The Annals of Applied Probability
2012, Vol. 22, No. 6, 2539–2559
DOI: 10.1214/12-AAP853
© Institute of Mathematical Statistics, 2012

ON THE RATE OF APPROXIMATION IN FINITE-ALPHABET LONGEST INCREASING SUBSEQUENCE PROBLEMS

BY CHRISTIAN HOUDRÉ[1] AND ZSOLT TALATA[2]

Georgia Institute of Technology and University of Kansas

The rate of convergence of the distribution of the length of the longest increasing subsequence, toward the maximal eigenvalue of certain matrix ensembles, is investigated. For finite-alphabet uniform and nonuniform i.i.d. sources, a rate of $\log n/\sqrt{n}$ is obtained. The uniform binary case is further explored, and an improved $1/\sqrt{n}$ rate obtained.

1. Introduction. In this paper, we consider the length of the longest increasing subsequence of a random word of size n for general i.i.d. sequences with alphabet of fixed size m. As $n \to \infty$, the limiting distribution of the normalized length has direct connections to random matrix theory. If the i.i.d. sequence is uniformly distributed, Tracy and Widom [17] proved that the limiting distribution is that of the largest eigenvalue of the $m \times m$ traceless Gaussian Unitary Ensemble (GUE); while for general i.i.d. sequences, Its, Tracy and Widom [9, 10] showed it to be the distribution of the largest eigenvalue of a direct sum of certain elements of GUEs.

Limiting distributions in similar problems have also been formulated as Brownian functionals [1, 5, 6, 8]. In particular, in [8], the length of the longest increasing subsequence is obtained as a random walk functional, and the limiting distribution, as a Brownian functional. This direct approach allows us to explore several questions of probabilistic and statistical nature in this problem, such as the investigation of the rate of convergence to the limiting distribution, which is done below.

To briefly describe the content of the paper, for general i.i.d. sequences we derive, in Section 4, an upper bound of order $\log n/\sqrt{n}$ on the rate of convergence, using strong approximation techniques. In the special case of uniform binary sequences ($m = 2$), the rate is sharpened in Section 5 to the order $1/\sqrt{n}$.

In previous and related studies, the rate of convergence of certain random walk functionals has been investigated. For example, in queuing theory, Glynn and Whitt [5] obtained a similar rate via the KMT technique. In that problem, although the functional of an m-dimensional random walk is similar, the random walks are mutually independent, which is not our case. Moreover, what is meant there by

Received February 2012.

[1]Supported in part by NSA Grant H98230-09-1-0017.

[2]Supported in part by NSF Grant DMS-09-06929.

MSC2010 subject classifications. 60C05, 60G15, 60G17, 62E17, 62E20.

Key words and phrases. Longest increasing subsequence, Brownian functional, approximation, rate of convergence.

rate of convergence is an almost sure upper bound on the deviation between the random walk and the Brownian functional. The order of that bound is given, but its constant factor may depend on the realization of the process. Here, the random walks are dependent, and by rate of convergence, we mean an upper bound on the deviation of the distribution functions.

Although the Skorokhod embedding of random walks usually provides a rate of $\mathcal{O}(n^{-1/4})$ [14], when the random walk is one dimensional and the functionals are the supremum or the local score, Etienne and Vallois [3] obtained a rate of $\mathcal{O}(\sqrt{\log n/n})$ using embedding techniques. It is not clear whether or not their results can be used or generalized to our problem.

To begin we introduce in the next section some notation and also summarize some of the interplay between the longest increasing subsequence problem and random matrix theory.

2. Longest increasing subsequences. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with values in the ordered alphabet $\mathcal{A} = \{\alpha_1, \ldots, \alpha_m\}$, where $\alpha_1 < \alpha_2 < \cdots < \alpha_m$. Let $p_r = \mathbb{P}(X_1 = \alpha_r)$, $r = 1, \ldots, m$, with $p_{\max} = \max_{1\le r\le m} p_r$ and let also k be the multiplicity of $p_{\max}$ among the probabilities p_r $(1 \le r \le m)$.

$$k = \#\{r : 1 \le r \le m, p_r = p_{\max}\}. \tag{2.1}$$

Finally, let LI_n be the length of the longest increasing subsequence of $X_1, \ldots, X_n$, that is,

$$LI_n = \max\{j : X_{i_1} \le X_{i_2} \le \cdots \le X_{i_j}, \text{ for some } 1 \le i_1 < i_2 < \cdots < i_j \le n\}.$$

Properly renormalized, LI_n is known to converge to the maximal eigenvalue of some matrix ensemble (see [9–11, 17]). In fact, in the notation of [8],

$$\frac{LI_n - np_{\max}}{\sqrt{np_{\max}}} \Rightarrow J_k, \tag{2.2}$$

where

$$\sqrt{p_{\max}} J_k = -\frac{1}{m}\sum_{r=1}^{m-1} r\sigma_r \tilde{B}^r(1) + \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{m-1}\le t_m=1\\ t_r=t_{r-1}, r\in I^*}} \sum_{r=1}^{m-1} \sigma_r \tilde{B}^r(t_r), \tag{2.3}$$

with $\sigma_r^2 = p_r + p_{r+1} - (p_r - p_{r+1})^2$, $r = 1, 2, \ldots, m-1$ and $I^* = \{r : p_r < p_{\max}, 1 \le r \le m\}$. Above, $(\tilde{B}^1(t), \ldots, \tilde{B}^{m-1}(t))^\top$ is an $(m-1)$-dimensional driftless Brownian motion with covariance matrix

$$t\begin{pmatrix} 1 & \rho_{1,2} & \rho_{1,3} & \cdots & \rho_{1,m-1} \\ \rho_{2,1} & 1 & \rho_{2,3} & \cdots & \rho_{2,m-1} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & 1 & \rho_{m-2,m-1} \\ \rho_{m-1,1} & \rho_{m-1,2} & \cdots & \rho_{m-1,m-2} & 1 \end{pmatrix},$$

where

$$\rho_{r,s} = \begin{cases} -\dfrac{p_r + \mu_r \mu_s}{\sigma_r \sigma_s}, & \text{if } s = r-1, \\ -\dfrac{p_s + \mu_r \mu_s}{\sigma_r \sigma_s}, & \text{if } s = r+1, \\ -\dfrac{\mu_r \mu_s}{\sigma_r \sigma_s}, & \text{if } |r-s| > 1, 1 \le r, s \le m-1, \end{cases}$$

and $\mu_r = p_r - p_{r+1}$, $1 \le r \le m-1$.

Next, let

$$H_m = \sqrt{2}\left\{-\frac{1}{m}\sum_{r=1}^{m-1} r\bar{B}^r(1) + \max_{\substack{0 \le t_1 \le \cdots \\ \le t_{m-1} \le 1}} \sum_{r=1}^{m-1} \bar{B}^r(t_r)\right\}, \tag{2.4}$$

where $(\bar{B}^1(t), \dots, \bar{B}^{m-1}(t))^\top$ is an $(m-1)$-dimensional driftless Brownian motion with covariance matrix

$$t\begin{pmatrix} 1 & -1/2 & & & \bigcirc \\ -1/2 & 1 & -1/2 & & \\ & \ddots & \ddots & \ddots & \\ \bigcirc & & -1/2 & 1 & -1/2 \\ & & & -1/2 & 1 \end{pmatrix}.$$

Comparing (2.4) and (2.3), it is immediate that if the distribution on the alphabet $\mathcal{A}$ is uniform, that is, $p_r = 1/m$, $r = 1, \dots, m$, then $k = m$, $\mu_r = 0$, $\sigma_r^2 = 2/m$, and thus $J_m = H_m$, and therefore

$$\frac{LI_n - n/m}{\sqrt{n/m}} \Rightarrow H_m.$$

Actually, similar results hold true for countable alphabets (see [8]) and our methodology also gives the rate in that case.

Let us now briefly recall the connections, originating in [1] and [6], between random matrix theory and the Brownian functionals encountered in the present paper.

An $m \times m$ element of the Gaussian Unitary Ensemble (GUE) is an $m \times m$ Hermitian random matrix $\{Y_{i,j}\}_{1 \le i,j \le m}$ with $Y_{i,i} \sim N(0,1)$ for $1 \le i \le m$, $\mathrm{Re}(Y_{i,j}) \sim N(0, 1/2)$ and $\mathrm{Im}(Y_{i,j}) \sim N(0, 1/2)$ for $1 \le i < j \le m$, and $Y_{i,i}$, $\mathrm{Re}(Y_{i,j})$, $\mathrm{Im}(Y_{i,j})$ are mutually independent for $1 \le i \le j \le m$.

Writing $x^{(m)} = (x_1, x_2, \dots, x_m) \in \mathbb{R}^m$ for any $m \ge 1$, letting $\Delta(x^{(m)}) = \Pi_{1 \le i < j \le m}(x_i - x_j)$ be the Vandermonde determinant, the following facts hold true.

