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

Jean-Christophe Breton∗† and Christian Houdré‡§
June 8, 2021

Abstract

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 both $n$ and $m$ converge together to infinity. If $m$ does not grow too fast and if the draws are uniform, 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 towards the Tracy-Widom distribution.

Key words: Longest increasing subsequence; GUE; Random words; Strong approximation; Tracy-Widom distribution; Young tableaux.

AMS 2000 Subject Classification. Primary: 60F05. Secondary: 60B12, 60C05, 60F15.

1 Introduction and results

Let $A_m = \{\alpha_1 < \alpha_2 < \cdots < \alpha_m\}$ 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 $A_m$. 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 (e.g., see [Fu] or [St]). 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 mathematical subfields (e.g. random matrices, queueing theory, percolation theory) have been investigated in numerous papers ([BDJ], [BS], [BM], [GW], [ITW1], [ITW2], [Jo], [TW3], …). 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 a $m \times m$ matrix from the traceless Gaussian Unitary Ensemble (GUE). More generally (see [Jo]), 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 a $m \times m$ traceless GUE matrix. For different random words such as non uniform or Markovian ones, things are more involved ([ITW1], [ITW2], [HL2], [HX], [CG]).

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

---

∗Laboratoire de Mathématiques, Image et Applications, Avenue Michel Crépeau, Université de La Rochelle, 17042 La Rochelle cedex, France. E-mail: jcbreton@univ-lr.fr
†Corresponding author
‡School of Mathematics, 686 Cherry Street, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA. E-mail: houdre@math.gatech.edu
§Research supported in part by NSA grant H98230-09-1-0017
Theorem 1  Let $V_k(n, m) = \sum_{i=1}^{k} R_i^n$ be the sum of the lengths $R_i^n$ of the first $k$ rows of the Young tableau. Then,

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

(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

$$I_{k,m} = \{ t = (t_{j,l} : 1 \leq j \leq k, 0 \leq l \leq m) : t_{j,j-1} = 0, t_{j,m-k+j} = 1, 1 \leq j \leq k, t_{j,l} \leq t_{j,l-1}, 2 \leq j \leq k, 2 \leq l \leq m \}.$$

Here, and in the sequel, the rows beyond the height of the tableau are considered to be of length zero. Letting $\Theta_k : \mathbb{R}^k \to \mathbb{R}^k$ be defined via $(\Theta_k(x))_j = \sum_{i=1}^{j} x_i$, $1 \leq j \leq k$, then the shape of the Young tableau is given by $\Theta_k^{-1}(V_1(n, m), \ldots, V_m(n, m))^T = (R_1^n, \ldots, R_m^n)^T$. Moreover, let $(\lambda_{GUE,m}^1, \lambda_{GUE,m}^2, \ldots, \lambda_{GUE,m}^m)$ be the spectrum, written in non-increasing order, of a $m \times m$ traceless element of the GUE, when the GUE is equipped with the measure

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

and $C_m = (2\pi)^{m/2} \prod_{j=1}^{m} j!$ (see [Me]). An important fact (see [Ba], [BJ], [Do], [GTW], [HL2], [OCY]) asserts that

$$
\frac{\sqrt{m-1}}{\sqrt{m}} \Theta_k^{-1} \left( \left( \max_{t \in I_{k,m}} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} (\hat{B}(t_{j,l}) - \hat{B}(t_{j,l-1})) \right)_{1 \leq k \leq m} \right)
\leq (\lambda_{GUE,m}^1, \lambda_{GUE,m}^2, \ldots, \lambda_{GUE,m}^m).
$$

(2)

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

$$
(\lambda_{GUE,m}^1, \lambda_{GUE,m}^2, \ldots, \lambda_{GUE,m}^m) \leq (\lambda_{GUE,m}^1, \lambda_{GUE,m}^2, \ldots, \lambda_{GUE,m}^m) + Z_m e_m,
$$

(3)

where $Z_m$ is a centered Gaussian random variable with variance $1/m$, independent of the vector $(\lambda_{GUE,m}^1, \lambda_{GUE,m}^2, \ldots, \lambda_{GUE,m}^m)$ and where $e_m = (1, 1, \ldots, 1)$, see [HX] 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 [TW1], [TW2] and also Theorem 1.4 in [Jo], with slight change in the notation):

Theorem 2  For each $r \geq 1$, there is a distribution $F_r$ on $\mathbb{R}^r$ such that:

$$
\left( m^{1/6} (\lambda_{GUE,m}^k - 2\sqrt{m}) \right)_{1 \leq k \leq r} \Rightarrow F_r, \quad m \to +\infty.
$$

(4)

Remark 3  The distribution $F_r$ is explicitly known (see (3.48) in [Jo]) 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 $r \geq 1$:

$$
\lim_{m \to +\infty} \lim_{n \to +\infty} \left( \frac{(V_k(n, m) - kn/m - 2k\sqrt{n})}{\sqrt{n}} \times m^{2/3} \right)_{1 \leq k \leq r} = F_r \Theta_r^{-1}.
$$

(5)
In fact, since \( \lim_{m \to +\infty} m^{1/6} Z_m = 0 \), in probability, and in view of (3), throughout, in studying weak asymptotics, one will be able to replace the correlated Brownian motions of (2) by uncorrelated (standard) ones.

Following universality argument in percolation models developed by Bodineau and Martin ([BMI]), 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 \( n \)th letter as in the case with the model studied in [HIL]). 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 lighten the notation, we shall still write \( m \) instead of \( m(n) \). A main result of this note is:

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

\[
\left( \frac{V_k(n, m) - kn/m - 2k\sqrt{n}}{n^{1/2} m^{-2/3}} \right)_{1 \leq k \leq r} \Rightarrow F_r \Theta_r^{-1}, \quad n \to +\infty.
\]

Remark 9, 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 \( F_r \) is the Tracy-Widom distribution \( F_{TW} \), we have:

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

\[
\frac{V_1(n, m) - (n/m) - 2n^{1/2}}{n^{1/2} m^{-2/3}} \Rightarrow F_{TW}, \quad n \to +\infty.
\]

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 \( \mathbb{P}(X_i^m = \alpha_j) = p_j^m \), let \( p_{\text{max}}^m = \max_{1 \leq j \leq m} p_j^m \), and let also \( J(m) = \{ j : p_j^m = p_{\text{max}}^m \} = \{ j_1, \ldots, j_{k(m)} \} \) with \( k(m) = \text{card} (J(m)) \). Now, from [HL1] and as \( n \to +\infty \), the behavior of the first row of the Young tableau in this non-uniform setting is given by:

\[
\frac{V_1(n, m) - p_{\text{max}}^m n}{\sqrt{p_{\text{max}}^m n}} \Rightarrow \sqrt{1 - k(m)p_{\text{max}}^m} k(m)^{-1} \sum_{j=1}^{k(m)} B_j(1) + \max_{0 = t_0 \leq t_1 \leq \ldots \leq t_{k(m)} - 1} \sum_{l=1}^{k(m)} (B_l'(t_l) - B_l'(t_{l-1})),
\]

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

\[
Z_k = \frac{1}{k} \sum_{j=1}^{k} B_j(1) \quad \text{and} \quad D_k = \max_{0 = t_0 \leq t_1 \leq \ldots \leq t_{k-1}} \sum_{l=1}^{k} (B_l'(t_l) - B_l'(t_{l-1})),
\]

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:

\[
k^{1/6} (D_k - 2\sqrt{k}) \Rightarrow F_{TW}, \quad k \to +\infty.
\]

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 \( p_{\text{max}}^m \in [0, 1] \), we can also assume that \( p_{\text{max}}^m \to p_{\text{max}} \). In this case, taking the limit first in \( n \) and next in \( m \) yields:

\[
\frac{V_1(n, m) - p_{\text{max}}^m n}{\sqrt{p_{\text{max}}^m n}} \Rightarrow (\sqrt{1 - kp_{\text{max}}} - 1) Z_k + D_k.
\]
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 \( (\sqrt{1 - k(m)} p_{max}^m - 1)^2 k(m)^{-2/3} \leq (k(m)p_{max}^m)^2(k(m)^{-2/3} \leq k(m)^{-2/3} \to 0 \), when \( m \to +\infty \):

\[
(\sqrt{1 - k(m)} p_{max}^m - 1)Z_{k(m)} k(m)^{1/6} \sim N(0, (\sqrt{1 - k(m)} p_{max}^m - 1)^2 k(m)^{-2/3}) \to 0.
\]

