



# GUE minors, maximal Brownian functionals and longest increasing subsequences

Florent Benaych-Georges, Christian Houdré

► **To cite this version:**

Florent Benaych-Georges, Christian Houdré. GUE minors, maximal Brownian functionals and longest increasing subsequences. *Markov Processes and Related Fields, Polymath*, 2015, 21, pp.109-126.

**HAL Id: hal-00917273**

<https://hal.archives-ouvertes.fr/hal-00917273v4>

Submitted on 26 Nov 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# GUE MINORS, MAXIMAL BROWNIAN FUNCTIONALS AND LONGEST INCREASING SUBSEQUENCES IN RANDOM WORDS

FLORENT BENAYCH-GEORGES AND CHRISTIAN HOUDRÉ

ABSTRACT. We present equalities in law between the spectra of the principal minors of a GUE matrix and some maximal functionals of independent Brownian motions. In turn, these results allow to recover the limiting shape (properly centered and scaled) of the RSK Young diagrams associated with a random word as a function of the spectra of these minors. Since the length of the top row of the diagrams is the length of the longest increasing subsequence of the random word, the corresponding limiting law also follows.

## 1. INTRODUCTION

It is by now well known that there exist strong and interesting connections between directed percolation and random matrices. The precise results we have in mind have their origins in the identity in law, due to Baryshnikov [2] and Gravner, Tracy and Widom [11], between the maximal eigenvalue of an  $M \times M$  element of the GUE and a certain maximal functional of standard  $M$ -dimensional Brownian motion originating in queuing theory, with Glynn and Whitt [10]. This first result has seen many extensions and complements. For example, Bougerol and Jeulin [5] as well as O’Connell and Yor [27] obtained identities in law between (different) multivariate Brownian functionals and the spectrum of the GUE whose equivalence is shown in Biane, Bougerol and O’Connell [3]. Various related representations have also been put forward and studied for instance in Doumerc [8], Johansson [18, 20], O’Connell [26] to name but a few authors and pieces of work.

Our interest in such representations comes from the identification, first obtained by Kerov [22], of the limiting length (properly centered and scaled) of the longest increasing subsequence of a random word, as the maximal eigenvalue of a certain random matrix. For

---

2000 *Mathematics Subject Classification.* 15A52;60F05.

*Key words and phrases.* Random matrices, Brownian motion, random words, longest increasing subsequences, RSK correspondence.

FBG: MAP5, Université Paris Descartes, 45, rue des Saints-Pères 75270 Paris Cedex 06, France. florent.benaych-georges@parisdescartes.fr. This work was done in part during this author’s stay at Georgia Institute of Technology.

CH: School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia, 30332-0160, USA. houdre@math.gatech.edu. This work was partially supported by the grant #246283 from the Simons Foundation and by a Simons Foundation Fellowship, grant #267336. Many thanks to the Laboratory MAS of École Centrale, Paris, and to the LPMA of the Université Pierre et Marie Curie for their hospitality while part of this research was carried out.

example, in the case of a word with i.i.d. uniformly distributed letters in an alphabet of size  $M$ , the limiting law is the maximal eigenvalue of the  $M \times M$  traceless GUE. Moreover, [22] showed that the whole normalized limiting shape of the RSK Young diagrams associated with the random word, is the spectrum of the  $M \times M$  traceless GUE. Since the length of the top row of the diagrams is the length of the longest increasing subsequence of the random word, the maximal eigenvalue result is recovered. (The asymptotic length result was rediscovered by Tracy and Widom [28] and the asymptotic shape one by Johansson [19], who positively answered a conjecture of Tracy and Widom. Extensions to non-uniform letters were also obtained by Its, Tracy and Widom [16, 17].)

Limiting laws expressed in terms of maximal Brownian functionals are also obtained in [12]. These last representations involve *dependent* Brownian motions and do not clearly recover the results of [28] or [16, 17], which themselves are mainly derived by analytical techniques. To resolve this issue, we provide below an extension of Baryshnikov's result [2] on the identification of the multivariate law of the maximal eigenvalues of the principal minors of a GUE matrix, with some maximal functionals of a standard multidimensional Brownian motion. This allows us to circumvent the analytical approach and provides a mixed combinatorial/probabilistic methodology for the solutions of these finite alphabet longest increasing subsequence problems. Our hope is that Theorem 1, below, will also be helpful to fully identify eigenvalues of random matrices as the limiting laws in the corresponding Markov random word problems (see Conjecture 7 in Kuperberg [23]). In the Markovian setting, the analytical methodology is lacking, in contrast to the probabilistic one, and to date the limiting laws are mainly only expressed as Brownian functionals (see [13]). Indeed, the multivariate functional appearing in Theorem 1 is exactly the one giving the limiting law of the shape of the RSK image of a Markov random word in [13], the only difference being that the Brownian motions in [13] are correlated, say, with covariance  $\Sigma$ . This correlation issue in maximal functionals often amounts to adding a condition on the trace of the random matrix (as in [28, 19, 16, 17]). However, for general Markov random words the full identification of these functionals via random matrices remains open. For Markov random words with, e.g., cyclic transition matrix, the length of the longest increasing subsequence will be asymptotically identified as the eigenvalue of some random matrices once one obtains a more general version of Theorem 1 below, with correlated Brownian motions. Our intuition is that to get such a generalization, one needs to consider the principal minors of more general random matrices, namely Gaussian Hermitian matrices with as diagonal, and up to scaling factor, a Gaussian vector with covariance  $\Sigma$ . Besides providing the final touch to an essentially probabilistic proof of the random word asymptotics problem, our results also allow us to shed new lights on the queuing framework by providing, for example, joint limiting laws involving departing times and service times of individual customers.

## 2. STATEMENTS AND PROOFS OF THE RESULTS

Throughout, fix a positive integer  $M$  and consider:

- an  $M \times M$  GUE matrix  $H = [h_{ij}]$ , i.e., a standard Gaussian variable on the space of  $M \times M$  Hermitian matrices endowed with the Euclidean scalar product given by  $X \cdot Y = \text{Tr } XY$ ,
- an  $M$ -dimensional standard Brownian motion  $B = (B_k(t))_{t \in [0,1], k=1, \dots, M}$ .

For each  $k = 1, \dots, M$ , denote by

$$(1) \quad \mu_1^k \geq \dots \geq \mu_k^k,$$

the eigenvalues of the principal  $k \times k$  minor of  $H$ . Next, introduce the set

$$\mathcal{P} := \{\pi : [0, 1] \rightarrow \{1, \dots, M\}, \text{ càdlàg, non-decreasing}\},$$

and for  $\pi \in \mathcal{P}$ , let

$$\Delta_\pi(B) := \int_0^1 dB_{\pi(t)}(t) = \sum_{j=1}^M (B_j(t_j) - B_j(t_{j-1})),$$

where  $0 = t_0 \leq t_1 \leq \dots \leq t_M = 1$  are such that

$$\pi(\cdot) = \sum_{j=1}^{M-1} j \times 1_{[t_{j-1}, t_j)}(\cdot) + M \times 1_{[t_{M-1}, t_M]}(\cdot).$$

To complete our notations, for  $\pi_1, \pi_2 \in \mathcal{P}$ , we write  $\pi_1 < \pi_2$  whenever  $\pi_1(t) < \pi_2(t)$  for all  $t \in [0, 1]$ . Let us now state our first result which, in particular, when  $\ell = 1$  below, identifies the joint law of the maximal eigenvalue of the principal minors of  $H$  and therefore extends Theorem 0.7 of [2].