First, from [17] and [8], $\lambda_1^{(m,0)} \stackrel{\mathcal{L}}{=} H_m$, where $\lambda_1^{(m,0)}$ is the largest eigenvalue of the $m \times m$ *traceless* GUE. Using the joint density of the eigenvalues of the traceless $m \times m$ GUE [15, 17], the distribution function of H_m can be computed directly, for all $m \geq 2$ and all $s \geq 0$, as

$$\mathbb{P}(H_m \leq s) = c_m^0 \int_{\{\max x_j \leq s\}} e^{-(1/2)\sum_{i=1}^m x_i^2} \Delta\big(x^{(m)}\big)^2 \lambda_m\big(dx^{(m)}\big), \tag{2.5}$$

where λ_m is the Lebesgue measure concentrated on the hyperplane $\mathcal{L}_m = \{x \in \mathbb{R}^m : \sum_{i=1}^m x_i = 0\}$, and where

$$(c_m^0)^{-1} = \int_{\mathbb{R}^m} e^{-(1/2)\sum_{i=1}^m x_i^2} \Delta\big(x^{(m)}\big)^2 \lambda_m\big(dx^{(m)}\big) = (2\pi)^{(m-1)/2} \prod_{i=0}^{m-1} i!.$$

Note that H_m is a.s. nonnegative, and so $\mathbb{P}(H_m \leq s) = 0$ for all $s < 0$.

Second, for all $k \geq 2$, J_k can be represented [7, 8] as

$$J_k = H_k + \sqrt{\frac{1 - kp_{\max}}{k}} Z, \tag{2.6}$$

where Z is a standard normal random variable and, moreover, H_k and Z are independent, while, $J_1 = \sqrt{1 - p_{\max}} Z$.

The distribution of J_k can be described [9, 10] as the largest eigenvalue of the direct sum of d mutually independent GUEs, each of size $k_j \times k_j$, $1 \leq j \leq d$, subject to the eigenvalue constraint $\sum_{i=1}^m \sqrt{p_i}\lambda_i = 0$. The k_j are the multiplicities of the probabilities having common values, the p_i are ordered in nonincreasing order and the eigenvalues are ordered in terms of the GUEs corresponding to the appropriate values of p_i.

As shown in [10], for any $k \geq 1$ and all $s \in \mathbb{R}$, J_k has distribution given by

$$\begin{aligned}\mathbb{P}(J_k \leq s)& \\ &= c_{k,p_{\max}} \int_{\{\max x_j \leq s\}} e^{-(1/2)[\sum_{i=1}^k x_i^2 + (p_{\max}/(1-kp_{\max}))(\sum_{i=1}^k x_i)^2]} \Delta\big(x^{(k)}\big)^2 \, dx^{(k)},\end{aligned}$$

where

$$c_{k,p_{\max}}^{-1} = \int_{\mathbb{R}^k} e^{-(1/2)[\sum_{i=1}^k x_i^2 + (p_{\max}/(1-kp_{\max}))(\sum_{i=1}^k x_i)^2]} \Delta\big(x^{(k)}\big)^2 \, dx^{(k)}.$$

Below, we study the rate of approximation in (2.2) and prove (see Section 4) that

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}\Big(\frac{LI_n - np_{\max}}{\sqrt{np_{\max}}} \geq x\Big) - \mathbb{P}(J_k \geq x) \right| \leq C(m,k) \frac{\log n}{\sqrt{n}},$$

where the constant $C(m,k)$ depends only on m and k.

3. Upper bounds on the density functions.

Our first results provide upper bounds on the density of the functionals J_k and H_k.

PROPOSITION 3.1. (i) *Let f_{H_k} be the probability density function of H_k. Then, for any $k = 2, 3, \ldots, m$,*

$$\sup_{x \in \mathbb{R}} f_{H_k}(x) \leq k^{3k} (2\pi e^2)^{k/2} \sqrt{\frac{e}{\pi}}.$$

(ii) *Let* f_{J_k} *be the probability density function of* J_k. *Then, for any* $k = 2, 3, \dots, m-1$,

$$\sup_{x\in\mathbb{R}} f_{J_k}(x) \le \min\left\{\sqrt{\frac{k}{2\pi(1-kp_{\max})}}, k^{3k}(2\pi e^2)^{k/2}\sqrt{\frac{e}{\pi}}\right\}$$

and for $k = 1$, $\sup_{x\in\mathbb{R}} f_{J_1}(x) = 1/\sqrt{2\pi(1-p_{\max})}$.

REMARK 3.2. The distribution function of the largest eigenvalue of the $k \times k$ GUE can be computed directly [15], and so does the one of the $k \times k$ traceless GUE in (2.5). Its derivative, the density function, is upper bounded by k times the density function of the one-dimensional marginal of the distribution of the eigenvalues of the $k \times k$ GUE. Both the joint density of the eigenvalues of the $k \times k$ GUE and its marginals have a determinantal representation using Hermite polynomials [15], Section 6.2, which provides an upper bound of order $k^{5/6}$ on the density of the largest eigenvalue of the $k \times k$ GUE. In turn, this gives the order of the constant with the rate of convergence $\log n/\sqrt{n}$, in the framework of [5].

For the joint and marginal density of the eigenvalues of the $k \times k$ traceless GUE, a determinantal representation does not seem to be available. One can still conjecture a polynomial upper bound on the supremum of the density of the largest eigenvalue of the $k \times k$ traceless GUE, but the authors' efforts did not lead to such a bound in part (i) of Proposition 3.1. Indeed, the traceless condition induces dependencies between the entries of the Gaussian random matrix making the analysis more delicate than in the GUE case.

PROOF. First for (ii) using (2.6), for $k > 1$,

$$\begin{aligned} f_{J_k}(x) &= \int_{\mathbb{R}} f_{H_k}(u) f_{\sqrt{(1-kp_{\max})/k}Z}(x-u)\,du \\ &\le \sup_{u\in\mathbb{R}} f_{H_k}(u), \end{aligned}$$

and use (i). Similarly, for $k < m$,

$$\sup_{x\in\mathbb{R}} f_{J_k}(x) \le \sup_{u\in\mathbb{R}} f_{\sqrt{(1-kp_{\max})/k}Z}(u) = \sqrt{\frac{k}{2\pi(1-kp_{\max})}}.$$

Now for (i), to upper bound the density of H_k, consider its cumulative distribution function. By (2.5),

$$\begin{aligned} \mathbb{P}(H_k \le s) = {}& \frac{1}{(2\pi)^{(k-1)/2}} \\ & \times \int_{\{\max x_j \le s\}} \exp\Biggl(-\frac{1}{2}\sum_{i=1}^k x_i^2 + 2\sum_{1\le i<j\le k}\log|x_i - x_j| \\ & \qquad - \sum_{i=1}^{k-1}\log i!\Biggr)\lambda_k(dx), \end{aligned}$$

and so

$$
\begin{aligned}
&\mathbb{P}(s \le H_k \le s+\varepsilon) \\
&\quad = \int_{\{s \le \max x_j \le s+\varepsilon\}} \frac{e^{-(1/2)(1-2/\nu^2)\sum_{i=1}^k x_i^2}}{(2\pi)^{(k-1)/2}} \\
&\qquad \times \exp\Biggl(-\frac{1}{\nu^2}\sum_{i=1}^k x_i^2 + 2\sum_{1\le i<j\le k}\log|x_i-x_j| - \sum_{i=1}^{k-1}\log i!\Biggr)\lambda_k(dx)
\end{aligned} \tag{3.1}
$$

for any $\nu > \sqrt{2}$. In order to dominate the first term of the integrand in (3.1), a bound (see [15], Appendix A.6) going back to Stieltjes, asserts that

$$
\begin{aligned}
&\frac{1}{2}\sum_{i=1}^k x_i^2 - \sum_{1\le i<j\le k}\log|x_i-x_j| \\
&\qquad \ge \frac{1}{4}k(k-1)(1+\log 2) - \frac{1}{2}\sum_{i=1}^k i\log i.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\frac{1}{\nu^2}\sum_{i=1}^k x_i^2 - 2\sum_{1\le i<j\le k}\log|x_i-x_j| \\
&\quad = 2\Biggl(\frac{1}{2\nu^2}\sum_{i=1}^k x_i^2 - \sum_{1\le i<j\le k}\log|x_i-x_j|\Biggr) \\
&\quad = 2\Biggl(\frac{1}{2}\sum_{i=1}^k\biggl(\frac{x_i}{\nu}\biggr)^2 - \sum_{1\le i<j\le k}\log\biggl|\frac{x_i}{\nu}-\frac{x_j}{\nu}\biggr|\Biggr) - \frac{k(k-1)}{2}2\log\nu \qquad (3.2) \\
&\quad \ge 2\Biggl(\frac{1}{4}k(k-1)(1+\log 2) - \frac{1}{2}\sum_{i=1}^k i\log i\Biggr) - k(k-1)\log\nu \\
&\quad = \frac{1}{2}k(k-1)(1+\log 2) - \sum_{i=1}^k i\log i - k(k-1)\log\nu.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
\sum_{i=1}^{k-1}\log i! &= \sum_{i=1}^{k-1}\sum_{j=1}^{i}\log j = \sum_{j=1}^{k-1}(k-j)\log j \\
&= k\log(k-1)! - \sum_{i=1}^{k-1} i\log i.
\end{aligned} \tag{3.3}
$$

