1 \le x_1, x_2 \le m \\ x_1 + x_2 \le m}} (\ell - k)!(k-1)! f_{k+1}(x_1) f_{k+1}(x_2) f_{\ell-k-1}(m - x_1 - x_2 + 1) \tag{A.1}$$

for any $0 < k < \ell - 1$, and by

$$f_{\ell-1}(m) = f_\ell(m) + \sum_{\substack{0 \le x_1, x_2 < m-1 \\ x_1 + x_2 < m-1}} (\ell-1)!(1 + (\ell! - 1) \cdot \mathbb{1}_{x_1 > 0})(1 + (\ell! - 1) \cdot \mathbb{1}_{x_2 > 0}) f_0(x_1) f_0(x_2) f_1(m - x_1 - x_2 - 1),$$

$$f_\ell(m) = \ell! \cdot f_0(m-1), \tag{A.2}$$

the $2m^{th}$ moment of the ℓ-anticommutator is

$$M_{2m} = \ell! \cdot f_0(m) \tag{A.3}$$

Proof. The proof follows similarly as with the 2-anticommutator. Let $f_0(m)$ be the number of non-crossing matchings with respect to all $(\ell, 2\ell m)$-configurations starting with $a^{(1)}_{i_1 i_2} a^{(2)}_{i_2 i_3} \cdots a^{(\ell)}_{i_\ell i_{\ell+1}}$ and let $f_k(m)$ be the number of such matchings where the first k terms are paired with the last k terms in a nested fashion (i.e. for $k = 3$, we would have configurations of the form $a^{(1)}_{i_1 i_2} a^{(2)}_{i_2 i_3} a^{(3)}_{i_3 i_4} \cdots a^{(3)}_{i_{2\ell m-2} i_{2\ell m-1}} a^{(2)}_{i_{2\ell m-1} i_{2\ell m}} a^{(1)}_{i_{2\ell m} i_1}$ and matchings such that $i_2 = i_{2\ell m}$, $i_3 = i_{2\ell m-1}$, and $i_4 = i_{2\ell m-2}$).

We first find the recurrence relation for $f_0(m)$. To ensure non-crossing matchings, $a^{(1)}_{i_1 i_2}$ must be paired with some $a^{(1)}_{i_{2\ell j} i_{2\ell j+1}}$ with $j \le m$ (in the case when $j = m$, we identify $2\ell m + 1$ as 1). When $j = m$, the number of non-crossing matchings is simply $f_1(m)$ by definition. When $j < m$, the number of non-crossing matchings within $a^{(1)}_{i_1 i_2} \cdots a^{(1)}_{i_{2\ell j} i_{2\ell j+1}}$ is $f_1(j)$, while the number of non-crossing matchings within the rest of the cyclic product for which we have no restrictions is $\ell! f_0(m-j)$, with the $\ell!$ accounting for different possible arrangements of the first ℓ terms. Multiplying these together and summing over all possible j's gives

$$f_0(m) = f_1(m) + \ell! \sum_{j=1}^{m-1} f_1(j) f_0(m-j). \tag{A.4}$$

Now turning to $f_k(m)$, we look separately at when $0 < k < \ell - 1$, when $k = \ell - 1$, and when $k = \ell$.

When $0 < k < \ell - 1$, we have either that $a^{(k+1)}_{i_{k+1} i_{k+2}}$ is paired with $a^{(k+1)}_{i_{2\ell m-k} i_{2\ell m-k+1}}$, or that $a^{(k+1)}_{i_{k+1} i_{k+2}}$ is paired with $a^{(k+1)}_{i_{2\ell x_1-k} i_{2\ell x_1-k+1}}$ and $a^{(n)}_{i_{2\ell m-k} i_{2\ell m-k+1}}$ is paired with $a^{(n)}_{i_{2\ell(m-x_2)+k+1} i_{2\ell(m-x_2)+k+2}}$, where $n \in \{k+1, \dots, \ell\}$, with both $x_1, x_2 \ge 1$ and $2\ell x_1 - k + 1 < 2\ell(m - x_2) + k + 1$, or $x_1 + x_2 \le m$. The first case is precisely the definition of $f_{k+1}(m)$. In the second case, the number of non-crossing matchings of terms between $a^{(k+1)}_{i_{k+1} i_{k+2}}$ and $a^{(k+1)}_{i_{2\ell x_1-k} i_{2\ell x_1-k+1}}$ is $f_{k+1}(x_1)$, the number of non-crossing matchings of terms between $a^{(n)}_{i_{2\ell(m-x_2)+k+1} i_{2\ell(m-x_2)+k+2}}$ and $a^{(n)}_{i_{2\ell m-k} i_{2\ell m-k+1}}$ is $(\ell - k)! f_{k+1}(x_2)$, with the $(\ell - k)!$ accounting for different possible arrangements of the last ℓ terms, and the number of non-crossing matchings of terms between $a^{(k+1)}_{i_{2\ell x_1-k} i_{2\ell x_1-k+1}}$ and $a^{(n)}_{i_{2\ell(m-x_2)+k+1} i_{2\ell(m-x_2)+k+2}}$ is $(k-1)! f_{\ell-k+1}(m - x_1 - x_2 + 1)$.

The last statement follows from viewing the $2\ell(m-x_1-x_2)+2k+2$ terms between $a^{(k+1)}_{i_{2\ell x_1-k}i_{2\ell x_1-k+1}}$ and $a^{(n)}_{i_{2\ell(m-x_2)+k+1}i_{2\ell(m-x_2)+k+2}}$ as $2\ell(m-x_1-x_2+1)$ terms where the $(l-k-1)$ terms on both end are matched to each other and hence fixed, with the $(k-1)!$ accounting for different permutations of the remaining $k+1$ of the first ℓ terms. We sum over all possible x_1 and x_2's to get the desired result:

$$f_k(m) = f_{k+1}(m) + \sum_{\substack{1\le x_1,x_2\le m\\ x_1+x_2\le m}} (\ell-k)!(k-1)!f_{k+1}(x_1)f_{k+1}(x_2)f_{\ell-k-1}(m-x_1-x_2+1). \tag{A.5}$$

When $k=\ell-1$, we either have that $a^{(\ell)}_{i_\ell i_{\ell+1}}$ is paired with $a^{(\ell)}_{i_{2\ell m-\ell+1}i_{2\ell m-\ell+2}}$, which is simply $f_\ell(m)$ by definition, or that $a^{(\ell)}_{i_\ell i_{\ell+1}}$ is paired with $a^{\ell}_{i_{2\ell x_1+\ell+1}i_{2\ell x_1+\ell+2}}$ and $a^{(n)}_{i_{2\ell m-\ell+1}i_{2\ell m-\ell+2}}$ is paired with $a^{(n)}_{i_{2\ell(m-x_2)-\ell}i_{2\ell(m-x_2)-\ell+1}}$, with both $x_1,x_2\ge 0$. The number of non-crossing matchings of terms between $a^{(\ell)}_{i_\ell i_{\ell+1}}$ and $a^{\ell}_{i_{2\ell x_1+\ell+1}i_{2\ell x_1+\ell+2}}$ is $(1+(\ell!-1)\mathbb{1}_{x_1>0})f_0(x_1)$, the number of non-crossing matchings of terms between $a^{(n)}_{i_{2\ell(m-x_2)-\ell}i_{2\ell(m-x_2)-\ell+1}}$ and $a^{(n)}_{i_{2\ell m-\ell+1}i_{2\ell m-\ell+2}}$ is $(1+(\ell!-1)\mathbb{1}_{x_2>0})f_0(x_2)$, with a factor of $\ell!$ when either $x_1,x_2>0$ due to different possible arrangements of the first ℓ terms starting at $a^{(n)}_{i_{\ell+1}i_{\ell+2}}$, where $n\in\{1,\dots,\ell-1\}$, and the number of non-crossing matchings of terms between $a^{\ell}_{i_{2\ell x_1+\ell+1}i_{2\ell x_1+\ell+2}}$ and $a^{(n)}_{i_{2\ell(m-x_2)-\ell}i_{2\ell(m-x_2)-\ell+1}}$ is $(\ell-1)!f_1(m-x_1-x_2-1)$. For the last statement, we view the $2\ell(m-x_1-x_2)-2\ell-2$ terms between $a^{\ell}_{i_{2\ell x_1+\ell+1}i_{2\ell x_1+\ell+2}}$ and $a^{(n)}_{i_{2\ell(m-x_2)-\ell}i_{2\ell(m-x_2)-\ell+1}}$ as $2\ell(m-x_1-x_2-1)$ terms with the first and last term matched with each other, with $(l-1)!$ accounting for different arrangements of the remaining $l-1$ terms. We once again sum over all possible x_1 and x_2's to reach the desired result:

$$\begin{aligned} &f_{\ell-1}(m) = f_\ell(m)+\\ &\sum_{\substack{0\le x_1,x_2<m-1\\ x_1+x_2<m-1}} (\ell-1)!(1+(\ell!-1)\cdot\mathbb{1}_{x_1>0})(1+(\ell!-1)\cdot\mathbb{1}_{x_2>0})f_0(x_1)f_0(x_2)f_1(m-x_1-x_2-1).\end{aligned} \tag{A.6}$$

Lastly, when $k=\ell$, with no matching conditions on the terms between the first and last ℓ terms, for each possible permutation of the next ℓ terms, we have $f_0(m-1)$ non-crossing matchings, amounting to $\ell!\cdot f_0(m-1)$ total non-crossing matchings.