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

\[
V_1(n, m) - p_{max}^m n - 2\sqrt{k(m)} p_{max}^m n k(m)^{2/3} \Rightarrow F_{TW}
\]

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 \( A_m \). More precisely, let \( p_{2nd}^m = \max(p_j^m : 1 \leq j \leq m) \):

**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 \( A_m \), goes to infinity when \( n \to +\infty \), in such a way that \( k(m(n))^7/10 / p_{max}^{3/10} = o(n^{3/10}(\log n)^{-3/5}) \). Assume moreover that

\[
(p_{2nd}^m)^2 (n^{11/10}/(\log n)^{1/5}) = o(p_{max}^m).
\]

Then

\[
V_1(n, m(n)) - p_{max}^m n - 2\sqrt{k(m(n)) p_{max}^m n} k(m(n))^{2/3} \Rightarrow F_{TW}.
\]

Let us stress again the fact that in the previous result, \( m \) is a function of \( n \), with the only requirement that \( k(m(n))^{7/10} / p_{max}^{3/10} = o(n^{3/10}(\log n)^{-3/5}) \). Note that in the uniform case, \( k(m) = m \) and \( p_{max}^m = 1/m \) and that in general \( 1/m \leq p_{max}^m \leq 1/k(m) \).

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 [BM] showed that the fluctuations of the last-passage directed percolation model with Gaussian \( iid \) weights actually extend to \( iid \) 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 [GW] in a related queueing model.

Here, we closely follow [BM] 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 strong integrability property, the strong approximation can be made more precise than in [BM]. But, this is not enough to obtain the fluctuations for \( m \) of larger order. Actually the Gaussian approximation is responsible for the condition \( m = o(n^{3/10}(\log n)^{-3/5}) \), which comes short of the corresponding polynomial order condition \( m = o(n^{3/7}) \) obtained in [BM]. However in contrast to [BM], 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 Skorohod embedding, Baik and Suidan [BS] derived, independently of [BM], similar convergence results (see Theorem 2 in [BS]), under the condition \( m = o(n^{3/14}) \). See also [Su] for related results (under \( m = o(n^{1/7}) \)) in percolation models using functional methods in the CLT.

Finally, note that [BM, BS, Su] 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 \((B^i(s))_{s \geq 0}, 1 \leq l \leq m\), be independent standard Brownian motions. For \( s > 0, m \geq 1 \) and \( k \geq 1 \), let

\[
L_k(s, m) = \sup_{t \in I_{k,m}(s)} \sum_{j=1}^{k} \sum_{l=1}^{m-k+j} (B^l(t,j,i) - B^l(t,j,i-1)),
\]

where \( I_{k,m}(s) = \{ st, t \in I_{k,m} \} \). For \( k = 1 \), \( L_1(s, m) \) coincides with the Brownian percolation model used in [BM], see also [GW] for a related queueing model. For \( s = 1, \Theta_m^{-1}((L_k(1, m))_{1 \leq k \leq m}) \) has the same law as the spectrum of a \( m \times m \) GUE matrix, see [Do] and [HX].

Since \((L_1(., m), \ldots, L_m(., n))\) is a continuous function of \( B^1, \ldots, B^m \), which are independent; Brownian scaling entails:

\[
(L_1(s, m), \ldots, L_m(s, m)) \overset{\mathcal{L}}{=} \sqrt{s}(L_1(1, m), \ldots, L_m(1, m)).
\]

Plugging (13) into (4) yields, as \( m \rightarrow +\infty \),

\[
\left( \frac{L_k(n, m) - 2k\sqrt{nm}}{n^{1/2}m^{-1/6}} \right)_{1 \leq k \leq r} \Rightarrow \mathcal{F}_r \Theta_r^{-1}.
\]

Combinatorics. Let

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

be Bernoulli random variables with parameter \( P(X_{i,j}^m = \alpha_j) = 1/m \) 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 \( j \neq j' \), \((X_{1,j}^m, \ldots, X_{n,j}^m)\) and \((X_{1,j'}^m, \ldots, X_{n,j'}^m)\) are identically distributed but no longer independent.

Recall again 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 \( X_1^m, \ldots, X_n^m \).

Let \( S_{k,j}^m = \sum_{i=1}^{k} X_{i,j}^m \) be the number of occurrences of \( \alpha_j \) among \((X_{i,j}^m)_{1 \leq i \leq k}\). An increasing subsequence of \((X_{i,j}^m)_{1 \leq i \leq k}\) consists of successive blocks, each one made of an identical letter, with the sequence of letters representing each block being strictly increasing. Since for \( 1 \leq k < l \leq n \) the number of occurrences of \( \alpha_j \) among \((X_{i,j}^m)_{k \leq i \leq l}\) is \( S_{l,j}^m - S_{k,j}^m \), and it follows that:

\[
V_k(n, m) = \max_{0 = t_0 \leq t_1 \leq \ldots \leq t_m = n} \left[ (S_{t_1}^{m,1} - S_0^{m,1}) + (S_{t_2}^{m,2} - S_{t_1}^{m,2}) + \cdots + (S_{t_m}^{m,m} - S_{t_{m-1}}^{m,m}) \right],
\]

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 [HL2]):

\[
V_k(n, m) = \max_{k \in I_{k,m}(n)} \sum_{j=1}^{k} \sum_{l=1}^{m-k+j} (S_{k,j}^{m,l} - S_{k,j,l-1}^{m,l}),
\]
where

\[ J_{r,m}(n) = \{ \mathbf{k} = (k_{j,l} : 1 \leq j \leq r, 0 \leq l \leq m) : k_{j,j-1} = 0, k_{j,m-r+j} = n, 1 \leq j \leq r, \\
  k_{j,l} \leq k_{j,l+1}, 1 \leq j \leq r, 1 \leq l \leq m-1; k_{j,l} \leq k_{j-1,l-1}, 2 \leq j \leq r, 1 \leq l \leq m \}. \]

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

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

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

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

\[
\begin{align*}
V_k(n,m) - kn/m &= 2k \sqrt{n} / m^{2/3} \\
&= \sigma_m \tilde{V}_k(n,m) - 2k \sqrt{n} / m^{2/3} \\
&= \tilde{V}_k(n,m) - 2k \sqrt{n} \sigma_m^{-1} / m^{2/3} (\sigma_m m^{2/3}) \\
&= \tilde{V}_k(n,m) - 2k \sqrt{nm} + 2k \sqrt{n} (\sigma_m^{-1} - m^{1/2}) / (m^{1/2} m^{-1/6}) \\
&= \tilde{V}_k(n,m) - 2k \sqrt{nm} / m^{1/2} m^{-1/6}. \tag{18}
\end{align*}
\]

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

\[
\tilde{V}_k(n,m) - 2k \sqrt{nm} / m^{1/2} m^{-1/6}.
\]

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

\[
\begin{align*}
\left| \tilde{V}_k(n,m) - L_k(n,m) \right| &= \sup_{\mathbf{t} \in I_{r,m}(n)} \left| \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} (S_{m,l}^{n,m} - S_{m,l}^{n,n-1}) - \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} (B^l(t_{j,i}) - B^l(t_{j,i-1})) \right| \\
&\leq \sup_{\mathbf{t} \in I_{r,m}(n)} \left| \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left( S_{m,l}^{n,m} - S_{m,l}^{n,n-1} \right) - \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left( B^l(t_{j,i}) - B^l(t_{j,i-1}) \right) \right| \\
&= \sup_{\mathbf{t} \in I_{r,m}(n)} \left| \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left( \tilde{S}_{m,l}^k - B^l(t_{j,i}) \right) \right| - \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left( \tilde{S}_{m,l}^k - B^l(t_{j,i-1}) \right) \\
&= \sup_{\mathbf{t} \in I_{r,m}(n)} \left| \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left( (S_{m,l}^{n,m} - B^l(t_{j,i})) + (B^l(t_{j,i}) - B^l(t_{j,i-1})) \right) \\
&- (S_{m,l}^{n,m} - B^l(t_{j,i-1})) - (B^l(t_{j,i-1}) - B^l(t_{j,i-1})) \right|.
\end{align*}
\]
\[
\begin{align*}
\leq \sup_{t \in \mathbb{R}} & \left\{ \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \left( |S_{i,t,j}^{m,l} - B^l([t,j,j])| + |B^l([t,j,l]) - B^l([t,j,l])| ight) \\
& + |S_{i,t,j}^{m,l} - B^l([t,j,l-1])| + |B^l([t,j,l-1]) - B^l([t,j,l-1])| \right\} \\
\leq 2k \sum_{l=1}^{m} (Y_{n,l} + W_{n,l}),
\end{align*}
\]

where we set
\[
Y_{n,1} = \max_{1 \leq i \leq n} |S_{i}^{m} - B^l(i)| \quad \text{and} \quad W_{n,l} = \sup_{0 \leq t \leq n} |B^l(s) - B^l(t)|.
\]