Combining (3.2) and (3.3) leads to

$$\exp\Biggl(-\frac{1}{\nu^2}\sum_{i=1}^{k} x_i^2 + 2\sum_{1\le i<j\le k}\log|x_i-x_j| - \sum_{i=1}^{k-1}\log i!\Biggr)$$

$$\tag{3.4} \le \exp\Biggl(-\frac{1}{2}k(k-1)(1+\log 2) + k(k-1)\log\nu - k\log(k-1)! + 2\sum_{i=1}^{k-1} i\log i + k\log k\Biggr).$$

Here, using Stirling's inequality (see Feller [4], page 54)

$$(k-1)! > \sqrt{2\pi(k-1)}(k-1)^{(k-1)}e^{-(k-1)+1/(12(k-1)+1)},$$

and

$$\sum_{i=1}^{k-1} i\log i \le \int_1^k i\log i\,di = \frac{k^2}{2}\log k - \frac{k^2}{4} + \frac{1}{4},$$

the exponent in (3.4) can be upper bounded by

$$\begin{aligned}
&-\frac{1}{2}k(k-1)(1+\log 2) + k(k-1)\log\nu - k(k-1)\log(k-1)\\
&\quad - k\Biggl(-(k-1) + \frac{1}{12(k-1)+1} + \log\sqrt{2\pi(k-1)}\Biggr) + k\log k\\
&\quad + k^2\log k - \frac{k^2}{2} + \frac{1}{2}\\
&\quad = k^2\Biggl(-\frac{1}{2}(1+\log 2) + \log\nu - \log(k-1) + 1 + \log k - \frac{1}{2}\Biggr)\\
&\qquad + k\Biggl(\frac{1}{2}(1+\log 2) - \log\nu + \log(k-1)\\
&\qquad\qquad - 1 - \frac{1}{12(k-1)+1} + \log\sqrt{2\pi(k-1)} + \log k\Biggr) + \frac{1}{2}\\
&\quad \le k^2\Biggl(\log\frac{\nu}{\sqrt{2}} + \log\frac{k}{k-1}\Biggr) + \frac{5}{2}k\log k + k\log\sqrt{\frac{4\pi}{\nu^2 e}} + \frac{1}{2}.
\end{aligned}$$

And the bound (3.4) becomes

$$\tag{3.5} \begin{aligned}
&\exp\Biggl(-\frac{1}{4}\sum_{i=1}^{k} x_i^2 + 2\sum_{1\le i<j\le k}\log|x_i-x_j| - \sum_{i=1}^{k-1}\log i!\Biggr)\\
&\quad \le \biggl(\frac{\nu}{\sqrt{2}}\biggr)^{k^2}\biggl(\frac{k}{k-1}\biggr)^{k^2} k^{5k/2}\biggl(\frac{4\pi}{\nu^2 e}\biggr)^{k/2}\sqrt{e} \le \biggl(\frac{\nu}{\sqrt{2}}\biggr)^{k^2} k^{5k/2}\biggl(\frac{4\pi e}{\nu^2}\biggr)^{k/2}\sqrt{e}.
\end{aligned}$$

Thus (3.1) becomes

$$\mathbb{P}(s \le H_k \le s+\varepsilon) \le \left(\frac{\nu}{\sqrt{2}}\right)^{k^2} k^{5k/2}\left(\frac{4\pi e}{\nu^2}\right)^{k/2} \times \sqrt{e}\int_{\{s\le \max x_j \le s+\varepsilon\}} \frac{e^{-(1/2)(1-2/\nu^2)\sum_{i=1}^k x_i^2}}{(2\pi)^{(k-1)/2}}\lambda_k(dx). \tag{3.6}$$

On the hyperplane $\mathcal{L}_k \subset \mathbb{R}^k$, the function

$$\frac{1}{(2\pi)^{(k-1)/2}[(1-2/\nu^2)^{-1}]^{(k-1)/2}} e^{-(1/2)(1-2/\nu^2)\sum_{i=1}^k x_i^2}$$

is the probability density function of the $(k-1)$-dimensional normal distribution with mean $(0,\dots,0) \in \mathbb{R}^{k-1}$ and covariance $[(1-2/\nu^2)^{-1}]I_{k-1}$, where I_{k-1} is the $(k-1)$-dimensional identity matrix. Therefore,

$$\begin{aligned}
&\int_{\{s\le \max x_j \le s+\varepsilon\}} \frac{e^{-(1/2)(1-2/\nu^2)\sum_{i=1}^k x_i^2}}{(2\pi)^{(k-1)/2}}\lambda_k(dx)\\
&\quad= \left(1-\frac{2}{\nu^2}\right)^{-(k-1)/2} \times \int_{\{s\le \max x_j \le s+\varepsilon\}} \frac{e^{-(1/2)(1-2/\nu^2)\sum_{i=1}^k x_i^2}}{(2\pi)^{(k-1)/2}[(1-2/\nu^2)^{-1}]^{(k-1)/2}}\lambda_k(dx)\\
&\quad\le \left(1-\frac{2}{\nu^2}\right)^{-(k-1)/2} \times \sum_{j=1}^k \int_{\{s\le x_j \le s+\varepsilon\}} \frac{e^{-(1/2)(1-2/\nu^2)\sum_{i=1}^k x_i^2}}{(2\pi)^{(k-1)/2}[(1-2/\nu^2)^{-1}]^{(k-1)/2}}\lambda_k(dx)\\
&\quad\le \left(1-\frac{2}{\nu^2}\right)^{-(k-1)/2} \sum_{j=1}^k \sqrt{2}\varepsilon \sup_{\{s\le x_j\le s+\varepsilon\}} \frac{e^{-(1/2)(1-2/\nu^2)x_j^2}}{\sqrt{2\pi(1-2/\nu^2)^{-1}}}\\
&\quad\le \frac{\varepsilon}{\sqrt{\pi}} k\left(1-\frac{2}{\nu^2}\right)^{-k/2+1}.
\end{aligned} \tag{3.7}$$

Using (3.7), (3.6) yields

$$\mathbb{P}(s \le H_k \le s+\varepsilon) \le \left(\frac{\nu^2}{2}\right)^{k^2/2}\left(1-\frac{2}{\nu^2}\right)^{-k/2+1} \varepsilon k^{5k/2+1}\left(\frac{4\pi e}{\nu^2}\right)^{k/2}\sqrt{\frac{e}{\pi}}.$$

Choosing $\nu > \sqrt{2}$ such that $1-2/\nu^2 = 1/k$ provides

$$\left(\frac{\nu^2}{2}\right)^{k^2/2} = \left(1-\frac{1}{k}\right)^{-k^2/2} \le e^{k/2},$$

leading to

$$\mathbb{P}(s \le H_k \le s+\varepsilon) \le \varepsilon k^{3k}(2\pi e^2)^{k/2}\sqrt{\frac{e}{\pi}},$$

and the proof is complete. □

4. Rate of convergence results. Below, we study the rate of convergence in (2.2) and show that:

THEOREM 4.1. *For any $n \in \mathbb{N}$, $m \in \mathbb{N}$, for $k = 2, 3, \dots, m$,*

$$\begin{aligned}\sup_{x\in\mathbb{R}}&\left|\mathbb{P}\left(\frac{LI_n - np_{\max}}{\sqrt{np_{\max}}} \ge x\right) - \mathbb{P}(J_k \ge x)\right| \\ &\le c(m-1)\left((m-1)^2\sigma_{\max}^2 + \left(\frac{k^{3k}}{\sqrt{p_{\max}}} \wedge \sqrt{\frac{k}{p_{\max}(1-kp_{\max})}}\right)\right)\frac{\log n}{\sqrt{n}},\end{aligned} \tag{4.1}$$

where $c > 0$ is an absolute constant and $\sigma_{\max} = \max_{1\le r\le m-1}\sigma_r$. For $k = 1$, (4.1) holds with the minimum replaced by $(p_{\max}(1-p_{\max}))^{-1/2}$.

REMARK 4.2. For $k = m$, the minimum in (4.14) is $m^{3m+1/2}$ since $\sqrt{k/(p_{\max}(1-kp_{\max}))}$ is then understood to be infinite. In general, the minimum is $k^{3k}/\sqrt{p_{\max}}$, if $p_{\max} \ge (k^{6k}-k)/(k^{6k+1})$ and $\sqrt{k/(p_{\max}(1-kp_{\max}))}$, if $p_{\max} \le (k^{6k}-k)/(k^{6k+1})$.

