IMS Collections
High Dimensional Probability V: The Luminy Volume
Vol. 5 (2009) 185–212

DOI: 10.1214/09-IMSCOLL513

On the longest increasing subsequence for finite and countable alphabets

Christian Houdré[1] **and Trevis J. Litherland**[1]

Georgia Institute of Technology

Abstract: Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of iid random variables with values in a finite ordered alphabet $\{\alpha_1, \ldots, \alpha_m\}$. Let LI_n be the length of the longest increasing subsequence of $X_1, X_2, \ldots, X_n$. Properly centered and normalized, the limiting distribution of LI_n is expressed as various functionals of m and $(m-1)$-dimensional Brownian motions. These expressions are then related to similar functionals appearing in queueing theory, allowing us to further describe asymptotic behaviors when, in turn, m grows without bound. The finite alphabet results are then used to treat the countable (infinite) alphabet case.

1. Introduction

The pursuit of a robust understanding of the asymptotics of the length of the longest increasing subsequence $L\sigma_n$ of a random permutation of length n – often known as "Ulam's Problem" – has given rise to a remarkable collection of results. The work of Logan and Shepp [22], and Vershik and Kerov [32], first showed that $\mathbb{E}L\sigma_n/\sqrt{n} \to 2$. Following this fundamental asymptotic result, Baik, Deift, and Johansson, in their landmark paper [2], determined the limiting distribution of $L\sigma_n$, properly centered and normalized. This problem has thus emerged as a nexus of once seemingly unconnected mathematical ideas. Indeed, the latter paper is, in particular, quite remarkable for the sheer breadth of mathematical machinery required, machinery calling upon an understanding of random matrix theory, the asymptotics of Toeplitz operators, Riemann-Hilbert Theory, as well as the Robinson-Schensted-Knuth correspondence, to obtain the limiting Tracy-Widom distribution.

Initial approaches to the problem relied heavily on combinatorial arguments. Most work of the last decade, however, such as Seppäläinen [28], have instead used interacting particle processes and so-called "hydrodynamical arguments" to show that $L\sigma_n/\sqrt{n} \to 2$ in expectation and in probability. Building on these ideas, Cator and Groeneboom [5] prove that $\mathbb{E}L\sigma_n/\sqrt{n} \to 2$ in a way that avoids both ergodic decomposition arguments and the subadditive ergodic theorem. Aldous and Diaconis [1] also connect these particle process concepts to the card game solitaire. Finally, Seppäläinen [29] uses these particle processes to solve an open asymptotics problem in queueing theory. Moving beyond the asymptotics of $\mathbb{E}L\sigma_n$, Cator and Groeneboom [6] use particle processes to directly obtain the cube-root asymptotics of the variance of $L\sigma_n$. Further non-asymptotic results for $\mathbb{E}L\sigma_n$ are to be found in Pilpel [26] and Groeneboom [12].

[1]Georgia Institute of Technology, School of Mathematics, Atlanta, Georgia, 30332-0160, e-mail: houdre@math.gatech.edu; trevisl@math.gatech.edu

Keywords and phrases: longest increasing subsequence, Brownian functional, Functional Central Limit Theorem, Tracy-Widom distribution.

AMS 2000 subject classifications: 60C05, 60F05, 60F17, 60G15, 60G17, 05A16.

The related problem of the asymptotics of LI_n, the length of the longest increasing subsequence of a sequence drawn independently and identically from a finite alphabet of size m, has developed along parallel lines. Tracy and Widom [31] as well as Johannson [19] have shown, in the uniform case, that the limiting distribution again enjoys a direct connection to the distribution of the largest eigenvalue in the Gaussian Unitary Ensemble paradigm. Its, Tracy, and Widom [17, 18] have further examined this problem in the inhomogeneous case, relating the limiting distribution to certain direct sums of GUEs. In another direction, Chistyakov and Götze [7] have solved the two-letter Markov case.

Problems from statistical physics have long inspired a lot of the research into these topics. Kuperberg [20], for instance, shows that certain quantum spin matrices are, in law, asymptotically equal to a traceless GUE matrix. The standard general overview of the subject of random matrices is Mehta [23], a work motivated and influenced by some of the origins of the subject in physics.

While the above achievements have undoubtedly stimulated further inquiry, one might still suspect that a more direct route to the limiting distribution of LI_n might be had, one whose methods reflect the essentially probabilistic nature of the problem. This paper proposes a step towards such an approach for the independent finite or infinite countable alphabet case, calling only upon some very well-known results of classical probablity theory described. Indeed, the sequel will show that the limiting distribution of LI_n can be constructed in a most natural manner as a Brownian functional. In the context of random growth processes, Gravner, Tracy, and Widom [11] have already obtained a Brownian functional of the form we derive. This functional appeared first in the work of Glynn and Whitt [10], in queueing theory, and its relation to the eigenvalues of the GUE has also been elucidated by Baryshnikov [3]. It is, moreover, remarked in [11] that the longest increasing subsequence problem could also be studied using a Brownian functional formulation.

We begin our study of this problem, in the next section, by expressing LI_n as a simple algebraic expression. Using this simple characterization, we then briefly determine, in Section 3, the limiting distribution of LI_n in the case of an m-letter alphabet with each letter drawn independently. Our result is expressed as a functional of an $(m-1)$-dimensional Brownian motion with correlated coordinates. Using certain natural symmetries, this limiting distribution is further expressed as various functionals of (standard) multidimensional Brownian motion. Some connections with the Brownian functional originating with the work of Glynn and Whitt in queueing theory are also investigated. This easily leads, from known random matrix results, to various asymptotic limits as m grows. Section 5 is devoted to developing the corresponding results for infinite countable alphabets. In Section 6, we finish the paper by indicating some open questions and future directions for research.

2. Combinatorics

Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of values taken from an m-letter ordered alphabet, $\alpha_1 < \alpha_2 < \cdots < \alpha_m$. Let a_k^r be the number of occurrences of α_r among $X_1, X_2, \ldots, X_k$, $1 \le k \le n$. Each increasing subsequence of $X_1, X_2, \ldots, X_n$ consists simply of runs of identical values, with the values of each successive run forming an increasing subsequence of α_r. Moreover, the number of occurrences of $\alpha_r \in \{\alpha_1, \ldots, \alpha_m\}$ among $X_{k+1}, \ldots, X_\ell$, where $1 \le k < \ell \le n$, is simply $a_\ell^r - a_k^r$. The length of the longest increasing subsequence of $X_1, X_2, \ldots, X_n$ is then given by

$$LI_n = \max_{\substack{0 \le k_1 \le \cdots \\ \le k_{m-1} \le n}} [(a^1_{k_1} - a^1_0) + (a^2_{k_2} - a^2_{k_1}) + \cdots + (a^m_n - a^m_{k_{m-1}})], \tag{2.1}$$

i.e.,

$$LI_n = \max_{\substack{0 \le k_1 \le \cdots \\ \le k_{m-1} \le n}} [(a^1_{k_1} - a^2_{k_1}) + (a^2_{k_2} - a^3_{k_2}) + \cdots + (a^{m-1}_{k_{m-1}} - a^m_{k_{m-1}}) + a^m_n], \tag{2.2}$$

where $a^r_0 = 0$. For $i = 1, \ldots, n$ and $r = 1, \ldots, m-1$, let

$$Z^r_i = \begin{cases} 1, & \text{if } X_i = \alpha_r, \\ -1, & \text{if } X_i = \alpha_{r+1}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.3}$$

and let $S^r_k = \sum_{i=1}^k Z^r_i$, $k = 1, \ldots, n$, and also $S^r_0 = 0$. Then clearly $S^r_k = a^r_k - a^{r+1}_k$. Hence,

$$LI_n = \max_{\substack{0 \le k_1 \le \cdots \\ \le k_{m-1} \le n}} \{S^1_{k_1} + S^2_{k_2} + \cdots + S^{m-1}_{k_{m-1}} + a^m_n\}. \tag{2.4}$$

Since $a^1_k, \ldots, a^m_k$ must evidently sum to k, we have

$$\begin{aligned} n &= \sum_{r=1}^m a^r_n \\ &= \sum_{r=1}^{m-1} \left(a^m_n + \sum_{j=r}^{m-1} S^j_n \right) + a^m_n \\ &= \sum_{r=1}^{m-1} r S^r_n + m a^m_n. \end{aligned}$$

Solving for a^m_n gives us

$$a^m_n = \frac{n}{m} - \frac{1}{m} \sum_{r=1}^{m-1} r S^r_n.$$

Substituting into (2.4), we finally obtain

$$LI_n = \frac{n}{m} - \frac{1}{m} \sum_{r=1}^{m-1} r S^r_n + \max_{\substack{0 \le k_1 \le \cdots \\ \le k_{m-1} \le n}} \{S^1_{k_1} + S^2_{k_2} + \cdots + S^{m-1}_{k_{m-1}}\}. \tag{2.5}$$

The expression (2.5) is of a *purely combinatorial nature or, in more probabilistic terms, is of a pathwise nature*. We now analyze (2.5) in light of the probabilistic nature of the sequence $X_1, X_2, \ldots, X_n$.

3. Probabilistic Development

We consider first the simpler case in which $X_1, X_2, \ldots, X_n, \ldots$ are iid, with each letter drawn uniformly from $\mathcal{A} = \{\alpha_1, \ldots, \alpha_m\}$.

Proposition 3.1. *Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of iid random variables drawn uniformly from the ordered finite alphabet $\mathcal{A} = \{\alpha_1, \ldots, \alpha_m\}$. Then*

$$\frac{LI_n - n/m}{\sqrt{2n/m}} \Rightarrow -\frac{1}{m}\sum_{i=1}^{m-1} i\tilde{B}^i(1) + \max_{\substack{0 \le t_1 \le \cdots \\ \le t_{m-1} \le 1}} \sum_{i=1}^{m-1} \tilde{B}^i(t_i), \tag{3.1}$$

where $(\tilde{B}^1(t), \ldots, \tilde{B}^{m-1}(t))_{0 \le t \le 1}$ is an $(m-1)$-dimensional (centered) Brownian motion with covariance matrix given by

$$t\begin{pmatrix} 1 & -1/2 & & & \bigcirc \\ -1/2 & 1 & -1/2 & & \\ & \ddots & \ddots & \ddots & \\ \bigcirc & & -1/2 & 1 & -1/2 \\ & & & -1/2 & 1 \end{pmatrix}. \tag{3.2}$$

Proof. For each fixed letter r, the sequence $Z_1^r, Z_2^r, \ldots Z_n^r, \ldots$ as defined in (2.3) is also formed of iid random variables with $\mathbb{P}(Z_1^r = 1) = \mathbb{P}(Z_1^r = -1) = 1/m$, and $\mathbb{P}(Z_1^r = 0) = 1 - 2/m$. Thus $\mathbb{E}Z_1^r = 0$, and $\mathbb{E}(Z_1^r)^2 = 2/m$, and so, $\mathrm{Var}S_n^r = 2n/m$, for $r = 1, 2, \ldots, m-1$. Defining $\hat{B}_n^r(t) = \frac{1}{\sqrt{2n/m}} S_{[nt]}^r + \frac{1}{\sqrt{2n/m}}(nt - [nt])Z_{[nt]+1}^r$, for $0 \le t \le 1$, and noting that the local maxima of $\hat{B}_n^i(t)$ occur at $t = k/n$, $k = 0, \ldots, n$, we have from (2.5) that

$$\frac{LI_n - n/m}{\sqrt{2n/m}} = -\frac{1}{m}\sum_{i=1}^{m-1} i\hat{B}_n^i(1) + \max_{\substack{0 \le t_1 \le \cdots \\ \le t_{m-1} \le 1}} [\hat{B}_n^1(t_1) + \cdots + \hat{B}_n^{m-1}(t_{m-1})]. \tag{3.3}$$

We can now invoke Donsker's Theorem since the measures $\mathbb{P}_n$ generated by $(\hat{B}_n^1(t), \ldots, \hat{B}_n^{m-1}(t))$ satisfy $\mathbb{P}_n(A) \to \mathbb{P}_\infty(A)$, for all Borel subsets A of the space of continuous functions $C([0,1]^{m-1})$ for which $\mathbb{P}_\infty(\partial A) = 0$, where $\mathbb{P}_\infty$ is the $(m-1)$-dimensional Wiener measure. Thus, by Donsker's Theorem and the Continuous Mapping Theorem we have that $(\hat{B}_n^1(t), \ldots, \hat{B}_n^{m-1}(t)) \Rightarrow (\tilde{B}^1(t), \ldots, \tilde{B}^{m-1}(t))$, where the Brownian motion on the right has a covariance structure which we now describe. First, $\mathrm{Cov}(Z_1^r, Z_1^s) = \mathbb{E}Z_1^r Z_1^s = 0$, for $|r - s| \ge 2$, and $\mathrm{Cov}(Z_1^r, Z_1^{r+1}) = \mathbb{E}Z_1^r Z_1^{r+1} = -1/m$, for $r = 1, 2, \ldots, m-1$. Then, as already noted, for each fixed r, $Z_1^r, Z_2^r, \ldots Z_n^r, \ldots$ are iid, and for fixed k, $Z_k^1, Z_k^2, \ldots, Z_k^{m-1}$ are dependent but identically distributed random variables. Moreover, it is equally clear that for any r and s, $1 \le r < s \le m-1$, the sequences $(Z_k^r)_{k \ge 1}$ and $(Z_\ell^s)_{\ell \ge 1}$ are also identical distributions of the Z_k^r and that Z_k^r and Z_ℓ^s are independent for $k \ne \ell$. Thus, $\mathrm{Cov}(S_n^r, S_n^s) = n\mathrm{Cov}(Z_1^r, Z_1^s)$. This result, together with the $2n/m$ normalization factor gives (3.2) as covariance matrix of $(\tilde{B}^1(t), \ldots, \tilde{B}^{m-1}(t))$.

