Bernoulli **16**(2), 2010, 471–492
DOI: 10.3150/09-BEJ218

Asymptotics for random Young diagrams when the word length and alphabet size simultaneously grow to infinity

JEAN-CHRISTOPHE BRETON[1] and CHRISTIAN HOUDRÉ[2]

[1]*Laboratoire de Mathématiques, Image et Applications, Avenue Michel Crépeau, Université de La Rochelle, 17042 La Rochelle cedex, France. E-mail: jcbreton@univ-lr.fr*
[2]*School of Mathematics, 686 Cherry Street, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA. E-mail: houdre@math.gatech.edu*

Given a random word of size n whose letters are drawn independently from an ordered alphabet of size m, the fluctuations of the shape of the random RSK Young tableaux are investigated, when n and m converge together to infinity. If m does not grow too fast and if the draws are uniform, then the limiting shape is the same as the limiting spectrum of the GUE. In the non-uniform case, a control of both highest probabilities will ensure the convergence of the first row of the tableau toward the Tracy–Widom distribution.

Keywords: GUE; longest increasing subsequence; random words; strong approximation; Tracy–Widom distribution; Young tableaux

1. Introduction and results

Let $\mathcal{A}_m = \{\alpha_1 < \alpha_2 < \cdots < \alpha_m\}$ be an ordered alphabet of size m and let a word be made of the random letters $X_1^m, \ldots, X_n^m$ (independently) drawn from $\mathcal{A}_m$. Recall that the Robinson–Schensted–Knuth (RSK) correspondence associates to a (random) word a pair of (random) Young tableaux of the same shape, having at most m rows (see, e.g., [8] or [21]). It is then well known that the length, $V_1(n,m)$, of the top row of these tableaux coincides with the length of the longest (weakly) increasing subsequence of $X_1^m, \ldots, X_n^m$. The behavior of $V_1(n,m)$ when n and/or m go to $+\infty$ and its connections to various areas of mathematics (e.g., random matrices, queueing theory, percolation theory) have been investigated in numerous papers ([1,2,4,9,15–17,25], ...). For instance, appropriately renormalized and for uniform draws, $V_1(n,m)$ converges in law, as n goes to infinity and m is fixed, to the largest eigenvalue of an $m \times m$ matrix from the traceless Gaussian Unitary Ensemble (GUE). More generally (see [17]), when $n \to +\infty$ (and m is fixed), the shape of the whole Young tableaux associated to a uniform random word converges, after renormalization, to the law of the spectrum of an $m \times m$ traceless GUE matrix. For different random words, such as non-uniform or Markovian ones, the situation is more involved ([6,13–16]).

For independently and uniformly drawn random words, the following result holds, where, below and in the sequel, "$\Rightarrow$" stands for convergence in distribution.

1350-7265

Theorem 1. *Let $V_k(n,m) = \sum_{i=1}^k R_n^i$ be the sum of the lengths R_n^i of the first k rows of the Young tableau. Then*

$$\left(\frac{V_k(n,m) - kn/m}{\sqrt{n}}\right)_{1 \le k \le m} \Rightarrow \frac{\sqrt{m-1}}{m}\left(\max_{\mathbf{t} \in I_{k,m}} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left(\hat{B}^l(t_{j,l}) - \hat{B}^l(t_{j,l-1})\right)\right)_{1 \le k \le m}, \quad (1)$$

where $(\hat{B}^1, \ldots, \hat{B}^m)$ is a multidimensional Brownian motion with covariance matrix having diagonal terms equal to 1 *and off-diagonal terms equal to* $-1/(m-1)$, *and where $I_{k,m}$ is defined by*

$$\begin{aligned} I_{k,m} = \{\mathbf{t} = (t_{j,l} : 1 \le j \le k, 0 \le l \le m) : t_{j,j-1} = 0, t_{j,m-k+j} = 1, 1 \le j \le k, \\ t_{j,l-1} \le t_{j,l}, 1 \le j \le k, 1 \le l \le m-1; t_{j,l} \le t_{j-1,l}, 2 \le j \le k, 1 \le l \le m-1\}. \end{aligned}$$

Here, and in the sequel, the rows beyond the height of the tableau are considered to be of length zero. If we let $\mathbf{\Theta}_k : \mathbb{R}^k \to \mathbb{R}^k$ be defined via $(\mathbf{\Theta}_k(\mathbf{x}))_j = \sum_{i=1}^j x_i$, $1 \le j \le k$, then the shape of the Young tableau is given by $\mathbf{\Theta}_m^{-1}((V_1(n,m), \ldots, V_m(n,m))^t) = (R_n^1, \ldots, R_n^m)^t$. Moreover, let $(\lambda_{\mathrm{GUE},m}^{1,0}, \lambda_{\mathrm{GUE},m}^{2,0}, \ldots, \lambda_{\mathrm{GUE},m}^{m,0})$ be the spectrum, written in non-increasing order, of an $m \times m$ traceless element of the GUE, where the GUE is equipped with the measure

$$\frac{1}{C_m} \prod_{1 \le i < j \le m} (x_i - x_j)^2 \prod_{j=1}^{m} \mathrm{e}^{-x_j^2/2}$$

and $C_m = (2\pi)^{m/2} \prod_{j=1}^m j!$ (see [19]). An important fact (see [3,5,7,10,13,20]) asserts that

$$\begin{aligned} &\frac{\sqrt{m-1}}{\sqrt{m}} \mathbf{\Theta}_m^{-1}\left(\left(\max_{\mathbf{t} \in I_{k,m}} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left(\hat{B}^l(t_{j,l}) - \hat{B}^l(t_{j,l-1})\right)\right)_{1 \le k \le m}\right) \\ &\stackrel{\mathcal{L}}{=} (\lambda_{\mathrm{GUE},m}^{1,0}, \lambda_{\mathrm{GUE},m}^{2,0}, \ldots, \lambda_{\mathrm{GUE},m}^{m,0}). \end{aligned} \quad (2)$$

In fact, if $(\lambda_{\mathrm{GUE},m}^{1}, \lambda_{\mathrm{GUE},m}^{2}, \ldots, \lambda_{\mathrm{GUE},m}^{m})$ is the (ordered) spectrum of an $m \times m$ element of the GUE, then

$$(\lambda_{\mathrm{GUE},m}^{1}, \lambda_{\mathrm{GUE},m}^{2}, \ldots, \lambda_{\mathrm{GUE},m}^{m}) \stackrel{\mathcal{L}}{=} (\lambda_{\mathrm{GUE},m}^{1,0}, \lambda_{\mathrm{GUE},m}^{2,0}, \ldots, \lambda_{\mathrm{GUE},m}^{m,0}) + Z_m e_m, \quad (3)$$

where Z_m is a centered Gaussian random variable with variance $1/m$, independent of the vector $(\lambda_{\mathrm{GUE},m}^{1,0}, \lambda_{\mathrm{GUE},m}^{2,0}, \ldots, \lambda_{\mathrm{GUE},m}^{m,0})$, and where $e_m = (1, 1, \ldots, 1)$; see [14] for simple proofs of (2) and (3).

Finally, recall that, as $m \to +\infty$, the asymptotic behavior of the spectrum of the GUE has been obtained by Tracy and Widom (see [23], [24] and also Theorem 1.4 in [17], with a slight change of the notation).

Theorem 2. *For each $r \geq 1$, there is a distribution $\mathbf{F}_r$ on $\mathbb{R}^r$ such that*

$$\big(m^{1/6}\big(\lambda^k_{\mathrm{GUE},m} - 2\sqrt{m}\big)\big)_{1\leq k\leq r} \Rightarrow \mathbf{F}_r, \qquad m \to +\infty. \tag{4}$$

Remark 3. The distribution $\mathbf{F}_r$ is explicitly known (see (3.48) in [17]) and its first marginal coincides with the Tracy–Widom distribution.

Since $Z_m m^{1/6} \Rightarrow 0$ as $m \to +\infty$, taking successively the limits in n and then in m, (1)–(4) entail, for each $r \geq 1$, that

$$\lim_{m\to+\infty}\lim_{n\to+\infty}\left(\frac{V_k(n,m) - kn/m - 2k\sqrt{n}}{\sqrt{n}} \times m^{2/3}\right)_{1\leq k\leq r} = \mathbf{F}_r \boldsymbol{\Theta}_r^{-1}. \tag{5}$$

Following universality argument in percolation models developed by Bodineau and Martin ([4]), we show below that the limits in n and m in (5) can be explicitly taken simultaneously when the size m of the alphabet does not grow too fast with respect to n. Doing so, we are dealing with growing ordered alphabets and, at each step, the n letters X_i^m, $1 \leq i \leq n$, are redrawn (and not just the nth letter as in the case with the model studied in [12]). In a way, we are thus giving the fluctuations of the shape of the Young tableau of a random word when the alphabets are growing and are reshuffled. In the sequel, m will be a function $m(n)$ of n. However, in order to simplify the notation, we shall still write m instead of $m(n)$. A main result of this note is the following.

Theorem 4. *Let m tend to infinity as $n \to +\infty$ in such a way that $m = \mathrm{o}(n^{3/10}(\log n)^{-3/5})$. Then, for each $r \geq 1$,*

$$\left(\frac{V_k(n,m) - kn/m - 2k\sqrt{n}}{n^{1/2}m^{-2/3}}\right)_{1\leq k\leq r} \Rightarrow \mathbf{F}_r \boldsymbol{\Theta}_r^{-1}, \qquad n \to +\infty.$$

Remark 8, below, briefly discusses the growth conditions on m. Since, again, the length of the first row of the Young tableau is the length $V_1(n,m)$ of the longest increasing subsequence and since the first marginal of $\mathbf{F}_r$ is the Tracy–Widom distribution F_{TW}, we have the following result.