We have now fully defined our recurrences for $f_0(k)$, which represents the number of non-crossing matchings with respect to $(\ell,2\ell m)$-configurations where the first ℓ terms are fixed to be $a^{(1)}_{i_1i_2}a^{(2)}_{i_2i_3}\cdots a^{(\ell)}_{i_\ell i_{\ell+1}}$. Applying any permutation to these ℓ terms preserves the non-crossing property of these matchings. Hence, we multiply $f_0(m)$ by $\ell!$ to obtain all possible non-crossing partitions with respect to $(\ell,2\ell m)$-configurations, and we arrive at the even moments being $M_{2k}=\ell!\cdot f_0(k)$.

□

A.2. **Bulk Moments of {GOE, k-checkerboard}.**

Proposition A.2. *The $2m^{th}$ bulk moment of* {GOE, k-checkerboard} *is $M_{2m}=2(1-\frac{1}{k})^m f(m)$, where $f(0)=f(1)=1$, $g(1)=1$, and*

$$f(m) \;=\; 2\sum_{j=1}^{m-1} g(j)f(m-j)+g(m), \tag{A.7}$$

and

$$g(m) \;=\; 2f(m-1)+\sum_{\substack{0\le x_1,x_2\le m-2\\ x_1+x_2\le m-2}}(1+\mathbb{1}_{x_1>0})(1+\mathbb{1}_{x_2>0})f(x_1)f(x_2)g(m-1-x_1-x_2) \qquad \text{(A.8)}$$

Proof. By a result in [Tao1], the limiting distribution of the bulk of {GOE, $(k,1)$-checkerboard} is given by the limiting distribution of {GOE, $(k,0)$-checkerboard}. Because in a contributing cyclic product, every term $c_{i_\ell i_{\ell+1}}$ from the $(k,0)$-checkerboard must be non-weight with the modular restriction $i_\ell \not\equiv i_{\ell+1}$ (mod k), then the $2m^{\text{th}}$ bulk moment of {GOE, k-checkerboard} is essentially the $2m^{\text{th}}$ moment {GOE, GOE}, except that we have to account for all the modular restrictions. Since the $2m$ non-weight terms are paired together, the probability that all the terms from the $(k,0)$-checkerboard are non-weights is $\left(1-\frac{1}{k}\right)^m$. This completes the proof.

□

A.3. **Bulk Moments of {k-checkerboard, j-checkerboard}.**

Corollary A.3. *The $2m^{th}$ bulk moment of* {k-checkerboard, j-checkerboard} *is $M_{2m}=2\left(1-\frac{1}{k}\right)^m \left(1-\frac{1}{j}\right)^m f(m)$, where*

$$f(m) \;=\; 2\sum_{j=1}^{m-1} g(j)f(m-j)+g(m) \qquad \text{(A.9)}$$

and

$$g(m) \;=\; 2f(m-1)+\sum_{\substack{0\le x_1,x_2\le m-2\\ x_1+x_2\le m-2}}(1+\mathbb{1}_{x_1>0})(1+\mathbb{1}_{x_2>0})f(x_1)f(x_2)g(m-1-x_1-x_2). \qquad \text{(A.10)}$$

Proof. The proof is essentially the same as the proof of Proposition A.2. □

Appendix B. Proof of Multiple Regimes

In this section, we prove the existence of multiple regimes of eigenvalues for {GOE, k-checkerboard} and {k-checkerbord, j-checkerboard}. Our method involves decomposing each checkerboard matrix into the sum of its mean matrix and perturbation matrix and applying Weyl's inequality to bound the eigenvalue of the matrix ensemble in terms of the eigenvalue of its components. For the sake of simplicity, throughout this section we assume that the weight $w=1$ and that $k|N$ for {GOE, k-checkerboard} and $jk|N$ for k{-checkerboard, j-checkerboard}.

Definition B.1 (Mean Matrix)**.** *The mean matrix $\overline{A}_N$ of the k-checkerboard matrix $A_N=(a_{ij})$ is given by*

$$\bar{a}_{ij}=\begin{cases}0, & \text{if } i\not\equiv j \mod k\\ 1, & \text{if } i\equiv j \mod k.\end{cases} \qquad \text{(B.1)}$$

We note that the rank of $\overline{A}_N$ is k.

Definition B.2 (Perturbation Matrix)**.** *The perturbation matrix $\tilde{A}_N$ of the k-checkerboard matrix $A_N=\tilde{A}_N=(\tilde{a}_{ij})$ is given by*

$$\tilde{a}_{ij}=\begin{cases}a_{ij}, & \text{if } i\not\equiv j \mod k\\ 0, & \text{if } i\equiv j \mod k.\end{cases} \qquad \text{(B.2)}$$

Thus, we can write the k-checkerboard matrix simply as $A_N=\overline{A}_N+\tilde{A}_N$. As shown in [BCDHMSTPY], all the eigenvalues of $\tilde{A}_N$ are $O(N^{1/2})$. Moreover, $\overline{A}_N$ has k eigenvalues at $\frac{N}{k}$ and $N-k$ eigenvalues at 0.

Lemma B.3. *Let $\tilde{A}_N$ be the perturbation matrix as defined above and B_N an $N \times N$ GOE matrix, then with probability $1 - o(1)$, all the eigenvalues of $\{\tilde{A}_N, B_N\}$ are $O(N)$.*

Proof. We know that the maximum eigenvalue of a matrix is equal to the operator norm of the matrix. Since $||\tilde{A}_N|| = O(N^{1/2})$ and $||B_N|| = O(N^{1/2})$, then by submultiplicativity of the matrix norm, $||\tilde{A}_N B_N|| \leq ||\tilde{A}_N||||B_N|| = O(N)$. Similarly, $||B_N \tilde{A}_N|| = O(N)$. By Weyl's inequality, $\lambda_N(\{\tilde{A}_N, B_N\}) \leq \lambda_N(\tilde{A}_N B_N) + \lambda_N(B_N \tilde{A}_N) = O(N)$. The argument for the smallest negative eigenvalues follows from considering $-\tilde{A}_N$ and $-B_N$. □

Lemma B.4. *Let $\overline{A}_N$ be the mean matrix as defined above and B_N the $N \times N$ GOE matrix, then the largest eigenvalue of $\{\overline{A}_N, B_N\}$ is bounded above by $\frac{4N^{3/2}}{k}$, the smallest eigenvalue is bounded below by $-\frac{4N^{3/2}}{k}$, and there are at least $N - 2k$ eigenvalues at 0.*

Proof. First, observe that $\text{rank}(\overline{A}_N B_N) \leq \min(\text{rank}(\overline{A}_N), \text{rank}(B_N)) = k$. Similarly, $\text{rank}(B_N \overline{A}_N) \leq k$. By the subadditivity of rank, $\text{rank}(\{\overline{A}_N, B_N\}) \leq 2k$. Thus, at least $N - 2k$ eigenvalues are 0. For the highest eigenvalues, we see that $||\{\overline{A}_N, B_N\}|| \leq 2||\overline{A}_N||||B_N|| = 2 \cdot \frac{N}{k} \cdot 2N^{1/2} = \frac{4N^{3/2}}{k}$. Similarly, the smallest eigenvalue is bounded below by $-\frac{4N^{3/2}}{k}$. □

Empirically, we observe that $\overline{A}_N$ has k blip eigenvalues at $\frac{N^{3/2}}{k} + O(N)$ and k blip eigenvalues at $-\frac{N^{3/2}}{k} + O(N)$. By assuming this, we are able to prove the existence of multiple regimes for $\{A_N, B_N\}$, as follows:

Lemma B.5. *Let A_N be a k-checkerboard matrix and B_N an $N \times N$ GOE matrix, then $\{A_N, B_N\}$ has a blip containing k eigenvalues at $\frac{N^{3/2}}{k} + O(N)$, a blip containing k eigenvalues at $-\frac{N^{3/2}}{k} + O(N)$, and $N - 2k$ eigenvalues of $O(N)$.*

Proof. First note that we can write $\{A_N, B_N\} = \{\overline{A}_N, B_N\} + \{\tilde{A}_N, B_N\}$. By Weyl's inequality, we see that

$$\lambda_{N-k+1}(\{A_N, B_N\}) \;\geq\; \lambda_{N-k+1}(\{\overline{A}_N, B_N\}) + \lambda_1(\{\tilde{A}_N, B_N\}) \;=\; \frac{1}{k}N^{3/2} + O(N) \tag{B.3}$$

and

$$\lambda_N(\{A_N, B_N\}) \;\leq\; \lambda_N(\{\overline{A}_N, B_N\}) + \lambda_N(\{\tilde{A}_N, B_N\}) \;=\; \frac{1}{k}N^{3/2} + O(N). \tag{B.4}$$

