Bernoulli **21**(3), 2015, 1494–1537
DOI: 10.3150/14-BEJ612

Simultaneous large deviations for the shape of Young diagrams associated with random words

CHRISTIAN HOUDRÉ* and JINYONG MA**

Georgia Institute of Technology, School of Mathematics, Atlanta, GA 30332-0160, USA.
E-mail: *houdre@math.gatech.edu; **jma@math.gatech.edu

We investigate the large deviations of the shape of the random RSK Young diagrams associated with a random word of size n whose letters are independently drawn from an alphabet of size $m = m(n)$. When the letters are drawn uniformly and when both n and m converge together to infinity, m not growing too fast with respect to n, the large deviations of the shape of the Young diagrams are shown to be the same as that of the spectrum of the traceless GUE. In the non-uniform case, a control of both highest probabilities will ensure that the length of the top row of the diagram satisfies a large deviation principle. In either case, both speeds and rate functions are identified. To complete our study, non-asymptotic concentration bounds for the length of the top row of the diagrams, that is, for the length of the longest increasing subsequence of the random word are also given for both models.

Keywords: large deviations; longest increasing subsequence; random matrices; random words; strong approximation; Young diagrams

1. Introduction and results

Let $\mathcal{A}_m = \{\alpha_1 < \alpha_2 < \cdots < \alpha_m\}$ be an ordered alphabet of size m, and let a word be made of the random letters $X_1^m, X_2^m, \ldots, X_n^m$, independently drawn from $\mathcal{A}_m$. The Robinson–Schensted–Knuth (RSK) correspondence associates with this random word a pair of Young diagrams, of the same shape, having at most m rows. Now for $i = 1, 2, \ldots, m$, let $R_i(n, m)$ denote the length of the ith row of the Young diagrams, and recall that $R_1(n, m)$, the length of the top row, coincides with the length of the longest increasing subsequence of the random word $X_1^m X_2^m \cdots X_n^m$. Appropriately renormalized, and for uniform draws, the shape $(R_i(n, m))_{i=1}^m$ of the Young diagrams converges, in law and with n, to the spectrum of an $m \times m$ element of the traceless GUE ([20,31]). In turn, any fixed size subset of this spectrum, also converges with m, and after proper normalization, to a multidimensional Tracy–Widom distribution ([30,32]). These iterated convergence results have further led (see [10]) to the study of the limiting shape when both the word length and alphabet size simultaneously grow to infinity. This is briefly recalled next.

Let the random matrix $\mathbf{X} = (\mathbf{X}_{ij})_{1 \le i,j \le m}$ be an element of the $m \times m$ GUE with rescaling such that $\mathrm{Re}(\mathbf{X}_{ij}) \sim N(0, 1/2)$ and $\mathrm{Im}(\mathbf{X}_{ij}) \sim N(0, 1/2)$, for $i \ne j$; and $\mathbf{X}_{ii} \sim N(0, 1)$ (see [1] and [25] for background on random matrices). Let $(\lambda_1^m, \lambda_2^m, \ldots, \lambda_m^m)$ be the non-increasing ordered spectrum of $\mathbf{X}$, and let $(\lambda_1^{m,0}, \lambda_2^{m,0}, \ldots, \lambda_m^{m,0})$ be the corresponding ordered spectrum of an element of the traceless GUE (i.e., of $\mathbf{X} - \mathrm{tr}(\mathbf{X})/m$). An important fact (e.g., [5,6,14,15]) asserts that

$$\begin{aligned}&\big(\lambda_1^{m,0}, \lambda_2^{m,0}, \dots, \lambda_m^{m,0}\big)\\ &\quad \stackrel{\mathcal{L}}{=} \frac{\sqrt{m-1}}{\sqrt{m}} \mathbf{\Theta}_m^{-1}\left(\left(\max_{\mathbf{t}\in I_{k,m}} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \big(\tilde{B}^l_{t_{j,l}} - \tilde{B}^l_{t_{j,l-1}}\big)\right)_{1\le k\le m}\right),\end{aligned} \tag{1.1}$$

where $\mathbf{\Theta}_m : \mathbb{R}^m \to \mathbb{R}^m$ is defined via $(\mathbf{\Theta}_m(\mathbf{x}))_j = \sum_{i=1}^j x_i, 1 \le j \le m$, and where $(\tilde{B}^j_t)_{1\le j\le m, t\in[0,1]}$ is a driftless m-dimensional Brownian motion with covariance matrix

$$t\begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}, \tag{1.2}$$

with $\rho = -1/(m-1)$, and where for $1 \le k \le m$,

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

By comparing the Brownian functionals in (1.1) with discrete functionals representing the shape of the random Young diagrams, and via a KMT approximation, under some growth conditions on m, the simultaneous asymptotic convergence of the shapes is obtained in [10].

A related strategy is pursued here in order to investigate the large deviations of the shape of the RSK Young diagrams. More precisely, we obtain a large deviation principle (LDP) for the length of the first r rows of the Young diagrams, when n and m simultaneously converge to infinity and when the size m of the alphabet does not grow too fast. To achieve our goals, we also rely on the techniques and results developed in [7] (see also [3]), where large deviations are obtained for the largest (or the rth largest) eigenvalue of the GOE. These methodologies further give the multidimensional large deviations for the first r eigenvalues of the ordered spectrum of the traceless GUE. In turn, when combined with a KMT approximation, these lead to large deviations for the shape of the diagrams. Clearly, the results presented below complement the weak convergence ones of [10] and as any LDP results they allow to precisely quantify the deviation from the typical (limiting) shape of Young diagrams.

Let us next put our work into context. For random permutations, the large deviations of the length of the longest increasing subsequence are described in [13] and [29], while, moderate deviations are given in [23] and [24]. Closer to our framework, in [18], following the comparison method of [4] and [9], large deviations for the last-passage directed percolation model close to the x-axis are mainly established for i.i.d. weights which are Gaussian or have finite exponential moments. The study of the length of the top row of the diagrams also corresponds to a last-passage percolation problem but with *dependent* (exchangeable in the uniform case) Bernoulli weights (see (2.3)). For uniform draws, our framework also deals with the other rows of the diagrams.

Here is one of our main results:

Theorem 1.1. *In the uniform case, let m and n simultaneously converge to infinity in such a way that* $m(n) = \mathrm{o}(n^{1/4})$. *Then, for any* $r \geq 1$,

$$\left(\frac{R_1(n,m(n)) - n/m(n)}{\sqrt{n}}, \ldots, \frac{R_r(n,m(n)) - n/m(n)}{\sqrt{n}}\right)$$

satisfies a large deviation principle with speed $m(n)$ *and good rate function* I_r *on the space* $\mathcal{L}^r := \{(x_1, x_2, \ldots, x_r) \in \mathbb{R}^r : x_1 \geq x_2 \geq \cdots \geq x_r\}$, *where*

$$I_r(x_1, x_2, \ldots, x_r) = \begin{cases} 2\sum_{i=1}^{r} \int_2^{x_i} \sqrt{(z/2)^2 - 1}\,\mathrm{d}z, & \text{if } x_1 \geq x_2 \geq \cdots \geq x_r \geq 2, \\ +\infty, & \text{otherwise.} \end{cases} \tag{1.3}$$

In other words, for all $x_1 \geq x_2 \geq \cdots \geq x_r \geq 2$,

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\left(\frac{R_1(n,m(n)) - n/m(n)}{\sqrt{n}} \geq x_1, \ldots, \frac{R_r(n,m(n)) - n/m(n)}{\sqrt{n}} \geq x_r\right) \\ &\quad = -2\sum_{i=1}^{r} \int_2^{x_i} \sqrt{(z/2)^2 - 1}\,\mathrm{d}z, \end{aligned} \tag{1.4}$$

while for any $x < 2$ *and* $1 \leq i \leq r$,

$$\lim_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\left(\frac{R_i(n,m(n)) - n/m(n)}{\sqrt{n}} \leq x\right) = -\infty. \tag{1.5}$$

Remark 1.1. (i) The rate function I_r in (1.3) is a good rate function. Moreover, it is continuous and increasing with respect to each individual variable on its effective domain $\mathcal{D}_{I_r} = \{(x_1, x_2, \ldots, x_r) \in \mathbb{R}^r : x_1 \geq x_2 \geq \cdots \geq x_r \geq 2\}$, given that the other variables are fixed. Therefore, to prove a large deviation principle (LDP) as in Theorem 1.1, it is enough to prove a limiting equality on rectangular subsets as in (1.4) or (1.5) instead of proving both the usual upper and lower bounds, that is, that for any closed set F in $\mathcal{L}^r = \{(x_1, x_2, \ldots, x_r) \in \mathbb{R}^r : x_1 \geq x_2 \geq \cdots \geq x_r\}$,

$$\limsup_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\left(X_r^n \in F\right) \leq -\inf_F I_r, \tag{1.6}$$

and that for any open set O in $\mathcal{L}^r$,

$$\liminf_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\left(X_r^n \in O\right) \geq -\inf_O I_r, \tag{1.7}$$

where

$$X_r^n = \left(\frac{R_i(n,m(n)) - n/m(n)}{\sqrt{n}}\right)_{1\leq i\leq r}.$$

(ii) The restriction $m = \mathrm{o}(n^{1/4})$ (or $m = \mathrm{o}(n^{1/6})$ below) is a technical one and there is no reason to believe it is sharp. One can envision that our results continue to hold under at least a condition such as $m = \mathrm{o}(\sqrt{n})$.

In Theorem 1.1, if at least one of the renormalized variables is on the left of its simultaneous asymptotic mean, by changing the convergence speed from m to m^2, a more accurate form of (1.5) is valid. Below, the closed form expression obtained for K was found after Satya Majumdar kindly suggested that the methodology developed in [26] would apply to our traceless GUE framework.