Corollary 5. *Let m tend to infinity as $n \to +\infty$ in such a way that $m = \mathrm{o}(n^{3/10}(\log n)^{-3/5})$. Then*

$$\frac{V_1(n,m) - (n/m) - 2n^{1/2}}{n^{1/2}m^{-2/3}} \Rightarrow F_{\mathrm{TW}}, \qquad n \to +\infty.$$

When the independent random letters are no longer uniformly drawn, a similar asymptotic behavior continues to hold for $V_1(n,m)$, as explained next. Let the X_i^m, $1 \leq i \leq n$, be independently and identically distributed with $\mathbb{P}(X_1^m = \alpha_j) = p_j^m$, let $p_{\max}^m = \max_{1\leq j\leq m} p_j^m$ and also let $J(m) = \{j : p_j^m = p_{\max}^m\} = \{j_1, \ldots, j_{k(m)}\}$ with $k(m) = \mathrm{card}\,(J(m))$. Now, from [11] and as $n \to +\infty$, the behavior of the first row of the Young tableau in this non-uniform setting is given by

$$\frac{V_1(n,m)-p_{\max}^m n}{\sqrt{p_{\max}^m n}} \Rightarrow \frac{\sqrt{1-k(m)p_{\max}^m}-1}{k(m)}\sum_{j=1}^{k(m)} B^j(1) \\ + \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{k(m)-1}\le t_{k(m)}=1}} \sum_{l=1}^{k(m)}\big(B^l(t_l)-B^l(t_{l-1})\big), \tag{6}$$

where $(B^1,\dots,B^{k(m)})$ is a standard $k(m)$-dimensional Brownian motion. For the limiting behavior in m of the right-hand side of (6), as explained next, two cases can arise, depending on the number of most probable letters in $\mathcal{A}_m$. Setting

$$Z_k=\frac{1}{k}\sum_{j=1}^{k} B^j(1) \quad \text{and} \quad D_k=\max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{k-1}\le t_k=1}} \sum_{l=1}^{k}\big(B^l(t_l)-B^l(t_{l-1})\big),$$

and combining (2), (3) and (4), as well as Remark 3, when $k=1$, and since, clearly, $Z_k \sim \mathcal{N}(0,1/k)$, we have

$$k^{1/6}\big(D_k-2\sqrt{k}\big) \Rightarrow F_{\mathrm{TW}}, \qquad k\to+\infty. \tag{7}$$

First, let $k(m)$ be bounded. Eventually extracting a subsequence, we can assume that $k(m)$ is equal to a fixed $k\in\mathbb{N}\setminus\{0\}$ and since $p_{\max}^m\in[0,1]$, we can also assume that $p_{\max}^m\to p_{\max}$. In this case, taking the limit first in n and next in m yields

$$\frac{V_1(n,m)-p_{\max}^m n}{\sqrt{p_{\max}^m n}} \Rightarrow \big(\sqrt{1-kp_{\max}}-1\big)Z_k+D_k. \tag{8}$$

The limiting distribution on the right-hand side of (8) depends on k. For instance, for $k=1$, we recover a Gaussian distribution, while for $k>1$ and specific choice of the $p_{\max}^m$ for which $\lim_{m\to+\infty} p_{\max}^m=0$, we recover (8) without the Gaussian term. Thus, in general, when $k(m)$ is bounded, there is no global asymptotics, but only convergence (to different distributions) along subsequences.

Next, let $k(m)\to+\infty$. In this case, in (6), the Gaussian contribution is negligible. Indeed, since $(\sqrt{1-k(m)p_{\max}^m}-1)^2 k(m)^{-2/3}\le (k(m)p_{\max}^m)^2 k(m)^{-2/3}\le k(m)^{-2/3}\to 0$, when $m\to+\infty$,

$$\big(\sqrt{1-k(m)p_{\max}^m}-1\big)Z_{k(m)}k(m)^{1/6}\sim\mathcal{N}\big(0,\big(\sqrt{1-k(m)p_{\max}^m}-1\big)^2 k(m)^{-2/3}\big)\Rightarrow 0.$$

Hence, plugging the convergence result (7) into (6) leads to

$$\frac{V_1(n,m)-p_{\max}^m n-2\sqrt{k(m)p_{\max}^m n}}{\sqrt{k(m)p_{\max}^m n}}k(m)^{2/3}\Rightarrow F_{\mathrm{TW}}, \tag{9}$$

where the limit is first taken as $n \to +\infty$ and then as $m \to +\infty$. In this non-uniform setting, we have the following counterpart to Corollary 5 with an additional control on the second largest probability for the letters of $\mathcal{A}_m$. More precisely, let $p_{2\mathrm{nd}}^m = \max(p_j^m < p_{\max}^m : 1 \le j \le m)$.

Theorem 6. *Let the size m of the alphabets vary with n and assume that $k(m(n))$, the number of most probable letters in $\mathcal{A}_m$, goes to infinity when $n \to +\infty$, in such a way that $k(m(n)) = \mathrm{o}(n^{3/10}(\log n)^{-3/5})$. Assume, moreover, that*

$$\bigl(p_{2\mathrm{nd}}^{m(n)}\bigr)^2 \frac{n^{11/10}}{(\log n)^{1/5}} = \mathrm{o}\bigl(p_{\max}^{m(n)}\bigr). \tag{10}$$

Then

$$\frac{V_1(n, m(n)) - p_{\max}^{m(n)} n - 2\sqrt{k(m(n)) p_{\max}^{m(n)} n}}{\sqrt{k(m(n)) p_{\max}^{m(n)} n}} k(m(n))^{2/3} \Rightarrow F_{\mathrm{TW}}. \tag{11}$$

Let us again stress the fact that in the previous result, m is a function of n, with the only requirement being that $k(m(n)) = \mathrm{o}(n^{3/10}(\log n)^{-3/5})$. Note that in the uniform case, $k(m) = m$ and $p_{\max}^m = 1/m$ and that, in general, $1/m \le p_{\max}^m \le 1/k(m)$.

Let us now put our results in context, relate them to the current literature and also describe the main steps in the arguments developed below.

Bodineau and Martin [4] showed that the fluctuations of the last-passage directed percolation model with Gaussian i.i.d. weights actually extend to i.i.d. weights with finite $(2+r)$th moment, $r > 0$. Their arguments rely, in part, on a KMT approximation which was already used by Glynn and Whitt [9] in a related queueing model.

Here, we closely follow [4] and take advantage of the representation (2) of the spectrum of a matrix in the GUE. Using Brownian scaling in those Brownian functionals, we can mix together n and m in the corresponding limit (4) (see (14) below). Then, exhibiting an expression similar to (2), but with *dependent* Bernoulli random variables, for the shape of the Young tableau (see (17)), we show via a Gaussian approximation that the Bernoulli functionals stay close to the Brownian functionals (see (19)), so as to share the same asymptotics.

Since we apply a Gaussian approximation to Bernoulli random variables with a strong integrability property, the strong approximation can be made more precise than in [4]. However, this is not enough to obtain the fluctuations for m of larger order. Actually, the Gaussian approximation is responsible for the condition $m = \mathrm{o}(n^{3/10}(\log n)^{-3/5})$, which falls short of the corresponding polynomial order condition $m = \mathrm{o}(n^{3/7})$ obtained in [4]. However, in contrast to [4], the stronger integrability property of the Bernoulli random variables and the stronger condition on m are required to control the constants appearing in the Gaussian approximation applied to a triangular scheme of different distributions.

Using Skorokhod embedding, Baik and Suidan [2] derived, independently of [4], similar convergence results (see Theorem 2 in [2]), under the condition $m = \mathrm{o}(n^{3/14})$. See also [22] for related results (under $m = \mathrm{o}(n^{1/7})$) in percolation models using functional methods in the CLT.

Finally, note that [2,4,22] deal with percolation models with i.i.d. random variables under enough polynomial integrability. In our setting, the lengths of the rows of the Young tableaux associated to random words are expressed in terms of dependent (exchangeable in the uniform case) Bernoulli random variables. We are thus working with much more specific random variables, but without complete independence.

The paper is organized as follows: Section 2 is devoted to the proof of Theorem 4, while we sketch the changes needed to prove Theorem 6 in Section 3. We conclude in Section 4 with some remarks on the convergence of whole shape of Young tableaux when the draws are non-uniform.

2. Proof of Theorem 4

Brownian scaling. Let $(B^l(s))_{s\geq 0}$, $1\leq l\leq m$, be independent standard Brownian motions. For $s>0$, $m\geq 1$ and $k\geq 1$, let

$$L_k(s,m)=\sup_{\mathbf{t}\in I_{k,m}(s)}\sum_{j=1}^{k}\sum_{l=j}^{m-k+j}\big(B^l(t_{j,l})-B^l(t_{j,l-1})\big), \tag{12}$$

where $I_{k,m}(s)=\{s\mathbf{t},\mathbf{t}\in I_{k,m}\}$. For $k=1$, $L_1(s,m)$ coincides with the Brownian percolation model used in [4]; see also [9] for a related queueing model. For $s=1$, $\sqrt{\frac{m-1}{m}}\boldsymbol{\Theta}_m^{-1}((L_k(1,m))_{1\leq k\leq m})$ has the same law as the spectrum of an $m\times m$ GUE matrix; see [7] and [14].