So this proves the existence of k blip eigenvalues at $\frac{N^{3/2}}{k}$. Similarly, we can use Weyl's inequality to show the existence of blip eigenvalues at $-\frac{N^{3/2}}{k}$. For the bulk, we see that

$$\lambda_{N-k}(\{A_N, B_N\}) \;\leq\; \lambda_{N-k}(\{\overline{A}_N, B_N\}) + \lambda_N(\{\tilde{A}_N, B_N\}) \;=\; O(N), \tag{B.5}$$

and

$$\lambda_{k+1}(\{A_N, B_N\}) \;\geq\; \lambda_{k+1}(\{\overline{A}_N, B_N\}) + \lambda_1(\{\tilde{A}_N, B_N\}) \;=\; O(N). \tag{B.6}$$

This completes the proof for the existence of three different regimes. □

Now we consider $\{A_N, B_N\}$, where A_N is a k-checkerboard matrix and B_N is a j-checkerboard matrix. We assume $\gcd(k, j) = 1, N \mid kj$. Then we can write $\{A_N, B_N\} = \{\tilde{A}_N, \tilde{B}_N\} + \{\overline{A}_N, \tilde{B}_N\} + \{\tilde{A}_N, \overline{B}_N\} + \{\overline{A}_N, \overline{B}_N\}$. Similarly, we see that all eigenvalues of $\{\tilde{A}_N, \tilde{B}_N\}$ are of $O(N)$. Empirically, $\{\overline{A}_N, \tilde{B}_N\}$ has k eigenvalues at $\frac{1}{k}\sqrt{1 - \frac{1}{j}}N^{3/2}$, k eigenvalues at $-\frac{1}{k}\sqrt{1 - \frac{1}{j}}N^{3/2}$, and the remaining $N - 2k$ eigenvalues at 0. Heuristically, this can be seen from the fact that $\overline{A}_N$ has k eigenvalues at $\frac{1}{k}N$ and the eigenvalues of $\tilde{B}_N$ are bounded above and below by $\pm 2\sqrt{1 - \frac{1}{j}}N^{1/2}$. Similarly, empirically $\{\tilde{A}_N, \overline{B}_N\}$ has j eigenvalues at $\frac{1}{j}\sqrt{1 - \frac{1}{k}}N^{3/2}$, j eigenvalues at $-\frac{1}{j}\sqrt{1 - \frac{1}{k}}N^{3/2}$, and the remaining $N - 2j$ eigenvalues are at 0. For $\{\overline{A}_N, \tilde{B}_N\} + \{\tilde{A}_N, \overline{B}_N\}$, we observe that the largest eigenvalue is of $O(N^{3/2})$ but larger than $\frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}$ and $\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}$, the smallest eigenvalue is of $O(N^{3/2})$ but smaller than $-\frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}$ and $-\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}$. Furthermore, There are $k-1$ eigenvalues at each of $\pm\frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}$, and $j-1$ eigenvalues at each of $\pm\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}$, and the remaining $N-2k-2j+3$ eigenvalues of $O(N)$.

Lemma B.6. *Let $\overline{A}_N$ and $\overline{B}_N$ be average matrices as defined above, then $\{\overline{A}_N, \overline{B}_N\}$ has 1 eigenvalue exactly at $\frac{2N^2}{jk}$ and $N-1$ eigenvalues at 0.*

Proof. Since j and k are relatively prime, then from matrix multiplication, we see that $\overline{A}_N\overline{B}_N$ and $\overline{B}_N\overline{A}_N$ are both the constant matrix where every entry is $\frac{N}{kj}$. Hence, $\{\overline{A}_N, \overline{B}_N\}$ is the constant matrix where every entry is $\frac{2N}{kj}$. Such matrix has 1 eigenvalue exactly at $\frac{2N^2}{kj}$ and $N-1$ eigenvalues at 0. $\square$

Lemma B.7. *Let A_N be an $N\times N$ k-checkerboard matrix, and B_N an $N\times N$ j-checkerboard matrix such that $\gcd(j,k)=1$ and $jk|N$. Assume without loss of generality that $2\le k<j$. Then the eigenvalues of $\{A_N, B_N\}$ are distributed as follows:*

1. *1 eigenvalue at $\frac{2}{jk}N^2 + O(N^{3/2})$,*
2. *$k-1$ eigenvalues at each of $\pm\frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}+O(N)$,*
3. *1 eigenvalue between $\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N)$ and $\frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}+O(N)$ and 1 eigenvalue between $-\frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}+O(N)$ and $-\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N)$,*
4. *$j-2$ eigenvalues at $\pm\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N)$,*
5. *1 eigenvalue between $O(N)$ (positive) and $\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N)$ and 1 eigenvalue between $O(N)$ (negative) and $-\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N)$,*
6. *The remaining $N-2k-2j+1$ eigenvalues of $O(N)$.*

Proof. By assumption, we have $2\le k<j$, so $\frac{1}{k}\sqrt{1-\frac{1}{j}} > \frac{1}{j}\sqrt{1-\frac{1}{k}}$. Since $\{A_N, B_N\} = \{\tilde{A}_N, \tilde{B}_N\} + \{\overline{A}_N, \tilde{B}_N\} + \{\tilde{A}_N, \overline{B}_N\} + \{\overline{A}_N, \overline{B}_N\}$, then

$$\begin{aligned}\lambda_N(\{A_N, B_N\}) &\ge \lambda_N(\{\overline{A}, \overline{B}\}) + \lambda_1(\{\overline{A}, \tilde{B}\} + \{\tilde{A}, \overline{B}\} + \{\tilde{A}, \tilde{B}\}) \\ &\ge \lambda_N(\{\overline{A}, \overline{B}\}) + \lambda_1(\{\overline{A}, \tilde{B}\}) + \lambda_1(\{\tilde{A}, \overline{B}\}) + \lambda_1(\{\tilde{A}, \tilde{B}\}) \;=\; \frac{2N^2}{jk} + O(N^{3/2}).\end{aligned} \tag{B.7}$$

This establishes the existence of the largest blip. Then, we establish the existence of the intermediary blip containing $k-1$ eigenvalues at $\frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}+O(N)$:

$$\begin{aligned}\lambda_{N-1}(\{A_N, B_N\}) &\le \lambda_{N-1}(\{\overline{A}_N, \overline{B}_N\}) + \lambda_N(\{\overline{A}, \tilde{B}\} + \{\tilde{A}, \overline{B}\} + \{\tilde{A}, \tilde{B}\}) \\ &\le \lambda_N(\{\overline{A}_N, \tilde{B}_N\} + \{\tilde{A}_N, \overline{B}_N\}) + \lambda_N(\{\tilde{A}_N, \tilde{B}_N\}) \;=\; \frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2} + O(N), \\ \lambda_{N-k+1}(\{A_N, B_N\}) &\ge \lambda_1(\{\overline{A}_N, \overline{B}_N) + \lambda_{N-k+1}(\{\overline{A}, \tilde{B}\} + \{\tilde{A}, \overline{B}\} + \{\tilde{A}, \tilde{B}\}) \\ &\ge \lambda_{N-k+1}(\{\overline{A}_N, \tilde{B}_N\} + \{\tilde{A}_N, \overline{B}_N\}) + \lambda_1(\{\tilde{A}_N, \tilde{B}_N) = \frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2} + O(N).\end{aligned} \tag{B.8}$$

Next, we show that there is one eigenvalue between $\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N)$ and $\frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}+O(N)$ as well as $j-2$ eigenvalues at $\frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N)$:

$$\begin{aligned}
\lambda_{N-k}(\{A_N,B_N\}) &\leq \lambda_{N-k+1}(\{\overline{A}_N,\tilde{B}_N\}+\{\tilde{A}_N,\overline{B}_N\})+\lambda_{N-1}(\{\overline{A}_N,\overline{B}_N\}+\{\tilde{A}_N,\tilde{B}_N\})\\
&= \frac{1}{k}\sqrt{1-\frac{1}{j}}N^{3/2}+O(N),\\
\lambda_{N-k-1}(\{A_N,B_N\}) &\leq \lambda_{N-k}(\{\overline{A}_N,\tilde{B}_N\}+\{\tilde{A}_N,\overline{B}_N\})+\lambda_{N-1}(\{\overline{A}_N,\overline{B}_N\}+\{\tilde{A}_N,\tilde{B}_N\})\\
&= \frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N),\\
\lambda_{N-k-j+2}(\{A_N,B_N\}) &\geq \lambda_{1}(\{\overline{A}_N,\overline{B}_N\}+\{\tilde{A}_N,\tilde{B}_N\})+\lambda_{N-j-k+2}(\{\overline{A}_N,\tilde{B}_N\}+\{\tilde{A}_N,\overline{B}_N\})\\
&= \frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N).
\end{aligned}\tag{B.9}$$