Theorem 1.2. *In the uniform case, let m and n simultaneously converge to infinity in such a way that $m(n) = o(n^{1/6})$. Then, for any $r \geq 1$,*

$$\left(\frac{R_1(n, m(n)) - n/m(n)}{\sqrt{n}}, \ldots, \frac{R_r(n, m(n)) - n/m(n)}{\sqrt{n}}\right)$$

satisfies a large deviation principle with speed $(m(n))^2$ and good rate function $K(x_r)$ on the space $\mathcal{L}^r := \{(x_1, x_2, \ldots, x_r) \in \mathbb{R}^r : x_1 \geq x_2 \geq \cdots \geq x_r\}$, where K is the rate function in the large deviation principle for the largest eigenvalue of the $m \times m$ traceless GUE, when on the left of its asymptotic mean. It is given by

$$K(x) := \inf_{\mu \in \mathcal{M}_0((-\infty, x])} I(\mu), \tag{1.8}$$

where I (see (A.5)) is the rate function in the large deviation principle for the spectral measure of the GUE, and $\mathcal{M}_0((-\infty, x])$ is the set of zero mean probability measures supported on $(-\infty, x]$. For $x \leq 0$, $K(x) = +\infty$, for $x \geq 2$, $K(x) = 0$, and for $0 < x < 2$,

$$\begin{aligned} K(x) = \frac{1}{48}\Big(&3\big(9\sqrt[3]{2}3^{2/3}\big(\sqrt{81x^2+12}-9x\big)^{2/3}-8\big)x^2 \\ &+9\sqrt[3]{2}\sqrt[6]{3}\big(\sqrt{81x^2+12}-9x\big)^{1/3} \\ &\quad\times\big(\sqrt{27x^2+4}\big(\sqrt{81x^2+12}-9x\big)^{1/3}-5\sqrt[3]{2}\sqrt[6]{3}\big)x \\ &-6\sqrt[3]{2}\,3^{2/3}\big(\sqrt{81x^2+12}-9x\big)^{2/3} \\ &-3\,2^{2/3}3^{5/6}\sqrt{27x^2+4}\big(\sqrt{81x^2+12}-9x\big)^{1/3} \\ &+16\log\big(\sqrt{81x^2+12}-9x\big) \\ &-48\log\big(2\sqrt[3]{3}-\sqrt[3]{2}\big(\sqrt{81x^2+12}-9x\big)^{2/3}\big)+60+32\log 6\Big). \end{aligned} \tag{1.9}$$

In other words, for all $x_r \leq x_{r-1} \leq \cdots \leq x_1$, with $x_r \leq 2$,

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{(m(n))^2} \log \mathbb{P}\bigg(&\frac{R_1(n, m(n)) - n/m(n)}{\sqrt{n}} \leq x_1, \ldots, \\ &\frac{R_r(n, m(n)) - n/m(n)}{\sqrt{n}} \leq x_r\bigg) = -K(x_r), \end{aligned} \tag{1.10}$$

while for all $2 \le x_r \le x_{r-1} \le \cdots \le x_1$,

$$\lim_{n\to\infty} \frac{1}{(m(n))^2} \log \mathbb{P}\biggl(\frac{R_1(n,m(n)) - n/m(n)}{\sqrt{n}} \le x_1, \ldots, \frac{R_r(n,m(n)) - n/m(n)}{\sqrt{n}} \le x_r \biggr) = 0. \tag{1.11}$$

The LDP for the longest increasing subsequence is now a simple consequence.

Corollary 1.1. *Let* m *and* n *simultaneously converge to infinity in such a way that* $m(n) = \mathrm{o}(n^{1/4})$, *then for any* $x \ge 2$,

$$\lim_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\biggl(\frac{R_1(n,m(n)) - n/m(n)}{\sqrt{n}} \ge x \biggr) = -2\int_2^x \sqrt{(z/2)^2 - 1}\,\mathrm{d}z,$$

and similarly, if $m(n) = \mathrm{o}(n^{1/6})$, *for any* $x \le 2$,

$$\lim_{n\to\infty} \frac{1}{(m(n))^2} \log \mathbb{P}\biggl(\frac{R_1(n,m(n)) - n/m(n)}{\sqrt{n}} \le x \biggr) = -K(x).$$

Remark 1.2. The methodologies developed in this paper allow to derive LDPs in related problems. Such is the case for last-passage directed percolation close to the x-axis, or for the departure time from many queues in series when the number of customers is a fractional power of the number of servers. In these two problems, similar (discrete) functional representations are available but with i.i.d. weights, so the large deviations rate functions should be the corresponding rate functions in the LDP for the largest eigenvalue of the GUE.

When the independent random letters are no longer uniformly drawn, let X_i^m, $1 \le i \le n$, be independently and identically distributed with $\mathbb{P}(X_1^m = \alpha_j) = p_j^m$, $1 \le j \le m$. Moreover, let $p_{\max}^m = \max_{1\le j\le m} p_j^m$, let $p_{2\mathrm{nd}}^m = \max\{p_j^m < p_{\max}^m : 1 \le j \le m\}$, and let also $J(m) = \{j : p_j^m = p_{\max}^m\}$, where $k(m) = \operatorname{card}(J(m))$, that is, $k(m)$ is the multiplicity of $p_{\max}^m$.

Theorem 1.3. *In the non-uniform case, let* $k(m(n))$ *and* n *simultaneously converge to infinity in such a way that* $k(m(n))^3/p_{\max}^m = \mathrm{o}(n)$, *and let*

$$\frac{n(p_{2\mathrm{nd}}^m)^2}{k(m(n))p_{\max}^m} = \mathrm{o}\bigl(\exp\bigl(-k\bigl(m(n)\bigr)^\alpha\bigr)\bigr), \qquad \text{for some } \alpha > 1, \tag{1.12}$$

where $p_{2\mathrm{nd}}$ *is the second highest probability. Then,*

$$\frac{R_1(n,m(n)) - np_{\max}^m}{\sqrt{nk(m(n))p_{\max}^m}}$$

satisfies a LDP on $\mathbb{R}$ *with speed* $k(m(n))$ *and good rate function* I_1.

In other words, for any $x \geq 2$,

$$\lim_{n\to\infty} \frac{1}{k(m(n))} \log \mathbb{P}\left(\frac{R_1(n, m(n)) - np^m_{\max}}{\sqrt{nk(m(n))p^m_{\max}}} \geq x \right) = -2 \int_2^x \sqrt{(z/2)^2 - 1}\, \mathrm{d}z, \tag{1.13}$$

while for any $x < 2$,

$$\lim_{n\to\infty} \frac{1}{k(m(n))} \log \mathbb{P}\left(\frac{R_1(n, m(n)) - np^m_{\max}}{\sqrt{nk(m(n))p^m_{\max}}} \leq x \right) = -\infty. \tag{1.14}$$

Remark 1.3. (i) Above, the condition $k(m(n))^3/p^m_{\max} = \mathrm{o}(n)$ matches exactly the condition $m = \mathrm{o}(n^{1/4})$ of Theorem 1.1, but the new condition (1.12) is not present there. A similar remark applies to Theorem 1.2 and to Theorem 1.4.

(ii) In contrast to our first theorem, the one just stated is only for the first row of the diagrams and not for the whole shape. For non-uniform draws, a LDP shape result is also possible under conditions involving all the distinct probabilities and their respective multiplicity. These conditions are rather involved and the corresponding proofs rather tedious; therefore only a first row result is given above. A similar remark applies to Theorem 1.2 and to Theorem 1.4.

When the renormalized variable is on the left of its simultaneous asymptotic mean, again a more accurate form of (1.14) is possible. Before presenting this statement, let us first recall a few facts. For the alphabet $\mathcal{A}_m$ with corresponding set of probabilities $\mathcal{P} = \{p^m_1, p^m_2, \ldots, p^m_m\}$, let $p^{(1)} > p^{(2)} > \cdots > p^{(l)}$, $1 \leq l \leq m$, be the distinct elements in $\mathcal{P}$, and let $d_1, \ldots, d_l$ be the corresponding multiplicities, with $\sum_{i=1}^l d_i = m$. Then $p^{(1)} = p^m_{\max}$ and $d_1 = k(m)$ as in the previous notations. Let $\mathcal{G}_m(d_1, \ldots, d_l)$ be the set of $m \times m$ random matrices $\mathbf{X}$ which are direct sums of mutually independent elements of the $d_i \times d_i$ GUE, $1 \leq i \leq l$. Moreover, let $p_{(1)} \geq p_{(2)} \geq \cdots \geq p_{(m)}$ be the non-increasing rearrangement of $\mathcal{P}$. The "generalized" $m \times m$ traceless GUE associated with $\mathcal{P}$ is the set, denoted by $\mathcal{G}^0(p^m_1, p^m_2, \ldots, p^m_m)$, of $m \times m$ random matrices $\mathbf{X^0}$, of the form

$$\mathbf{X}^0_{i,j} = \begin{cases} \mathbf{X}_{i,i} - \sqrt{p_{(i)}} \sum_{h=1}^m \sqrt{p_{(h)}}\, \mathbf{X}_{h,h}, & \text{if } i = j, \\ \mathbf{X}_{i,j}, & \text{otherwise}, \end{cases} \tag{1.15}$$

where $\mathbf{X} \in \mathcal{G}_m(d_1, \ldots, d_l)$. Finally, let $\tilde{\lambda}^0_1$ be the largest eigenvalue of the diagonal block corresponding to $p^{(1)} = p^m_{\max}$ in $\mathbf{X^0}$.

Theorem 1.4. *Let* $k(m(n))$ *and* n *simultaneously converge to infinity in such a way that* $k(m(n))^5/p^m_{\max} = \mathrm{o}(n)$, *let also*

$$\frac{n(p^m_{\mathrm{2nd}})^2}{k(m(n))p^m_{\max}} = \mathrm{o}\bigl(\exp\bigl(-k\bigl(m(n)\bigr)^\alpha\bigr)\bigr), \qquad \textit{for some } \alpha > 2, \tag{1.16}$$

where $p_{2\mathrm{nd}}$ *is the second highest probability, and let,*

$$\lim_{n\to\infty} k\big(m(n)\big) p_{\max}^m = \eta, \tag{1.17}$$

for some $0 \le \eta \le 1$. *Then,*

$$\frac{R_1(n, m(n)) - np_{\max}^m}{\sqrt{nk(m(n))p_{\max}^m}}$$

satisfies a LDP on $\mathbb{R}$ *with speed* $(k(m(n)))^2$ *and good rate function* K_η, *where* K_η *is the rate function of* $\tilde{\lambda}_1^0$ *when on the left of its asymptotic mean.*

In other words, for any $x \le 2$,

$$\lim_{n\to\infty} \frac{1}{(k(m(n)))^2} \log \mathbb{P}\left(\frac{R_1(n, m(n)) - np_{\max}^m}{\sqrt{nk(m(n))p_{\max}^m}} \le x \right) = -K_\eta(x), \tag{1.18}$$

while for any $x \ge 2$,

$$\lim_{n\to\infty} \frac{1}{(k(m(n)))^2} \log \mathbb{P}\left(\frac{R_1(n, m(n)) - np_{\max}^m}{\sqrt{nk(m(n))p_{\max}^m}} \le x \right) = 0. \tag{1.19}$$