**Theorem 1.** *The following equality in law holds true:*

$$\left( \sum_{i=1}^{\ell} \mu_i^k \right)_{1 \leq \ell \leq k \leq M} \stackrel{\text{law}}{=} \left( \sup \left\{ \sum_{i=1}^{\ell} \Delta_{\pi_i}(B); \pi_1, \dots, \pi_\ell \in \mathcal{P}, \pi_1 < \dots < \pi_\ell \leq k \right\} \right)_{1 \leq \ell \leq k \leq M}.$$

This theorem has also a process version where the matrix  $H$  is replaced by an Hermitian Brownian motion, where  $B$  is taken up to time  $t$  and not to time 1.

In the forthcoming proof, and throughout,  $\implies$  indicates convergence in distribution.

*Proof.* Let  $w_{N,M} := [w_{ij}]$  ( $i = 1, \dots, N$ ,  $j = 1, \dots, M$ ) be an array of i.i.d. geometric random variables with parameter  $q \in (0, 1)$ , i.e., with law  $\sum_{k \geq 0} q^k (1-q) \delta_k$ . Such variables have mean  $e := q/(1-q)$  and variance  $v := q/(1-q)^2$ .

Applying the RSK correspondence to  $w_{N,M}$  (see e.g., [25] for an introduction to the RSK correspondence applied to arrays of integers) gives a pair  $(P, Q)$  of Young diagrams with the same shape. Let us denote the shape of these Young diagrams by

$$\lambda_1^M \geq \dots \geq \lambda_M^M,$$

where the exponent  $M$  is here to emphasize the dependence on the dimension  $M$  of the GUE matrix  $H$  (the dependence on  $N$  is implicit). In the same way, one can of course define, for each  $k = 1, \dots, M$ , the shape

$$\lambda_1^k \geq \dots \geq \lambda_k^k,$$

of the Young diagrams obtained by applying the RSK correspondence to the array  $w_{N,k}$ , which is the array  $w_{N,M}$  where all but the first  $k$  columns have been removed. Note that

$$(2) \quad \begin{pmatrix} & & & \lambda_1^1 & & & \\ & & & & \lambda_2^2 & & \\ & & \lambda_1^3 & & \lambda_2^3 & & \lambda_3^3 \\ & & \dots & & \dots & & \dots \\ \lambda_1^M & & \dots & & \dots & & \lambda_M^M \end{pmatrix},$$

is a Gelfand-Tsetlin pattern, i.e., satisfies the interlacing inequalities  $\lambda_i^k \geq \lambda_i^{k-1} \geq \lambda_{i+1}^k$  ( $1 \leq i < k \leq M$ ). This can be seen from the fact that if  $(P_k, Q_k)$  denotes the pair associated to  $w_{N,k}$  by the RSK correspondence, then  $P_{k-1}$  can be deduced from  $P_k$  by removing all the boxes filled with the number  $k$ .

Let us now define the random variables

$$(3) \quad \xi_i^k := \frac{\lambda_i^k - eN}{\sqrt{vN}}.$$

We then have the following lemma.

**Lemma 2.** *As  $N$  tends to infinity,*

$$\begin{pmatrix} & & & \xi_1^1 & & & \\ & & & & \xi_2^2 & & \\ & & \xi_1^3 & & \xi_2^3 & & \xi_3^3 \\ & & \dots & & \dots & & \dots \\ \xi_1^M & & \dots & & \dots & & \xi_M^M \end{pmatrix} \implies \begin{pmatrix} & & & \mu_1^1 & & & \\ & & & & \mu_2^2 & & \\ & & \mu_1^3 & & \mu_2^3 & & \mu_3^3 \\ & & \dots & & \dots & & \dots \\ \mu_1^M & & \dots & & \dots & & \mu_M^M \end{pmatrix},$$

where the  $\mu_i^k$ s are the ones introduced in (1).

*Proof.* This lemma is stated as Proposition 3.12 in [21] (with slightly different notation). However, for the convenience of the reader, we shall say a few words about its proof:

By a well known property of the geometric law, conditionally on  $S := \sum_{i,j} w_{i,j}$ , the law of  $w_{N,M}$  is the uniform law on the set of  $N \times M$  integer arrays summing up to  $S$ . But the RSK correspondence  $w \mapsto (P(w), Q(w))$  is a bijection, when defined on this set, and the information carried out by the Gelfand-Tsetlin pattern  $\lambda$  of (2) is the same as the one carried out by  $P(w)$ . It follows that conditionally on the event  $S = s$ , the law of the

Gelfand-Tsetlin pattern  $\lambda$  of (2) can be recovered as follows: First choose  $\lambda_1^M \geq \dots \geq \lambda_M^M$  according to its particular conditional law, and then  $\lambda$  is uniformly distributed on the set of integer valued Gelfand-Tsetlin patterns with bottom row  $\lambda_1^M \geq \dots \geq \lambda_M^M$ . Therefore, by conditioning, it is clear that the unconditional law of  $\lambda$  can be recovered exactly in the same way.

Moreover, it can also be proved [2, Proposition 4.7] that conditionally on the value of its bottom row  $\mu_1^M \geq \dots \geq \mu_M^M$ , the law of  $(\mu_i^k)_{1 \leq i \leq k \leq M}$  is the uniform law of real valued Gelfand-Tsetlin patterns with bottom row  $\mu_1^M \geq \dots \geq \mu_M^M$ .

Thus as  $M$  is fixed, up to standard analysis (see e.g. [2, Lem. 2.5]), it suffices to prove that the convergence in distribution stated in the lemma is true when restricted to the bottom rows:

$$(4) \quad (\xi_1^M, \dots, \xi_M^M) \implies (\mu_1^M, \dots, \mu_M^M).$$

To prove (4), one only needs to notice that the joint distribution of its right hand side has density

$$(5) \quad \frac{1}{\text{normalizing constant}} \prod_{1 \leq i < j \leq M} (\mu_j^M - \mu_i^M)^2 e^{-\frac{1}{2} \sum_i (\mu_i^M)^2},$$

(this is a well known fact, see e.g. [1, Th. 2.5.2]), whereas it has been proved by Johansson in [18] that the joint density of the left hand side of (4) with respect to the counting measure is

$$(6) \quad \frac{1}{\text{normalizing constant}} \prod_{1 \leq i < j \leq M} (\xi_i^M - \xi_j^M + j - i)^2 \prod_{i=1}^M \frac{(\xi_i^M + M)!}{(\xi_i^M + M - i)!}.$$

It can easily be seen that the quantities of (5) and (6) agree at the limit.  $\square$

Let us now give a few more definitions. Fix some positive integers  $k, N$ .