Since $(L_1(\cdot,m),\dots,L_m(\cdot,m))$ is a continuous function of $B^1,\dots,B^m$, which are independent, Brownian scaling entails that

$$(L_1(s,m),\dots,L_m(s,m))\stackrel{\mathcal{L}}{=}\sqrt{s}(L_1(1,m),\dots,L_m(1,m)). \tag{13}$$

Plugging (13) into (4) yields

$$\left(\frac{L_k(n,m)-2k\sqrt{nm}}{n^{1/2}m^{-1/6}}\right)_{1\leq k\leq r}\Rightarrow \mathbf{F}_r\boldsymbol{\Theta}_r^{-1}. \tag{14}$$

Combinatorics. Let

$$X_{i,j}^m=\begin{cases}1, & \text{if } X_i^m=\alpha_j,\\ 0, & \text{otherwise,}\end{cases}$$

be Bernoulli random variables with parameter $\mathbb{P}(X_i^m=\alpha_j)=1/m$ and variance $\sigma_m^2=(1/m)(1-1/m)$. For a fixed $1\leq j\leq m$, the $X_{i,j}^m$s are independent and identically distributed, while for $j\neq j'$, $(X_{1,j}^m,\dots,X_{n,j}^m)$ and $(X_{1,j'}^m,\dots,X_{n,j'}^m)$ are identically distributed, but no longer independent.

Again, recall that the length of the first row of the Young tableau of a random word is the length of the longest (weakly) increasing subsequence of $X_1^m,\dots,X_n^m$.

Let $S_k^{m,j}=\sum_{i=1}^k X_{i,j}^m$ be the number of occurrences of α_j among $(X_i^m)_{1\leq i\leq k}$. An increasing subsequence of $(X_i^m)_{1\leq i\leq k}$ consists of successive blocks, each one made of an identical letter, with the sequence of letters representing each block being strictly increasing. Since, for $1\leq k<$

$l \le n$ the number of occurrences of α_j among $(X_i^m)_{k\le i\le l}$ is $S_l^{m,j} - S_k^{m,j}$, it follows that

$$V_1(n,m) = \max_{\substack{0=l_0\le l_1\le\cdots\\ \le l_{m-1}\le l_m=n}} [(S_{l_1}^{m,1} - S_0^{m,1}) + (S_{l_2}^{m,2} - S_{l_1}^{m,2}) + \cdots + (S_n^{m,m} - S_{l_{m-1}}^{m,m})], \tag{15}$$

with the convention that $S_0^{m,1} = 0$. More involved combinatorial arguments yield the following expression for $V_k(n,m)$ (see Theorem 5.1 in [13]):

$$V_k(n,m) = \max_{\mathbf{k}\in J_{k,m}(n)} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} (S_{k_{j,l}}^{m,l} - S_{k_{j,l-1}}^{m,l}), \tag{16}$$

where

$$\begin{aligned} J_{r,m}(n) = \{\mathbf{k} = (k_{j,l} : 1 \le j \le r, 0 \le l \le m) : k_{j,j-1} = 0, k_{j,m-r+j} = n, 1 \le j \le r,\\ k_{j,l-1} \le k_{j,l}, 1 \le j \le r, 1 \le l \le m-1; k_{j,l} \le k_{j-1,l}, 2 \le j \le r, 1 \le l \le m-1\}. \end{aligned}$$

For $\mathbf{t} \in I_{r,m}(n)$, set $[\mathbf{t}] = ([t_{j,l}] : 1 \le j \le n, 0 \le l \le m) \in J_{r,m}(n)$ and thus

$$V_k(n,m) = \sup_{\mathbf{t}\in I_{k,m}(n)} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \big(S_{[t_{j,l}]}^{m,l} - S_{[t_{j,l-1}]}^{m,l}\big), \tag{17}$$

which is to be compared with (12) for Brownian functionals.

Centering and reducing. Let $\widetilde{X}_{i,j}^m = (X_{i,j}^m - 1/m)/\sigma_m$ and $\widetilde{S}_k^{m,l} = \sum_{i=1}^k \widetilde{X}_{i,l}^m$, and, replacing $X_{i,j}^m$ by $\widetilde{X}_{i,j}^m$, similarly define $\widetilde{V}_k(n,m)$. Clearly, $V_k(n,m) = \sigma_m \widetilde{V}_k(n,m) + kn/m$, hence,

$$\begin{aligned} &\frac{V_k(n,m) - kn/m - 2k\sqrt{n}}{\sqrt{n}} \times m^{2/3}\\ &\quad= \frac{\sigma_m \widetilde{V}_k(n,m) - 2k\sqrt{n}}{\sqrt{n}} \times m^{2/3}\\ &\quad= \frac{\widetilde{V}_k(n,m) - 2k\sqrt{n}\sigma_m^{-1}}{\sqrt{n}} \times (\sigma_m m^{2/3})\\ &\quad= \frac{\widetilde{V}_k(n,m) - 2k\sqrt{nm} + 2k\sqrt{n}(\sigma_m^{-1} - m^{1/2})}{n^{1/2}m^{-1/6}} \times (m^{1/2}\sigma_m). \end{aligned}$$

Note that $\sigma_m^{-1} - m^{1/2} \sim 1/\sqrt{m}$, and that $m^{1/6}m^{1/2}\sigma_m \sim m^{1/6}$, and so the limit under study is the same as that of

$$\frac{\widetilde{V}_k(n,m) - 2k\sqrt{nm}}{n^{1/2}m^{-1/6}}. \tag{18}$$

Bound. Next, and as in [4], we bound the difference between $\widetilde{V}_k(n,m)$ and $L_k(n,m)$. This bound holds true for any Brownian motions $(B_t^{m,j})_{t\geq 0}$, but it will only be correctly controlled for a special choice of the Brownian motions and for copies of the random variables $\widetilde{X}_{i,j}^m$ given by a coupling (using a strong approximation result, see Proposition 7 below).

$$
\begin{aligned}
&|\widetilde{V}_k(n,m) - L_k(n,m)| \\
&= \left| \sup_{\mathbf{t}\in I_{k,m}(n)} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left(S^{m,l}_{[t_{j,l}]} - S^{m,l}_{[t_{j,l-1}]}\right) - \sup_{\mathbf{t}\in I_{k,m}(n)} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left(B^l(t_{j,l}) - B^l(t_{j,l-1})\right) \right| \\
&\leq \sup_{\mathbf{t}\in I_{k,m}(n)} \left| \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left(S^{m,l}_{[t_{j,l}]} - S^{m,l}_{[t_{j,l-1}]}\right) - \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left(B^l(t_{j,l}) - B^l(t_{j,l-1})\right) \right| \\
&= \sup_{\mathbf{t}\in I_{k,m}(n)} \left| \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left(S^{m,l}_{[t_{j,l}]} - B^l(t_{j,l})\right) - \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left(S^{m,l}_{[t_{j,l-1}]} - B^l(t_{j,l-1})\right) \right| \\
&= \sup_{\mathbf{t}\in I_{k,m}(n)} \Bigg| \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \Big(\left(S^{m,l}_{[t_{j,l}]} - B^l([t_{j,l}])\right) + \left(B^l([t_{j,l}]) - B^l(t_{j,l})\right) \qquad (19) \\
&\qquad\qquad - \left(S^{m,l}_{[t_{j,l-1}]} - B^l([t_{j,l-1}])\right) - \left(B^l([t_{j,l-1}]) - B^l(t_{j,l-1})\right)\Big) \Bigg| \\
&\leq \sup_{\mathbf{t}\in I_{k,m}(n)} \Bigg\{ \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \Big(\left|S^{m,l}_{[t_{j,l}]} - B^l([t_{j,l}])\right| + |B^l([t_{j,l}]) - B^l(t_{j,l})| \\
&\qquad\qquad + \left|S^{m,l}_{[t_{j,l-1}]} - B^l([t_{j,l-1}])\right| + |B^l([t_{j,l-1}]) - B^l(t_{j,l-1})|\Big) \Bigg\} \\
&\leq 2k \sum_{l=1}^{m} (Y_n^{m,l} + W_n^l),
\end{aligned}
$$

where we set

$$
Y_n^{m,l} = \max_{1\leq i\leq n} |S_i^{m,l} - B^l(i)| \quad \text{and} \quad W_n^l = \sup_{\substack{0\leq s,t\leq n \\ |s-t|\leq 1}} |B^l(s) - B^l(t)|.
$$

Gaussian approximation. From now on, we assume that for each n and $l \in [1,m]$ (recall that $m = m(n)$), the random variables $\widetilde{X}_{i,l}^m$, $1 \leq i \leq n$, and the Brownian motion $(B^l(s))_{s\in[0,n+1]}$ appearing in $Y_n^{m,l}$ and W_n^l (rewritten as $(B^{m,l}(s))_{s\in[0,n+1]}$), are given by the following result, which is a compilation of strong approximation results of Komlós, Major and Tusnády and of Sakhanenko, and for which we refer to [18] (Theorem 2.1, Corollary 3.2) and the references therein. In the sequel, we write $B^{m,l}$ and $W_n^{m,l}$, instead of B^l and W_n^l, to insist on the dependence in m of the random variables given by the following proposition.