Next, the functional in (3.3) is a bounded continuous functional on $C([0,1]^{m-1})$. (This type of facts will be used throughout the paper.) Hence, by a final application of the Continuous Mapping Theorem,

$$\frac{LI_n - n/m}{\sqrt{2n/m}} \Rightarrow -\frac{1}{m}\sum_{i=1}^{m-1} i\tilde{B}^i(1) + \max_{\substack{0 \le t_1 \le \cdots \\ \le t_{m-1} \le 1}} \sum_{i=1}^{m-1} \tilde{B}^i(t_i). \tag{3.4}$$

□

Remark 3.1. (i) Above, we obtained the limiting distribution of LI_n as a Brownian functional. Tracy and Widom [31] already obtained the limiting distribution of LI_n as the law of the largest eigenvalue of the traceless Gaussian Unitary Ensemble (GUE) of $m \times m$ Hermitian matrices. Johansson [19] generalized this result to encompass all m eigenvalues. Gravner, Tracy, and Widom [11] in their study of random growth processes make a connection between the distribution of the largest eigenvalue in the $m \times m$ GUE and a Brownian functional essentially equivalent, up to a normal random variable, to the right hand side of (3.4). (This will become clear as we refine our understanding of (3.4) in the sequel.)
(ii) For $m = 2$, (3.1) simply becomes

$$\frac{LI_n - n/2}{\sqrt{n}} \Rightarrow -\frac{1}{2}B(1) + \max_{0\le t\le 1} B(t), \tag{3.5}$$

where B is standard one-dimensional Brownian motion. A well-known result of Pitman [27] implies that, up to a factor of 2, the functional in (3.5) is identical in law to the radial part of a three-dimensional standard Brownian motion at time $t = 1$. Specifically, Pitman shows that the two processes $(2\max_{0\le s\le t} B(s) - B(t))_{t\ge 0}$ and $(\sqrt{(B^1(t))^2 + (B^2(t))^2 + (B^3(t))^2})_{t\ge 0}$, have the same law, where $(B^1(t), B^2(t), B^3(t))_{t\ge 0}$ is a standard 3-dimensional Brownian motion.

It is trivial to show that the functional in (3.5) does indeed have the same distribution as that of the largest eigenvalue of a 2×2 matrix of the form

$$\begin{pmatrix} X & Y + iZ \\ Y - iZ & -X \end{pmatrix},$$

where X, Y, and Z are centered independent normal random variables, all with variance $1/4$.

It is instructive to express (3.1) in terms of an $(m-1)$-dimensional standard Brownian motion $(B^1(t), \dots, B^{m-1}(t))$. It is not hard to check that we can express $\tilde{B}^i(t)$, $i = 1, \dots, m-1$, in terms of the $B^i(t)$ as follows:

$$\tilde{B}^i(t) = \begin{cases} B^1(t), & i = 1, \\ \sqrt{\frac{i+1}{2i}}B^i(t) - \sqrt{\frac{i-1}{2i}}B^{i-1}(t), & 2 \le i \le m-1. \end{cases} \tag{3.6}$$

Substituting (3.6) back into (3.1), we obtain a more symmetric expression for our limiting distribution:

$$\frac{LI_n - n/m}{\sqrt{n}} \Rightarrow \frac{1}{\sqrt{m}} \max_{\substack{0\le t_1\le\cdots \\ \le t_{m-1}\le t_m=1}} \sum_{i=1}^{m-1}\left[-\sqrt{\frac{i}{i+1}}B^i(t_{i+1}) + \sqrt{\frac{i+1}{i}}B^i(t_i)\right]. \tag{3.7}$$

The above Brownian functional is similar to one introduced by Glynn and Whitt [10], in the context of a queueing problem:

$$D_m = \max_{\substack{0=t_0\le t_1\le\cdots \\ \le t_{m-1}\le t_m=1}} \sum_{i=1}^{m}\left[B^i(t_i) - B^i(t_{i-1})\right], \tag{3.8}$$

where $(B^1(t), \dots, B^m(t))$ is an m-dimensional standard Brownian motion. Gravner, Tracy, and Widom [11], in studying a one-dimensional discrete space and discrete time process, have shown that its limiting distribution is equal to both that of D_m and also that of the largest eigenvalue $\lambda_1^{(m)}$ of an $m \times m$ Hermitian matrix taken from a GUE. That is, D_m and $\lambda_1^{(m)}$ are in fact identical in law. Independently, Baryshnikov [3], studying closely related problems of queueing theory and of monotonous paths on the integer lattice, has shown that the *process* $(D_m)_{m\geq 1}$ has the same law as the *process* $(\lambda_1^{(m)})_{m\geq 1}$, where $\lambda_1^{(m)}$ is the largest eigenvalue of the matrix consisting of the first m rows and m columns of an infinite matrix in the Gaussian Unitary Ensemble.

Remark 3.2. It is quite clear that $LI_n \geq n/m$, since at least one of the m letters must lie on a substring of length at least n/m. Hence, the limiting functional in (3.1) must be supported on the positive real line. We can also see directly that the functional on the right hand side of (3.7) is non-negative. Indeed, for consider the more general Brownian functional of the form

$$\max_{\substack{0\leq t_1\leq\cdots\\ \leq t_{m-1}\leq t_m=1}} \sum_{i=1}^{m-1} \left[\beta_i B^i(t_{i+1}) - \eta_i B^i(t_i)\right],$$

where $0 \leq \beta_i \leq \eta_i$, for $i = 1, 2, \ldots, m-1$. Now for any fixed $t_{i+1} \in (0,1]$, $i = 1, \ldots, m-1$, $\max_{0\leq t_i\leq t_{i+1}}[\beta_i B^i(t_{i+1}) - \eta_i B^i(t_i)]$ is at least as large as the maximum value at the two extremes, that is, when $t_i = 0$ or $t_i = t_{i+1}$. These two values are simply $\beta_i B^i(t_{i+1})$ and $(\beta_i - \eta_i)B^i(t_{i+1})$. Since $0 \leq \beta_i \leq \eta_i$, at least one of these two values is non-negative. Hence, we can successively find $t_{m-1}, t_{m-2}, \ldots, t_1$ such that *each* term of the functional is non-negative. Thus the whole functional must be non-negative. Taking $\beta_i = \sqrt{i/(i+1)}$ and $\eta_i = \sqrt{(i+1)/i}$, the result holds for (3.7). The functional of Glynn and Whitt in (3.8) does not succumb to the same analysis since the $i = 1$ term demands that $t_0 = 0$.

Let us now turn our attention to the m-letter case wherein each letter $\alpha_r \in \mathcal{A}$ occurs with probability $0 < p_r < 1$, independently, and the p_r need not be equal as in the previous uniform case. For the non-uniform case, Its, Tracy, and Widom in [17] and [18] obtained the limiting distribution of LI_n. Reordering the probabilities such that $p_1 \geq p_2 \geq \cdots \geq p_m$, and grouping those probabilities having identical values $p_{(j)}$ of multiplicity k_j, $j = 1, \ldots, d$, (so that $\sum_{j=1}^d k_j = m$ and $\sum_{j=1}^d p_{(j)} k_j = 1$), they show that the limiting distribution is identical to the distribution of the largest eigenvalue of the direct sum of d mutually independent $k_j \times k_j$ GUEs, whose eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_m) = (\lambda_1^{k_1}, \lambda_2^{k_1}, \ldots, \lambda_{k_1}^{k_1}, \ldots, \lambda_1^{k_d}, \lambda_2^{k_d}, \ldots, \lambda_{k_d}^{k_d})$ satisfy $\sum_{i=1}^m \sqrt{p_i}\lambda_i = 0$. With the above ordering of the probabilities, the limiting distribution simplifies to an integral involving only p_1 and k_1.

We now state our own result in terms of functionals of Brownian motion. The form of the functional described, right below, is rather heavy. It is however this form which is needed to tackle, in the next section, the countable alphabet case. Nicer representations involving independent Brownian motions and more akin to those coming out of random matrix theory will be obtained afterwards. These nicer representations will also permit to transfer to our framework, and as m grows, asymptotic results available in random matrix theory.

Theorem 3.1. *Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of iid random variables taking values in an ordered finite alphabet $\mathcal{A} = \{\alpha_1, \ldots, \alpha_m\}$, such that $\mathbb{P}(X_1 = \alpha_r) = p_r$,* *for $r = 1, \ldots, m$, where $0 < p_r < 1$ and $\sum_{r=1}^m p_r = 1$. Then*

$$\frac{LI_n - np_{max}}{\sqrt{n}} \Rightarrow -\frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i \tilde{B}^i(1) + \max_{\substack{0=t_0\le t_1\le\cdots \\ \le t_{m-1}\le t_m=1 \\ t_i=t_{i-1},\ i\in I^*}} \sum_{i=1}^{m-1}\sigma_i \tilde{B}^i(t_i), \tag{3.9}$$

where $p_{max} = \max_{1\le r\le m} p_r$, $\sigma_r^2 = p_r + p_{r+1} - (p_r - p_{r+1})^2$, $r = 1, \ldots, m-1$, $I^* = \{r \in \{1, \ldots, m\} : p_r < p_{max}\}$, and where $(\tilde{B}^1(t), \ldots, \tilde{B}^{m-1}(t))_{0\le t\le 1}$ is an $(m-1)$-dimensional Brownian motion with covariance matrix given by

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

with

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

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

Proof. As before, we begin with the expression for LI_n displayed in (2.5), noting that for each letter α_r, $1 \le r \le m-1$, $(Z_k^r)_{k\ge 1}$ forms a sequence of iid random variables, and that moreover Z_k^r and Z_ℓ^s are independent for $k \ne \ell$, and for any r and s. Now, however, for each fixed k, the Z_k^r are no longer identically distributed; indeed,

$$\begin{cases} \mu_r := \mathbb{E}Z_1^r = p_r - p_{r+1}, & 1 \le r \le m-1, \\ \sigma_r^2 := \mathrm{Var} Z_1^r = p_r + p_{r+1} - (p_r - p_{r+1})^2, & 1 \le r \le m-1. \end{cases} \tag{3.10}$$

Since $0 < p_r < 1$, we have $\sigma_r^2 > 0$ for all $1 \le r \le m-1$. We are thus led to define our Brownian approximation by

$$\hat{B}_n^r(t) := \frac{S_{[nt]}^r - \mu_r[nt]}{\sigma_r\sqrt{n}} + (nt - [nt])\frac{Z_{[nt]+1}^r - \mu_r}{\sigma_r\sqrt{n}}, \quad 0 \le t \le 1, \quad 1 \le r \le m-1. \tag{3.11}$$

Again noting that the local maxima of $\hat{B}_n^i(t)$ occur on the set $\{t : t = k/n, k = 0, \ldots, n\}$, (2.5) becomes

$$\begin{aligned} LI_n = \frac{n}{m} - \frac{1}{m}\sum_{i=1}^{m-1} i\left[\sigma_i \hat{B}_n^i(1)\sqrt{n} + \mu_i n\right] \\ + \max_{\substack{0=t_0\le t_1\le\cdots \\ \le t_{m-1}\le t_m=1}} \left\{\sum_{i=1}^{m-1}\left[\sigma_i \hat{B}_n^i(t_i)\sqrt{n} + \mu_i t_i n\right]\right\}. \end{aligned} \tag{3.12}$$

Next,

$$\sum_{i=1}^{m-1} i\mu_i = \sum_{i=1}^{m-1}\sum_{j=i}^{m-1}\mu_j = \sum_{i=1}^{m-1}\sum_{j=i}^{m-1}(p_j - p_{j+1})$$
$$= \sum_{i=1}^{m-1}(p_i - p_m) = (1-p_m) - (m-1)p_m$$
$$= 1 - mp_m.$$

Hence, (3.12) becomes

$$LI_n = \frac{n}{m} - \frac{(1-mp_m)n}{m} - \frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i \hat{B}_n^i(1)\sqrt{n}$$
$$+ \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{m-1}\le t_m=1}} \sum_{i=1}^{m-1}\left[\sigma_i \hat{B}_n^i(t_i)\sqrt{n} + \mu_i t_i n\right], \tag{3.13}$$

and, dividing through by $\sqrt{n}$, we obtain

$$\frac{LI_n}{\sqrt{n}} = p_m\sqrt{n} - \frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i\hat{B}_n^i(1)$$
$$+ \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{m-1}\le t_m=1}} \sum_{i=1}^{m-1}\left[\sigma_i\hat{B}_n^i(t_i) + \mu_i t_i\sqrt{n}\right]. \tag{3.14}$$

Let $t_0 = 0$, and let $\Delta_i = t_i - t_{i-1}$, $i = 1, \ldots, m-1$. Since

$$\sum_{i=1}^{m-1}\mu_i t_i = \sum_{i=1}^{m-1}\mu_i\sum_{j=1}^{i}\Delta_j = \sum_{i=1}^{m-1}\Delta_i\sum_{j=i}^{m-1}\mu_j = \sum_{i=1}^{m-1}\Delta_i(p_i - p_m),$$

(3.14) becomes

$$\frac{LI_n}{\sqrt{n}} = p_m\sqrt{n} - \frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i\hat{B}_n^i(1)$$
$$+ \max_{\substack{\Delta_i\ge 0\\ \sum_{i=1}^{m-1}\Delta_i\le 1}} \left\{\sum_{i=1}^{m-1}\sigma_i\hat{B}_n^i(t_i) + \sqrt{n}\sum_{i=1}^{m-1}\Delta_i(p_i - p_m)\right\}, \tag{3.15}$$

where $t_i = \sum_{j=1}^{i}\Delta_j$.