Remark 1.4. The rate function K_η is given by

$$K_\eta(x) = \sup_{y\le 0}\left(xy - yS(y) + J\big(S(y)\big) + \frac{\eta y^2}{2} \right),$$

where J is the rate function (with speed m^2) of the largest eigenvalue of the $m \times m$ GUE, and for each $y \le 0$, $S(y)$ is the unique solution to $J'(t) = y$ with $t \le 2$. For $x \ge 2$, $J(x) = 0$, while for $x \le 2$, the following closed form expression for J is obtained in [11],

$$\begin{aligned} J(x) = \frac{1}{216}\Big(& -x\big(-72x + x^3 + 30\sqrt{12 + x^2} + x^2\sqrt{12 + x^2}\big) \\ & - 216 \log\Big(\frac{1}{6}\big(x + \sqrt{12 + x^2}\big)\Big)\Big). \end{aligned} \tag{1.20}$$

In particular, $K_0 = J$ and $K_1 = K$. In fact the relationship between the spectrum of GUE and traceless GUE implies that

$$K(x) = \sup_{y\le 0}\left(xy - J^*(y) + \frac{y^2}{2} \right),$$

where $*$ denotes the Legendre transform. For any $0 \le \eta \le 1$, $K_\eta(x) = 0$, when $x \ge 2$, while for $0 \le \eta < 1$ and $x \in (-\infty, 2)$, $K_\eta(x) > 0$ and is asymptotically equivalent to

$$\frac{x^2}{2(1-\eta)} + \log\left(-\frac{x}{1-\eta} \right),$$

as $x \to -\infty$. For $\eta = 1$, when $0 < x < 2$, $K_1(x) = K(x)$ is positive and finite. As $x \to 0$, $K(x) \sim -\log x$, while as $x \to 2$, $K(x) \sim C(2-x)^3$, for some positive constant C.

To complement the previous results, we provide corresponding concentration results. These rely in part on deviations inequalities for the largest eigenvalue of the $m \times m$ GUE matrix, obtained respectively, in [2] and [21]. Comparing the forthcoming result with Corollary 1.1, we see that, in this case, the deviation rates match the fluctuation results. In turn these rates match the order of the tails of the Tracy–Widom distribution. Our results are only stated for the top row of the diagrams, but the techniques easily apply to the whole shape when combined with deviation inequalities for the whole spectrum of the GUE.

Theorem 1.5. *In the uniform model, let* $0 < \alpha < 1/4$, *and let* $m \le An^\alpha$, *for some* $A > 0$. *Then, for any* $0 < \varepsilon < 1$,

$$\mathbb{P}\left(\frac{R_1(n,m) - n/m}{\sqrt{n/m}} \ge 2\sqrt{m}(1+\varepsilon)\right) \le C(A,\alpha)\exp\left(-\frac{m\varepsilon^{3/2}}{C(A,\alpha)}\right), \tag{1.21}$$

where

$$C(A,\alpha) = C\max\left(A^{10/3},1\right)\frac{1+\alpha}{1-4\alpha}\exp\left(\frac{1+\alpha}{1-4\alpha}\right),$$

for some absolute constant $C > 0$.

Likewise, let $0 < \alpha < 1/6$, *and let* $m \le An^\alpha$, *for some* $A > 0$. *Then, for any* $0 < \varepsilon < 1$,

$$\mathbb{P}\left(\frac{R_1(n,m) - n/m}{\sqrt{n/m}} \le 2\sqrt{m}(1-\varepsilon)\right) \le \bar{C}(A,\alpha)\exp\left(-\frac{m^2\varepsilon^3}{\bar{C}(A,\alpha)}\right), \tag{1.22}$$

where

$$\bar{C}(A,\alpha) = C\max\left(A^4,1\right)\frac{1+\alpha}{1-6\alpha}\exp\left(\frac{1+\alpha}{1-6\alpha}\right),$$

for some absolute constant $C > 0$.

Again, in the non-uniform case, similar results hold but under a further control of the second highest probability.

Theorem 1.6. *In the non-uniform model, let* $\alpha > 3$, *let* $k(m(n))^\alpha / p^m_{\max} \le An$, *for some* $A > 0$ *and let*

$$\frac{n(p^m_{2\text{nd}})^2}{k(m(n))p^m_{\max}} \le B\exp\left(-k\left(m(n)\right)\right), \tag{1.23}$$

for some $B > 0$. *Then, for any* $0 < \varepsilon < 1$,

$$\mathbb{P}\left(\frac{R_1(n,m) - np^m_{\max}}{\sqrt{nk(m)p^m_{\max}}} \ge 2(1+\varepsilon)\right) \le C(A,B,\alpha)\exp\left(-\frac{k(m)\varepsilon^{3/2}}{C(A,B,\alpha)}\right), \tag{1.24}$$

where

$$C(A,B,\alpha)=C\max\big(A^{10/3\alpha},1\big)\max(\sqrt{B},1)\frac{\alpha+2}{\alpha-3}\exp\left(\frac{\alpha+2}{\alpha-3}\right),$$

for some absolute constant $C>0$.

Likewise, let $\alpha>5$, *let* $k(m(n))^\alpha/p^m_{\max}\le An$, *for some* $A>0$, *and let*

$$\frac{n(p^m_{2\text{nd}})^2}{k(m(n))p^m_{\max}}\le B\exp\big(-k\big(m(n)\big)^2\big), \tag{1.25}$$

for some $B>0$. *Then, for any* $0<\varepsilon<1$,

$$\mathbb{P}\left(\frac{R_1(n,m)-np^m_{\max}}{\sqrt{nk(m)p^m_{\max}}}\le 2(1-\varepsilon)\right)\le \bar{C}(A,B,\alpha)\exp\left(-\frac{k(m)^2\varepsilon^3}{\bar{C}(A,B,\alpha)}\right), \tag{1.26}$$

where

$$\bar{C}(A,B,\alpha)=C\max\big(A^{4/\alpha},1\big)\max(\sqrt{B},1)\frac{\alpha+2}{\alpha-5}\exp\left(\frac{\alpha+2}{\alpha-5}\right),$$

for some absolute constant $C>0$.

2. Proof of Theorem 1.1 and Theorem 1.2

As in [10], let

$$X^m_{i,j}=\begin{cases}1, & \text{if } X^m_i=\alpha_j,\\ 0, & \text{otherwise,}\end{cases} \tag{2.1}$$

be Bernoulli random variables with parameter $1/m$. For a fixed j, $1\le j\le m$, the $X^m_{i,j}$'s are i.i.d. while for $j\ne j'$, $(X^m_{1,j},\dots,X^m_{n,j})$ and $(X^m_{1,j'},\dots,X^m_{n,j'})$ are identically distributed but no longer independent.

Let $S^{m,j}_k=\sum_{i=1}^k X^m_{i,j}$ be the number of occurrences of α_j among $(X^m_i)_{1\le i\le k}$. Since for $1\le k<l\le n$, the number of occurrences of α_j among $(X^m_i)_{k+1\le i\le l}$ is $S^{m,j}_l-S^{m,j}_k$,

$$R_1(n,m)=\sup_{0=l_0\le l_1\le\cdots\le l_m=n}\sum_{j=1}^m\big(S^{m,j}_{l_j}-S^{m,j}_{l_{j-1}}\big),$$

with the convention that $S^{m,j}_0=0$.

Moreover, letting $V_k(n,m)=\sum_{i=1}^k R_i(n,m)$, combinatorial arguments yield (see Theorem 3.1 in [15])

$$V_k(n,m)=\sup_{\mathbf{t}\in I_{k,m}(n)}\sum_{j=1}^k\sum_{l=j}^{m-k+j}\big(S^{m,l}_{[t_{j,l}]}-S^{m,l}_{[t_{j,l-1}]}\big),\qquad 1\le k\le m, \tag{2.2}$$

where

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

Let $\tilde{X}^m_{i,j} = (X^m_{i,j} - 1/m)/\sigma_m$, with $\sigma^2_m = (1/m)(1-1/m)$, let $\tilde{S}^{m,j}_k = \sum_{i=1}^k \tilde{X}^m_{i,j}$. Similarly define $\tilde{V}_k(n,m)$, $1 \le k \le m$ and let $\tilde{R}_k(n,m) = \tilde{V}_k(n,m) - \tilde{V}_{k-1}(n,m)$, $2 \le k \le m$, while $\tilde{R}_1(n,m) = \tilde{V}_1(n,m)$. Clearly $V_k(n,m) = \sigma_m \tilde{V}_k(n,m) + kn/m$, and

$$\frac{R_k(n,m) - n/m}{\sqrt{n}} = \sqrt{1 - \frac{1}{m}} \frac{\tilde{R}_k(n,m)}{\sqrt{nm}}.$$

Let

$$\tilde{V}_k(n,m) = \sup_{\mathbf{t} \in I_{k,m}(n)} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \big(\tilde{S}^{m,l}_{[t_{j,l}]} - \tilde{S}^{m,l}_{[t_{j,l-1}]}\big), \qquad 1 \le k \le m, \tag{2.3}$$

with

$$\operatorname{Cov}\big(\tilde{S}^{m,i}_\ell, \tilde{S}^{m,j}_\ell\big) = \begin{cases} \ell, & \text{if } i = j, \\ \rho\ell, & \text{otherwise,} \end{cases} \tag{2.4}$$

and $\rho = -1/(m-1)$.

