On the Limiting Shape of Young Diagrams Associated With Markov Random Words

Christian Houdré * Trevis J. Litherland †

September 25, 2012

Abstract

Let \((X_n)_{n \geq 0}\) be an irreducible, aperiodic, homogeneous Markov chain, with state space an ordered finite alphabet of size \(m\). Using combinatorial constructions and weak invariance principles, we obtain the limiting shape of the associated RSK Young diagrams as a multidimensional Brownian functional. Since the length of the top row of the Young diagrams is also the length of the longest (weakly) increasing subsequence of \((X_k)_{1 \leq k \leq n}\), the corresponding limiting law follows. We relate our results to a conjecture of Kuperberg by showing that, under a cyclic condition, a spectral characterization of the Markov transition matrix delineates precisely when the limiting shape is the spectrum of the \(m \times m\) traceless GUE. For \(m = 3\), all cyclic Markov chains have such a limiting shape, a fact previously only known for \(m = 2\). However, this is no longer true for \(m \geq 4\). In arbitrary dimension, we also study reversible Markov chains and obtain a characterization of symmetric Markov chains for which the limiting shape is the spectrum of the traceless GUE. To finish, we explore, in this general setting, connections between various limiting laws and spectra of Gaussian random matrices, focusing in particular on the relationship between the terminal points of the Brownian motions, the diagonals of the random matrix, and the scaling of the off-diagonals, a scaling we conjecture to be a function of the spectrum of the covariance matrix governing the Brownian motion.

AMS 2000 Subject Classification: 60C05, 60F05, 60F17, 60G15, 60G17, 05A16

Keywords: Longest increasing subsequence, Brownian functional, Functional Central Limit Theorem, Tracy-Widom distribution, Markov chains, Young diagrams, Random Matrices, GUE.

1 Introduction

The limiting distribution of the length of the longest increasing subsequence of a random permutation, was first determined by Baik, Deift, and Johansson [1], who identified it as the Tracy-Widom distribution, which describes the limiting behavior of the largest eigenvalue of the Gaussian unitary ensemble (GUE), as the size of the GUE goes to infinity.

The identification of the limiting distribution of \(LI_n\), the length of the longest (weakly) increasing subsequence of a random word of length \(n\), whose letters are iid and chosen uniformly from an ordered,
an $m$-letter alphabet, was first made by Tracy and Widom [28]. They showed that the limiting distribution of $LI_n$, properly centered and normalized, is that of the largest eigenvalue of the traceless $m \times m$ GUE. In the non-uniform iid case, Its, Tracy, and Widom [20, 21] described the corresponding limiting distribution as that of the largest eigenvalue of one of the diagonal blocks (corresponding to the highest probability) in a direct sum of certain independent GUE matrices. The number and respective dimensions of these matrices are determined by the multiplicities of the probabilities of choosing the letters, and the direct sum is subject again to an overall zero-trace type of condition.

The well-known Robinson-Schensted-Knuth (RSK) correspondence between sequences and pairs of Young diagrams led Tracy and Widom [28] to conjecture that the (necessarily $m$-row and of the same shape) Young diagrams of a random word generated by an $m$-letter, uniform iid sequence have a limiting shape given by the joint distribution of the eigenvalues of a $m \times m$ traceless element of the GUE. Since the length of the longest row of the Young diagrams is precisely $LI_n$, this appears to be a natural generalization. Johansson [22] proved this conjecture using orthogonal polynomial methods. Further, Okounkov [27], and Borodin, Okounkov, and Olshankii [6], as well as Johansson [22], also answered a conjecture of Baik, Deift, and Johansson [2, 3] regarding the limiting shape of the Young diagrams associated to a random permutation of $\{1, 2, \ldots, n\}$. In particular, as $n$ grows without bound, the lengths $\lambda_1, \lambda_2, \ldots, \lambda_k$ of the first $k$ rows of the Young diagrams, appropriately centered and scaled, have the same limiting law as the $k$ largest eigenvalues of a $n \times n$ element of the GUE, a result first proved, for $k = 2$, in [2, 3].

Extensions to the non-uniform iid case were addressed in Its, Tracy, and Widom [20, 21], who focused primarily on the top row of the Young diagrams. Here the obvious conjecture is that the limiting shape has rows whose suitably centered and normalized lengths have a joint distribution which is that of the whole spectrum of the direct sum of GUE matrices described above. Below, we prove this result as a special case of the Markovian framework.

Kuperberg [24] conjectured that if the word is generated by an irreducible, doubly-stochastic, cyclic Markov chain, then the limiting distribution of the shape is still that of the joint distribution of the eigenvalues of a traceless $m \times m$ element of the GUE. For $m = 2$, this was shown to be true by Chistyakov and Götze [9] (see also [18]), who, in view of further simulations, expressed doubts concerning the validity for $m \geq 4$. For $m = 3$, we will show that the conjecture holds as well. However, for $m \geq 4$, this is no longer the case. Indeed, some, but not all, cyclic Markov chains lead to a limiting law as the $k$ largest eigenvalues of a $n \times n$ element of the GUE, a result first proved, for $k = 2$, in [2, 3].

The precise class of homogeneous Markov chains with which Kuperberg’s conjecture is concerned is more specific than the ones we shall study. The irreducibility of the chain is a basic property we certainly must demand: each letter has to occur at some point following the occurrence of any given letter. Moreover, the doubly-stochastic hypothesis ensures that we have a uniform stationary distribution. However, the cyclic (also called circulant) criterion, i.e., the Markov transition matrix $P$ has entries satisfying $p_{i,j} = p_{i+1,j+1}$, for $1 \leq i, j \leq m$ (where $m + 1 = 1$), is more restrictive: cyclicity implies but is not equivalent to $P$ being doubly stochastic. Starting from a free-probability perspective, Kuperberg was led to introduce this cyclicity restriction via simulations [24] which appear to show that at least some irreducible, doubly-stochastic, non-cyclic Markov chains do not produce such limiting behavior.

The paper is organized in the following manner. In Section 2, we present a combinatorial formulation of the $LI_n$ problem and so obtain a functional of combinatorial quantities which describes the shape of the entire Young diagrams with $n$ cells, along with a concise expression for the associated asymptotic covariance structure. In Section 3, we apply Markovian Invariance Principles to express
the limiting shape of the Young diagrams as a Brownian functional for all irreducible, aperiodic, homogeneous Markov chains (without the cyclic or even the doubly-stochastic constraint.) Using this functional we are then able to answer Kuperberg’s conjecture. In Section 4, we investigate, in further detail, various symmetries exhibited by the Brownian functional and in particular obtain, for $m$ arbitrary, a necessary and sufficient condition for a cyclic Markov chain to have the same limiting law as in the iid uniform case. Still for $m$ arbitrary, a related characterization is also obtained for reversible Markov chains with symmetric transition matrices, i.e. in the doubly stochastic case. In Section 5, we further explore connections between the various Brownian functionals obtained as limiting laws and eigenvalues of random matrices. In particular, we relate the behavior of the terminal points of the Brownian motions to the diagonals of certain Gaussian random matrices, conjecturing that the off-diagonals scale according to some yet to be determined function of the spectrum of the covariance matrix governing the Brownian motions. Finally, in Section 6, we conclude with a brief discussion of natural extensions and complements to some of the ideas and results presented in the paper.

2 Combinatorics

The combinatorial development for the $m$-letter alphabet, as obtained in [17], resulted in $m-1$ partial sums $S^r_n$, $1 \leq r \leq m-1$. Using an even more straightforward development which involves $m$ quantities instead, we can obtain more symmetric expressions for $LI_n$. This is done next, and will prove useful when studying the shape of the whole RSK Young diagrams.

Let $a^r_k$ be the number of occurrences of $\alpha_r$ among $(X_i)_{1 \leq i \leq k}$. Each increasing subsequence of $(X_i)_{1 \leq i \leq k}$ consists simply of consecutive identical values, with these values forming an increasing subsequence of $\alpha_r$. Moreover, the number of occurrences of $\alpha_r \in \{\alpha_1, \ldots, \alpha_m\}$ among $(X_i)_{k+1 \leq i \leq \ell}$, where $1 \leq k < \ell \leq n$, is simply $a^r_\ell - a^r_k$. The length of the longest increasing subsequence of $X_1, X_2, \ldots, X_n$ is thus given by

$$LI_n = \max_{0 \leq k_1 \leq \cdots \leq k_{m-1} \leq n} [(a^1_{k_1} - a^1_0) + (a^2_{k_2} - a^2_{k_1}) + \cdots + (a^m_{k_{m-1}} - a^m_{k_{m-2}})],$$

(2.1)

$$LI_n = \max_{0 \leq k_1 \leq \cdots \leq k_{m-1} \leq n} [(a^1_{k_1} - a^1_0) + (a^2_{k_2} - a^2_{k_1}) + \cdots + (a^m_{k_{m-1}} - a^m_{k_{m-2}}) + a^m_n],$$

(2.2)

where $a^r_0 = 0$.

Moving beyond the purely combinatorial setting, assume that $(X_k)_{k \geq 0}$ is an infinite sequence generated by an irreducible homogeneous Markov chain having a stationary distribution $(\pi_1, \pi_2, \ldots, \pi_m)$. (For no $k \geq 0$ is the law of $X_k$ necessarily assumed to be the stationary distribution.) For each $1 \leq r \leq m$, set $T^r_k = a^r_k - \pi_r k$, for $k \geq 1$, and $T^r_0 = 0$. Beginning again with (2.1), we find that
For a uniform stationary distribution, \( \pi \) setting
\[
LI_n = \max_{0 \leq k_1 \leq \ldots \leq k_{m-1} \leq n} \left[ (a_{k_1}^1 - a_0^1) + (a_{k_2}^2 - a_{k_1}^2) + \cdots + (a_{k_m}^m - a_{k_{m-1}}^m) \right]
\]
\[
= \max_{0 \leq k_1 \leq \ldots \leq k_{m-1} \leq n} \left[ (T_{k_1}^1 + \pi_1 k_1) - (T_{k_0}^1 + \pi_1 k_0) + ((T_{k_2}^2 + \pi_2 k_2) - (T_{k_1}^2 + \pi_2 k_1)) + \cdots + \left( (T_{k_{m-1}}^m + \pi_m k_{m-1}) - (T_{k_{m-1}}^m + \pi_m k_{m-1}) \right) \right]
\]
\[
= \max_{0 \leq k_1 \leq \ldots \leq k_{m-1} \leq n} \left[ \left( T_{k_1}^1 - T_{k_0}^1 \right) + \left( T_{k_2}^2 - T_{k_1}^2 \right) + \cdots + \left( T_{k_{m-1}}^m - T_{k_{m-1}}^m \right) \right]
+ \pi_1(k_1 - k_0) + \pi_2(k_2 - k_1) + \cdots + \pi_m(k_{m-1} - k_{m-1}) \right].
\] (2.3)

Setting \( \pi_{max} = \max\{\pi_1, \pi_2, \ldots, \pi_m\} \), (2.3) becomes
\[
LI_n - \pi_{max} n = \max_{0=0 \leq k_0 \leq k_1 \leq \ldots \leq k_{m-1} \leq k_m=n} \sum_{r=1}^{m} \left[ (T_{kr}^r - T_{kr-1}^r) + (\pi_r - \pi_{max})(k_r - k_{r-1}) \right].
\] (2.4)

For a uniform stationary distribution, \( \pi_{max} = \pi_r = 1/m \), for all \( r \), and (2.4) simplifies to
\[
LI_n - \frac{n}{m} = \max_{0=0 \leq k_0 \leq k_1 \leq \ldots \leq k_{m-1} \leq k_m=n} \sum_{r=1}^{m} (T_{kr}^r - T_{kr-1}^r).
\] (2.5)

To introduce a random walk formalism into the picture, we next set, for \( i = 1, \ldots, n \) and \( r = 1, 2, \ldots, m \),
\[
W_i^r = \begin{cases} 
1, & \text{if } X_i = \alpha_r, \\
0, & \text{otherwise}. 
\end{cases}
\] (2.6)

Clearly, \( a_k^r = \sum_{i=1}^{k} W_i^r \), and so \( T_k^r = \sum_{i=1}^{k} (W_i^r - \pi_r) \), for \( 1 \leq r \leq m \).

To understand the limiting law of (2.4) or (2.5), we must have a more precise description of the underlying Markovian structure. To that end, let \( p_{r,s} = \mathbb{P}(X_{k+1} = \alpha_s|X_k = \alpha_r) \), and let \( P = (p_{r,s}) \) be the associated Markov transition matrix. In this setting,
\[
(p_{1}^{n+1}, p_{2}^{n+1}, \ldots, p_{m}^{n+1}) = (p_{1}^{n}, p_{2}^{n}, \ldots, p_{m}^{n}) P.
\]

Moreover, as usual, let \( p_{r,s}^{(k)} \) denote the \( k \)-step transition probability from \( \alpha_r \) to \( \alpha_s \); its associated transition matrix is simply \( P^k \).

Assume now that the law of \( X_0 \) is the stationary distribution. Thus, by construction, \( \mathbb{E}T_k^r = 0 \) for all \( 1 \leq r \leq m \) and \( 1 \leq k \leq n \), and our primary task is to describe the covariance structure of these random variables \( T_k^r \). First, for \( k \geq 1 \),
\[
\text{Var}T_k^r = \sum_{i=1}^{k} \text{Var}W_i^r + \sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^{k} \text{Cov}(W_{i_1}^r, W_{i_2}^r) + \sum_{i_1=2}^{k-1} \sum_{i_2=1}^{i_1-1} \text{Cov}(W_{i_1}^r, W_{i_2}^r).
\] (2.7)
Next, by stationarity, and since $W_i^T$ is, simply, a Bernoulli random variable with parameter $\pi_r$, (2.11) becomes

$$\text{Var}T_k^T = \sum_{i=1}^k \text{Var}W_i^T + \sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^k \text{Cov}(W_0^T, W_{i_2-i_1}^T) + \sum_{i_1=1}^k \sum_{i_2=1}^{k-1} \text{Cov}(W_0^T, W_{i_1-i_2}^T)$$

$$= k\pi_r(1 - \pi_r) + \sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^k \left(\pi_r\pi_r^{i_2-i_1} - \pi_r^2\right) + \sum_{i_1=1}^k \sum_{i_2=1}^{k-1} \left(\pi_r\pi_r^{i_1-i_2} - \pi_r^2\right)$$

$$= k\pi_r - k^2\pi_r^2 + \pi_r \sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^k e_r P^{i_2-i_1} e_r^T + \pi_r \sum_{i_1=1}^k \sum_{i_2=1}^{k-1} e_r P^{i_1-i_2} e_r^T, \quad (2.8)$$

where $e_r = (0, 0, \ldots, 0, 1, 0, \ldots)$ is the $r^{th}$ standard basis vector of $\mathbb{R}^m$. Setting

$$Q_k = \sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^k P^{i_2-i_1} = \sum_{i=1}^k (k - i)P^i, \quad (2.9)$$

we can rewrite (2.8) in the simple form

$$\text{Var}T_k^T = k\pi_r - k^2\pi_r^2 + 2\pi_r e_r Q_k e_r^T. \quad (2.10)$$

Our description of the covariance structure can now be completed using the above results. For $r_1 \neq r_2$ and $k \geq 1$,

$$\text{Cov}(T_k^{r_1}, T_k^{r_2}) = \sum_{i=1}^k \text{Cov}(W_i^{r_1}, W_i^{r_2}) + \sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^k \text{Cov}(W_{i_1}^{r_1}, W_{i_2}^{r_2}) + \sum_{i_1=1}^k \sum_{i_2=1}^{k-1} \text{Cov}(W_{0}^{r_1}, W_{i_2-i_1}^{r_2})$$

$$= \sum_{i=1}^k \text{Cov}(W_i^{r_1}, W_i^{r_2}) + \sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^k \text{Cov}(W_0^{r_1}, W_{i_2-i_1}^{r_2}) + \sum_{i_1=1}^k \sum_{i_2=1}^{k-1} \text{Cov}(W_0^{r_2}, W_{i_1-i_2}^{r_1})$$

$$= -k\pi_{r_1}\pi_{r_2} + \sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^k \left(\pi_{r_1}\pi_{r_2}^{i_2-i_1} - \pi_{r_1}\pi_{r_2}\right) + \sum_{i_1=1}^k \sum_{i_2=1}^{k-1} \left(\pi_{r_2}\pi_{r_2}^{i_1-i_2} - \pi_{r_1}\pi_{r_2}\right)$$

$$= -k^2\pi_{r_1}\pi_{r_2} + \pi_{r_1} \sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^k e_{r_1} P^{i_2-i_1} e_{r_2}^T + \pi_{r_2} \sum_{i_1=1}^k \sum_{i_2=1}^{k-1} e_{r_2} P^{i_1-i_2} e_{r_1}^T$$

$$= -k^2\pi_{r_1}\pi_{r_2} + \pi_{r_1} e_{r_1} P_{k} e_{r_2}^T + \pi_{r_2} e_{r_2} P_{k} e_{r_1}^T. \quad (2.11)$$

**Remark 2.1** Both (2.10) and (2.11) appear to be asymptotically quadratic in $k$. However, since $Q_k = \sum_{i=1}^k (k - i)P^i$, cancellations will show that when the Markov chain is irreducible and aperiodic, the order of the variance is, in fact, linear in $k$. 


In order to further analyze the asymptotics of $Q_k$, we first examine the Jordan decomposition of $P$ for a very general class of transition matrices.

**Proposition 2.1** Let $P$ be the $m \times m$ transition matrix of an irreducible, aperiodic, homogeneous Markov chain. Let $P = S^{-1} \Lambda S$ be the Jordan decomposition of $P$, where the rows of $S$ consist of the generalized left-eigenvectors of $P$ with, moreover, the first row of $S$ being the stationary distribution $(\pi_1, \pi_2, \ldots, \pi_m)$, and where $\Lambda = \text{diag}(1, \bar{\Lambda})$, with $\bar{\Lambda}$ being the $(m-1) \times (m-1)$ (lower) Jordan-block matrix. Then the first column of $S^{-1}$ is $(1, 1, \ldots, 1)^T$.

**Proof.** Since $P = S^{-1} \Lambda S$, then $PS^{-1} = S^{-1} \Lambda$. Denoting the first column of $S^{-1}$ by $c_1$, we have $PC_1 = c_1$. But since the rows of $P$ sum to 1, we see that $c_1 = (1, 1, \ldots, 1)^T$ satisfies $PC_1 = c_1$. Moreover, $c_1$ must be unique, up to normalization, since the irreducibility of $P$ implies that the eigenvalue $\lambda_1 = 1$ has multiplicity 1. Finally, since the inner product of the first row of $S$ and the first column of $S^{-1}$ is 1, the correct normalization is indeed $(1, 1, \ldots, 1)^T$.

Returning to $Q_k$, as given in (2.9), and using Proposition 2.1 we then obtain:

**Theorem 2.1** Let $(X_n)_{n \geq 0}$ be a sequence generated by an $m$-letter, aperiodic, irreducible, homogeneous Markov chain with state space $\mathcal{A}_m = \{a_1 < \cdots < a_m\}$, transition matrix $P$, and stationary distribution $(\pi_1, \pi_2, \ldots, \pi_m)$. Let also the law of $X_0$ be the stationary distribution. Moreover, for $1 \leq r \leq m$, let $T_k^r = a_k^r - \pi_k$, for $k \geq 1$, and $T_0^r = 0$, where $a_k^r$ is the number of occurrences of $a_r$ among $(X_i)_{1 \leq i \leq k}$. Then, for $1 \leq r \leq m$,

$$
\lim_{k \to \infty} \frac{\text{Var} T_k^r}{k} = \pi_r \left(1 + 2e_r S^{-1} D S e_r^T \right),
$$

(2.12)

and for $r_1 \neq r_2$,

$$
\lim_{k \to \infty} \frac{\text{Cov}(T_k^{r_1}, T_k^{r_2})}{k} = \pi_{r_1} e_{r_1} S^{-1} D S e_{r_2}^T + \pi_{r_2} e_{r_2} S^{-1} D S e_{r_1}^T,
$$

(2.13)

where $P = S^{-1} \Lambda S = S^{-1} \text{diag}(1, \bar{\Lambda}) S$ is the Jordan decomposition of $P$ in Proposition 2.1 and $D = \text{diag}(-1/2, \bar{\Lambda} (I - \bar{\Lambda})^{-1})$. That is, the asymptotic covariance matrix of $(T_k^1, T_k^2, \ldots, T_k^m)$ is given by

$$
\Sigma = \Pi + \Pi (S^{-1} DS + (S^{-1} DS)^T) \Pi,
$$

(2.14)

where $\Pi = \text{diag}(\pi_1, \pi_2, \ldots, \pi_m)$.

**Proof.** Beginning with (2.9), we employ the Jordan decomposition of $P$ in Proposition 2.1 to find that

$$
Q_k = \sum_{i=1}^{k-1} (k-i) (S^{-1} \Lambda S)^i = S^{-1} \left[ \sum_{i=1}^{k-1} (k-i) \Lambda^i \right] S = S^{-1} \text{diag} \left( h(1), \sum_{i=1}^{k-1} (k-i) \bar{\Lambda}^i \right) S,
$$

(2.15)

where $h(\lambda) := \sum_{k=1}^{n-1} (n-k) \lambda^k$. Now $h(1) = (k-1)/2$ is quadratic in $k$, while for $\lambda \neq 1$,

$$
h(\lambda) = k \frac{\lambda}{1-\lambda} + \frac{\lambda(\lambda^k - 1)}{(1-\lambda)^2},
$$

6
so that $h(\lambda)$ is linear in $k$. Now, by the aperiodicity of $P$, the diagonals of the lower-triangular matrix $\tilde{\Lambda}$ are all less than unity in modulus. We can thus write the matrix sum in (2.15) as

$$h(\tilde{\Lambda}) := k\tilde{\Lambda}(I - \tilde{\Lambda})^{-1} + o(k).$$

The matrix $Q_k$ can now be expressed as the sum of terms which are, respectively, quadratic and linear in $k$. Recalling, moreover, that the first row of $S$ contains the stationary distribution, and that the first column of $S^{-1}$ is $(1, 1, \ldots, 1)^T$, we have

$$Q_k = S^{-1}\text{diag}(h(1), h(\tilde{\Lambda})),
= \frac{k^2}{2} S^{-1}\text{diag}(1, 0, \ldots, 0)S + k S^{-1}\text{diag} \left(-\frac{1}{2}, \tilde{\Lambda}(I - \tilde{\Lambda})^{-1}\right) S + o(k)
= \frac{k^2}{2} \begin{pmatrix}
\pi_1 & \pi_2 & \cdots & \pi_m \\
\pi_1 & \pi_2 & \cdots & \pi_m \\
\vdots & \vdots & \ddots & \vdots \\
\pi_1 & \pi_2 & \cdots & \pi_m 
\end{pmatrix} + k S^{-1}DS + o(k).$$

Starting with the variance in (2.10), we now find that, for each $1 \leq r \leq m$,

$$\text{Var} \ T_r^{T_k} = k\pi_r - k^2\pi_r^2 + 2\pi_r e_r Q_k e_r^T
= k\pi_r - k^2\pi_r^2 + 2\pi_r \left(\frac{k^2}{2}\pi_r + ke_r S^{-1}DSe_r^T\right) + o(k)
= k\pi_r \left(1 + 2e_r S^{-1}DSe_r^T\right) + o(k),$$

from which the asymptotic result (2.12) follows immediately.