Recalling that $t_m := 1$, and setting $\Delta_m = 1 - t_{m-1}$, (3.15) enjoys a more symmetric representation as

$$\frac{LI_n}{\sqrt{n}} = -\frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i\hat{B}_n^i(1)$$
$$+ \max_{\substack{\Delta_i\ge 0\\ \sum_{i=1}^{m}\Delta_i = 1}} \left[\sum_{i=1}^{m-1}\sigma_i\hat{B}_n^i(t_i) + \sqrt{n}\sum_{i=1}^{m}\Delta_i p_i\right]. \tag{3.16}$$

Next,

$$\begin{aligned}\frac{LI_n - np_{max}}{\sqrt{n}} &= -\frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i \hat{B}_n^i(1) \\ &+ \max_{\substack{\Delta_i \geq 0 \\ \sum_{i=1}^m \Delta_i = 1}} \left[\sum_{i=1}^{m-1} \sigma_i \hat{B}_n^i(t_i) + \sqrt{n}\sum_{i=1}^m \Delta_i(p_i - p_{max})\right], \end{aligned} \tag{3.17}$$

where $p_{max} = \max_{1\leq i\leq m} p_i$. Clearly, if $\Delta_i > 0$ for any i such that $p_i < p_{max}$, then

$$\sqrt{n}\sum_{i=1}^m \Delta_i(p_i - p_{max}) \xrightarrow{a.s.} -\infty.$$

Intuitively, then, we should demand that $\Delta_i = 0$ for $i \in I^* := \{i \in \{1, 2, \ldots, m\} : p_i < p_{max}\}$. Indeed, we now show that in fact

$$\frac{LI_n - np_{max}}{\sqrt{n}} = -\frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i \hat{B}_n^i(1) + \max_{\substack{0=t_0\leq t_1\leq\cdots \\ \leq t_{m-1}\leq t_m=1 \\ t_i=t_{i-1},\ i\in I^*}} \sum_{i=1}^{m-1} \sigma_i \hat{B}_n^i(t_i) + E_n, \tag{3.18}$$

where the remainder term E_n is a random variable converging to zero in probability as $n \to \infty$.

To see this, let us introduce the following notation. Writing $t = (t_1, t_2, \ldots, t_{m-1}, t_m)$, let $T = \{t : 0 \leq t_1 \leq \cdots \leq t_{m-1} \leq t_m = 1\}$ and let $T^* = \{t \in T : t_i = t_{i-1}, i \in I^*\}$. Setting $C_n(t) = \sum_{i=1}^{m-1} \sigma_i \hat{B}_n^i(t_i)$ and $R(t) = \sum_{i=1}^m (t_i - t_{i-1})(p_{max} - p_i)$, we can rewrite the terms involving max in (3.17) and (3.18) as

$$\max_{t\in T} \left[C_n(t) - \sqrt{n}R(t)\right]$$

and

$$\max_{t\in T^*} C_n(t).$$

By the compactness of T and T^* and the continuity of $C_n(t)$ and $R(t)$, we see that for each n and each $\omega \in \Omega$, there is a $\tau^n \in T$ and a $\tau_*^n \in T^*$ such that

$$C_n(\tau^n) - \sqrt{n}R(\tau^n) = \max_{t\in T}\left[C_n(t) - \sqrt{n}R(t)\right],$$

and

$$C_n(\tau_*^n) = \max_{t\in T^*} C_n(t).$$

(Note that the piecewise-linear nature of $C_n(t)$ and the linear nature of $R(t)$ imply that the arguments maximizing the above must lie on a finite set; thus, the measurability of τ^n and τ_*^n is trivial.)

Now we first claim that the set of optimizing arguments $\{\tau^n\}_{n=1}^\infty$ a.s. does not have an accumulation point lying outside of T^*. Suppose the contrary, namely that for each ω in a set A of positive measure, there is a subsequence $(\tau^{n_k})_{k=1}^\infty$ of $(\tau^n)_{n=1}^\infty$ such that $d(\tau^{n_k}, T^*) > \epsilon$, for some $\epsilon > 0$, where the metric d is the one induced by the L_∞-norm over T, *i.e.*, by $\|t\|_\infty = \max_{1\leq i\leq m} |t_i|$.

Then, since $T^* \subset T$, it follows that, for all n,

$$C_n(\tau^n) - \sqrt{n}R(\tau^n) \geq C_n(\tau_*^n).$$

Now if $p_{max} = p_m$, then $t = (0, \dots, 0, 1) \in T^*$, and if for some $1 \le j \le m-1$ we have $p_{max} = p_j > \max_{j+1 \le i \le m} p_i$, then $t = (0, \dots, 0, 1, \dots, 1) \in T^*$, where there are $j-1$ zeros in t. Hence $C_{n_k}(\tau_*^{n_k}) \ge C_{n_k}(0, \dots, 0, 1, \dots, 1) = \sum_{i=j}^{m-1} \sigma_i \hat{B}^i_{n_k}(1)$, where the sum is taken to be zero for $j = m$. Given $0 < \delta < 1$, by the Central Limit Theorem, we can find a sufficiently negative real α such that

$$\begin{aligned}\mathbb{P}\left(C_{n_k}(\tau^{n_k}) - \sqrt{n_k}R(\tau^{n_k}) \ge \alpha\right) &\ge \mathbb{P}\left(C_{n_k}(\tau_*^{n_k}) \ge \alpha\right) \\ &\ge \mathbb{P}\left(\sum_{i=j}^{m-1} \sigma_i \hat{B}^i_{n_k}(1) \ge \alpha\right) \\ &> 1 - \delta,\end{aligned}$$

for n_k large enough. In particular, this implies that

$$\mathbb{P}\left(A \cap \{C_{n_k}(\tau^{n_k}) - \sqrt{n_k}R(\tau^{n_k}) \ge \alpha\}\right) > \frac{1}{2}\mathbb{P}(A), \tag{3.19}$$

for n_k large enough.

Next, note that for any $t \in T$, we can modify its components t_i to obtain an element of T^*, by collapsing certain consecutive t_is to single values, where $i \in \{j-1, j, \dots, \ell\}$ and $\{j, j+1, \dots, \ell\} \subset I^*$. With this observation, it is not hard to see that by replacing such maximal consecutive sets of components $\{t_i\}_{i=j-1}^{\ell}$ with their median values, we must have

$$d(\tau^{n_k}, T^*) = \max_{\{(j,\ell):\{j,j+1,\dots,\ell\} \subset I^*\}} \frac{(\tau_\ell^{n_k} - \tau_{j-1}^{n_k})}{2}.$$

Writing $p_{(2)}$ for the largest of the $p_i < p_{max}$, we see that for all k, and for almost all $\omega \in A$,

$$\begin{aligned}R(\tau^{n_k}) &= \sum_{i=1}^{m} (\tau_i^{n_k} - \tau_{i-1}^{n_k})(p_{max} - p_i) \\ &= \sum_{i \in I^*} (\tau_i^{n_k} - \tau_{i-1}^{n_k})(p_{max} - p_i) \\ &\ge (p_{max} - p_{(2)}) \sum_{i \in i^*} (\tau_i^{n_k} - \tau_{i-1}^{n_k}) \\ &\ge 2(p_{max} - p_{(2)}) d(\tau^{n_k}, T^*) \ge 2(p_{max} - p_{(2)})\epsilon.\end{aligned}$$

Now by Donsker's Theorem and the Continuous Mapping Theorem, we have that

$$\max_{t \in T} C_n(t) \Rightarrow \max_{t \in T} \sum_{i=1}^{m-1} \sigma_i \tilde{B}^i(t_i),$$

as $n_k \to \infty$, where $(\tilde{B}^1(t), \dots, \tilde{B}^{m-1}(t))$ is an $(m-1)$-dimensional Brownian motion described in greater detail below. The point here is simply that this limiting functional exists. Moreover,

$$\max_{t \in T} C_n(t) \ge C_n(\tau^n),$$

hence, given $0 < \delta < 1$, if M is chosen large enough, then

$$\begin{aligned}\mathbb{P}\left(C_{n_k}(\tau^{n_k}) \le M\right) &\ge \mathbb{P}\left(\max_{t \in T} C_{n_k}(t) \le M\right) \\ &> 1 - \delta,\end{aligned}$$

for n_k large enough.

We can next see how the boundedness of $R(\tau^{n_k})$ on A influences that of the whole expression $C_{n_k}(\tau^{n_k}) - \sqrt{n_k}R(\tau^{n_k})$ via the following estimates. Given $M > 0$ as above, if k is large enough, then

$$n_k \geq ((M - \alpha + 1)/(2(p_{max} - p_{(2)})\epsilon))^2,$$

and also

$$\begin{aligned}
&\mathbb{P}\left(A \cap \{C_{n_k}(\tau^{n_k}) - \sqrt{n_k}R(\tau^{n_k}) \leq \alpha - 1\}\right) \\
&\quad = \mathbb{P}\left(A \cap \{C_{n_k}(\tau^{n_k}) \leq \alpha - 1 + \sqrt{n_k}R(\tau^{n_k})\}\right) \\
&\quad \geq \mathbb{P}\left(A \cap \{C_{n_k}(\tau^{n_k}) \leq \alpha - 1 + \sqrt{n_k}(2(p_{max} - p_{(2)})\epsilon)\}\right) \\
&\quad \geq \mathbb{P}\left(A \cap \{C_{n_k}(\tau^{n_k}) \leq M\}\right) \\
&\quad > \frac{1}{2}\mathbb{P}(A).
\end{aligned}$$

But this contradicts (3.19); thus, our optimal parameter sequences $(\tau^n)_{n=1}^{\infty}$ must a.s. have their accumulation points in T^*.