- An *up-right path with values in*  $\{1, \dots, N\} \times \{1, \dots, k\}$  is a map  $\pi$  defined on a set  $\{1, \dots, p\}$ , for a positive integer  $p$ , taking values in  $\{1, \dots, N\} \times \{1, \dots, k\}$  and having two components  $\pi_1, \pi_2$ , such that for all  $i \in \{2, \dots, p\}$ , we have equality of the sets

$$\{\pi_1(i) - \pi_1(i-1), \pi_2(i) - \pi_2(i-1)\} = \{0, 1\}.$$

- The *support* of an up-right path  $\pi : \{1, \dots, p\} \rightarrow \{1, \dots, N\} \times \{1, \dots, k\}$  is the set  $\{\pi(1), \dots, \pi(p)\}$ .
- The *starting point* (resp. *ending point*) of an up-right path  $\pi : \{1, \dots, p\} \rightarrow \{1, \dots, N\} \times \{1, \dots, k\}$  is  $\pi(1)$  (resp.  $\pi(p)$ ). The *starting abscissa* (resp. *ending abscissa*) of  $\pi$  is the first coordinate of its starting point (resp. ending point).
- For  $\pi$  an up-right path with values in  $\{1, \dots, N\} \times \{1, \dots, k\}$  and  $P$  a point of the support of  $\pi$ , we denote by  $\pi^{[P]}$  (resp.  $\pi^{(P]}$ ) the path obtained by removing all points in  $\pi$  strictly before (resp. strictly after) the point  $P$  (these last definitions are unambiguous since, as defined, an up-right path is always one-to-one).

- For  $\pi, \pi'$  up-right paths with values in  $\{1, \dots, N\} \times \{1, \dots, k\}$  such that the starting point of  $\pi'$  is just to the right or just above the ending point of  $\pi$ , we denote by  $\pi \cup \pi'$  their (naturally defined) concatenation.
- For  $\pi, \pi'$  up-right paths with values in  $\{1, \dots, N\} \times \{1, \dots, k\}$ , we write  $\pi < \pi'$ , if for any  $n \in \{1, \dots, N\}$ , the intersection of the support of  $\pi$  with  $\{n\} \times \{1, \dots, k\}$  is located strictly below the intersection of the support of  $\pi'$  with  $\{n\} \times \{1, \dots, k\}$ .

We will also need the following lemmas.

**Lemma 3.** *Let  $1 \leq \ell \leq k$  and let  $\pi_1, \dots, \pi_\ell$  be up-right paths with values in  $\{1, \dots, N\} \times \{1, \dots, k\}$  and having pairwise disjoint supports. Then, there exist up-right paths  $\pi'_1, \dots, \pi'_\ell$  with values in  $\{1, \dots, N\} \times \{1, \dots, k\}$ , having pairwise disjoint supports and starting abscissas all equal to 1 such that*

$$\cup_{i=1}^{\ell} \text{support}(\pi_i) \subset \cup_{i=1}^{\ell} \text{support}(\pi'_i).$$

*Proof.* Note first that if  $\ell = k$ , the lemma is clearly true by choosing  $\pi'_1, \dots, \pi'_k$  to be the maximal horizontal lines with starting abscissas all equal to 1. So let us assume that  $\ell \leq k - 1$ , and present a trivial claim which will be useful in the sequel:

**Claim :** The hypothesis  $\ell \leq k - 1$  implies that at least one point of the set  $\{1\} \times \{1, \dots, k\}$  is the starting point of no  $\pi_i$  or that at least one  $\pi_i$  has starting point in  $\{1\} \times \{1, \dots, k\}$  and second point just above its starting point.

Next, let  $\text{sa}_{\max}$  denote the maximum of the starting abscissas of  $\pi_1, \dots, \pi_\ell$  and  $\ell_{\text{sa}_{\max}}$  denote the number of  $i$ 's such that  $\pi_i$  has starting abscissa  $\text{sa}_{\max}$ . Let us prove the lemma by induction (for the lexical order) on  $(\text{sa}_{\max}, \ell_{\text{sa}_{\max}})$ . If  $\text{sa}_{\max} = 1$ , then there is nothing to do (as all the  $\pi_i$ 's have starting abscissas equal to 1). This begins the induction. So let us now assume that  $\text{sa}_{\max} \geq 2$ . Let  $i_0$  be such that  $\pi_{i_0}$  has starting abscissa  $\text{sa}_{\max}$ . Let  $P$  be the point in  $\{1, \dots, N\} \times \{1, \dots, k\}$  just to the left of the starting point of  $\pi_{i_0}$ .

- If  $P$  is on none of the supports of the  $\pi_i$ 's, then one can extend  $\pi_{i_0}$  by adding  $P$  at its beginning and conclude by the induction hypothesis.
- If  $P$  is on the support of a path  $\pi_{i_1}$ , then two cases can occur:
  - If  $\pi_{i_1}$  does not end at  $P$ : then the point after  $P$  in  $\pi_{i_1}$  must be the point  $Q$  just above  $P$ , and replacing  $\pi_{i_0}$  by  $\pi_{i_1}^{[P]} \cup \pi_{i_0}$  and  $\pi_{i_1}$  by  $\pi_{i_1}^{[Q]}$  allows to conclude by the induction hypothesis.
  - If  $\pi_{i_1}$  ends at  $P$ : then replacing  $\pi_{i_0}$  by  $\pi_{i_1} \cup \pi_{i_0}$  and  $\pi_{i_1}$  by a point of the set  $\{1\} \times \{1, \dots, k\}$  allows to conclude (note that replacing  $\pi_{i_1}$  by a point of the set  $\{1\} \times \{1, \dots, k\}$  is possible by the claim stated above: if a point of the set  $\{1\} \times \{1, \dots, k\}$  is the starting point of no  $\pi_i$  then this is trivial, while if a  $\pi_{i_2}$  has starting point in  $\{1\} \times \{1, \dots, k\}$  and second point just above its starting point, one can remove its starting point from  $\pi_{i_2}$  to get back to the first case).

□

**Lemma 4.** *Let  $1 \leq \ell \leq k$  and let  $\pi_1, \dots, \pi_\ell$  be up-right paths with values in  $\{1, \dots, N\} \times \{1, \dots, k\}$ , pairwise disjoint supports and starting abscissas all equal to 1. Then there exist up-right paths  $\pi'_1, \dots, \pi'_\ell$  with values in  $\{1, \dots, N\} \times \{1, \dots, k\}$ , having pairwise disjoint supports, starting abscissas all equal to 1 ending abscissas all equal to  $N$ , and such that*

$$\cup_{i=1}^{\ell} \text{support}(\pi_i) \subset \cup_{i=1}^{\ell} \text{support}(\pi'_i).$$

*Proof.* The proof is quite similar to the proof of the previous lemma. The difference lies in that the hypotheses require all the starting abscissas to be equal to one, while the conclusion requires all the ending abscissas to be equal to  $N$ .