In particular, (4.1) implies the following result which should be contrasted with Theorem 4 and Theorem 6 of [2].

COROLLARY 4.3. *If k is fixed and $m \to \infty$ as $n \to \infty$ in such a way that $m = o(n^{1/4}\log^{-1/2} n)$, then*

$$\frac{LI_n - np_{\max}}{\sqrt{np_{\max}}} \Rightarrow J_k.$$

PROOF. This immediately follows from Theorem 4.1 since $\sigma_{\max}^2 \le 2/m$ and $1/m \le p_{\max} \le 1/k$. □

PROOF OF THEOREM 4.1. Set $L_n = (LI_n - np_{\max})/\sqrt{n}$, and for $i = 1, \dots, n$ and $r = 1, \dots, m-1$, set also

$$Z_i^r = \begin{cases} 1, & \text{if } X_i = \alpha_r, \\ -1, & \text{if } X_i = \alpha_{r+1}, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, $\operatorname{Var} Z_i^r = \sigma_r^2$ and $\mathbb{E} Z_i^r = \mu_r$. Set $\tilde{S}_0^r = 0$ and

$$\tilde{S}_j^r = \sum_{i=1}^{j} \frac{Z_i^r - \mu_r}{\sigma_r}, \qquad j = 1, \dots, n,$$

and then L_n can be written (see the proof of Theorem 3.1 in [8]), as

$$L_n = -\frac{1}{m} \sum_{r=1}^{m-1} r\sigma_r \frac{\tilde{S}_n^r}{\sqrt{n}} + \max_{\substack{0=j_0 \le j_1 \le \cdots \\ \le j_{m-1} \le j_m = n \\ j_r = j_{r-1}, r \in I^*}} \sum_{r=1}^{m-1} \sigma_r \frac{\tilde{S}_{j_r}^r}{\sqrt{n}} + E_n, \tag{4.2}$$

where for the remainder term E_n we have for any $\varepsilon > 0$,

$$\mathbb{P}(|E_n| \ge \varepsilon) < \varepsilon\bigl(1 + (m-1)^2 \sigma_{\max}^2\bigr). \tag{4.3}$$

Letting

$$\widetilde{H}_{n,k} = -\frac{1}{m} \sum_{r=1}^{m-1} r\sigma_r \tilde{B}^r(1) + \max_{\substack{0=j_0 \le j_1 \le \cdots \\ \le j_{m-1} \le j_m = n \\ j_r = j_{r-1}, r \in I^*}} \sum_{r=1}^{m-1} \sigma_r \tilde{B}^r\biggl(\frac{j_r}{n}\biggr), \tag{4.4}$$

we have for any $\varepsilon > 0$

$$\begin{aligned} &\mathbb{P}\bigl(|L_n/\sqrt{p_{\max}} - J_k| > 2\varepsilon\bigr) \\ &\quad \le \mathbb{P}\bigl(|L_n - \widetilde{H}_{n,k}| > \varepsilon\sqrt{p_{\max}}\bigr) + \mathbb{P}\bigl(|\widetilde{H}_{n,k} - \sqrt{p_{\max}} J_k| > \varepsilon\sqrt{p_{\max}}\bigr). \end{aligned} \tag{4.5}$$

Using Lemmas 4.4 and 4.5 below, (4.5) can be upper bounded by

$$\begin{aligned} &\frac{1 + (m-1)^2\sigma_{\max}^2}{2} \varepsilon\sqrt{p_{\max}} + \exp\biggl(-\frac{\xi\varepsilon\sqrt{np_{\max}}}{16(m-1)}\biggr) \sum_{r=1}^{m-1} \biggl(1 + \frac{\sigma_r\sqrt{n}}{1 - |\mu_r|}\biggr) \\ &\qquad + 4(m-1)n \exp\biggl(\frac{-\varepsilon^2 n p_{\max}}{8\sigma_{\max}^2 (m-1)^2}\biggr). \end{aligned} \tag{4.6}$$

Now, from Proposition 3.1,

$$\begin{aligned} &\bigl|\mathbb{P}\bigl(L_n/\sqrt{p_{\max}} \ge x\bigr) - \mathbb{P}(J_k \ge x)\bigr| \\ &\qquad \le \mathbb{P}\bigl(\bigl|L_n/\sqrt{p_{\max}} - J_k\bigr| \ge 2\varepsilon\bigr) + \mathbb{P}(x - 2\varepsilon \le J_k \le x + 2\varepsilon) \\ &\qquad \le \frac{1 + (m-1)^2\sigma_{\max}^2}{2} \varepsilon\sqrt{p_{\max}} + \exp\biggl(-\frac{\xi\varepsilon\sqrt{np_{\max}}}{16(m-1)}\biggr) \sum_{r=1}^{m-1} \biggl(1 + \frac{\sigma_r\sqrt{n}}{1 - |\mu_r|}\biggr) \\ &\qquad\quad + 4(m-1)n \exp\biggl(\frac{-\varepsilon^2 n p_{\max}}{8\sigma_{\max}^2 (m-1)^2}\biggr) \\ &\qquad\quad + 4\varepsilon \min\biggl\{\sqrt{\frac{k}{2\pi(1 - kp_{\max})}}, k^{3k}(2\pi e^2)^{k/2}\sqrt{\frac{e}{\pi}}\biggr\}. \end{aligned} \tag{4.7}$$

With

$$\varepsilon = \frac{16(m-1)}{\xi\sqrt{p_{\max}}}\frac{\log n}{\sqrt{n}},$$

the right-hand side of (4.7) becomes

$$\frac{\log n}{\sqrt{n}}\Bigg(\frac{8(m-1)(1+(m-1)^2\sigma_{\max}^2)}{\xi} + \frac{1}{\log n}\sum_{r=1}^{m-1}\bigg(\frac{1}{\sqrt{n}} + \frac{\sigma_r}{1-|\mu_r|}\bigg)$$
$$+ 4(m-1)n^{-(32/\sigma_{\max}^2\xi^2)\log n - \log\log n/\log n + 3/2}$$
$$+ \frac{64(m-1)}{\xi\sqrt{p_{\max}}}\min\Bigg\{\sqrt{\frac{k}{2\pi(1-kp_{\max})}}, k^{3k}(2\pi e^2)^{k/2}\sqrt{\frac{e}{\pi}}\Bigg\}\Bigg),$$

which yields the claim of the theorem. □

LEMMA 4.4. *For any* $\varepsilon > 0$,

$$\mathbb{P}(|L_n - \widetilde{H}_{n,k}| > \varepsilon)$$
$$\le (1+(m-1)^2\sigma_{\max}^2)\frac{\varepsilon}{2} + \exp\bigg(-\frac{\xi\varepsilon\sqrt{n}}{16(m-1)}\bigg)\sum_{r=1}^{m-1}\bigg(1 + \frac{\sigma_r\sqrt{n}}{1-|\mu_r|}\bigg),$$

where $\xi > 0$ *is an absolute constant.*

PROOF. Comparing (4.2) to (4.4),

$$|L_n - \widetilde{H}_{n,k}|$$
$$\le |E_n| + \sum_{r=1}^{m-1}\bigg(1+\frac{r}{m}\bigg)\sigma_r \max_{0\le j\le n}\bigg|\frac{\tilde{S}_j^r}{\sqrt{n}} - \tilde{B}^r\bigg(\frac{j}{n}\bigg)\bigg| \tag{4.8}$$
$$\le |E_n| + 2\sum_{r=1}^{m-1}\sigma_r \max_{0\le j\le n}\bigg|\frac{\tilde{S}_j^r}{\sqrt{n}} - \tilde{B}^r\bigg(\frac{j}{n}\bigg)\bigg|.$$

For any $\delta > 0$ and $0 \le r \le m-1$,

$$\mathbb{P}\bigg(\max_{0\le j\le n}\bigg|\frac{\tilde{S}_j^r}{\sqrt{n}} - \tilde{B}^r\bigg(\frac{j}{n}\bigg)\bigg| > \delta\bigg) = \mathbb{P}\Big(\max_{0\le j\le n}|\tilde{S}_j^r - \tilde{B}^r(j)| > \delta\sqrt{n}\Big). \tag{4.9}$$