Thus, given $\epsilon > 0$, there is an integer N_ϵ such that the set $A_{n,\epsilon} = \{d(\tau^k, T^*) < \epsilon^3, k \geq n\}$ satisfies $\mathbb{P}(A_{n,\epsilon}) \geq 1 - \epsilon$, for all $n \geq N_\epsilon$. Now for each τ^n define $\hat{\tau}^n \in T^*$ to be the (not necessarily unique) point of T^* which is closest in the L^∞-distance to τ^n. Recalling that

$$E_n = C_n(\tau^n) - \sqrt{n}R(\tau^n) - C_n(\tau_*^n) \geq 0,$$

and noting that $R(t) \geq 0$, for all $t \in T$, we can estimate the remainder term E_n as follows: for $n \geq N_\epsilon$,

$$\begin{aligned}
\mathbb{P}(E_n \geq \epsilon) &= \mathbb{P}(\{E_n \geq \epsilon\} \cap A_{n,\epsilon}) + \mathbb{P}\left(\{E_n \geq \epsilon\} \cap A_{n,\epsilon}^c\right) \\
&\leq \mathbb{P}(\{E_n \geq \epsilon\} \cap A_{n,\epsilon}) + \mathbb{P}\left(A_{n,\epsilon}^c\right) \\
&\leq \mathbb{P}(\{E_n \geq \epsilon\} \cap A_{n,\epsilon}) + \epsilon \\
&= \mathbb{P}\left(\{C_n(\tau^n) - \sqrt{n}R(\tau^n) - C_n(\tau_*^n) \geq \epsilon\} \cap A_{n,\epsilon}\right) + \epsilon \\
&\leq \mathbb{P}\left(\{C_n(\tau^n) - \sqrt{n}R(\tau^n) - C_n(\hat{\tau}^n) \geq \epsilon\} \cap A_{n,\epsilon}\right) + \epsilon \\
&\leq \mathbb{P}(\{C_n(\tau^n) - C_n(\hat{\tau}^n) \geq \epsilon\} \cap A_{n,\epsilon}) + \epsilon \\
&\leq \mathbb{P}\left(\left|\sum_{i=1}^{m-1} \sigma_i(\hat{B}_n^i(\tau_i^n) - \hat{B}_n^i(\hat{\tau}_i^n))\right| \geq \epsilon\right) + \epsilon. \qquad (3.20)
\end{aligned}$$

To further bound the right-hand side of (3.20), note that for all $n \geq 1$ and all $1 \leq i \leq m-1$, we have $\mathrm{Var}(\hat{B}_n^i(t_i) - \hat{B}_n^i(s_i)) = |t_i - s_i|$. Then, let $(s,t) \in T \times T$ be such that $\|t - s\|_\infty \leq \epsilon^3$. Using the Bienaymé-Chebyshev inequality, we find that for n large enough,

$$\begin{aligned}
\mathbb{P}\left(\left|\sum_{i=1}^{m-1} \sigma_i(\hat{B}_n^i(t_i) - \hat{B}_n^i(s_i))\right| \geq \epsilon\right) &\leq \epsilon^{-2}(m-1)^2 \max_{1\leq i\leq m-1} \sigma_i^2 \|t - s\|_\infty \\
&\leq \epsilon^{-2}(m-1)^2 \max_{1\leq i\leq m-1} \sigma_i^2 \epsilon^3 \\
&= \epsilon(m-1)^2 \max_{1\leq i\leq m-1} \sigma_i^2.
\end{aligned}$$

Since $\|\tau^n - \hat{\tau}^n\| < \epsilon^3$, for $n \geq N_\epsilon$, this can be used to bound (3.20):

$$\begin{aligned}\mathbb{P}\left(|E_n| \geq \epsilon\right) &< \mathbb{P}\left(\left|\sum_{i=1}^{m-1} \sigma_i(\hat{B}_n^i(\tau_i^n) - \hat{B}_n^i(\hat{\tau}_i^n))\right| \geq \epsilon\right) + \epsilon \\ &\leq \epsilon\left\{(m-1)^2 \max_{1\leq i\leq m-1} \sigma_i^2 + 1\right\}.\end{aligned}$$

Finally, ϵ being arbitrary, we have indeed shown that $E_n \to 0$ in probability.

Applying Donsker's Theorem, the Continuous Mapping Theorem, and Slutsky's (or the converging-together) Theorem, e.g., see [4, 8], to (3.18) we finally have:

$$\frac{LI_n - p_{max}n}{\sqrt{n}} \Rightarrow -\frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i \tilde{B}^i(1) + \max_{\substack{0=t_0\leq t_1\leq\cdots \\ \leq t_{m-1}\leq t_m=1 \\ t_i=t_{i-1},\ i\in I^*}} \sum_{i=1}^{m-1} \sigma_i \tilde{B}^i(t_i), \tag{3.21}$$

where $(\tilde{B}^1(t), \ldots, \tilde{B}^{m-1}(t))$ is an $(m-1)$-dimensional Brownian motion covariance matrix, $t(\rho_{r,s})_{r,s}$, where

$$\rho_{r,s} = \begin{cases} 1, & r = s, \\ -\frac{p_r+\mu_r\mu_s}{\sigma_r\sigma_s}, & s = r-1, \\ -\frac{p_s+\mu_r\mu_s}{\sigma_r\sigma_s}, & s = r+1, \\ -\frac{\mu_r\mu_s}{\sigma_r\sigma_s}, & |r-s| > 1, \quad 1 \leq r, s \leq m-1. \end{cases}$$

Now for $t = \ell/n$, and $1 \leq r \leq s \leq m-1$, the covariance structure above is computed as follows:

$$\begin{aligned}\mathrm{Cov}(\hat{B}_n^r(t), \hat{B}_n^s(t)) &= \mathrm{Cov}\left(\sum_{i=1}^{\ell} \frac{Z_i^r - \mu_r}{\sigma_r\sqrt{n}}, \sum_{i=1}^{\ell} \frac{Z_i^s - \mu_s}{\sigma_s\sqrt{n}}\right) \\ &= \frac{1}{n\sigma_r\sigma_s}\mathrm{Cov}\left(\sum_{i=1}^{\ell}(Z_i^r - \mu_r), \sum_{i=1}^{\ell}(Z_i^s - \mu_s)\right) \\ &= \frac{1}{n\sigma_r\sigma_s}\sum_{i=1}^{\ell}\mathrm{Cov}(Z_i^r - \mu_r, Z_i^s - \mu_s) \\ &= \frac{\ell}{n\sigma_r\sigma_s}\mathrm{Cov}(Z_1^r - \mu_r, Z_1^s - \mu_s) \\ &= t\begin{cases} \frac{1}{\sigma_r\sigma_s}\,\sigma_r\sigma_s, & s = r, \\ \frac{1}{\sigma_r\sigma_s}(0 - \mu_r\mu_s - \mu_r\mu_s + \mu_r\mu_s), & s > r+1, \\ \\ \frac{1}{\sigma_r\sigma_s}(-p_s - \mu_r\mu_s - \mu_r\mu_s + \mu_r\mu_s), & s = r+1, \end{cases} \\ &= t\begin{cases} 1 & s = r, \\ -\frac{\mu_r\mu_s}{\sigma_r\sigma_s} & s > r+1, \\ -\frac{(p_s+\mu_r\mu_s)}{\sigma_r\sigma_s} & s = r+1, \end{cases}\end{aligned}$$

using the properties of the Z_k^r noted at the beginning of the proof. □

We now study (3.9) on a case-by-case basis. First, let $I^* = \emptyset$, that is, let $p_i = 1/m$, for $i = 1, \ldots, m$. Then $\sigma_i^2 = 2p_i = 2/m$, for all $i \in \{1, 2, \ldots, m\}$. Hence, simply rescaling (3.9) by $\sqrt{2/m}$ recovers the uniform result in (3.1).

Next, consider the case where $p_{max} = p_j$, for *precisely one* $j \in \{1, \dots, m\}$. We then have $I^* = \{1, 2, \dots, m\} \backslash \{j\}$. This forces us to set $0 = t_0 = t_1 = \cdots = t_{j-1}$ and $t_j = t_{j+1} = \cdots = t_{m-1} = t_m = 1$, in the maximizing term in (3.9) . Then, $(LI_n - np_{max})/\sqrt{n}$ converges to a centered normal random variable. Intuitively, this result is not surprising since the longest increasing subsequence is, asymptotically, a string consisting primarily of the most frequently occurring letter, a string whose length is approximated by a binomial random variable with parameters n and p_{max}. The variance of the limiting normal distribution is, in fact, equal to $p_{max}(1-p_{max})$.

Corollary 3.1. *If* $p_{max} = p_j$ *for precisely one* $j \in \{1, \dots, m\}$, *then*

$$\frac{LI_n - np_{max}}{\sqrt{n}} \Rightarrow -\frac{1}{m} \sum_{i=1}^{m-1} i\sigma_i \tilde{B}^i(1) + \sum_{i=j}^{m-1} \sigma_i \tilde{B}^i(1), \tag{3.22}$$

where the last term in (3.22) *is not present if* $j = m$.

To consolidate our analysis, we treat the general case for which p_{max} occurs exactly k times among $\{p_1, p_2, \dots, p_m\}$, where $2 \le k \le m - 1$. Not only will we recover the natural analogues of the previous cases, but we will also express our results in terms of another functional of Brownian motion which is more symmetric. Combining the $2 \le k \le m-1$ case at hand with the $k = 1$ case previously examined, we have the following:

Corollary 3.2. *Let* $p_{max} = p_{j_1} = p_{j_2} = \cdots = p_{j_k}$ *for* $1 \le j_1 < j_2 < \cdots < j_k \le m$, *for some* $1 \le k \le m - 1$, *and let* $p_i < p_{max}$, *otherwise. Then*

$$\frac{LI_n - np_{max}}{\sqrt{n}} \Rightarrow \sqrt{p_{max}(1 - p_{max})} \max_{\substack{0=t_0\le t_1\le\cdots \\ \le t_{k-1}\le t_k=1}} \sum_{\ell=1}^{k} \left[\tilde{B}^\ell(t_\ell) - \tilde{B}^\ell(t_{\ell-1})\right], \tag{3.23}$$

where the k-dimensional Brownian motion $(\tilde{B}^1(t), \tilde{B}^2(t), \dots, \tilde{B}^k(t))_{0\le t\le 1}$ *has covariance matrix*

$$t \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \rho & 1 & \rho \\ \rho & \cdots & \cdots & \rho & 1 \end{pmatrix}, \tag{3.24}$$

with $\rho = -p_{max}/(1 - p_{max})$.

Proof. Let $p_{max} = p_{j_1} = p_{j_1} = \cdots = p_{j_k}$, with $1 \le j_1 < j_2 < \cdots < j_k \le m$ and $2 \le k \le m-1$, *i.e.*, let $I^* = \{1, 2, \dots, m\} \backslash \{j_1, j_2, \dots, j_k\}$. Set $j_0 = 1$ and $j_{k+1} = m$. Then (3.18) becomes

$$\begin{aligned} \frac{LI_n - np_{max}}{\sqrt{n}} &= -\frac{1}{m} \sum_{i=1}^{m-1} i\sigma_i \hat{B}_n^i(1) + \max_{\substack{0=t_0\le t_1\le\cdots \\ \le t_{m-1}\le t_m=1 \\ t_i=t_{i-1},\ i\in I^*}} \sum_{i=1}^{m-1} \sigma_i \hat{B}_n^i(t_i) + E_n \\ &= -\frac{1}{m} \sum_{i=1}^{m-1} i\sigma_i \hat{B}_n^i(1) + \max_{\substack{0=t_{j_0}\le t_{j_1}\le\cdots \\ \le t_{j_k}\le t_{j_{k+1}}=1}} \sum_{\ell=0}^{k} \sum_{i=j_\ell}^{j_{\ell+1}-1} \sigma_i \hat{B}_n^i(t_{j_l}) + E_n \end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i \hat{B}_n^i(1) \\
&\quad + \max_{\substack{0=t_{j_0}\le t_{j_1}\le\cdots \\ \le t_{j_k}\le t_{j_{k+1}}=1}} \left[\sum_{\ell=1}^{k-1}\sum_{i=j_\ell}^{j_{\ell+1}-1} \sigma_i \hat{B}_n^i(t_{j_\ell}) + \sum_{i=j_k}^{m-1}\sigma_i \hat{B}_n^i(1)\right] + E_n \\
&= \left[-\frac{1}{m}\sum_{i=1}^{m-1} i\sigma_i \hat{B}_n^i(1) + \sum_{i=j_k}^{m-1}\sigma_i \hat{B}_n^i(1)\right] \\
&\quad + \max_{\substack{0=t_{j_0}\le t_{j_1}\le\cdots \\ \le t_{j_k}\le t_{j_{k+1}}=1}} \sum_{\ell=1}^{k-1}\sum_{i=j_\ell}^{j_{\ell+1}-1} \sigma_i \hat{B}_n^i(t_{j_\ell}) + E_n. \qquad (3.25)
\end{aligned}
$$

We immediately recognize that the first term on the right hand side of (3.25) corresponds to $k = 1$ and a binomial random variable. Using the definition of the $\hat{B}_n^i$ in (3.11), (3.25) can then be rewritten as

$$
\begin{aligned}
&\frac{a_n^{j_k} - np_{max}}{\sqrt{n}} + \max_{\substack{0=t_{j_0}\le t_{j_1}\le\cdots \\ \le t_{j_k}\le t_{j_{k+1}}=1}} \sum_{\ell=1}^{k-1}\sum_{i=j_\ell}^{j_{\ell+1}-1} \sigma_i \hat{B}_n^i(t_{j_\ell}) + E_n \\
&= \frac{a_n^{j_k} - np_{max}}{\sqrt{n}} + \max_{\substack{0=t_{j_0}\le t_{j_1}\le\cdots \\ \le t_{j_k}\le t_{j_{k+1}}=1}} \sum_{\ell=1}^{k-1}\sum_{i=j_\ell}^{j_{\ell+1}-1} \sigma_i \left(\frac{S^i_{[nt_{j_\ell}]} - \mu_i\,[nt_{j_\ell}]}{\sigma_i\sqrt{n}}\right) + E_n \\
&= \frac{a_n^{j_k} - np_{max}}{\sqrt{n}} \\
&\quad + \frac{1}{\sqrt{n}} \max_{\substack{0=t_{j_0}\le t_{j_1}\le\cdots \\ \le t_{j_k}\le t_{j_{k+1}}=1}} \sum_{\ell=1}^{k-1}\left(\left(a^{j_\ell}_{[nt_{j_\ell}]} - a^{j_{\ell+1}}_{[nt_{j_\ell}]}\right) - [nt_{j_\ell}]\left(p_{j_\ell} - p_{j_{\ell+1}}\right)\right) + E_n \qquad (3.26)
\end{aligned}
$$