An identical development shows that, for $r_1 \neq r_2$, (2.11) simplifies to

$$\text{Cov}(T_{r_1}^{T_k}, T_{r_2}^{T_k}) = -k^2\pi_{r_1}\pi_{r_2} + \pi_{r_1} e_{r_1} Q_k e_{r_2}^T + \pi_{r_2} e_{r_2} Q_k e_{r_1}^T
= -k^2\pi_{r_1}\pi_{r_2} + \pi_{r_1} \left(\frac{k^2}{2}\pi_{r_2} + ke_{r_2} S^{-1}DSe_{r_1}^T\right)
+ \pi_{r_2} \left(\frac{k^2}{2}\pi_{r_1} + ke_{r_1} S^{-1}DSe_{r_2}^T\right) + o(k)
= k \left(\pi_{r_1} e_{r_1} S^{-1}DSe_{r_2}^T + \pi_{r_2} e_{r_2} S^{-1}DSe_{r_1}^T\right) + o(k),$$

from which the asymptotic result (2.13) follows, and so does (2.14).

**Remark 2.2** To see that (2.12) and (2.13) both recover the covariance results for the iid case investigated by the authors in [17], let $P$ be the transition matrix whose rows each consist of the stationary distribution $(\pi_1, \pi_2, \ldots, \pi_m)$. In this case $\lambda_2 = \cdots = \lambda_m = 0$, and so $D = \text{diag}(-1/2, 0, \ldots, 0)$. Hence,
\[ e_{r_1} S^{-1} D S e_{r_2}^T = (1, \ldots, *) D (\pi_{r_2}, \ldots, \pi_1) = -\frac{\pi_{r_2}}{2}, \]

for all \( r_1 \) and \( r_2 \), and so, for each \( r \),

\[
\lim_{k \to \infty} \frac{\text{Var} T^r_k}{k} = \pi_r \left( 1 + 2 \left( -\frac{\pi_r}{2} \right) \right) = \pi_r (1 - \pi_r),
\]

while, for \( r_1 \neq r_2 \),

\[
\lim_{k \to \infty} \frac{\text{Cov}(T^r_k, T^{r_2}_k)}{k} = \pi_{r_1} \left( -\frac{\pi_{r_2}}{2} \right) + \pi_{r_2} \left( -\frac{\pi_{r_1}}{2} \right) = -\pi_{r_1} \pi_{r_2}.
\]

In the uniform iid case, \( \pi_r = 1/m \), for all \( 1 \leq r \leq m \). Hence, for \( r_1 \neq r_2 \), the asymptotic correlation between \( T^r_k \) and \( T^{r_2}_k \) is given by \((-1/(m^2))/((1/m)(1-1/m)) = -1/(m-1)\), so that the covariance matrix is indeed the permutation-symmetric one obtained in [17]. There is, moreover, another Brownian functional representation for the iid uniform case in [17] in which the Brownian motions have a tridiagonal covariance matrix.

### 3 The Limiting Shape of the Young diagrams

Thus far, our results have centered on \( LI_n \) alone, essentially ignoring the larger question of the structure of the entire Young diagrams. The present section extends the combinatorial development of the previous section to answer the question of the limiting shape of the Young diagrams. Our first result in this direction is a purely combinatorial expression generalizing (2.1).

**Theorem 3.1** Let \( R^1_n, R^2_n, \ldots, R^n_n \) be the lengths of the first \( 1 \leq r \leq m \) rows of the RSK Young diagrams generated by the sequence \((X_k)_{1 \leq k \leq n}\) whose elements belong to an ordered alphabet \( A_m = \{\alpha_1 < \cdots < \alpha_m\} \). Then, for each \( 1 \leq r \leq m \), the sum of the lengths of the first \( r \) rows of the Young diagrams is given by

\[
\sum_{j=1}^{r} R^j_n = \max_{k, r \in J_r} \sum_{j=1}^{r} \sum_{\ell=j-1}^{m-r+j} (a^\ell_{k_j, \ell} - a^\ell_{k_{j+1}, \ell+1}), \quad (3.1)
\]

where \( J_r = \{ (k_j, \ell, 1 \leq j \leq r, 0 \leq \ell \leq m) : k_{j_1, \ell} = 0, 1 \leq j \leq r, 0 \leq \ell \leq j - 1; k_{j, \ell} = n, 1 \leq j \leq r, m - r + j \leq \ell \leq m; k_{j_1, \ell_1} \leq k_{j_1, \ell_1} \leq k_{j_1, \ell_1} \leq j \leq r, 1 \leq \ell \leq m; k_{j_1, \ell_1+1} \leq k_{j_1, \ell_1} \leq \ell \leq m - 1 \}, \)

and where \( a^\ell_k \) is the number of occurrences of \( \alpha_\ell \) among \((X_i)_{1 \leq i \leq k}\).

**Proof.** Recall that the sum of the lengths of the first \( r \) rows of the RSK Young diagrams generated by a sequence \((X_k)_{1 \leq k \leq n}\), whose letters arise from an \( m \)-letter alphabet, has an interpretation in terms of the length of certain increasing sequences. Indeed, the sum \( R^1_n + R^2_n + \cdots + R^n_n \) is equal to the maximum sum of the lengths of \( r \) disjoint, increasing subsequences of \((X_k)_{1 \leq k \leq n}\), where by \textit{disjoint} it is meant that each element of \((X_k)_{1 \leq k \leq n}\) occurs in at most one of the \( r \) subsequences. (See Lemma 1 of Section 3.2 in [12].) More general results of this sort, involving partial orderings of the alphabet and associated antichains, are known as Greene’s Theorem [16]. However, such results are not enough for our purpose. Below we need a different way of reconstructing disjoint subsequences.
We begin by examining an arbitrary collection of \( r \) disjoint, increasing subsequences of \( (X_k)_{1 \leq k \leq n} \), and show that we can always map these \( r \) subsequences onto another collection of \( r \) disjoint, increasing subsequences whose properties will be amenable to our combinatorial analysis.

Specifically, with the number of rows \( r \) fixed, suppose that, for each \( 1 \leq j \leq r \), we have an increasing subsequence \( (X^j_{k_j})_{1 \leq \ell \leq n_j} \) of length \( n_j \leq n \), and that the \( r \) subsequences are disjoint.

We first construct the new subsequence \( (\tilde{X}^1_{k_1^r})_{1 \leq \ell \leq \tilde{n}_1} \) as follows. First, place all \( \alpha_1 \)s occurring among the \( r \) original subsequences into \( (\tilde{X}^1_{k_1^r})_{1 \leq \ell \leq \tilde{n}_1} \), if there are any. If the last \( \alpha_1 \) occurs at the \( n^\text{th} \) index, then \( (\tilde{X}^1_{k_1^r})_{1 \leq \ell \leq \tilde{n}_1} \) is complete. Otherwise, place all \( \alpha_2 \)s which occur after the final \( \alpha_1 \) into \( (\tilde{X}^1_{k_1^r})_{1 \leq \ell \leq \tilde{n}_1} \), if there are any. If the last \( \alpha_2 \) occurs at the \( n^\text{th} \) index, then \( (\tilde{X}^1_{k_1^r})_{1 \leq \ell \leq \tilde{n}_1} \) is complete. Otherwise, continue adding, successively, \( \alpha_3, \ldots, \alpha_{m-r+1} \) in the same manner. Thus, \( (\tilde{X}^1_{k_1^r})_{1 \leq \ell \leq \tilde{n}_1} \) consists of a weakly increasing sequence of length \( \tilde{n}_1 \) having values in \( \{\alpha_1, \ldots, \alpha_{m-r+1}\} \).

Next, we construct the new subsequence \( (\tilde{X}^2_{k_2^r})_{1 \leq \ell \leq \tilde{n}_2} \) similarly. By considering only those letters among the \( r \) original subsequences which have not already been moved to the first new subsequence, start with the smallest available letter, \( \alpha_2 \), and continue adding, successively, \( \alpha_3, \ldots, \alpha_{m-r+2} \). Note that, crucially, all \( \alpha_2 \)s added to \( (\tilde{X}^2_{k_2^r})_{1 \leq \ell \leq \tilde{n}_2} \) occur before the last index at which \( \alpha_1 \) was added to the first subsequence. More generally, each \( \alpha_j, 2 \leq j \leq m - r + 2 \), added to \( (\tilde{X}^2_{k_2^r})_{1 \leq \ell \leq \tilde{n}_2} \) occurs before the last \( \alpha_{j-1} \) was added to the first subsequence. Thus, \( (\tilde{X}^2_{k_2^r})_{1 \leq \ell \leq \tilde{n}_2} \) consists of a weakly increasing subsequence of length \( \tilde{n}_2 \) having values in \( \{\alpha_2, \ldots, \alpha_{m-r+2}\} \).

The construction of \( (\tilde{X}^j_{k_j^r})_{1 \leq \ell \leq \tilde{n}_j} \), for \( 3 \leq j \leq r \), continues in the same manner, with \( (\tilde{X}^j_{k_j^r})_{1 \leq \ell \leq \tilde{n}_j} \), constructed from among the entries of the \( r \) original subsequences which were not moved into any of the first \( j-1 \) new subsequences, so that \( (\tilde{X}^j_{k_j^r})_{1 \leq \ell \leq \tilde{n}_j} \), consists of a weakly increasing sequence of length \( \tilde{n}_j \) having values in \( \{\alpha_j, \ldots, \alpha_{m-r+j}\} \). It is possible that beyond some \( j \geq 2 \) the new subsequences may be empty.

We claim that, indeed, the construction of the \( r^\text{th} \) new subsequence exhausts the set of available entries. Indeed, without loss of generality, assume that after we have created the \( (r-1)^\text{th} \) new subsequence, the set of available entries is non-empty, and designate the location of the final \( \alpha_\ell \) to be included in the \( j^\text{th} \) new subsequence by \( k_{j,\ell} \), for \( 1 \leq j \leq r \) and \( 1 \leq \ell \leq m \). (If no \( \alpha_\ell \) was available for inclusion, set \( k_{j,\ell} = k_{j,\ell-1} \), where \( k_{j,0} = 0 \), for all \( 1 \leq j \leq r \).) Clearly, all \( \alpha_1, \alpha_2, \ldots, \alpha_{r-1} \) have been included in the first \( r-1 \) new subsequences. If \( r = m \), we are done: simply put the remaining \( \alpha_\ell s \) into the \( r^\text{th} \) new subsequence. If \( r < m \), we may still ask whether there was, for some \( r+1 \leq \ell \leq m \), an \( \alpha_\ell \) from among the available entries which occurred before \( k_{r,\ell-1} \). Assume that there is such an \( \alpha_\ell \). Now by construction, \( k_{j+1,\ell-r+j} \leq k_{j,\ell-r+j-1} \), for \( 1 \leq j \leq r-1 \). Hence, there exist letters \( \alpha_{j_1} < \alpha_{j_2} < \cdots < \alpha_{j_r} < \alpha_{\ell-1} \) among the original subsequences which occurred after \( k_{r,\ell-1} \), and, moreover, each letter must come from a different subsequence. But since each original subsequence was increasing, none of them could have contained an \( \alpha_\ell \) before \( k_{r,\ell-1} \), and we have a contradiction.

To better understand this construction, consider the first row of Figure H which shows an initial sequence of length \( n = 12 \), with \( m = 4 \) letters, broken into \( r = 3 \) disjoint, increasing subsequences of lengths \( n_1 = 3, n_2 = 4, \) and \( n_3 = 3 \), and so with total length 10. The final three rows of the diagram show the results of the operations described above, producing 3 new increasing subsequences of length \( \tilde{n}_1 = 4, \tilde{n}_2 = 3, \) and \( \tilde{n}_3 = 3 \).
Figure 1: Transformation of $r = 3$ subsequences.
Hence, if we wish to find \( r \) disjoint, increasing subsequences whose length sum is maximal, it suffices to consider only those disjoint, increasing subsequences for which the final occurrence of the letter \( \alpha \) in the subsequence \( i \) happens after the final occurrence in the subsequence \( j \), whenever \( i < j \). Because such ranges do not overlap, if we wish to count the number of \( \alpha \)'s in the \( j \)th subsequence, it suffices to simply count the number of \( \alpha \)'s in \((X_k)_{1 \leq k \leq n}\) over that range.

Indeed, returning to the fundamental combinatorial objects of our development, the \( a_{jk}^\ell \), we see that since \( a_{jk}^\ell - a_{jk}^{\ell-1} \) counts the number of \( \alpha \)'s in the range \( \ell + 1, \ldots, k \), we can describe the valid index ranges over which to search for the maximal sum as

\[
J_{r,m} = \left\{(k_{j,\ell},1 \leq j \leq r,0 \leq \ell \leq m) : k_{j,\ell} = 0,1 \leq j \leq r,0 \leq \ell \leq j - 1; k_{j,\ell} = n,1 \leq j \leq r,m - r + j \leq \ell \leq m;k_{j,\ell-1} \leq k_{j,\ell},1 \leq j \leq r,1 \leq \ell \leq m;k_{j+1,\ell+1} \leq k_{j,\ell},1 \leq j \leq r - 1,0 \leq \ell \leq m - 1\right\}.
\]

The constraints on the \( k_{j,\ell} \) follow simply from the fact that each subsequence is increasing and that, moreover, the intervals associated with a given letter do not overlap. Figure 2 indicates the relative positions of each range, for \( r = 4 \) and \( m = 7 \).

Since the first possible letter of each subsequence grows from \( \alpha_1 \) to \( \alpha_r \), and the last possible letter grows from \( \alpha_{m+r-1} \) to \( \alpha_m \), the result is proved.

We are now ready to apply our asymptotic covariance results (Theorem 2.1), along with a Brownian sample-path approximation, to the combinatorial expression \((3.1)\), and so obtain a Brownian functional expression for the limiting shape of the Young diagrams for all irreducible, aperiodic, homogeneous Markov chains.

Indeed, for each \( 1 \leq r \leq m \), let the sum of the first \( r \) rows of the Young diagrams be given by

\[
V_n^r := \sum_{j=1}^{r} R_{nj}^r = \max_{k_{j,\ell} \in J_{r,m}} \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} (a_{k_{j,\ell}}^\ell - a_{k_{j,\ell-1}}^\ell), \tag{3.2}
\]

where the index set \( J_{r,m} \) is defined as in Theorem 3.1 Define as before

\[
T_k^r = \sum_{i=1}^{k} (W_i^r - \pi_r) = a_k^r - \pi_r k,
\]

and so rewrite \((3.2)\) as

Figure 2: Schematic diagram of \( J_{r,m} \), for \( r = 4, m = 7 \).
The constraints

Next, let \( \tau \) be a permutation of the indices 1, 2, \ldots, \( m \) such that \( \pi_{\tau(1)} \geq \pi_{\tau(2)} \geq \cdots \geq \pi_{\tau(m)} > 0 \). Moreover, we demand that if \( \pi_{\tau(i)} = \pi_{\tau(j)} \) for \( i < j \), then \( \tau(i) < \tau(j) \). (The permutation so defined is thus unique.) Let \( \nu_r = \sum_{j=1}^{r} \pi_{\tau(j)} \) be the sum of the \( r \) largest values among \( \pi_1, \pi_2, \ldots, \pi_m \). We obtain, below, the limiting distribution of \( (V_n^r - \nu_r) / \sqrt{n} \) as a Brownian functional.

To introduce Brownian sample-path approximations, and for each \( 1 \leq r \leq m \), we first define the asymptotic variance of \( T_n^r \) as in (3.12), by

\[
\sigma_r^2 := \lim_{n \to \infty} \frac{\text{Var} T_n^r}{n} = e_r \Sigma e_r^T,
\]

and, for \( r_1 \neq r_2 \), the asymptotic covariance of \( T_n^{r_1} \) and \( T_n^{r_2} \) by

\[
\sigma_{r_1, r_2} := \lim_{n \to \infty} \frac{\text{Cov}(T_n^{r_1}, T_n^{r_2})}{n} = e_{r_1} \Sigma e_{r_2}^T,
\]

where \( \Sigma \) is the covariance matrix of Theorem 2.1 associated with the transition matrix \( P \). For each \( 1 \leq r \leq m \), we then let

\[
\hat{B}_n^r(t) = \frac{T_n^r(nt) + (nt - [nt])(W_n^{r+1} - \pi_r)}{\sigma_r \sqrt{n}},
\]

for \( 0 \leq t \leq 1 \). This rescaling of \([0, n]\) to \([0, 1]\) calls for us to define a new parameter set over which we will maximize a functional arising from the expressions in (3.5). Indeed, for any positive integers \( s \) and \( d \), with \( s \leq d \), define the set

\[
I_{s,d} = \left\{ (t_{j,\ell}, 1 \leq j \leq s, 0 \leq \ell \leq d) : t_{j,\ell} = 0, 1 \leq j \leq s, 0 \leq \ell \leq j - 1; \right. \\
\left. t_{j,\ell} = 1, 1 \leq j \leq s, d - s + j \leq \ell \leq d; \right. \\
\left. t_{j,\ell-1} \leq t_{j,\ell}, 1 \leq j \leq s, 1 \leq \ell \leq d; \right. \\
\left. t_{j+1,\ell+1} \leq t_{j,\ell}, 1 \leq j \leq s - 1, 0 \leq \ell \leq d - 1 \right\}.
\]

Note that the constraints \( t_{j,j-1} = 0 \) and \( t_{j,d-s+j} = 1 \), for \( 1 \leq j \leq s \), force many of the \( t_{j,\ell} \) to be zero or one. We will denote the \( s \times (d + 1) \)-tuple elements of \( I_{s,d} \), by \((t_{\ldots})\). Figure 3 shows the structure of \( I_{s,d} \), for \( s = 4 \) and \( d = 7 \). The locations of \( t_{j,\ell} \) are indicated by the horizontal lines within the diagram.

With this notation, (3.3) becomes

\[
\frac{V_n^r}{\sqrt{n}} - n\nu_r = \max_{(t_{\ldots}) \in I_{s,d}} \left\{ \sum_{j=1}^{r} \sum_{\ell=j}^{m} \sigma_{\ell} \left( \hat{B}_n^\ell(t_{j,\ell}) - \hat{B}_n^\ell(t_{j,\ell-1}) \right) + \sum_{j=1}^{r} \sum_{\ell=j}^{m} \sqrt{n}(\pi_\ell - \pi_{\tau(j)}) (t_{j,\ell} - t_{j,\ell-1}) \right\}.
\]

(3.7)
Theorem 3.2 Let \((X_n)_{n \geq 0}\) be an irreducible, aperiodic, homogeneous Markov chain with finite state space \(A_m = \{\alpha_1 < \cdots < \alpha_m\}\), transition matrix \(P\), and stationary distribution \((\pi_1, \pi_2, \ldots, \pi_m)\). Let \(\Sigma = (\sigma_{r,s})_{1 \leq r, s \leq m}\) be the associated asymptotic covariance matrix, as given in (2.14), and let the law of \(X_0\) be given by the stationary distribution. Let \(\tau\) be the permutation of \(\{1, 2, \ldots, m\}\) such that \(\pi_\tau(i) \geq \pi_\tau(i+1)\), and \(\tau(i) < \tau(j)\) whenever \(\pi_\tau(i) = \pi_\tau(j)\) and \(i < j\). For each \(1 \leq r \leq m\), let \(V_r\) be the sum of the lengths of the first \(r\) rows of the associated Young diagrams, and let \(\nu_r = \sum_{j=1}^{r} \pi_{\tau(j)}\). Finally, let \(d_r\) be the multiplicity of \(\pi_{\tau(r)}\), and let

\[
m_r = \begin{cases} 0, & \text{if } \pi_{\tau(r)} = \pi_{\tau(1)}, \\ \max \{i : \pi_{\tau(i)} > \pi_{\tau(r)}\}, & \text{otherwise}. \end{cases}
\]

Then, for each \(1 \leq r \leq m\),

\[
\frac{V_r - n\nu_r}{\sqrt{n}} \Rightarrow V_\infty^r := \sum_{i=1}^{m_r} \sigma_{\tau(i)} B^{\tau(i)}(1) + \max_{I_{r-m_r,d_r}} \sum_{j=1}^{r-m_r (d_r+m_r-r+j)} \sum_{\ell=j}^{\sigma_{\tau(m_r+\ell)} (B^{\tau(m_r+\ell)}(t_{j,\ell}) - B^{\tau(m_r+\ell)}(t_{j,\ell-1}))},
\]

where the first sum on the right-hand side of (3.8) is understood to be 0, if \(m_r = 0\). Above, \(\sigma_r^2 = \sigma_{r,r}\), and \((\tilde{B}^1(t), \tilde{B}^2(t), \ldots, \tilde{B}^m(t))\) is an \(m\)-dimensional Brownian motion, with covariance matrix \(\tilde{\Sigma} = (\tilde{\sigma}_{r,s})_{1 \leq r, s \leq m}\) given by

\[
(\tilde{\sigma}_{r,s}) = t(\sigma_{r,s})/\sigma_r \sigma_s,
\]

for \(1 \leq r, s \leq m\). Moreover, for any \(1 \leq k \leq m\),
\[
\left( \frac{V_1^n - n\nu_1}{\sqrt{n}}, \frac{V_2^n - n\nu_2}{\sqrt{n}}, \ldots, \frac{V_k^n - n\nu_k}{\sqrt{n}} \right) \implies \left( V_1^\infty, V_2^\infty, \ldots, V_k^\infty \right). \tag{3.10}
\]

**Remark 3.1** The critical indices \(d_r\) and \(m_r\) in Theorem 3.2 are chosen so that
\[
\pi_{\tau(m_r)} > \pi_{\tau(m_r+1)} = \cdots = \pi_{\tau(m_r+d_r)} > \pi_{\tau(m_r+d_r+1)}.
\]
Thus, the functional in (3.8) consists of a sum of \(m_r\) Gaussian random variables and a maximal functional involving only \(d_r\) of the \(m\) one-dimensional Brownian motions.