**Gaussian approximation.** From now on, we assume that for each \( n \) and \( l \in [1, m] \) (recall that \( m = m(n) \)), the random variables \( X_{i,t}^{m,n} \), \( 1 \leq i \leq n \), and the Brownian motion \((B^l(s))_{s \in [0,n]}\), appearing in \( Y_{n,1} \) and \( W_{n,l} \) (rewritten as \((B^l(s))_{s \in [0,n]}\)), are given by the following result, which is a compilation of strong approximation results of Komlós, Major, Tusnády and of Sakhanenko and for which we refer to [L] (Th. 2.1, Cor 3.2) and the references therein. In the sequel, we write \( B_{n}^{m,l} \) and \( W_{n,l} \), instead of \( B^l \) and \( W_{n,l} \), to insist on the dependence in \( m \) of the random variables given by the forthcoming proposition.

**Proposition 7** Let \((X_{n})_{n} \geq 1 \) be a sequence of iid 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 \((\tilde{X}_{n})_{1 \leq n \leq N} \) having the same law as \((X_{n})_{1 \leq n \leq N} \), and independent Gaussian variables \((\tilde{Y}_{n})_{1 \leq n \leq N} \) having same expectations and variances as \((X_{n})_{1 \leq n \leq N} \) such that for every \( x > 0 \):

\[
\mathbb{P} \left( \max_{1 \leq k \leq N} \left| \sum_{j=1}^{k} \tilde{X}_{j} - \sum_{j=1}^{k} \tilde{Y}_{j} \right| \geq x \right) \leq (1 + c_2(F)N^{1/2}) \exp(-c_1(F)x),
\]

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

\[
\lambda(F) = \sup \{ \lambda > 0 : \lambda \mathbb{E}[|X_{1} - \mathbb{E}[X_{1}]|^3 \exp(\lambda|X_{1} - \mathbb{E}[X_{1}]|)] \leq \mathbb{E}[|X_{1} - \mathbb{E}[X_{1}]|^2] \}.
\]

The strong approximation entails the following bound for the tail of \( Y_{n,1}^{m} \):

\[
\mathbb{P}(Y_{n,1}^{m} \geq x) \leq (1 + c_2(F)n^{1/2}) \exp(-c_1(F)n),
\]

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

\[
\mathbb{E}[|X_{1,1}^{m} - \mathbb{E}[X_{1,1}^{m}]|^2] = \frac{1}{m} \left( 1 - \frac{1}{m} \right) \lambda \leq \frac{1}{m} \left( 1 - \frac{1}{m} \right)^{\frac{\lambda}{2}} \leq \frac{1}{m} \left( 1 - \frac{1}{m} \right) \left( 1 - \frac{2}{m} + \frac{2}{m^2} \right) \lambda = \lambda \mathbb{E}[(X_{1,1}^{m} - \mathbb{E}[X_{1,1}^{m}])^2] \leq \lambda \mathbb{E}[(X_{1,1}^{m} - \mathbb{E}[X_{1,1}^{m}])^3 \exp(\lambda|X_{1,1}^{m} - \mathbb{E}[X_{1,1}^{m}]|)]
\]

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

\[
\frac{1}{2} \mathbb{E} \left[ |X_{1,1}^{m} - \mathbb{E}[X_{1,1}^{m}]|^3 \exp \left( \frac{1}{2} |X_{1,1}^{m} - \mathbb{E}[X_{1,1}^{m}]| \right) \right] \leq \frac{1}{2} \mathbb{E} \left[ |X_{1,1}^{m} - \mathbb{E}[X_{1,1}^{m}]|^2 \right] \leq \mathbb{E}[|X_{1,1}^{m} - \mathbb{E}[X_{1,1}^{m}]|^2].
\]

Thus, \( c_1(m) \) and \( c_2(m) \) behave like \( 1/\sqrt{m} \). Note also that the bound in (20) is non-trivial for \( x \geq \tilde{a}_n := \log(1 + c_2(m)n^{1/2})/c_1(m) \).
Remark 8 In order to obtain KMT bounds in our framework, we first apply Proposition 7, individually for each $1 \leq l \leq m$ to construct $(\tilde{X}^m_{i,l}) : 1 \leq i \leq n$, and the Brownian motion $B^l$ satisfying (20) on some probability space $(\Omega_l, \mathcal{F}_l, \mathbb{P}_l)$. At this point, on the probability space $(\Omega_1 \times \cdots \times \Omega_m, \mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_m, \mathbb{P} \otimes \cdots \otimes \mathbb{P}_m)$ the Brownian motions $B^l, 1 \leq l \leq m$, are rendered independent and so are, for different $l$, the independent Gaussian random variables $(Y^l_i)_{1 \leq i \leq n}$. Next, we explain that this suffices and that we can consider the Brownian motion $B^l$, $1 \leq l \leq m$, to be correlated via $\Sigma(m)$, given by $\Sigma_{i,i}^m = 1$, $\Sigma_{i,j}^m = -1/(m - 1)$. Indeed, setting $\mathbb{b}(1/m)$ for the centered Bernoulli distribution $b(1/m)$ normalized with unit variance, we have
\[
\left((\tilde{X}^m_{i,1})_{1 \leq i \leq n}, \ldots, (\tilde{X}^m_{i,m})_{1 \leq i \leq n}, (Y^1_i)_{1 \leq i \leq n}, \ldots, (Y^m_i)_{1 \leq i \leq n}\right) \sim \mathbb{b}(1/m)^{\otimes nm} \otimes \mathcal{N}(0, 1)^{\otimes nm}. \tag{21}
\]

Now, consider the vector
\[
\left((U^m_{i,1})_{1 \leq i \leq n}, \ldots, (U^m_{i,m})_{1 \leq i \leq n}, (V^1_i)_{1 \leq i \leq n}, \ldots, (V^m_i)_{1 \leq i \leq n}\right) \sim Q \tag{22}
\]
where $U^m_{i,l} \sim \mathbb{b}(1/m)$, $V^l_i \sim \mathcal{N}(0, 1)$ and where both $(U^m_{i,l})_{1 \leq i \leq n}$, $(V^l_i)_{1 \leq i \leq m}$ are correlated by $\Sigma(m)$. The distribution $Q$, in (22), is absolutely continuous with respect to the distribution given in (21), and let us denote by $f$ its Radon-Nikodým derivative. Then, consider the probability space $(\Omega_1 \times \cdots \times \Omega_m, \mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_m, \mathbb{P}_f)$ where
\[
d\mathbb{P}_f = f((\tilde{X}^m_{i,1})_{1 \leq i \leq n}, \ldots, (\tilde{X}^m_{i,m})_{1 \leq i \leq n}, (Y^1_i)_{1 \leq i \leq n}, \ldots, (Y^m_i)_{1 \leq i \leq n})d\mathbb{P},
\]
and where $\mathbb{P} = \mathbb{P}_1 \otimes \cdots \otimes \mathbb{P}_m$. Observe that under $\mathbb{P}_f$, the vector
\[
\left((\tilde{X}^m_{i,1})_{1 \leq i \leq n}, \ldots, (\tilde{X}^m_{i,m})_{1 \leq i \leq n}, (Y^1_i)_{1 \leq i \leq n}, \ldots, (Y^m_i)_{1 \leq i \leq n}\right)
\]
has distribution $Q$ and, therefore, both $(\tilde{X}^m_{i,l})_{1 \leq i \leq m}$ and $(Y^l_i)_{1 \leq i \leq m}$ are correlated via $\Sigma(m)$. In turn, the Brownian motions $B^l$, $1 \leq l \leq m$, are also correlated via $\Sigma(m)$. In the sequel, up to applying this argument, we assume that the Brownian motions are correlated via $\Sigma(m)$, but we keep our notations unchanged (see also our next comment).