By symmetry argument, we can use Weyl's inequality to establish the existence of their negative counterpart. Finally, for the bulk, we have

$$\begin{aligned}
\lambda_{N-j-k+1}(\{A_N,B_N\}) &\leq \lambda_{N-j-k+2}(\{\overline{A}_N,\tilde{B}_N\}+\{\tilde{A}_N,\overline{B}_N\})+\lambda_{N-1}(\{\overline{A}_N,\overline{B}_N\}+\{\tilde{A}_N,\tilde{B}_N\})\\
&\leq \frac{1}{j}\sqrt{1-\frac{1}{k}}N^{3/2}+O(N),\\
\lambda_{N-j-k}(\{A_N,B_N\}) &\leq \lambda_{N-j-k+1}(\{\overline{A}_N,\tilde{B}_N\}+\{\tilde{A}_N,\overline{B}_N\})+\lambda_{N-1}(\{\overline{A}_N,\overline{B}_N\}+\{\tilde{A}_N,\tilde{B}_N\})\\
&= O(N).
\end{aligned}\tag{B.10}$$

By symmetry argument, we can use Weyl's inequality to bound the bulk from the below. This completes the proof. □

Note that in the proof of Lemma B.7, Weyl's inequality fails to provide an accurate bound on the four eigenvalues between different regimes. Empirically, we observe that among those eigenvalues, the positive ones belong to the regimes corresponding to their lower bound, and the negative ones belong to the regimes corresponding to their upper bound.

APPENDIX C. ALMOST SURE CONVERGENCE

The traditional way to show weak convergence of empirical spectral measures to a limiting spectral measure (in probability or almost-surely) is to show that the variance (for convergence in probability) or fourth moment (for almost-sure convergence) of the m^{th} moment, averaged over the $N\times N$ ensemble, is $O(\frac{1}{N})$ for probability or $O(\frac{1}{N^2})$ for almost-sure. In the case of the blip spectral measure, neither of these methods gives the desired result. However, for the anticommutator of a j-checkerboard and k-checkerboard matrices, there is only one eigenvluae in the far away blip so the measure is only a single delta spike whose distribution and is described in Theorem 3.17 through their moments. As such, for fixed j and k the variance and fourth moment over the ensemble of the general m^{th} moment do not go to 0 which means that we cannot use the previous methods. We therefore define a modified spectral measure which averages over the eigenvalues of many matrices in order to extend standard techniques, in particular we link the moments to the moments of the $k\times k$ Gaussian Wigner matrix using similar methods as [BCDHMSTPY].

In order to facilitate the proof of the main convergence result (Theorem C.5) we first introduce some new notation. In all that follows we fix k and suppress k-dependence in our notation. Let Ω_N be the probability space of $N\times N$ k-checkerboard matrices with the natural probability measure. Then we define the product probability space

$$\Omega := \prod_{N \in \mathbb{N}} \Omega_N. \tag{C.1}$$

By Kolmogorov's extension theorem (see [Tao2]), this is equipped with a probability measure which agrees with the probability measures on Ω_N when projected to the N^{th} coordinate. Given $\{A_N\}_{N\in\mathbb{N}} \in \Omega$, we denote by A_N the $N \times N$ matrix given by projection to the N^{th} coordinate. In what follows, we suppress the subscript $N \in \mathbb{N}$ on elements of Ω, writing them as $\{A_N\}$.

Remark C.1. [KKMSX] *employs a similar construction using product space, while* [HM] *views elements of Ω as infinite matrices and the projection map $\Omega \to \Omega_N$ as simply choosing the upper left $N \times N$ minor.*

We treated the m^{th} moment of an empirical spectral measure $\mu_{A,N}^{(m)}$ as a random variable on Ω_N, but similarly to [BCDHMSTPY] we may treat it as a random variable on Ω. To highlight this, we can define the random variable $X_{m,N}$ on Ω

$$X_{m,N}(\{A_N\}) := \mu_{A_N,N}^{(m)}. \tag{C.2}$$

These have centered r^{th} moment

$$X_{m,N}^{(r)} := \mathbb{E}[(X_{m,N} - \mathbb{E}[X_{m,N}])^r]. \tag{C.3}$$

From the discussion at the beginning of this section, we wish to average over a growing number of matrices of the same size, since we reduce to $k \times k$ Gaussian Winger matrices, so it is advantageous to work over $\Omega^{\mathbb{N}}$; this again is equipped with a natural probability measure by Kolmogorov's extension theorem. Its elements are sequences of sequences of matrices, and we denote them by $\overline{A} = \{A^{(i)}\}_{i\in\mathbb{N}}$ where $A^{(i)} \in \Omega$. We now give a more abstract definition of the averaged blip spectral measure.

Definition C.2. *Fix a function $g : \mathbb{N} \to \mathbb{N}$. The **averaged empirical blip spectral measure** associated to $\overline{A} \in \Omega^{\mathbb{N}}$ is*

$$\mu_{N,g,\overline{A}} := \frac{1}{g(N)} \sum_{i=1}^{g(N)} \mu_{A_N^{(i)},N}. \tag{C.4}$$

In other words, we project onto the N^{th} coordinate in each copy of Ω and then average over the first $g(N)$ of these $N \times N$ matrices.

Remark C.3. *If one wishes to avoid defining an empirical spectral measure which takes eigenvalues of multiple matrices, one may use the construction of a $\mathbb{N} \times \mathbb{N}$ block matrix with independent $N \times N$ checkerboard matrix blocks.*

Analogously to $X_{m,N}$, we denote by $Y_{m,N,g}$ the random variable on $\Omega^{\mathbb{N}}$ defined by the moments of the averaged empirical blip spectral measure

$$Y_{m,N,g}(\overline{A}) := \mu_{N,g,\overline{A}}^{(m)}. \tag{C.5}$$

The centered r^{th} moment (over $\Omega^{\mathbb{N}}$) of this random variable is denoted by $Y_{m,N,g}^{(r)}$.

We now prove almost-sure weak convergence of the averaged blip spectral measures under the assumption that g grows as a power of N. Recall the following definition.

Definition C.4. *A sequence of random measures $\{\mu_N\}_{N\in\mathbb{N}}$ on a probability space Ω converges **weakly almost-surely** to a fixed measure μ if, with probability 1 over $\Omega^{\mathbb{N}}$, we have*

$$\lim_{N\to\infty} \int f d\mu_N = \int f d\mu \tag{C.6}$$

for all $f \in \mathcal{C}_b(\mathbb{R})$ (continuous and bounded functions).

Theorem C.5. *Let $g : \mathbb{N} \to \mathbb{N}$ be such that there exists an $\delta > 0$ for which $g(N) = \omega(N^\delta)$. Then, as $N \to \infty$, the averaged empirical spectral measures $\mu_{N,g,\overline{A}}$ of the j-checkerboard and k-checkerboard anticommutator ensemble converge weakly almost-surely to the measure with moments $M_{k,m} = \frac{1}{k}\mathbb{E}Tr[C^m]$, the limiting expected moments computed in Theorem 3.17.*

Proof. For simplicity of notation, we can fix k and simply denote $M_{k,m}$ by M_m. By the triangle inequality, we have

$$|Y_{m,N,g} - M_m| \;\leq\; |Y_{m,N,g} - \mathbb{E}[Y_{m,N,g}]| + |\mathbb{E}[Y_{m,N,g}] - M_m|. \tag{C.7}$$

From Theorem 3.17, we know that $\mathbb{E}[X_{m,N}] \to M_m$, and it follows that $\mathbb{E}[Y_{m,N,g}] \to M_m$. Hence to show that $Y_{m,N,g} \to M_m$ almost surely, it suffices to show that $|Y_{m,N,g} - \mathbb{E}[Y_{m,N,g}]| \to 0$ almost surely as $N \to \infty$. We show that the limit as $N \to \infty$ of all moments over Ω_N of any arbitrary moment of the empirical spectral measure exists, and that we may always choose a sufficiently high moment[1] such that the standard method of Chebyshev's inequality and the Borel-Cantelli lemma gives that $|Y_{m,N,g} - \mathbb{E}[Y_{m,N,g}]| \to 0$. Finally, the moment convergence theorem gives almost-sure weak convergence to the limiting averaged blip spectral measure.

Lemma C.6. *Let $X_{m,N}$ be as defined in* (C.2)*. Then for any $t \in \mathbb{N}$, the r^{th} centered moment of $X_{m,N}$ satisfies*

$$X_{m,N}^{(r)} \;=\; \mathbb{E}[(X_{m,N} - \mathbb{E}[X_{m,N}])^r] \;=\; O_{m,r}(1) \tag{C.8}$$

as N goes to infinity.