Let  $\text{ea}_{\min}$  denote the minimum of the ending abscissas of  $\pi_1, \dots, \pi_\ell$  and  $\ell_{\text{ea}_{\min}}$  denote the number of  $i$ 's such that  $\pi_i$  has ending abscissa  $\text{ea}_{\min}$ . Let us prove the lemma by induction (for the lexical order) on  $(\text{ea}_{\min}, \ell_{\text{ea}_{\min}})$ . If  $\text{ea}_{\min} = N$ , then there is nothing to do (as all the  $\pi_i$ 's have ending abscissas equal to  $N$ ). This begins the induction. So let us next assume that  $\text{ea}_{\min} \leq N - 1$ . Let  $i_0$  be such that  $\pi_{i_0}$  have ending abscissa  $\text{ea}_{\min}$ . Let  $P$  be the point in  $\{1, \dots, N\} \times \{1, \dots, k\}$  just to the right of the ending point of  $\pi_{i_0}$ .

- If  $P$  is on none of the supports of the  $\pi_i$ 's, then one can extend  $\pi_{i_0}$  by adding  $P$  to its end and conclude by the induction hypothesis.
- If  $P$  is on the support of a path  $\pi_{i_1}$ , then  $P$  is not the starting point of  $\pi_{i_1}$  and the point preceding  $P$  in  $\pi_{i_1}$  is the point  $Q$  just below  $P$ . Hence replacing  $\pi_{i_0}$  by  $\pi_{i_0} \cup \pi_{i_1}^{[P]}$  and  $\pi_{i_1}$  by  $\pi_{i_1}^{[Q]}$  allows to conclude by the induction hypothesis.

□

**Lemma 5.** *Let  $1 \leq \ell \leq k$  and let  $\pi_1, \dots, \pi_\ell$  be up-right paths with values in  $\{1, \dots, N\} \times \{1, \dots, k\}$  having pairwise disjoint supports, starting abscissas all equal to 1 and ending abscissas all equal to  $N$ . Then, the collection  $(\pi_i)_{1 \leq i \leq \ell}$  can be re-indexed in such a way that  $\pi_1 < \dots < \pi_\ell$ .*

*Proof.* It suffices to re-index the collection  $(\pi_i)_{1 \leq i \leq \ell}$  according to the order of their starting abscissas and to use the definition of up-right paths combined with the disjoint supports.

□

**Lemma 6.** *For each  $1 \leq \ell \leq k \leq M$ ,*

$$(7) \quad \lambda_1^k + \dots + \lambda_\ell^k = \max_{\pi_1, \dots, \pi_\ell} \sum_{r=1}^{\ell} \sum_{(i,j) \in \pi_r} w_{ij},$$

where the max is over collections  $\{\pi_1 < \dots < \pi_\ell\}$  of up-right paths in the set

$$\{1, \dots, N\} \times \{1, \dots, k\}$$

starting in the subset  $\{1\} \times \{1, \dots, k\}$  and ending in the subset  $\{N\} \times \{1, \dots, k\}$ .

*Proof.* By [25, Th. 1.1.1] and [9, Chap. 3, Lemma 1], (7) is true when the paths are only required to be pairwise disjoint, without any condition on the starting and ending points.



Then, the previous three lemmas allow to claim that any set of  $\ell$  pairwise disjoint paths can be changed into a set of  $\ell$  pairwise disjoint paths  $\pi_1 < \dots < \pi_\ell$ , with starting abscissas all equal to 1 and ending abscissas all equal to  $N$ , in such a way that the union of the supports of the new paths contains the union of the supports of the former ones. To finish the proof, it then suffices to notice that since the  $w_{ij}$ 's are non-negative, enlarging the union of the supports of the paths never decreases the total weight.  $\square$

To complete the proof of Theorem 1, note that any up-right path  $\pi_r$  as described in the previous lemma is a concatenation of at most  $k$  paths with fixed second coordinate and has length between  $N$  and  $N + M$ . Moreover, by Donsker theorem (see [4, 10]), the  $M$ -dimensional process

$$\left( \frac{1}{\sqrt{vN}} \sum_{i=1}^{\lfloor Nt \rfloor} (w_{ik} - e) \right)_{k=1, \dots, M},$$

converges in distribution (for the Skorohod topology) to the  $M$ -dimensional Brownian motion  $B$ . To finish the proof, apply both Lemma 2 and Lemma 6.  $\square$

### 3. APPLICATIONS OF THEOREM 1

At first, Let us present and prove some corollaries of Theorem 1 which motivated, in large part, the present study.

Let us briefly recall the framework of the works on random words cited in the Introduction. Let  $(X_i)_{i \geq 1}$  be a sequence of i.i.d. random variables on a totally ordered finite alphabet  $\mathcal{A}$  of cardinality  $k$ . Denote the elements of  $\mathcal{A}$  by  $\alpha_1, \dots, \alpha_k$  listed in such a way that if  $p_i := \mathbb{P}(X_1 = \alpha_i)$ ,  $i = 1, \dots, k$ , then  $p_1 \geq \dots \geq p_k$  (therefore this indexing of the letters in  $\mathcal{A}$  has nothing to do with the order used on  $\mathcal{A}$ ). Next, decompose the alphabet  $\mathcal{A}$  into subsets  $\mathcal{A}_1, \dots, \mathcal{A}_n$  in such a way that  $\alpha_i$  and  $\alpha_j$  belong to the same  $\mathcal{A}_m$ ,  $m = 1, \dots, n \leq k$ , if and only if  $p_i = p_j$ . Finally, let  $\text{LI}_N$  be the length of the longest (weakly) increasing subsequence of the random word

$$X_1 \cdots \cdots X_N.$$

When combined with [12], the following corollary provides an alternative approach to [22, Chap. 3, Sec. 3.4, Th. 2] and [28] for uniform letters, or to [16]<sup>1</sup>.

---

<sup>1</sup>On the same year, the same authors published a second paper on the subject, [17]. This paper connects the analysis of the length of the longest weakly increasing subsequence of inhomogeneous random words to a Riemann-Hilbert problem and an associated system of integrable partial differential equations. In particular, it shows that the Poissonization of the distribution function of this length can be identified as the Jimbo-Miwa-Ueno tau function.