Proposition 7. *Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. random variables with common distribution F having finite exponential moments. Then, on a common probability space and for every N, one can construct a sequence $(\widetilde{X}_n)_{1\leq n\leq N}$ having the same law as $(X_n)_{1\leq n\leq N}$, and independent Gaussian variables $(Y_n)_{1\leq n\leq N}$ having the same expectations and variances as $(X_n)_{1\leq n\leq N}$, such that for every $x > 0$,*

$$\mathbb{P}\left(\max_{1\leq k\leq N}\left|\sum_{j=1}^{k}\widetilde{X}_j - \sum_{j=1}^{k} Y_j\right| \geq x\right) \leq \left(1 + c_2(F)N^{1/2}\right)\exp(-c_1(F)x),$$

where $c_1(F)$ and $c_2(F)$ are positive constants (depending on F). Moreover, $c_1(F) = c_3\lambda(F)$ and $c_2(F) = \lambda(F)\operatorname{Var}(X_1)^{1/2}$, where c_3 is an absolute constant and $\lambda(F)$ is given by

$$\lambda(F) = \sup\left\{\lambda > 0 : \lambda\mathbb{E}\left[|X_1 - \mathbb{E}[X_1]|^3\exp(\lambda|X_1 - \mathbb{E}[X_1]|)\right] \leq \mathbb{E}\left[|X_1 - \mathbb{E}[X_1]|^2\right]\right\}.$$

The strong approximation entails the following bound for the tail of $Y_n^{m,l}$:

$$\mathbb{P}(Y_n^{m,l} \geq x) \leq \left(1 + c_2(m)n^{1/2}\right)\exp(-c_1(m)x), \tag{20}$$

where $c_1(m) = c_3\lambda(\widetilde{X}_{1,1}^m)$ and $c_2(m) = \lambda(\widetilde{X}_{1,1}^m)\operatorname{Var}(\widetilde{X}_{1,1}^m)^{1/2}$. Observe that $\lambda(\widetilde{X}_{1,1}^m) = \sigma_m \times \lambda(X_{1,1}^m - \mathbb{E}[X_{1,1}^m])$ and note that $\lambda(X_{1,1}^m) \in [2^{-1}, 2]$. Indeed, for $\lambda \geq 2$,

$$\begin{aligned}
\mathbb{E}[|X_{1,1}^m - \mathbb{E}[X_{1,1}^m]|^2] &= \frac{1}{m}\left(1 - \frac{1}{m}\right) \\
&\leq \frac{1}{m}\left(1 - \frac{1}{m}\right)\frac{\lambda}{2} \\
&\leq \frac{1}{m}\left(1 - \frac{1}{m}\right)\left(1 - \frac{2}{m} + \frac{2}{m^2}\right)\lambda \\
&= \lambda\mathbb{E}[|\widetilde{X}_{1,1}^m - \mathbb{E}[X_{1,1}^m]|^3] \\
&\leq \lambda\mathbb{E}[|X_{1,1}^m - \mathbb{E}[X_{1,1}^m]|^3\exp(\lambda|\widetilde{X}_{1,1}^m - \mathbb{E}[X_{1,1}^m]|)],
\end{aligned}$$

while, since $|X_{1,1}^m - \mathbb{E}[X_{1,1}^m]| \leq 1$,

$$\begin{aligned}
\tfrac{1}{2}\mathbb{E}\left[|X_{1,1}^m - \mathbb{E}[X_{1,1}^m]|^3\exp\left(\tfrac{1}{2}|X_{1,1}^m - \mathbb{E}[X_{1,1}^m]|\right)\right] &\leq \tfrac{1}{2}\exp\left(\tfrac{1}{2}\right)\mathbb{E}\left[|X_{1,1}^m - \mathbb{E}[X_{1,1}^m]|^2\right] \\
&\leq \mathbb{E}\left[|X_{1,1}^m - \mathbb{E}[X_{1,1}^m]|^2\right].
\end{aligned}$$

Thus, $c_1(m)$ and $c_2(m)$ behave like $1/\sqrt{m}$. Also, note that the bound in (20) is non-trivial for $x \geq \tilde{a}_n := \log(1 + c_2(m)n^{1/2})/c_1(m)$.

Approximating sets. Let $A_1^n = \{\max_{l\le m} Y_n^{m,l} > a_n\}$, for some $a_n = Cc_1(m)^{-1}(\log n)^2) \ge \tilde{a}_n$, where C is some finite constant. We have

$$\begin{aligned}
\mathbb{P}(A_1^n) &= \mathbb{P}\Big(\bigcup_{l\le m}\{Y_n^{m,l} > a_n\}\Big)\\
&\le \sum_{l\le m}\mathbb{P}(Y_n^{m,l} > a_n)\\
&\le m\mathrm{e}^{-c_1(m)a_n}\big(1 + c_2(m)n^{1/2}\big)\\
&\sim \sqrt{mn}\mathrm{e}^{-c_1(m)a_n}\\
&= \sqrt{mn}\mathrm{e}^{-(c_3C(\log n)^2)/2} \to 0, \qquad n\to+\infty.
\end{aligned}$$

Let $A_2^n = \{\max_{1\le l\le m} W_n^{m,l} > b_n\}$, for $b_n = \log n$. Standard estimates (including reflection principle, Brownian scaling and Gaussian tail estimates) lead to

$$\begin{aligned}
\mathbb{P}(A_2^n) &= \mathbb{P}\Big(\bigcup_{l\le m}\{W_n^{m,l} > b_n\}\Big)\\
&\le \sum_{l\le m}\mathbb{P}(W_n^{m,l} > b_n)\\
&\le m\mathbb{P}(W_n^{m,1} > b_n)\\
&= m\mathbb{P}\Big(\sup_{\substack{0\le s,t\le n\\ |s-t|\le 1}} |B_s^{m,1} - B_t^{m,1}| > b_n\Big).
\end{aligned}$$

However,

$$\begin{aligned}
\sup_{\substack{0\le s,t\le n\\ |s-t|\le 1}} |B_s^{m,1} - B_t^{m,1}| &\le \sup_{0\le i\le n-2}\sup_{i\le s,t\le i+2}|B_s^{m,1} - B_t^{m,1}|\\
&\le \sup_{0\le i\le n-2}\Big(\sup_{i\le t\le i+2} B_t^{m,1} - \inf_{i\le s\le i+2} B_s^{m,1}\Big)
\end{aligned}$$

and so

$$\begin{aligned}
\mathbb{P}(A_2^n) &\le m\mathbb{P}\Big(\sup_{0\le i\le n-2}\Big(\sup_{i\le t\le i+2} B_t^{m,1} - \inf_{i\le s\le i+2} B_s^{m,1}\Big) > b_n\Big)\\
&\le mn\mathbb{P}\Big(\sup_{t\in[0,2]} B_t^{m,1} - \inf_{s\in[0,2]} B_s^{m,1} > b_n\Big)\\
&\le mn\Big(\mathbb{P}\Big(\sup_{t\in[0,2]} B_t^{m,1} > b_n/2\Big) + \mathbb{P}\Big(\sup_{s\in[0,2]} B_s^{m,1} > b_n/2\Big)\Big) \qquad (21)\\
&\le 2mn\mathbb{P}(|B_2^{m,1}| > b_n/2)\\
&\le 4mn\exp(-b_n^2/16) \to 0, \qquad n\to+\infty.
\end{aligned}$$

Final bound. From (14), the approximation of $(\widetilde{V}_k(n,m))_{1\le k\le r}$ by $(L_k(n,m))_{1\le k\le r}$ will imply the theorem if

$$\mathbb{P}\left(\sum_{k=1}^{r}|\widetilde{V}_k(n,m)-L_k(n,m)|\ge c_n\right)\to 0,\qquad n\to+\infty, \tag{22}$$

for some

$$c_n=\mathrm{o}(n^{1/2}m^{-1/6}). \tag{23}$$

Since $\lim_{n\to+\infty}(\mathbb{P}(A_1^n)+\mathbb{P}(A_2^n))=0$, it is enough to prove that

$$\lim_{n\to+\infty}\mathbb{P}\left(\left\{\sum_{k=1}^{r}|\widetilde{V}_k(n,m)-L_k(n,m)|\ge c_n\right\}\cap(A_1^n)^c\cap(A_2^n)^c\right)=0. \tag{24}$$

But

$$\begin{aligned}
&\mathbb{E}\left[\sum_{k=1}^{r}|\widetilde{V}_k(n,m)-L_k(n,m)|\mathbf{1}_{(A_1^n)^c\cap(A_2^n)^c}\right]\\
&\le\sum_{k=1}^{k}2rm\mathbb{E}\big[(Y_n^{m,1}+W_n^{m,1})\mathbf{1}_{(A_1^n)^c\cap(A_2^n)^c}\big]\\
&\le 2r^2m(\mathbb{E}[Y_n^{m,1}\mathbf{1}_{Y_n^{m,1}\le a_n}]+b_n)\\
&\le 2r^2m(\mathbb{E}[Y_n^{m,1}-\tilde{a}_n\mathbf{1}_{Y_n^{m,1}\le a_n}]+\tilde{a}_n+b_n)\\
&\le 2r^2m(\mathbb{E}[Y_n^{m,1}-\tilde{a}_n\mathbf{1}_{\tilde{a}_n\le Y_n^{m,1}\le a_n}]+\tilde{a}_n+b_n)\\
&\le 2r^2m\left(\int_{\tilde{a}_n}^{a_n}\mathbb{P}(Y_n^{m,1}\ge x)\,\mathrm{d}x+\tilde{a}_n+b_n\right)\\
&\le 2r^2m\left(\int_{\tilde{a}_n}^{a_n}\mathrm{e}^{-c_1(m)x}\big(1+c_2(m)n\big)\,\mathrm{d}x+\tilde{a}_n+b_n\right)\\
&\le 2r^2m\left(\frac{1+c_2(m)n}{c_1(m)}\mathrm{e}^{-c_1(m)\tilde{a}_n}+\tilde{a}_n+b_n\right)\\
&\le 2r^2m\left(\frac{1}{c_1(m)}+\tilde{a}_n+b_n\right)\\
&\le 2r^2m^{3/2}\left(\frac{2(1+\log(1+c_2(m)n^{1/2}))}{c_3}+b_n\right).
\end{aligned}$$

Finally,

$$
\begin{aligned}
&\mathbb{P}\left(\left\{\sum_{k=1}^{r}|\widetilde{V}_k(n,m)-L_k(n,m)|\geq c_n\right\}\cap(A_1^n)^c\cap(A_2^n)^c\right)\\
&\quad\leq\frac{2r^2m^{3/2}}{c_n}\left(\frac{2(1+\log(1+c_2(m)n^{1/2}))}{c_3}+\log n\right)=\mathrm{O}\left(\frac{m^{3/2}\log n}{c_n}\right).
\end{aligned}
\tag{25}
$$

A choice of c_n ensuring that the bound in (25) goes to zero as $n\to+\infty$ and also compatible with (23) is possible when $m\log n=\mathrm{o}(n^{1/2}m^{-1/6})$, that is, when $m=\mathrm{o}(n^{3/10}(\log n)^{-3/5})$. Finally, (22) and (24) hold true, completing the proof of Theorem 4.