**Remark 3.2** Another, perhaps more natural, way of describing the covariance structure of the \(m\)-dimensional Brownian motion in Theorem 3.2 is to note that \((\sigma_1 \tilde{B}^1(t), \sigma_2 \tilde{B}^2(t), \ldots, \sigma_m \tilde{B}^m(t))\) has covariance matrix \(t\Sigma\).

Let us now examine the case \(r = 1\). Here, as previously noted, \(V_1^n = LI_n\). Since \(m_1 = 0\), (3.8) becomes
\[
\frac{LI_n - n\pi_{\text{max}}}{\sqrt{n}} \implies \max_{(t, \ldots) \in I_{1, d_1}} \sum_{\ell=1}^{d_1} \sigma_{\tau(\ell)} \left( \tilde{B}^{\tau(\ell)}(t_1, \ell) - \tilde{B}^{\tau(\ell)}(t_1, \ell-1) \right), \tag{3.11}
\]
where we have written \(\pi_{\text{max}}\) for \(\pi_{\tau(1)}\). The functional in (3.11) is similar to the one obtained in the iid case in [17], the essential difference being, not in the form of the Brownian functional, but rather in the covariance structure of the Brownian motions.

To see precisely where this difference comes into play, note that if the transition matrix \(P\) is cyclic, then the covariance matrix of the Brownian motion is also cyclic. Consider then the 3-letter aperiodic, homogeneous, doubly-stochastic Markov case. Since the Brownian covariance matrix is symmetric, and, moreover, degenerate, an additional circularity constraint forces it to have the permutation-symmetric structure seen in the iid uniform case. In particular, \(LI_n\) will have as a limiting law, up to a scaling factor, the maximal eigenvalue of the traceless \(3 \times 3\) GUE:
\[
\frac{LI_n - n/3}{\sqrt{n}} \implies \sigma \max_{(t, \ldots) \in I_{1, 3}} \sum_{\ell=1}^{3} \left( \tilde{B}^{\ell}(t_1, \ell) - \tilde{B}^{\ell}(t_1, \ell-1) \right), \tag{3.12}
\]
where \(\sigma = \sigma_\ell\), for all \(1 \leq \ell \leq 3\), and with the Brownian covariance matrix given by
\[
\tilde{\Sigma} = \begin{pmatrix}
1 & -1/2 & -1/2 \\
-1/2 & 1 & -1/2 \\
-1/2 & -1/2 & 1
\end{pmatrix} t,
\]
and where we have used the fact that \(\tau(\ell) = \ell\), for all \(1 \leq \ell \leq 3\).

However, when \(m \geq 4\), the cyclicity constraint does not force the Brownian covariance matrix to have the permutation-symmetric structure, as the following example shows for \(m = 4\).
Example 3.1 Consider the following doubly-stochastic, aperiodic, cyclic transition matrix:

\[ P = \begin{pmatrix}
0.4 & 0.3 & 0.2 & 0.1 \\
0.1 & 0.4 & 0.3 & 0.2 \\
0.2 & 0.1 & 0.4 & 0.3 \\
0.3 & 0.2 & 0.1 & 0.4
\end{pmatrix}. \]  

(3.13)

While the doubly-stochastic nature of \( P \) ensures that the stationary distribution is uniform, the covariance matrix of the limiting Brownian motion, at three-decimal accuracy, is computed to be

\[ \tilde{\Sigma} = \begin{pmatrix}
1.000 & -0.357 & -0.287 & -0.357 \\
-0.357 & 1.000 & -0.357 & -0.287 \\
-0.287 & -0.357 & 1.000 & -0.357 \\
-0.357 & -0.287 & -0.357 & 1.000
\end{pmatrix}, \]  

(3.14)

and \( \sigma_r^2 = \sigma^2 = 0.263 \), for each \( 1 \leq r \leq 4 \). Thus,

\[ \frac{LI_n - n/4}{\sqrt{n}} \Rightarrow \sigma \max_{(t, \ldots) \in I_{1,4}} \sum_{\ell=j}^4 \left( \tilde{B}^\ell(t, \ell) - \tilde{B}^\ell(t, \ell-1) \right), \]  

(3.15)

for \( 1 \leq r \leq 4 \). However, while the form of the functional is the same as in the iid uniform case (up to the constant), the covariance structure of the Brownian motion in (3.14) differs from that of the uniform iid case, i.e., from

\[ \begin{pmatrix}
1 & -1/3 & -1/3 & -1/3 \\
-1/3 & 1 & -1/3 & -1/3 \\
-1/3 & -1/3 & 1 & -1/3 \\
-1/3 & -1/3 & -1/3 & 1
\end{pmatrix}, \]  

(3.16)

and so the limiting distribution in (3.15) is not that of the uniform iid case.

We thus see that Kuperberg’s conjecture regarding the shape of the RSK Young diagrams for random sequences generated by aperiodic, homogeneous, and cyclic matrices [24] is not true for general \( m \)-alphabets. By simply extending the first-row analysis above to the second and third rows, we see that it is true for \( m = 3 \). However, it fails for \( m \geq 4 \). In the next section we shall see that for the cyclic case the structure of \( \Sigma \) can be described in a way which delineates precisely when the limiting law is the spectrum of the traceless GUE.

In the more general doubly stochastic case, we have the following corollary:

**Corollary 3.1** Let the transition matrix \( P \) of Theorem 3.2 be doubly stochastic. Then, for every \( 1 \leq r \leq m, \ m_r = 0, d_r = m, \) and

\[ \frac{V_r - rn/m}{\sqrt{n}} \Rightarrow \max_{(t, \ldots) \in I_{r,m}} \sum_{j=1}^r \sum_{\ell=j}^{m-r+j} \sigma_\ell \left( \tilde{B}^\ell(t, j, \ell) - \tilde{B}^\ell(t, j, \ell-1) \right). \]  

(3.17)

If, moreover, the matrix \( P \) has all entries of \( 1/m \), then
\[
\frac{V_r^n - rn/m}{\sqrt{n}} \Rightarrow \sqrt{m-1} \max_{(t_\ldots) \in I_{r,m}} \sum_{j=1}^r \sum_{\ell=j}^{m-r+j} \left( \tilde{B}^\ell(t_j,\ell) - \tilde{B}^\ell(t_j,\ell-1) \right), \tag{3.18}
\]

and the covariance matrix in (3.9) has all its off-diagonals equal to \(-1/(m-1)\).

**Proof.** For each \(1 \leq r \leq m\), \(\pi_r = 1/m\), and so \(\nu_r = r/m\), \(m_r = 0\), and the multiplicity \(d_r = m\). Moreover, the permutation \(\tau\) is simply the identity permutation. This proves (3.17). If, moreover, all the transition probabilities are \(1/m\), then the multinomial nature of the underlying combinatorial quantities \(a^r_k\) tells us that \(\sigma^2_r = (1/m)(1-1/m)\), for each \(1 \leq r \leq m\), and that \(\rho_{r_1,r_2} = -1/(m-1)\), for each \(r_1 \neq r_2\), thus proving (3.18).

To see that the functional in (3.17) is generally different from the uniform iid case, even for \(m = 3\), consider the following non-cyclic example:

**Example 3.2** Let a doubly-stochastic (but non-cyclic), aperiodic Markov chain have transition matrix

\[
P = \begin{pmatrix}
 0.4 & 0.6 & 0.0 \\
 0.6 & 0.0 & 0.4 \\
 0.0 & 0.4 & 0.6
\end{pmatrix}. \tag{3.19}
\]

Again, the doubly-stochastic nature of \(P\) ensures that the stationary distribution is uniform, and in the present example, the asymptotic covariance matrix, at three-decimal accuracy, is computed to be

\[
\begin{pmatrix}
 0.459 & 0.049 & -0.506 \\
 0.049 & 0.086 & -0.136 \\
 -0.506 & -0.136 & 0.642
\end{pmatrix}. \tag{3.20}
\]

Note that, even though we have a uniform stationary distribution, the asymptotic variances (i.e., the diagonals of (3.20)) have dramatically different values. Moreover, according to Remark 2.2, in the uniform iid case, the only possibility for the Brownian covariance matrix is that the off-diagonals have value \(-1/2\). However, the Brownian motion covariance matrix obtained from (3.20) is

\[
\begin{pmatrix}
 1.000 & 0.246 & -0.935 \\
 0.246 & 1.000 & -0.577 \\
 -0.935 & -0.577 & 1.000
\end{pmatrix} t. \tag{3.21}
\]

Not only are the off-diagonals different from \(-1/2\), but in some cases are even positive. In short, the functional in (3.17) has a distribution which differs from any iid case (even non-uniform).

**Remark 3.3** Generalizing a result of Baryshnikov [4] and of Gravner, Tracy, and Widom [15] on the representation of the maximal eigenvalue of an \(m \times m\) element of the GUE, Doumerc [11] found a Brownian functional expression for all the eigenvalues of an \(m \times m\) element of the GUE. Our expression in (3.18) is similar, with the exception that our \(m\)-dimensional Brownian motion is constrained by a zero-sum condition, and, moreover, has a different covariance structure. (We note, moreover, that the parameters over which his Brownian functional is maximized in [11] might be intended to range over a slightly larger set which corresponds to our \(I_{r,m}\).) Using a path-transformation technique relating the joint distribution of a certain transformation of \(n\) continuous processes to the joint distribution of
the processes conditioned never to leave the Weyl chamber, O’Connell and Yor [26] employed queuing-theoretic arguments to obtain different types of representations for the entire spectrum of the \( m \times m \) element of the GUE. In a study of much more general transformations of this type, Bougerol and Jeulin [7] obtained this result as a special case and moreover these two representations are shown to be equivalent to each other in Biane, Bougerol and O’Connell [5]. In both these works, the maximal and minimal eigenvalues have the same representations and it is also the same as the one given by our Brownian functionals. However, our representation for the rest of the spectrum is different from theirs and the equivalence with the works cited above is not that immediate. Indeed, in our case, a single supremum is taken over a rather involved set while in theirs, sequence suprema and infima are taken over much simpler sets. Another piece of work of interest is reported by Kirillov [23], where minima and maxima can be disentangled via classes of linear transformations (see Theorem 3.3 and Theorem 3.5 in [22]).

If \( d_r = 1 \), i.e., if the \( r^{th} \) most probable state is unique, then the following result can be viewed as lying at the other extreme from Corollary 3.1.

**Corollary 3.2** Let \( 1 \leq r \leq m \), and let \( d_r = 1 \) in Theorem 3.2. Then

\[
\frac{V_r}{\sqrt{n}} - n\nu_r \Rightarrow \sum_{i=1}^{r} \sigma(\tau(i)) \tilde{B}(\tau(i)).
\]

**Proof.** If \( d_r = 1 \), then \( m_r = r - 1 \), and so the maximal term of (3.8) contains only one summand, namely \( \sigma(\tau(r)) \tilde{B}(\tau(r)) \). Including this term in the first summation term of (3.8) proves (3.22).

**Remark 3.4** The maximal term of the functional in (3.8) is that of the doubly-stochastic, \( d_r \)-letter case. Indeed, the maximal term involves precisely \( d_r \) Brownian motions over the \( r - m_r \) rows. Such a functional would arise in a doubly-stochastic \( d_r \)-letter situation with a covariance matrix consisting of the sub-matrix of the original \( \Sigma \) corresponding to the \( d_r \) Brownian motions, as in Corollary 3.1. The Gaussian term corresponds to the functional of Corollary 3.2. That is, in some sense, the limiting law of (3.8) interpolates between these two extreme cases.

**Proof.** (Theorem 3.2) Since the \( r = m \) case is trivial (\( V_n^m \) is then identically equal to \( n \)), assume that \( r < m \). Recall the approximating functional (3.7):

\[
\frac{V_r}{\sqrt{n}} - n\nu_r = \max_{r,m} \left\{ \sum_{j=1}^{m-r+j} \sigma(\tau(\ell)) \left( \tilde{B}(t_{\ell,j}) - \tilde{B}(t_{\ell,j-1}) \right) + \sum_{j=1}^{m-r+j} \sum_{\ell=j}^{t_{\ell,j}} \sigma(\tau(\ell)) \left( \tilde{B}(t_{\ell,j}) - \tilde{B}(t_{\ell,j-1}) \right) \right\}.
\]

(3.23)

Introducing the notation \( \Delta t_{j,\ell} := [t_{j,\ell-1}, t_{j,\ell-1}] \) and \( M_n^\ell(\Delta t_{j,\ell}) := M_n^\ell(t_{j,\ell}) - M_n^\ell(t_{j,\ell-1}) \), for any \( m \)-dimensional process \( M(t) = (M^1(t), M^2(t), \ldots, M^m(t)) \), \( t \in [0,1] \), we can rewrite (3.23) more compactly as
\[
\frac{V^r_n - n\nu_r}{\sqrt{n}} = \max_{I_{r,m}} \left\{ \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} \sigma^\ell \tilde{B}^\ell_n(\Delta t_{j,\ell}) - \sqrt{n} \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} (\pi_{\tau(j)} - \pi_\ell) |\Delta t_{j,\ell}| \right\}.
\] (3.24)

The main idea of the proof to follow will be to show that the second summation of (3.24) can, in effect, be eliminated by choosing the \(\Delta t_{j,\ell}\) in an appropriate manner. Now some of the coefficients \((\pi_{\tau(j)} - \pi_\ell)\) are zero; such terms do not cause any problems. Intuitively, however, the remaining terms should have \(|\Delta t_{j,\ell}| = 0\). Defining the restricted set of parameters \(I_{r,m}^* = \{(t_{j,\ell}) \in I_{r,m} : \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} (\pi_{\tau(j)} - \pi_\ell) |\Delta t_{j,\ell}| = 0, 1 \leq \ell \leq m\}\), we see that, provided \(I_{r,m}^* \neq \emptyset\),

\[
\max_{I_{r,m}^*} \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} \left( \sigma^\ell \tilde{B}^\ell_n(\Delta t_{j,\ell}) - \sqrt{n} (\pi_{\tau(j)} - \pi_\ell) |\Delta t_{j,\ell}| \right) \geq \max_{I_{r,m}^*} \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} \sigma^\ell \tilde{B}^\ell_n(\Delta t_{j,\ell}).
\] (3.25)

Moreover, by the Invariance Principle and the Continuous Mapping Theorem,

\[
\max_{I_{r,m}^*} \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} \sigma^\ell \tilde{B}^\ell_n(\Delta t_{j,\ell}) = \max_{I_{r,m}^*} \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} \sigma^\ell \tilde{B}^\ell_n(\Delta t_{j,\ell}).
\] (3.26)

We claim that, indeed, \(I_{r,m}^* \neq \emptyset\), and that, moreover,

\[
\max_{I_{r,m}^*} \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} \left( \sigma^\ell \tilde{B}^\ell_n(\Delta t_{j,\ell}) - \sqrt{n} (\pi_{\tau(j)} - \pi_\ell) |\Delta t_{j,\ell}| \right) = \max_{I_{r,m}^*} \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} \sigma^\ell \tilde{B}^\ell_n(\Delta t_{j,\ell}).
\] (3.27)

We will prove that \(I_{r,m}^* \neq \emptyset\) by creating a bijection between \(I_{r,m}^*\) and \(I_{r-m_r, d_r}\). To this end, for \(1 \leq i \leq m_r\), let \(\hat{I}_{r(i),i} = [u_{r(i),i-1}, u_{r(i),i}] = [0, 1]\). Next, choose any \((u_{r(i)}) \in I_{r-m_r, d_r}\), and define further intervals \(\hat{I}_{r(r+i),\ell} = \Delta u_{j,\ell}\), for \(1 \leq j \leq r - m_r\), \(1 \leq \ell \leq d_r\).

We now create a partition of these intervals in a manner which relies on the ideas used in the proof of Theorem 3.1. Consider the set of points \(\{u_{j,\ell}\}_{1 \leq j \leq r-m_r, 1 \leq \ell \leq d_r}\), and order them as \(s_0 := 0 < s_1 < \cdots < s_{k-1} < s_k := 1\), for some integer \(k\), and let \(\Delta s_q = [s_{q-1}, s_q]\), for all \(1 \leq q \leq k\).

Trivially, for each \(1 \leq q \leq k\), and for each \(1 \leq i \leq m_r\), \(\Delta s_q \subset \hat{I}_{r(i),i}\). Moreover, for each \(1 \leq j \leq r - m_r\), there exists a unique \(\ell(j, q)\) such that \(\Delta s_q \subset \hat{I}_{r(m_r + j),\ell(j, q)}\). For each \(q\), consider the set of indices \(A_q := \{\tau(1), \ldots, \tau(m_r)\} \cup \{\tau(m_r + \ell(1,q)), \ldots, \tau(m_r + \ell(r-m_r,q))\}\), and order these \(r\) elements of \(A_q\) as \(1 \leq \ell(1,q) < \cdots < \ell(r,q) \leq m\).

Using these partitions, we examine, with foresight, the following functional of a general \(m\)-dimensional process \((M(t))_{t \geq 0}\):

\[
\sum_{i=1}^{m_r} M^{(i)}(1) + \sum_{j=1}^{r-m_r} \sum_{\ell=j}^{r-m_r+d_r-1} M^{(m_r+\ell)}(\Delta u_{j,\ell})
\] (3.28)
\begin{align}
\sum_{i=1}^{m_r} \left( \sum_{q=1}^{\kappa} M^{r(i)}(\Delta s_q) \right) + \sum_{j=1}^{(r-m_r)} \sum_{\ell=j}^{(r-m_r+d_r-1)} \left( \sum_{q: \Delta s_q \subseteq I_r(m_r+j), \ell} M^{r(m_r+\ell)}(\Delta s_q) \right) \\
= \sum_{q=1}^{\kappa} \left( \sum_{i=1}^{m_r} M^{r(i)}(\Delta s_q) + \sum_{j=1}^{(r-m_r)} M^{r(m_r+\ell(j,q))}(\Delta s_q) \right) \\
= \sum_{q=1}^{\kappa} \sum_{j=1}^{r} M^{\ell(j,q)}(\Delta s_q) = \sum_{j=1}^{r} \sum_{q=1}^{\kappa} M^{\ell(j,q)}(\Delta s_q) \\
= \sum_{j=1}^{r} \sum_{\ell=1}^{r} M^{\ell(j,q)}(\Delta t_{j,\ell}), \tag{3.29}
\end{align}

where, for each 1 \leq j \leq r, and for each 1 \leq \ell \leq m, t_{j,\ell} := \max\{s_q : \ell \geq \ell(j,q)\}. (That is, for each j, we collapse together intervals \Delta s_q corresponding to the same component \(M^\ell.\) Now, since our functional in (3.29) has non-trivial summands only for \(\ell\) such that \(\pi_{\tau(\ell)} = \pi_{\tau(\ell)}\), we have shown that \((t,.)\) \(\in I^*_r,m\).

The following example illustrates this argument. Suppose we have an alphabet of size \(m = 8\), with

\((\pi_1, \pi_2, \ldots, \pi_8) = (0.07, 0.1, 0.2, 0.06, 0.2, 0.06, 0.1, 0.2).\)

Then,

\[\pi_{\tau(1)} = \pi_{\tau(2)} = \pi_{\tau(3)} = 0.2, \quad m_1 = m_2 = m_3 = 0, \quad d_1 = d_2 = d_3 = 3,\]

\[\pi_{\tau(4)} = \pi_{\tau(5)} = 0.1, \quad m_4 = m_5 = 3, \quad d_4 = d_5 = 2,\]

\[\pi_{\tau(6)} = 0.07, \quad m_6 = 5, \quad d_6 = 1,\]

\[\pi_{\tau(7)} = \pi_{\tau(8)} = 0.06, \quad m_7 = m_8 = 6, \quad d_7 = d_8 = 2.\]

In particular, note that the two largest, distinct probability values are 0.2 and 0.1, of multiplicities 3 and 2, respectively. Next, consider the case \(r = 4\). We now show how \(I_{r-m, d_r} = I_{4-3,2} = I_{1,2}\) corresponds to an element of \(I^*_r,m = I^*_4,8\). Figure 4 shows a typical element of the unconstrained index set \(I_{4,8}\).

Now \(\tau(1) = 3, \tau(2) = 5, \tau(3) = 8, \tau(4) = 2,\) and \(\tau(5) = 7.\) Our construction begins with the amalgamation of \(m_r = m_4 = 3\) rows, corresponding to the three indices for which \(\pi_i\) is strictly less than \(\pi_{\tau(r)} = \pi_{\tau(4)} = 0.1,\) with \(I_{1,2}\). This is shown in Figure 5.

Finally, we simply reorder each vertical column in the original order of the indices, as shown in Figure 6. We see that, first of all, we have constructed an element of \(I_{4,8}.\) Moreover, since we have three rows whose indices are associated with the maximum value, and a remaining row whose indices are associated with \(\pi_{\tau(4)},\) we indeed have an element of \(I^*_4,8\). Note that the \(4 \times 4 = 16\) free indices in \(I_{4,8}\) (corresponding to the locations of the 16 vertical bars in Figure 4) have been reduced to a \textit{single} one in \(I^*_4,8.\)

In addition, we may essentially reverse this construction, starting with an element of \(I^*_r,m \neq \emptyset,\) and so obtain an element of \(I_{r-m, d_r}.\) Indeed, from the definitions of \(I^*_r,m\) and \(\nu_r\) we know that

\[\nu_r = \sum_{j=1}^{r} \pi_{\tau(j)} = \sum_{j=1}^{r} \sum_{\ell=j}^{m-r+j} \pi_{\ell} |\Delta t_{j,\ell}|,\]
Figure 4: A typical element of $I_{4,8}$.

Figure 5: Amalgamating 3 rows with $I_{1,2}$.