**Corollary 7.** *Let  $p_{\max} := p_1$ ,  $k_1 := \#\mathcal{A}_1$  and let  $H = [h_{ij}]$  be a  $k_1 \times k_1$  GUE matrix with largest eigenvalue  $\mu_{\max}$ . Then, as  $N$  tends to infinity,*

$$(8) \quad \frac{\text{LI}_N - Np_{\max}}{\sqrt{Np_{\max}}} \Longrightarrow \frac{\sqrt{1 - k_1 p_{\max}} - 1}{k_1} \sum_{j=1}^{k_1} h_{jj} + \mu_{\max}.$$

*Proof.* From Theorem 1, with the notation introduced above, and if  $B$  is now a  $k_1$ -dimensional standard Brownian motion,

$$(9) \quad (\mu_1^j)_{1 \leq j \leq k_1} \cup (h_{jj})_{1 \leq j \leq k_1} \stackrel{\text{law}}{=} \left( \max_{0=t_0 \leq \dots \leq t_j=1} \sum_{i=1}^j (B_i(t_i) - B_i(t_{i-1})) \right)_{1 \leq j \leq k_1} \cup (B_j(1))_{1 \leq j \leq k_1},$$

by also noticing that  $h_{11} = \mu_1^1 = B_1(1)$  and that for all  $j = 2, \dots, k_1$ ,  $h_{jj} = \sum_{i=1}^j \mu_i^j - \sum_{i=1}^{j-1} \mu_i^{j-1} \stackrel{\text{law}}{=} \sum_{i=1}^j B_i(1) - \sum_{i=1}^{j-1} B_i(1) = B_j(1)$ . Next, Corollary 3.3 in [12] asserts that

$$(10) \quad \frac{\text{LI}_N - Np_{\max}}{\sqrt{Np_{\max}}} \Longrightarrow \frac{\sqrt{1 - k_1 p_{\max}} - 1}{k_1} \sum_{j=1}^{k_1} B_j(1) + \max_{0=t_0 \leq \dots \leq t_{k_1}=1} \sum_{i=1}^{k_1} (B_i(t_i) - B_i(t_{i-1})),$$

combining (9) and (10) gives (8).  $\square$

Denote by  $\lambda_1 \geq \dots \geq \lambda_k$  the shape of the Young diagrams obtained by applying the RSK correspondence to the random word

$$X_1 \cdots \cdots X_N,$$

and let

$$\xi_i = \frac{\lambda_i - Np_i}{\sqrt{Np_i}},$$

$1 \leq i \leq k$ , be the corresponding rescaled variables. Next, introduce independent GUE matrices  $H_1, \dots, H_n$ , where each  $H_j$  has size  $k_j := \#\mathcal{A}_j$ , and let

$$H := \begin{pmatrix} H_1 & & \\ & \ddots & \\ & & H_n \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \tilde{H}_1 & & \\ & \ddots & \\ & & \tilde{H}_n \end{pmatrix} := H - \text{Tr}(HJ)J,$$

where

$$J = \text{diag}(\sqrt{p_1}, \dots, \sqrt{p_k}).$$

**Remark 8.** Note that  $J$  is a unit vector of the space  $\mathcal{H}_{k_1} \times \dots \times \mathcal{H}_{k_n}$  endowed with the Euclidean inner product structure, so  $H - \text{Tr}(HJ)J$  is the orthogonal projection onto  $J^\perp$ , so that its law is the law of  $H$  conditioned to belong to  $J^\perp$ .

The following corollary makes full use of Theorem 1 and, when combined with [15] or [13], provides an alternative approach to [22], [19], [16], [17] (see also, Bufetov [6] and Méliot [24]).

Let us now define the random vector  $(\mu_1, \dots, \mu_k)$  by

$$(\mu_1, \dots, \mu_k) := (\text{ordered spectrum of } \tilde{H}_1, \dots, \text{ordered spectrum of } \tilde{H}_n).$$

**Corollary 9.** *As  $N \rightarrow \infty$ ,*

$$(\xi_1, \dots, \xi_k) \Longrightarrow (\mu_1, \dots, \mu_k).$$

**Remark 10.** The limiting law of  $\text{LI}_N$ , rescaled, is simply the law of  $\mu_1$  and is given by

$$(11) \quad \lambda_{\max}(H_1) - p_1 \text{Tr}(H_1) - \sqrt{p_1(1 - k_1 p_1)} Z,$$

where  $Z$  is a standard normal random variable, independent of  $H_1$ . Note also that this law only depends on  $p_1$  and  $k_1$ .

*Proof.* First,

$$\text{Tr}(HJ) = \sum_{j=1}^n \sqrt{p_{(j)}} \text{Tr} H_j,$$

where for all  $j$ ,  $p_{(j)} := p_\ell$ , for  $\ell \in \mathcal{A}_j$ . So for each  $i$ , we have

$$\tilde{H}_i = H_i - \left( \sqrt{p_i} \sum_{j=1}^n \sqrt{p_{(j)}} \text{Tr} H_j \right) I,$$

where  $I$  is the corresponding identity matrix. Then, Theorem 3.1 and Remark 3.2 (iv) in [15] together with Theorem 1 allow to conclude.  $\square$

In case the i.i.d. random variables generating the random word are replaced by an (irreducible, aperiodic) homogeneous Markov chain, with state space  $\mathcal{A}$  of cardinality  $k$ , the corresponding limiting laws can also be given in terms of maximal Brownian functionals similar to those in Theorem 1 (see [13]). However, an important difference is that now the standard Brownian motion  $B$  is replaced by a correlated one  $\tilde{B}$  with, say, covariance matrix  $\Sigma$  instead of  $I$ . The possible identification of (the law of) these functionals as (the law of) maximal eigenvalues (or spectra) of random matrices has not been fully accomplished yet, although various cases are done. In particular, for cyclic transition matrices  $P$ , in which case the stationary distribution is the uniform one, there is a curious dichotomy between alphabets of size at most three and of size four or more. Indeed for  $k \leq 3$ , the cyclic hypothesis forces  $\Sigma$  to have a permutation-symmetric structure seen in the i.i.d. uniform case. For example, for  $k = 3$ ,  $\Sigma$  is, a rescaled version of,

$$(12) \quad \Sigma_u := \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix},$$

and so (up to a multiplicative constant) and with  $k = k_1 = 3$ ,  $p_{\max} = p_1 = 1/3$ , (8) continues to hold for cyclic Markov chains. For  $k \geq 4$ , the cyclicity constraint on  $P$  forces

$\Sigma$  to be cyclic but does no longer force the permutation-symmetric structure, and, say, for  $k = 4$ ,  $\Sigma$  might differ from, a rescaled version of,

$$(13) \quad \Sigma_u := \begin{pmatrix} 1 & -1/3 & -1/3 & -1/3 \\ -1/3 & 1 & -1/3 & -1/3 \\ -1/3 & -1/3 & 1 & -1/3 \\ -1/3 & -1/3 & -1/3 & 1 \end{pmatrix}.$$

In fact, if

$$(14) \quad P = \begin{pmatrix} p_1 & p_2 & p_3 & p_4 \\ p_4 & p_1 & p_2 & p_3 \\ p_3 & p_4 & p_1 & p_2 \\ p_2 & p_3 & p_4 & p_1 \end{pmatrix},$$

then  $\Sigma = (\sigma_{i,j})_{1 \leq i,j \leq 4}$  is a rescaled version of  $\Sigma_u$  if and only if  $p_3^2 = p_2 p_4$ . Nevertheless, see [13], for  $k \geq 2$ , and if  $\sigma = \sigma_{i,i}$ ,  $i = 1, \dots, k$ ,

$$(15) \quad \frac{\text{LI}_N - N/k}{\sigma\sqrt{N}} \implies \sup \left\{ \Delta_\pi(\tilde{B}); \pi \in \mathcal{P}, \pi \leq k \right\} = \max_{0=t_0 \leq \dots \leq t_k=1} \sum_{j=1}^k \left( \tilde{B}_j(t_j) - \tilde{B}_j(t_{j-1}) \right).$$

Assuming that in addition to be cyclic,  $P$  is also symmetric (for  $k = 2$  the cyclic and symmetric assumptions are the same, and Chistyakov and Götze [7], see also [14], showed that the corresponding limiting law is the maximal eigenvalue of the  $2 \times 2$  traceless GUE) a diagonalization argument, combined with (15), leads to the following result.