Remark 8.

- In the above proof, the condition $m=\mathrm{o}(n^{3/10}(\log n)^{-3/5})$ is needed only once, to ensure the compatibility of (23) with the bound (25). However, this is essential to make the Gaussian approximation work.
- When $m=[n^a]$, the growth condition $m=\mathrm{o}(n^{3/10}(\log n)^{-3/5})$ can be rewritten as $a<3/10$ and this growth condition remains true, in particular, when m is of subpolynomial order. The condition $a<3/10$ is stronger than its counterpart $a<3/7$ in [4] and this seems to be due to the fact that we work with a triangular array of random variables.
- For the top line of the tableau, our result falls short of a result of Johansson in [17], which asserts the convergence of $V_1(n,n^a)$ (properly scaled and normalized) toward the Tracy–Widom distribution. More precisely, setting $a_n\ll b_n$ for $a_n=\mathrm{o}(b_n)$, Theorem 1.7 in [17] actually gives, in our notation, for $\sqrt{n}\ll m$,

$$\frac{V_1(n,m)-n/m-2\sqrt{n}}{n^{1/6}}\Rightarrow F_{\mathrm{TW}},$$

for $(\log n)^{2/3}\ll m\ll\sqrt{n}$,

$$\frac{V_1(n,m)-n/m-2\sqrt{n}}{n^{1/2}m^{-2/3}}\Rightarrow F_{\mathrm{TW}}$$

and, for $\sqrt{n}/m\to l$,

$$\frac{V_1(n,m)-n/m-2\sqrt{n}}{(1+l)^{2/3}n^{1/6}}\Rightarrow F_{\mathrm{TW}}.$$

In the middle limit above, [17], Theorem 1.7, requires $(\log n)^{2/3}=\mathrm{o}(m)$, while we do not require a lower bound condition on m. Besides, our Theorem 4 applies to the shape of the whole Young tableau.

3. Proof of Theorem 6

In this section, we sketch the changes needed in the previous arguments in order to prove Theorem 6. Note that in the uniform setting, the representation (16) for $V_k(n,m)$ is a maximum taken over the most probable letters. This is trivially true since, in this case, all the letters have the same probability. But this property, which appears to be fundamental when we center and normalize the $X^m_{i,j}$, is no longer true in the non-uniform setting. However, we shall approximate $V_1(n,m)$ below by a random variable $V_1'(n,m)$ defined as a maximum taken over only the most probable letters, as in (16); see (28). Part of the remaining work is then to show that we can suitably control this approximation, and this is done in Lemma 9. This control is at the root of the extra condition (10) in Theorem 6.

Let us revise our notation for the non-uniform setting. In this section, X_i^m, $1 \le i \le n$, are independently and identically distributed with $\mathbb{P}(X_1^m = \alpha_j) = p_j^m$. Set $p^m_{\max} = \max_{1\le j\le m} p_j^m$ and $J(m) = \{j : p_j^m = p^m_{\max}\} = \{j_1, \dots, j_{k(m)}\}$, with $k(m) = \operatorname{card}(J(m))$, and also set $\sigma_m^2 = p^m_{\max}(1 - p^m_{\max})$. Finally, note that since $k(m(n))p^m_{\max} \le 1$ and $k(m(n)) \to +\infty$, it follows that $p^{m(n)}_{\max} \to 0$ as $n \to +\infty$.

Brownian scaling. Let $(B^l(s))_{s\ge 0}$, $1 \le l \le k(m)$, be independent standard Brownian motions. For $s > 0$, $m \ge 1$ and $k \ge 1$, let

$$L_1(s,k(m)) = \sup_{\mathbf{t}\in I_{k(m)}(s)} \sum_{l=1}^{k(m)} \bigl(B^l(t_l) - B^l(t_{l-1})\bigr), \tag{26}$$

where $I_{k(m)}(s) = \{\mathbf{t} : 0 \le t_1 \le \cdots \le t_{l-1} \le t_l \le \cdots \le t_{k(m)} = s\}$. Recall that $L_1(1,k(m))$ has the same law as the largest eigenvalue of a $k(m)\times k(m)$ GUE matrix (see (2), (3), (4) and Remark 3 for $k=1$) and so

$$k^{1/6}\bigl(L(1,k) - 2\sqrt{k}\bigr) \Rightarrow F_{\mathrm{TW}}.$$

By Brownian scaling, $L_1(s,m) \stackrel{\mathcal{L}}{=} \sqrt{s}L_1(1,m)$, so that when $n \to +\infty$,

$$\frac{L_1(n,k(m(n))) - 2\sqrt{nk(m(n))}}{n^{1/2}k(m(n))^{-1/6}} \Rightarrow F_{\mathrm{TW}}. \tag{27}$$

Combinatorics revisited. Let

$$X^m_{i,j} = \begin{cases} 1, & \text{when } X_i^m = \alpha_j, \\ 0, & \text{otherwise}, \end{cases}$$

be Bernoulli random variables with parameter $\mathbb{P}(X_i^m = \alpha_j) = p_j^m$ and variance $(\sigma_j^m)^2 = p_j^m(1 - p_j^m)$. For a fixed $1 \le j \le m$, the $X^m_{i,j}$s are independent and identically distributed. Since the expression (15) has a purely combinatorial nature, we still have

$$V_1(n,m) = \max_{\substack{0=l_0\le l_1\le\cdots\\ \le l_{m-1}\le l_m=n}} \left(\sum_{j=1}^{m}\sum_{i=l_{j-1}+1}^{l_j} X^m_{i,j}\right),$$

with the convention that $\sum_{i=l_{j-1}+1}^{l_j} X_{i,j}^m = 0$ whenever $l_{j-1} = l_j$.

In fact, for most draws, the maximum in V_1 is attained on the sums $\sum_{j\in J(m)} \sum_{i=l_{j-1}+1}^{l_j} X_{i,j}^m$ corresponding to the most probable letters, that is, letting

$$V_1'(n,m) = \max_{\substack{0=l_0\le l_1\le\cdots\\ \le l_{m-1}\le l_m=n\\ l_{j-1}=l_j \text{ for } j\notin J(m)}} \left(\sum_{j=1}^{n} \sum_{i=l_{j-1}+1}^{l_j} X_{i,j}^m\right), \tag{28}$$

we have, with large probability, $V_1(n,m) = V_1'(n,m)$. However, it is not always true that $V_1(n,m) = V_1'(n,m)$, for instance, in the case where the n letters drawn are letters with associated probability strictly less than $p_{\max}^m$, $V_1'(n,m) = 0$ while there is an $l = (l_j)_{j=0,\dots,m}$ with $0 = l_0 \le l_1 \le \cdots \le l_{m-1} \le l_m = n$ such that $\sum_{j=1}^m \sum_{i=l_{j-1}+1}^{l_j} X_{i,j}^m > 0$, ensuring that $V_1(n,m) > 0$. In the sequel, we prove Theorem 6 by first showing that the statement of the theorem is true for $V_1'(n,m)$ instead of $V_1(n,m)$ and then by controlling the error made when $V_1'(n,m)$ is replaced by $V_1(n,m)$.