Alternatively, if instead of using Proposition 7 one uses Theorem 4.1 of [Li], each random variable (vector) in this theorem should be $m$ (the alphabet size)-dimensional and there should be $n$ (the number of letters in the word) many of them. Then, one constructs a KMT-approximation with independent copies of these vectors but where each vector has its coordinates taken correlated with the same correlation matrix as each random vector associated with the $i$th letter of the word. More precisely, using the terminology of [Li], one can construct $n$ random vectors, which are independent copies of the $X_i$’s there, denoted by $\tilde{X}_i$, in such a way that they remain independent of each other but with the same covariance matrix as the $X_i$’s, as well as $n$ independent Gaussian random vectors $Y_i$’s with the same covariance matrix, for which a KMT approximation hold. In our case, let
\[
X_{i,j} = X^m_{i,j} - \mathbb{E}[X^m_{i,j}],
\]
where $X^m_{i,j}$ is defined at the beginning of Section 2. Now, for fixed $i$, $X^m_{i,j}$ and $X^m_{i,k}, j \neq k$ are orthogonal random variables and so the covariance matrix $\Sigma = (\Sigma_{k,\ell})_{1 \leq k,\ell \leq m}$ of the $X_{i,j}$, where $i$ is a fixed letter of the word while $k$ and $\ell$ run over the alphabet of size $m$ is such that
\[
\Sigma_{k,\ell} = \mathbb{E}[X_{i,k}X_{i,\ell}] = \begin{cases} 
\frac{m-1}{m^2}, & \text{if } k = \ell, \\
-\frac{1}{m^2}, & \text{if } k \neq \ell.
\end{cases}
\]
Then, taking $X_i = (X_{i,1}, \ldots, X_{i,m})$ in Theorem 4.1 of [Li], one can further take $\tilde{X}_i$ in the same theorem to have correlated entries with covariance matrix $\Sigma$. Now, $\Sigma$ has eigenvalues $1/m$ with multiplicity $m$. Therefore, the uniform non-degeneracy conditions (4.1) of Theorem 4.1 (which seems to contain a typo with an unnecessary extra $D^2$) are satisfied and so its conclusions apply. Using this dependent version of the KMT approximation might lead to a different alphabet-growth rate $\alpha$, after evaluating the various parameters.
Approximating sets. Let \( A_1^n = \{ \max_{1 \leq m} Y_{n,m}^l > a_n \} \), for some \( a_n = Cc_1(m)^{-1}(\log n)^2 \leq \tilde{a}_n \) where \( C \) is some finite constant. We have
\[
\mathbb{P}(A_1^n) = \mathbb{P}\left( \bigcup_{l \leq m} \{ Y_{n,m}^l > a_n \} \right)
\leq \sum_{l \leq m} \mathbb{P}(Y_{n,m}^l > a_n)
\leq me^{-c_1(m)a_n}(1 + c_2(m)n^{1/2})
\sim \sqrt{mn}e^{-c_1(m)a_n}
= \sqrt{mn}e^{-(c_3C(\log n)^2)/2} \rightarrow 0, \quad n \rightarrow +\infty.
\]

Let \( A_2^n = \{ \max_{1 \leq t \leq m} W_{n,m}^l > b_n \} \), for \( b_n = \log n \). Standard estimates (including reflection principle, Brownian scaling and Gaussian tail estimates) lead to :
\[
\mathbb{P}(A_2^n) = \mathbb{P}\left( \bigcup_{l \leq m} \{ W_{n,m}^l > b_n \} \right)
\leq \sum_{l \leq m} \mathbb{P}(W_{n,m}^l > b_n)
\leq m \mathbb{P}(W_{n,m}^l > b_n)
= m \mathbb{P}\left( \sup_{0 \leq s,t \leq \eta} |B_{s}^{m,1} - B_{t}^{m,1}| > b_n \right).
\]

But,
\[
\sup_{0 \leq s,t \leq \eta} |B_{s}^{m,1} - B_{t}^{m,1}| \leq \sup_{0 \leq s \leq \eta - 2, t \leq s, t \leq t + 2} \sup_{|s-t| \leq 1} |B_{s}^{m,1} - B_{t}^{m,1}|
\leq \sup_{0 \leq s \leq \eta - 2} \left( \sup_{t \leq \eta + 2} B_{s}^{m,1} - \inf_{t \leq \eta + 2} B_{s}^{m,1} \right),
\]

and so
\[
\mathbb{P}(A_2^n) \leq m \mathbb{P}\left( \sup_{0 \leq t \leq \eta - 2} \left( \sup_{t \leq \eta + 2} B_{s}^{m,1} - \inf_{t \leq \eta + 2} B_{s}^{m,1} \right) > b_n \right)
\leq mn \mathbb{P}\left( \sup_{t \leq [0,2]} B_{t}^{m,1} - \inf_{s \leq [0,2]} B_{s}^{m,1} > b_n \right)
\leq mn \left( \mathbb{P}\left( \sup_{t \leq [0,2]} B_{t}^{m,1} > b_n/2 \right) + \mathbb{P}\left( \sup_{s \leq [0,2]} B_{s}^{m,1} > b_n/2 \right) \right)
\leq 2mn \exp(-b_n^2/16) \rightarrow 0, \quad n \rightarrow +\infty.
\]

Final bound. Since the Brownian motions \( B^l, 1 \leq l \leq m, \) are correlated via \( \Sigma^{(m)} \) (see Remark 8), combining (2) and (3) and the observation made after (12), we have that \( (L_k(n,m))_{1 \leq k \leq r} \), for independent Brownian motions, and \( (L_k(n,m))_{1 \leq k \leq r} \), for Brownian motions correlated via \( \Sigma^{(m)} \), only differ by \( Z_m e_m \). But since \( Z_m m^{1/6} \Rightarrow 0 \), the limiting result (14) still applies for \( (L_k(n,m))_{1 \leq k \leq r} \) with Brownian motions correlated via \( \Sigma^{(m)} \) (see the paragraph after (5)). As a consequence, the approximation of \( (\tilde{V}_k(n,m))_{1 \leq k \leq r} \) by \( (L_k(n,m))_{1 \leq k \leq r} \) will imply the theorem if
\[
\mathbb{P}\left( \sum_{k=1}^r \left| \tilde{V}_k(n,m) - L_k(n,m) \right| \geq c_n \right) \rightarrow 0, \quad n \rightarrow +\infty,
\]
for some
\[ c_n = o(n^{1/2}m^{-1/6}). \tag{25} \]
Since \( \lim_{n \to +\infty} \left( \mathbb{P}(A_1^n) + \mathbb{P}(A_2^n) \right) = 0 \), it is enough to prove that
\[
\lim_{n \to +\infty} \mathbb{P} \left( \left\{ \sum_{k=1}^{r} \tilde{V}_k(n, m) - L_k(n, m) \geq c_n \right\} \cap (A_1^n)^c \cap (A_2^n)^c \right) = 0. \tag{26}
\]
But
\[
\mathbb{E} \left[ \sum_{k=1}^{r} \left| \tilde{V}_k(n, m) - L_k(n, m) \right| 1_{(A_1^n)^c \cap (A_2^n)^c} \right] \\
\leq 2r^2m \left( \mathbb{E} [Y_n^{m,1} 1_{Y_n^{m,1} \leq a_n}] + b_n \right) \\
\leq 2r^2m \left( \mathbb{E} [Y_n^{m,1} - \tilde{a}_n] 1_{Y_n^{m,1} \leq a_n} + \tilde{a}_n + b_n \right) \\
\leq 2r^2m \left( \mathbb{E} [(Y_n^{m,1} - \tilde{a}_n) 1_{\tilde{a}_n \leq Y_n^{m,1} \leq a_n}] + \tilde{a}_n + b_n \right) \\
\leq 2r^2m \left( \int_{\tilde{a}_n}^{a_n} \mathbb{P}(Y_n^{m,1} \geq x) dx + \tilde{a}_n + b_n \right) \\
\leq 2r^2m \left( \int_{\tilde{a}_n}^{a_n} e^{-c_1(m)x}(1 + c_2(m)n^{1/2}) dx + \tilde{a}_n + b_n \right) \\
\leq 2r^2m \left( \frac{1 + c_2(m)n^{1/2}}{c_1(m)} e^{-c_1(m)\tilde{a}_n} + \tilde{a}_n + b_n \right) \\
\leq 2r^2m \left( \frac{1}{c_1(m)} + \tilde{a}_n + b_n \right) \\
\leq 2r^2m^{3/2} \left( \frac{2(1 + \log(1 + c_2(m)n^{1/2}))}{c_3} + b_n \right).
\]
Finally,
\[
\mathbb{P} \left( \left\{ \sum_{k=1}^{r} \tilde{V}_k(n, m) - L_k(n, m) \geq c_n \right\} \cap (A_1^n)^c \cap (A_2^n)^c \right) \\
\leq \frac{2r^2m^{3/2}}{c_n} \left( \frac{2(1 + \log(1 + c_2(m)n^{1/2}))}{c_3} + \log n \right) = O \left( \frac{m^{3/2} \log n}{c_n} \right). \tag{27}
\]
A choice of \( c_n \) ensuring that the bound in (27) goes to zero as \( n \to +\infty \) and also compatible with (25) is possible when \( m^{3/2} \log n = o(n^{1/2}m^{-1/6}) \), i.e., when \( m = o(n^{3/10}(\log n)^{-3/5}) \). Finally, (24) and (26) hold true, achieving the proof of Theorem 4.