**Proposition 11.** *Let  $P := (p_{i,j})_{1 \leq i,j \leq k}$  be cyclic and symmetric, i.e.,  $P = (p(j-i))_{1 \leq i,j \leq k}$ , where  $p$  is a  $k$ -periodic function defined on  $\mathbb{Z}$  such that  $p(r) = p(-r)$ , for all  $r \in \mathbb{Z}$ . Let*

$$(16) \quad \lambda_\ell := \sum_{r=1}^k p(r) \cos(2\pi(\ell-1)r/k) \quad (1 \leq \ell \leq k),$$

and let  $(B_j)_{j=2,\dots,k}$  be a  $(k-1)$ -dimensional standard Brownian motion on  $[0, 1]$ . Then,

$$(17) \quad \frac{\text{LI}_N - N/k}{\sigma\sqrt{N}} \implies \max_{0=t_0 \leq \dots \leq t_k=1} \left\{ \sqrt{\frac{2}{k}} \sum_{j=1}^k \sum_{r=1}^{\lfloor \frac{k-1}{2} \rfloor} \sqrt{\frac{1+\lambda_{r+1}}{1-\lambda_{r+1}}} \left( \cos\left(\frac{2\pi jr}{k}\right) (B_{2r}(t_j) - B_{2r}(t_{j-1})) + \sin\left(\frac{2\pi jr}{k}\right) (B_{2r+1}(t_j) - B_{2r+1}(t_{j-1})) \right) - \frac{1}{\sqrt{k}} \sqrt{\frac{1+\lambda_{\frac{k}{2}+1}}{1-\lambda_{\frac{k}{2}+1}}} \left( 2 \sum_{j=1}^k (B_k(t_j) - B_k(t_{j-1})) - B_k(1) \right) \right\},$$

where the last term, above, is only present for  $k$  even.

*Proof.* Since  $P$  is symmetric, it can be diagonalized as  $P = S\Lambda S^\top$ , where  $\Lambda$  is the diagonal matrix formed with its eigenvalues  $(\lambda_\ell)_{1 \leq \ell \leq k}$  (we will see below that these are the quantities defined at (16)) and where  $S$  is a matrix formed by the orthonormal column eigenvectors  $(u_\ell)_{1 \leq \ell \leq k}$  where  $u_1^\top = (1/\sqrt{k}, \dots, 1/\sqrt{k})$ . Next, by Theorem 4.3 in [13],  $\Sigma$ , the covariance matrix of the  $k$ -dimensional correlated Brownian motion  $\tilde{B}$ , is given by  $\Sigma = S\Lambda_\Sigma S^\top$ , where  $\Lambda_\Sigma$  is the diagonal matrix with diagonal entries  $0, (1 + \lambda_2)/(1 - \lambda_2), \dots, (1 + \lambda_k)/(1 - \lambda_k)$ . Therefore,  $\tilde{B} = S\sqrt{\Lambda_\Sigma}B$ , where now  $B$  is a standard  $k$ -dimensional Brownian motion. Next, the symmetric and cyclic structures imply that the eigenvalues of  $P$  are in fact the  $\lambda_\ell$ 's defined at (16):  $\lambda_\ell = \sum_{r=1}^k p(r) \cos(2\pi(\ell - 1)r/k)$ ,  $1 \leq \ell \leq k$ , (clearly they are not all simple since  $\lambda_\ell = \lambda_{k-\ell+2}$ ,  $\ell = 2, \dots, k$ ). The corresponding orthonormal column eigenvectors are

$$v_\ell := (v_{j,\ell})_{1 \leq j \leq k} = (\sqrt{2} \cos(2\pi(\ell - 1)j/k)/\sqrt{k})_{1 \leq j \leq k}, \quad \ell = 1, 2, \dots, \lfloor k/2 \rfloor + 1,$$

and

$$w_\ell := (w_{j,\ell})_{1 \leq j \leq k} = (\sqrt{2} \sin(2\pi(\ell - 1)j/k)/\sqrt{k})_{1 \leq j \leq k}, \quad \ell = 2, 3, \dots, \lfloor (k - 1)/2 \rfloor + 1.$$

Clearly,  $v_1 = u_1$  is an eigenvector corresponding to the simple eigenvalue 1, while if  $k$  is even,  $v_{(k/2)+1} = (1/\sqrt{k}, -1/\sqrt{k}, \dots, 1/\sqrt{k}, -1/\sqrt{k})$  is an eigenvector corresponding to the simple eigenvalue  $\sum_{r=1}^k p(r) \cos(2\pi(k/2 + 1 - 1)r/k) = \sum_{r=1}^k (-1)^r p(r)$ . Moreover, for  $\ell = 2, 3, \dots, \lfloor (k - 1)/2 \rfloor + 1$ ,  $v_\ell$  and  $w_\ell$  share the same eigenvalue  $\lambda_\ell$ . Therefore,

$$S = \left( v_1, v_2, w_2, \dots, v_{\lfloor \frac{k-1}{2} \rfloor + 1}, w_{\lfloor \frac{k-1}{2} \rfloor + 1}, v_{\frac{k}{2} + 1} \right),$$

where, above, the last column is only present if  $k$  is even. Next, from the transformation  $\tilde{B} = S\sqrt{\Lambda_\Sigma}B$ , and since  $(\sqrt{\Lambda_\Sigma}B)_\ell = \sqrt{(1 + \lambda_{\lfloor \ell/2 \rfloor + 1})/(1 - \lambda_{\lfloor \ell/2 \rfloor + 1})}B_\ell$ ,  $\ell = 2, \dots, k$ , and  $(\sqrt{\Lambda_\Sigma}B)_1 = 0$ , then for  $j = 1, \dots, k$ ,

$$(18) \quad \tilde{B}_j = \sum_{\ell=2}^k u_{j,\ell} \sqrt{\frac{1 + \lambda_{\lfloor \frac{\ell}{2} \rfloor + 1}}{1 - \lambda_{\lfloor \frac{\ell}{2} \rfloor + 1}}} B_\ell,$$

where  $u_{j,\ell} = v_{j, \lfloor \ell/2 \rfloor + 1}$  or  $u_{j,\ell} = w_{j, \lfloor \ell/2 \rfloor + 1}$ , for  $\ell$  even or odd. Therefore, for  $j = 1, \dots, k$ ,

$$(19) \quad \tilde{B}_j = \sqrt{\frac{2}{k}} \sum_{r=1}^{\lfloor \frac{k-1}{2} \rfloor} \sqrt{\frac{1 + \lambda_{r+1}}{1 - \lambda_{r+1}}} \left( \cos\left(\frac{2\pi r j}{k}\right) B_{2r} + \sin\left(\frac{2\pi r j}{k}\right) B_{2r+1} \right) \\ + \frac{(-1)^{j+1}}{\sqrt{k}} \sqrt{\frac{1 + \lambda_{\frac{k}{2}+1}}{1 - \lambda_{\frac{k}{2}+1}}} B_k,$$