Figure 6: Reordering vertically to obtain an element in $I_{4,8}^*$.
for any \((t_\ast) \in I_{r,m}^*\). However, we also have

\[
\sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} \pi_\ell |\Delta t_{j,\ell}| = 1_{\{m_r > 0\}} \left( \sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} 1_{\{\pi_\ell(t) \geq \pi_\tau(m_r)\}} \pi_\ell |\Delta t_{j,\ell}| \right)
+ \sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} 1_{\{\pi_\ell(t) < \pi_\tau(m_r)\}} \pi_\ell |\Delta t_{j,\ell}| 
+ 1_{\{m_r = 0\}} \sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} |\Delta t_{j,\ell}|
\leq 1_{\{m_r > 0\}} ((\pi_\tau(1) + \cdots + \pi_\tau(m_r)) + (r - m_r)\pi_\tau(r)) + 1_{\{m_r = 0\}} r\pi_\tau(1)
= \nu_r,
\]

with equality holding throughout if and only if \(m_r = 0\) or \(m_r > 0\) and \(\sum_{j=1}^{r} |\Delta t_{j,\ell}| = 1\), for all \(\ell\) such that \(\pi_\tau(\ell) \geq \pi_\tau(m_r)\), and that, moreover, \(\sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} 1_{\{\pi_\ell(t) = \pi_\tau(r)\}} |\Delta t_{j,\ell}| = r - m_r\). If \(m_r > 0\), then, for any \((t_\ast) \in I_{r,m}^*\), we may start with (3.29), and use again the permutation of the indices employed there. We thus obtain the first term of (3.28), which corresponds to the condition \(\sum_{j=1}^{r} |\Delta t_{j,\ell}| = 1\), for all \(\ell\) such that \(\pi_\tau(\ell) \geq \pi_\tau(m_r)\), and also the second term of (3.28), which corresponds to the other condition \(\sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} 1_{\{\pi_\ell(t) = \pi_\tau(r)\}} |\Delta t_{j,\ell}| = r - m_r\). If \(m_r = 0\) the same reasoning holds, except that the first term in (3.28) is taken to be zero.

Having thus established a bijection between \(I_{r,m}^*\) and \(I_{r-m_r, d_r}\), we may thus maximize over these two parameter sets, and so, for any process \((M(t))_{t \geq 0}\), obtain the general result

\[
\sum_{j=1}^{m_r} M^\tau_{\ast}(i)(1) + \max_{I_{r-m_r, d_r}} \sum_{j=1}^{(r-m_r) \cdot (r-m_r+d_r)-1} \sum_{\ell, j} M^\tau(m_r+\ell) (\Delta u_{j,\ell}) = \max_{I_{r,m}^*} \sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} M^\ell(j,\ell) (\Delta t_{j,\ell}). \tag{3.30}
\]

We now proceed to show that (3.27) holds. First, fix \(c > 0\), and, for each \(1 \leq \ell \leq m_r\), set

\[
c_\ell = \begin{cases} c, & \text{if } \pi_\ell \leq \pi_\tau(r), \\ 0, & \text{otherwise}. \end{cases} \tag{3.31}
\]

Next, let \(\widehat{M}_n^\ell(t) = \sigma_\ell \widehat{B}_n^\ell(t) - c_\ell t\), and let \(M^\ell(t) = \sigma_\ell \widehat{B}_n^\ell(t) - c_\ell t\). Then, for \(n\) large enough, namely, for \(n > c/\left(\pi_\tau(r) - \pi_\tau(r+1)\right)\), we have that, almost surely, for any \((t_\ast) \in I_{r,m}\),

\[
\sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} \widehat{M}_n^\ell(\Delta t_{j,\ell}) \geq \sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} \left( \sigma_\ell \widehat{B}_n^\ell(\Delta t_{j,\ell}) - \sqrt{n} (\pi_\tau(j) - \pi_\ell) |\Delta t_{j,\ell}| \right). \tag{3.32}
\]

Hence, almost surely, both

\[
\max_{I_{r,m}} \sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} \widehat{M}_n^\ell(\Delta t_{j,\ell}) \geq \max_{I_{r,m}} \sum_{j=1}^{r} \sum_{\ell=1}^{m-r+j} \left( \sigma_\ell \widehat{B}_n^\ell(\Delta t_{j,\ell}) - \sqrt{n} (\pi_\tau(j) - \pi_\ell) |\Delta t_{j,\ell}| \right), \tag{3.33}
\]
and
\[
\max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \hat{M}^\ell_n(\Delta t_{j,\ell}) = \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \sigma_\ell \hat{B}^\ell_n(\Delta t_{j,\ell}). \tag{3.34}
\]

Now choose any \( z > 0 \). Then
\[
\mathbb{P} \left( \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \left( \sigma_\ell \hat{B}^\ell_n(\Delta t_{j,\ell}) - \sqrt{n} \left( \pi_{\tau(j)} - \pi_\ell \right) |\Delta t_{j,\ell}| \right) - \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \sigma_\ell \hat{B}^\ell_n(\Delta s_{j,\ell}) > z \right)
\leq \mathbb{P} \left( \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \hat{M}^\ell_n(\Delta t_{j,\ell}) - \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \hat{M}^\ell_n(\Delta t_{j,\ell}) > z \right), \tag{3.35}
\]
so that
\[
\limsup_{n \to \infty} \mathbb{P} \left( \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \left( \sigma_\ell \hat{B}^\ell_n(\Delta t_{j,\ell}) - \sqrt{n} \left( \pi_{\tau(j)} - \pi_\ell \right) |\Delta t_{j,\ell}| \right) - \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \sigma_\ell \hat{B}^\ell_n(\Delta s_{j,\ell}) > z \right)
\leq \limsup_{n \to \infty} \mathbb{P} \left( \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \hat{M}^\ell_n(\Delta t_{j,\ell}) - \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} \hat{M}^\ell_n(\Delta t_{j,\ell}) > z \right)
\leq \mathbb{P} \left( \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) - \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) > z \right), \tag{3.36}
\]
by the Invariance Principle and the Continuous Mapping Theorem. Next, for any \( 0 \leq \varepsilon \leq 1 \), let
\[
I_{r,m}(\varepsilon) = \{ (t_{j,\ell}) \in I_{r,m} : \sum_{j,\ell} |\Delta t_{j,\ell}| 1_{\{ \pi_\ell < \pi_{\tau(j)} \}} \leq \varepsilon r \}.
\]
Thus, \( I_{r,m}^* = I_{r,m}(0) \subset I_{r,m}(\varepsilon) \subset I_{r,m}(1) = I_{r,m} \). We bound (3.36) using this family of subsets as follows:
\[
\mathbb{P} \left( \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) - \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) > z \right)
\leq \mathbb{P} \left( \max_{I_{r,m}(\varepsilon)} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) - \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) > z \right)
\leq \mathbb{P} \left( \max_{I_{r,m} \setminus I_{r,m}(\varepsilon)} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) - \max_{I_{r,m}} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) > z \right)
\leq \mathbb{P} \left( \max_{I_{r,m} \setminus I_{r,m}(\varepsilon)} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) - \max_{I_{r,m} \setminus I_{r,m}(\varepsilon)} \sum_{j=1}^r \sum_{\ell=1}^{m-r+j} M^\ell(\Delta t_{j,\ell}) > z \right)
\]
and thus we have in fact shown, with the help of (3.36), that with probability one,

\[ \max_{I_r,m(e)} \max_{j=1}^{r} \min_{\ell=1}^{r} \tilde{B}^{\ell}(\Delta t_{j,\ell}) - \max_{I_r,m} \max_{j=1}^{r} \min_{\ell=1}^{r} \tilde{B}^{\ell}(\Delta s_{j,\ell}) > z \]

we have in fact shown, with the help of (3.36), that with probability one,

\[ \max_{I_r,m(e)} \max_{j=1}^{r} \min_{\ell=1}^{r} \tilde{B}^{\ell}(\Delta t_{j,\ell}) - \max_{I_r,m} \max_{j=1}^{r} \min_{\ell=1}^{r} \tilde{B}^{\ell}(\Delta s_{j,\ell}) > z + \varepsilon r \]

Finally, we can obtain the convergence of the joint distribution in (3.10) in the following manner.

\[ \sum_{j=1}^{r} \sum_{\ell=1}^{r} \sigma_{\ell} \tilde{B}_{n}^{\ell}(\Delta t_{j,\ell}) - \sqrt{n} (\pi_{r(j)} - \pi_{\ell}) |\Delta t_{j,\ell}| \to 0. \]

Since

\[ \max_{I_r,m(e)} \max_{j=1}^{r} \min_{\ell=1}^{r} \sigma_{\ell} \tilde{B}_{n}^{\ell}(\Delta s_{j,\ell}) = \max_{I_r,m} \max_{j=1}^{r} \min_{\ell=1}^{r} \sigma_{\ell} \tilde{B}_{n}^{\ell}(\Delta s_{j,\ell}), \]

by the Converging Together Lemma, we have proved (3.27). Equation (3.3) of the theorem follows from the bijection between \( I_{r,m} \) and \( I_{r-m,d} \) described in the general result (3.30).

Finally, we can obtain the convergence of the joint distribution in (3.10) in the following manner. Given any \((\theta_1, \theta_2, \ldots, \theta_r) \in \mathbb{R}^r\), we have

\[ \sum_{k=1}^{r} \theta_k \left( \frac{V^k_m - nV^k_k}{\sqrt{n}} \right) \]

\[ = \sum_{k=1}^{r} \theta_k \left( \max_{j=1}^{m-k+j} \sum_{\ell=j}^{r} \left( \sigma_{\ell} \tilde{B}_{n}^{\ell}(\Delta t_{j,\ell}) - \sqrt{n} (\pi_{r(j)} - \pi_{\ell}) |\Delta t_{j,\ell}| \right) \right) \]
\[
\sum_{k=1}^{r} \theta_k \left( \max_{I_{k,m}} \sum_{j=1}^{k} \sum_{\ell=j}^{m-k+j} \left( \sigma_\ell \tilde{B}_n^\ell(\Delta t_{j,\ell}) - \sqrt{n} (\pi_\tau(j) - \pi_\ell) \left| \Delta t_{j,\ell} \right) \right) - \max_{I_{k,m}} \sum_{j=1}^{k} \sum_{\ell=j}^{m-k+j} \sigma_\ell \tilde{B}_n^\ell(\Delta s_{j,\ell}) \right) + \sum_{k=1}^{r} \theta_k \left( \max_{I_{k,m}} \sum_{j=1}^{k} \sum_{\ell=j}^{m-k+j} \sigma_\ell \tilde{B}_n^\ell(\Delta s_{j,\ell}) \right). \tag{3.40}
\]

Now from (3.38), the first summation on the right-hand side of (3.40) converges to zero in probability, as \(n \to \infty\). Moreover, the second summation is a continuous functional of \((\tilde{B}_n^1, \tilde{B}_n^2, \ldots, \tilde{B}_n^m)\), and so, by the Invariance Principle and Continuous Mapping Theorem, converges. Then the Converging Lemma, along with the bijection result (3.30), gives

\[
\sum_{k=1}^{r} \theta_k \left( \frac{V_n^k - n\nu_k}{\sqrt{n}} \right) \Rightarrow \sum_{k=1}^{r} \theta_k \left( \max_{I_{k,m}} \sum_{j=1}^{k} \sum_{\ell=j}^{m-k+j} \sigma_\ell \tilde{B}_n^\ell(\Delta s_{j,\ell}) \right) = \sum_{k=1}^{r} \theta_k V_{\infty}^k. \tag{3.41}
\]

Since (3.41) holds for arbitrary \((\theta_1, \theta_2, \ldots, \theta_r) \in \mathbb{R}^r\), by the Cramèr-Wold Theorem, we have the joint convergence result (3.10).

Since the shape of the Young diagrams is more naturally expressed in terms of the \(R_n^k\), rather than of the \(V_n^k\), we may restate the results of the previous theorem as follows:

**Theorem 3.3** Let \((X_n)_{n \geq 0}\) be an irreducible, aperiodic, homogeneous Markov chain with finite state space \(\mathcal{A}_m = \{\alpha_1 < \cdots < \alpha_m\}\), and with stationary distribution \((\pi_1, \pi_2, \ldots, \pi_m)\). Then, in the notations of Theorem 3.3,

\[
\left( \frac{R_n^1 - n\pi_{\tau(1)}}{\sqrt{n}}, \frac{R_n^2 - n\pi_{\tau(2)}}{\sqrt{n}}, \ldots, \frac{R_n^m - n\pi_{\tau(m)}}{\sqrt{n}} \right) \Rightarrow (R_\infty^1, R_\infty^2, \ldots, R_\infty^m), \tag{3.42}
\]

where

\[
R_\infty^1 = \max_{l_1, d_1} \sum_{\ell=1}^{d_1} \sigma_{\tau(\ell)} \left( \tilde{B}_{\tau(\ell)}(t_{1,\ell}) - \tilde{B}_{\tau(\ell)}(t_{1,\ell-1}) \right), \tag{3.43}
\]

and, for each \(2 \leq k \leq m\),

\[
R_\infty^k = \sum_{i=m_{k-1}+1}^{m_k} \sigma_{\tau(i)} \tilde{B}_{\tau(i)}(1) + \max_{l_{k-m_k,d_k}} \sum_{j=1}^{k-m_k} \sum_{\ell=j}^{d_k} \sigma_{\tau(m_k+\ell)} \tilde{B}_{\tau(m_k+\ell)}(\Delta t_{j,\ell}) - \max_{l_{k-1-m_{k-1},d_{k-1}}} \sum_{j=1}^{k-1-m_{k-1}} \sum_{\ell=j}^{d_{k-1}} \sigma_{\tau(m_{k-1}+\ell)} \tilde{B}_{\tau(m_{k-1}+\ell)}(\Delta t_{j,\ell}), \tag{3.44}
\]

24
where we use the notation $\tilde{B}^s(\Delta t_{j,\ell}) = \tilde{B}^s(t_{j,\ell}) - \tilde{B}^s(t_{j,\ell-1})$, for any $1 \leq s \leq m$, $1 \leq j \leq k$, and $1 \leq \ell \leq m$, and where the first sum on the right-hand side of (3.44) is understood to be 0, if $m_k = m_{k-1}$.

**Proof.** First, $R_n^1 = V_n^1$, and, for each $2 \leq k \leq m$, $R_n^k = V_n^k - V_n^{k-1}$. Expressing these equalities at the multivariate level, we have

$$
\left(\frac{R_n^1 - n\pi_{\tau(1)}}{\sqrt{n}}, \frac{R_n^2 - n\pi_{\tau(2)}}{\sqrt{n}}, \ldots, \frac{R_n^m - n\pi_{\tau(m)}}{\sqrt{n}}\right)
= \left(\frac{V_n^1 - n\pi_{\tau(1)}}{\sqrt{n}}, \frac{V_n^2 - V_n^1 - n\pi_{\tau(2)}}{\sqrt{n}}, \ldots, \frac{V_n^m - V_n^{m-1} - n\pi_{\tau(m)}}{\sqrt{n}}\right)
= \left(\frac{V_n^1 - n\nu_1}{\sqrt{n}}, \frac{V_n^2 - n\nu_2}{\sqrt{n}}, \ldots, \frac{V_n^m - n\nu_m}{\sqrt{n}}\right) - \left(0, \frac{V_n^1 - n\nu_1}{\sqrt{n}}, \ldots, \frac{V_n^{m-1} - n\nu_{m-1}}{\sqrt{n}}\right)
\implies (V_n^1, V_n^2, \ldots, V_n^m) - (0, V_n^1, \ldots, V_n^{m-1}) := (R_n^{\infty,1}, R_n^{\infty,2}, \ldots, R_n^{\infty,m}),
$$

where the weak convergence follows immediately from the Continuous Mapping Theorem, since the transformation is linear. Equations (3.43) and (3.44) follow simply from the Brownian expressions for $(V_n^1, V_n^2, \ldots, V_n^m)$ in Theorem 3.2.

If all $m$ letters have unique stationary probabilities, then we have the following corollary to Theorem 3.3.

**Corollary 3.3** If the stationary distribution of Theorem 3.3 is such that each $\pi_{\tau}$ is unique, then

$$
\left(\frac{R_n^1 - n\pi_{\tau(1)}}{\sqrt{n}}, \frac{R_n^2 - n\pi_{\tau(2)}}{\sqrt{n}}, \ldots, \frac{R_n^m - n\pi_{\tau(m)}}{\sqrt{n}}\right) \implies N((0,0,\ldots,0), \Sigma).
$$

In other words, the limiting distribution is identical in law to the spectrum of the diagonal matrix $D = \text{diag}\{Z_1, Z_2, \ldots, Z_m\}$, where $(Z_1, Z_2, \ldots, Z_m)$ is a centered normal random vector with covariance matrix $\Sigma$.

**Proof.** Now, for all $1 \leq k \leq m$, $d_k = 1$, and $m_k = k - 1$, so that

$$
R_n^{\infty,1} = \max_{I_1d_1} \sum_{\ell=1}^{d_1} \sigma_{\tau(\ell)} \left(\tilde{B}^{\tau(\ell)}(t_{1,\ell}) - \tilde{B}^{\tau(\ell)}(t_{1,\ell-1})\right) = \sigma_{\tau(1)} \tilde{B}^{\tau(1)}(1),
$$

and, for each $2 \leq k \leq m$,

$$
R_n^{\infty,k} = \sum_{i=m_{k-1}+1}^{m_k} \sigma_{\tau(i)} \tilde{B}^{\tau(i)}(1) + \max_{I_kd_k} \sum_{j=1}^{k-m_k} \sum_{\ell=j}^{k-m_k-d_k} \sigma_{\tau(m_k+j)} \tilde{B}^{\tau(m_k+j)}(\Delta t_{j,\ell})
$$

25
\[
\sum_{d} k-1-m_{k-1} (d_{k-1}+m_{k-1}-k+1+j) \sum_{\ell=1}^{\sigma_{k-1+\ell}} \tilde{B}^{\tau(k-1+\ell)}(\Delta t_{j,\ell})
\]

Moreover, the joint law result for \( (R_{1,\infty}^1, R_{2,\infty}^2, \ldots, R_{m,\infty}^m) \) holds as well, and this is clearly a multivariate normal distribution, with mean \((0,0,\ldots,0)\) and covariance matrix \(\Sigma\). Since the spectrum of a diagonal matrix consists of its diagonal elements, the final claim of the corollary holds.

\[\square\]

**Remark 3.5** The joint law of \( (R_{1,\infty}^1, R_{2,\infty}^2, \ldots, R_{m,\infty}^m) \) in the iid uniform alphabet case is identical to the joint law of the eigenvalues of an \( m \times m \) traceless GUE matrix. Corollary 3.3 also gives a spectral characterization for the unique probability case, in particular, for a non-uniform iid alphabet with unique stationary probabilities. This is consistent with the characterization of the limiting law of \( LI_{\infty} = \sum_{i=1}^{\infty} I_{i} \) in the non-uniform iid case, due to Its, Tracy, and Widom \[20, 21\], as that of the largest eigenvalue of the block associated with the most probable letters among a direct sum of independent GUE matrices whose dimensions correspond to the multiplicities \( d_r \) of Theorem 3.2 and 3.3 subject to the condition that \( \sum_{r=1}^{m} \sqrt{\tau(r)} X_r = 0 \), where \( X_1, X_2, \ldots, X_m \) are the diagonal elements of the random matrix.

**Remark 3.6** The difference between the zero-trace condition \( \sum_{r=1}^{m} X_r = 0 \) and the generalized traceless condition \( \sum_{r=1}^{m} \sqrt{\tau(r)} X_r = 0 \) amounts to nothing more than a difference in the choice of scaling for each row \( R_r^m \). We will find it more natural to express our results using the normalization associated with the zero-trace condition \( \sum_{r=1}^{m} X_r = 0 \).

### 4 Fine Structure of the Brownian Functional

So far, we have seen that the limiting shape of the RSK random Young diagrams generated by an aperiodic, irreducible, homogeneous Markov chain can be expressed as a Brownian functional. The form of this functional is similar to the iid case; the essential difference is in the covariance structure of the Brownian motion. We begin our study of the consequences of this difference.

In the iid uniform \( m \)-alphabet case, Johansson \[22\] proved that the limiting shape of the Young diagrams had a joint law which is that of the spectrum of an \( m \times m \) traceless GUE matrix. An immediate consequence of this result is that the limiting shape of the Young diagrams contains simple symmetries, e.g., for each \( 1 \leq r \leq m \), \( R_{r,\infty}^r = -R_{m+1-r,\infty}^m \). Now, as was seen in Corollary 3.1 of Theorem 3.2 the form of the Brownian functional in the doubly stochastic case involved only the maximal term. We will see that there is also a pleasing symmetry to the limiting shape of Young diagrams in the doubly stochastic case by examining a natural bijection between the parameter set \( I_{r,m} \) and \( I_{m-r,m} \),

\[\text{26}\]
for any \(1 \leq r \leq m - 1\). Indeed, this result will follow as a corollary to the following, more general, theorem:

**Theorem 4.1** The limiting functionals of Theorem 3.2 enjoy the following symmetry property: for every \(1 \leq r \leq m - 1\),

\[
V'_\infty := \sum_{i=1}^{m_r} \sigma_{\tau(i)} \tilde{B}^\tau(i)(1) + \max_{t(\cdot) \in I_{r-m_r,dr}} \sum_{j=1}^{r-m_r} \sum_{\ell=j}^{m_r+d_r-r+j} \sigma_{\tau(m_r+\ell)} \tilde{B}^\tau(m_r+\ell)(\Delta t_{j,\ell}) \equiv \sum_{i=m_r+d_r+1}^{m} \sigma_{\tau(i)} \tilde{B}^\tau(i)(1) + \max_{u(\cdot) \in I_{m_r+d_r-r,dr}} \sum_{j=1}^{m_r+d_r-r} \sum_{\ell=j}^{r-m_r+j} \sigma_{\tau(m_r+\ell)} \tilde{B}^\tau(m_r+\ell)(\Delta u_{j,\ell}),
\]

where \(\tilde{B}^\ell(\Delta) := \tilde{B}^\ell(t) - \tilde{B}^\ell(s)\), for \(\Delta = [s, t]\), and where the non-maximal terms on the left and right-hand sides of (4.1) are identically zero if \(m_r = 0\), or \(m_r + d_r = m\), respectively.

**Remark 4.1** Recall that, from the definitions of \(m_r\) and \(d_r\), the non-maximal summation terms on the left and right-hand sides of (4.1) reflect the letters which have, respectively, greater and smaller stationary probabilities than \(\pi_r\). Recall, moreover, that the maximal terms are associated with the indices having the same stationary probability as \(\pi_r\). The maximal term on the left-hand side of (4.1) involves a summation over \(r - m_r\) rows, while the one on the right-hand side involves \(m_r + 1 - r\) rows. Thus, in a sense, the two maximal terms in (4.1) split \(d_r = m_r + 1 - m_r\) rows between themselves.