Setting $a_n^{j_{k+1}} = n - \sum_{\ell=1}^{k} a_n^{j_\ell}$, we note that the random vector $(a_n^{j_1}, a_n^{j_2}, \ldots, a_n^{j_{k+1}})$ has a multinomial distribution with parameters n and $(p_{max}, p_{max}, \ldots, p_{max}, 1 - kp_{max})$. It is thus natural to introduce a new Brownian motion approximation as follows:

$$
\check{B}_n^\ell(t) = \frac{a^{j_\ell}_{[nt_{j_\ell}]} - [nt_{j_\ell}]\,p_{max}}{\sqrt{np_{max}(1-p_{max})}}, \quad 1 \le \ell \le k. \qquad (3.27)
$$

Substituting (3.27) into (3.26) gives

$$
\begin{aligned}
&\sqrt{p_{max}(1-p_{max})}\left\{\check{B}_n^k(1) + \max_{\substack{0=t_0\le t_1\le\cdots \\ \le t_{k-1}\le t_k=1}} \sum_{\ell=1}^{k-1}\left[\check{B}_n^\ell(t_\ell) - \check{B}_n^{\ell+1}(t_\ell)\right]\right\} + E_n \\
&= \sqrt{p_{max}(1-p_{max})} \max_{\substack{0=t_0\le t_1\le\cdots \\ \le t_{k-1}\le t_k=1}} \sum_{\ell=1}^{k}\left[\check{B}_n^\ell(t_\ell) - \check{B}_n^\ell(t_{\ell-1})\right] + E_n. \qquad (3.28)
\end{aligned}
$$

By Donsker's Theorem, $(\check{B}_n^1(t), \check{B}_n^2(t), \ldots, \check{B}_n^k(t))$ converges jointly to a k-dimensional Brownian motion $(\tilde{B}^1(t), \tilde{B}^2(t), \ldots, \tilde{B}^k(t))$. This Brownian motion has the covariance structure given by (3.24), with $\rho = -p_{max}/(1 - p_{max})$, a fact which follows immediately from the covariance of the multinomial distribution, where the covariance of any two distinct $a_r^{j_\ell}$ is simply $-rp_{max}^2$, for $1 \le r \le n$. This, together with our analysis of the unique p_{max} case, proves the corollary. □

Remark 3.3. The above results provide a Brownian functional equivalent to the GUE result of Its, Tracy, and Widom [17] (described in detail in the comments preceding Theorem 3.1). Note that the limiting distribution in (3.23) depends only on k and p_{max}; neither the specific values of $j_1, j_2, \ldots, j_k$ nor the remaining values of p_i are material, a fact already noted in [17]. Also, it follows from generic results on Brownian functionals that this limiting law has a density, which in the uniform case is supported on the positive real line, while supported on all of $\mathbb{R}$ in the non-uniform case.

We have already seen in (3.7) that the limiting distribution for the uniform case has a nice representation as a functional of standard Brownian motion. We now also express the limiting distribution in (3.23) as a functional of standard Brownian motion. This new functional extends to the uniform case, although its form is different from that of (3.7). This limiting random variable can be viewed as the sum of a normal one and of a maximal eigenvalue type one.

Corollary 3.3. *Let $p_{max} = p_{j_1} = p_{j_2} = \cdots = p_{j_k}$, for $1 \le j_1 < j_2 < \cdots < j_k \le m$, and some $1 \le k \le m$, and let $p_i < p_{max}$, otherwise. Then*

$$\frac{LI_n - np_{max}}{\sqrt{n}} \Rightarrow \sqrt{p_{max}}\Big\{\frac{\sqrt{1-kp_{max}}-1}{k}\sum_{j=1}^{k} B^j(1)$$

$$+ \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{k-1}\le t_k=1}} \sum_{\ell=1}^{k}\left[B^\ell(t_\ell) - B^\ell(t_{\ell-1})\right]\Big\}, \tag{3.29}$$

where $(B^1(t), B^2(t), \ldots, B^k(t))_{0\le t\le 1}$ is a standard k-dimensional Brownian motion.

Proof. Let us first examine the non-uniform case $1 \le k \le m-1$. Recall that $\rho = -p_{max}/(1-p_{max})$. Now the covariance matrix in (3.24) has eigenvalues $\lambda_1 = 1-\rho = 1/(1-p_{max})$ of multiplicity $k-1$ and $\lambda_2 = 1+(k-1)\rho = (1-kp_{max})/(1-p_{max}) < \lambda_1$ of multiplicity 1. From the symmetries of the covariance matrix, it is not hard to see that we can write each Brownian motion $\tilde{B}^i(t)$ as a linear combination of standard Brownian motions $(B^1(t), \ldots, B^k(t))$ as follows:

$$\tilde{B}^i(t) = \beta B^i(t) + \eta \sum_{j=1, j\ne i}^{k} B^j(t), \qquad i = 1, \ldots, k, \tag{3.30}$$

where

$$\beta = \frac{(k-1)\sqrt{\lambda_1}+\sqrt{\lambda_2}}{k}, \qquad \eta = \frac{-\sqrt{\lambda_1}+\sqrt{\lambda_2}}{k}. \tag{3.31}$$

Substituting (3.30) and (3.31) into (3.23), and noting that $\beta - \eta = \sqrt{\lambda_1} = 1/\sqrt{1-p_{max}}$, we find that

$$\begin{aligned}
&\sqrt{p_{max}(1-p_{max})} \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{k-1}\le t_k=1}} \sum_{\ell=1}^{k}\Big[\tilde B^\ell(t_\ell)-\tilde B^\ell(t_{\ell-1})\Big]\\
&= \sqrt{p_{max}(1-p_{max})} \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{k-1}\le t_k=1}} \sum_{\ell=1}^{k}\Big\{\beta\left[B^\ell(t_\ell)-B^\ell(t_{\ell-1})\right]\\
&\qquad + \eta \sum_{j=1, j\ne\ell}^{k}\left[B^j(t_\ell)-B^j(t_{\ell-1})\right]\Big\}\\
&= \sqrt{p_{max}(1-p_{max})} \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{k-1}\le t_k=1}} \sum_{\ell=1}^{k}\Big\{(\beta-\eta)\left[B^\ell(t_\ell)-B^\ell(t_{\ell-1})\right]\\
&\qquad + \eta \sum_{j=1}^{k}\left[B^j(t_\ell)-B^j(t_{\ell-1})\right]\Big\}\\
&= \sqrt{p_{max}(1-p_{max})} \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{k-1}\le t_k=1}} \Big\{\sum_{\ell=1}^{k}(\beta-\eta)\left[B^\ell(t_\ell)-B^\ell(t_{\ell-1})\right]\\
&\qquad + \eta \sum_{\ell=1}^{k}\sum_{j=1}^{k}\left[B^j(t_\ell)-B^j(t_{\ell-1})\right]\Big\}\\
&= \sqrt{p_{max}(1-p_{max})}\Big\{\eta\sum_{j=1}^{k}B^j(1)+(\beta-\eta)\max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{k-1}\le t_k=1}}\sum_{\ell=1}^{k}\left[B^\ell(t_\ell)-B^\ell(t_{\ell-1})\right]\Big\}\\
&= \sqrt{p_{max}}\Big\{\frac{\sqrt{1-kp_{max}}-1}{k}\sum_{j=1}^{k}B^j(1)+\max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{k-1}\le t_k=1}}\sum_{\ell=1}^{k}\left[B^\ell(t_\ell)-B^\ell(t_{\ell-1})\right]\Big\}.
\end{aligned}$$

(3.32)

To complete the proof, let $k = m$. Then from Proposition 3.1,

$$\frac{LI_n - n/m}{\sqrt{n}} \Rightarrow \sqrt{\frac{2}{m}}\Big\{-\frac{1}{m}\sum_{i=1}^{m-1} i\tilde B^i(1) + \max_{\substack{0\le t_1\le\cdots\\ \le t_{m-1}\le 1}}\sum_{i=1}^{m-1}\tilde B^i(t_i)\Big\}, \tag{3.33}$$

where the $(m-1)$-dimensional Brownian motion $(\tilde B^1(t), \dots, \tilde B^{m-1}(t))$ has a tridiagonal covariance matrix given by (3.2). Now, this last Brownian motion can be obtained from a standard m-dimensional Brownian motion $(B^1(t), \dots, B^m(t))$ via the a.s. transformations

$$\tilde B^i(t) = \frac{1}{\sqrt{2}}(B^i(t) - B^{i+1}(t)), \qquad 1 \le i \le m-1.$$

It is easily verified that $(\tilde B^1(t), \dots, \tilde B^{m-1}(t))$ so obtained does indeed have the covariance structure given by (3.2). Substituting these independent Brownian motions into (3.33), gives the following a.s equalities:

$$\frac{LI_n - n/m}{\sqrt{n}} \Rightarrow \sqrt{\frac{2}{m}}\Big\{-\frac{1}{m}\sum_{i=1}^{m-1} i\tilde{B}^i(1) + \max_{\substack{0\le t_1\le\cdots\\ \le t_{m-1}\le 1}} \sum_{i=1}^{m-1} \tilde{B}^i(t_i)\Big\}$$

$$= \sqrt{\frac{1}{m}}\Big\{-\frac{1}{m}\sum_{i=1}^{m-1} i[B^i(1) - B^{i+1}(1)]$$

$$+ \max_{\substack{0\le t_1\le\cdots\\ \le t_{m-1}\le 1}} \sum_{i=1}^{m-1} [B^i(t_i) - B^{i+1}(t_i)]\Big\}$$

$$= \sqrt{\frac{1}{m}}\Big\{-\frac{1}{m}\sum_{i=1}^{m} B^i(1) + B^m(1)$$

$$+ \max_{\substack{0\le t_1\le\cdots\\ \le t_{m-1}\le 1}} \sum_{i=1}^{m}[B^i(t_i) - B^i(t_{i-1})] - B^m(1)\Big\}$$

$$= \sqrt{\frac{1}{m}}\Big\{-\frac{1}{m}\sum_{i=1}^{m} B^i(1) + \max_{\substack{0\le t_1\le\cdots\\ \le t_{m-1}\le 1}} \sum_{i=1}^{m}[B^i(t_i) - B^i(t_{i-1})]\Big\}, \tag{3.34}$$

which give (3.29), with $k = m$ and $p_{max} = 1/m$. □

We have already obtained several representations for the limiting law in the uniform case. Yet one more pleasing functional for the limiting distribution of LI_n is described in the following

Theorem 3.2. *Let $p_{max} = p_1 = p_2 = \cdots = p_m = 1/m$. Then*

$$\frac{LI_n - n/m}{\sqrt{n}} \Rightarrow \frac{\tilde{H}_m}{\sqrt{m}},$$

where

$$\tilde{H}_m = \sqrt{\frac{m-1}{m}} \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{m-1}\le t_m=1}} \sum_{i=1}^{m} \left[\tilde{B}^i(t_i) - \tilde{B}^i(t_{i-1})\right], \tag{3.35}$$

and where $(\tilde{B}^1(t), \tilde{B}^2(t), \ldots, \tilde{B}^m(t))_{0\le t\le 1}$ is an m-dimensional Brownian motion having covariance matrix (3.24), with $\rho = -1/(m-1)$, and thus such that $\sum_{i=1}^m \tilde{B}^i(t) = 0$, for all $0 \le t \le 1$.

Proof. We show that the functional being maximized in (3.35) has the same covariance structure as the functional being maximized in (3.7), a result which we restate as:

$$\frac{LI_n - n/m}{\sqrt{n}} \Rightarrow \frac{1}{\sqrt{m}} \max_{\substack{0\le t_1\le\cdots\\ \le t_{m-1}\le t_m=1}} \sum_{i=1}^{m-1} \left[\beta_i B^i(t_{i+1}) - \eta_i B^i(t_i)\right], \tag{3.36}$$

where $\beta_i = \sqrt{i/(i+1)}$ and $\eta_i = \sqrt{(i+1)/i}$. From this it will immediately follow that the maxima, over all $0 \le t_1 \le t_2 \le \cdots \le t_{m-1} \le 1$, in both expressions have the same law, clinching the proof.