Next, $\tilde{V}_k(n,m)$ can be approximated by

$$\tilde{L}_k(n,m) = \sup_{\mathbf{t} \in I_{k,m}(n)} \sum_{j=1}^{k} \sum_{l=j}^{m-k+j} \big(\tilde{B}^l_{t_{j,l}} - \tilde{B}^l_{t_{j,l-1}}\big), \qquad 1 \le k \le m, \tag{2.5}$$

where $(\tilde{B}^j)_{1 \le j \le m}$ is a driftless m-dimensional Brownian motion with covariance matrix given in (1.2), and

$$\tilde{L}_k(n,m) \stackrel{\mathcal{L}}{=} \sqrt{n}\tilde{L}_k(1,m).$$

More precisely, inspired by [9],

$$\big|\tilde{V}_k(n,m) - \tilde{L}_k(n,m)\big| \le 2k \sum_{l=1}^{m} \big(Y^{m,l}_n + W^l_n\big), \tag{2.6}$$

where

$$Y^{m,l}_n = \max_{1 \le i \le n} \big|\tilde{S}^{m,l}_i - \tilde{B}^l_i\big| \quad \text{and} \quad W^l_n = \sup_{\substack{0 \le s,t \le n \\ |s-t| \le 1}} \big|\tilde{B}^l_s - \tilde{B}^l_t\big|.$$

Since

$$\big(\tilde{R}_k(n,m)\big)_{1\le k\le m} = \mathbf{\Theta}_m^{-1}\big(\big(\tilde{V}_k(n,m)\big)_{1\le k\le m}\big),$$

for any $\varepsilon > 0$, and from (2.6),

$$\begin{aligned}
&\mathbb{P}\big(\big|\tilde{R}_k(n,m) - \big(\tilde{L}_k(n,m) - \tilde{L}_{k-1}(n,m)\big)\big| \ge \sqrt{mn}\varepsilon\big)\\
&\le \mathbb{P}\left(2(2k-1)\sum_{l=1}^{m}\big(Y_n^{m,l} + W_n^l\big) \ge \sqrt{mn}\varepsilon\right)\\
&\le \mathbb{P}\left(\sum_{l=1}^{m} Y_n^{m,l} \ge \frac{\sqrt{mn}\varepsilon}{4(2k-1)}\right) + \mathbb{P}\left(\sum_{l=1}^{m} W_n^l \ge \frac{\sqrt{mn}\varepsilon}{4(2k-1)}\right) \qquad (2.7)\\
&\le \sum_{l=1}^{m}\left(\mathbb{P}\left(Y_n^{m,l} \ge \frac{\sqrt{mn}\varepsilon}{m(8k-4)}\right) + \mathbb{P}\left(W_n^l \ge \frac{\sqrt{mn}\varepsilon}{m(8k-4)}\right)\right)\\
&= m\mathbb{P}\left(Y_n^{m,1} \ge \frac{\sqrt{n}\varepsilon}{\sqrt{m}(8k-4)}\right) + m\mathbb{P}\left(W_n^1 \ge \frac{\sqrt{n}\varepsilon}{\sqrt{m}(8k-4)}\right),
\end{aligned}$$

for $1 \le k \le m$, and with the convention that $\tilde{L}_0(n,m) = 0$.

From Sakhanenko's version of the KMT inequality as stated, for example, in Theorem 2.1 and Corollary 3.2 of [22],

$$\mathbb{P}\left(Y_n^{m,1} \ge \frac{\sqrt{n}\varepsilon}{\sqrt{m}(8k-4)}\right) \le \big(1 + c_2(m)\sqrt{n}\big)\exp\left(-c_1(m)\frac{\sqrt{n}\varepsilon}{\sqrt{m}(8k-4)}\right), \qquad (2.8)$$

where, as $m \to +\infty$, $c_1(m) \sim C_1/\sqrt{m}$ and $c_2(m) \sim C_2/\sqrt{m}$, for some absolute constants $C_1 > 0$ and $C_2 > 0$. Moreover,

$$\begin{aligned}
\mathbb{P}\left(W_n^1 \ge \frac{\sqrt{n}\varepsilon}{\sqrt{m}(8k-4)}\right) &\le 2n\mathbb{P}\left(\big|\tilde{B}_2^1\big| \ge \frac{\sqrt{n}\varepsilon}{\sqrt{m}(16k-8)}\right)\\
&= 4n\mathbb{P}\left(\tilde{B}_2^1 \ge \frac{\sqrt{n}\varepsilon}{\sqrt{m}(16k-8)}\right) \qquad (2.9)\\
&\le 4\mathrm{e}n\exp\left(-\frac{n\varepsilon^2}{4\mathrm{e}m(16k-8)^2}\right).
\end{aligned}$$

Combining (2.8) and (2.9), letting $\varepsilon < 1$, and since $m(n) = \mathrm{o}(n^{1/4})$ (or simply, $m(n) = \mathrm{o}(\sqrt{n})$, to get a meaningful bound),

$$\mathbb{P}\big(\big|\tilde{R}_k(n,m) - \big(\tilde{L}_k(n,m) - \tilde{L}_{k-1}(n,m)\big)\big| \ge \sqrt{mn}\varepsilon\big) \le C_3\sqrt{mn}\exp\left(-\frac{\sqrt{n}\varepsilon}{C_3 m}\right), \qquad (2.10)$$

for $1 \le k \le r$, and where C_3 is a positive constant depending on k, which for fixed r can be chosen only depending on r. For any $x_1 \ge x_2 \ge \cdots \ge x_r > 2$, $r \ge 1$, and $0 < \varepsilon < \min(1, x_r - 2)$,

$$\begin{aligned} &\mathbb{P}\biggl(\frac{\tilde{R}_1(n,m)}{\sqrt{mn}} \ge x_1, \frac{\tilde{R}_2(n,m)}{\sqrt{mn}} \ge x_2, \ldots, \frac{\tilde{R}_r(n,m)}{\sqrt{mn}} \ge x_r\biggr) \\ &\quad\le \mathbb{P}\biggl(\frac{\tilde{L}_1(n,m) - \tilde{L}_0(n,m)}{\sqrt{mn}} \ge x_1 - \varepsilon, \ldots, \frac{\tilde{L}_r(n,m) - \tilde{L}_{r-1}(n,m)}{\sqrt{mn}} \ge x_r - \varepsilon\biggr) \\ &\qquad + \sum_{i=1}^{r} \mathbb{P}\biggl(\frac{\tilde{R}_i(n,m) - (\tilde{L}_i(n,m) - \tilde{L}_{i-1}(n,m))}{\sqrt{mn}} \ge \varepsilon\biggr) \end{aligned} \tag{2.11}$$

and

$$\begin{aligned} &\mathbb{P}\biggl(\frac{\tilde{R}_1(n,m)}{\sqrt{mn}} \ge x_1, \frac{\tilde{R}_2(n,m)}{\sqrt{mn}} \ge x_2, \ldots, \frac{\tilde{R}_r(n,m)}{\sqrt{mn}} \ge x_r\biggr) \\ &\quad\ge \mathbb{P}\biggl(\frac{\tilde{L}_1(n,m) - \tilde{L}_0(n,m)}{\sqrt{mn}} \ge x_1 + \varepsilon, \ldots, \frac{\tilde{L}_r(n,m) - \tilde{L}_{r-1}(n,m)}{\sqrt{mn}} \ge x_r + \varepsilon\biggr) \\ &\qquad - \sum_{i=1}^{r} \mathbb{P}\biggl(\frac{(\tilde{L}_i(n,m) - \tilde{L}_{i-1}(n,m)) - \tilde{R}_i(n,m)}{\sqrt{mn}} \ge \varepsilon\biggr), \end{aligned} \tag{2.12}$$

with again the convention that $\tilde{L}_0(n,m) = 0$.

Combining (1.1) with Theorem A.1 of the Appendix, when m and n simultaneously converge to infinity, the large deviations for $(\tilde{L}_k(n,m))_{1 \le k \le r}$ are then given by:

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\biggl(\frac{\tilde{L}_1(n,m(n))}{\sqrt{m(n)n}} \ge x_1, \ldots, \frac{\tilde{L}_r(n,m(n)) - \tilde{L}_{r-1}(n,m(n))}{\sqrt{m(n)n}} \ge x_r\biggr) \\ &\quad = -2\sum_{i=1}^{r} \int_2^{x_i} \sqrt{(z/2)^2 - 1}\, \mathrm{d}z, \end{aligned} \tag{2.13}$$

for all $x_1 \ge x_2 \ge \cdots \ge x_r > 2$, while for any $x < 2$ and $1 \le k \le r$,

$$\lim_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\biggl(\frac{\tilde{L}_k(n,m(n)) - \tilde{L}_{k-1}(n,m(n))}{\sqrt{m(n)n}} \le x\biggr) = -\infty. \tag{2.14}$$

This implies that,

$$\begin{aligned} &\mathbb{P}\biggl(\frac{\tilde{L}_1(n,m(n))}{\sqrt{m(n)n}} \ge x_1 \pm \varepsilon, \ldots, \frac{\tilde{L}_r(n,m(n)) - \tilde{L}_{r-1}(n,m(n))}{\sqrt{m(n)n}} \ge x_r \pm \varepsilon\biggr) \\ &\quad = \exp\bigl(-m(n)\bigl(I_r(x_1 \pm \varepsilon, \ldots, x_r \pm \varepsilon) + \mathrm{o}(1)\bigr)\bigr), \end{aligned}$$

where o(1) indicates a quantity converging to zero as n converges to infinity. Combining this fact with (2.10), for any $1 \le k \le r$,

$$\begin{aligned}
&\frac{\mathbb{P}(|\tilde{R}_k(n,m) - (\tilde{L}_k(n,m) - \tilde{L}_{k-1}(n,m))| \ge \sqrt{mn}\varepsilon)}{\mathbb{P}(\tilde{L}_1(n,m) \ge \sqrt{mn}(x_1 \pm \varepsilon), \ldots, \tilde{L}_r(n,m) - \tilde{L}_{r-1}(n,m) \ge \sqrt{mn}(x_r \pm \varepsilon))} \\
&\quad \le C_3\sqrt{mn}\exp\biggl(-\frac{\sqrt{n}\varepsilon}{C_3 m} + m\bigl(I_r(x_1 \pm \varepsilon, \ldots, x_r \pm \varepsilon) + \mathrm{o}(1)\bigr)\biggr) \qquad (2.15)\\
&\quad = C_3\sqrt{mn}\exp\biggl(\frac{\sqrt{n}}{m}\biggl(-\frac{\varepsilon}{C_3} + \frac{m^2}{\sqrt{n}}\bigl(I_r(x_1 \pm \varepsilon, \ldots, x_r \pm \varepsilon) + \mathrm{o}(1)\bigr)\biggr)\biggr) \longrightarrow 0,
\end{aligned}$$

as $m, n \to \infty$, $m = \mathrm{o}(n^{1/4})$. From (2.11) and (2.15), and since $m = \mathrm{o}(n^{1/4})$,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\biggl(\frac{\tilde{R}_1(n,m(n))}{\sqrt{m(n)n}} \ge x_1, \ldots, \frac{\tilde{R}_r(n,m(n))}{\sqrt{m(n)n}} \ge x_r\biggr) \\
&\quad \le \limsup_{n\to\infty} \frac{1}{m(n)} \log 2\mathbb{P}\biggl(\frac{\tilde{L}_1(n,m(n))}{\sqrt{m(n)n}} \ge x_1 - \varepsilon, \ldots, \\
&\qquad\qquad \frac{\tilde{L}_r(n,m(n)) - \tilde{L}_r(n,m(n))}{\sqrt{m(n)n}} \ge x_r - \varepsilon\biggr) \qquad (2.16)\\
&\quad = -I_r(x_1 - \varepsilon, \ldots, x_r - \varepsilon).
\end{aligned}$$