Proof. Firstly, we have

$$\begin{aligned}\mathbb{E}[(X_{m,N} - \mathbb{E}[X_{m,N}])^r] \;&=\; \mathbb{E}\left[\sum_{\ell=0}^{r}\binom{r}{\ell}(X_{m,N})^\ell\,(\mathbb{E}[X_{m,N}])^{r-\ell}\right] \\ &=\; \sum_{\ell=0}^{r}\binom{r}{\ell}\mathbb{E}\left[(X_{m,N})^\ell\right](\mathbb{E}[X_{m,N}])^{r-\ell}.\end{aligned} \tag{C.9}$$

From the moments given in Section 3.17, we have $\mathbb{E}[X_{m,N}] = O_m(1)$ hence $(\mathbb{E}[X_{m,N}])^{r-\ell} = O_{m,r,\ell}(1)$ for all ℓ. As such, it suffices to show that $\mathbb{E}[(X_{m,N})^\ell] = O_{m,\ell}(1)$. By (3.17), we have that

$$\begin{aligned}&\mathbb{E}[X_{m,N,1}]^l \\ &= \mathbb{E}\left[\frac{1}{k_1}\sum_{\alpha=2n}^{8n} c_\alpha\left(\frac{k}{N^{3/2}}\right)^\alpha\left(\frac{1}{N^m}\sum_{i=0}^{m}\binom{m}{i}\left(-\frac{N^{3/2}}{k}\right)^{m-i}\mathrm{Tr}[\{A_N,B_N\}^{\alpha+i}]\right)^l\right] \\ &= \mathbb{E}\left[\sum_{\substack{2n\leq\alpha_1\leq 8n\\ 0\leq i_1\leq m}}\sum_{\substack{2n\leq\alpha_2\leq 8n\\ 0\leq i_2\leq m}}\cdots\sum_{\substack{2n\leq\alpha_l\leq 8n\\ 0\leq i_l\leq m}}\frac{1}{N^{m\ell}}\prod_{\nu=1}^{l}c_{\alpha_v}\binom{m}{i_v}(-1)^{m-i_v}\left(\frac{N^{3/2}}{k}\right)^{l-i_v}\mathrm{Tr}[\{A_N,B_N\}^{\alpha+i_v}]\right] \\ &= \sum_{\substack{2n\leq\alpha_1\leq 8n\\ 0\leq i_1\leq m}}\sum_{\substack{2n\leq\alpha_2\leq 8n\\ 0\leq i_2\leq m}}\cdots\sum_{\substack{2n\leq\alpha_l\leq 8n\\ 0\leq i_l\leq m}}\frac{1}{N^{m\ell}}\prod_{\nu=1}^{l}c_{\alpha_v}\binom{m}{i_v}(-1)^{m-i_v}\left(\frac{N^{3/2}}{k}\right)^{l-i_v}\mathbb{E}\left[\prod_{v=1}^{l}\mathrm{Tr}[\{A_N,B_N\}^{\alpha+i_v}]\right].\end{aligned} \tag{C.10}$$

[1]Note the difference between this and the standard technique of, for instance, [HM], which uses only the fourth moment.

Now, recall that

$$\mathbb{E}\left[\prod_{v=1}^{\ell} \mathrm{Tr} A^{2n+i_v}\right] = \sum_{\alpha_1^1,\dots,\alpha_{2n+i_1}^1 \le N} \cdots \sum_{\alpha_1^\ell,\dots,\alpha_{2n+i_\ell}^\ell \le N} \mathbb{E}\left[\prod_{j=1}^{\ell} a_{\alpha_1^j,\alpha_2^j} \dots a_{\alpha_{2n+i_j}^j,\alpha_1^j}\right]. \tag{C.11}$$

We have now reached a combinatorial problem similar to the one we encounter in Section 2. For each j, since the length of the cyclic product $a_{\alpha_1^j,\alpha_2^j} \dots a_{\alpha_{2n+i_j}^j,\alpha_1^j}$ is fixed at $2n+i_j$, we can choose the number of blocks (determining the class), the location of the blocks (determining the configuration), the matchings and indexings. By Lemma 3.13 and 3.14, we have that the main contribution from configurations of length $(2n+i_j)$ in B_j-class is $\frac{(2n+i_j)^{B_j}}{B_j!}$. By the same arguments made in Section 2, the number of ways we can choose the number of blocks having one a and two a's as well as the number of ways to choose matchings across the ℓ cyclic products are independent of N, j's and i_j's, so for simplicity, we are denoting them as C. Finally, the contribution from choosing the indices of all the blocks and w's is $O_k(N^{2n\ell+i_1+\cdots+i_\ell-B_1-\cdots-B_\ell})$. As such, if $B_1,\dots,B_\ell \ge m$, the total contribution is $O_{m,k}(1)$. If there exists $B_{j'} < m$, then the overall contribution is

$$CN^{\ell m-B_1-\cdots-B_\ell} \prod_{u=1}^{\ell}\left[\sum_{j_u=0}^{2n}\binom{2n}{j_u}\sum_{i_u=0}^{m+j_u}\binom{m+j_u}{i_u}(-1)^{m-i_u}\frac{(2n+i_u)^{B_u}}{B_u!}\right] = 0, \tag{C.12}$$

since the sum over $j_u = j'$ is equal to 0 by by Lemma 3.16. As such, the total contribution of $\mathbb{E}[X_{m,N}^\ell]$ is simply $O_{m,\ell}(1)$ (suppressing k), as desired. □

We apply the following theorem (Theorem 1.2 of [Fer]) with $X = X_{m,N} - \mathbb{E}[X_{m,N}]$, $s = g(N)$ and $\mu_i = X_{m,N}^{(i)}$.

Theorem C.7. *Let $r \in \mathbb{N}$ and let $X_1,\dots,X_s$ be i.i.d. copies of some mean-zero random variable X with absolute moments $\mathbb{E}[|X|^\ell] < \infty$ for all $\ell \in \mathbb{N}$. Then*

$$\mathbb{E}\left[\left(\sum_{i=1}^{s} X_i\right)^r\right] = \sum_{1\le m\le \frac{r}{2}} B_{m,r}(\mu_2,\mu_3,\dots,\mu_r)\binom{s}{m} \tag{C.13}$$

where the μ_i are the moments of X and $B_{m,r}$ is a function independent of s, the details of which are given in [Fer].

We must first show boundedness of the absolute moments of $X_{m,N}$. By Cauchy-Schwarz,

$$\left(\int |x^{2\ell+1}| d\mu_{X_{m,N}}\right)^2 \le \int |x|^2 d\mu_{X_{m,N}} \cdot \int |x|^{4\ell} d\mu_{X_{m,N}}, \tag{C.14}$$

where $\mu_{X_{m,N}}$ is the probability measure on Ω given by the density of $X_{m,N}$. Since, for fixed N, the even moments of $X_{m,N}$ are finite by (C.10), the previous bound shows that all odd absolute moments are finite as well. Hence Theorem C.7 applies, yielding

$$\mathbb{E}\left[\left(\sum_{i=1}^{g(N)} X_{m,N,i} - \mathbb{E}[X_{m,N,i}]\right)^r\right] = \sum_{1\le m\le \frac{r}{2}} B_{m,r}(X_{m,N}^{(2)}, X_{m,N}^{(3)},\dots,X_{m,N}^{(r)})\binom{g(N)}{m}. \tag{C.15}$$

where the $X_{m,N,i}$ are i-indexed i.i.d. copies of $X_{m,N}$. By Lemma C.6, for sufficiently high N, $X_{m,N}^{(t)}$ are uniformly bounded above by some constant K for $1 \le t \le m$, so there exists C such that

$B_{m,r}(X^{(2)}_{m,N}, X^{(3)}_{m,N}, \dots, X^{(r)}_{m,N}) < C$ for all sufficiently large N and for all $1 \le m \le r/2$. Hence

$$\mathbb{E}\left[\left(\sum_{i=1}^{g(N)} X_{m,N,i} - \mathbb{E}[X_{m,N,i}]\right)^r\right] \;\le\; \sum_{1\le m\le \frac{r}{2}} C\binom{g(N)}{m}. \tag{C.16}$$

As such, we have

$$Y^{(r)}_{m,N,g} = \frac{1}{g(N)^r}\mathbb{E}\left[\left(\sum_{i=1}^{g(N)} X_{m,N,i} - \mathbb{E}[X_{m,N,i}]\right)^r\right] \le \sum_{1\le m\le \frac{r}{2}} \frac{C}{g(N)^r}\binom{g(N)}{m} = O\left(\frac{1}{g(N)^{r/2}}\right). \tag{C.17}$$

Since $g(N) = \omega(N^\delta)$, we may choose r sufficiently large so that

$$Y^{(r)}_{m,N,g} \;=\; O\left(\frac{1}{N^2}\right). \tag{C.18}$$

Then by Chebyshev's inequality,

$$\mathbb{P}(|Y_{m,N,g} - \mathbb{E}[Y_{m,N,g}]| > \epsilon) \;\le\; \frac{\mathbb{E}[(Y_{m,N,g} - \mathbb{E}[Y_{m,N,g}])^r]}{\epsilon^r} \;=\; \frac{Y^{(r)}_{m,N,g}}{\epsilon^r} \;=\; O\left(\frac{1}{N^2}\right). \tag{C.19}$$

We now apply the following (see for example [Can]).