Centering and reducing. Let $\widetilde{X}_{i,j}^m = (X_{i,j}^m - p_j^m)/\sigma_j^m$ be the corresponding centered and normalized scaled Bernoulli random variables and let $\widetilde{S}_l^{m,j} = \sum_{i=1}^l \widetilde{X}_{i,j}^m$. Also, let

$$\begin{aligned}\widetilde{V}_1'(n,m) &= \max_{\substack{0=l_0\le l_{j_1}\le\cdots\\ \le l_{j_{k(m)-1}}\le l_{j_{k(m)}}=n}} \left(\sum_{j\in J(m)} \sum_{i=l_{j-1}+1}^{l_j} \widetilde{X}_{i,j}^m\right) = \max_{\substack{0=l_0\le l_{j_1}\le\cdots\\ \le l_{j_{k(m)-1}}\le l_{j_{k(m)}}=n}} \left(\sum_{j\in J(m)} (\widetilde{S}_{l_j}^{m,j} - \widetilde{S}_{l_{j-1}}^{m,j})\right)\\ &= \sup_{\mathbf{t}\in I_{k(m(n))}(n)} \sum_{\delta=1}^{k(m(n))} \left(\widetilde{S}_{[t_{j,l}]}^{m,j_\delta} - \widetilde{S}_{[t_{j,l-1}]}^{m,j_\delta}\right),\end{aligned}$$

which is to be compared to (26). Since $V_1'(n,m) - np_{\max}^m = \sigma_m \widetilde{V}_1'(n,m)$, we have

$$\begin{aligned}k(m)^{1/6}\frac{V_1'(n,m) - np_{\max}^m - 2\sqrt{nk(m)\sigma_m^2}}{\sqrt{p_{\max}^m n}}\frac{\sqrt{p_{\max}^m}}{\sigma_m}\\ = k(m)^{1/6}\frac{\widetilde{V}_1(n,m) - 2\sqrt{nk(m)}}{\sqrt{n}}.\end{aligned}$$

Since $\sigma_m \sim \sqrt{p_{\max}^m}$ and

$$\frac{2\sqrt{k(m)np_{\max}^m} - 2\sqrt{k(m)n\sigma_m^2}}{\sqrt{n\sigma_m^2}} = \frac{2\sqrt{k(m)}}{\sigma_m}\frac{p_{\max}^m - \sigma_m^2}{\sqrt{p_{\max}^m} + \sqrt{\sigma_m^2}} \sim \frac{2\sqrt{k(m)}}{\sqrt{p_{\max}^m}}\frac{(p_{\max}^m)^2}{\sqrt{p_{\max}^m}} \le 2\sqrt{p_{\max}^m} \to 0,$$
$$n \to +\infty,$$

it remains to show that

$$k(m)^{1/6}\frac{\widetilde{V}_1'(n,m) - 2\sqrt{nk(m)}}{\sqrt{n}} \Rightarrow F_{\mathrm{TW}}, \tag{29}$$

for which we shall use (27).

Sketch of proof of (29). Roughly speaking, the proof of (29) follows along the same lines as the corresponding proof of the convergence of (18), changing only m to $k(m)$. We show that if $k(m(n)) = \mathrm{o}(n^{3/10}(\log n)^{-3/5})$, then for some Brownian motions given via strong approximation, we have

$$|\widetilde{V}_1'(n,m) - L_1(n, k(m(n)))| \leq \sum_{l=1}^{k(m(n))} (Y_n^{m,l} + W_n^{m,l}),$$

where

$$Y_n^{m,l} = \max_{1\leq i\leq n} |S_i^{m,l} - B^{m,l}(i)| \quad \text{and} \quad W_n^{m,l} = \sup_{\substack{0\leq s,t\leq n \\ |s-t|\leq 1}} |B^{m,l}(s) - B^{m,l}(t)|.$$

Indeed, setting $A_1^n = \{\max_{l\leq k(m(n))} Y_n^{m,l} > a_n\}$ for some $a_n = \mathrm{O}(c_1(k(m(n)))^{-1}(\log n)^2) \geq \tilde{a}_n := \log(1 + c_2(k(m(n)))n^{1/2})/c_1(k(m(n)))$ and setting $A_2^n = \{\max_{1\leq l\leq k(m(n))} W_n^{m,l} > b_n\}$ for some $b_n = \mathrm{O}(\log n)$, we show that

$$\mathbb{P}(A_1^n) \to 0, \qquad \mathbb{P}(A_2^n) \to 0, \qquad \text{when } n \to +\infty.$$

From (27), the approximation of $\widetilde{V}_1(n, k(m(n)))$ by $L_1(n, k(m(n))))$ will imply the theorem if

$$\mathbb{P}\big(|\widetilde{V}_1'(n, k(m(n))) - L_1(n, k(m(n)))| \geq c_n\big) \to 0, \qquad n \to +\infty, \tag{30}$$

for some

$$c_n = \mathrm{o}(n^{1/2}k(m(n))^{-1/6}). \tag{31}$$

Since $\lim_{n\to+\infty}(\mathbb{P}(A_1^n) + \mathbb{P}(A_2^n)) = 0$ and

$$\begin{aligned} &\mathbb{P}\big(\{|\widetilde{V}_1'(n, k(m(n))) - L_1(n, k(m(n)))| \geq c_n\} \cap (A_1^n)^c \cap (A_2^n)^c\big) \\ &\quad \leq \frac{2k(m(n))^{3/2}}{c_n}\left(\frac{2(1 + \log(1 + c_2(k(m(n)))n^{1/2}))}{c_3} + \log n\right), \end{aligned} \tag{32}$$

a choice of c_n, ensuring that the bound in (32) goes to zero and is compatible with (31), is possible since $k(m(n)) = \mathrm{o}(n^{3/10}(\log n)^{-3/5})$. This proves (29) and thus the statement (11) of Theorem 6, but for $V_1'(n,m)$ instead of $V_1(n,m)$.

Control of the error $V_1(n,m) - V_1'(n,m)$. Clearly, $V_1(n,m) - V_1'(n,m) \geq 0$ and is, in fact, zero with a large probability, so that we expect $\mathbb{E}[V_1(n,m) - V_1'(n,m)]$ to be small. Actually, we show the following.

Lemma 9. *For some absolute constant $C > 0$, we have*

$$\mathbb{E}[|V_1(n,m) - V_1'(n,m)|] \leq Cnp_{2\mathrm{nd}}^m, \tag{33}$$

where $p_{2\mathrm{nd}}^m$ stands for the second largest probability for the letters of $\mathcal{A}_m$.

The conclusion in (11) holds true when

$$\lim_{n\to+\infty}\left(\mathbb{E}[|V_1(n,m)-V_1'(n,m)|]\times\frac{k(m(n))^{2/3}}{\sqrt{k(m(n))\,p_{\max}^{m(n)}n}}\right)=0. \tag{34}$$

However, with the help of (33), the conclusion in (34) is then valid when $\lim_{n\to+\infty} p_{2\mathrm{nd}}^{m(n)} \times k(m(n))^{1/6}n^{1/2}/(p_{\max}^{m(n)})^{1/2}=0$ and, since $k(m(n))=\mathrm{o}(n^{3/10}(\log n)^{-3/5})$, this will follow from (10).

It remains to prove Lemma 9, that is, to give an explicit bound on $\mathbb{E}[|V_1(n,m)-V_1'(n,m)|]$. To do so, rewrite $V_1(n,m)=\max_{l\in I(m)} Z(l)$ and $V_1'(n,m)=\max_{l\in I^*(m)} Z(l)$, where $I(m)=\{l=(l_j)_{1\le j\le m}: l_{j-1}\le l_j, l_0=0, l_m=n\}$, $I^*(m)=\{l\in I(m): l_{j-1}=l_j \text{ for } j\notin J(m)\}$ and

$$Z(l)=\sum_{j=1}^m Y_j(l),\qquad Y_j(l)=\sum_{i=l_{j-1}+1}^{l_j} X_{i,j}^m.$$

Clearly, since $I^*(m)\subset I(m)$, we have $V_1'(n,m)\le V_1(n,m)$. Moreover, since the $X_{i,j}^m$ are Bernoulli random variables with parameter p_j^m and since the X_i's are independent, we have $Y_j(l)\sim\mathcal{B}(l_j-l_{j-1},p_j)$ and $\sum_{j\in J(m)}Y_j(l)\sim\mathcal{B}(\sum_{j\in J(m)}l_j-l_{j-1},p_{\max}^m)$, where $\mathcal{B}(n,p)$ stands for the binomial distribution with parameters n and p.

If $l\in I^*(m)$, $Z(l)=\sum_{j\in J(m)}Y_j(l)\sim\mathcal{B}(n,p_{\max}^m)$, since, in this case, $n=\sum_{j=1}^m(l_j-l_{j-1})=\sum_{j\in J(m)}(l_j-l_{j-1})$. If $l\notin I^*(m)$, we rewrite $Z(l)$ as

$$Z(l)=Z(\tilde{l})+R(l),$$

where $\tilde{l}\in I^*(m)$ and $R(l)$ is an error term. Indeed, let $J_l=\{j\notin J(m): l_{j-1}<l_j\}$ and for $j\in J_l$, define

$$\theta(j)=\begin{cases}\max A_j, & \text{if } A_j\ne\varnothing,\\ \min B_j, & \text{otherwise,}\end{cases}$$

where $A_j=\{k\in J(m): k\le j\}$ and where $B_j=\{k\in J(m): k\ge j\}$. Now,

$$\begin{aligned}Z(l)&=\sum_{j\in J(m)}\sum_{i=l_{j-1}+1}^{l_j}X_{i,j}^m+\sum_{j\in J_l}\sum_{i=l_{j-1}+1}^{l_j}X_{i,j}^m\\
&=\sum_{j\in J(m)}\sum_{i=l_{j-1}+1}^{l_j}X_{i,j}^m+\sum_{j\in J_l}\sum_{i=l_{j-1}+1}^{l_j}X_{i,\theta(j)}^m \qquad (35)\\
&\quad+\sum_{j\in J_l}\sum_{i=l_{j-1}+1}^{l_j}\big(X_{i,j}^m-X_{i,\theta(j)}^m\big). \qquad (36)\end{aligned}$$

Define $\tilde{l} \in I^*(m)$ by $\tilde{l}_j = \tilde{l}_{j-1}$ if $j \notin J(m)$ and $\tilde{l}_j = l_{k-1}$ for $j \in J(m)$, where $k = \min\{l > j : l \in J(m)\}$, with the conventions that $\min \varnothing = m+1$ and that $\tilde{l}_{j_0-1} = 0$ for $j_0 = \min J(m)$. We then have

$$\sum_{j \in J(m)} \sum_{i=l_{j-1}+1}^{l_j} X^m_{i,j} + \sum_{j \in J_l} \sum_{i=l_{j-1}+1}^{l_j} X^m_{i,\theta(j)} = Z(\tilde{l}).$$