In summary, the functional on the right-hand side of (4.1) corresponds to the sum of the lengths of the \(m - r\) bottom rows of the Young diagrams.

**Proof.** Without loss of generality, we may assume that \(\tau(j) = j\), for all \(1 \leq j \leq m\). Fix \(1 \leq r \leq m - 1\), and for any point \(t\) in the index set \(I_{r-m_r,dr}\), define \(\Delta t_{j+m_r,\ell} = [t_{j,\ell-1}, t_{j,\ell}]\), for \(1 \leq j \leq r - m_r\) and \(1 \leq \ell \leq d_r\). Furthermore, for each \(1 \leq j \leq m_r\) or \(m_r+1 < j \leq m\), set \(\Delta t_{j,\ell} = [0, 1]\), for \(j = \ell\), \(\Delta t_{j,\ell} = \{0\}\), for \(0 \leq \ell < j\), and \(\Delta t_{j,\ell} = \{1\}\), for \(j < \ell \leq m\). Next, as in the proof of Theorem 3.2 consider the set of points \(\{t_{j,\ell}\}_{1 \leq j \leq r - m_r, 1 \leq \ell \leq d_r}\), and order them as \(s_0 := 0 < s_1 < \cdots < s_{\kappa-1} < s_\kappa := 1\), for some integer \(\kappa\), and let \(\Delta s_q = [s_{q-1}, s_q]\), for each \(1 \leq q \leq \kappa\).

Now, for each \(1 \leq q \leq \kappa\), let \(A_q\) consist of the indices \(\ell\) for which \(\Delta s_q \cap \Delta t_{j,\ell} \neq \emptyset\). Then, almost surely,

\[
\sum_{i=1}^{m_r} \sigma_{\tau(i)} \tilde{B}^\tau(i)(1) + \sum_{j=1}^{r-m_r} \sum_{\ell=j}^{m_r+d_r-r+j} \sigma_{\tau(m_r+\ell)} \tilde{B}^\tau(m_r+\ell)(\Delta t_{j,\ell}) = \sum_{j=1}^{r} \sum_{\ell=1}^{m_r} \sigma_{\tau(\ell)} \tilde{B}^\tau(\Delta t_{j,\ell}) = \sum_{j=1}^{r} \sum_{\ell=1}^{m_r} \sigma_{\tau(\ell)} \tilde{B}^\tau(\Delta t_{j,\ell} \cap \Delta s_q) = \sum_{j=1}^{r} \sum_{\ell=1}^{m_r} \sigma_{\tau(\ell)} \tilde{B}^\tau(\Delta s_q),
\]

Now by the “stairstep” properties of \(I_{r,m}\) there are precisely \(r\) elements in each \(A_q\). Letting \(\hat{A}_q = \{1, \ldots, m\} \setminus A_q\), for each \(1 \leq q \leq \kappa\), we thus see that each \(\hat{A}_q\) contains exactly \(m - r\) elements.
Let \( \tilde{\ell}_{j,q} \) be the \( j \)th smallest element of \( \tilde{A}_q \). We claim that for each \( 1 \leq j \leq m - r \), the sequence \( \tilde{\ell}_{j,1}, \tilde{\ell}_{j,2}, \ldots, \tilde{\ell}_{j,K} \) is weakly decreasing.

Indeed, fix \( 1 \leq j \leq m - r \) and \( 1 \leq q \leq \kappa - 1 \), and suppose that \( \tilde{\ell}_{j,q} \) is less than all the elements of \( A_q \). Then, by the properties of \( I_{r,m} \), the least element of \( A_{q+1} \) is no smaller, so that the \( j \)th smallest element of \( \tilde{A}_q \), \( \tilde{\ell}_{j,q+1} \) is also \( \tilde{\ell}_{j,q} \). Next, suppose that \( \tilde{\ell}_{j,q} \) is greater than \( k \) elements of \( A_q \). Thus, \( \tilde{\ell}_{j,q} = j + k \). Then there are at most \( k \) elements of \( A_{q+1} \) which are less than or equal to \( \tilde{\ell}_{j,q} \), by the properties of \( I_{r,m} \). But this implies that there are at least \( j \) elements of \( \tilde{A}_{q+1} \) which are less than or equal to \( \tilde{\ell}_{j,q} \). Thus, \( \tilde{\ell}_{j,q+1} \leq \tilde{\ell}_{j,q} \), and the claim is proved.

Moreover, since each \( A_q \) contains \( \{1, 2, \ldots, m_r\} \), we see that necessarily each \( \tilde{A}_q \) contains \( \{m_r + d_r + 1, m_r + d_r + 2, \ldots, m \} \).

For each \( 1 \leq j \leq m - r \), we may now amalgamate the intervals \( \Delta s_q \) to obtain a partition of the unit interval. Specifically, for each \( 1 \leq j \leq m - r \), and each \( 1 \leq \ell \leq m \), let \( \tilde{u}_{j,\ell} \) be the smallest \( s_q \) such that \( \tilde{\ell}_{j,q+1} \leq \ell \). (We define \( \tilde{u}_{j,0} = 1 \), for all \( 1 \leq j \leq m - r \).

Finally, and most crucially, recall that \( \sum_{\ell=1}^{m} \sigma_{\ell} \bar{B}^\ell(t) = 0 \), for all \( t \). Then since \( (\bar{B}^1, \bar{B}^2, \ldots, \bar{B}^m) \) \( \subseteq \) \( (-\bar{B}^1, -\bar{B}^2, \ldots, -\bar{B}^m) \),

\[
\sum_{j=1}^{r} \sum_{q=1}^{\kappa} \sum_{\ell \in \tilde{A}_q} \sigma_{\ell} \bar{B}^{\ell}(\Delta s_q) = \sum_{j=1}^{r} \sum_{q=1}^{\kappa} \sum_{\ell \in \tilde{A}_q} \left( -\sigma_{\ell} \bar{B}^{\ell}(\Delta s_q) \right) \\
= -\sum_{i=m_r + d_r + 1}^{m} \sigma_i \bar{B}^i(1) - \sum_{j=1}^{m_r + d_r - r} \sum_{\ell=1}^{m} \sigma_{m_r + \ell} \bar{B}^{m_r + \ell}(\Delta u_{j,\ell}) \\
\subseteq \sum_{i=m_r + d_r + 1}^{m} \sigma_i \bar{B}^i(1) + \sum_{j=1}^{m_r + d_r - r} \sum_{\ell=1}^{m} \sigma_{m_r + \ell} \bar{B}^{m_r + \ell}(\Delta u_{j,\ell}), \tag{4.3}
\]

where \( \Delta u_{j,\ell} = [u_{j,\ell-1}, u_{j,\ell}] \). But, by the way we ordered each \( A_q \), we must have \( \Delta u_{j_1,\ell} \cap \Delta u_{j_2,\ell} = \emptyset \), for any \( j_1 \neq j_2 \). Thus, \( u \in I_{m_r + d_r - r,d_r} \), and so we may restrict the summation over \( \ell \) in (4.3) to \( \ell = j, \ldots, r - m_r + j \), since the remaining terms are zero. Equation (4.1) follows immediately by taking the maxima over \( I_{r-r,d_r} \) and \( I_{m_r + d_r - r,d_r} \) over the left-hand and right-hand sides, respectively, of (4.3).

For doubly stochastic transition matrices, the symmetry is even more apparent:

**Corollary 4.1** Let the transition matrix \( P \) of Theorem 3.2 be doubly stochastic. Then, for every \( 1 \leq r \leq m - 1 \),

\[
V^r := \max_{t(\cdot) \in I_{r,m}} \sum_{j=1}^{m-r+j} \sum_{\ell=j}^{m-r} \sigma_\ell \left( \bar{B}^\ell(t_{j,\ell}) - \bar{B}^\ell(t_{j,\ell-1}) \right) \\
\subseteq \max_{u(\cdot) \in I_{m-r,m}} \sum_{j=1}^{m-r+j} \sum_{\ell=j}^{m-r} \sigma_\ell \left( \bar{B}^\ell(u_{j,\ell}) - \bar{B}^\ell(u_{j,\ell-1}) \right) := V^m - r, \tag{4.4}
\]

and so
$$\lim_{n \to \infty} \frac{\sum_{j=1}^{r} R_j^n - rn/m}{\sqrt{n}} \overset{\mathcal{L}}{=} \lim_{n \to \infty} \frac{rn/m - \sum_{j=m-r+1}^{m} R_j^n}{\sqrt{n}}. \quad (4.5)$$

Moreover,

$$(V_{\infty}^1, \ldots, V_{\infty}^r) \overset{\mathcal{L}}{=} (V_{\infty}^{m-1}, \ldots, V_{\infty}^{m-r}). \quad (4.6)$$

**Proof.** Since $m_r = 0$ and $d_r = m$ for all $1 \leq r \leq m$, the non-maximal terms on both sides of (4.1) disappear, and we have (4.4).

To prove (4.5), recall that $V_m^n = \sum_{j=1}^{m} R_j^n = n$. Then, from the result just proved,

$$\frac{V_{n}^{m-r} - (m-r)n/m}{\sqrt{n}} = \frac{\sum_{j=1}^{m-r} R_j^n - (m-r)n/m}{\sqrt{n}} = \frac{(n - \sum_{j=m-r+1}^{m} R_j^n) - (m-r)n/m}{\sqrt{n}} = \frac{rn/m - \sum_{j=m-r+1}^{m} R_j^n}{\sqrt{n}} = \Longrightarrow V_{\infty}^{m-r} \overset{\mathcal{L}}{=} V_{\infty}^r, \quad (4.7)$$

and we have established the claimed symmetry.

Finally, the extension of (4.4) to (4.6) follows from a standard Cramér-Wold argument. ■

**Remark 4.2** Since $R_m^r = -V_{\infty}^{m-1}$, almost surely, Corollary 4.1 states that $R_m^r \overset{\mathcal{L}}{=} -R_{\infty}^1$. From the symmetry of the Brownian motion, we thus see that $R_m^r$ may be represented as a minimal Brownian functional:

$$R_m^r = \min_{\ell_1, m} \sum_{\ell=1}^{m} \sigma_\ell \left( \tilde{B}^\ell(t_{1,\ell}) - \tilde{B}^\ell(t_{1,\ell-1}) \right),$$

a result also noted in [26].

Turning again to the cyclic case, recall that, for $m \geq 4$, the limiting shape of the Young diagrams in general differs from that of the iid uniform case. The following proposition characterizes the asymptotic covariance matrices of such Markov chains.

**Proposition 4.1** Let $P$ be the $m \times m$ transition matrix of an aperiodic, irreducible, cyclic Markov chain on an $m$-letter, ordered alphabet, $A_m = \{\alpha_1 < \alpha_2 < \cdots < \alpha_m\}$, with

$$P = \begin{pmatrix}
    a_1 & a_m & \cdots & a_3 & a_2 \\
    a_2 & a_1 & \cdots & a_3 \\
    \vdots & \ddots & \ddots & \ddots & \vdots \\
    a_{m-1} & \cdots & a_1 & a_m \\
    a_m & a_{m-1} & \cdots & a_2 & a_1
\end{pmatrix}. \quad (4.8)$$
For $1 \leq j \leq m$, let $\lambda_j = \sum_{k=1}^{m} a_k \omega^{(k-1)(j-1)}$ be an eigenvalue of $P$, where $\omega = \exp(2\pi i/m)$ is the $m$th principal root of unity. Let also $\gamma_j = \lambda_j/(1 - \lambda_j)$, for $2 \leq j \leq m$, and $\beta_j = \cos(2\pi j/m)$, for $0 \leq j \leq m$. Then, the asymptotic covariance matrix $\Sigma$ is given by:

For $m = 2m_0 + 1$,

$$\Sigma = \frac{m-1}{m^2} M^{(1)} + \frac{4}{m^2} \sum_{j=2}^{m_0+1} \text{Re}(\gamma_j) M^{(j)},$$

and for $m = 2m_0$,

$$\Sigma = \frac{m-1}{m^2} M^{(1)} + \frac{4}{m^2} \sum_{j=2}^{m_0} \text{Re}(\gamma_j) M^{(j)} + \frac{2}{m^2} \gamma_{m_0+1} M^{(m_0+1)},$$

where $M^{(j)}$ is an $m \times m$ Toeplitz matrix with entries $(M^{(j)})_{k,\ell} = \beta_{(j-1)(k-\ell)}$, for $2 \leq j \leq m$, and $(M^{(1)})_{k,\ell} = \delta_{k,\ell} - (1 - \delta_{k,\ell})/(m-1)$, for $j = 1$.

**Proof.** It is straightforward, and classical, to verify that, for each $1 \leq j \leq m$, $(1, \omega^{j-1}, \omega^{2(j-1)}, \ldots, \omega^{(m-1)(j-1)})$ is a left eigenvector of $P$, with eigenvalue $\lambda_j = \sum_{k=1}^{m} a_k \omega^{(k-1)(j-1)}$. We can thus write our standard Jordan decomposition of $P$, which in this case is a true diagonalization, as $P = S^{-1}AS$, where $A = \text{diag}(1, \lambda_2, \ldots, \lambda_m)$,

$$S = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{m-1} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & \omega^{m-1} & \omega^{2(m-1)} & \cdots & \omega^{(m-1)^2} \end{pmatrix},$$

and

$$S^{-1} = \frac{1}{m} \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \cdots & \omega^{-(m-1)} \\ 1 & \omega^{-2} & \omega^{-4} & \cdots & \omega^{-(2(m-1))} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & \omega^{-(m-1)} & \omega^{-(2(m-1))} & \cdots & \omega^{-(m-1)^2} \end{pmatrix}.$$  

In the present cyclic, and hence, doubly stochastic case, we know that $\Sigma = (1/m) (I + S^{-1}DS + (S^{-1}DS)^T)$, where, as usual, $D = \text{diag}(\gamma_1, \gamma_2, \ldots, \gamma_m) = \text{diag}(-\gamma_1/2, \lambda_2/(1 - \lambda_2), \ldots, \lambda_m/(1 - \lambda_m))$. We can then compute the entries of $S^{-1}DS$ as follows:

$$(S^{-1}DS)_{j_1,j_2} = \sum_{k,\ell} (S^{-1})_{j_1,k}(D)_{k,\ell}(S)_{\ell,j_2}$$

$$= \sum_{k,\ell} \frac{1}{m} \omega^{-(j_1-1)(k-1)} \delta_{k,\ell} \gamma_k \omega^{(j_2-1)(\ell-1)}$$

$$= \sum_{k,\ell} \frac{1}{m} (\omega^{-(j_1-1)}(k-1)) \delta_{k,\ell} \gamma_k \omega^{(j_2-1)(\ell-1)}$$

$$= \sum_{k,\ell} \frac{1}{m} \omega^{-(j_1-1)(k-1)} \delta_{k,\ell} \gamma_k \omega^{(j_2-1)(\ell-1)}$$

30
\[ \sum_{k=1}^{m} \gamma_k \omega(j_2-j_1)(k-1) = \frac{1}{m} \left( -\frac{1}{2} + \sum_{k=2}^{m} \gamma_k \omega(j_2-j_1)(k-1) \right), \tag{4.13} \]

for all \( 1 \leq j_1, j_2 \leq m \). The entries of the asymptotic covariance matrix \( \Sigma \) can thus be written as

\[ \sigma_{j_1,j_2} = \frac{1}{m} \left( \delta_{j_1,j_2} + (S^{-1}DS)_{j_1,j_2} + (S^{-1}DS)_{j_2,j_1} \right) \]
\[ = \frac{1}{m} \left( \delta_{j_1,j_2} + \frac{1}{m} \left( -1 + \sum_{k=2}^{m} \gamma_k (\omega(j_2-j_1)(k-1) + \omega(j_1-j_2)(k-1)) \right) \right) \]
\[ = \frac{m-1}{m^2} M_{j_1,j_2}^{(1)} + \frac{2}{m^2} \sum_{k=2}^{m} \gamma_k \beta_{j_2-j_1}(k-1), \tag{4.14} \]

for all \( 1 \leq j_1, j_2 \leq m \).

Next, note that since \( \lambda_{m+2-k} = \bar{\lambda}_k \), i.e., the complex conjugate of \( \lambda_{m+2-k} \), we have \( \gamma_{m+2-k} = \bar{\gamma}_k \), for all \( 2 \leq k \leq m \). Moreover, since \( \beta_{j_2-j_1}(k-1) = \beta_{j_2-j_1}((m+2-k)-1) \), we can write (4.14) more symmetrically as (4.9) or (4.10), depending on whether \( m \) is odd or even, respectively, and in the latter case, we also use that \( \gamma_{m+1} \) is real, since \( \omega^{m+1} = -1 \).

Let us again examine the cases \( m = 3 \) and \( m = 4 \). In the former case, we have

\[ M^{(1)} = \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix}. \]

But for \( m = 3 \), \( \beta_1 = -1/2 = \beta_2 \), and so \( M^{(1)} = M^{(2)} \). Hence

\[ \Sigma = \frac{2}{9} M^{(1)} + \frac{4}{9} \text{Re}(\gamma_2) M^{(2)} = \frac{2}{9} (1 + 2 \text{Re}(\gamma_2)) M^{(1)}. \tag{4.15} \]

Hence, for \( m = 3 \), cyclicity always produces a rescaled version of the uniform iid case, with the rescaling factor given by \( 1 + 2 \text{Re}(\gamma_2) \).

For \( m = 4 \), however,

\[ M^{(1)} = \begin{pmatrix} 1 & -1/3 & -1/3 & -1/3 \\ -1/3 & 1 & -1/3 & -1/3 \\ -1/3 & -1/3 & 1 & -1/3 \\ -1/3 & -1/3 & -1/3 & 1 \end{pmatrix}, \]

and \( \beta_1 = 0, \beta_2 = -1, \text{ and } \beta_3 = 0 \). Thus,

\[ M^{(2)} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}. \]
and
\[ M^{(3)} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix}. \]

In this case, we have
\[ \Sigma = \frac{3}{16} M^{(1)} + \frac{4}{16} \text{Re}(\gamma_2) M^{(2)} + \frac{2}{16} \gamma_3 M^{(3)}. \]

Next, note that \( 2M^{(2)} + M^{(3)} = 3M^{(1)} \). Then, if \( \text{Re}(\gamma_2) = \gamma_3 \),
\[ \Sigma = \frac{3}{16} M^{(1)} + \frac{4}{16} \text{Re}(\gamma_2) M^{(2)} + \frac{2}{16} \gamma_3 M^{(3)} \\
= \frac{3}{16} M^{(1)} + \frac{2}{16} (2\text{Re}(\gamma_2) M^{(1)}) \\
= \frac{3}{16} (1 + 2\text{Re}(\gamma_2)) M^{(1)}, \tag{4.16} \]
so that there is still a rescaled version of the iid case in a non-iid cyclic setting. Indeed, since we know that \( \lambda_2 = a_1 + ia_2 - a_3 - ia_4 = (a_1 - a_3) + i(a_2 - a_4) \) and \( \lambda_3 = a_1 - a_2 + a_3 - a_4 \), we find that
\[ \text{Re}(\gamma_2) = \frac{1 - a_2 - 2a_3 - a_4}{(a_2 + 2a_3 + a_4)^2 + (a_2 - a_4)^2} - 1, \]
and \( \gamma_3 = 1/(2(a_2 + a_4)) - 1 \). A short calculation then shows that \( \text{Re}(\gamma_2) = \gamma_3 \) if and only if \( a_3^2 = a_2a_4 \). We thus have a complete characterization of all 4-letter, cyclic Markov chains whose Young diagrams have the same limiting shape as the uniform iid case. In particular, choosing \( a_2 = a_4 = a \), for some \( 0 < a < 1/3 \), leads to \( a_3 = a \) and \( a_1 = 1 - 3a \). If, moreover, \( a = 1/4 \), we have again the iid uniform case. For \( a \neq 1/4 \), however, we may view the Markov chain as a “lazy” version of the uniform iid case.

Note that the scaling factor in both (4.15) and (4.16) is \( 1 + 2\text{Re}(\gamma_2) \). The following theorem shows that, in fact, such a scaling factor occurs for general \( m \), and gives a spectral characterization of cyclic transition matrices which lead to an iid limiting shape.

**Theorem 4.2** Let \( P \) be the \( m \times m \) transition matrix of an aperiodic, irreducible, cyclic Markov chain on an \( m \)-letter, ordered alphabet given in Proposition 4.1. Then the asymptotic covariance matrix \( \Sigma \) is a rescaled version of the iid uniform covariance matrix \( \Sigma_{\text{iidu}} := ((m - 1)/m^2)M^{(1)} \) if and only if
\[ \text{Re} \left( \frac{\lambda_j}{1 - \lambda_j} \right) = \gamma, \quad \text{for all } 2 \leq j \leq m, \tag{4.17} \]
for some real constant \( \gamma \). Moreover, the scaling is then given by
\[ \Sigma = (1 + 2\gamma)\Sigma_{\text{iidu}}. \tag{4.18} \]
Proof. We first claim that the system of matrix equations

\[ \sum_{j=2}^{m} b_j M^{(j)} = M^{(1)} \]  

(4.19)

has a unique solution \( b_j = 1/(m-1) \), for all \( 2 \leq j \leq m \). Indeed, revisiting (4.14), we can express each \( M^{(j)} \) as

\[ M^{(j)} = \tilde{M}^{(j)} + \tilde{M}^{(-j)} = \tilde{M}^{(j)} + \tilde{M}^{(m-j+1)}; \]

(4.20)

where \( (\tilde{M}^{(j)})_{k,\ell} = \omega^{(j-1)(\ell-k)/2} \), for all \( 1 \leq k, \ell \leq m \), so that (4.19) becomes

\[ M^{(1)} = \sum_{j=2}^{m} b_j \left( \tilde{M}^{(j)} + \tilde{M}^{(m-j+1)} \right) = \sum_{j=2}^{m} (b_j + b_{m-j+1}) \tilde{M}^{(j)} = \sum_{j=2}^{m} \tilde{b}_j \tilde{M}^{(j)}, \]

(4.21)

where \( \tilde{b}_j := (b_j + b_{m-j+1})/2 \), for \( 2 \leq j \leq m \).