Lemma C.8 (Borel-Cantelli)**.** *Let B_i be a sequence of events with $\sum_i \mathbb{P}(B_i) < \infty$. Then*

$$\mathbb{P}\left(\bigcap_{j=1}^{\infty}\bigcup_{\ell=j}^{\infty} B_\ell\right) \;=\; 0. \tag{C.20}$$

Define the events

$$B_N^{(m,d,g)} \;:=\; \left\{A \in \Omega^{\mathbb{N}} : |Y_{m,N,g}(A) - \mathbb{E}[Y_{m,N,g}]| \;\ge\; \frac{1}{d}\right\}. \tag{C.21}$$

Then $\mathbb{P}(B_N^{(m,d,g)}) \;\le\; \frac{C_m d^r}{N^2}$, so for fixed m, d, the conditions of the Borel-Cantelli lemma are satisfied. Hence

$$\mathbb{P}\left(\bigcap_{j=1}^{\infty}\bigcup_{\ell=j}^{\infty} B_\ell^{(m,d,g)}\right) \;=\; 0. \tag{C.22}$$

Taking a union of these measure-zero sets over $d \in \mathbb{N}$ we have

$$\mathbb{P}\left(Y_{m,N,g} \neq \mathbb{E}[Y_{m,N,g}] \text{ for infinitely many } N\right) \;=\; 0, \tag{C.23}$$

and taking the union over $m \in \mathbb{Z}_{\ge 0}$,

$$\mathbb{P}\left(\exists m \text{ such that } Y_{m,N,g} \neq \mathbb{E}[Y_{m,N,g}] \text{ for infinitely many } N\right) \;=\; 0. \tag{C.24}$$

Therefore with probability 1 over $\Omega^{\mathbb{N}}$, $|Y_{m,N,g} - \mathbb{E}[Ym,N,g]| \to 0$ for each m. This, together with (C.7) and the discussion following it, yields that the moments $\mu^{(m)}_{N,g} = Y_{m,N,g} \to M_m$ almost surely. We now use the following to show almost-sure weak convergence of measures (see for example [Ta]).

Theorem C.9 (Moment Convergence Theorem)**.** *Let μ be a measure on $\mathbb{R}$ with finite moments $\mu^{(m)}$ for all $m \in \mathbb{Z}_{\ge 0}$, and $\mu_1, \mu_2, \dots$ a sequence of measures with finite moments $\mu_n^{(m)}$ such that $\lim_{n\to\infty} \mu_n^{(m)} = \mu^{(m)}$ for all $m \in \mathbb{Z}_{\ge 0}$. If in addition the moments $\mu^{(m)}$ uniquely characterize a measure, then the sequence μ_n converges weakly to μ.*

Since the Moment Convergence Theorem C.9 is usually used when measures are assumed to be probability measures, this is somewhat stronger. A full proof of this is given in Appendix D of [BCDHMSTPY].

Theorem C.10 (Carleman's condition)**.** *Let μ be a measure with all moments $\mu^{(m)}$ finite for all $m \geq 0$ If*

$$\sum_{n\geq 1}(\mu^{(2n)})^{-\frac{1}{2n}} = \infty \tag{C.25}$$

then μ is the unique measure with moments $\mu^{(m)}$.

To show Carleman's condition is satisfied for the limiting moments M_m, we show that M_m are bounded above by the moments of the Gaussian. The odd moments vanish, and by Theorem 3.17 the even moments are given by

$$M_{2m} \ = \ \frac{2}{k}\mathbb{E}[\text{Tr } A^{2m}] \ = \sum_{1\leq i_1,\dots,i_{2m}\leq k} 2\mathbb{E}[a_{i_1i_2}a_{i_2i_3}\dots a_{i_{2m}i_1}], \tag{C.26}$$

and as $\mathbb{E}[a_{i_1i_2}a_{i_2i_3}\dots a_{i_ni_1}]$ is maximized when all $a_{i_\ell i_{\ell+1}}$ are equal,

$$M_{2m} \ \leq \sum_{1\leq i_1,\dots,i_{2m}\leq k} 2(2m-1)!! \ = \ 2k^{2m}(2m-1)!!. \tag{C.27}$$

These are bounded by the moments of $\mathcal{N}(0,2k)$ so Carleman's condition is satisfied, thus we let $\overline{\mu}$ be the unique measure determined by the moments M_m. Choose $\overline{A} \in \Omega^{\mathbb{N}}$. Then the preceding argument showed that, with probability 1 over $\overline{A}$ chosen from $\Omega^{\mathbb{N}}$, all moments $\mu^{(m)}_{N,g,\overline{A}}$ of the measures $\mu_{N,g,\overline{A}}$ converge to M_m. Then by Theorem C.9 the measures $\mu_{N,g,\overline{A}}$ converge weakly to $\overline{\mu}$ with probability 1, completing the proof. □

Remark C.11. *These same methods can be used to prove the convergence of the intermediate blip. This case has some slight subtleties since we square the distribution and then center it. However when we consider the squared and centered distribution we are able to apply the same methods from this appendix and use Carleman's condition to similarly get convergence for the intermediate blip.*

Appendix D. Almost-Sure Convergence of the Bulk

Theorem D.1. *For A_N and B_N both $N \times N$ $(k,0)$-checkerboard matrices we get that for any fixed ℓ*

$$\lim_{N\to\infty} \mathit{Var}[\nu^{(\ell)}_{\{A_N,B_N\}}] \ = \ O\left(\frac{1}{N^2}\right). \tag{D.1}$$

Proof. We know that by the eigenvalue trace lemma, we have

$$\begin{aligned}\text{Var}\left[\nu^{(\ell)}_{\{A_N,B_N\}}\right] \ &= \ \left|\mathbb{E}[(\nu_{\{A_N,B_N\}^{(\ell)}})^2] - \left[\mathbb{E}[\nu_{\{A_N,B_N\}}]\right]^2\right| \\ &= \frac{1}{N^{2\ell+2}}\left|\mathbb{E}[\text{Tr}(\{A_N,B_N\}^\ell)^2] - (\mathbb{E}[\text{Tr}\{A_N,B_N\}^\ell])^2\right| \ = \ \frac{1}{N^{2\ell+2}}\sum_{I,I'}|\mathbb{E}[\zeta_I\zeta_{I'}] - \mathbb{E}[\zeta_I]\mathbb{E}[\zeta_{I'}]|\end{aligned} \tag{D.2}$$

where the ζ_I and $\zeta_{I'}$ are stand ins for a product $c_{i_1,i_2}c_{i_2,i_3}\cdots c_{i_{2\ell},i_1}$ where every c is a or b and is some expansion of $(AB+BA)^\ell$. So then we know that from the proof in [BCDHMSTPY] that for any choices of A's and B's this is $O(1/N^2)$ and since we consider ℓ as fixed we know that there are 2^ℓ different configurations which are constant and for each configuration from the paper we know that they are $O_m(1/N^2)$ which means that if we add up all of these different cases and configurations we get that it is still $O(1/N^2)$ which proves the theorem. This theorem proves convergence when combined with Chebyshev's inequality and the Borel-Cantelli lemma. □

By Chebyshev's inequality we get the first inequality and by Theorem D.1 we get that the sum of variances is finite, giving,

$$\sum_{N=1}^{\infty}\text{Pr}\left(\left|\nu^{(\ell)}_{\{A_N,B_N\}} - \mathbb{E}[\nu^{(\ell)}_{\{A_N,B_N\}}] > \epsilon\right|\right) \ \leq \ \frac{1}{\epsilon^2}\sum_{N=1}^{\infty}\text{Var}(\nu_{\{A_N,B_N\}}) \ < \ \infty. \tag{D.3}$$

So then by the Borel-Cantelli lemma and Theorem D.1 we get that the moments converge almost surely.