Applying Sakhanenko's version of the KMT inequality [12–14, 16] to the partial sums $\tilde{S}_j^r$, $j = 0,\dots,n$, of the i.i.d. random variables $(Z_i^r - \mu_r)/\sigma_r$, $i = 1,\dots,n$, (4.9) can be upper bounded by

$$(1 + C_2\sqrt{n})\exp(-C_1\delta\sqrt{n}). \tag{4.10}$$

Above, $C_1 = \xi\lambda_r$ and $C_2 = \lambda_r$, where ξ is an absolute constant and

$$\lambda_r = \sup\left\{\lambda : \lambda\mathbb{E}\left(\left|\frac{Z_i^r - \mu_r}{\sigma_r}\right|^3 \exp\left\{\lambda\left|\frac{Z_i^r - \mu_r}{\sigma_r}\right|\right\}\right) \le \mathbb{E}\left(\frac{Z_i^r - \mu_r}{\sigma_r}\right)^2\right\}$$
$$= \sigma_r \sup\{\lambda : \lambda\mathbb{E}(|Z_i^r - \mu_r|^3 \exp\{\lambda|Z_i^r - \mu_r|\}) \le \operatorname{Var} Z_i^r\}.$$

Since $|Z_i^r - \mu_r| \le 2$, choosing $\lambda = 1/4$ gives

$$\begin{aligned}
&\frac{1}{4}\mathbb{E}\left(|Z_i^r - \mu_r|^3 \exp\left\{\frac{1}{4}|Z_i^r - \mu_r|\right\}\right)\\
&\qquad = \mathbb{E}\left((Z_i^r - \mu_r)^2 \frac{|Z_i^r - \mu_r|}{4} \exp\left\{\frac{|Z_i^r - \mu_r|}{4}\right\}\right)\\
&\qquad \le \mathbb{E}\left((Z_i^r - \mu_r)^2 \frac{1}{2} \exp\left\{\frac{1}{2}\right\}\right)\\
&\qquad \le \operatorname{Var} Z_i^r,
\end{aligned}$$

which implies that $\lambda_r \ge \sigma_r/4$. Next, for any $\lambda > 1/\min\{1 - \mu_r, 1 + \mu_r\}$,

$$\begin{aligned}
\operatorname{Var} Z_i^r &= p_r(1 - \mu_r)^2 + p_{r+1}(1 + \mu_r)^2\\
&< \lambda \min\{|1 - \mu_r|, |1 + \mu_r|\}\big(p_r(1 - \mu_r)^2 + p_{r+1}(1 + \mu_r)^2\big)\\
&\le \lambda(p_r|1 - \mu_r|^3 + p_{r+1}|1 + \mu_r|^3)\\
&\le \lambda\mathbb{E}(|Z_i^r - \mu_r|^3)\\
&\le \lambda\mathbb{E}(|Z_i^r - \mu_r|^3 \exp\{\lambda|Z_i^r - \mu_r|\}),
\end{aligned}$$

which implies that $\lambda_r \le \sigma_r/\min\{1 - \mu_r, 1 + \mu_r\} = \sigma_r/(1 - |\mu_r|)$. Thus, the upper bound (4.10) becomes

$$\left(1 + \frac{\sigma_r\sqrt{n}}{1 - |\mu_r|}\right) \exp\left(-\frac{\xi\sigma_r}{4}\delta\sqrt{n}\right). \tag{4.11}$$

Combining (4.11) and (4.3) with (4.8),

$$\begin{aligned}
&\mathbb{P}(|L_n - \widetilde{H}_{n,k}| > \varepsilon)\\
&\quad \le \mathbb{P}\left(|E_n| + 2\sum_{r=1}^{m-1} \sigma_r \max_{0 \le j \le n}\left|\frac{\tilde{S}_j^r}{\sqrt{n}} - \tilde{B}^r\left(\frac{j}{n}\right)\right| > \varepsilon\right)\\
&\quad \le \mathbb{P}\left(|E_n| > \frac{\varepsilon}{2}\right) + \sum_{r=1}^{m-1} \mathbb{P}\left(\max_{0 \le j \le n}\left|\frac{\tilde{S}_j^r}{\sqrt{n}} - \tilde{B}^r\left(\frac{j}{n}\right)\right| > \frac{\varepsilon}{4\sigma_r(m-1)}\right)\\
&\quad \le \frac{\varepsilon}{2}(1 + (m-1)^2\sigma_{\max}^2) + \sum_{r=1}^{m-1}\left(1 + \frac{\sigma_r\sqrt{n}}{1 - |\mu_r|}\right) \exp\left(-\frac{\xi\sigma_r}{4}\frac{\varepsilon}{4\sigma_r(m-1)}\sqrt{n}\right),
\end{aligned}$$

and the proof is complete. □

LEMMA 4.5. *For any* $\varepsilon > 0$,

$$\mathbb{P}\big(\big|\widetilde{H}_{n,k} - \sqrt{p_{\max}} J_k\big| > \varepsilon\big) \le 4(m-1)n \exp\biggl(\frac{-\varepsilon^2 n}{8\sigma_{\max}^2 (m-1)^2}\biggr).$$

PROOF. Comparing (2.3) and (4.4),

$$\begin{aligned}
&\big|\widetilde{H}_{n,k} - \sqrt{p_{\max}} J_k\big| \\
&\quad \le \sum_{r=1}^{m-1}\Bigl(1+\frac{r}{m}\Bigr)\sigma_r \max_{0\le j\le n-1}\sup_{0\le t\le 1/n}\biggl|\tilde{B}^r\Bigl(\frac{j}{n}+t\Bigr) - \tilde{B}^r\Bigl(\frac{j}{n}\Bigr)\biggr| \qquad (4.12)\\
&\quad \le 2\sum_{r=1}^{m-1}\sigma_r \max_{0\le j\le n-1}\sup_{0\le t\le 1/n}\biggl|\tilde{B}^r\Bigl(\frac{j}{n}+t\Bigr) - \tilde{B}^r\Bigl(\frac{j}{n}\Bigr)\biggr|.
\end{aligned}$$

Here, for any $\delta > 0$ and $0 \le r \le m-1$,

$$\begin{aligned}
&\mathbb{P}\biggl(\max_{0\le j\le n-1}\max_{0\le t\le 1/n}\biggl|\tilde{B}^r\Bigl(\frac{j}{n}+t\Bigr) - \tilde{B}^r\Bigl(\frac{j}{n}\Bigr)\biggr| > \delta\biggr) \\
&\quad \le \sum_{j=0}^{n-1}\mathbb{P}\biggl(\max_{0\le t\le 1/n}\biggl|\tilde{B}^r\Bigl(\frac{j}{n}+t\Bigr) - \tilde{B}^r\Bigl(\frac{j}{n}\Bigr)\biggr| > \delta\biggr) \\
&\quad = \sum_{j=0}^{n-1}\mathbb{P}\Bigl(\max_{0\le t\le 1}|\tilde{B}^r(t)| > \delta\sqrt{n}\Bigr) \qquad (4.13)\\
&\quad = \sum_{j=0}^{n-1} 2\mathbb{P}\big(|N(0,1)| > \delta\sqrt{n}\big) \\
&\quad \le 4n\exp(-\delta^2 n/2),
\end{aligned}$$

where, above, we have used standard Gaussian estimates. Using (4.12) and (4.13), we finally get

$$\begin{aligned}
&\mathbb{P}\big(\big|\widetilde{H}_{n,k} - \sqrt{p_{\max}} J_k\big| > \varepsilon\big) \\
&\quad \le \mathbb{P}\biggl(2\sum_{r=1}^{m-1}\sigma_r \max_{0\le j\le n-1}\sup_{0\le t\le 1/n}\biggl|\tilde{B}^r\Bigl(\frac{j}{n}+t\Bigr) - \tilde{B}^r\Bigl(\frac{j}{n}\Bigr)\biggr| > \varepsilon\biggr) \\
&\quad \le \sum_{r=1}^{m-1}\mathbb{P}\biggl(\max_{0\le j\le n-1}\sup_{0\le t\le 1/n}\biggl|\tilde{B}^r\Bigl(\frac{j}{n}+t\Bigr) - \tilde{B}^r\Bigl(\frac{j}{n}\Bigr)\biggr| > \frac{\varepsilon}{2\sigma_r(m-1)}\biggr) \\
&\quad \le 4(m-1)n\exp\biggl(\frac{-\varepsilon^2 n}{8\sigma_{\max}^2(m-1)^2}\biggr).
\end{aligned}$$

□

5. Uniform binary letters. In general, we do not known whether or not the bound in Theorem 4.1 can be sharpened to $\mathcal{O}(1/\sqrt{n})$. As shown below, with a more direct proof, for binary alphabets with uniform distribution this is possible.

Note that for binary alphabets with nonuniform distribution, that is, for $m=2$ and $k=1$, the limiting distribution J_1 is a normal random variable with zero mean and variance $1-p_{\max}$. Although the proof of Theorem 4.1 simplifies in this special case, it still yields

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{LI_n-np_{\max}}{\sqrt{np_{\max}}}\geq x\right)-\bar{\Phi}\left(\frac{x}{\sqrt{1-p_{\max}}}\right)\right| \leq c\left(\sigma_1^2+\frac{1}{\sqrt{p_{\max}(1-p_{\max})}}\right)\frac{\log n}{\sqrt{n}},$$

where $\bar{\Phi}$ is the standard normal survival function.