Let $(\tilde{B}^1(t), \tilde{B}^2(t), \ldots, \tilde{B}^m(t))$ be an m-dimensional Brownian motion with a permutation-invariant covariance matrix described by $\mathrm{Cov}(\tilde{B}^i(t), \tilde{B}^j(t)) = -t/$

$(m-1), i \neq j$, and $\mathrm{Var}\tilde{B}^i(t) = t$. Then, $\mathbb{E}(\sum_{i=1}^m \tilde{B}^i(t))^2 = 0$, for all $0 \le t \le 1$, so that $\sum_{i=1}^m \tilde{B}^i(t)$ is identically equal to zero.

Let $t = (t_1, t_2, \ldots, t_{m-1})$ be a fixed collection of t_i from the Weyl chamber $T = \{(t_1, t_2, \ldots, t_{m-1}) : 0 \le t_1 \le t_2 \le \cdots \le t_{m-1} \le 1\}$. Setting $t_m = 1$, and also setting

$$X_t = \sqrt{\frac{m-1}{m}} \sum_{i=1}^m \left[\tilde{B}^i(t_i) - \tilde{B}^i(t_{i-1})\right], \tag{3.37}$$

we then have

$$\begin{aligned} \mathrm{Cov}(X_t, X_s) &= \frac{m-1}{m} \sum_{1 \le i,j \le m} \mathrm{Cov}(\tilde{B}^i(t_i) - \tilde{B}^i(t_{i-1}), \tilde{B}^j(s_j) - \tilde{B}^j(s_{j-1})) \\ &= \frac{m-1}{m} \sum_{i=1}^m [t_i \wedge s_i - t_i \wedge s_{i-1} - t_{i-1} \wedge s_i + t_{i-1} \wedge s_{i-1}] \\ &\quad - \frac{1}{m} \sum_{i \neq j} [t_i \wedge s_j - t_i \wedge s_{j-1} - t_{i-1} \wedge s_j + t_{i-1} \wedge s_{j-1}]. \end{aligned} \tag{3.38}$$

We can rewrite (3.38) in a clear way by setting $T_1 = [0, t_1]$ and $T_i = (t_i, t_{i+1}]$, $i = 2, \ldots, m-1$, and similarly $S_1 = [0, s_1]$ and $S_i = (s_i, s_{i+1}]$, $i = 2, \ldots, m-1$. Letting *Leb* denote the Lebesgue measure on $[0,1]$, a case-by-case analysis of the relative positions of t_i, t_{i-1}, s_i, and s_{i-1} quickly yields that

$$\begin{aligned} \mathrm{Cov}(X_t, X_s) &= \frac{m-1}{m} \sum_{i=1}^m Leb(T_i \cap S_i) - \frac{1}{m} \sum_{i \neq j} Leb(T_i \cap S_j) \\ &= \frac{m-1}{m} \sum_{i=1}^m Leb(T_i \cap S_i) - \frac{1}{m} \left[1 - \sum_{i=1}^m Leb(T_i \cap S_i)\right] \\ &= -\frac{1}{m} + \sum_{i=1}^m Leb(T_i \cap S_i). \end{aligned} \tag{3.39}$$

To complete the proof, we show that

$$Y_t = \sum_{i=1}^{m-1} \left[\beta_i B^i(t_{i+1}) - \eta_i B^i(t_i)\right], \tag{3.40}$$

has the same covariance structure as X_t, where $\beta_i = \sqrt{i/(i+1)}$ and $\eta_i = \sqrt{(i+1)/i}$. Using the independence of the components of the Brownian motion, we also have

$$\begin{aligned} \mathrm{Cov}(Y_t, Y_s) &= \sum_{i=1}^{m-1} \mathrm{Cov}\left(\beta_i B^i(t_{i+1}) - \eta_i B^i(t_i), \beta_i B^i(s_{i+1}) - \eta_i B^i(s_i)\right) \\ &= \sum_{i=1}^{m-1} \left[\frac{i}{i+1}(t_{i+1} \wedge s_{i+1}) - t_{i+1} \wedge s_i - t_i \wedge s_{i+1} + \frac{i+1}{i}(t_i \wedge s_i)\right] \\ &= \sum_{i=1}^m \frac{i-1}{i} t_i \wedge s_i - \sum_{i=1}^{m-1} \left[t_{i+1} \wedge s_i + t_i \wedge s_{i+1} - \frac{i+1}{i} t_i \wedge s_i\right] \\ &= \frac{m-1}{m} - \sum_{i=1}^{m-1} \left[t_{i+1} \wedge s_i + t_i \wedge s_{i+1} - 2(t_i \wedge s_i)\right]. \end{aligned} \tag{3.41}$$

As before, a simple case-by-case analysis of the summands in (3.41) reveals that

$$\begin{aligned}\operatorname{Cov}(Y_t, Y_s) &= \frac{m-1}{m} - \left[1 - \sum_{i=1}^{m} Leb(T_i \cap S_i)\right] \\ &= -\frac{1}{m} + \sum_{i=1}^{m} Leb(T_i \cap S_i), \end{aligned} \tag{3.42}$$

completing the proof. □

We are now at a point where we can more clearly see the similarities between the functional D_m of Glynn and Whitt in (3.8) and that of (3.7), which we have shown to have the same law as $\tilde{H}_m$ in (3.35). Indeed, the only difference between the functionals is simply that in (3.8) the Brownian motions are independent, while in (3.35) they are subject to the zero-sum constraint. Gravner, Tracy, and Widom [11] have already remarked that random words could be studied via such Brownian functionals. In fact, a restatement of Corollary 3.3 shows that, in law, D_m and $\tilde{H}_m$ differ by a centered normal random variable, as indicated by the next proposition and corollary. This, in turn, will allow us to clearly state (in the next section) asymptotic results for $\tilde{H}_m$ from the known corresponding results for D_m.

Proposition 3.2. *Let*

$$H_m = \sqrt{2}\left\{-\frac{1}{m}\sum_{i=1}^{m-1} i\tilde{B}^i(1) + \max_{\substack{0\le t_1\le\cdots\\ \le t_{m-1}\le 1}} \sum_{i=1}^{m-1} \tilde{B}^i(t_i)\right\}, \tag{3.43}$$

$m \geq 2$, and let $H_1 \equiv 0$ a.s., where $(\tilde{B}^1(t), \ldots, \tilde{B}^{m-1}(t))$ is an $(m-1)$-dimensional Brownian motion with tridiagonal covariance matrix given by (3.2)*. Let*

$$D_m = \max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{m-1}\le t_m=1}} \sum_{i=1}^{m} \left[B^i(t_i) - B^i(t_{i-1})\right],$$

where $(B^1(t), \ldots, B^m(t))_{0\le t\le 1}$ is a standard m-dimensional Brownian motion, related to $(\tilde{B}^1(t), \ldots, \tilde{B}^{m-1}(t))_{0\le t\le 1}$ by the almost sure identities

$$\tilde{B}^i(t) = \frac{1}{\sqrt{2}}(B^i(t) - B^{i+1}(t)), \qquad 1 \le i \le m-1.$$

Then $D_m = Z_m + H_m$ a.s., where Z_m is a centered normal random variable with variance $1/m$, which is given by $Z_m = (1/m)\sum_{i=1}^{m} B^i(1)$.

Combining the previous proposition with Proposition 3.1 and Theorem 3.2, the following is immediate:

Corollary 3.4. *For each $m \geq 1$, $\tilde{H}_m \stackrel{\mathcal{L}}{=} D_m - Z_m$, where $\mathcal{L}$ denotes equality in distribution.*

Finally, the above relationship between $\tilde{H}_m$ (resp., H_m) and D_m as well as Corollary 3.3 allows to further express the limiting distribution in a rather compact form.

Proposition 3.3. *Let $p_{max} = p_{j_1} = p_{j_2} = \cdots = p_{j_k}$, for $1 \le j_1 < j_2 < \cdots < j_k \le m$, and some $1 \le k \le m$, and let $p_i < p_{max}$, otherwise. Then*

$$\frac{LI_n - np_{max}}{\sqrt{n}} \Rightarrow \sqrt{p_{max}} H_k + \sqrt{p_{max}(1 - kp_{max})} Z_k. \tag{3.44}$$

Remark 3.4. One can also write the limiting law of Proposition 3.3 in terms of the functional D_k. Indeed, we have

$$\frac{LI_n - np_{max}}{\sqrt{np_{max}}} \Rightarrow D_k + (\sqrt{1 - kp_{max}} - 1)Z_k,$$

so that the limiting law is expressed as the sum of a centered normal random variable and of the maximal eigenvalue of a $k \times k$ element of the GUE.

The covariance structure of $X_t = \sqrt{(m-1)/m}\sum_{i=1}^{m}(\tilde{B}^i(t_i) - \tilde{B}^i(t_{i-1}))$, as given in (3.39), gives the L^2-distance between any X_t and X_s:

$$\mathbb{E}(X_t - X_s)^2 = 2\left[1 - \sum_{i=1}^{m} Leb(T_i \cap S_i)\right]. \tag{3.45}$$

In turn, (3.45) is useful, in applying entropy bounds, e.g., see Section 14 of Lifshits [21], to show that

$$\mathbb{E}\left(\max_{\substack{0=t_0\le t_1\le\cdots\\ \le t_{m-1}\le t_m=1}} X_t\right) \le K\sqrt{m-1},$$

for some absolute constant K. A lower bound of similar order ($\sqrt{m-1}$) can also be obtained by directly minimizing the above maximum.

In ways much deeper than the bounds just described, the behavior of D_m has been well-studied. In particular, it is known that $D_m/\sqrt{m} \to 2$ a.s. and in L^1, as $m \to \infty$ (see [3, 9, 10, 13, 24, 25, 29]), and that $(D_m - 2\sqrt{m})m^{1/6} \Rightarrow F_2$, as $m \to \infty$, where F_2 is the Tracy-Widom distribution (see [3, 11, 31, 30]). From these results, the decompositions of the previous section and some standard tools such as the Gaussian concentration inequality, asymptotics of H_m or $\tilde{H}_m$ follow.

Proposition 3.4. *We have that*

$$\frac{H_m}{\sqrt{m}} \to 2$$

a.s. and in L^1, as $m \to \infty$. Moreover,

$$\left(\frac{H_m}{\sqrt{m}} - 2\right) m^{2/3} \Rightarrow F_2, \tag{3.47}$$

where F_2 is the Tracy-Widom distribution. The same statements hold for $\tilde{H}_m$ in place of H_m.

Remark 3.5. (i) In the conclusion to [31], Tracy and Widom already derived (3.47) by applying a scaling argument to the limiting distribution of the uniform alphabet case. In our case we can moreover assert that a.s. and in the mean,

$$\lim_{k\to+\infty}\lim_{n\to+\infty}\frac{LI_n - np_{max}}{\sqrt{knp_{max}}} = 2,$$

and that

$$\left(\frac{LI_n - np_{max}}{\sqrt{knp_{max}}} - 2\right) k^{2/3} \Rightarrow F_2,$$

where the weak limit is first taken over n and then over k, and where p_{max} depends on k, and necessarily decreases to zero, as $k \to \infty$.

(ii) Using scaling, subadditivity, and concentration arguments found in Hambly, Martin, and O'Connell [13] and in O'Connell and Yor [24], one could prove directly that $\tilde{H}_m/\sqrt{m} \to 2$ a.s. This could be accomplished by studying, as do these authors, a process version of $\tilde{H}_m$.

4. Countable Infinite Alphabets

Let us now study the problem of describing LI_n for an ordered, countably infinite alphabet $\mathcal{A} = \{\alpha_n\}_{n\geq 1}$, where $\alpha_1 < \alpha_2 < \cdots < \alpha_m < \cdots$. Let $(X_i)_{i=1}^n$, $X_i \in \mathcal{A}$, be an iid sequence, with $\mathbb{P}(X_1 = \alpha_r) = p_r > 0$, for $r \geq 1$.

The central idea in the first part of our approach is to introduce two new sequences derived from $(X_i)_{i=1}^n$. Fix $m \geq 1$. The first sequence, which we shall term the *capped sequence*, is defined by taking $T_i^m = X_i \wedge \alpha_m$, for $i \geq 1$. The second one, $(Y_i^m)_{i=1}^{N_{n,m}}$, the *reduced sequence*, consists of the subsequence of $(X_i)_{i=1}^n$ of length $N_{n,m}$, for which $X_i \leq \alpha_m$, for $i \geq 1$. Thus, the capped sequence $(T_i^m)_{i=1}^n$ is obtained by setting to α_m all letter values greater than α_m, while the reduced sequence $(Y_i^m)_{i=1}^{N_{n,m}}$ is obtained by eliminating letter values greater than α_m altogether.