Let $\alpha^m_{i,j} := X^m_{i,j} - X^m_{i,\theta(j)}$ be the random variables taking the values -1, 0 and $+1$ with respective probabilities $p^m_{\max}$, $1 - p^m_{\max} - p^m_j$ and p^m_j. Independently, let $\varepsilon^m_{i,j}$ be Bernoulli random variables with parameter $q^m_j = (p^m_{2\mathrm{nd}} - p^m_j)/(1 - p^m_{\max} - p^m_j) \in (0,1)$, where $p^m_{2\mathrm{nd}} = \max(p^m_j < p^m_{\max} : 1 \le j \le m)$, and define

$$\beta^m_{i,j} = \begin{cases} -1, & \alpha^m_{i,j} = -1, \\ 0, & \alpha^m_{i,j} = 0 \text{ and } \varepsilon^m_{i,j} = 0, \\ +1, & \alpha^m_{i,j} = +1 \text{ or } \alpha^m_{i,j} = 0 \text{ and } \varepsilon^m_{i,j} = 1. \end{cases}$$

Note that $\mathbb{P}(\beta^m_{i,j} = +1) = p^m_{2\mathrm{nd}}$ and that $\alpha^m_{i,j} \le \beta^m_{i,j}$, so that

$$R(l) \le \widetilde{R}(l) = \sum_{j \in J_l} \sum_{i=l_{j-1}+1}^{l_j} \beta^m_{i,j}.$$

Since $Z(l) \le Z(\tilde{l}) + \widetilde{R}(l)$, we have

$$\begin{aligned} \max_{l \in I(m)} Z(l) &\le \max_{l \in I(m)} Z(\tilde{l}) + \max_{l \in I(m)} \widetilde{R}(l) \\ &\le \max_{l \in I^*(m)} Z(l) + \max_{l \in I(m)} \widetilde{R}(l). \end{aligned}$$

Next, observe that for $l \in I^*(m)$, $\widetilde{R}(l) = 0$. However, since the event $\{\widetilde{R}(l) < 0, \forall l \notin I^*(m)\}$ is non-negligible, we cannot change $\max_{l \in I(m)} \widetilde{R}(l)$ into $\max_{l \notin I^*(m)} \widetilde{R}(l)$. We obtain

$$0 \le \max_{l \in I(m)} Z(l) - \max_{l \in I^*(m)} Z(l) \le \max_{l \in I(m)} \widetilde{R}(l).$$

The random variable $\widetilde{R}(l)$ is the sum of $\sum_{j \in J_l}(l_j - l_{j-1})$ i.i.d. random variables so that $\max_{l \in I(m)} \widetilde{R}(l)$ is distributed according to $(\max_{1 \le k \le n} \sum_{i=1}^k \beta^m_i)^+$, where $(\beta^m_i)_i$ are i.i.d. with

$$\mathbb{P}(\beta^m_1 = -1) = p^m_{\max}, \qquad \mathbb{P}(\beta^m_1 = 0) = 1 - p^m_{\max} - p^m_{2\mathrm{nd}}, \qquad \mathbb{P}(\beta^m_1 = +1) = p^m_{2\mathrm{nd}}. \quad (37)$$

We are now interested in bounding $\mathbb{E}[(\max_{1 \le k \le n} \sum_{i=1}^k \beta^m_i)^+]$.

Let $(\varepsilon^m_i)_i$ be i.i.d. Bernoulli random variables with parameter $p^m_{\max} + p^m_{2\mathrm{nd}}$ and let, independently, $(Y^m_i)_i$ be i.i.d. Rademacher random variables with parameter $p^m_{2\mathrm{nd}}/(p^m_{2\mathrm{nd}} + p^m_{\max})$ (i.e.,

$\mathbb{P}(Y_i^m = 1) = 1 - \mathbb{P}(Y_i^m = -1) = p_{2\mathrm{nd}}^m/(p_{2\mathrm{nd}}^m + p_{\max}^m))$. Then β_i^m and $\varepsilon_i^m Y_i^m$ have the same distribution and we have

$$\mathbb{E}\left[\left(\max_{1\le k\le n}\sum_{i=1}^k \beta_i^m\right)^+\right] = \mathbb{E}\left[\left(\max_{1\le k\le n}\sum_{i=1}^k \varepsilon_i^m Y_i^m\right)^+\right]$$
$$= \mathbb{E}\left[\mathbb{E}\left[\left(\max_{1\le k\le n}\sum_{i=1}^k \varepsilon_i^m Y_i^m\right)^+ \middle| \mathcal{G}_n\right]\right],$$

where $\mathcal{G}_n = \sigma(\varepsilon_i^m : 1 \le i \le n)$. However, since $(\varepsilon_i^m)_i$ is independent of $(Y_i^m)_i$, we have

$$\mathbb{E}\left[\left(\max_{1\le k\le n}\sum_{i=1}^k \varepsilon_i^m Y_i^m\right)^+ \middle| \mathcal{G}_n\right] = \mathbb{E}\left[\left(\max_{1\le k\le \ell}\sum_{i=1}^k Y_i^m\right)^+ \middle| \mathcal{G}_n\right],$$

where $\ell = \sum_{i=1}^n \varepsilon_i^m$ has a $\mathcal{B}(n, p_{\max}^m + p_{2\mathrm{nd}}^m)$ distribution. But

$$\mathbb{E}\left[\left(\max_{1\le k\le \ell}\sum_{i=1}^k Y_i^m\right)^+ \middle| \mathcal{G}_n\right] = \sum_{k=1}^{+\infty}\left(\mathbb{P}\left(\left(\max_{1\le j\le \ell}\sum_{i=1}^j Y_i^m\right)^+ \ge k\right)\right)$$
$$= \sum_{k=0}^{+\infty}\left(1 - \mathbb{P}\left(\max_{1\le j\le \ell}\sum_{i=1}^j Y_i^m \le k\right)\right)$$
$$= \sum_{k=0}^{\ell-1}\left(1 - \mathbb{P}\left(\max_{1\le j\le \ell}\sum_{i=1}^j Y_i^m \le k\right)\right)$$
$$= \ell - U_\ell,$$

where $U_\ell = \sum_{k=0}^{\ell-1} u_{\ell,k}$ and $u_{\ell,k} = \mathbb{P}(\max_{1\le j\le \ell}\sum_{i=1}^j Y_i^m \le k)$. With the latest notation, we are now investigating $\gamma_n = \mathbb{E}[\ell - U_\ell]$. For simplicity, in the sequel, we set $p_{*,m} := p_{2\mathrm{nd}}^m/(p_{2\mathrm{nd}}^m + p_{\max}^m)$ and $q_{*,m} := 1 - p_{*,m}$.

The elements of the sequence $(u_{\ell,k})_{1\le k\le \ell-1}$ satisfy the following induction relations:

$$u_{\ell,k} = q_{*,m} u_{\ell-1,k+1} + p_{*,m} u_{\ell-1,k-1}, \qquad k \ge 1, \qquad u_{\ell,0} = q_{*,m} u_{\ell-1,1},$$

and $u_{\ell,k} = 1$ for $k \ge \ell$. From these, we derive $U_\ell = 2q_{*,m} - q_{*,m} u_{\ell-1,0} + U_{\ell-1}$ and, since $U_1 = u_{1,0} = q_{*,m}$, $U_\ell = (2\ell - 1) q_{*,m} - q_{*,m}\sum_{k=1}^{\ell-1} u_{k,0}$.

In order to compute $\sum_{k=1}^{\ell-1} u_{k,0}$, we introduce the hitting time $\tau_1^m = \min(k \ge 1 : \sum_{i=1}^k Y_i^m = 1)$ of the random walk $(\sum_{i\le j} Y_i^m)_j$. We then have

$$\mathbb{P}(\tau_1^m \le k) = \mathbb{P}\left(\max_{i\le k}\sum_{j=1}^i Y_j^m \ge 1\right) = 1 - \mathbb{P}\left(\max_{i\le k}\sum_{j=1}^i Y_j^m \le 0\right) = 1 - u_{k,0}$$

so that $\sum_{k=1}^{\ell-1} u_{k,0} = \sum_{k=1}^{\ell-1} \mathbb{P}(\tau_1^m \geq k+1) = \sum_{k=2}^{\ell} \mathbb{P}(\tau_1^m \geq k) = -1 + \sum_{k=1}^{\ell} \mathbb{P}(\tau_1^m \geq k)$ and

$$\begin{aligned} U_\ell &= 2\ell q_{*,m} - q_{*,m} \sum_{k=1}^{\ell} \mathbb{P}(\tau_1^m \geq k) \\ &= 2\ell q_{*,m} - q_{*,m} \sum_{i=1}^{+\infty} (i \wedge \ell) \mathbb{P}(\tau_1^m = i) \\ &= 2\ell q_{*,m} - q_{*,m} \mathbb{E}[\tau_1^m \wedge \ell | \mathcal{G}_n]. \end{aligned}$$

Next,

$$\mathbb{E}\left[\left(\max_{i \leq k \leq \ell} \sum_{i=1}^{k} Y_i^m\right)^+ \middle| \mathcal{G}_n\right] = \ell(1 - 2q_{*,m}) + q_{*,m} \mathbb{E}[\tau_1^m \wedge \ell | \mathcal{G}_n]$$

and we have

$$\begin{aligned} \gamma_n &:= \mathbb{E}[\ell - U_\ell] \\ &= \mathbb{E}\big[\ell(1 - 2q_{*,m}) + q_{*,m} \mathbb{E}[\tau_1^m \wedge \ell | \mathcal{G}_n]\big] \\ &\leq (1 - 2q_{*,m} + q_{*,m}) \mathbb{E}[\ell] \\ &= n p_{2nd}^m. \end{aligned}$$

This completes the proof of Lemma 9.