In this section, $m=2$ and assume $\mathbb{P}(X_i=\alpha_1)=\mathbb{P}(X_i=\alpha_2)=1/2$, $i\in\mathbb{N}$. Let

$$Z_i=\begin{cases}1, & \text{if } X_i=\alpha_1,\\ -1, & \text{if } X_i=\alpha_2,\end{cases}$$

and let $S_0=0$, $S_k=\sum_{i=1}^k Z_i$, $k\geq 1$. Define

$$\widehat{B}_n(t)=\frac{S_{[nt]}}{\sqrt{n}}+(nt-[nt])\frac{Z_{[nt]+1}}{\sqrt{n}},\qquad 0\leq t\leq 1.$$

Then,

$$\frac{LI_n-n/2}{\sqrt{n}}=-\frac{\widehat{B}_n(1)}{2}+\max_{t\in[0,1]}\widehat{B}_n(t),$$

and (2.4) becomes

$$\frac{LI_n-n/2}{\sqrt{n}}\Longrightarrow -\frac{B(1)}{2}+\max_{t\in[0,1]}B(t),$$

where B is a standard Brownian motion.

THEOREM 5.1. *For any $n\in\mathbb{N}$,*

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{LI_n-n/2}{\sqrt{n}}\geq x\right)-\mathbb{P}\left(-\frac{B(1)}{2}+\max_{t\in[0,1]}B(t)\geq x\right)\right|\leq\frac{24}{\sqrt{n}}.$$

PROOF. Note that $\max_{t\in[0,1]}\widehat{B}_n(t)=\max_{k=0,\dots,n}S_k/\sqrt{n}$. Let

$$\bar{F}(m,b):=\mathbb{P}\Big(\max_{t\in[0,1]}B(t)\geq m,\ B(1)\leq b\Big),\qquad m,b\in\mathbb{R}$$

and

$$\bar{F}_n(i,j):=\mathbb{P}\Big(\max_{k=0,\dots,n}S_k\geq i,\ S_n\leq j\Big) = \mathbb{P}\left(\max_{t\in[0,1]}\widehat{B}_n(t)\geq\frac{i}{\sqrt{n}},\ \widehat{B}_n(1)\leq\frac{j}{\sqrt{n}}\right),\qquad i,j\in\mathbb{Z}.$$

By the reflection principle, for any $m \geq 0$, $b \leq m$

$$\begin{aligned}\bar{F}(m,b) &= \mathbb{P}\Big(\max_{t\in[0,1]} B(t) \geq m, B(1) \geq m + (m-b)\Big) \\ &= \mathbb{P}\big(B(1) \geq 2m - b\big) = \bar{\Phi}(2m-b),\end{aligned}$$

and for any $i \geq 0$, $j \leq i$

$$\begin{aligned}\bar{F}_n(i,j) &= \mathbb{P}\Big(\max_{k=0,\dots,n} S_k \geq i, S_n \geq i + (i-j)\Big) \\ &= \mathbb{P}(S_n \geq 2i - j) = \bar{\Phi}_n\Big(2\frac{i}{\sqrt{n}} - \frac{j}{\sqrt{n}}\Big),\end{aligned}$$

where

$$\bar{\Phi}(z) = \mathbb{P}\big(B(1) \geq z\big), \qquad \bar{\Phi}_n(z) = \mathbb{P}\big(S_n/\sqrt{n} \geq z\big), \qquad z \in \mathbb{R}.$$

As is well known (e.g., see [18]),

$$\sup_{z\in\mathbb{R}} |\bar{\Phi}(z) - \bar{\Phi}_n(z)| \leq \frac{0.7975}{\sqrt{n}}. \tag{5.1}$$

Next, the joint probability density function of $(\max_{t\in[0,1]} B(t), B(1))$ is

$$f(m,b) = -\frac{\partial^2 \bar{F}(m,b)}{\partial m\,\partial b} = 2\bar{\Phi}''(2m-b)$$

if $m \geq 0$, $b \leq m$, and zero elsewhere. For any $x \geq 0$, we thus have

$$\begin{aligned}&\mathbb{P}\Big(\max_{t\in[0,1]} B(t) - \frac{B(1)}{2} < x\Big) \\ &\quad= \int_0^{2x}\int_{2m-2x}^{m} f(m,b)\,db\,dm \\ &\quad= \int_0^{2x}\int_{2m-2x}^{m} 2\bar{\Phi}''(2m-b)\,db\,dm \\ &\quad= -2\int_0^{2x} [\bar{\Phi}'(2m-b)]_{b=2m-2x}^{b=m}\,dm \\ &\quad= -2\int_0^{2x} \bar{\Phi}'(m) - \bar{\Phi}'(2x)\,dm = 2\bar{\Phi}(0) - 2\bar{\Phi}(2x) + 2\cdot 2x\bar{\Phi}'(2x) \\ &\quad= 2\bar{\Phi}(0) - 2\bar{\Phi}(2x) - 4x\frac{1}{\sqrt{2\pi}}e^{-(2x)^2/2} = 1 - 2\bar{\Phi}(2x) - 4x\frac{1}{\sqrt{2\pi}}e^{-2x^2}.\end{aligned} \tag{5.2}$$

Observe that S_n is even if n is even, and S_n is odd if n is odd. In the sequel, *assume that* n *is even*, in the other case the computation is similar, and omitted. The joint probability mass function of $(\max_{k=0,\dots,n} S_k, S_n)$ is then

$$p(i,j) = \bar{F}_n(i,j) - \bar{F}_n(i+1,j) - \bar{F}_n(i,j-2) + \bar{F}_n(i+1,j-2)$$

for j even, $i \geq 0$, $j \leq i$, and zero elsewhere.

For any $x \geq 0$, with the notation $l = [x\sqrt{n}]$, we thus have

$$
\begin{aligned}
&\mathbb{P}\left(\max_{t\in[0,1]} \widehat{B}_n(t) - \frac{\widehat{B}_n(1)}{2} < x\right)\\
&\quad = \mathbb{P}\left(\max_{k=0,\dots,n} S_k - \frac{S_n}{2} < l\right)\\
&\quad = \sum_{i=0}^{2l-2} \sum_{\substack{j=2i-2l+2\\ j\,\mathrm{even}}}^{i} p(i,j)\\
&\quad = \sum_{i=0}^{2l-2} [\bar{F}_n(i,i) - \bar{F}_n(i,2i-2l) - \bar{F}_n(i+1,i) + \bar{F}_n(i+1,2i-2l)]\\
&\quad = \sum_{i=0}^{2l-2} \left[\bar{\Phi}_n\left(\frac{i}{\sqrt{n}}\right) - \bar{\Phi}_n\left(\frac{2l}{\sqrt{n}}\right) - \bar{\Phi}_n\left(\frac{i+2}{\sqrt{n}}\right) + \bar{\Phi}_n\left(\frac{2l+2}{\sqrt{n}}\right)\right]\\
&\quad = \bar{\Phi}_n(0) + \bar{\Phi}_n\left(\frac{1}{\sqrt{n}}\right) - \bar{\Phi}_n\left(\frac{2l-1}{\sqrt{n}}\right) - \bar{\Phi}_n\left(\frac{2l}{\sqrt{n}}\right)\\
&\qquad - (2l-2)\left[\bar{\Phi}_n\left(\frac{2l}{\sqrt{n}}\right) - \bar{\Phi}_n\left(\frac{2l+2}{\sqrt{n}}\right)\right]\\
&\quad = \bar{\Phi}_n(0) + \bar{\Phi}_n\left(\frac{2}{\sqrt{n}}\right) - 2\bar{\Phi}_n\left(\frac{2l}{\sqrt{n}}\right) - (2l-2)\mathbb{P}(S_n = 2l);
\end{aligned}
\tag{5.3}
$$

where in the last step we used the fact that $\bar{\Phi}_n$ is constant on the intervals $[\frac{i}{\sqrt{n}}, \frac{i+2}{\sqrt{n}})$, when i is a nonnegative even integer.