Now, clearly, each \( M^{(j)} \) is cyclic, so that in solving (4.21) we need only examine the \( m \) entries in the first rows of the matrices. We can thus reduce (4.21) to the \( m \times (m-1) \) system of equations

\[
\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
\omega & \omega^2 & \omega^3 & \cdots & \omega^{m-1} \\
\omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{2(m-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\omega^{m-1} & \omega^{2(m-1)} & \omega^{3(m-1)} & \cdots & \omega^{(m-1)^2}
\end{pmatrix}
\begin{pmatrix}
\tilde{b}_2 \\
\tilde{b}_3 \\
\vdots \\
\tilde{b}_m
\end{pmatrix}
= \begin{pmatrix}
1 \\
\frac{-1}{m-1} \\
\vdots \\
\frac{-1}{m-1}
\end{pmatrix}.
\]

(4.22)

Since each of the last \( m-1 \) rows of the matrix in (4.22) sums to \(-1\), it is clear that \( \tilde{b}_j = 1/(m-1) \) is a solution to the system. To see that this solution is, in fact, unique, consider the \((m-1) \times (m-1)\) sub-matrix consisting of the last \( m-1 \) rows of the matrix in (4.22), namely,

\[
\begin{pmatrix}
\omega & \omega^2 & \omega^3 & \cdots & \omega^{m-1} \\
\omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{2(m-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\omega^{m-1} & \omega^{2(m-1)} & \omega^{3(m-1)} & \cdots & \omega^{(m-1)^2}
\end{pmatrix}.
\]

(4.23)

Now this matrix, which is very closely related to the Fourier matrix which arises in discrete Fourier transform problems, is in fact invertible, and can be shown to have one eigenvalue of \(-1\), and \( m-2 \) eigenvalues of the form \( \pm \sqrt{m} \) and \( \pm i \sqrt{m} \), so that the modulus of the determinant is \( m^{(m-2)/2} \neq 0 \). Thus, the solution \( \tilde{b}_j = 1/(m-1) \) is unique, and since \( b_j = (b_j + b_{m-j+1})/2 = b_{m-j+1} \), for all \( 2 \leq j \leq m \), we conclude that \( b_j = 1/(m-1) \) as well, for all \( 2 \leq j \leq m \), and the claim is proved.

We can now use Proposition 4.1 to simplify the asymptotic covariance matrix decomposition as follows:

\[ \Sigma = \frac{m-1}{m^2} M^{(1)} + \frac{2}{m^2} \sum_{k=2}^{m} \gamma_k M^{(k)} \]

33
where \( \gamma = \text{Re}(\gamma_j) \), for all \( 2 \leq j \leq m \). If the real parts of \( \gamma_j \) are not all identical, then the uniqueness of the solution of (4.19) implies that no such simplification is possible, and the theorem is proved. ■

Remark 4.3 To see that the condition in (4.17) is not vacuous for any \( m \), recall that for \( m = 4 \), the “lazy” chain has the iid limiting shape. This is true for general \( m \): if \( a_2 = a_3 = \cdots = a_m = a \), for some \( 0 < a < 1/(m-1) \), then \( \lambda_j = 1-(m-1)a \), for all \( 2 \leq j \leq m \). Trivially, then, \( \gamma_j = 1/((m-1)a) - 1 := \gamma \), for all \( 2 \leq j \leq m \), so that the conditions of Theorem 4.2 are satisfied, and the scaling factor is given by \( 1 + 2\gamma = (2 - (m-1)a)/((m-1)a) \). Even in the \( m = 4 \) case, however, we saw that there were other, more general, cyclic transition matrices which gave rise to the iid limiting distribution.

The previous proposition indicates precisely when we may expect the limiting shape of a cyclic Markov chain to be spectrum of the traceless GUE. Now the first-order behavior of all rows of the Young diagrams is \( n/m + O(\sqrt{n}) \) for cyclic Markov chains. Although this differs from the first-order behavior in the non-uniform iid case, one may still ask whether the limiting shape for some cyclic Markov chains might still be that of some non-uniform iid case. In fact, this can never occur: cyclicity ensures that the asymptotic covariance matrix is also cyclic, and thus cannot be equal to the asymptotic covariance matrix of any non-uniform iid case.

Another class of Markov chains whose asymptotic covariance matrices can be easily studied is the class of reversible Markov chains, i.e., those with transition matrices such that \( \pi_i P_{i,j} = \pi_j P_{j,i} \), for all \( 1 \leq i, j \leq m \). The following theorem describes the asymptotic covariance matrix of such Markov chains and, in the doubly stochastic case, gives necessary and sufficient conditions for recovering a rescaled uniform iid asymptotic covariance matrix.

Theorem 4.3 Let \( P \) be the transition matrix of an aperiodic, irreducible, reversible Markov chain on an \( m \)-letter, ordered alphabet \( A_m = \{\alpha_1 < \alpha_2 < \cdots < \alpha_m\} \). Then the asymptotic covariance matrix \( \Sigma \) is given by

\[
\Sigma = \Pi^{1/2} S^T (I + 2D) S \Pi^{1/2}, \tag{4.25}
\]

where \( P = (\Pi^{1/2})^{-1} A (\Pi^{1/2}) \) is the diagonalization of \( P \). If, moreover, \( P \) is doubly stochastic, i.e., if \( P \) is symmetric, then \( \Sigma \) is a rescaled version of the iid uniform case with transition matrix \( P_{\text{uid}} \) if and only if \( P = \alpha P_{\text{uid}} + (1 - \alpha) I \), for \( 0 < \alpha \leq m/(m-1) \).

Proof. It is elementary that \( P \) is similar to a symmetric matrix \( Q \). Specifically, \( P = \Pi^{-1/2} Q \Pi^{1/2} \), where \( \Pi \) is the diagonal matrix containing the stationary distribution. Since \( Q \) is symmetric, it is
diagonalizable with orthogonal eigenvectors. Writing \( Q = S^T \Lambda S \), we have \( P = (S \Pi^{1/2})^{-1} \Lambda (S \Pi^{1/2}) \). But then

\[
\Sigma = \Pi + \Pi (S \Pi^{1/2})^{-1} D (S \Pi^{1/2}) + (S \Pi^{1/2})^{-1} D (S \Pi^{1/2})^T \Pi \\
= \Pi + 2 \Pi^{1/2} S^T D S \Pi^{1/2} \\
= \Pi^{1/2} (I + 2 S^T D S) \Pi^{1/2} \\
= \Pi^{1/2} S^T (I + 2D) S \Pi^{1/2}.
\]

(4.26)

where \( D = \text{diag}(1/2, \lambda_2/(1 - \lambda_2), \ldots, \lambda_m/(1 - \lambda_m)) \). Writing \( \Lambda_{\Sigma} := (I + 2D) \), we find that

\[
\Lambda_{\Sigma} = \text{diag} \left( 0, \frac{1 + \lambda_2}{1 - \lambda_2}, \ldots, \frac{1 + \lambda_m}{1 - \lambda_m} \right).
\]

In the doubly stochastic case, (4.26) constitutes the diagonalization of \( \Sigma \), since \( \Pi = (1/m)I \). If \( \Sigma = a \Sigma_{\text{uiid}} \), for some \( a > 0 \), then the spectrum of \( \Sigma \) must consist of 0 and \( m - 1 \) identical positive eigenvalues. Therefore, \( \lambda_2 = \lambda_3 = \cdots = \lambda_m = \lambda \), so that \( \Lambda = \text{diag}(1, \lambda, \lambda, \ldots, \lambda) \). Thus \( \Lambda_{\Sigma} = (1 + \lambda)(I - \Lambda)/(1 - \lambda)^2 \), so that (4.26) now becomes

\[
\Sigma = \frac{1}{m} S^T \Lambda_{\Sigma} S = \frac{1 + \lambda}{m(1 - \lambda)^2} S^T (I - \Lambda) S = \frac{1 + \lambda}{m(1 - \lambda)^2} (I - P).
\]

(4.27)

In particular, \( \Sigma_{\text{uiid}} = (1/m)(I - P_{\text{uiid}}) \). Since \( \Sigma = a \Sigma_{\text{uiid}} \), for some \( a > 0 \), it follows that

\[
P = a \frac{(1 - \lambda)^2}{1 + \lambda} P_{\text{uiid}} + \left( 1 - a \frac{(1 - \lambda)^2}{1 + \lambda} \right) I.
\]

(4.28)

After setting \( \alpha = a(1 - \lambda)^2/(1 + \lambda) \), and checking that \( P \) is irreducible for \( 0 < \alpha \leq m/(m - 1) \), the theorem is proved.

\[\blacksquare\]

**Remark 4.4** Recall that in the 4-letter cyclic case, where \( P \) had initial column entries \( a_1, a_2, a_3, \) and \( a_4 \), the criterion for obtaining a rescaling of the covariance matrix in the iid case was that \( a_2^2 = a_2 a_4 \). If we further demand that \( P \) be symmetric, then \( a_2 = a_4 \), and the criterion if refined to \( a_2 = a_3 = a_4 = a \), which is consistent with Theorem 4.3 as \( P = 4a P_{\text{uiid}} + (1 - 4a) I \), for \( 0 < 4a \leq 4/3 \).

We now seek to relate iid non-uniform limiting shapes (which are spectra of direct sums of GUEs) to those of general Markov chains having the same stationary distribution. The following interpolation result describes the asymptotic covariance matrix for a one-parameter convex class of Markov chains, the covariance matrix being given as a linear combination of the "base" covariance matrix and that of a closely-related iid case.

**Proposition 4.2** For any \( m \geq 3 \), let \( P_0 \) be the \( m \times m \) transition matrix of an irreducible, aperiodic, Markov chain, and let its associated asymptotic covariance matrix be given by

\[
\Sigma_0 = \Pi_0 + \Pi_0 (S_0^{-1} D_0 S_0) + (S_0^{-1} D_0 S_0)^T \Pi_0.
\]

(4.29)
in the standard notations of Theorem 2.1. Then, for $0 < \delta \leq 1$, the transition matrix $P = (1 - \delta)I_m + \delta P_0$ has an asymptotic covariance matrix given by

$$
\Sigma = \frac{1}{\delta} (\Sigma_0 + (1 - \delta)\Sigma_{\Pi_0}),
$$

(4.30)

where $\Sigma_{\Pi_0}$ is the covariance matrix associated with the iid Markov chain having the same stationary distribution as $P_0$.

**Proof.** Using the standard notations of Theorem 2.1 we will write

$$
\Sigma = \Pi + \Pi(S^{-1}DS) + (S^{-1}DS)^T\Pi
$$
in terms of the decomposition $\Sigma_0$ in (4.29). Now, clearly, the stationary distribution under $P$ is that of $P_0$, so that $\Pi = \Pi_0$. We will thus write the stationary distribution simply as $(\pi_1, \pi_2, \ldots, \pi_m)$. Moreover, the eigenvectors are also unchanged, so that $S = S_0$. However, for each eigenvalue $\lambda_{k,0}$ of $P_0$, we have that $\lambda_k = (1 - \delta) + \delta\lambda_{k,0}$ is an eigenvalue of $P$, for $1 \leq k \leq m$. Thus, for each $2 \leq k \leq m$, the diagonal entries of $D$ are given by

$$
\gamma_k := \frac{\lambda_k}{1 - \lambda_k} = \frac{(1 - \delta) + \delta\lambda_{k,0}}{\delta(1 - \lambda_{k,0})} = \frac{1 - \delta}{\delta} + \gamma_{k,0},
$$

where $\gamma_{k,0}$ are the diagonal entries of $D_0$. We can thus decompose $D$ as follows:

$$
D = \text{diag}(-1/2, \gamma_2, \ldots, \gamma_m)
= \text{diag}(-1/2, 0, \ldots, 0) + \left(\frac{1 - \delta}{\delta}\right) \text{diag}(0, 1, \ldots, 1) + \left(\frac{1}{\delta}\right) \text{diag}(0, \gamma_2, \ldots, \gamma_m, 0)
= \text{diag} \left(-\left(\frac{1 - \delta}{2\delta}\right), 0, \ldots, 0\right) + \left(\frac{1 - \delta}{\delta}\right) I_m + \left(\frac{1}{\delta}\right) D_0.
$$

(4.31)

Next, recall from Proposition 2.1 that the first column of $S^{-1}$ is $(1, 1, \ldots, 1)^T$. Hence,

$$
S^{-1}DS = S_0^{-1}DS_0
= \left(\begin{array}{ccc}
1 & \ast & \ldots & \ast \\
\vdots & \vdots & \ddots & \vdots \\
1 & \ast & \ldots & \ast
\end{array}\right) \left(\begin{array}{cccc}
\frac{-1 - \delta}{2\delta} & 0 & \ldots & 0 \\
0 & \ddots & \vdots & \vdots \\
0 & \ldots & 0 & 0
\end{array}\right) \left(\begin{array}{cccc}
\pi_1 & \pi_2 & \cdots & \pi_m \\
\vdots & \vdots & \ddots & \vdots \\
\pi_1 & \pi_2 & \cdots & \pi_m
\end{array}\right)
\left(\begin{array}{cccc}
\pi_1 & \pi_2 & \cdots & \pi_m \\
\vdots & \vdots & \ddots & \vdots \\
\pi_1 & \pi_2 & \cdots & \pi_m
\end{array}\right)
= -\left(\frac{1 - \delta}{2\delta}\right) S_0^{-1}I_m S_0 + \left(\frac{1 - \delta}{\delta}\right) S_0^{-1}D_0 S_0
\left(\begin{array}{cccc}
1 & \ast & \ldots & \ast \\
\vdots & \vdots & \ddots & \vdots \\
1 & \ast & \ldots & \ast
\end{array}\right)
\left(\begin{array}{cccc}
\pi_1 & \pi_2 & \cdots & \pi_m \\
\vdots & \vdots & \ddots & \vdots \\
\pi_1 & \pi_2 & \cdots & \pi_m
\end{array}\right)
+ \left(\frac{1 - \delta}{\delta}\right) I_m + \left(\frac{1}{\delta}\right) S_0^{-1}D_0 S_0.
$$

(4.32)
which gives us

\[
\Pi S^{-1}DS = \Pi_0 S^{-1}DS
\]

\[
= -\left(\frac{1-\delta}{2\delta}\right)\begin{pmatrix}
\pi_1 & 0 & \cdots & 0 \\
0 & \pi_2 & \cdots & \vdots \\
0 & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & \pi_m
\end{pmatrix}
\begin{pmatrix}
\pi_1 & \pi_2 & \cdots & \pi_m \\
\pi_1 & \pi_2 & \cdots & \pi_m \\
\vdots & \vdots & \ddots & \vdots \\
\pi_1 & \pi_2 & \cdots & \pi_m
\end{pmatrix}
\]

\[
+ \left(\frac{1-\delta}{\delta}\right)\Pi_0 + \left(\frac{1}{\delta}\right)\Pi_0 S_0^{-1}D_0 S_0
\]

\[
= -\left(\frac{1-\delta}{2\delta}\right)\begin{pmatrix}
\pi_1^2 & \pi_1 \pi_2 & \cdots & \pi_1 \pi_m \\
\pi_2 \pi_1 & \pi_2^2 & \cdots & \pi_2 \pi_m \\
\vdots & \vdots & \ddots & \vdots \\
\pi_m \pi_1 & \pi_m \pi_2 & \cdots & \pi_m^2
\end{pmatrix}
\]

\[
+ \left(\frac{1-\delta}{\delta}\right)\Pi_0 + \left(\frac{1}{\delta}\right)\Pi_0 S_0^{-1}D_0 S_0.
\] (4.33)

Finally, we can express \(\Sigma\) as

\[
\Sigma = \Pi + \Pi(S^{-1}DS) + (S^{-1}DS)^T \Pi
\]

\[
= \left(\frac{1}{\delta}\right)\Pi_0 + \left(1 - \frac{1}{\delta}\right)\Pi_0 + \Pi_0 (S^{-1}DS) + (\Pi_0 (S^{-1}DS))^T
\]

\[
= \left(\frac{1}{\delta}\right)\Sigma_0 + \left(1 - \frac{1}{\delta}\right)\Pi_0 - 2\Pi_0
\]

\[
+ \left(1 - \frac{1}{\delta}\right)\begin{pmatrix}
\pi_1^2 & \pi_1 \pi_2 & \cdots & \pi_1 \pi_m \\
\pi_2 \pi_1 & \pi_2^2 & \cdots & \pi_2 \pi_m \\
\vdots & \vdots & \ddots & \vdots \\
\pi_m \pi_1 & \pi_m \pi_2 & \cdots & \pi_m^2
\end{pmatrix}
\]

\[
= \left(\frac{1}{\delta}\right)\Sigma_0 + \left(1 - \frac{1}{\delta}\right)(-\Sigma_0 \Pi_0)
\]

\[
= \frac{1}{\delta} (\Sigma_0 + (1 - \delta)\Sigma_0),
\] (4.34)

and we are done.

Thus far we have expressed our limiting laws in terms of Brownian functionals whose Brownian motions have a non-trivial covariance structure arising directly from the specific nature of the transition matrix. It is of interest to instead express the limiting laws in terms of standard Brownian motions.

Since the asymptotic covariance matrix \(\Sigma\) is non-negative definite, we can find an \(m \times m\) matrix \(C\) such that \(\Sigma = CC^T\). Clearly, we then have

37
Theorem 3.2 as

Unfortunately, the maximal term does not in general succumb to any significant simplifications. How-
so that

\[ C_k = \sum_{r=0}^{m} \nu_r \]

In order to simplify notation, we will assume that \( \tau(\ell) = \ell \), for all \( \ell \), and so write our main result in Theorem 3.2 as

\[
\frac{V_r \sqrt{\pi_i} \nu_r}{n} = \sum_{k=1}^{m} \sigma_k B^k(1) + \max_{I_{r-m_r, \ell}} \sum_{j=1}^{m} \sum_{\ell=1}^{r-m_r} (d_r+m_r-r+j) \sigma_{m_r+\ell} B^{m_r+\ell}(\Delta t_{j,\ell}) := V_{r,\ell}.
\] (4.36)

Simply substituting (4.35) into (4.36) immediately yields

\[
V_r = \sum_{k=1}^{m} \left( \sum_{i=1}^{m} C_{k,i} B^i(1) \right) + \max_{I_{r-m_r, \ell}} \sum_{j=1}^{m} \sum_{\ell=1}^{r-m_r} (d_r+m_r-r+j) \left( \sum_{i=1}^{m} C_{m_r+\ell,i} B^i(\Delta t_{j,\ell}) \right)
\] (4.37)

Now the first term in (4.37) is simply a Gaussian term whose variance can be computed explicitly. Unfortu-

Unfortunately, the maximal term does not in general succumb to any significant simplifications. How-

\[
V_{r,\ell} = \sum_{i=1}^{m} \sqrt{\pi_i} B^i(1) - \sum_{k=1}^{m} \nu_m r - \sum_{k=1}^{m} (\sqrt{\pi_i} \pi_k) B^i(1)
\]
\[
\sum_{i=1}^{m_r} \sqrt{\pi_i} B_i^j(1) - \nu_r \sum_{i=1}^{m_r} \sqrt{\pi_i} B_i^j(1) \right) + \sqrt{\pi_r} \max_{I_{r-m_r,d_r}} \sum_{j=1}^{r-m_r} \sum_{\ell=1}^{d_r} B_{m_r+\ell}(\Delta t_{j,\ell}) \\
\left. + \right. + \sqrt{\pi_r} \right\} \\
(4.38)
\]

Note that the first two Gaussian term of (4.38) are independent of the remaining two Gaussian-maximal expression terms.

Following Glynn and Whitt [13] and Barishnykov [4], who studied the Brownian functional

\[
D_m = \max_{I_m} \sum_{\ell=1}^{m} B_\ell^j(\Delta t_j),
\]

we define the following, more general, Brownian functional:

\[
D_{r,m} := \max_{I_{r,m}} \sum_{j=1}^{r} \sum_{\ell=1}^{(m-r+j)} B^j_\ell(\Delta t_{j,\ell}),
\]

(4.39)

where \(1 \leq r \leq m\). Clearly, the maximal term in (4.38) has just such a form. We also remark that \(D_{r,m}\) corresponds to the sum of the \(r\) largest eigenvalues of an \(m \times m\) GUE matrix.

To better understand (4.38), we may, without much loss in generality, focus on the first block, that is, values of \(r\) such that \(m_r = 0\). The first Gaussian term of (4.38) thus vanishes, and, writing \(\pi_{\max}\) for \(\pi_r\), we have

\[
V_r^\infty = -r \pi_{\max} \sum_{i=d_1+1}^{m} \sqrt{\pi_i} B_i^j(1) + \sqrt{\pi_{\max}} \left( -r \pi_{\max} \sum_{i=1}^{d_1} B_i^j(1) + D_{r,d_r} \right). \tag{4.40}
\]

In the uniform iid case, the first Gaussian term of (4.40) itself vanishes, since \(d_r = d_1 = m\), and we have

\[
V_r^\infty = \frac{1}{\sqrt{m}} \left( -\frac{r}{m} \sum_{i=1}^{m} B_i^j(1) + D_{r,m} \right) := H_{r,m}. \tag{4.41}
\]

Furthermore, and still specializing (4.40) to \(r = 1\),

\[
\frac{LI_n - n \pi_{\max}}{\sqrt{n}} \Rightarrow -\pi_{\max} \sum_{i=d_1+1}^{m} \sqrt{\pi_i} B_i^j(1) + \sqrt{\pi_{\max}} \left( -\pi_{\max} \sum_{i=1}^{d_1} B_i^j(1) + D_{1,d_1} \right) \\
= -\pi_{\max} \sum_{i=d_1+1}^{m} \sqrt{\pi_i} B_i^j(1) + \sqrt{\pi_{\max}} \left( \frac{1}{d_1} - \pi_{\max} \right) \sum_{i=1}^{d_1} B_i^j(1) + \sqrt{\pi_{\max}} H_{1,d_1}. \tag{4.42}
\]
One can easily compute the variance of the Gaussian terms in (4.41) to be \( \pi_{\text{max}}(1/d_1 - \pi_{\text{max}}) \). For \( d_1 = 1 \), this becomes \( \pi_{\text{max}}(1 - \pi_{\text{max}}) \), and since \( H_{1,1} = 0 \) a.s., the limiting distribution is \( N(0, \pi_{\text{max}}(1 - \pi_{\text{max}})) \). For \( d_1 = m \), the variance of the Gaussian term vanishes, and we again recover (4.41), namely, \( H_{1,m}/\sqrt{m} \). Both of these results are consistent with those of the authors’ previous paper [17].