\[ \square \]

**Remark 9**

- In the above proof, the condition \( m = o(n^{3/10}(\log n)^{-3/5}) \) is needed only once, to ensure the compatibility of (25) with the bound (27). However, this is essential to make the Gaussian approximation work.
- When \( m = [n^a] \), the growth condition \( m = o(n^{3/10}(\log n)^{-3/5}) \) rewrites as \( a < 3/10 \), and this growth condition remains true, in particular, when \( m \) is of sub-polynomial order. The condition \( a < 3/10 \) is stronger than its counterpart \( a < 3/7 \) in [BM] 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 is short of a result of Johansson in [Jo] which asserts the convergence of \( V_1(n, n^a) \) (properly scaled and normalized) towards the Tracy-Widom distribution. More precisely, setting \( a_n \ll b_n \) for \( a_n = o(b_n) \), Th. 1.7 in [Jo] actually gives in our notations: for \( \sqrt{n} \ll m, \)

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

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

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

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

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

In the middle limit above, [Jo, Th. 1.7] requires \((\log n)^{3/2} = o(m)\) 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 \( V_1(n, m) \) 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 \( X_{\alpha} \), 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 only over most probable letters as in (16), see (30). Part of the remaining work is then to show that we can suitably control this approximation and this is done in Lemma 10. 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, \( X^m_\alpha \), \( 1 \leq \alpha \leq n, \) are independently and identically distributed with \( \mathbb{P}(X^m_\alpha = \alpha_j) = p^m_{\alpha_j} \) and \( J(m) = \{ j : p^m_{\alpha_j} = p^m_{\alpha_{\max}} \} = \{ j_1, \ldots, j_{k(m)} \} \), with \( k(m) = \text{card}(J(m)) \), and set also \( \sigma^2_m = p^m_{\alpha_{\max}}(1 - p^m_{\alpha_{\max}}) \). Finally, note that since \( k(m(n))p^m_{\alpha_{\max}} \leq 1 \) and \( k(m(n)) \to +\infty \), it follows that \( p^m_{\alpha_{\max}} \to 0 \), as \( n \to +\infty \).

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

\[
L_1(s, k(m)) = \sup_{t \in J_{k(m)}(s)} \sum_{l=1}^{k(m)} (B^l(t_l) - B^l(t_{l-1})),
\]

where \( J_{k(m)}(s) = \{ t : 0 \leq t_1 \leq \cdots \leq t_{l-1} \leq t_l \leq \cdots \leq t_{k(m)} = s \} \). 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:

\[
k^{1/6}(L(1, k) - 2\sqrt{k}) \Rightarrow F_{TW}.
\]

By Brownian scaling, \( L_1(s, m) \overset{d}{=} \sqrt{s}L_1(1, m) \), so that when \( n \to +\infty \):

\[
\frac{L_1(n, k(m(n))) - 2\sqrt{n k(m(n))}}{n^{1/2}k(m(n))^{-1/6}} \Rightarrow F_{TW}.
\]
Combinatorics revisited. Let

\[ X_{i,j}^m = \begin{cases} 1 & \text{when } X_{i,j}^m = \alpha_j \\ 0 & \text{otherwise,} \end{cases} \]

be Bernoulli random variables with parameter \( \mathbb{P}(X_{i,j}^m = \alpha_j) = p_{i,j}^m \) and variance \((\sigma_{i,j}^m)^2 = p_{i,j}^m(1 - p_{i,j}^m)\). For a fixed \(1 \leq j \leq m\), the \( X_{i,j}^m \)'s are independent and identically distributed. Since the expression (15) has a purely combinatorial nature, we still have

\[
V_1(n, m) = \max_{0 = l_0 \leq l_1 \leq \cdots \leq l_{m-1} \leq l_{m} = n} \left( \sum_{j=1}^{m} \sum_{i=l_{j-1}+1}^{l_{j}} X_{i,j}^m \right),
\]

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

\[
V_1'(n, m) = \max_{0 = l_0 \leq l_1 \leq \cdots \leq l_{m-1} \leq l_{m} = n} \left( \sum_{j=1}^{m} \sum_{i=l_{j-1}+1}^{l_{j}} X_{i,j}^m \right),
\]

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 if the \( n \) letters drawn are letters with associated probability strictly less than \( p_{j,m}^{\text{max}} \), \( V_1'(n, m) = 0 \) while there is a \( l = (l_j)_{j=0,\ldots,m} \) with \( 0 = l_0 \leq l_1 \leq \cdots \leq l_{m-1} \leq l_{m} = n \) 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 \( \tilde{X}_{i,j}^m = (X_{i,j}^m - p_{i,j}^m)/\sigma_{i,j}^m \) be the corresponding centered and normalized scaled Bernoulli random variables and let \( \tilde{S}_{i,j}^m = \sum_{i=1}^{l_{j}} \tilde{X}_{i,j}^m \). Let also,

\[
\tilde{V}_1'(n, m) = \max_{0 = l_0 \leq l_1 \leq \cdots \leq l_{k(m)} = n} \left( \sum_{j \in J(m)} \sum_{i=l_{j-1}+1}^{l_{j}} \tilde{X}_{i,j}^m \right)
\]

\[
= \max_{0 = l_0 \leq l_1 \leq \cdots \leq l_{k(m)} = n} \left( \sum_{j \in J(m)} \left( \tilde{S}_{i,j}^m \right) \right)
\]

\[
= \max_{0 = l_0 \leq l_1 \leq \cdots \leq l_{k(m)} = n} \left( \sum_{j \in J(m)} \left( \tilde{S}_{i,j}^m \right) \right)
\]

\[
= \sup_{0 \leq l_{m} = n} \left( \sum_{j \in J(m)} \left( \tilde{S}_{i,j}^m \right) \right)
\]

which is to be compared to (28). Since \( V_1'(n, m) - np_{i,j}^{\text{max}} = \sigma_{i,j}^m \tilde{V}_1'(n, m) \), we have

\[
k(m)^{1/6} V_1'(n, m) - np_{i,j}^{\text{max}} - 2 \sqrt{nk(m)\sigma_{i,j}^m} \frac{p_{i,j}^{\text{max}}}{\sigma_{i,j}^m}
\]

\[
= k(m)^{1/6} \tilde{V}_1'(n, m) - 2 \sqrt{nk(m)\sigma_{i,j}^m}
\]

Since \( \sigma_{i,j}^m \sim \sqrt{p_{i,j}^{\text{max}}} \) and

\[
2 \sqrt{k(m)np_{i,j}^{\text{max}} - 2 \sqrt{k(m)n\sigma_{i,j}^2}} = \frac{2 \sqrt{k(m)n\sigma_{i,j}^2}}{\sqrt{n} \sigma_{i,j}^m}
\]

\[
= \frac{2 \sqrt{k(m)n\sigma_{i,j}^2}}{\sigma_{i,j}^m} \sqrt{p_{i,j}^{\text{max}} + \sqrt{\sigma_{i,j}^2}}
\]
it remains to show
\[ k(m)^{1/6} \frac{\tilde{V}_1'(n, m) - 2\sqrt{nk(m)}}{\sqrt{n}} \rightarrow F_{TW} \] for which we shall use (29).