where the last term on the right of (19) is only present for  $k$  even. With (19), the sum on the right hand side of (15) becomes:

$$(20) \quad \sqrt{\frac{2}{k}} \sum_{j=1}^k \sum_{r=1}^{\lfloor \frac{k-1}{2} \rfloor} \sqrt{\frac{1+\lambda_{r+1}}{1-\lambda_{r+1}}} \left( \cos\left(\frac{2\pi jr}{k}\right) B_{2r}(t_j) + \sin\left(\frac{2\pi jr}{k}\right) B_{2r+1}(t_j) \right. \\ \left. - \cos\left(\frac{2\pi jr}{k}\right) B_{2r}(t_{j-1}) - \sin\left(\frac{2\pi jr}{k}\right) B_{2r+1}(t_{j-1}) \right) \\ + \frac{1}{\sqrt{k}} \sqrt{\frac{1+\lambda_{\frac{k}{2}+1}}{1-\lambda_{\frac{k}{2}+1}}} \sum_{j=1}^k (-1)^{j+1} (B_k(t_j) - B_k(t_{j-1})),$$

an expression only involving standard Brownian motions and where, again, the last term

$$(21) \quad \frac{1}{\sqrt{k}} \sqrt{\frac{1+\lambda_{\frac{k}{2}+1}}{1-\lambda_{\frac{k}{2}+1}}} \sum_{j=1}^k (-1)^{j+1} (B_k(t_j) - B_k(t_{j-1})) \\ = \frac{-1}{\sqrt{k}} \sqrt{\frac{1+\lambda_{\frac{k}{2}+1}}{1-\lambda_{\frac{k}{2}+1}}} \left( B_k(1) + \sum_{j=1}^{k-1} 2(-1)^j B_k(t_j) \right) \\ (22) \quad = \frac{-1}{\sqrt{k}} \sqrt{\frac{1+\lambda_{\frac{k}{2}+1}}{1-\lambda_{\frac{k}{2}+1}}} \left( 2 \sum_{j=1}^k (B_k(t_j) - B_k(t_{j-1})) - B_k(1) \right),$$

is only present if  $k$  is even. □

**Remark 12.** Let us specialize the previous results to instances where further simplifications and identifications occur.

(i) For  $k = 3$ , and up to the factor  $\sqrt{2(1+\lambda_2)/(k(1-\lambda_2))} = \sqrt{2(1+3p_1)/(3(3-3p_1))}$ , the right-hand side of (17) becomes

$$\max_{0=t_0 \leq t_1 \leq t_2 \leq t_3=1} \sum_{j=1}^3 \left( \cos\left(\frac{2\pi j}{3}\right) (B_2(t_j) - B_2(t_{j-1})) + \sin\left(\frac{2\pi j}{3}\right) (B_3(t_j) - B_3(t_{j-1})) \right) \\ = \max_{0 \leq t_1 \leq t_2 \leq 1} \left( B_2(1) + \sqrt{3} B_3(t_1) - \frac{\sqrt{3}}{2} B_3(t_2) - \frac{3}{2} B_2(t_2) \right) \\ \stackrel{\text{law}}{=} \sqrt{\frac{2}{3}} \max_{0 \leq t_1 \leq t_2 \leq t_3=1} \sum_{j=1}^2 \left( -\sqrt{\frac{j}{j+1}} B_j(t_{j+1}) + \sqrt{\frac{j}{j+1}} B_j(t_j) \right) \\ \stackrel{\text{law}}{=} \sqrt{\frac{2}{3}} \left( \max_{0=t_0 \leq t_1 \leq t_2 \leq t_3=1} \sum_{j=1}^3 (B_j(t_j) - B_j(t_{j-1})) - \frac{1}{3} \sum_{j=1}^3 B_j(1) \right),$$

where the last equality, in law, follows either by using, in (15), the simple linear transformation

$$(23) \quad \tilde{B}_j = \sqrt{\frac{2(1+\lambda_2)}{3(1-\lambda_2)}} \left( \sqrt{\frac{2}{3}} B_j - \sqrt{\frac{1}{6}} \sum_{i=1, i \neq j}^3 B_i \right), \quad j = 1, 2, 3,$$

which, by comparing covariances, is easily verified from (19); or, still by comparing covariances, by arguments such as those in the proof of Theorem 3.2 in [12]. Therefore, with the help of Theorem 1, and up to a scaling factor, the limiting law of  $\text{LI}_N$  is that of the maximal eigenvalue of the  $3 \times 3$  traceless GUE.

(ii) For  $k = 4$ ,  $\Sigma$ , the covariance matrix of  $\tilde{B} = (\tilde{B}_j)_{j=1, \dots, 4}$  is given, up to a scaling constant, by:

$$(24) \quad \Sigma := \begin{pmatrix} 2\eta_2 + \eta_3 & -\eta_3 & -2\eta_2 + \eta_3 & -\eta_3 \\ -\eta_3 & 2\eta_2 + \eta_3 & -\eta_3 & -2\eta_2 + \eta_3 \\ -2\eta_2 + \eta_3 & -\eta_3 & 2\eta_2 + \eta_3 & -\eta_3 \\ -\eta_3 & -2\eta_2 + \eta_3 & -\eta_3 & 2\eta_2 + \eta_3 \end{pmatrix}.$$

where  $\eta_2 = (1 + \lambda_2)/(1 - \lambda_2)$ ,  $\lambda_2 = p_1 - p_3$ , and  $\eta_3 = (1 + \lambda_3)/(1 - \lambda_3)$ ,  $\lambda_3 = p_1 - 2p_2 + p_3$ . Clearly,  $\Sigma$  can differ from  $\Sigma_u$ , e.g., let  $2\eta_2 = \eta_3$ , i.e., let

$$P = \begin{pmatrix} p_1 & p_2 & \frac{p_2(1-2p_2)}{1+2p_2} & p_2 \\ p_2 & p_1 & p_2 & \frac{p_2(1-2p_2)}{1+2p_2} \\ \frac{p_2(1-2p_2)}{1+2p_2} & p_2 & p_1 & p_2 \\ p_2 & \frac{p_2(1-2p_2)}{1+2p_2} & p_2 & p_1 \end{pmatrix}.$$

Then, and up to the multiplicative constant  $4\eta_2$ ,  $\Sigma$  becomes:

$$\Sigma := \begin{pmatrix} 1 & -1/2 & 0 & -1/2 \\ -1/2 & 1 & -1/2 & 0 \\ 0 & -1/2 & 1 & -1/2 \\ -1/2 & 0 & -1/2 & 1 \end{pmatrix},$$

which is clearly different from, a rescaled version of, (13). In fact, if  $\Sigma = \Sigma_u$ , then clearly

$$(25) \quad \tilde{B}_j = \frac{\sqrt{3}}{2} B_j - \frac{1}{2\sqrt{3}} \sum_{i=1, i \neq j}^4 B_i, \quad j = 1, 2, 3, 4.$$

Conversely, and clearly, for a linear transformation such as

$$\tilde{B}_j = \alpha_j B_j - \sum_{i=1, i \neq j}^4 \beta_i B_i, \quad j = 1, 2, 3, 4,$$

to lead to  $\Sigma$ , one needs  $\Sigma$  to be permutation-symmetric and, up to a multiplicative constant, the right-hand side of (17) becomes equal in law to

$$\max_{0=t_0 \leq t_1 \leq t_2 \leq t_3 \leq t_4=1} \sum_{j=1}^4 (B_j(t_j) - B_j(t_{j-1})) - \frac{1}{4} \sum_{j=1}^4 B_j(1),$$

and corresponds to the matrix  $P$  in (14) with  $p_2 = p_3 = p_4$ .