5 Connections to Random Matrix Theory

For iid uniform \( m \)-letter alphabets, the limiting law of the shape of the Young diagrams corresponds to the joint distribution of the eigenvalues of an \( m \times m \) matrix from the traceless GUE [22]. In the non-uniform iid case, we further noted that Its, Tracy, and Widom [20, 21] have described the limiting law of the length of the longest increasing subsequence as that of the maximal eigenvalue of a random matrix consisting of independent diagonal blocks, each of which is a matrix from the GUE. The size of each block depends upon the multiplicity of the corresponding stationary probability. In addition, there is a zero-trace condition involving the stationary probabilities on the composite matrix (see [19] for the RSK diagrams result).

As a first step in extending these connections between Brownian functionals and spectra of random matrices, recall the general case when the stationary probabilities are all distinct (see Remark 3.5). Our Brownian functionals then have no true maximal terms, so that the limiting shape, \( (R_1^2, R_2^2, \ldots, R_\infty^2) \) is simply multivariate normal, with covariance matrix \( \Sigma \) (or, more precisely, the matrix obtained by permuting the rows and columns of \( \Sigma \) using \( \tau \), the permutation of \( \{1, 2, \ldots, m\} \) previously defined). Trivially, this limiting law corresponds to the spectrum of a diagonal matrix whose elements are multivariate normal with the same covariance matrix \( \Sigma \).

We can see that this general result is consistent with the iid non-uniform case having distinct probabilities. Indeed, each block is of size 1, and is rescaled so that the variance is \( \pi_\tau(i)(1 - \pi_\tau(i)) \), for \( 1 \leq i \leq m \). Because of this rescaling, instead of having a generalized zero-trace condition, as in the non-rescaled matrices used in [20, 21], our condition is rather a true zero-trace condition. This zero-trace condition is clear, since the covariance matrix for any iid case (uniform and non-uniform alike) is that of a multinomial distribution with parameters \( (n = 1; \pi_\tau(1), \pi_\tau(2), \ldots, \pi_\tau(m)) \), and any \( (Y_1, Y_2, \ldots, Y_m) \) having such a distribution of course satisfies \( \sum_{i=1}^m Y_i = 1 \), so that \( \text{Var}(\sum_{i=1}^m Y_i) = 0 \), which implies the zero-trace condition for \( (R_1^1, R_2^1, \ldots, R_\infty^1) \).

Next, consider the case when each stationary probability has multiplicity no greater than 2. One may conjecture that \( (R_1^1, R_2^1, \ldots, R_\infty^1) \) is equal in law to the spectrum of a direct sum of certain \( 2 \times 2 \) and/or \( 1 \times 1 \) random matrices. Specifically, let \( \kappa \leq m \) be the number of distinct probabilities among the stationary distributions. Then the composite matrix consists of a direct sum of \( \kappa \) GUE matrices which are as follows. First, the overall diagonal \( (X_1, X_2, \ldots, X_m) \) of the matrix has a \( N(0, \Sigma) \) distribution. Next, if \( d_r = 1 \), then the GUE matrix is simply the \( 1 \times 1 \) matrix \( (X_r) \). Finally, if \( d_r = 2 \), then the GUE matrix is the \( 2 \times 2 \) matrix

\[
\begin{pmatrix}
X_{m_r+1} & Y_{m_r+1} + iZ_{m_r+1} \\
Y_{m_r+1} - iZ_{m_r+1} & X_{m_r+2}
\end{pmatrix},
\]

whose off-diagonal random variables \( Y_{m_r+1} \) and \( Z_{m_r+1} \) are iid, centered, normal random variables, independent of all other random variables in the overall matrix, with variance

\[
(\sigma^2_{m_r+1} - 2\rho_{m_r+1,m_r+2}\sigma_{m_r+1}\sigma_{m_r+2} + \sigma^2_{m_r+2})/4.
\]
Such a conjecture would imply the following marginal result regarding a single block of such a matrix, which without loss of generality we take to be the first block. In fact, this result is a genuine extension of the connection between maximal Brownian functionals and random matrices beyond the standard GUE, traceless GUE, and purely diagonal cases already discussed.

**Theorem 5.1** Let $d_1 = 2$ and $\tau(r) = r$, for all $1 \leq r \leq m$. Then $(R_1^1, R_2^1) = (V_1^1, V_2^1 - V_1^1)$ is distributed as the spectrum $\{ (\lambda_1, \lambda_2) : \lambda_1 \geq \lambda_2 \}$ of the $2 \times 2$ Gaussian Hermitian matrix

$$M := \begin{pmatrix} X_1 & Y_1 + iZ_1 \\ Y_1 - iZ_1 & X_2 \end{pmatrix},$$

where $(X_1, X_2)$ is a pair of centered, bivariate normal random variables with covariance matrix

$$\Sigma_2 = \begin{pmatrix} \tilde{\sigma}_1^2 & \tilde{\rho}_1 \tilde{\sigma}_2 \\ \tilde{\rho}_1 \tilde{\sigma}_2 & \tilde{\sigma}_2^2 \end{pmatrix},$$

and $Y_1$ and $Z_1$ are iid, centered normal random variables, independent of $(X_1, X_2)$, with variance

$$(\tilde{\sigma}_1^2 - 2\tilde{\rho}_1 \tilde{\sigma}_2 + \tilde{\sigma}_2^2)/4. \tag{5.3}$$

**Proof.** We prove the equivalent result that $(V_1^1, V_2^1)$ is distributed as the pair $(\lambda_1, \lambda_1 + \lambda_2)$. Now, by definition,

$$(V_1^1, V_2^1) = \left( \max_{0 \leq t \leq 1} (\tilde{\sigma}_1 \tilde{B}^1(t) + \tilde{\sigma}_2 \tilde{B}^2(t)), \tilde{\sigma}_1 \tilde{B}^1(1) + \tilde{\sigma}_1 \tilde{B}^2(1) \right)$$

$$= \left( \tilde{\sigma}_2 \tilde{B}^2(1) + \max_{0 \leq t \leq 1} (\tilde{\sigma}_1 \tilde{B}^1(t) - \tilde{\sigma}_2 \tilde{B}^2(t)), \tilde{\sigma}_1 \tilde{B}^1(1) + \tilde{\sigma}_1 \tilde{B}^2(1) \right). \tag{5.4}$$

We now simplify (5.4), by introducing new Brownian motions and then decomposing the resulting expression into two independent parts. To do so, begin by defining the new variances and correlation coefficients $\sigma_1^2 := \tilde{\sigma}_2^2$, $\sigma_2^2 := \tilde{\sigma}_1^2 - 2\tilde{\rho}_1 \tilde{\sigma}_2 + \tilde{\sigma}_2^2$, and $\rho := (\tilde{\rho}_1 - \tilde{\sigma}_2)/\sqrt{\tilde{\sigma}_1^2 - 2\rho \tilde{\sigma}_1 \tilde{\sigma}_2 + \tilde{\sigma}_2^2}$. Then it is easily verified that $B_1^1(t) := \tilde{B}_2^2(t)$, and $B_2^2(t) := (\tilde{\sigma}_1 \tilde{B}_1^1(t) - \tilde{\sigma}_2 \tilde{B}_2^2(t))/\sigma_2$, where the equalities are pointwise, are (dependent) standard Brownian motions, and (5.4) becomes

$$(V_1^1, V_2^1) = \left( \sigma_1 B_1^1(1) + \sigma_2 \max_{0 \leq t \leq 1} B_2^2(t), 2\sigma_1 B_1^1(1) + \sigma_2 B_2^2(1) \right)$$

$$= \left( (\sigma_1 B_1^1(1) - \rho \sigma_1 B_2^2(1)) + \sigma_2 \left( \rho \frac{\sigma_1}{\sigma_2} B_2^2(1) + \max_{0 \leq t \leq 1} B_2^2(t) \right), \right.$$  

$$2(\sigma_1 B_1^1(1) - \rho \sigma_1 B_2^2(1)) + (\sigma_2 + 2\rho \sigma_1) B_2^2(1) \right). \tag{5.5}$$

Note that $B_1^1(t) - \rho B_2^2(t)$ is independent of $B_2^2(t)$ and has variance $\sigma_1^2(1 - \rho^2)t$. Introducing the Brownian functional

$$U(\beta) = \left( \beta - \frac{1}{2} \right) B_2^2(1) + \max_{0 \leq t \leq 1} B_2^2(t), \tag{5.6}$$

$\beta \in \mathbb{R}$, and using $\sigma_1^2, \sigma_2^2$, and $\rho$ above, (5.5) becomes

41
\( V_1, V_2^\infty \) \( \overset{L}{=} \sigma_1 \sqrt{1 - \rho^2} Z(1, 2) + \left( \sigma_2 \left( \frac{1}{2} - \rho \frac{\sigma_1}{\sigma_2} \right), (\sigma_2 + 2 \rho \sigma_1) B^2(1) \right) \)

\[
= \frac{\tilde{\sigma}_1 \tilde{\sigma}_2 \sqrt{1 - \rho^2}}{\sqrt{\tilde{\sigma}_1^2 - 2 \rho \tilde{\sigma}_1 \tilde{\sigma}_2 + \tilde{\sigma}_2^2}} Z(1, 2)
+ \left( \sqrt{\tilde{\sigma}_1^2 - 2 \tilde{\sigma}_1 \tilde{\sigma}_2 + \tilde{\sigma}_2^2} \right) \left. \frac{\tilde{\sigma}_1^2 - \tilde{\sigma}_2^2}{2 \sqrt{\tilde{\sigma}_1^2 - 2 \rho \tilde{\sigma}_1 \tilde{\sigma}_2 + \tilde{\sigma}_2^2}} U \right), \tag{5.7}
\]

where \( Z \) is a standard normal random variable independent of the sigma-field generated by \( B^2 \).

Turning now to the eigenvalues’ distributions, we first consider the centered, multivariate normal random variables \( (W_1, W_2) \), having covariance matrix

\[
\begin{pmatrix}
\sigma_1^2 & \rho \sigma_1 \sigma_2 \\
\rho \sigma_1 \sigma_2 & \sigma_2^2
\end{pmatrix},
\]

and let \( W_3 \) and \( W_4 \) be two iid, centered, normal random variables, independent of \( (W_1, W_2) \), with variance \( \sigma_2^2 \). Then it is classical that

\[
\left( W_2, \sqrt{W_3^2 + W_4^2} \right) \overset{L}{=} \sigma_2 (B(1), 2 \max_{0 \leq t \leq 1} B(t) - B(1)),
\]

or, equivalently,

\[
\left( W_2, \beta W_2 + \frac{1}{2} \sqrt{W_3^2 + W_4^2} \right) \overset{L}{=} \sigma_2 (B(1), U(\beta)), \tag{5.8}
\]

where \( B \) is a standard Brownian motion, and \( U(\beta), \beta \in \mathbb{R} \), is defined in terms of \( B \), rather than in terms of \( B^2 \), as in (5.6). Then consider the random variable

\[
\tilde{\lambda} := W_1 + \sqrt{W_2^2 + W_3^2 + W_4^2}
= \left( W_1 - \rho \frac{\sigma_1}{\sigma_2} \right) + \left( \rho \frac{\sigma_1}{\sigma_2} + \sqrt{W_2^2 + W_3^2 + W_4^2} \right). \tag{5.9}
\]

Using (5.8), and noting that the variance of the first term in (5.9) is \( \sigma_1^2 (1 - \rho^2) \), it is easy to see that

\[
\tilde{\lambda} \overset{L}{=} \sigma_1 \sqrt{1 - \rho^2} Z + 2 \rho \sigma_1 U \left( \frac{\rho \sigma_1}{2 \sigma_2} \right), \tag{5.10}
\]

where \( Z \) is a standard normal random variable independent of \( B \).

We now apply this result to the eigenvalues of the matrix \( M \) in (5.1), namely, to

\[
\lambda_1 = \left( \frac{X_1 + X_2}{2} \right) + \sqrt{\left( \frac{X_1 - X_2}{2} \right)^2 + Y_1^2 + Z_1^2}, \tag{5.11}
\]

and

\[
\lambda_2 = \left( \frac{X_1 + X_2}{2} \right) - \sqrt{\left( \frac{X_1 - X_2}{2} \right)^2 + Y_1^2 + Z_1^2}. \tag{5.12}
\]
Letting \( W_1 = (X_1 + X_2)/2, W_2 = (X_1 - X_2)/2, W_3 = Y_1, \) and \( W_4 = Z_1, \) we have

\[
(\lambda_1, \lambda_1 + \lambda_2) = \left( W_1 + \sqrt{W_2^2 + W_3^2 + W_4^2}, 2W_1 \right)
\]

\[
= \left( \left( W_1 - \rho \frac{\sigma_1}{\sigma_2} W_2 \right) + 2 \left( \rho \frac{\sigma_1}{\sigma_2} W_2 + \frac{1}{2} \sqrt{W_2^2 + W_3^2 + W_4^2} \right), \right.
\]

\[
2 \left( W_1 - \rho \frac{\sigma_1}{\sigma_2} W_2 \right) + 2 \rho \frac{\sigma_1}{\sigma_2} W_2 \right),
\]

\[
(5.13)
\]

where \( \sigma_1^2 = (\sigma_1^2 + 2 \rho \sigma_1 \sigma_2 + \sigma_2^2)/4, \) \( \sigma_2^2 = (\sigma_1^2 - 2 \rho \sigma_1 \sigma_2 + \sigma_2^2)/4, \) and \( \rho \sigma_1^2 \sigma_2 = (\sigma_1^2 - \sigma_2^2)/4. \) Noting that the variance of \( W_1 - (\rho \sigma_1/\sigma_2) W_2 \) is \( \sigma_1^2 (1 - \rho^2) = \sigma_1^2 (1 - \rho^2), \) and that, moreover, \( \beta := \rho \sigma_1 / 2 \sigma_2 = (\sigma_1^2 - \sigma_2^2)/(2 \sqrt{\sigma_1^2 - 2 \rho \sigma_1 \sigma_2 + \sigma_2^2}), \) we find that

\[
(\lambda_1, \lambda_1 + \lambda_2) = \sigma_1 \sqrt{1 - \rho^2} Z(1, 2) + \left( 2 \sigma_2 \sigma \left( \frac{\rho \sigma_1}{2 \sigma_2} \right), 2 \rho \frac{\sigma_1}{\sigma_2} B^2(1) \right)
\]

\[
= \sigma_1 \sqrt{1 - \rho^2} Z(1, 2) + \sigma_2 (U(\beta), 4 \beta B^2(1))
\]

\[
\overset{\xi}{=} (V_{1, \infty}^1, V_{\infty}^2),
\]

and we have our identity in law.

The preceding theorem marks a first step in describing the increasingly robust connection between Gaussian random matrices and maximal functionals of several Brownian motions. We next turn to another very striking relationship between a different class of 2 \( \times \) 2 Gaussian random matrices and maximal functionals of several Brownian motions. We next turn to another very striking relationship between a different class of 2 \( \times \) 2 Gaussian random matrices and maximal functionals of several Brownian motions.

By a standard Brownian bridge terminating at \( b \) we mean a process \((\hat{B}(t; b))_{0 \leq t \leq 1}\) equal in law to \((B(t) - tB(1) + bt)_{0 \leq t \leq 1}\), where \( B \) is a standard Brownian motion.

**Theorem 5.2** Let \((\hat{B}^1(t; b_1))_{0 \leq t \leq 1}\) and \((\hat{B}^2(t; b_2))_{0 \leq t \leq 1}\) be two independent standard Brownian bridges terminating at \( b_1 \) and \( b_2 \), respectively, and let

\[
M_{bb} := \begin{pmatrix}
   b_1 & Y_1 + iZ_1 \\
   Y_1 - iZ_1 & b_2
\end{pmatrix},
\]

where \( b_1 \) and \( b_2 \) are real constants, while \( Y_1 \) and \( Z_1 \) are independent centered normal random variables with variance 1/2. Then the largest eigenvalue of \( M_{bb} \) has the same law as

\[
V_{bb} := \max_{0 \leq t \leq 1} ((\hat{B}^1(t; b_1) - \hat{B}^1(0; b_1)) + (\hat{B}^2(t; b_2) - \hat{B}^2(t; b_2))).
\]

(5.16)
Proof. We first simplify the familiar-looking expression (5.16) and obtain

\[
V_{bb} := \max_{0 \leq t \leq 1} \left( (\dot{B}^1(t; b_1) - \dot{B}^1(0; b_1)) + (\dot{B}^2(1; b_2) - \dot{B}^2(t; b_2)) \right)
\]

\[
= \max_{0 \leq t \leq 1} \left( \dot{B}^1(t; b_1) + (b_2 - \dot{B}^2(t; b_2)) \right)
\]

\[
= b_2 + \max_{0 \leq t \leq 1} \left( \dot{B}^1(t; b_1) - \dot{B}^2(t; b_2) \right). \tag{5.17}
\]

Since a linear combination of two independent Brownian bridges is again a Brownian bridge, by elementary rescaling we obtain the equality in law

\[
V_{bb} \overset{d}{=} b_2 + \sqrt{2} \max_{0 \leq t \leq 1} \dot{B} \left( t; \frac{b_1 - b_2}{\sqrt{2}} \right), \tag{5.18}
\]

where \( \dot{B} \) is a standard Brownian bridge terminating at \((b_1 - b_2)/\sqrt{2}\).

Next, recall the elementary result (which may be easily obtained from the Reflection Principle) stating that the maximum \( M(b) \) of a standard Brownian bridge terminating at \( b \) has a distribution given by

\[
P(M(b) \geq a) = \exp \left( -\frac{2}{\sqrt{2}} (a - b) \right), \text{ for any } a \geq \max(b, 0).
\]

Applying this result to (5.18), we find that

\[
P(V_{bb} \geq a) = P\left( \max_{0 \leq t \leq 1} \dot{B} \left( t; \frac{b_1 - b_2}{\sqrt{2}} \right) \geq \frac{a - b_2}{\sqrt{2}} \right)
\]

\[
= \exp \left( -2 \left( \frac{a - b_2}{\sqrt{2}} \right) \left( \frac{a - b_2}{\sqrt{2}} - \frac{b_1 - b_2}{\sqrt{2}} \right) \right)
\]

\[
= \exp \left( -(a - b_1)(a - b_2) \right), \tag{5.19}
\]

for all \( a \geq \max(b_1, b_2) \).

To complete the proof of the claim, we simply note that the eigenvalues of \( M_{bb} \) satisfy the characteristic equation

\[
(\lambda - b_1)(\lambda - b_2) - Y_1^2 - Z_1^2 = 0.
\]

Writing \( \lambda_1 \) for the largest eigenvalue of \( M_{bb} \), we see that the event \( \{\lambda_1 \geq a\} \), for \( a \geq \max(b_1, b_2) \), may be rewritten as the event \( \{Y_1^2 + Z_1^2 \geq (a - b_1)(a - b_2)\} \). But \( Y_1^2 + Z_1^2 \) is distributed as a (rescaled) Chi-squared random variable with two degrees of freedom and is in fact exponential with parameter 1. Thus,

\[
P(\lambda_1 \geq a) = P(Y_1^2 + Z_1^2 \geq (a - b_1)(a - b_2))
\]

\[
= \exp \left( -(a - b_1)(a - b_2) \right), \tag{5.20}
\]

and our claim is proved.
Remark 5.1 The significance of Theorem 5.2 lies in the fact that we can de-condition \( b_1 \) and \( b_2 \) to obtain again the connection between maximal Brownian functionals and both the GUE or the traceless GUE, as shown next. Indeed, replace \( b_1 \) and \( b_2 \) by iid standard normal random variables \( X_1 \) and \( X_2 \) which are independent of \((Y_1, Z_1)\) in the random matrix setting. Then, the joint density of \((V_{bb}, X_1, X_2)\) is given by

\[
g_3(v, x_1, x_2) = (2v - x_1 - x_2)\varphi(x_1)\varphi(x_2)\exp(-(v - x_1)(v - x_2)),
\]

for \( x_1, x_2 \in \mathbb{R} \) and \( v \geq \max(x_1, x_2) \), where \( \varphi \) is the standard normal density.

Let us first examine the constant-trace case by letting \( w_1 = x_1 + x_2 \) and \( w_2 = x_1 - x_2 \) and then conditioning the joint law in (5.21) on \( w_1 \). An elementary calculation shows that the conditional density is simply

\[
h(v|w_1) = \frac{4}{\sqrt{\pi}} \left(v - \frac{w_1}{2}\right)^2 \exp\left(-\left(v - \frac{w_1}{2}\right)^2\right),
\]

for \( v \geq w_1/2 \). In particular, for \( w_1 = 0 \), we do indeed recover the traceless case:

\[
h(v|w_1 = 0) = \frac{4}{\sqrt{\pi}} v^2 e^{-v^2},
\]

for \( v \geq 0 \), which is the density associated with the square root of a (rescaled) Chi-squared random variable with three degrees of freedom, as expected.

Next, let us return to the joint density in (5.21) and successively integrate out \( x_1 \) and \( x_2 \). Writing \( \Phi \) for the standard normal cumulative distribution function, we find that the joint density of \((V_{bb}, X_2)\) is given by

\[
g_2(v, x_2) = \varphi(v) \left(\varphi(x_2) + v\Phi(x_2)\right),
\]

for \(-\infty < x_2 \leq v < +\infty\). Integrating once more, we obtain the density of the largest eigenvalue of the \( 2 \times 2 \) GUE,

\[
g_1(v) = \varphi(v) \left((1 + v^2)\Phi(v) + v\varphi(v)\right),
\]

for all \( v \in \mathbb{R} \). We have thus recovered the traceless and standard GUE distributions from our conditional result.

The results above are all ultimately based on the trivial fact that we can solve a quadratic equation. While larger random matrices do not in general yield to such straightforward analyses, there are nonetheless certain classes that can be linked to distributions of maximal Brownian functionals. The following development consolidates several of these and leads to a more general conjecture.