**Sketch of proof of (31).** Roughly speaking, the proof of (31) is along the same lines of the corresponding proof of the convergence of (18), changing only \( m \) into \( k(m) \). We show that when \( k(m(n)) = o(n^{3/10}(\log n)^{-3/5}) \), then for some Brownian motions given via strong approximation, we have
\[ \left| \tilde{V}_1'(n, m) - L_1(n, k(m(n))) \right| \leq \sum_{i=1}^{k(m(n))} (Y_{n,i}^m + W_{n,i}^m), \]
where
\[ Y_n^m = \max_{1 \leq i \leq n} |S_i^m - B_i^m| \quad \text{and} \quad W_n^m = \sup_{0 \leq s, t \leq n} |B^m(s) - B^m(t)|. \]

Indeed, setting \( A_1^n = \{ \max_{1 \leq i \leq k(m(n))} Y_{n,i}^m > a_n \} \), for some \( a_n = O(c_1(k(m(n)))^{-1}(\log n)^2) \), where \( \tilde{a}_n := \log(1 + c_2(k(m(n)))n^{1/2})/c_1(k(m(n))) \), and setting \( A_2^n = \{ \max_{1 \leq i \leq k(m(n))} W_{n,i}^m > b_n \} \), for some \( b_n = O(\log n) \), we show that
\[ P(A_1^n) \rightarrow 0, \quad P(A_2^n) \rightarrow 0, \quad \text{when} \quad n \rightarrow +\infty. \]

From (29), the approximation of \( \tilde{V}_1(n, k(m(n))) \) by \( L_1(n, k(m(n))) \) will imply the theorem if
\[ P\left( \left| \tilde{V}_1'(n, k(m(n))) - L_1(n, k(m(n))) \right| \geq c_n \right) \rightarrow 0, \quad n \rightarrow +\infty, \quad \text{(32)} \]
for some
\[ c_n = o(n^{1/2}k(m(n))^{-1/6}). \quad \text{(33)} \]

Since \( \lim_{n \rightarrow +\infty} (P(A_1^n) + P(A_2^n)) = 0 \) and
\[ P\left( \left\{ \left| \tilde{V}_1'(n, k(m(n))) - L_1(n, k(m(n))) \right| \geq c_n \right\} \cap (A_1^n)^c \cap (A_2^n)^c \right) \leq \frac{2k(m(n))^{3/2}}{c_n} \left( \frac{2(1 + \log(1 + c_2(k(m(n)))n^{1/2}))}{c_3} + \log n \right), \quad \text{(34)} \]
a choice of \( c_n \), ensuring that the bound in (34) goes to zero and is compatible with (33), is possible since \( k(m(n)) = o(n^{3/10}(\log n)^{-3/5}) \). This proves (31) 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 \( E[V_1(n, m) - V_1'(n, m)] \) to be small. Actually we show:

**Lemma 10** For some absolute constant \( C > 0 \), we have
\[ E[|V_1(n, m) - V_1'(n, m)|] \leq Cnp_{2nd}^m, \quad \text{(35)} \]
where \( p_{2nd}^m \) stands for the second largest probability for the letters of \( A_m \).
The conclusion in (11) holds true when
\[
\lim_{n \to \infty} \left( \mathbb{E}[|V_1(n, m) - V'_1(n, m)|] \times \frac{k(m(n))^{2/3}}{\sqrt{k(m(n))p_{\max}} n} \right) = 0. \quad (36)
\]
But with the help of (35), the conclusion in (36) is then valid when \( \lim_{n \to \infty} \frac{p(n(n))^{1/6} n^{1/2}}{p_{\max}^{1/2}} = 0 \)
and, since \( k(m(n)) = o(p^{3/10}(\log n)^{-3/5}) \), this will follow from (10).

It remains to prove Lemma 10, i.e. to give an explicit bound on \( \mathbb{E}[|V_1(n, m) - V'_1(n, m)|] \). To do so, rewrite \( V_1(n, m) = \max_{i \in I(m)} Z(l) \) and \( V'_1(n, m) = \max_{i \in I^*(m)} Z(l) \) where \( I(m) = \{ l = (l_j)_{1 \leq j \leq m} : l_{j-1} \leq l_j, l_0 = 0, l_m = n \} \), \( I^*(m) = \{ l \in I(m) : l_{j-1} = l_j \text{ for } j \notin J(m) \} \) and
\[
Z(l) = \sum_{j=1}^{m} Y_j(l), \quad Y_j(l) = \sum_{i = j_{l+1}}^{l_j} X_{i,j}^m.
\]
Clearly, since \( I^*(m) \subset I(m) \), we have \( V'_1(n, m) \leq V_1(n, m) \). Moreover, since the \( X_{i,j}^m \) are Bernoulli random variables with parameter \( p_{j,m}^m \) and since the \( X_i \)s are independent, we have \( Y_j(l) \sim \mathcal{B}(l_{j-1}, p_{j}) \) and \( \sum_{j \in J(m)} Y_j(l) \sim \mathcal{B}(\sum_{j \in J(m)} l_j - l_{j-1}, p_{\max}) \), 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 \( J_1 = \{ j \notin J(m) : l_{j-1} < l_j \} \) and for \( j \in J_1 \), define
\[
\theta(j) = \begin{cases} \max A_j & \text{if } A_j \neq \emptyset \\ \min B_j & \text{otherwise,} \end{cases}
\]
where \( A_j = \{ k \in J(m) : k \leq j \} \) and where \( B_j = \{ k \in J(m) : k \geq j \} \). Now,
\[
Z(l) = \sum_{j \in J(m)} \sum_{i = j_{l+1}}^{l_j} X_{i,j}^m + \sum_{j \in J_1} \sum_{i = j_{l+1}}^{l_j} X_{i,j}^m \]
\[
= \sum_{j \in J(m)} \sum_{i = j_{l+1}}^{l_j} X_{i,j}^m + \sum_{j \in J_1} \sum_{i = j_{l+1}}^{l_j} X_{i,\theta(j)}^m \]
\[
+ \sum_{j \in J(m)} \sum_{i = j_{l+1}}^{l_j} (X_{i,j}^m - X_{i,\theta(j)}^m). \quad (37)
\]
Define \( \tilde{l} \in I^*(m) \) by \( \tilde{l}_j = \hat{l}_j - 1 \) if \( j \notin J(m) \) and \( \tilde{l}_j = l_{j-1} \) for \( j \in J(m) \), where \( k = \min\{ l > j : l \in J(m) \} \), with the convention that \( \min \emptyset = m + 1 \), and that \( \tilde{l}_{j_0 - 1} = 0 \) for \( j_0 = \min J(m) \). We then have
\[
\sum_{j \in J(m)} \sum_{i = j_{l+1}}^{l_j} X_{i,j}^m + \sum_{j \in J_1} \sum_{i = j_{l+1}}^{l_j} X_{i,\theta(j)}^m = Z(\tilde{l}). \quad (38)
\]
Let \( r_{i,j}^m := X_{i,j}^m - X_{i,j}^m \) be the random variables taking the values \(-1, 0\) and \(+1\) with respective probabilities \( p_{\max}^m, 1 - p_{\max}^m - p_j^m \) and \( p_j^m \). Independently, let \( r_{i,j}^m \) be Bernoulli random variables with parameter \( q_j^m = (p_{2nd}^m - p_j^m)/(1 - p_{\max}^m - p_j^m) \in (0, 1) \) where \( p_{2nd}^m = \max(p_j^m : 1 \leq j \leq m) \) and define
\[
\beta_{i,j}^m = \begin{cases} -1 & \alpha_{i,j}^m = -1 \\ 0 & \alpha_{i,j}^m = 0 \text{ and } \epsilon_{i,j}^m = 0 \\ +1 & \alpha_{i,j}^m = +1 \text{ or } \alpha_{i,j}^m = 0 \text{ and } \epsilon_{i,j}^m = 1. \end{cases}
\]
Note that \( \mathbb{P}(\beta^m_{i,j} = +1) = p^m_{2nd} \) and that \( \alpha^m_{i,j} \leq \beta^m_{i,j} \), so that
\[
R(l) \leq \tilde{R}(l) = \sum_{j \in J_i} \sum_{i = l_j - 1 + 1}^{l_j} \beta^m_{i,j}.
\]

Since \( Z(l) \leq Z(\hat{l}) + \tilde{R}(l) \), we have
\[
\max_{l \in I(m)} Z(l) \leq \max_{l \in I(m)} Z(\hat{l}) + \max_{l \in I(m)} \tilde{R}(l) \leq \max_{l \in I(m)} Z(l) + \max_{l \in I(m)} \tilde{R}(l).
\]

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

The random variable \( \tilde{R}(l) \) is the sum of \( \sum_{j \in J_i} (l_j - l_j - 1) \) iid random variables, so that \( \max_{l \in I(m)} \tilde{R}(l) \) is distributed according to \( \left( \max_{1 \leq k \leq n} \sum_{i=1}^{k} \beta^m_{i} \right)^+ \) where \( \beta^m_{i} \) are iid with
\[
\mathbb{P}(\beta^m_{i} = -1) = p^m_{\text{max}}, \quad \mathbb{P}(\beta^m_{i} = 0) = 1 - p^m_{\text{max}} - p^m_{2nd}, \quad \mathbb{P}(\beta^m_{i} = +1) = p^m_{2nd}.
\]

We are now interested in bounding \( \mathbb{E} \left[ \left( \max_{1 \leq k \leq n} \sum_{i=1}^{k} \beta^m_{i} \right)^+ \right] \).