Likewise, from (2.12) and (2.15),

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\biggl(\frac{\tilde{R}_1(n,m(n))}{\sqrt{m(n)n}} \ge x_1, \ldots, \frac{\tilde{R}_r(n,m(n))}{\sqrt{m(n)n}} \ge x_r\biggr) \\
&\quad \ge \liminf_{n\to\infty} \frac{1}{m(n)} \log \frac{1}{2}\mathbb{P}\biggl(\frac{\tilde{L}_1(n,m(n))}{\sqrt{m(n)n}} \ge x_1 + \varepsilon, \ldots, \\
&\qquad\qquad \frac{\tilde{L}_r(n,m(n)) - \tilde{L}_r(n,m(n))}{\sqrt{m(n)n}} \ge x_r + \varepsilon\biggr) \qquad (2.17)\\
&\quad = -I_r(x_1 + \varepsilon, \ldots, x_r + \varepsilon).
\end{aligned}$$

Now letting $\varepsilon \to 0$,

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\biggl(\frac{\tilde{R}_1(n,m(n))}{\sqrt{m(n)n}} \ge x_1, \ldots, \frac{\tilde{R}_r(n,m(n))}{\sqrt{m(n)n}} \ge x_r\biggr) \\
&\quad = -2\sum_{i=1}^{r} \int_2^{x_i} \sqrt{(z/2)^2 - 1}\,\mathrm{d}z,
\end{aligned}$$

for any $x_1 \geq x_2 \geq \cdots \geq x_r > 2$. Next, assume that $x_1 \geq x_2 \geq \cdots \geq x_k > x_{k+1} = \cdots = x_r = 2$, $1 \leq k \leq r$, with the convention that $k = r$ corresponds to $x_1 \geq x_2 \geq \cdots \geq x_r > 2$. Under the conditions given in Theorem 1.1, for any $\varepsilon > 0$,

$$\liminf_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\left(\frac{\tilde{R}_1(n, m(n))}{\sqrt{m(n)n}} \geq x_1, \ldots, \frac{\tilde{R}_r(n, m(n))}{\sqrt{m(n)n}} \geq x_r \right) \geq -2 \sum_{i=1}^{k} \int_2^{x_i} \sqrt{(z/2)^2 - 1}\, \mathrm{d}z - 2 \sum_{i=k+1}^{r} \int_2^{2+\varepsilon} \sqrt{(z/2)^2 - 1}\, \mathrm{d}z.$$

Letting $\varepsilon \to 0$, gives

$$\liminf_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\left(\frac{\tilde{R}_1(n, m(n))}{\sqrt{m(n)n}} \geq x_1, \ldots, \frac{\tilde{R}_r(n, m(n))}{\sqrt{m(n)n}} \geq x_r \right) \geq -2 \sum_{i=1}^{r} \int_2^{x_i} \sqrt{(z/2)^2 - 1}\, \mathrm{d}z, \tag{2.18}$$

while,

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\left(\frac{\tilde{R}_1(n, m(n))}{\sqrt{m(n)n}} \geq x_1, \ldots, \frac{\tilde{R}_r(n, m(n))}{\sqrt{m(n)n}} \geq x_r \right) \\ &\leq \limsup_{n\to\infty} \frac{1}{m(n)} \log \mathbb{P}\left(\frac{\tilde{R}_1(n, m(n))}{\sqrt{m(n)n}} \geq x_1, \ldots, \frac{\tilde{R}_k(n, m(n))}{\sqrt{m(n)n}} \geq x_k \right) \\ &= -2 \sum_{i=1}^{r} \int_2^{x_i} \sqrt{(z/2)^2 - 1}\, \mathrm{d}z. \end{aligned} \tag{2.19}$$

Combining (2.18) and (2.19) proves (1.4).

Fix $x < 2$ and let $0 < \varepsilon < \min(1, 2 - x)$, then

$$\begin{aligned} \mathbb{P}\left(\frac{\tilde{R}_k(n, m)}{\sqrt{mn}} \leq x \right) &\leq \mathbb{P}\left(\frac{\tilde{L}_k(n, m) - \tilde{L}_{k-1}(n, m)}{\sqrt{mn}} \leq x + \varepsilon \right) \\ &\quad + \mathbb{P}\left(\frac{|\tilde{R}_k(n, m) - (\tilde{L}_k(n, m) - \tilde{L}_{k-1}(n, m))|}{\sqrt{mn}} \geq \varepsilon \right), \end{aligned} \tag{2.20}$$

for any $1 \leq k \leq r$. From (2.14),

$$\frac{1}{m} \log \mathbb{P}\left(\frac{\tilde{L}_k(n, m) - \tilde{L}_{k-1}(n, m)}{\sqrt{mn}} \leq x + \varepsilon \right) \longrightarrow -\infty,$$

and from (2.10),

$$\frac{1}{m}\log\mathbb{P}\left(\frac{|\tilde{R}_k(n,m)-(\tilde{L}_k(n,m)-\tilde{L}_{k-1}(n,m))|}{\sqrt{mn}}\geq\varepsilon\right)\leq\frac{\log(C_3\sqrt{mn})}{m}-\frac{\sqrt{n}\varepsilon}{C_3m^2}\longrightarrow-\infty, \tag{2.21}$$

as $m,n\to\infty$, $m=\mathrm{o}(n^{1/4})$. Thus for any $x<2$ and $1\leq k\leq r$,

$$\lim_{n\to\infty}\frac{1}{m(n)}\log\mathbb{P}\left(\frac{\tilde{R}_k(n,m(n))}{\sqrt{m(n)n}}\leq x\right)=-\infty, \tag{2.22}$$

which proves (1.5) in Theorem 1.1.

Proof of Theorem 1.2. First, (1.11) is just a direct consequence of (1.4). Next, we prove (1.10). Fix $y_1\geq y_2\geq\cdots\geq y_r$, with $y_r<2$. If $K(y_r)<+\infty$, then there exists $\delta>0$ such that $K(y_r-\delta)<+\infty$ and such that for any $0<\varepsilon<\min(1,\delta,2-y_r)$,

$$\begin{aligned}&\mathbb{P}\left(\frac{\tilde{R}_1(n,m)}{\sqrt{mn}}\leq y_1,\ldots,\frac{\tilde{R}_r(n,m)}{\sqrt{mn}}\leq y_r\right)\\&\quad\leq\mathbb{P}\left(\frac{\tilde{L}_1(n,m)-\tilde{L}_0(n,m)}{\sqrt{mn}}\leq y_1+\varepsilon,\ldots,\frac{\tilde{L}_r(n,m)-\tilde{L}_{r-1}(n,m)}{\sqrt{mn}}\leq y_r+\varepsilon\right)\\&\qquad+\sum_{i=1}^{r}\mathbb{P}\left(\frac{|\tilde{R}_i(n,m)-(\tilde{L}_i(n,m)-\tilde{L}_{i-1}(n,m))|}{\sqrt{mn}}\geq\varepsilon\right)\end{aligned} \tag{2.23}$$

and

$$\begin{aligned}&\mathbb{P}\left(\frac{\tilde{R}_1(n,m)}{\sqrt{mn}}\leq y_1,\ldots,\frac{\tilde{R}_r(n,m)}{\sqrt{mn}}\leq y_r\right)\\&\quad\geq\mathbb{P}\left(\frac{\tilde{L}_1(n,m)-\tilde{L}_0(n,m)}{\sqrt{mn}}\leq y_1-\varepsilon,\ldots,\frac{\tilde{L}_r(n,m)-\tilde{L}_{r-1}(n,m)}{\sqrt{mn}}\leq y_r-\varepsilon\right)\\&\qquad-\sum_{i=1}^{r}\mathbb{P}\left(\frac{|\tilde{R}_i(n,m)-(\tilde{L}_i(n,m)-\tilde{L}_{i-1}(n,m))|}{\sqrt{mn}}\geq\varepsilon\right),\end{aligned} \tag{2.24}$$

with once more the convention that $\tilde{L}_0(n,m)=0$.

Combining (1.1) with Corollary A.1, when m and n simultaneously converge to infinity,

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{m(n)^2}\log\mathbb{P}\left(\frac{\tilde{L}_1(n,m(n))}{\sqrt{m(n)n}}\leq y_1,\ldots,\frac{\tilde{L}_r(n,m(n))-\tilde{L}_{r-1}(n,m(n))}{\sqrt{m(n)n}}\leq y_r\right)\\&\quad=-K(y_r),\end{aligned} \tag{2.25}$$

for all $y_r \le y_{r-1} \le \cdots \le y_1$ with $y_r < 2$. Thus

$$\mathbb{P}\biggl(\frac{\tilde{L}_1(n,m(n))}{\sqrt{m(n)n}} \le y_1 \pm \varepsilon, \dots, \frac{\tilde{L}_r(n,m(n)) - \tilde{L}_{r-1}(n,m(n))}{\sqrt{m(n)n}} \le y_r \pm \varepsilon\biggr)$$
$$= \exp\bigl(-m(n)^2\bigl(K(y_r \pm \varepsilon) + \mathrm{o}(1)\bigr)\bigr),$$

where $\mathrm{o}(1)$ is meant for an expression converging to zero as n converges to infinity. Combining this last fact with (2.10), for any $1 \le k \le r$,

$$\begin{aligned}&\frac{\mathbb{P}(|\tilde{R}_k(n,m) - (\tilde{L}_k(n,m) - \tilde{L}_{k-1}(n,m))| \ge \sqrt{mn}\varepsilon)}{\mathbb{P}(\tilde{L}_1(n,m) \le \sqrt{mn}(y_1 \pm \varepsilon), \dots, \tilde{L}_r(n,m) - \tilde{L}_{r-1}(n,m) \le \sqrt{mn}(y_r \pm \varepsilon))}\\ &\quad\le C_3\sqrt{mn}\exp\biggl\{\frac{\sqrt{n}}{m}\biggl(-\frac{\varepsilon}{C_3} + \frac{m^3}{\sqrt{n}}\bigl(K(y_r \pm \varepsilon) + \mathrm{o}(1)\bigr)\biggr)\biggr\} \longrightarrow 0,\end{aligned} \tag{2.26}$$

as $m, n \to \infty$, $m = \mathrm{o}(n^{1/6})$. Repeating previous arguments, letting $\varepsilon \to 0$, and since $m = \mathrm{o}(n^{1/6})$,

$$\lim_{n\to\infty} \frac{1}{m(n)^2} \log \mathbb{P}\biggl(\frac{\tilde{R}_1(n,m(n))}{\sqrt{m(n)n}} \le y_1, \dots, \frac{\tilde{R}_r(n,m(n))}{\sqrt{m(n)n}} \le y_r\biggr) = -K(y_r), \tag{2.27}$$

for $y_r \le y_{r-1} \le \cdots \le y_1$, with $y_r < 2$ and $K(y_r) < +\infty$.