Let us compare (5.2) and (5.3). Since for any $x \geq 0$, $2x \in [\frac{2l}{\sqrt{n}}, \frac{2l+2}{\sqrt{n}})$, by (5.1),

$$
\sup_{x\geq 0}\left|2\bar{\Phi}(2x) - 2\bar{\Phi}_n\left(\frac{2l}{\sqrt{n}}\right)\right| = \sup_{x\geq 0}|2\bar{\Phi}(2x) - 2\bar{\Phi}_n(2x)| \leq \frac{1.595}{\sqrt{n}}. \tag{5.4}
$$

Moreover, from symmetry considerations, we know that

$$
\bar{\Phi}_n(0) + \bar{\Phi}_n\left(\frac{2}{\sqrt{n}}\right) = \frac{1}{2} + \frac{1}{2}\mathbb{P}(S_n = 0) + \frac{1}{2} - \mathbb{P}(S_n = 0) = 1 - \frac{1}{2}\mathbb{P}(S_n = 0).
$$

Thus,

$$
1 - \left(\bar{\Phi}_n(0) + \bar{\Phi}_n\left(\frac{2}{\sqrt{n}}\right)\right) = \frac{1}{2}\mathbb{P}(S_n = 0) = \frac{1}{2}\binom{n}{n/2}2^{-n}. \tag{5.5}
$$

Using Stirling's formula

$$
\sqrt{2\pi}\, n^{n+1/2} e^{-n} e^{1/(12n+1)} \leq n! \leq \sqrt{2\pi}\, n^{n+1/2} e^{-n} e^{1/(12n)}, \tag{5.6}
$$

the rightmost term in (5.5) is dominated by

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{(1/2)n(1/2)n}}\frac{n^n}{((1/2)n)^{n/2}((1/2)n)^{n/2}}e^{1/(12n)-1/(6n+1)-1/(6n+1)}2^{-n} \\ &\qquad \le \frac{1}{\sqrt{2\pi}}\frac{2}{\sqrt{n}} \le \frac{0.8}{\sqrt{n}}. \end{aligned} \tag{5.7}$$

Combining (5.4), (5.7) and Lemma 5.2 below will complete the proof. □

LEMMA 5.2. *For any* $n \in \mathbb{N}$,

$$\sup_{x\ge 0}\left|4x\frac{1}{\sqrt{2\pi}}e^{-2x^2} - (2[x\sqrt{n}]-2)\mathbb{P}(S_n = 2[x\sqrt{n}])\right| \le \frac{21}{\sqrt{n}}. \tag{5.8}$$

PROOF. First, consider the range $x \ge \sqrt{n}/6$. In this case, both terms on the left-hand side of (5.8) vanish exponentially fast as $n \to \infty$. Indeed,

$$4x\frac{1}{\sqrt{2\pi}}e^{-2x^2} < \frac{4}{\sqrt{2\pi}}e^{-x^2} \le \frac{4}{\sqrt{2\pi}}e^{-n/36}.$$

Using $\sqrt{n}e^{-n/36} \le \sqrt{18}e^{-18/36} \le \sqrt{18/e}$, we have

$$4x\frac{1}{\sqrt{2\pi}}e^{-2x^2} \le \frac{4\sqrt{18}}{\sqrt{2\pi e}}\frac{1}{\sqrt{n}} = \frac{12}{\sqrt{\pi e}}\frac{1}{\sqrt{n}}. \tag{5.9}$$

If $x \ge \sqrt{n}/2 + 1/\sqrt{n}$, $\mathbb{P}(S_n = 2[x\sqrt{n}]) = 0$. For $\sqrt{n}/6 \le x < \sqrt{n}/2 + 1/\sqrt{n}$, recalling the notation $l = [x\sqrt{n}]$, for $n/6 \le l \le n/2$,

$$\begin{aligned} \mathbb{P}(S_n = 2l) &= \binom{n}{(n+2l)/2}2^{-n} = \binom{n}{n/2+l}2^{-n} \\ &\le \binom{n}{n/2+n/6}2^{-n} = \binom{n}{2n/3}2^{-n}. \end{aligned} \tag{5.10}$$

Using Stirling's formula (5.6) again, (5.10) can be upper bounded by

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{(2/3)n(1/3)n}}\frac{n^n}{((2/3)n)^{2n/3}((1/3)n)^{n/3}}e^{1/(12n)-1/(8n+1)-1/(4n+1)}2^{-n} \\ &\qquad \le \frac{1}{\sqrt{2\pi}}\frac{9}{\sqrt{2n}}\frac{1}{[(2/3)^{2/3}(1/3)^{1/3}2]^n}e^{-123/(540n)} \le \frac{9}{2\sqrt{\pi}}\frac{1}{\sqrt{n}}e^{-n/18}. \end{aligned}$$

Thus,

$$(2l-2)\mathbb{P}(S_n = 2l) \le \frac{9}{2\sqrt{\pi}}\frac{1}{\sqrt{n}}ne^{-n/18}.$$

Since $ne^{-n/18} \le 18/e$, we have

$$(5.11)\qquad (2l-2)\mathbb{P}(S_n = 2l) \le \frac{81}{\sqrt{\pi}e}\frac{1}{\sqrt{n}}.$$

Hence, (5.9) and (5.11) gives the bound (5.8).

Next, consider the range $0 \le x < \sqrt{n}/6$, with the notation $l = [x\sqrt{n}]$, $0 \le l < n/6$. The left-hand side of (5.8) can be upper bounded by

$$(5.12)\qquad \begin{aligned} &\left|4x\frac{1}{\sqrt{2\pi}}e^{-2x^2} - 4\frac{l}{\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-2(l/\sqrt{n})^2}\right| \\ &\qquad + \left|4\frac{l}{\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-2(l/\sqrt{n})^2} - (2l-2)\mathbb{P}(S_n = 2l)\right|. \end{aligned}$$

Since the function xe^{-2x^2} is monotone on the intervals $[0, 1/2)$ and $[1/2, \infty)$,

$$(5.13)\qquad \begin{aligned} &4x\frac{1}{\sqrt{2\pi}}e^{-2x^2} - 4\frac{l}{\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-2(l/\sqrt{n})^2} \\ &\quad \le \frac{4}{\sqrt{2\pi}}\left(\frac{l}{\sqrt{n}} + \frac{1}{\sqrt{n}}\right)e^{-2(l/\sqrt{n}+1/\sqrt{n})^2} - \frac{4}{\sqrt{2\pi}}\frac{l}{\sqrt{n}}e^{-2(l/\sqrt{n})^2} \\ &\quad \le \frac{4}{\sqrt{2\pi}}\left(\frac{l}{\sqrt{n}} + \frac{1}{\sqrt{n}}\right)e^{-2(l/\sqrt{n}+1/\sqrt{n})^2} - \frac{4}{\sqrt{2\pi}}\frac{l}{\sqrt{n}}e^{-2(l/\sqrt{n}+1/\sqrt{n})^2} \\ &\quad \le \frac{4}{\sqrt{2\pi}}\frac{1}{\sqrt{n}}e^{-2(l/\sqrt{n}+1/\sqrt{n})^2} \\ &\quad \le \frac{4}{\sqrt{2\pi}}\frac{1}{\sqrt{n}}. \end{aligned}$$

On the other hand,

$$(5.14)\qquad \begin{aligned} &4\frac{l}{\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-2(l/\sqrt{n})^2} - 4x\frac{1}{\sqrt{2\pi}}e^{-2x^2} \\ &\quad \le \frac{4}{\sqrt{2\pi}}\frac{l}{\sqrt{n}}e^{-2(l/\sqrt{n})^2} - \frac{4}{\sqrt{2\pi}}\left(\frac{l}{\sqrt{n}} + \frac{1}{\sqrt{n}}\right)e^{-2(l/\sqrt{n}+1/\sqrt{n})^2} \\ &\quad \le \frac{4}{\sqrt{2\pi}}\frac{l}{\sqrt{n}}e^{-2(l/\sqrt{n})^2}(1 - e^{-4l/n-2/n}) \\ &\qquad - \frac{4}{\sqrt{2\pi}}\frac{1}{\sqrt{n}}e^{-2(l/\sqrt{n}+1/\sqrt{n})^2} \\ &\quad \le \frac{4}{\sqrt{2\pi}}\frac{l}{\sqrt{n}}e^{-2l^2/n}(1 - e^{-6l/n}) - \frac{4}{\sqrt{2\pi}}\frac{1}{\sqrt{n}}. \end{aligned}$$

Using $1 - e^{-t} \leq t$ $(t \in \mathbb{R})$ with $t = 6l/n$, and also $te^{-t} \leq 1/e$ $(t \in \mathbb{R})$ with $t = 2l^2/n$, the right-most term in (5.14) is dominated by

$$\begin{aligned}
&\frac{12}{\sqrt{2\pi}} \frac{1}{\sqrt{n}} \frac{2l^2}{n} e^{-2l^2/n} - \frac{4}{\sqrt{2\pi}} \frac{1}{\sqrt{n}} \\
&\quad \leq \left(\frac{12}{e\sqrt{2\pi}} - \frac{4}{\sqrt{2\pi}} \right) \frac{1}{\sqrt{n}} \\
&\quad \leq \left(\frac{3}{e} - 1 \right) \frac{4}{\sqrt{2\pi}} \frac{1}{\sqrt{n}}.
\end{aligned} \tag{5.15}$$

From (5.13) and (5.15) we get the following bound for the first term in (5.12):

$$\left| 4x \frac{1}{\sqrt{2\pi}} e^{-2x^2} - 4 \frac{l}{\sqrt{n}} \frac{1}{\sqrt{2\pi}} e^{-2(l/\sqrt{n})^2} \right| \leq \frac{4}{\sqrt{2\pi}} \frac{1}{\sqrt{n}}. \tag{5.16}$$