(iii) Finally, it is easy to see that the properties just described continue to hold for arbitrary dimension  $k \geq 4$ . In arbitrary dimension, if  $\Sigma = \Sigma_u$  (the  $k$ -dimensional version of the matrix defined at (12) and (13)), then the linear transformation corresponding to (25) is given by

$$\tilde{B}_j = \sqrt{\frac{k-1}{k}} B_j - \sqrt{\frac{1}{k(k-1)}} \sum_{i=1, i \neq j}^k B_i, \quad j = 1, \dots, k,$$

Conversely, for a linear transformation such as

$$\tilde{B}_j = \alpha_j B_j - \sum_{i=1, i \neq j}^k \beta_i B_i, \quad j = 1, 2, \dots, k,$$

to lead to  $\Sigma$ , one needs  $\Sigma$  to be permutation-symmetric. In either instance, and up to a multiplicative constant, the right-hand side of (17) has the same law as

$$\max_{0=t_0 \leq \dots \leq t_k=1} \sum_{j=1}^k (B_j(t_j) - B_j(t_{j-1})) - \frac{1}{k} \sum_{j=1}^k B_j(1),$$

which, in turn, via Theorem 1, is the law of the maximal eigenvalue of an element of the  $k \times k$  traceless GUE.

## REFERENCES

- [1] G. Anderson, A. Guionnet, O. Zeitouni *An Introduction to Random Matrices*. Cambridge studies in advanced mathematics, 118 (2009).
- [2] Y. Baryshnikov *GUEs and queues*, Probab. Theory Relat. Fields 119, 256–274 (2001).
- [3] P. Biane, P. Bougerol, N. O’Connell *Littelman paths and Brownian paths*, Duke Math. J., vol. 130, pp. 127–167, 2005.
- [4] P. Billingsley *Convergence of probability measures* Wiley, 1999.
- [5] P. Bougerol, T. Jeulin *Paths in Weyl chambers and random matrices* Probab. Theory Related Fields, vol. 124, no. 4, pp. 517–543, 2002.
- [6] A. Bufetov *A central limit theorem for the extremal characters of the infinite symmetric group*, Functional Analysis and Its Applications, vol. 46, no. 2, pp. 83–93, 2012.
- [7] G.P. Chistyakov, F. Götze *Distribution of the shape of Markovian random words*, Probab. Theory Relat. Fields, vol. 129, no. 1, pp. 18–36, 2004.
- [8] Y. Doumerc *A note on representations of classical Gaussian matrices* Séminaire de Probabilités XXXVII., Lecture Notes in Math., No. 1832, Springer, Berlin, (2003), 370–384.
- [9] W. Fulton *Young Tableaux*, CUP, Cambridge, (1997).



- [10] P. Glynn, W. Whitt *Departures form many queues in series*, Ann. Appl. Probability, 1, 546–572, 1991.
- [11] J. Gravner, C.A. Tracy, H. Widom *Limit theorems for height fluctuations in a class of discrete space and time growth models*, J. Stat. Phys. 102, (2001), 1085–1132.
- [12] C. Houdré, T. Litherland *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.
- [13] C. Houdré, T. Litherland *On the limiting shape of Young diagrams associated with Markov random words*, arXiv 1110.4570, (2011).
- [14] C. Houdré, T. Litherland *Asymptotics for the length of the longest increasing subsequence of a binary Markov random word*, in: Malliavin calculus and stochastic analysis: A Festschrift in honor of David Nualart, New York, NY: Springer. Springer Proceedings in Mathematics & Statistics 34, 511-524, 2013.
- [15] C. Houdré, H. Xu *On the limiting shape of Young diagrams associated with inhomogeneous random words*, in: High Dimensional Probability VI: The Banff volume Progress in Probability, 66, Birkhauser (2013)
- [16] A. Its, C.A. Tracy, H. Widom *Random words, Toeplitz determinants, and integrable systems. I. Random matrix models and their applications*, 245–258, Math. Sci. Res. Inst. Publ., 40, Cambridge Univ. Press, Cambridge, 2001.
- [17] A. Its, C.A. Tracy, H. Widom *Random words, Toeplitz determinants, and integrable systems. II. Advances in nonlinear mathematics and science. Phys. D.*, vol. 152-153 (2001), 199–224.
- [18] K. Johansson *Shape fluctuations and random matrices* Comm. Math. Phys. 209, (2000), 437–476.
- [19] K. Johansson *Discrete orthogonal polynomial ensembles and the Plancherel measure*. Ann. of Math. (2) 153 (2001), no. 1, 259–296.
- [20] K. Johansson *Non-intersecting paths, random tilings and random matrices*, Probab. Theor. Rel. Fields, 123, (2002), 225–280.
- [21] K. Johansson, E. Nordenstam *Eigenvalues of GUE minors*, Elec. J. Probab. Vol. 11 (2006), Paper no. 50, pages 1342–1371.
- [22] S. Kerov *Asymptotic Representation Theory of the Symmetric Group and its Applications in Analysis*, Vol. 219. AMS, Translations of Mathematical Monographs, 2003. (Russian edition: D. Sci thesis, 1994)
- [23] G. Kuperberg *Random words, quantum statistics, central limits, random matrices*. Methods Appl. Anal. 9 (1), 99–118 (2002).
- [24] P.-L. Méliot *Fluctuations of central measures on partitions* Proceedings of the 24th International Conference on Formal Power series and Algebraic Combinatorics (Nagoya, Japan), pp. 387-398, 2012.
- [25] P. van Moerbeke *Random and Integrable Models in Mathematics and Physics in Random Matrices, Random Processes and Integrable Systems*, edited by J. Harnad, CRM Series in mathematical Physics, 2011.
- [26] N. O’Connell *A path transformation for random walks and the Robinson-Schensted correspondence* Trans. Amer. Math. Soc. 355, (2003), 3669–3697.
- [27] N. O’Connell, M. Yor *A representation for non-colliding random walks* Electron. Comm. Probab., vol. 7, (2002), 1–12.
- [28] C. Tracy, H. Widom *On the distribution of the lengths of the longest increasing monotone subsequences in random words*. Probab. Theor. Rel. Fields. 119 (2001), 350–380.