Let $LI_{n,m}^{cap}$ and $LI_{n,m}^{red}$ to be the lengths of the longest increasing subsequence of $(T_i^m)_{i=1}^n$ and $(Y_i^m)_{i=1}^{N_{n,m}}$, respectively. Now on the one hand, any subsequence of the reduced sequence is again a subsequence of the original sequence $(X_i)_{i=1}^n$. On the other hand, any increasing subsequence of $(X_i)_{i=1}^n$ is again an increasing subsequence of the capped one. These two observations lead to the pathwise bounds

$$LI_{n,m}^{red} \leq LI_n \leq LI_{n,m}^{cap}, \tag{4.1}$$

for all $m \geq 1$ and $n \geq 1$.

These bounds suggest that the behavior of the iid infinite case perhaps mirrors that of the iid finite-alphabet case. Indeed, we do have the following result, which amounts to an extension of Theorem 3.1 (or, more precisely, of Proposition 3.3) to the iid infinite-alphabet case.

Theorem 4.1. *Let $(X_i)_{i\geq 1}$ be a sequence of iid random variables taking values in the ordered alphabet $\mathcal{A} = \{\alpha_n\}_{n\geq 1}$. Let $\mathbb{P}(X_1 = \alpha_j) = p_j$, for $j \geq 1$. Let $p_{max} = p_{j_1} = p_{j_2} = \cdots = p_{j_k}$, $1 \leq j_1 < j_2 < \cdots < j_k$, $k \geq 1$, and let $p_i < p_{max}$, otherwise. Then*

$$\frac{LI_n - np_{max}}{\sqrt{n}} \Rightarrow \sqrt{p_{max}}H_k + \sqrt{p_{max}(1-kp_{max})}Z_k := R(p_{max}, k).$$

The proof of the theorem relies on an understanding of the limiting distributions of $LI_{n,m}^{red}$ and $LI_{n,m}^{cap}$. To this end, let us introduce some more notation. For a finite m-alphabet, and for $V_1, \ldots, V_n$ iid with $\mathbb{P}(V_1 = \alpha_r) = q_r > 0$, let $LI_n(q) := LI_n(q_1, \ldots, q_m)$ denote the length of the longest increasing subsequence of $(V_i)_{i=1}^n$. For each $m \geq 1$, let also $\pi_m = \sum_{r=1}^m p_r$.

First, let us choose m large enough so that $1 - \pi_{m-1} < p_{max}$. Next, observe that, from the capping at α_m, $LI_{n,m}^{cap}$ is distributed as $LI_n(\tilde{p})$, where $\tilde{p} = (p_1, \ldots, p_{m-1}, 1 - \pi_{m-1})$. But since m is chosen large enough, the maximal probability among the entries of $\tilde{p}$ is then p_{max}, of multiplicity k, as for the original infinite alphabet. By Theorem 3.1, we thus have

$$\frac{LI_n(\tilde{p}) - np_{max}}{\sqrt{n}} \Rightarrow R(p_{max}, k), \tag{4.2}$$

as $n \to \infty$.

Turning to $LI_{n,m}^{red}$, suppose that the number of elements $N_{n,m}$ of the reduced subsequence $(Y_i^m)_{i=1}^{N_{n,m}}$ is equal to j. Since only the elements of $(X_i)_{i=1}^n$ which are at most α_m are left, $LI_{n,m}^{red}$ must be distributed as $LI_j(\hat{p})$, where $\hat{p} = (p_1/\pi_m, \ldots, p_m/\pi_m)$. From the way m is chosen, the maximal probability among the entries of $\hat{p}$ is then p_{max}/π_m, of multiplicity k. Invoking again the finite-alphabet result of Theorem 3.1, we find that

$$\frac{LI_n(\hat{p}) - n(p_{max}/\pi_m)}{\sqrt{n}} \Rightarrow R\left(\frac{p_{max}}{\pi_m}, k\right), \tag{4.3}$$

as $n \to \infty$.

We now relate the two limiting expressions in (4.2) and (4.3) by the following elementary lemma.

Lemma 4.1. *Let $k \geq 1$ be an integer, and let $(q_m)_{m=1}^\infty$ be a sequence of reals in $(0, 1/k]$ converging to $q \geq 0$. Then $R(q_m, k) \Rightarrow R(q, k)$, as $m \to \infty$.*

Proof. Assume $q > 0$. Then

$$\begin{aligned} R(q_m, k) &= \sqrt{q_m}\left\{\sqrt{1 - q_m k} Z_k + H_k\right\} \\ &= \sqrt{q_m}\{\sqrt{1 - qk} Z_k + H_k\} \\ &\quad + \sqrt{q_m}\{\sqrt{1 - q_m k} - \sqrt{1 - qk}\} Z_k \\ &= \sqrt{\frac{q_m}{q}} R(q, k) + c_m Z_k, \end{aligned} \tag{4.4}$$

where $c_m = \sqrt{1 - q_m k} - \sqrt{1 - qk}$. Since $q_m \to q$ as $m \to \infty$, $c_m \to 0$, and so $c_m Z_k \Rightarrow 0$, as $m \to \infty$. This gives the result. The degenerate case, $q = 0$, is clear. ☐

The main idea developed in the proof of Theorem 4.1 is now to use the basic inequality (4.1) in conjunction with a conditioning argument for $LI_{n,m}^{red}$, in order to apply Lemma 4.1, *i.e.*, to use $R(p_{max}/\pi_m, k) \Rightarrow R(p_{max}, k)$, as $m \to \infty$, since $\pi_m \to 1$.

Proof of Theorem 4.1. First, fix an arbitrary $s > 0$. As previously noted in Remark 3.3, $R(p_{max}, k)$ is absolutely continuous, with density supported on $\mathbb{R}$ ($\mathbb{R}^+$ in the uniform case), and so s is a continuity point of its distribution function. Next, choose $0 < \epsilon_1 < 1$, and $0 < \delta < 1$, and again note that $(1 + \delta)s$ is also necessarily a continuity point for $R(p_{max}, k)$.

With this choice of ϵ_1, pick $\beta > 0$ such that $\mathbb{P}(Z \geq \beta) < \epsilon_1/2$, where Z is a standard normal random variable. Finally, pick ϵ_2 such that $0 < \epsilon_2 < \epsilon_1 \mathbb{P}(R(p_{max}, k) < (1 + \delta)s)$. Such a choice of ϵ_2 can always be made since the support of $R(p_{max}, k)$ includes $\mathbb{R}^+$.

We have seen that, for m large enough, we can bring some finite-alphabet results to bear on the infinite case. In fact, we need a few more technical requirements to complete our proof. Setting $\sigma_m^2 = \pi_m(1 - \pi_m)$, we choose large enough m so that:

(i) $1 - \pi_{m-1} < p_{max}$,

(ii) $(s + p_{max}\beta\sigma_m/\pi_m)/\sqrt{\pi_m - \beta\sigma_m} < (1 + \delta)s$, and

(iii) $|\mathbb{P}(R(p_{max}, k) < (1 + \delta)s) - \mathbb{P}(R(p_{max}/\pi_m, k) < (1 + \delta)s)| < \epsilon_2/2$.

The conditions (i) and (ii) are clearly satisfied, since $\pi_m \to 1$ and $\sigma_m \to 0$, as $m \to \infty$. The condition (iii) is also satisfied, as seen by applying Lemma 4.1 to $R(p_{max}/\pi_m, k)$, with $\pi_m \to 1$, and since $(1+\delta)s$ is also a continuity point for $R(p_{max}, k)$.

Now recall that $LI_{n,m}^{cap}$ is distributed as $LI_n(\tilde{p})$, where $\tilde{p} = (p_1, \dots, p_{m-1}, 1 - \pi_{m-1})$. Hence, we have from (4.1) and (4.2) that

$$\frac{LI_n - np_{max}}{\sqrt{n}} \le \frac{LI_{n,m}^{cap} - np_{max}}{\sqrt{n}} \Rightarrow R(p_{max}, k), \tag{4.5}$$

as $n \to \infty$, and so

$$\begin{aligned} \mathbb{P}\left(\frac{LI_n - np_{max}}{\sqrt{n}} \le s\right) &\ge \mathbb{P}\left(\frac{LI_{n,m}^{cap} - np_{max}}{\sqrt{n}} \le s\right) \\ &\to \mathbb{P}(R(p_{max}, k) \le s), \end{aligned} \tag{4.6}$$

as $n \to \infty$.

More work is required to make use of the left-hand minorization in (4.1) (*i.e.*, $LI_{n,m}^{red} \le LI_n$.) Recall that if the length $N_{n,m}$ of the reduced sequence is equal to j, then $LI_{n,m}^{red}$ must be distributed as $LI_j(\hat{p})$, where $\hat{p} = (p_1/\pi_m, \dots, p_m/\pi_m)$. Now the essential observation is that $N_{n,m}$ is distributed as a binomial random variable with parameters π_m and n. It is thus natural to focus on the values of j close to $\mathbb{E}N_{n,m} = n\pi_m$. Writing the variance of $N_{n,m}$ as $n\sigma_m^2$, where, as above, $\sigma_m^2 = \pi_m(1-\pi_m)$, and setting

$$\gamma_{n,m,j} := \mathbb{P}(N_{n,m} = j) = \binom{n}{j} \pi_m^j (1-\pi_m)^{n-j},$$

we have

$$\begin{aligned} &\mathbb{P}\left(\frac{LI_{n,m}^{red} - np_{max}}{\sqrt{n}} \le s\right) \\ &= \sum_{j=0}^{n} \mathbb{P}\left(\frac{LI_{n,m}^{red} - np_{max}}{\sqrt{n}} \le s | N_{n,m} = j\right) \gamma_{n,m,j} \\ &= \sum_{j=0}^{n} \mathbb{P}\left(\frac{LI_j(\hat{p}) - np_{max}}{\sqrt{n}} \le s\right) \gamma_{n,m,j} \\ &= \sum_{j=0}^{n} \mathbb{P}\left(\frac{LI_j(\hat{p}) - \frac{p_{max}}{\pi_m} j}{\sqrt{j}} \le \sqrt{\frac{n}{j}} \left(s + \frac{p_{max}}{\sqrt{n}}\left(n - \frac{j}{\pi_m}\right)\right)\right) \gamma_{n,m,j} \\ &\le \sum_{j=\lceil n\pi_m - \beta\sigma_m\sqrt{n}\rceil}^{n} \mathbb{P}\left(\frac{LI_j(\hat{p}) - \frac{p_{max}}{\pi_m} j}{\sqrt{j}} \le \sqrt{\frac{n}{j}} \left(s + \frac{p_{max}}{\sqrt{n}}\left(n - \frac{j}{\pi_m}\right)\right)\right) \gamma_{n,m,j} \\ &\qquad + \sum_{j=0}^{\lceil n\pi_m - \beta\sigma_m\sqrt{n}\rceil - 1} \gamma_{n,m,j} \\ &< \sum_{j=\lceil n\pi_m - \beta\sigma_m\sqrt{n}\rceil}^{n} \mathbb{P}\left(\frac{LI_j(\hat{p}) - \frac{p_{max}}{\pi_m} j}{\sqrt{j}} \le \sqrt{\frac{n}{j}} \left(s + \frac{p_{max}}{\sqrt{n}}\left(n - \frac{j}{\pi_m}\right)\right)\right) \gamma_{n,m,j} \\ &\qquad + \epsilon_1, \end{aligned} \tag{4.7}$$

for sufficiently large n, where (4.7) follows from the Central Limit Theorem and our choice of β, and where, as usual, $\lceil \cdot \rceil$ is the ceiling function.