Let \( B \) be \( m \)-dimensional standard Brownian motion and introduce the notation

\[
\text{diag}(A) = \begin{pmatrix} A_{1,1} \\ \vdots \\ A_{m,m} \end{pmatrix},
\]

for any \( m \times m \) matrix \( A \) and
Diag \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} v_1 & 0 \\ & \ddots & \vdots \\ 0 & & v_m \end{pmatrix},

for any vector \(v = (v_1, \ldots, v_m)^T\). Now recall the Brownian functional introduced by Glynn and Whitt [13], namely,

\[ D_m = \max_{k=1}^m B^k(\Delta t_k), \]

where the maximum is taken over the usual Weyl chamber, and from results of Baryshnikov [4] and Gravner, Tracy, and Widom [15], \(D_m \equiv \lambda_{\max}(M)\), for \(M\) an \(m \times m\) element of the GUE. To extend this identity to a broader class of random matrices, first let \(A = aI + \frac{1}{m}b^T\), where \(a \in \mathbb{R}\), \(b \in \mathbb{R}^m\), and where \(1_m = (1, 1, \ldots, 1)^T \in \mathbb{R}^m\). Next, letting \(\tilde{B} = AB\), the corresponding maximal Brownian functional associated with \(\tilde{B}\) is such that

\[ \tilde{D}_m := \max_{k=1}^m \tilde{B}^k(\Delta t_k) \]

\[ = \max_{k=1}^m \left( aB^k(\Delta t_k) + \sum_{j=1}^m b_j B^j(\Delta t_k) \right) \]

\[ = aD_m + \sum_{j=1}^m b_j \sum_{k=1}^m B^j(\Delta t_k) \]

\[ = aD_m + b^T B(1) \]

\[ \equiv a\lambda_{\max}(M) + b^T \text{diag}(M) \]

\[ = \lambda_{\max}(aM + b^T \text{diag}(M)I) \]

\[ = \lambda_{\max}(\tilde{M}) \]

where \(\tilde{M} = aM + b^T \text{diag}(M)I\).

This rank-one perturbation result, when combined with Theorem 3.3 or Corollary 3.1 and taking \(a = 1\) and \(b_i = -1/m\), \(i = 1, \ldots, m\), recovers [28] only using combinatorial and probabilistic techniques and therefore bypassing the analytical ones present there. It also shares, as shown next, its amenability to analysis with two further cases: the cyclic case for \(m = 2\) or \(3\), and also the case \(A = aP\), where \(P\) is an arbitrary permutation matrix and \(a\) an arbitrary real. Write the general \(m \times m\) cyclic matrix \(A\) as \(A = \sum_{k=0}^{m-1} a_k P^k\), where \(P\) is the permutation matrix \(P = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}\). Such a cyclic also
yields a cyclic covariance matrix $\Sigma := AA^T$, and moreover,
\[
\tilde{D}_m := \max \sum_{k=1}^m \sum_{j=1}^m A_{kj} B_j^j (\Delta t_k) = \max \sum_{k=1}^m \sum_{j=1}^m a_{k-j} B_j^j (\Delta t_k) = \max \sum_{r=0}^{m-1} a_m \sum_{k=1}^m B_k^{r+1} (\Delta t_k).
\]

We claim that for all three cases above,
\[
\tilde{D}_m \leq \lambda_{\max} \left( f(\sigma(A)) M_0 + \begin{pmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_m \end{pmatrix} \right),
\]
where $f$ is an as-yet unspecified function of $\sigma(A)$, the spectrum of $A$, $(X_1, \ldots, X_m)^T \sim N(0, \Sigma)$, $\Sigma = AA^T$, and $M_0$ is a “GUE matrix” with all diagonal entries of 0. Writing $\tilde{M} = f(\sigma(A)) M_0 + \begin{pmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_m \end{pmatrix}$, we then have $\tilde{D}_m \leq \lambda_{\max}(\tilde{M})$. Let us first look at the simple case of $A = aP$, where $P$ is an arbitrary permutation matrix with $P_{i,j} = \delta_{i,\sigma(j)}$, $\sigma$ a permutation of $\{1, 2, \ldots, m\}$. Then
\[
\tilde{B} = aPB \leq aB, \text{ and so } \tilde{D}_m = \max \sum_{k=1}^m \begin{pmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_m \end{pmatrix},
\]
where $(X_1, \ldots, X_m)^T \sim N(0, \Sigma)$, $\Sigma = AA^T = a^2 I$. But since $\tilde{M} = aM_0 + a \text{Diag}(\text{diag}(M)) = aM$, $M \in \text{GUE}$, we simply recover the base case $\tilde{D}_m \leq a\lambda_{\max}(M)$ and see that $f(\sigma(A)) = a$.

Next, consider the degenerate cyclic case where $a_0 = a_1 = \cdots = a_{m-1} = a$, so that $A = a_{1,m}1_{m}^T$. But this is just a special case of the rank-one perturbation $A = aI + b b^T$, so that $\tilde{D}_m \leq a_{1,m}^2 B(1) \sim N(0, ma^2)$, and $\tilde{M} = 0M + a_{1,m}^T B(1)^T = a \sum_{k=1}^m X_k I$, so that $\lambda_{\max}(\tilde{M}) = a \sum_{k=1}^m X_k \sim N(0, ma^2)$. In this pure Gaussian case, $f(\sigma(A)) = 0$.

More generally, as previously shown, for $A = aI + b b^T$, $\tilde{D}_m \leq \lambda_{\max}(aM + (b^T 1_m) \text{diag}(M)I)$
\[
= \lambda_{\max}(aM_0 + (a + b^T 1_m) \text{diag}(M)I)
= \lambda_{\max}(aM_0 + A \text{diag}(M)I),
\]
and since $A \text{diag}(M) \sim N(0, \Sigma)$, $f(\sigma(A)) = a$.

Let us now look at the low-dimensional cyclic cases in turn. For the $2 \times 2$ cyclic case, let $A = a_0 I + a_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (a_0 - a_1) I + a_1 1_2 1_2^T$. Then
\[
\tilde{M} = (a_0 - a_1) M + a_1 1_2 1_2^T \text{diag}(M)I
= (a_0 - a_1) M_0 + (a_0 - a_1) \text{Diag}(\text{diag}(M)) + a_1 1_2 1_2^T \text{diag}(M)I
= (a_0 - a_1) M_0 + \text{Diag}((a_0 - a_1) I + a_1 1_2 1_2^T) \text{diag}(M))
= (a_0 - a_1) M_0 + \text{Diag}(A \text{diag}(M)),
\]

47
and here \( A \text{diag}(M) \sim N(0, \Sigma) \) and so \( f(\sigma(A)) = a_0 - a_1 \). Finally, consider the 3-dimensional cyclic case. Let \( A = \begin{pmatrix} a_0 & a_2 & a_1 \\ a_1 & a_0 & a_2 \\ a_2 & a_1 & a_0 \end{pmatrix} \), then \( \Sigma = AA^T = \sum_{k=1}^{3} a_k^2 I + \sum_{k \neq \ell} a_k a_\ell (13_3^T - I) = (\sum_{k=1}^{3} a_k^2 - \sum_{k \neq \ell} a_k a_\ell) I + (\sum_{k \neq \ell} a_k a_\ell) 13_3^T \). But we can also write \( \Sigma = (\alpha I + \beta_1 13_3^T)(\alpha I + \beta_1 13_3^T)^T = \alpha^2 I + (2\alpha \beta + 3\beta^2) 13_3^T \). Thus \( \alpha^2 = \sum_{k=1}^{3} a_k^2 - \sum_{k \neq \ell} a_k a_\ell \), therefore \( 3\beta^2 + 2\alpha \beta = \sum_{k \neq \ell} a_k a_\ell \) and so \( \beta = -\alpha/3 \pm 1/3 \sqrt{\sum_{k=1}^{3} a_k^2 - 3 \sum_{k \neq \ell} a_k a_\ell} \). Now, take \( \alpha = \sqrt{\sum_{k=1}^{3} a_k^2 - \sum_{k \neq \ell} a_k a_\ell} \) and \( \beta = -\alpha/3 + \sqrt{\sum_{k=1}^{3} a_k^2 - 3 \sum_{k \neq \ell} a_k a_\ell}/3 \). But, \( \bar{D}_3 = \lambda_{\max}(\alpha M_0 + \text{Diag}(\alpha I + \beta_1 13_3^T) \text{diag}(M)) \), so that \( \alpha I + \beta_1 13_3^T \text{diag}(M) \sim N(0, \Sigma) \), where \( f(\sigma(A)) = \alpha \). Noting that \( M_0 \) and the matrix \( \begin{pmatrix} X_1 & 0 \\ 0 & \ddots \\ 0 & X_m \end{pmatrix} \) are independent, we can summarize all the above results in the following theorem:

**Theorem 5.3** Let \( B = (B(t))_{0 \leq t \leq 1} \) be a standard \( m \)-dimensional Brownian motion, \( A \) an \( m \times m \) real matrix, and \( \bar{B} = AB \), so that \( \bar{B} \) has covariance matrix \( \Sigma t = AA^T t \). Then, in the cases noted in the table below, the maximal functional \( \bar{D}_m = \max \sum_{k=1}^{m} \bar{B}^k(\Delta t_k) \) is equal in law to the largest eigenvalue of an \( m \times m \) Hermitian Gaussian random matrix having diagonals distributed as \( N(0, \Sigma) \) and off-diagonals with total variance given by \( f(\sigma(A)) \), independent of each other and of the diagonal entries.

| Case                        | \( \sigma(A) \)                      | \( f(\sigma(A)) \)          |
|-----------------------------|--------------------------------------|-----------------------------|
| \( m \times m, A = aP, a \in \mathbb{R}, P \text{ perm.} \) | \{a, a, \ldots, a\}                   | \( a \)                     |
| \( m \times m, A = aI + 1_m b^t \) | \{a + b^t 1_m, a, \ldots, a\} | \( a \)                     |
| \( 2 \times 2 \text{ cyclic } A = \sum_{k=0}^{1} a_k P^k \) | \{a_0 + a_1, a_0 - a_1\}     | \( a_0 - a_1 \)             |
| \( 3 \times 3 \text{ cyclic } A = \sum_{k=0}^{2} a_k P^k \) | \{a_0 + a_1 + a_2, a_0 + a_1 u + a_2 u^2, a_0 + a_1 u^2 + a_2 u\}, \( u = e^{2i\pi/3} \) | \( \alpha = \sqrt{\sum_{k=0}^{3} a_k^2 - \sum_{k \neq \ell} a_k a_\ell} = |a_0 + a_1 u + a_2 u^2| = |a_0 + a_1 u^2 + a_2 u| \) |

### 6 Concluding Remarks

In this paper, we have obtained the limiting shape of Young diagrams generated by an aperiodic, irreducible, homogeneous Markov chain on a finite state alphabet. The following remarks indicate natural directions in which our results in some cases can, and in other cases, may hope to, be extended.

- Our limiting theorems have all been proved assuming that the initial distribution is the stationary one. However, such results as Theorem 2 of Derriennic and Lin [10] allow to extend our framework to initial distributions started at a specified state. Indeed, in this case, i.e., if for some \( k = 1, \ldots, m \), \( \mathbb{P}(X_0 = a_k) = 1 \), the asymptotic covariance matrix is still given by (2.14), and, for example, Theorem 3.2 remains valid. For an arbitrary initial distribution, what is needed in this non-stationary context is an invariance principle. More generally, our results continue to hold for \( k \)-th-order Markov chains, and in fact, they extend to any sequence for which both an asymptotic covariance matrix and an
invariance principle exist.

- Our limiting theorems have only been proved for finite alphabets. However, from the authors’ previous work [17], it is known that for countably infinite iid alphabets, $LI_n$ has a limiting law corresponding to that of a non-uniform, finite-alphabet. Hence, for a countably infinite-alphabet Markov chain (subject to additional constraints), we might still be able to obtain limiting laws of the form developed in this paper.

- By using appropriate existing concentration inequalities, one can expect to establish the convergence of the moments of the rows of the diagrams.

- One field in which the connection between Brownian functionals and random matrix theory has been exploited is in Queuing Theory. The development below, following O’Connell and Yor [25], shows how Brownian functionals of the sort we have studied arise as generalizations of standard queuing models.

Let $A(s, t]$ and $S(s, t]$, $-\infty < s < t < \infty$, be two independent Poisson point process on $\mathbb{R}$, with intensity measures $\lambda$ and $\mu$, respectively, with $0 < \lambda < \mu$. Here $A$ represents the arrivals process, and $S$ the service time process, at a queue consisting of a single server. The condition $\lambda < \mu$ ensures that the queue length

$$Q(t) = \sup_{-\infty < s \leq t} \{A(s, t] - S(s, t]\}, \quad (6.1)$$

is a.s. finite, for any $t \in \mathbb{R}$. Then, defining the departure process

$$D(s, t] = A(s, t] - (Q(t) - Q(s)), \quad (6.2)$$

which is simply the number of arrivals during $(s, t]$ less the change in the queue length during $(s, t]$, the classical problem is to determine the distribution of $D(s, t]$. The answer to this problem is given by Burke’s Theorem [8] (see Theorem 1 of [25]):

**Theorem 6.1** $D$ is a Poisson process with intensity $\lambda$, and $\{D(s, t], s \leq t\}$ is independent of $\{Q(s), s \geq t\}$.

That is, $D$ has the same law as the arrivals process $A$. Moreover, since the queue length after time $t$ is independent of the process $D$ up to time $t$, one may take the departures from the first queue and use them as inputs to a second queue, and observe that the departure process from the second queue also has the law of $A$. Proceeding in this way, one generalizes to a tandem queue of $n$ servers, each taking the departures from the previous queue as its arrivals process.

One can further generalize this model to a Brownian queue in tandem in the following manner. Let $B, B^1, B^2, \ldots, B^n$ be independent, standard Brownian motions on $\mathbb{R}$, and write $B_k(s, t] = B^k(t) - B^k(s)$, for each $k$ and $s < t$, and similarly for $B$. Let $m > 0$ be a constant, and define, in complete analogy to (6.1) and (6.2),

$$q_1(t) = \sup_{-\infty < s \leq t} \{B(s, t] + B^1(s, t] - m(t-s)\}, \quad (6.3)$$

and, for $s < t$,
\[ d_1(s, t) = B(s, t) - (q_1(t) - q_1(s)). \] (6.4)

For \( k = 2, 3, \ldots, n \), let

\[ q_k(t) = \sup_{-\infty < s \leq t} \left\{ d_{k-1}(s, t) + B^k(s, t) - m(t - s) \right\}, \] (6.5)

and, for \( s < t \),

\[ d_k(s, t) = d_{k-1}(s, t) - (q_k(t) - q_k(s)). \] (6.6)

Here \( B \) is the arrivals process for the first queue, \( d_{k-1} \) is the arrivals process for the \( k^{th} \) queue \((k \geq 2)\), and \( m(t) - B^k(t) \) is the service process for the \( k^{th} \) queue, for all \( k \). Using the ideas employed in Burke’s Theorem, it can be shown that the generalized queue lengths \( q_1(0), q_2(0), \ldots, q_n(0) \) are iid random variables. Moreover, they are exponentially distributed with mean \( 1/m \).

Using the definitions in (6.3)-(6.6), and a simple inductive argument, one finds that

\[ \sum_{k=0}^{n} q_k(0) = \sup_{t > 0} \{ B(-t, 0) - mt + L_n(t) \}, \] (6.7)

where

\[ L_n(t) = \sup_{0 \leq s_1 \leq \cdots \leq s_{m-1} \leq t} \{ B^1(-t, -s_{n-1}) + \cdots + B^n(-s_1, 0) \}. \] (6.8)

By Brownian rescaling, we observe that

\[ L_n(t) \overset{\mathcal{D}}{=} \sqrt{t} \sup_{0 \leq s_1 \leq \cdots \leq s_{m-1} \leq t} \{ B^1(-t, -s_{n-1}) + \cdots + B^n(-s_1, 0) \} \overset{\mathcal{D}}{=} \sqrt{t} V^1_{\infty}, \] (6.9)

where the functional \( V^1_{\infty} \) is as in Theorem 3.2 with associated \( n \times n \) covariance matrix \( \Sigma = tI_n \) and parameter set \( I_{1,n} \). Thus, \( L_n(t) \) may be thought of as a process version of this \( V^1_{\infty} \).

The generalized Brownian queues in (6.3)-(6.6) involved independent Brownian motions. These can be replaced by Brownian motions \( B^1, \ldots, B^n \) for which \((\sigma_1 B^1(t), \ldots, \sigma_n B^n(t))\) has (nontrivial) covariance matrix \( t\Sigma \). Whether or not we keep the initial arrival process \( B \) independent of \((B^1, \ldots, B^n)\), we no longer have that \( q_1(0), q_2(0), \ldots, q_n(0) \) are iid random variables, due to the dependence among the service times \( mt - B^k(t) \), but we do still have the identity (6.7) and (6.9) relating the total occupancy of the queue at time zero to \( V^1_{\infty} \). More importantly, our generalizations of the Brownian functionals \( L_n(t) \) above can be used to describe the joint law of the input/output of each queue.

- An important topic connecting much of random matrix theory to other problems, such as the shape of random Young diagrams, is the field of orthogonal polynomials. (See, e.g., [22].) It would be of great interest to see what, if any, classes of orthogonal polynomials are associated with the present paper.
References

[1] Baik, J., Deift, P., and Johansson, K., “On the distribution of the length of the longest increasing subsequence of random permutations,” J. Amer. Math. Soc., vol. 12, no. 4, pp. 1119-1178, 1999.

[2] Baik, J., Deift, P., and Johansson, K., “On the distribution of the length of the second row of a Young diagram under Plancherel measure,” Geom. Funct. Anal., vol. 10, no. 4, pp. 702-731, 2000.

[3] Baik, J., Deift, P., and Johansson, K., ”Addendum to: “On the distribution of the length of the second row of a Young diagram under Plancherel measure”,” Geom. Funct. Anal., vol. 10, no. 6, pp. 1606-1607, 2000.

[4] Baryshnikov, Y., “GUEs and queues,” Probab. Theory Related Fields, vol. 119, no. 2, pp. 256-274, 2001.

[5] Biane, P., Bougerol, P., and O’Connell, N., “Littelmann paths and Brownian paths,” Duke Math. J., vol. 130, pp. 127-167, 2005.

[6] Borodin, A., Okounkov, A., and Olshanski, G, “Asymptotics of Plancherel measures for symmetric groups,” J. Amer. Math. Soc., vol. 13, no. 3, pp. 481-515 (electronic), 2000.

[7] Bougerol, P. and Jeulin, T., “Paths in Weyl chambers and random matrices,” Probab. Theory Related Fields, vol. 124, no. 4, pp. 517-543, 2002.

[8] Burke, P.J., “The output of a queueing system,” Operations. Res., vol. 4, pp. 699-704 (1957), 1956.

[9] Chistyakov, G.P. and Götze, F., “Distribution of the shape of Markovian random words,” Probab. Theory Related Fields, vol. 129, no. 1, pp. 18-36, 2004.

[10] Derriennic, Y. and Lin, M., “The central limit theorem for Markov chains with normal transition operators, started at a point,” Probab. Theory Related Fields, vol. 119, no. 4, pp. 508-528, 2001.

[11] Doumerc, Y., “A note on representations of eigenvalues of classical Gaussian matrices,” in Séminaire de Probabilités XXXVII, vol. 1832 of Lecture Notes in Math., pp. 370-384, Berlin: Springer, 2003.

[12] Fulton, W., Young diagrams, vol. 35 of London Mathematical Society Student Texts. Cambridge: Cambridge University Press, 1997. With applications to representation theory and geometry.

[13] Glynn, P.W. and Whitt, W., “Departure from many queues in series,” Ann. Appl. Probab., vol. 1, no. 4, pp. 546-572, 1991.

[14] Gordin, M.I., “The central limit theorem for stationary processes,” Dokl. Akad. Nauk SSSR, vol. 188, pp. 739-741, 1969.

[15] Gravner, J., Tracy, C., and Widom, H., “Limit theorems for height fluctuations in a class of discrete space and time growth models,” J. Statist. Phys., vol. 102, no. 5-6, pp. 1085-1132, 2001.

[16] Greene, C., “Some partitions associated with a partially ordered set,” J. Combin. Theory Ser. A, vol. 20, no. 1, pp. 69-79, 1976.
[17] Houdré C. and Litherland, T., “On the longest increasing subsequence for finite and countable alphabets,” in High Dimensional Probability V: The Luminy Volume (Beachwood, Ohio, USA: Institute of Mathematical Statistics), 185-212, 2009.

[18] Houdré C. and Litherland, T., “Asymptotics for the length of the longest increasing subsequence of a binary Markov random word,” ArXiv #math.PR/1110.1324. To appear in: Malliavin Calculus and Stochastic Analysis: A Festschrift in Honor of David Nualart.

[19] Houdré C. and Xu, H., “On the limiting shape of Young diagrams associated with inhomogeneous random words,” ArXiv #math.PR/0901.4138v2. To appear in: High Dimensional Probability VI (2012).

[20] Its, A.R., Tracy, C. A., and Widom, H., “Random words, Toeplitz determinants, and integrable systems. I,” in Random matrix models and their applications, vol. 40 of Math. Sci. Res. Inst. Publ., pp. 245-258, Cambridge: Cambridge Univ. Press, 2001.

[21] Its, A.R., Tracy, C. A., and Widom, H., “Random words, Toeplitz determinants, and integrable systems. II,” Phys. D, vol 152/153, pp. 199-224, 2001. Advances in Nonlinear Mathematics and Science.

[22] Johansson, K., “Discrete orthogonal polynomial ensembles and the Plancherel measure,” Ann. of Math. (2), vol. 153, no.1, pp. 259-296, 2001.

[23] Kirillov, A., “Introduction to tropical combinatorics,” Physics and Combinatorics, 2000 (Nagoya), pp. 82-150, World Sci. Publ., River Edge, NJ, 2001.

[24] Kuperberg, G., “Random words, quantum statistics, central limits, random matrices,” Methods. Appl. Anal., vol. 9, no. 1, pp. 99-118, 2002.

[25] O’Connell, N. and Yor, M., “Brownian analogues of Burke’s theorm,” Stochastic Process. Appl., vol. 96, no. 2, pp. 285-304, 2001.

[26] O’Connell, N. and Yor, M., “A representation for non-colliding random walks,” Electron. Comm. Probab., vol. 7, pp. 1-12 (electronic), 2002.

[27] Okounkov, A., “Random matrices and random permutations,” Internat. Math. Res. Notices., no. 20, pp. 1043-1095, 2000.

[28] Tracy, C. A. and Widom, H., “On the distributions of the lengths of the longest monotone subsequences in random words,” Probab. Theory Related Fields, vol. 119, no. 3, pp. 350-380, 2001.