Let \((\epsilon^m_{i})_i\) be iid Bernoulli random variables with parameter \(p^m_{\text{max}} + p^m_{2nd}\) and let, independently, \((Y^m_{i})_i\) be iid Rademacher random variables with parameter \(p^m_{2nd}/(p^m_{2nd} + p^m_{\text{max}})\) (i.e. \(\mathbb{P}(Y^m_{i} = 1) = 1 - \mathbb{P}(Y^m_{i} = -1) = p^m_{2nd}/(p^m_{2nd} + p^m_{\text{max}})\)). Then \(\beta^m_{i}\) and \(\epsilon^m_{i}Y^m_{i}\) have the same distribution and we have
\[
\mathbb{E} \left[ \left( \max_{1 \leq k \leq n} \sum_{i=1}^{k} \epsilon^m_{i}Y^m_{i} \right)^+ \right] = \mathbb{E} \left[ \left( \max_{1 \leq k \leq n} \sum_{i=1}^{k} \epsilon^m_{i}Y^m_{i} \right)^+ \mid G_n \right] = \mathbb{E} \left[ \left( \max_{1 \leq k \leq n} \sum_{i=1}^{k} \epsilon^m_{i}Y^m_{i} \right)^+ \mid G_n \right]
\]
where \(G_n = \sigma(\epsilon^m_{i} : 1 \leq i \leq n)\). But since \((\epsilon^m_{i})_i\) is independent of \((Y^m_{i})_i\), we have
\[
\mathbb{E} \left[ \left( \max_{1 \leq k \leq n} \sum_{i=1}^{k} \epsilon^m_{i}Y^m_{i} \right)^+ \mid G_n \right] = \mathbb{E} \left[ \left( \max_{1 \leq k \leq \ell} \sum_{i=1}^{k} Y^m_{i} \right)^+ \mid G_n \right]
\]
where \(\ell = \sum_{i=1}^{n} \epsilon^m_{i}\) has a \(\mathcal{B}(n, p^m_{\text{max}} + p^m_{2nd})\) distribution. But
\[
\mathbb{E} \left[ \left( \max_{1 \leq k \leq \ell} \sum_{i=1}^{k} Y^m_{i} \right)^+ \mid G_n \right] = \sum_{k=0}^{+\infty} \left( \mathbb{P} \left( \left( \max_{1 \leq j \leq \ell} \sum_{i=1}^{j} Y^m_{i} \right)^+ \geq k \right) \right) = \sum_{k=0}^{+\infty} \left( 1 - \mathbb{P} \left( \max_{1 \leq j \leq \ell} \sum_{i=1}^{j} Y^m_{i} \leq k \right) \right) = \ell - U_{\ell}
\]
where \(U_{\ell} = \sum_{k=0}^{\ell-1} u_{\ell,k}\) and \(u_{\ell,k} = \mathbb{P}(\max_{1 \leq j \leq \ell} \sum_{i=1}^{j} Y^m_{i} \leq k)\). With the latest notation, we are now investigating \(\gamma_n = \mathbb{E} [\ell - U_{\ell}]\). For simplicity, in the sequel, we set \(p_{*, m} := p^m_{2nd}/(p^m_{2nd} + p^m_{\text{max}})\) and \(q_{*, m} := 1 - p_{*, m}\).
The elements of the sequence \((u_{\ell,k})_{1 \leq k \leq \ell-1}\) satisfy the following induction relations:

\[
u_{\ell,k} = q_{s,m}u_{\ell-1,k+1} + p_{s,m}u_{\ell-1,k-1}, \quad k \geq 1, \quad u_{\ell,0} = q_{s,m}u_{\ell-1,1},
\]

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

In order to compute \(E_{\tau_1}^{\mathcal{G}_n}\), we introduce the hitting time \(\tau_1^m = \min\{k \geq 1 : \sum_{i=1}^k Y_i^m = 1\}\) of the random walk \((\sum_{i \leq j} Y_i^m)\). We then have

\[
P(\tau_1^m \leq k) = P\left( \max_{i \leq k} \sum_{j=1}^i Y_j^m \geq 1 \right) = 1 - P\left( \max_{i \leq k} \sum_{j=1}^i Y_j^m \leq 0 \right) = 1 - u_{k,0},
\]

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

\[
U_\ell = 2\ell q_{s,m} - q_{s,m} \sum_{k=1}^{\ell} P(\tau_1^m \geq k)
\]

\[
= 2\ell q_{s,m} - q_{s,m} \sum_{i=1}^{+\infty} (i \wedge \ell)P(\tau_1^m = i)
\]

\[
= 2\ell q_{s,m} - q_{s,m}E[\tau_1^m \wedge \ell | \mathcal{G}_n].
\]

Next,

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

and we have

\[
\gamma_n := E[\ell - U_\ell]
\]

\[
= E[\ell(1 - 2q_{s,m}) + q_{s,m}E[\tau_1^m \wedge \ell | \mathcal{G}_n]]
\]

\[
= E[\ell(1 - 2q_{s,m})] + q_{s,m}E[\tau_1^m \wedge \ell]
\]

\[
= E[\ell(1 - 2q_{s,m})] + q_{s,m}(E[\ell : \tau_1^m \geq n] + E[\tau_1^m \wedge \ell : \tau_1^m < n])
\]

\[
= E[\ell(1 - 2q_{s,m})] + q_{s,m}(E[\ellP(\tau_1^m \geq n) + E[\tau_1^m \wedge \ell : \tau_1^m < n]).
\]

From [Fe, p. 352], the law of \(\tau_1^m\) is given for \(j \geq 0\) by \(P(\tau_1^m = 2j) = 0\) and

\[
P(\tau_1^m = 2j + 1) = \frac{1}{2j + 1} \binom{2j + 1}{j + 1} p_{s,m}^{j+1} q_{s,m}^j.
\]

Observe that the sum \(\sum_{j=0}^{+\infty} \frac{1}{2j + 1} \binom{2j + 1}{j + 1} p_{s,m}^{j+1} q_{s,m}^j\) converges uniformly with respect to \(n\) since \(q_{s,m} \leq 1\) and \(p_{s,m} \to 0\), when \(m = m(n) \to +\infty\) so that,

\[
\lim_{n \to +\infty} P(\tau_1^m(n) < n) = \lim_{n \to +\infty} \left( \sum_{j=0}^{n-1} \frac{1}{2j + 1} \binom{2j + 1}{j + 1} p_{s,m}^{j+1} q_{s,m}^j \right) m_{p_{m,s} q_{m,s}} = 0,
\]

and \(P(\tau_1^m \geq n) = 1\). Taking \(n \to +\infty\) in (41) and recalling the definition of \(p_{m,s}, q_{m,s}\), we obtain \(\gamma_n \sim n P_{2nd}\), which concludes the proof of Lemma 10.