Theorem D.2. *For A_N and B_N both $N \times N$ palindromic Toeplitz matrices we get that for any fixed ℓ*

$$\lim_{N\to\infty} \mathbb{E}[|M_m(\{A_N, B_N\}) - \mathbb{E}[M_m(\{A_N, B_N\})]|^4] \;=\; O_m\left(\frac{1}{N^2}\right) \tag{D.4}$$

Proof. Expanding this yields

$$\begin{aligned}&\mathbb{E}[M_m(\{A_N, B_N\})^4] \;-\; 4\mathbb{E}[M_m(\{A_N, B_N\})^3]\mathbb{E}[M_m(\{A_N, B_N\})]\\ &\;+\; 6\mathbb{E}[M_m(\{A_N, B_N\})^2]\mathbb{E}[M_m(\{A_N, B_N\})]^2 \;-\; 3\mathbb{E}[M_m(\{A_N, B_N\})]\mathbb{E}[M_m(\{A_N, B_N\})]^3\end{aligned} \tag{D.5}$$

As the odd moments are all 0 their expected value is always 0l so we only need to consider even moments. So then we can write the terms as

$$\mathbb{E}[M_{2m}(\{A_N, B_N\})^4] \;=\; \frac{1}{N^{8m+4}} \sum_i \sum_j \sum_k \sum_\ell \mathbb{E}[c_{is}c_{js}c_{ks}c_{\ell s}] \tag{D.6}$$

where

$$\mathbb{E}[c_{is}c_{js}c_{\ell s}c_{ks}] = \mathbb{E}[c_{|i_1-i_2|}\cdots c_{|i_{4m}-i_1|}c_{|j_1-j_2|}\cdots c_{|j_{4m}-j_1|}c_{|k_1-k_2|}\cdots c_{|k_{4m}-k_1|}c_{|\ell_1-\ell_2|}\cdots c_{|\ell_{4m}-\ell_1|}], \tag{D.7}$$

where the c's are all a's or b's and a's can only match with a's and b's can only match with b's, however, it suffices to allow matches to be free since the upper bound of $O(1/N^2)$ holds in any case. Then we know that from [HM], given a fixed expansion of a's and b's of the terms in the binomial expansion are $O(N^{8m+2})$. So then since we know that there are 2^{8m} expansions of the anticommutator and m is a fixed constant we get that the total contribution of all of the terms in the anticommutator expansion are $2^{8m} \cdot O(N^{8m+2}) = O(N^{8m+2})$ since 2^{8m} is a fixed constant. □

Remark D.3. *Note that the proof of Theorem D.2 also applies when A_N is a palindromic Toeplitz and B_N is a GOE. This is due to the fact that all matchings of a palindromic Toeplitz matrix and GOE anticommutator are valid in the palindromic Toeplitz case, so whenever a degree of freedom is lost in the palindromic Toeplitz and palindromic-Toeplitz anticommutator it is also lost in the palindromic Toeplitz - GOE case since the GOE case is strictly more restricted and having the same indices implies their differences are equal. So the same argument proves convergence for the case where A_N is palindromic Toeplitz and B_N is a GOE.*

Appendix E. Polynomial Weight Functions for Intermediary Blips

In theory, the exact choice of the polynomial weight function shouldn't affect the moment of the intermediary blips, as long as it satisfies all the required conditions. For the sake of completeness, we include here the expression for the weight functions for the intermediary blips of $\{k$-checkerboard, j-checkerboard$\}$. Let $w_s = \frac{(-1)^{s+1}}{k}\sqrt{1-\frac{1}{j}}$ for $s \in \{1,2\}$ and $w_s = \frac{(-1)^{s+1}}{j}\sqrt{1-\frac{1}{k}}$ for $s \in \{3,4\}$. Then, the weight function for the intermediary blip at $w_s N^{3/2}$ is

$$g_s^{2n}(x) \;=\; \frac{x^{2n}\prod_{i=1;i\neq s}^{4}\left(x - \frac{w_i}{w_s}\right)^{2n}\left(x - \frac{w_5\sqrt{N}}{w_s}\right)^{10n}(x-A)^{2n}}{B^{2n}\left(1 - \frac{w_5\sqrt{N}}{w_s}\right)^{10n}(1-A)^{2n}}, \tag{E.1}$$

where $A = 1 + \left(1 + \sum_{i=1;i\neq s}^{4}\frac{1}{1-w_i/w_s}\right)^{-1}$ and $B = \prod_{i=1;i\neq s}^{4}\left(1 - \frac{w_i}{w_s}\right)$. It's clear that g_s^{2n} has zeros of order $2n$ at 0, $\frac{w_i}{w_s}$ for $i \neq s$, and zero of order $10n$ at $\frac{w_5\sqrt{N}}{w_s}$. We prove that g_s^{2n} vanishes at $O\left(\frac{1}{\sqrt{N}}\right)$, $\frac{w_i}{w_s} + O\left(\frac{1}{\sqrt{N}}\right)$ for $i \neq s$, and $\frac{w_5\sqrt{N}}{w_s} + O\left(\frac{1}{\sqrt{N}}\right)$, is equal to 1 at $1 + O\left(\frac{1}{\sqrt{N}}\right)$, and has a critical point at 1. The key to the proof is the evaluation of the expression $\lim_{N\to\infty}(1 + O(1/\sqrt{N}))^{2n}$. Since $\lim_{y\to\infty}(1+1/y)^y = e$

and $n \ll \log\log(N)$, then $1 \le \lim_{N\to\infty}(1+O(1/\sqrt{N}))^{2n} \le \lim_{N\to\infty}\left((1+O(1/\sqrt{N}))^{O(\sqrt{N})}\right)^{2n/O(\sqrt{N})} = \lim_{N\to\infty} e^{2n/O(\sqrt{N})} = 1$. Hence, $\lim_{N\to\infty}(1+O(1/\sqrt{N}))^{2n}$
$= 1$.

Now, we first evaluate the function at $x = 1 + O\left(\frac{1}{N}\right)$ as $N \to \infty$,

$$\begin{aligned}
&\lim_{N\to\infty} g_s^{2n}\left(1+O\left(\frac{1}{\sqrt{N}}\right)\right) \\
&= \lim_{N\to\infty}\left(1+O\left(\frac{1}{\sqrt{N}}\right)\right)^{2n} \cdot \frac{\left(1+O\left(\frac{1}{\sqrt{N}}\right)-\frac{w_5\sqrt{N}}{w_s}\right)^{10n}}{\left(1-\frac{w_5\sqrt{N}}{w_s}\right)^{10n}} \cdot \frac{\left(1+O\left(\frac{1}{\sqrt{N}}\right)-A\right)^{2n}}{(1-A)^{2n}} \\
&\cdot \frac{\prod_{i=1;i\neq s} 4\left(1+O\left(\frac{1}{\sqrt{N}}\right)-\frac{w_i}{w_s}\right)^{2n}}{B^{2n}} \\
&= \lim_{N\to\infty}\frac{\prod_{i=1;i\neq s}^{4}\left(1+O\left(\frac{1}{\sqrt{N}}\right)-\frac{w_i}{w_s}\right)^{2n}}{B^{2n}} \;=\; 1. \qquad \text{(E.2)}
\end{aligned}$$

Note that we repeatedly apply the evaluation $\lim_{N\to\infty}(1+O(1/\sqrt{N}))^{2n} = 1$ above to simplify the expression. As an example, $\left(\frac{1+O(1/\sqrt{N})-A}{1-A}\right)^{2n} = \left(1+\frac{O(1/\sqrt{N})}{1-A}\right)^{2n} = \left(1+O(1/\sqrt{N})\right)^{2n} = 1$.

Then, we consider the evaluation of the function at $O\left(\frac{1}{\sqrt{N}}\right)$, and the evaluation of the function at other vanishing points similarly follows,

$$\begin{aligned}
&\lim_{N\to\infty} g_s^{2n}\left(O\left(\frac{1}{\sqrt{N}}\right)\right) \\
&= \lim_{N\to\infty}\frac{\prod_{i=1;i\neq s}^{4}\left(O\left(\frac{1}{\sqrt{N}}\right)-\frac{w_i}{w_s}\right)^{2n}}{B^{2n}} \cdot \frac{\left(O\left(\frac{1}{\sqrt{N}}\right)-\frac{w_5\sqrt{N}}{w_s}\right)^{10n}}{\left(1-\frac{w_5\sqrt{N}}{w_s}\right)^{10n}} \cdot \frac{\left(O\left(\frac{1}{\sqrt{N}}\right)-A\right)^{2n}}{(1-A)^{2n}} \cdot \left(O\left(\frac{1}{\sqrt{N}}\right)\right)^{2n} \\
&= \lim_{N\to\infty} C_N \cdot \left(O\left(\frac{1}{\sqrt{N}}\right)\right)^{2n} \;=\; 0. \qquad \text{(E.3)}
\end{aligned}$$

Finally, we prove that g_s^{2n} has a critical point at 1. Using logarithmic derivative, we obtain

$$\frac{(g_s^{2n})'(x)}{g_s^{2n}(x)} \;=\; 2n\cdot\left(\frac{1}{x}+\sum_{i=1;i\neq s}^{4}\frac{1}{x-\frac{w_i}{w_s}}+\frac{5}{x-\frac{w_5\sqrt{N}}{w_s}}+\frac{1}{x-A}\right) \qquad \text{(E.4)}$$

Since $g_s^{2n}(1) = 1$ and $A = 1 + \left(1+\sum_{i=1;i\neq s}^{4}\frac{1}{1-\frac{w_i}{w_s}}\right)$, then

$$(g_s^{2n})'(1) \;=\; \lim_{N\to\infty} 2n\cdot\left(1+\sum_{i=1;i\neq s}\frac{1}{1-\frac{w_i}{w_s}}+\frac{5}{1-\frac{w_5\sqrt{N}}{w_s}}+\frac{1}{1-A}\right) \;=\; \lim_{N\to\infty}\frac{10n}{1-\frac{w_5\sqrt{N}}{w_s}} \;=\; 0. \qquad \text{(E.5)}$$

Thus, g_s^{2n} is the desired weight function for the intermediary blip at $w_s N^{3/2}$.