Now for fixed $y_1 \ge y_2 \ge \cdots \ge y_r$, $y_r < 2$, let us tackle the case $K(y_r) = +\infty$. Then,

$$\begin{aligned}&\mathbb{P}\biggl(\frac{\tilde{R}_1(n,m)}{\sqrt{mn}} \le y_1, \dots, \frac{\tilde{R}_r(n,m)}{\sqrt{mn}} \le y_r\biggr)\\ &\quad\le \mathbb{P}\biggl(\frac{\tilde{L}_r(n,m) - \tilde{L}_{r-1}(n,m)}{\sqrt{mn}} \le y_r + \varepsilon\biggr)\\ &\qquad+ \mathbb{P}\biggl(\frac{|\tilde{R}_r(n,m) - (\tilde{L}_r(n,m) - \tilde{L}_{r-1}(n,m))|}{\sqrt{mn}} \ge \varepsilon\biggr).\end{aligned} \tag{2.28}$$

As m and n simultaneously converge to infinity with $m = \mathrm{o}(n^{1/6})$, the second term on the right of (2.28) is exponentially negligible with speed m^2, that is,

$$\begin{aligned}&\frac{1}{m^2}\log \mathbb{P}\biggl(\frac{|\tilde{R}_k(n,m) - (\tilde{L}_k(n,m) - \tilde{L}_{k-1}(n,m))|}{\sqrt{mn}} \ge \varepsilon\biggr)\\ &\quad\le \frac{\log(C_3\sqrt{mn})}{m^2} - \frac{\sqrt{n}\varepsilon}{C_3 m^3} \longrightarrow -\infty,\end{aligned} \tag{2.29}$$

while the first term is, from (2.25), dominated by $\mathrm{e}^{-m(n)^2 K(y_r+\varepsilon)}$. Thus (2.27), in this case, follows by letting $\varepsilon \to 0$.

Now let $2 = y_r \le y_{r-1} \le \cdots \le y_1$, then for any $\varepsilon > 0$,

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{m(n)^2} \log \mathbb{P}\biggl(\frac{\tilde{R}_1(n,m(n))}{\sqrt{m(n)n}} \le y_1, \dots, \frac{\tilde{R}_r(n,m(n))}{\sqrt{m(n)n}} \le y_r\biggr) \\
&\quad \ge \liminf_{n\to\infty} \frac{1}{m(n)^2} \log \mathbb{P}\biggl(\frac{\tilde{R}_1(n,m(n))}{\sqrt{m(n)n}} \le y_1, \dots, \frac{\tilde{R}_r(n,m(n))}{\sqrt{m(n)n}} \le 2-\varepsilon\biggr) \\
&\quad = -K(2-\varepsilon).
\end{aligned} \tag{2.30}$$

Again, letting $\varepsilon \to 0$, and since K is continuous (see the Appendix for a proof),

$$\liminf_{n\to\infty} \frac{1}{m(n)^2} \log \mathbb{P}\biggl(\frac{\tilde{R}_1(n,m(n))}{\sqrt{m(n)n}} \le y_1, \dots, \frac{\tilde{R}_r(n,m(n))}{\sqrt{m(n)n}} \le y_r\biggr) \ge -K(2) = 0.$$

Clearly,

$$\limsup_{n\to\infty} \frac{1}{m(n)^2} \log \mathbb{P}\biggl(\frac{\tilde{R}_1(n,m(n))}{\sqrt{m(n)n}} \le y_1, \dots, \frac{\tilde{R}_r(n,m(n))}{\sqrt{m(n)n}} \le y_r\biggr) \le 0,$$

which proves the case $y_r = 2$, and finishes the proof of the first part of Theorem 1.2. Lemma A.1 of the Appendix gives a proof of (1.8).

When $x \le 0$, $\mathcal{M}_0((-\infty, x])$ is empty so $K(x) = +\infty$ and when $x \ge 2$, the semicircular probability measure belongs to $\mathcal{M}_0((-\infty, x])$, thus $K(x) = 0$. When $0 < x < 2$, the closed form expression of K given via (1.9) can indeed be derived using the techniques developed in [26]. Denote by μ_0 the zero mean probability measure supported on $(-\infty, x]$, minimizing

$$I(\mu) = \frac{1}{2}\int y^2 \mu(\mathrm{d}y) - \iint \log|t-y| \mu(\mathrm{d}t)\mu(\mathrm{d}y) - \frac{3}{4} \tag{2.31}$$

(the existence and uniqueness of μ_0 follows from Theorem 1.3 of Chapter 1 of [28]. Moreover, in view of Theorem 2.5 of Chapter IV of [28], μ_0 is absolutely continuous with continuous density ρ_0, while from Theorem 1.10 and Theorem 1.11 of Chapter IV in [28], its support is a finite interval). Let us now proceed to explicitly find ρ_0. To do so, consider the Lagrange function

$$E(\mu) = I(\mu) + c_1\biggl(\int \mu(\mathrm{d}y) - 1\biggr) + c_2 \int y\mu(\mathrm{d}y),$$

where the Lagrange multipliers c_1 and c_2 correspond to the constraints that μ is a zero mean probability measure. Let $[L', x]$ be the support of ρ_0, and for any continuous function h supported on $[L', x]$ such that $h(y) \ge -\rho_0(y)$, let

$$\begin{aligned}
E(\rho_0 + \varepsilon h) &= \frac{1}{2}\int y^2\bigl(\rho_0(y) + \varepsilon h(y)\bigr)\,\mathrm{d}y \\
&\quad - \iint \log|t-y|\bigl(\rho_0(t) + \varepsilon h(t)\bigr)\bigl(\rho_0(y) + \varepsilon h(y)\bigr)\,\mathrm{d}t\,\mathrm{d}y - \frac{3}{4} \\
&\quad + c_1\biggl(\int \bigl(\rho_0(y) + \varepsilon h(y)\bigr)\,\mathrm{d}y - 1\biggr) + c_2 \int y\bigl(\rho_0(y) + \varepsilon h(y)\bigr)\,\mathrm{d}y.
\end{aligned} \tag{2.32}$$

Thus

$$\left.\frac{\mathrm{d}E(\rho_0+\varepsilon h)}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0}=0,$$

gives

$$\int\left(\frac{y^2}{2}-2\int\log|t-y|\rho_0(t)\,\mathrm{d}t+c_1+c_2y\right)h(y)\,\mathrm{d}y=0, \tag{2.33}$$

for any continuous function h such that $h(y)\geq-\rho_0(y)$. Let

$$g(y)=\frac{y^2}{2}-2\int\log|t-y|\rho_0(t)\,\mathrm{d}t+c_1+c_2y,$$

which is a continuous function on $[L',x]$. Let $h(y)=g^+(y)$, then (2.33) yields

$$\int_{g(y)\geq 0}g(y)^2\,\mathrm{d}y=0,$$

thus $g(y)\leq 0$ for $y\in[L',x]$. Likewise, letting

$$h(y)=\begin{cases}0, & \text{if } g(y)>0,\\ g(y), & \text{if } -\rho_0(y)\leq g(y)\leq 0,\\ -\rho_0(y), & \text{if } g(y)<-\rho_0(y),\end{cases} \tag{2.34}$$

then (2.33) yields $g(y)\geq 0$ for $y\in[L',x]$. Thus,

$$\frac{y^2}{2}-2\int\log|t-y|\rho_0(t)\,\mathrm{d}t+c_1+c_2y=0, \tag{2.35}$$

for any $y\in[L',x]$. In turn, differentiating (2.35) with respect to y further gives,

$$y-2\,\text{p.v.}\int\frac{\rho_0(t)}{y-t}\,\mathrm{d}t+c_2=0, \tag{2.36}$$

where p.v. denotes the Cauchy principal value.

Let $L=L'-x$ and $f_x(t)=\rho_0(t+x)$ be supported on $[L,0]$, then the finite Hilbert transform

$$\frac{1}{\pi}\,\text{p.v.}\int_{L'}^{x}\frac{\rho_0(t)}{y-t}\,\mathrm{d}t=\frac{y+c_2}{2\pi},$$

becomes

$$\frac{1}{\pi}\,\text{p.v.}\int_{L}^{0}\frac{f_x(t)}{y-t}\,\mathrm{d}t=\frac{x+y+c_2}{2\pi},$$

for any $y\in[L,0]$. From Section 4.3 of [33], this finite Hilbert transform can be inverted as

$$f_x(y)=\frac{1}{\pi\sqrt{(y-L)(-y)}}\left(\text{p.v.}\int_L^0\frac{\sqrt{(t-L)(-t)}}{t-y}\frac{x+t+c_2}{2\pi}\,\mathrm{d}t+c_3\right), \tag{2.37}$$

where $L \le y \le 0$. Moreover,

$$\begin{aligned} &\text{p.v.} \int_L^0 \frac{\sqrt{(t-L)(-t)}}{t-y} \frac{x+t+c_2}{2\pi} \, \mathrm{d}t \\ &= \frac{1}{16}\big(4c_2(L-2y)+L^2+4L(x+y)-8y(x+y)\big). \end{aligned} \tag{2.38}$$

Since $f_x(L)=0$,

$$c_3 = \tfrac{1}{16}\big(4L(c_2+x)+3L^2\big),$$

which when plugged into (2.37) yields

$$f_x(y) = \frac{\sqrt{y(L-y)}(2c_2+L+2(x+y))}{4\pi y}.$$

Now, the two constraints $\int \mathrm{d}\mu_0(y)=1$ and $\int y\,\mathrm{d}\mu_0(y)=0$, yield

$$\int_L^0 y f_x(y)\,\mathrm{d}y + x = 0, \qquad \int_L^0 f_x(y)\,\mathrm{d}y = 1,$$

leading to

$$L = \frac{2\,2^{2/3}(\sqrt{81x^2+12}-9x)^{2/3} - 4\,6^{1/3}}{3^{2/3}(\sqrt{81x^2+12}-9x)^{1/3}} \tag{2.39}$$

and

$$\begin{aligned} c_2 &= \frac{2\,3^{2/3} - \sqrt[3]{6}(\sqrt{81x^2+12}-9x)^{2/3}}{2^{2/3}(\sqrt{81x^2+12}-9x)^{1/3}} \\ &\quad - \frac{\sqrt[3]{2}\,3^{2/3}(\sqrt{81x^2+12}-9x)^{2/3} + (6\,2^{2/3}\sqrt[3]{3}/(\sqrt{81x^2+12}-9x)^{2/3}) + 6}{18x} - x. \end{aligned} \tag{2.40}$$