\section{Concluding remarks}

A natural question to handle 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, $X^m_i$, $1 \leq i \leq n$, are independently and identically distributed with $P(X^m_i = \alpha_j) = p^m_{ij}$. In order to simplify the notations, we assume (without loss of generality) that the ordered letters $\alpha_1^m < \cdots < \alpha_n^m$ have moreover non-increasing probabilities (i.e. $p^m_1 \geq p^m_2 \geq \cdots \geq p^m_n$). Let $d^m_i = \text{card}\{j : p^m_{ij} = p^m_i\}$ be the multiplicity of $p^m_i$ and let $m^m_r = \max\{i : p^m_i > p^m_r\}$ be the number of letters (strictly) more probable than $p^m_r$. Let $J_r(m) = \{i : p^m_i = p^m_r\} = \{m_r + 1, \ldots, m_r + d^m_r\}$ be the indices of the letters with probability $p^m_r$. We recover our previous notation, $r = 1$, with $(k(m) = d^m_1$ and $J(m) = J_1(m)$. Since the expression (16) has a purely combinatorial nature, it still holds true that

$$V_r(n, m) = \max_{k \in J_r(m)} \left( \sum_{i=1}^{r} \sum_{l=0}^{d^m_i} \sum_{i=k_{j-1}+1}^{k_l} X^m_{i,l} \right).$$

Let $\nu^m_k = \sum_{i=1}^{k} p^m_i$. Note that from Theorem 5.2 in [HL2], when $m$ is fixed and $n \to +\infty$, we have for each $1 \leq r \leq m$:

$$\left( \frac{V_k(n, m) - \nu^m_k n}{\sqrt{n}} \right)_{1 \leq k \leq r} \Rightarrow (V^k_\infty)_{1 \leq k \leq r},$$

where the limit is given in Section 6 of [HL2] by $V^r_\infty = Z(m, r) + \sqrt{p^m_r D_r - m^m_{\infty}} d^m_r$ with $Z(m, r) \sim \mathcal{N}(0, \nu^m_r)$, for $\nu_r^m = \nu^m_{m^m_r}(1 - \nu^m_{m^m_r}) + (p^m_r (r - m^m_r))^2$ and,

$$D_{r,m} = \max_{t \in I_{r,m}} \left( \sum_{j=1}^{r} \sum_{l=0}^{d^m_j} (B^j(t,j,l) - B^j(t,j,l-1)) \right),$$

for

$$I_{r,m} = \{t = (t_j,l, 1 \leq j \leq r, 0 \leq l \leq m) : t_{j,0} = 0, t_{j,m_r+j} = 1, 1 \leq j \leq r, t_{j,l-1} \leq t_{j,l}, 1 \leq j \leq r, 1 \leq l \leq m - 1, t_{j,l} \leq t_{j-l+1,l}, 2 \leq j \leq r, 1 \leq l \leq m - 1\}.$$

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, [GW] and [Ba]). 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 rewrites as

$$(m^{1/6}(D_{k,m} - k \sqrt{m}))_{1 \leq k \leq r} \Rightarrow \mathbf{F}_r \Theta_r^{-1}, \quad m \to +\infty.$$ (43)

Arguing like in the previous sections, we would like to derive the fluctuations of $(V_k(n, m))_{1 \leq k \leq r}$ with respect to $n$ and $m$ simultaneously from (42) and (43). However, in the non-uniform case, this is not that transparent since, for each $r \geq 1$, the behavior of $m^m_\infty$ and of $d^m_\infty$, with respect to $m$, is not that clear cut. In particular, $r - m^m_\infty$ may not be stationary and (43) can no longer be used for $D_{r-m^m_\infty, d^m_\infty}$. Besides, the random fluctuations of $\sqrt{p^m_r D_{r-m^m_\infty, d^m_\infty}}$ in $V^r_\infty$ are of order $(p^m_r)^{1/2}(d^m_r)^{1/6}$ which, in general, does not dominate those of $Z(m, r) \sim \mathcal{N}(0, \nu_r^m)$. Thus, for general non-uniform alphabets, we cannot infer which part of the law of $V^r_\infty = Z(m, r) + \sqrt{p^m_r D_{r-m^m_\infty, d^m_\infty}}$ 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, either 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 performed. Both authors would like to Boris Bukh and Alperen Özdémit for their comments which led to the extended Remark 8 of the present version.
References

[BDJ] J. Baik, P. Deift, K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. J. Amer. Math. Soc., vol. 12, pp. 1119–1178, 1999.

[BS] J. Baik, T. Suidan. A GUE central limit theorem and universality of directed first and last passage percolation site. Int. Math. Res. Not. no. 6, pp. 325–337, 2005.

[Ba] Y. Baryshnikov. GUEs and Queues. Probab. Theor. and Relat. Fields, vol. 119, pp. 256–274, 2001.

[BM] T. Bodineau, J. Martin. A universality property for last-passage percolation paths close to the axis. Elect. Comm. Probab. vol. 10, pp. 105–112, 2005.

[BJ] P. Bougerol, T. Jeulin. Paths in Weyl chambers and random matrices. Probab. Theor. and Relat. Fields, vol. 124 no. 4, pp. 517–543, 2002.

[CG] Y. Baryshnikov. GUEs and Queues. Probab. Theor. and Relat. Fields, vol. 119, pp. 256–274, 2001.

[BM] T. Bodineau, J. Martin. A universality property for last-passage percolation paths close to the axis. Elect. Comm. Probab. vol. 10, pp. 105–112, 2005.

[BJ] P. Bougerol, T. Jeulin. Paths in Weyl chambers and random matrices. Probab. Theor. and Relat. Fields, vol. 124 no. 4, pp. 517–543, 2002.

[CG] Y. Baryshnikov. GUEs and Queues. Probab. Theor. and Relat. Fields, vol. 119, pp. 256–274, 2001.

[Do] W. Doumerc. A note on representations of eigenvalues of classical Gaussian matrices. Séminaire de Probabilité XXXVII, Lecture Notes in Math., no. 1832, Springer, Berlin pp. 370-384, 2003.

[Fo] W. Feller. An Introduction to Probability Theory and its Applications, 3rd Edition vol. 1, Wiley ed., 1967.

[Fu] W. Fulton. Young tableaux with applications to representation theory and geometry. Cambridge University Press, 1997.

[GW] P. W. Glynn, W. Whitt. Departure from many queues in series. Ann. App. Probab., vol. 1, pp. 546–572, 1991.

[GTW] J. Gravner, C. Tracy, H. Widom. Limit theorems for height fluctuations in a class of discrete space and time growth models. J. Stat. Phys., vol. 102 nos 5-6, pp. 1085–1132, 2001.

[HIL1] C. Houdré, T. Litherland. On the longest increasing subsequence for finite and countable alphabets. High dimensional probability V: the Luminy volume, pp. 185–212, Inst. Math. Stat. (IMS) Collect., 5, Inst. Math. Statist., Beachwood, OH, 2009.

[HIL2] C. Houdré, T. Litherland. On the limiting shape of Markovian random Young tableaux. Preprint arXiv:0810.2982, 2008

[HIL] C. Houdré, Ü. Işlak, T. Litherland. Longest increasing subsequence asymptotics with growing alphabets. Preprint. 2015.

[HX] C. Houdré, H. Xu. On the limiting shape of random Young tableaux associated to inhomogeneous words, arXiv:0901.4138, 2009.

[ITW1] A. Its, C. Tracy, H. Widom. Random words, Toeplitz determinants, and integrable systems I. Random matrix models and their applications, Math. Sci. Res. Inst. Publ., vol. 40, pp. 245–258, Cambridge Univ. Press, Cambridge, 2001.

[ITW2] A. Its, C. Tracy, H. Widom. Random words, Toeplitz determinants and integrable systems. II. Advances in nonlinear mathematics and science. Phys. D 152/153 pp. 199–224, 2001.

[Jo] K. Johansson. Discrete polynomials ensembles and the Plancherel measure. Ann. Math. vol. 153, pp. 259–296, 2001.

[Li] M. Lifshits. Lecture notes on strong approximation. Pub. IRMA Lille, vol. 53, no 13, 2000.

[Me] M. L. Mehta. Random matrices, 2nd ed. Academic Press, San Diego, 1991

[OCY] N. O’Connell, M. Yor. Brownian analogues of Burke’s theorem. Stoch. Proc. Appl. vol. 96, no. 2, pp. 285–304, 2001.

[St] R. Stanley. Enumerative combinatorics, vol. 2. Cambridge University Press, Cambridge, 1999.

[Su] T. Suidan. A remark on a theorem of Chatterjee and last passage percolation. J. Phys. A: Math. Gen. vol. 39, pp. 8977–8981, 2006.

[TW1] C. Tracy, H. Widom. Level-spacing distributions and the Airy kernel. Commun. Math. Phys., vol. 159, pp. 151–174, 1994.

[TW2] C. Tracy, H. Widom. Correlation functions, Cluster functions and spacing distributions for random matrices. J. Statist. Phys., vol. 92, no 5-6, pp. 809–835, 1998.

[TW3] C. Tracy, H. Widom. On the distribution of the length of the longest monotone subsequences in random word. Probab. Theory Relat. Fields, vol. 119, pp. 350–380, 2001.