Next, note that for $\lceil n\pi_m - \beta\sigma_m\sqrt{n}\rceil \le j \le n$, and making use of the condition (ii),

$$\begin{aligned}
&\sqrt{\frac{n}{j}}\left(s+\frac{p_{max}}{\sqrt{n}}\left(n-\frac{j}{\pi_m}\right)\right) \\
&\quad< \sqrt{\frac{n}{n\pi_m-\beta\sigma_m\sqrt{n}}}\left(s+\frac{p_{max}}{\sqrt{n}}\left(n-\frac{n\pi_m-\beta\sigma_m\sqrt{n}}{\pi_m}\right)\right) \\
&\quad= \frac{1}{\sqrt{\pi_m-\beta\sigma_m/\sqrt{n}}}\left(s+\frac{p_{max}\beta\sigma_m}{\pi_m}\right) \\
&\quad\le \frac{1}{\sqrt{\pi_m-\beta\sigma_m}}\left(s+\frac{p_{max}\beta\sigma_m}{\pi_m}\right) \\
&\quad< s(1+\delta). \qquad (4.8)
\end{aligned}$$

Hence, for sufficiently large n, we have

$$\begin{aligned}
&\sum_{j=\lceil n\pi_m-\beta\sigma_m\sqrt{n}\rceil}^{n} \mathbb{P}\left(\frac{LI_j(\hat{p})-\frac{p_{max}}{\pi_m}j}{\sqrt{j}} \le \sqrt{\frac{n}{j}}\left(s+\frac{p_{max}}{\sqrt{n}}\left(n-\frac{j}{\pi_m}\right)\right)\right)\gamma_{n,m,j} \\
&\qquad + \epsilon_1 \\
&\le \sum_{j=\lceil n\pi_m-\beta\sigma_m\sqrt{n}\rceil}^{n} \mathbb{P}\left(\frac{LI_j(\hat{p})-\frac{p_{max}}{\pi_m}j}{\sqrt{j}} \le s(1+\delta)\right)\gamma_{n,m,j}+\epsilon_1. \qquad (4.9)
\end{aligned}$$

Now from the condition (iii), and from the weak convergence, as $j \to \infty$, of $(LI_j(\hat{p}) - (p_{max}/\pi_m)j)/\sqrt{j}$ to $R(p_{max}/\pi_m, k)$, we find that, for j large enough,

$$\begin{aligned}
&\left|\mathbb{P}\left(\frac{LI_j(\hat{p})-\frac{p_{max}}{\pi_m}j}{\sqrt{j}} \le (1+\delta)s\right) - \mathbb{P}(R(p_{max},k) \le (1+\delta)s)\right| \\
&\le \left|\mathbb{P}\left(\frac{LI_j(\hat{p})-\frac{p_{max}}{\pi_m}j}{\sqrt{j}} \le (1+\delta)s\right) - \mathbb{P}\left(R\left(\frac{p_{max}}{\pi_m},k\right) \le (1+\delta)s\right)\right| \\
&\quad + \left|\mathbb{P}(R(p_{max},k) \le (1+\delta)s) - \mathbb{P}\left(R\left(\frac{p_{max}}{\pi_m},k\right) \le (1+\delta)s\right)\right| \\
&< \frac{\epsilon_2}{2}+\frac{\epsilon_2}{2} \\
&< \epsilon_1\mathbb{P}(R(p_{max},k) \le (1+\delta)s), \qquad (4.10)
\end{aligned}$$

and so,

$$\mathbb{P}\left(\frac{LI_j(\hat{p})-\frac{p_{max}}{\pi_m}j}{\sqrt{j}} \le (1+\delta)s\right) \le (1+\epsilon_1)\mathbb{P}\left(R(p_{max},k) \le (1+\delta)s\right). \qquad (4.11)$$

Now since $\lceil n\pi_m - \beta\sigma_m\sqrt{n}\rceil \to \infty$, as $n \to \infty$, with the help of (4.9) and (4.11), (4.7) becomes

$$\mathbb{P}\left(\frac{LI_{n,m}^{red} - np_{max}}{\sqrt{n}} \leq s\right)$$

$$\leq \sum_{j=\lceil n\pi_m - \beta\sigma_m\sqrt{n}\rceil}^{n} (1+\epsilon_1)\mathbb{P}\left(R(p_{max},k) \leq (1+\delta)s\right)\gamma_{n,m,j} + \epsilon_1$$

$$\leq (1+\epsilon_1)\mathbb{P}\left(R(p_{max},k) \leq (1+\delta)s\right) + \epsilon_1. \tag{4.12}$$

From (4.1) we know that $LI_{n,m}^{red} \leq LI_n$ a.s., and so

$$\mathbb{P}\left(\frac{LI_n - np_{max}}{\sqrt{n}} \leq s\right) \leq \mathbb{P}\left(\frac{LI_{n,m}^{red} - np_{max}}{\sqrt{n}} \leq s\right)$$

$$\leq (1+\epsilon_1)\mathbb{P}\left(R(p_{max},k) \leq (1+\delta)s\right) + \epsilon_1, \tag{4.13}$$

for large enough n. But since ϵ_1 and δ are arbitrary, (4.13) and (4.6) together show that

$$\mathbb{P}\left(\frac{LI_n - np_{max}}{\sqrt{n}} \leq s\right) \to \mathbb{P}(R(p_{max},k) \leq s), \tag{4.14}$$

for all $s > 0$.

The proof for $s < 0$ is similar. Indeed, since necessarily $p_{max} < 1/k$, $R(p_{max},k)$ describes the limiting distribution of the longest increasing subsequence for a non-uniform alphabet, and so is supported on $\mathbb{R}$. Then, one needs only to examine quantities of the form $\mathbb{P}(R(p_{max},k) \leq (1-\delta)s)$, instead of $\mathbb{P}(R(p_{max},k) \leq (1+\delta)s)$, as we have done throughout the proof for $s > 0$. These changes lead to the resulting statement. □

Remark 4.1. As an alternative to the above proof, one could certainly adopt the finite-alphabet development of the previous sections so as to express LI_n, for countable infinite alphabets, in terms of approximations to functionals of Brownian motion. More precisely,

$$\begin{aligned} LI_n &= \sup_{m\geq 2} \max_{\substack{0\leq k_1\leq\cdots \\ \leq k_{m-1}\leq n}} \left\{S_{k_1}^1 + S_{k_2}^2 + \cdots + S_{k_{m-1}}^{m-1} + a_n^m\right\} \\ &= \sup_{m\geq 2}\left\{\frac{n}{m} - \frac{1}{m}\sum_{r=1}^{m-1} rS_n^r + \max_{\substack{0\leq k_1\leq\cdots \\ \leq k_{m-1}\leq n}} \sum_{r=1}^{m-1} S_{k_r}^r\right\}, \end{aligned}$$

where a_n^m counts the number of occurrences of the letter α_m among $(X_i)_{1\leq i\leq n}$, and $S_k^r = \sum_{i=1}^{k} Z_i^r$ is the sum of independent random variables defined as in (2.3). After centering and normalizing the S_k^r, as was done to obtain (3.12) in the non-uniform finite alphabet development, one could then try to apply Donsker's Theorem to obtain a Brownian functional, which we now know to be distributed as $R(p_{max},k)$.

5. Concluding Remarks

Our development of the general finite-alphabet case leads us to consider several new directions in which to pursue this method and raises a number of interesting questions. These include the following.

- Extending our fixed finite-alphabet case to that of having each X_n take values in $\{1, 2, \ldots, m_n\}$ is an important first step. Fruitful approaches to such asymptotic questions would nicely close the circle of ideas initiated here. Such a study is already under consideration (see [15]).
- As we have noted throughout the paper, there is a pleasing if still rather mysterious connection between our limiting distribution results and those of random matrix theory. This connection deserves to be further explored. Recall, for instance, Baryshnikov's observation [3] that the process $(D_m)_{m\geq 1}$ is identical in law to the process $(\lambda_1^{(m)})_{m\geq 1}$ consisting of the largest eigenvalues of the $m \times m$ minors of an infinite GUE matrix. This fact is consistent with an interleaving-eigenvalue result from basic linear algebra, namely, that if $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of an $n\times n$ symmetric matrix A, and if $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n-1}$ are the eigenvalues of the matrix consisting of the first $(n-1)$ rows and columns of A, then $\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \cdots \geq \mu_{n-1} \geq \lambda_n$. We thus see the consistency between the $D_m \leq D_{m+1}$ a.s. fact noted above and that of $\lambda_1 \geq \mu_1$.
- Pursuing our analysis further, one might hope to find ways in which we can derive the densities of our limiting functionals in a direct manner. Its, Tracy, and Widom [17] have obtained clear expressions of the limiting distributions. While we have obtained our limiting distributions in a straightforward way, in turn, these densities do not clearly follow from our approach.
- In another direction, our independent-letter paradigm can be extended to various types of dependent cases, foremost of which would be the Markov case. This will be presented elsewhere [16], where the framework of [14] is, moreover, further extended.
- Various other types of subsequence problems can be tackled by the methodologies used in the present paper. To name but a few, comparisons for unimodal sequences, alternating sequences, and sequences with blocks will deserve further similar studies.

Acknowledgements

C. H. would like to thank Zhan Shi for discussions and encouragements on this project. Both authors would like to thank the organizers of the Special Program on High-Dimensional Inference and Random Matrices at SAMSI. Their hospitality and support, through the grant DMS-0112069, greatly facilitated the completion of this paper. Finally, many thanks to the referee for numerous comments.

References

[1] Aldous, D. and Diaconis, P. (1999). Longest increasing subsequences: From patience sorting to the Baik-Deift-Johansson theorem. *Bull. Amer. Math. Soc. (N.S.)* **36**(4) 413–432.

[2] Baik, J., Deift, P. and Johansson, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12**(4) 1119–1178.

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

[4] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. John Wiley and Sons, Inc.

[5] Cator, E. and Groeneboom, P. (2005). Hammersley's process with sources and sinks. *Ann. Probab.* **33** 879–903.
[6] Cator, E. and Groeneboom, P. (2006). Second class particles and cube root asymptotics for Hammersley's process. *Ann. Probab.* **34** 1273–1295.
[7] Chistyakov, G. P. and Götze, F. (2004). Distribution of the shape of Markovian random words. *Probab. Theory and Related Fields* **129** 18–36.
[8] Durrett, R. (1996). *Probability: Theory and Examples*, 2nd ed. Duxbury Press.
[9] Geman, S. (1980). A limit theorem for the norm of random matrices. *Ann. Probab.* **8** 252–261.
[10] Glynn, P. W. and Whitt, W. (1991). Departure from many queues in series. *Ann. Appl. Probab.* **1** 546–572.
[11] Gravner, J., Tracy, C. and Widom, H. (2001). Limit theorems for height fluctuations in a class of discrete space and time growth models. *J. Stat. Phys.* **102**(5,6) 1085–1132.
[12] Groeneboom, P. (2002). Hydrodynamical methods for analyzing longest increasing subsequences. *J. Comput. Appl. Math.* **142** 83–105.
[13] Hambly, B. M., Martin, J. B. and O'Connell, N. (2002). Concentration results for a Brownian directed percolation problem. *Stochastic Process. Appl.* **102** 207–220.
[14] Houdré, C., Lember, J. and Matzinger, H. (2006). On the longest common increasing binary subsequenc. *C.R. Acad. Sci. Paris, Ser. I* **343** 589–594.
[15] Houdré, C. and Litherland, T. (2009). Longest increasing subsequence asymptotics for growing alphabets. To appear.
[16] Houdré, C. and Litherland, T. (2008). On the limiting shape of random Young tableaux for Markovian words. Available at arXiv:0810.2982.
[17] Its, A. R., Tracy, C. and Widom, H. (2001). Random words, Toeplitz determinants, and integrable systems. I. Random matrix models and their applications. *Math. Sci. Res. Inst. Publ.* **40** 245–258. Cambridge Univ. Press, Cambridge.
[18] Its, A. R., Tracy, C. and Widom, H. (2001). Random words random words, Toeplitz determinants, and integrable systems. II. Advances in nonlinear mathematics and science. *Phys. D* 152–153, 199–224.
[19] Johansson, K. (2001). Discrete orthogonal polynomial ensembles and the plancherel measure. *Ann. of Math.* **153** 199–224.
[20] Kuperberg, G. (2002). Random words, quantum statistics, central limits, and random matrices. *Methods. Appl. Anal.* **9**(1) 99–118.
[21] Lifshits, M. A. (1995). *Gaussian Random Functions. Mathematics and Its Applications* **322**. Kluwer Academic Publishers, Dordrecht.
[22] Logan, B. F. and Shepp, L. A. (1977). A variational problem for random Young tableaux. *Adv. Math.* **26** 206–222.
[23] Mehta, M. L. (2004). *Random Matrices*, 3rd ed. Academic Press.
[24] O'Connell, N. and Yor, M. (2001). Brownian analogues of Burke's theorem. *Stochastic Process. Appl.* **96** 285–304.
[25] O'Connell, N. and Yor, M. (2002). A Representation for non-colliding random walks. *Electron. Comm. Probab.* **7** 1–12.
[26] Pilpel, S. (1990). Descending subsequences of random permutations. *J. Combin. Theory Ser. A* **53** 96–116.
[27] Pitman, J. W. (1975). One-dimensional Brownian motion and the three-dimensional Bessel process. *Adv. Appl. Prob.* **7** 511–526.
[28] Seppäläinen, T. (1996). A microscopic model for Burgers equation and the

longest increasing subsequence. *Electron. J. Probab.* **1**(5) 1–51.

[29] Seppäläinen, T. (1997). A scaling limit for queues in series. *Ann. Appl. Probab.* **7** 855–872.

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

[31] Tracy, C. and Widom, H. (2001). On the distributions of the lengths of the longest monotone subsequences in random words. *Probab. Theory and Related Fields* **119** 350–380.

[32] Vershik, A. and Kerov, S. (1977). Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables. *Soviet Math. Dokl.* **18** 527–531.