Integrate (2.35) with respect to μ_0 to get

$$\iint \log|y-t|\mu_0(\mathrm{d}t)\mu_0(\mathrm{d}y) = \frac{1}{4}\int y^2\mu_0(\mathrm{d}y) + \frac{c_1}{2},$$

while c_1 is determined by substituting $y=x$ in (2.35),

$$c_1 = -\frac{x^2}{2} + 2\int \log|x-t|\mu_0(\mathrm{d}t) - c_2 x.$$

Finally,

$$\begin{aligned} I(\mu_0) &= \frac{1}{2}\int y^2\mu_0(\mathrm{d}y) - \iint \log|t-y|\mu_0(\mathrm{d}t)\mu_0(\mathrm{d}y) - \frac{3}{4} \\ &= \frac{1}{4}\int_L^0 (x+y)^2 f_x(y)\,\mathrm{d}y - \int_L^0 \log(-y) f_x(y)\,\mathrm{d}y + \frac{x^2}{4} + \frac{c_2 x}{2} - \frac{3}{4}. \end{aligned} \tag{2.41}$$

Finally, inserting the above values of L and c_2 into (2.41) provides the closed form expression for K. □

3. Proof of Theorem 1.3 and Theorem 1.4

Recall, see (2.2), that

$$R_1(n,m) = V_1(n,m) = \sup_{0=l_0\le l_1\le\cdots\le l_m=n} \sum_{j=1}^{m} \left(S_{l_j}^{m,j} - S_{l_{j-1}}^{m,j}\right)$$

and let

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

where $J(m) = \{j : p_j^m = p_{\max}^m\}$ as defined before. Then, Lemma 9 in [10] asserts that:

$$\mathbb{E}\left|V_1(n,m) - V_1'(n,m)\right| \le Cnp_{2\text{nd}}^m, \tag{3.1}$$

where $p_{2\text{nd}}$ is the second highest probability and $C > 0$ some absolute constant.

In order to prove Theorem 1.3, we first need a lemma.

Lemma 3.1. *Let $k(m(n))$ converge to infinity with n in such a way that $k(m(n))^3/p_{\max}^m = \mathrm{o}(n)$, then for any $x \ge 2$,*

$$\lim_{n\to\infty} \frac{1}{k(m(n))} \log \mathbb{P}\left(\frac{V_1'(n,m(n)) - np_{\max}^m}{\sqrt{nk(m(n))p_{\max}^m}} \ge x \right) = -2\int_2^x \sqrt{(z/2)^2 - 1}\,\mathrm{d}z, \tag{3.2}$$

and for any $x < 2$,

$$\lim_{n\to\infty} \frac{1}{k(m(n))} \log \mathbb{P}\left(\frac{V_1'(n,m(n)) - np_{\max}^m}{\sqrt{nk(m(n))p_{\max}^m}} \le x \right) = -\infty. \tag{3.3}$$

Proof. As in the proof of Theorem 1.1, for any $j \in J(m)$, set $\tilde{X}_{i,j}^m = (X_{i,j}^m - p_{\max}^m)/\sigma_m$, where $\sigma_m^2 = p_{\max}^m(1 - p_{\max}^m)$, and set $\tilde{S}_\ell^{m,j} = \sum_{i=1}^{\ell} \tilde{X}_{i,j}^m$. Hence,

$$\frac{V_1'(n,m) - np_{\max}^m}{\sqrt{nk(m(n))p_{\max}^m}} = \left(\sqrt{1 - p_{\max}^m}\right) \frac{\tilde{V}_1'(n,m)}{\sqrt{nk(m(n))}},$$

with the obvious notation for $\tilde{V}_1'(n,m)$. Since $k(m(n))p_{\max}^m \le 1$, as $n \to \infty$, $p_{\max}^m \to 0$, so (3.2) can be reduced to,

$$\lim_{n\to\infty} \frac{1}{k(m(n))} \log \mathbb{P}\left(\frac{\tilde{V}_1'(n,m(n))}{\sqrt{nk(m(n))}} \ge x \right) = -I_1(x), \tag{3.4}$$

for any $x \geq 2$. Moreover, (3.3) can be reduced to,

$$\lim_{n\to\infty} \frac{1}{k(m(n))} \log \mathbb{P}\left(\frac{\tilde{V}_1'(n, m(n))}{\sqrt{nk(m(n))}} \leq x \right) = -\infty, \tag{3.5}$$

for any $x < 2$. Now,

$$\tilde{V}_1'(n, m) = \sup_{\substack{0=l_0\leq l_1\leq\cdots\leq l_m=n \\ l_{j-1}=l_j \text{ for } j\notin J(m)}} \sum_{j=1}^{m} \left(\tilde{S}_{l_j}^{m,j} - \tilde{S}_{l_{j-1}}^{m,j} \right), \tag{3.6}$$

where

$$\operatorname{Cov}\left(\tilde{S}_\ell^{m,i}, \tilde{S}_\ell^{m,j} \right) = \begin{cases} \ell, & \text{if } i = j, \\ \rho_1 \ell, & \text{otherwise,} \end{cases} \tag{3.7}$$

where $\rho_1 = -p_{\max}^m/(1 - p_{\max}^m)$. From its very definition, $\tilde{V}_1'(n, m)$ only depends on $(\tilde{S}_\ell^{m,j})_{j\in J(m)}$ and can thus be approximated, via KMT, by the Brownian functional $F(n, k)$ with $k = \operatorname{card}(J(m))$ (from here onward, m is short for $m(n)$ and k is short for $k(m(n))$), where

$$F(n, k) = \sup_{0=t_0\leq t_1\leq\cdots\leq t_k=n} \sum_{r=1}^{k} \left(\tilde{B}_{t_r}^{(r)} - \tilde{B}_{t_{r-1}}^{(r)} \right), \tag{3.8}$$

where $(\tilde{B}^{(r)})_{1\leq r\leq k}$ is a centered k-dimensional Brownian motion with covariance matrix

$$\begin{pmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_1 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1 & \rho_1 & \cdots & 1 \end{pmatrix} t.$$

Moreover,

$$F(n, k) \stackrel{\mathcal{L}}{=} \sqrt{n} F(1, k), \tag{3.9}$$

while from Corollary 3.2 and Corollary 3.3 in [16],

$$\begin{aligned} \sqrt{1 - p_{\max}^m} F(1, k) &\stackrel{\mathcal{L}}{=} \frac{\sqrt{1 - k p_{\max}^m} - 1}{k} \sum_{j=1}^{k} B_1^{(j)} \\ &\quad + \sup_{0=t_0\leq t_1\leq\cdots\leq t_k=1} \sum_{r=1}^{k} \left(B_{t_r}^{(r)} - B_{t_{r-1}}^{(r)} \right), \end{aligned} \tag{3.10}$$

where $(B^{(j)})_{1\leq j\leq k}$ is a standard k-dimensional Brownian motion on $[0, 1]$. The first weighted sum in (3.10) is a Gaussian random variable with variance at most $1/k$, and as well known (see the introductory section and the cited references therein):

$$\sup_{0=t_0\le t_1\le\cdots\le t_k=1}\sum_{r=1}^{k}\big(B_{t_r}^{(r)}-B_{t_{r-1}}^{(r)}\big)\stackrel{\mathcal{L}}{=}\lambda_1^k, \tag{3.11}$$

where λ_1^k is the largest eigenvalue of a $k\times k$ element of the GUE. Next, since $\lambda_1^k/\sqrt{k}$ satisfies a LDP with rate function I_1 and since the contribution of the Gaussian term is negligible, we get via Theorem A.1 of the Appendix:

$$\lim_{k\to\infty}\frac{1}{k}\log\mathbb{P}\big(F(1,k)\ge\sqrt{k}x\big)=-I_1(x). \tag{3.12}$$

Now, as in the proof of Theorem 1.1,

$$\begin{aligned}&\mathbb{P}\big(\big|\tilde{V}_1'(n,m)-F(n,k)\big|\ge\sqrt{nk}\varepsilon\big)\\&\quad\le k\mathbb{P}\biggl(Y_n^{m,l}\ge\frac{\sqrt{n}\varepsilon}{4\sqrt{k}}\biggr)+k\mathbb{P}\biggl(W_n^l\ge\frac{\sqrt{n}\varepsilon}{4\sqrt{k}}\biggr),\end{aligned} \tag{3.13}$$

where l is any element of $J(m)$ and where

$$Y_n^{m,l}=\max_{1\le i\le n}\big|\tilde{S}_i^{m,l}-\tilde{B}_i^{(l)}\big|\quad\text{and}\quad W_n^l=\sup_{\substack{0\le s,t\le n\\|s-t|\le 1}}\big|\tilde{B}_s^{(l)}-\tilde{B}_t^{(l)}\big|.$$

As in getting (2.8),

$$\mathbb{P}\biggl(Y_n^{m,1}\ge\frac{\sqrt{n}\varepsilon}{4\sqrt{k}}\biggr)\le\big(1+c_2\big(p_{\max}^m\big)\sqrt{n}\big)\exp\biggl(-c_1\big(p_{\max}^m\big)\frac{\sqrt{n}\varepsilon}{4\sqrt{k}}\biggr), \tag{3.14}$$

where $c_1(p_{\max}^m)\sim C_1\sqrt{p_{\max}^m}$ and $c_2(p_{\max}^m)\sim C_2\sqrt{p_{\max}^m}$, for some constants C_1 and C_2, while from (2.9),

$$\mathbb{P}\biggl(W_n^1\ge\frac{\sqrt{n}\varepsilon}{4\sqrt{k}}\biggr)\le C_3n\exp\biggl(-\frac{n\varepsilon^2}{C_3k}\biggr), \tag{3.15}$$

for some positive constant C_3. Combining (3.14) and (3.15), letting $\varepsilon<1$, and since $k(m(n))^3/p_{\max}^m=\mathrm{o}(n)$ (or simply, $k(m(n))/p_{\max}^m=\mathrm{o}(n)$, to get a meaningful bound),

$$\mathbb{P}\big(\big|\tilde{V}_1'(n,m)-F(n,k)\big|\ge\sqrt{nk}\varepsilon\big)\le C_4k\sqrt{np_{\max}^m}\exp\biggl(-\frac{\sqrt{np_{\max}^m}\varepsilon}{C_4\sqrt{k}}\biggr), \tag{3.16}$$

for some positive constant C_4. From (3.12), for any $x>2$ and $0<\varepsilon<\min(1,x-2)$,

$$\mathbb{P}\big(F(n,k)\ge\sqrt{nk}(x\pm\varepsilon)\big)=\exp\big\{-k\big(I_1(x\pm\varepsilon)+\mathrm{o}(1)\big)\big\}. \tag{3.17}$$

Hence,

$$\frac{\mathbb{P}(|\tilde{V}_1'(n,m) - F(n,k)| \geq \sqrt{nk}\varepsilon)}{\mathbb{P}(F(n,k) \geq \sqrt{nk}(x \pm \varepsilon))}$$
$$\leq C_4 k\sqrt{np_{\max}^m}\exp\left(\sqrt{\frac{np_{\max}^m}{k}}\left(-\frac{\varepsilon}{C_4} + \sqrt{\frac{k^3}{np_{\max}^m}}\big(I_1(x \pm \varepsilon) + \mathrm{o}(1)\big)\right)\right) \longrightarrow 0,$$

since $k(m(n))^3/p_{\max}^m = \mathrm{o}(n)$, and, again, as in the proof of Theorem 1.1, this leads to (3.4) for any $x > 2$. Arguments similar to those developed at the end of the proof of Theorem 1.1 show that (3.4) is valid for any $x \geq 2$.