Appendix F. Lower even moments of {GOE, k-BCE} and {k-BCE, k-BCE}

We provide here explicit expression of lower even moments of {GOE, k-BCE} and {k-BCE, k-BCE} based on genus expansion formulae, where distributions are rescaled and normalized to have mean zero and variance one. This means that the second method of both distributions are 1. Theoretically, with enough computing power, we should be able to obtain closed form expressions for any even moments for the two distribution. However, in reality, the computation is extremely complicated and time-consuming. Hence, we only provide the moments of $\{\text{GOE}, k\text{-BCE}\}$ up to the 10^{th} moment and $\{k\text{-BCE}, k\text{-BCE}\}$ up to the 8^{th} moment.

TABLE 2. Lower Even Moments of $\{\text{GOE}, k\text{-BCE}\}$

Moment	Value
4^{th}	$\frac{5}{2} + \frac{1}{2k^2}$
6^{th}	$\frac{33}{4} + \frac{19}{4k^2}$
8^{th}	$\frac{249}{8} + \frac{34}{k^2} + \frac{27}{8k^4}$
10^{th}	$\frac{2033}{16} + \frac{875}{4k^2} + \frac{1043}{16k^4}$

TABLE 3. Lower Even Moments of $\{k\text{-BCE}, k\text{-BCE}\}$

Moment	Value
4^{th}	$\frac{10k^4 + 86k^2 + 48}{4k^4 + 8k^2 + 4}$
6^{th}	$\frac{66k^6 + 1890k^4 + 9084k^2 + 3360}{8k^6 + 24k^4 + 24k^2 + 8}$
8^{th}	$\frac{498k^8 + 33236k^6 + 529634k^4 + 1759064k^2 + 499968}{16k^8 + 64k^6 + 96k^4 + 64k^2 + 16}$

REFERENCES

[NR] A. Nica, and S. Roland, Commutators of free random variables. (1998): 553-592.

[BasBo1] A. Basak and A. Bose, Balanced random Toeplitz and Hankel matrices, Electronic Comm. in Prob. 15 (2010), 134–148.

[BasBo2] A. Basak and A. Bose, Limiting spectral distribution of some band matrices, Periodica Mathematica Hun- garica 63 (2011), no. 1, 113–150.

[BBDLMSWX] K. Blackwell, N. Borde, C. Devlin VI, N. Luntzlara, R. Ma, S. J. Miller, M. Wang, and W. Xu, Distribution of Eigenvalues of Random Real Symmetric Block Matrices (2019), Submitted

[BCDHMSTPY] P. Burkhardt, P. Cohen, J. DeWitt, M. Hlavacek, S.J. Miller, C. Sprunger, Y.N.T. Vu, R.V. Peski, and K. Yang, Random matrix ensembles with split limiting behavior. Random Matrices: Theory and Applications, 7.03 (2018): 1850006.

[BFMT] O. Barrett, F. W. K. Firk, S. J. Miller, and C. Turnage-Butterbaugh, From Quantum Systems to L-Functions: Pair Correlation Statistics and Beyond, in Open Problems in Mathematics (editors John Nash Jr. and Michael Th. Rassias), Springer-Verlag, 2016, pages 123–171.

[BLMST] O. Beckwith, V. Luo, S. J. Miller, K. Shen, and N. Triantafillou, Distribution of eigenvalues of weighted, structured matrix ensembles, Integers: Electronic Journal Of Combinatorial Number Theory 15 (2015), paper A21, 28 pages.

[BHS1] A. Bose, R. S. Hazra, and K. Saha, Patterned random matrices and notions of independence, Technical report R3/2010 (2010), Stat-Math Unit, Kolkata.

[BHS2] A. Bose, R. S. Hazra, and K. Saha, Patterned random matrices and method of moments, in Proceedings of the International Congress of Mathematicians Hyderabad, India, 2010, 2203–2230. (Invited article). World Scientific, Singapore and Imperial College Press, UK.

[Can] F.P. Cantelli, "Sulla probabilità come limite della frequenza", Atti Accad. Naz. Lincei 26:1 (1917) pp.39–45.

[Cou] R. Couillet and M. Debbah. Random matrix methods for wireless communications. Cambridge University Press, 2011.

[DFJKRSSW] T. Dunn , H. Fleischmann, F. Jackson, S. Khunger, S. J. Miller, L. Reifenberg, A. Shashkov, and S. Willis, Limiting Spectral Distributions of Families of Block Matrix Ensembles, The PUMP Journal of Undergraduate Research (5 (2022), 122-147)

[Fer] D. Ferge, Moment equalities for sums of random variables via integer partitions and Faà di Bruno's formula, Turkish Journal of Mathematics 38 (2014), no. 3, 558–575; doi:10.3906/mat-1301-6.

[Fo] P. J. Forrester, Log-Gases and Random Matrices, London Mathematical Society Monographs 34, Princeton University Press, Princeton, NJ 2010.

[FM] F. W. K. Firk and S. J. Miller, Nuclei, Primes and the Random Matrix Connection, Symmetry 1 (2009), 64–105.

[For] P.J. Forrester. Log-gases and random matrices. Number 34. Princeton University Press, 2010.

[GKMN] L. Goldmakher, C. Khoury, S. J. Miller and K. Ninsuwan, On the spectral distribution of large weighted random regular graphs, to appear in Random Matrices: Theory and Applications. http://arxiv. org/abs/1306.6714.

[HJ] R. Horn and C. Johnson, Matrix Analysis, Cambridge University Press, 1985

[HZ] J. Harer and D. Zagier. The Euler characteristic of the moduli space of curves. Invent. Math., 85(3):457–485, 1986.

[HM] C. Hammond and S. J. Miller "Distribution of Eigenvalues of Real Symmetric Toeplitz Matrices" Journal of Theoretical Probability 18 (2005)

[KKMSX] G. Kopp et. al, Journal of Theoretical Probability (26 (2013), no. 4, 1020–1060).

[McK] B. McKay, The expected eigenvalue distribution of a large regular graph, Linear Algebra Appl. 40 (1981), 203–216.

[Me] M. Meckes, The spectra of random abelian G-circulant matrices, ALEA Lat. Am. J. Probab. Math. Stat. 9 (2012) no. 2, 435–450.

[Meh] M.L. Mehta. Random Matrices, volume 142. Academic press, 2004.

[Mez] F. Mezzadri, N.C. Snaith, Recent Perspectives in Random Matrix Theory and Number Theory. Cambridge University Press, 2005.

[MMS] A. Massey, S.J. Miller, and J. Sinsheimer. "Distribution of eigenvalues of real symmetric palindromic Toeplitz matrices and circulant matrices." Journal of Theoretical Probability 20 (2007): 637-662.

[MS] A. J. Mingo, and R. Speicher. *Free probability and random matrices*. Vol. 35. New York: Springer, 2017.

[RS] Z. Rudnick, and P. Sarnak, Zeros of principal L-functions and random matrix theory, Duke Math. J. 81 (1996), 269–322.

[Ta] L. Takacs, A Moment Convergence Theorem, The Amer. Math. Monthly 98 (Oct 1991), no. 8, 742-746

[Tao2] T. Tao, An Introduction to Measure Theory. Graduate Studies in Mathematics. Vol. 126 Providence: American Mathematical Society. p.195. ISBN 978-0-8218-6919-2.

[Tao1] T. Tao, 254a, notes 4: The semi-circular law, https://terrytao.wordpress.com/2010/02/ 02/254a-notes-4-the-semi-circular-law/, Posted:2010-02-02, Accessed:2016-08-04.

[Vas] V. Vasilchuk, "On the asymptotic distribution of the commutator and anticommutator of random matrices." Journal of Mathematical Physics 44.4 (2003): 1882-1908.

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

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

Email address: glenn.bruda@ufl.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA, GAINESVILLE, FL 32611

Email address: bf8@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

Email address: raul.marquez02@utrgv.edu

SCHOOL OF MATHEMATICAL AND STATISTICAL SCIENCES, UNIVERSITY OF TEXAS RIO GRANDE VALLEY, BROWNSVILLE, TX 78520

Email address: sjm1@williams.edu, Steven.Miller.MC.96@aya.yale.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

Email address: bp492@cam.ac.uk

DEPARTMENT OF PURE MATHEMATICS AND MATHEMATICAL STATISTICS, UNIVERSITY OF CAMBRIDGE, CAMBRIDGE, UK

Email address: vismay.sharan@yale.edu

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CT 06511

Email address: ds15@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

Email address: sw21@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

Email address: jjw3@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267