4. Concluding remarks

A natural question to consider next would be to derive a result similar to Theorem 4 for non-uniformly distributed letters. The special case of the longest increasing subsequence (i.e., $r = 1$) is dealt with in Theorem 6. Let us investigate what happens for the whole shape of the Young tableau.

First, let us slightly expand our notation. In this section, X_i^m, $1 \leq i \leq n$, are independently and identically distributed with $\mathbb{P}(X_1^m = \alpha_j) = p_j^m$. In order to simplify the notation, we assume (without loss of generality) that the ordered letters $\alpha_1^m < \cdots < \alpha_m^m$ have, moreover, non-increasing probabilities (i.e., $p_1^m \geq p_2^m \geq \cdots \geq p_m^m$). Let $d_i^m = \text{card}\,\{j : p_j^m = p_i^m\}$ be the multiplicity of p_i^m and let $m_r^m = \max\{i : p_i^m > p_r^m\}$ be the number of letters (strictly) more probable than α_r^m. Let $J_r(m) = \{i : p_i^m = p_r^m\} = \{m_r + 1, \ldots, m_r + d_r^m\}$ be the indices of the letters with the same probability p_r^m. We recover our previous notation, $r = 1$, with $k(m) = d_1^m$ and $J(m) = J_1(m)$. Since the expression (16) has a purely combinatorial nature, it still holds true that

$$V_r(n, m) = \max_{\mathbf{k} \in J_{r,m}(n)} \left(\sum_{j=1}^{r} \sum_{l=j}^{m-r+j} \sum_{i=k_{j,l-1}+1}^{k_{j,l}} X_{i,l}^m \right).$$

Let $\nu_k^m = \sum_{i=1}^k p_i^m$. Note that, from Theorem 5.2 in [13], when m is fixed and $n \to +\infty$, we have for each $1 \le r \le m$ that

$$\left(\frac{V_k(n,m) - \nu_k^m n}{\sqrt{n}}\right)_{1\le k\le r} \Longrightarrow (V_\infty^k)_{1\le k\le r}, \tag{38}$$

where the limit is given in Section 6 of [13] by $V_\infty^r = Z(m,r) + \sqrt{p_r^m} D_{r-m_r^m, d_r^m}$, with $Z(m,r) \sim \mathcal{N}(0, \upsilon_r^m)$ for $\upsilon_r^m = \nu_{m_r^m}^m(1-\nu_{m_r^m}^m) + (p_r^m(r-m_r^m))^2$, and

$$D_{r,m} = \max_{\mathbf{t}\in I_{r,m}} \left(\sum_{j=1}^{r} \sum_{l=j}^{(m-r+j)} \big(B^l(t_{j,l}) - B^l(t_{j,l-1})\big) \right)$$

for

$$\begin{aligned} I_{r,m} = \{\mathbf{t} = (t_{j,l}, 1\le j\le r,\ 0\le l\le m)\colon\ & t_{j,j-1} = 0, t_{j,m-r+j} = 1, 1\le j\le r, \\ & t_{j,l-1} \le t_{j,l}, 1\le j\le r, 1\le l\le m-1, t_{j,l} \le t_{j-1,l}, 2\le j\le r, 1\le l\le m-1\}. \end{aligned}$$

Note that $D_{r,m}$ is a natural generalization of the Brownian functional $L_1(s,k)$ used in Section 3 (see also, in a queuing context, [9] and [3]). In particular, $D_{r,m}$ is equal in distribution to the sum of the r largest eigenvalues of an $m \times m$ matrix from the GUE and Theorem 2 can be rewritten as

$$\big(m^{1/6}\big(D_{k,m} - k\sqrt{m}\big)\big)_{1\le k\le r} \Rightarrow \mathbf{F}_r \mathbf{\Theta}_r^{-1}, \qquad m \to +\infty. \tag{39}$$

Arguing as in the previous sections, we would like to derive the fluctuations of $(V_k(n,m))_{1\le k\le r}$ with respect to n and m simultaneously from (38) and (39). However, in the non-uniform case, this is not that transparent since, for each $r \ge 1$, the behavior of m_r^m and of d_r^m, with respect to m, is not that clear cut. In particular, $r - m_r^m$ may not be stationary and (39) can no longer be used for $D_{r-m_r^m, d_r^m}$. Besides, the random fluctuations of $\sqrt{p_r^m} D_{r-m_r^m, d_r^m}$ in V_∞^r are of order $(p_r^m)^{1/2}(d_r^m)^{1/6}$, which, in general, does not dominate those of $Z(m,r) \sim \mathcal{N}(0, \upsilon_r^m)$. Thus, for general non-uniform alphabets, we cannot infer which part of the law of $V_\infty^r = Z(m,r) + \sqrt{p_r^m} D_{r-m_r^m, d_r^m}$ will drive the fluctuations. We can imagine that, taking simultaneous limits in n and m, the fluctuations of $V_r(n, m(n))$, properly centered and normalized, are either Gaussian, or driven by $\mathbf{F}_r$ (as in Theorem 4) or given by an interpolation between these distributions, depending on the alphabets considered.

Acknowledgments

The first author thanks the School of Mathematics of the Georgia Institute of Technology where this work was initiated. The second author would like to thank the Équipe Modal'X of the University Paris X for its hospitality and support while part of this research was being carried out. Research of the second author was supported in part by NSA Grant H98230-09-1-0017.

References

[1] Baik, J., Deift, P. and Johansson, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** 1119–1178. MR1682248

[2] Baik, J. and Suidan, T. (2005). A GUE central limit theorem and universality of directed first and last passage percolation site. *Int. Math. Res. Not.* **6** 325–337. MR2131383

[3] Baryshnikov, Y. (2001). GUEs and Queues. *Probab. Theory Related Fields* **119** 256–274. MR1818248

[4] Bodineau, T. and Martin, J. (2005). A universality property for last-passage percolation paths close to the axis. *Electron. Comm. Probab.* **10** 105–112. MR2150699

[5] Bougerol, P. and Jeulin, T. (2002). Paths in Weyl chambers and random matrices. *Probab. Theory Related Fields* **124** 517–543. MR1942321

[6] Chistyakov, G.P. and Götze, F. (2004). Distribution of the shape of Markovian random words. *Probab. Theory Related Fields* **129** 18–36. MR2052861

[7] Doumerc, Y. (2003). A note on representations of eigenvalues of classical Gaussian matrices. In *Séminaire de Probabilité XXXVII. Lecture Notes in Math.* **1832** 370–384. Berlin: Springer. MR2053054

[8] Fulton, W. (1997). *Young Tableaux with Applications to Representation Theory and Geometry*. Cambridge Univ. Press. MR1464693

[9] Glynn, P.W. and Whitt, W. (1991). Departure from many queues in series. *Ann. Appl. Probab.* **1** 546–572. MR1129774

[10] Gravner, J., Tracy, C. and Widom, H. (2001). Limit theorems for height fluctuations in a class of discrete space and time growth models. *J. Stat. Phys.* **102** 1085–1132. MR1830441

[11] Houdré, C. and Litherland, T. (2008). On the longest increasing subsequence for finite and countable alphabets. Preprint. Available at arXiv:math.PR/0612364.

[12] Houdré, C. and Litherland, T. (2008). Longest increasing subsequence asymptotics with growing alphabets. Preprint.

[13] Houdré, C. and Litherland, T. (2008). On the limiting shape of Markovian random Young tableaux. Preprint. Available at arXiv:0810.2982.

[14] Houdré, C. and Xu, H. (2009). On the limiting shape of random Young tableaux associated to inhomogeneous words. Available at arXiv:0901.4138.

[15] Its, A., Tracy, C. and Widom, H. (2001). Random words, Toeplitz determinants, and integrable systems I. In *Random Matrix Models and Their Applications*. *Math. Sci. Res. Inst.* **40** 245–258. Cambridge Univ. Press. MR1842789

[16] Its, A., Tracy, C. and Widom, H. (2001). Random words, Toeplitz determinants and integrable systems II. Advances in nonlinear mathematics and science. *Phys. D* **152/153** 199–224. MR1837910

[17] Johansson, K. (2001). Discrete polynomials ensembles and the Plancherel measure. *Ann. of Math. (2)* **153** 259–296. MR1826414

[18] Lifshits, M. (2000). Lecture notes on strong approximation. *Pub. IRMA Lille* **53** 13.

[19] Mehta, M.L. (1991). *Random Matrices*, 2nd ed. San Diego: Academic Press. MR1083764

[20] O'Connell, N. and Yor, M. (2001). Brownian analogues of Burke's theorem. *Stochastic Process Appl.* **96** 285–304. MR1865759

[21] Stanley, R. (1999). *Enumerative Combinatorics* **2**. Cambridge Univ. Press. MR1676282

[22] Suidan, T. (2006). A remark on a theorem of Chatterjee and last passage percolation. *J. Phys. A: Math. Gen.* **39** 8977–8981. MR2240468

[23] Tracy, C. and Widom, H. (1994). Level-spacing distributions and the Airy kernel. *Commun. Math. Phys.* **159** 151–174. MR1257246

[24] Tracy, C. and Widom, H. (1998). Correlation functions, cluster functions and spacing distributions for random matrices. *J. Statist. Phys.* **92** 809–835. MR1657844

[25] Tracy, C. and Widom, H. (2001). On the distribution of the length of the longest monotone subsequences in random word. *Probab. Theory Relat. Fields* **119** 350–380. MR1821139

Received December 2008 and revised May 2009