To control the second term in (5.12), let us recall (see, e.g., Feller [4], page 182) that

$$\mathbb{P}(S_n = 2l) = \frac{2}{\sqrt{n}} \frac{1}{\sqrt{2\pi}} e^{-(2l/\sqrt{n})^2/2} e^{\varepsilon_n}, \tag{5.17}$$

where

$$-\frac{3l^2}{n^2} - \frac{1}{4n} - \frac{1}{360n^3} \leq \varepsilon_n \leq \frac{2l^4}{n^3} - \frac{1}{4n} + \frac{1}{20n^3} \qquad \text{if } l < n/6. \tag{5.18}$$

Hence, for the second term in (5.12), we have

$$\begin{aligned}
&\left| 4 \frac{l}{\sqrt{n}} \frac{1}{\sqrt{2\pi}} e^{-2(l/\sqrt{n})^2} - (2l - 2)\mathbb{P}(S_n = 2l) \right| \\
&\quad \leq \frac{2}{\sqrt{2\pi}} \frac{2l}{\sqrt{n}} e^{-2(l/\sqrt{n})^2} |1 - e^{\varepsilon_n}| + 2\mathbb{P}(S_n = 2l) \\
&\quad \leq \frac{4}{\sqrt{2\pi}} \frac{l}{\sqrt{n}} e^{-2(l/\sqrt{n})^2} |1 - e^{\varepsilon_n}| + \frac{4}{\sqrt{2\pi}} \frac{1}{\sqrt{n}} e^{-2(l/\sqrt{n})^2} e^{\varepsilon_n}.
\end{aligned} \tag{5.19}$$

If $l \leq n^{5/8}/3$, (5.18) becomes

$$-\frac{1}{3n^{1/3}} - \frac{1}{4n} - \frac{1}{360n^3} \leq \varepsilon_n \leq \frac{2}{81\sqrt{n}} - \frac{1}{4n} + \frac{1}{20n^3},$$

and using $|e^z - 1| \leq \max\{|z|, |z + \frac{z^2}{2} \frac{1}{1-|z|}|\}$, $|z| < 1$, (5.19) can be upper bounded by

$$\begin{aligned}
&\frac{4}{\sqrt{2\pi}} \frac{l}{\sqrt{n}} e^{-2(l/\sqrt{n})^2} \frac{2}{3\sqrt{n}} + \frac{4}{\sqrt{2\pi}} \frac{1}{\sqrt{n}} e^{-2(l/\sqrt{n})^2} e^{2/(81\sqrt{n})} \\
&\quad \leq \frac{4}{\sqrt{2\pi}} \frac{1}{2\sqrt{e}} \frac{2}{3\sqrt{n}} + \frac{4}{\sqrt{2\pi}} \frac{1}{\sqrt{n}} e^{2/81} \\
&\quad \leq \frac{4}{\sqrt{2\pi}} \frac{1.23}{\sqrt{n}},
\end{aligned} \tag{5.20}$$

using $ze^{-2z^2} \le 1/(2\sqrt{e})$ with $z = l/\sqrt{n}$.

If $n^{5/8}/3 < l < n/6$, let us consider (5.18) again and apply to (5.19) the trivial upper bound

$$\begin{aligned}
&\frac{4}{\sqrt{2\pi}}\frac{l}{\sqrt{n}}e^{-2(l/\sqrt{n})^2}e^{2l^4/n^3} + \frac{4}{\sqrt{2\pi}}\frac{1}{\sqrt{n}}e^{-2(l/\sqrt{n})^2}e^{2l^4/n^3} \\
&\quad = \frac{4}{\sqrt{2\pi}}\frac{l+1}{\sqrt{n}}e^{-2l^2/n+2l^4/n^3} \qquad (5.21)\\
&\quad \le \frac{8}{6\sqrt{2\pi}}\frac{1}{\sqrt{n}}ne^{-2l^2/n+2l^4/n^3}.
\end{aligned}$$

In this range of l, it is easy to show that $-\frac{2l^2}{n} + \frac{2l^4}{n^3} \le -\frac{35}{18}n^{1/4}$, thus (5.21) is itself dominated by

$$\frac{8}{6\sqrt{2\pi}}\frac{1}{\sqrt{n}}ne^{-35n^{1/4}/18}.$$

Using $ne^{-35n^{1/4}/18} \le (72/35e)^4$, we further get the upper bounds

$$\frac{8}{6\sqrt{2\pi}}\left(\frac{72}{35e}\right)^4\frac{1}{\sqrt{n}} \le \frac{0.44}{\sqrt{2\pi}}\frac{1}{\sqrt{n}}. \qquad (5.22)$$

Since (5.20) is larger than (5.22), when $0 \le x < \sqrt{n}/6$, (5.16) and (5.20) give the following upper bound for (5.12):

$$\begin{aligned}
&\left|4x\frac{1}{\sqrt{2\pi}}e^{-2x^2} - 4\frac{l}{\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-2(l/\sqrt{n})^2}\right| \\
&\quad + \left|4\frac{l}{\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-2(l/\sqrt{n})^2} - (2l-2)\mathbb{P}(S_n = 2l)\right| \\
&\quad \le \frac{4\cdot 2.23}{\sqrt{2\pi}}\frac{1}{\sqrt{n}},
\end{aligned}$$

which is less than what we had obtained for $x \ge \sqrt{n}/6$. □

REFERENCES

[1] BARYSHNIKOV, Y. (2001). GUEs and queues. *Probab. Theory Related Fields* **119** 256–274. MR1818248

[2] BRETON, J.-C. and HOUDRÉ, C. (2010). Asymptotics for random Young diagrams when the word length and alphabet size simultaneously grow to infinity. *Bernoulli* **16** 471–492. MR2668911

[3] ETIENNE, M. P. and VALLOIS, P. (2004). Approximation of the distribution of the supremum of a centered random walk. Application to the local score. *Methodol. Comput. Appl. Probab.* **6** 255–275. MR2073541

[4] FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications*, *Vol.* 1, 3rd ed. Wiley, New York. MR0228020

[5] GLYNN, P. W. and WHITT, W. (1991). Departures from many queues in series. *Ann. Appl. Probab.* **1** 546–572. MR1129774

[6] GRAVNER, J., TRACY, C. A. 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

[7] HOUDRÉ, C. and LITHERLAND, T. On the limiting shape of young diagrams associated with Markov random words. Preprint. Available at arXiv:1110.4570.

[8] HOUDRÉ, C. and LITHERLAND, T. J. (2009). On the longest increasing subsequence for finite and countable alphabets. In *High Dimensional Probability V*: *The Luminy Volume*. *Inst. Math. Stat. Collect.* **5** 185–212. IMS, Beachwood, OH. MR2797948

[9] ITS, A. R., TRACY, C. A. and WIDOM, H. (2001). Random words, Toeplitz determinants, and integrable systems. I. In *Random Matrix Models and Their Applications*. *Math. Sci. Res. Inst. Publ.* **40** 245–258. Cambridge Univ. Press, Cambridge. MR1842789

[10] ITS, A. R., TRACY, C. A. and WIDOM, H. (2001). Random words, Toeplitz determinants and integrable systems. II. *Phys. D* **152/153** 199–224. MR1837910

[11] JOHANSSON, K. (2001). Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. of Math.* (2) **153** 259–296. MR1826414

[12] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32** 111–131. MR0375412

[13] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1976). An approximation of partial sums of independent RV's, and the sample DF. II. *Z. Wahrsch. Verw. Gebiete* **34** 33–58. MR0402883

[14] LIFSHITS, M. (2000). Lecture notes on strong approximation. *Pub. IRMA Lille* **53** No. 13, 25 pp.

[15] MEHTA, M. L. (2004). *Random Matrices*, 3rd ed. *Pure and Applied Mathematics* (*Amsterdam*) **142**. Elsevier/Academic Press, Amsterdam. MR2129906

[16] SAKHANENKO, A. I. (1984). Rate of convergence in the invariance principle for variables with exponential moments that are not identically distributed. In *Limit Theorems for Sums of Random Variables*. *Trudy Inst. Mat.* **3** 4–49. "Nauka" Sibirsk. Otdel., Novosibirsk. MR0749757

[17] TRACY, C. A. and WIDOM, H. (2001). On the distributions of the lengths of the longest monotone subsequences in random words. *Probab. Theory Related Fields* **119** 350–380. MR1821139

[18] VAN BEEK, P. (1972). An application of Fourier methods to the problem of sharpening the Berry–Esseen inequality. *Z. Wahrsch. Verw. Gebiete* **23** 187–196. MR0329000

SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332-0160
USA
E-MAIL: houdre@math.gatech.edu

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF KANSAS
405 SNOW HALL
1460 JAYHAWK BOULEVARD
LAWRENCE, KANSAS 66045-7523
USA
E-MAIL: talata@math.ku.edu