The proof of (3.5) is similar to the uniform case. First, from (3.10) and (3.11), for any fixed $x < 2$,

$$\lim_{k\to\infty} \frac{1}{k} \log \mathbb{P}\big(F(1,k) \leq \sqrt{k}x\big) = -\infty. \tag{3.18}$$

Moreover, for any $0 < \varepsilon < \min(1, 2 - x)$,

$$\begin{aligned} &\mathbb{P}\big(\tilde{V}_1'(n,m) \leq \sqrt{nk}x\big) \\ &\quad \leq \mathbb{P}\big(F(n,k) \leq \sqrt{nk}(x+\varepsilon)\big) + \mathbb{P}\big(\big|\tilde{V}_1'(n,m) - F(n,k)\big| \geq \sqrt{nk}\varepsilon\big), \end{aligned} \tag{3.19}$$

while $\mathbb{P}(|\tilde{V}_1'(n,m) - F(n,k)| \geq \sqrt{nk}\varepsilon)$ is exponentially negligible with speed $k(m)$. Therefore, (3.5) holds true under the condition $k(m(n))^3/p_{\max}^m = \mathrm{o}(n)$. □

Proof of Theorem 1.3. First, so as not to further burden the notations, below m is short for $m(n)$ and k is short for $k(m(n))$. Next, set $X = (V_1(n,m) - np_{\max}^m)/\sqrt{nkp_{\max}^m}$, $Y = (V_1(n,m) - V_1'(n,m))/\sqrt{nkp_{\max}^m}$ and $Z = (V_1'(n,m) - np_{\max}^m)/\sqrt{nkp_{\max}^m}$. Then, for any $x > 2$ and $0 < \varepsilon < x - 2$,

$$\mathbb{P}(X \geq x) \leq \mathbb{P}(Z \geq x - \varepsilon) + \mathbb{P}\big(|Y| \geq \varepsilon\big) \tag{3.20}$$

and

$$\mathbb{P}(X \geq x) \geq \mathbb{P}(Z \geq x + \varepsilon) - \mathbb{P}\big(|Y| \geq \varepsilon\big). \tag{3.21}$$

Moreover, from (3.1),

$$\mathbb{P}\big(|Y| \geq \varepsilon\big) \leq \frac{Cp_{2\mathrm{nd}}^m\sqrt{n}}{\varepsilon\sqrt{kp_{\max}^m}}, \tag{3.22}$$

and from Lemma 3.1,

$$\mathbb{P}(Z \geq x \pm \varepsilon) = \exp\big(-k\big(I_1(x \pm \varepsilon) + \mathrm{o}(1)\big)\big).$$

Under the condition (1.12),

$$\frac{\mathbb{P}(|Y| \geq \varepsilon)}{\mathbb{P}(Z \geq x \pm \varepsilon)} \leq \frac{Cp_{2\mathrm{nd}}^m\sqrt{n}}{\varepsilon\sqrt{kp_{\max}^m}} \exp\big(k\big(I_1(x \pm \varepsilon) + \mathrm{o}(1)\big)\big) \to 0, \qquad \text{as } n \to \infty. \tag{3.23}$$

Letting ε go to 0, and repeating the arguments of the proof of Theorem 1.1, establishes (1.13), for any $x \geq 2$, under the conditions given in Theorem 1.3. For (1.14), for any $x < 2$ and $0 < \varepsilon < 2 - x$,

$$\mathbb{P}(X \leq x) \leq \mathbb{P}(Z \leq x + \varepsilon) + \mathbb{P}\big(|Y| \geq \varepsilon\big).$$

From (3.3), $\mathbb{P}(Z \leq x + \varepsilon)$ is exponentially negligible with speed $k(m)$, and from arguments as in getting (3.23), $\mathbb{P}(|Y| \geq \varepsilon)$ is also exponentially negligible with speed $k(m)$, which proves (1.14). □

Proof of Theorem 1.4 and Remark 1.4. Again, below, m is short for $m(n)$ and k is short for $k(m(n))$. First, (1.19) is a direct consequence of (1.13), so let us prove (1.18). When on the left of its simultaneous asymptotic mean, $V_1'(n, m)$ can be approximated by $F(n, k)$ (see (3.8)). Hence, the rate function K_η should be the corresponding "left" rate function of the Brownian functional $F(1, k)$ (see (3.9)) with speed k^2. From the right-hand side of (3.10), it is clear that this new rate function depends on $\eta = \lim_{n\to\infty} k p_{\max}^m$; let us denote it by K_η. Moreover, for $F(1, k)$, and from [17] or [6],

$$\sqrt{1 - p_{\max}^m} F(1, k) \stackrel{\mathcal{L}}{=} \tilde{\lambda}_1^0,$$

where $\tilde{\lambda}_1^0$ is the largest eigenvalue of the diagonal block corresponding to $p_{\max}^m$ in $\mathbf{X^0}$, and where $\mathbf{X^0}$ is an element of $\mathcal{G}^0(p_1^m, p_2^m, \ldots, p_m^m)$ (see (1.15)). So the rate function K_η should also be the corresponding "left" rate function for $\tilde{\lambda}_1^0$ with speed k^2.

Again, from [17],

$$\lambda_1^k \stackrel{\mathcal{L}}{=} \tilde{\lambda}_1^0 + \sqrt{p_{\max}^m} g, \tag{3.24}$$

where λ_1^k is the largest eigenvalue of an element of the $k \times k$ GUE and where g is a standard normal random variable which is independent of $\tilde{\lambda}_1^0$.

Let

$$J(x) = \begin{cases} \inf\limits_{\mu \in \mathcal{M}((-\infty, x])} I(\mu), & \text{if } x \in (-\infty, 2], \\ 0, & \text{if } x \in [2, +\infty), \end{cases} \tag{3.25}$$

$$G_\eta(x) = \begin{cases} \dfrac{x^2}{2\eta}, & \text{if } x \in (-\infty, 0], \\ 0, & \text{if } x \in [0, +\infty), \end{cases} \tag{3.26}$$

be the respective rate function for λ_1^k, with speed k^2 and with $I(\mu)$ given in (A.5), and for the Gaussian term. Now, see [11], when $x \leq 2$,

$$\begin{aligned} J(x) = \tfrac{1}{216}\Big(&-x\big(-72x + x^3 + 30\sqrt{12 + x^2} + x^2\sqrt{12 + x^2}\big) \\ &- 216 \log\big(\tfrac{1}{6}\big(x + \sqrt{12 + x^2}\big)\big)\Big). \end{aligned} \tag{3.27}$$

Hence,

$$J'(x) = \tfrac{1}{54}\big(-x^3 + 36x - \big(12 + x^2\big)^{3/2}\big), \tag{3.28}$$

$$J''(x) = \tfrac{1}{18}\big(12 - x^2 - x\sqrt{12 + x^2}\big). \tag{3.29}$$

Clearly, for $x \in (-\infty, 2)$, $0 < J''(x) < 1$ and by a Taylor expansions for J and J', and for $x < -5$,

$$J(x) = \frac{x^2}{2} + \log(-x) + \frac{3}{4} + e_1(x), \tag{3.30}$$

$$J'(x) = x + \frac{1}{x} + e_2(x), \tag{3.31}$$

with $|e_1(x)| \le 2/x^2$ and $|e_2(x)| \le 4/|x|^3$.

From (3.24), it is well known (see [12,27]) that,

$$J(x) = K_\eta \,\square\, G_\eta(x) := \inf_{y\in\mathbb{R}}\big(K_\eta(y) + G_\eta(x - y)\big), \tag{3.32}$$

and taking Legendre transforms:

$$K_\eta(x) = \big(J^*(y) - G^*_\eta(y)\big)^*(x),$$

where

$$G^*(y) = \begin{cases} \dfrac{\eta y^2}{2}, & \text{if } y \le 0, \\ +\infty, & \text{if } y > 0, \end{cases}$$

so that

$$K_\eta(x) = \sup_{y\le 0}\left(xy - J^*(y) + \frac{\eta y^2}{2}\right). \tag{3.33}$$

Therefore, for $0 < \eta < 1$, K_η interpolates between $K_0 = J$ and $K_1 = K$. From the very definition of the Legendre transform,

$$J^*(y) = \sup_{x\in\mathbb{R}}\big(xy - J(x)\big),$$

there exists, for each $y \le 0$, a unique solution, denoted by $S(y)$, to $J'(x) = y$ for $x \in (-\infty, 2]$. Clearly, the function S is increasing on $(-\infty, 0]$, with $S(0) = 2$, $\lim_{y\to-\infty} S(y) = -\infty$, and with

$$S'(y) = \frac{1}{J''(S(y))},$$

for $y < 0$. Thus, for $y \le 2$,

$$J^*(y) = yS(y) - J\big(S(y)\big),$$

and,

$$K_\eta(x) = \sup_{y\le 0}\left(xy - yS(y) + J\big(S(y)\big) + \frac{\eta y^2}{2}\right).$$

For $y \le 0$, let

$$H_{x,\eta}(y) := xy - yS(y) + J\big(S(y)\big) + \frac{\eta y^2}{2},$$

then

$$H'_{x,\eta}(y) = x - S(y) + \eta y, \qquad H''_{x,\eta}(y) = -\frac{1}{J''(S(y))} + \eta,$$

so $H''_{x,\eta}(y) < 0$ for $y \in (-\infty, 0)$, $x \in \mathbb{R}$ and $0 \le \eta \le 1$. When $x \ge 2$, for any $0 \le \eta \le 1$, $H'_{x,\eta}(y) > 0$ for $y < 0$ with $H'_{x,\eta}(0) \ge 0$, thus $K_\eta(x) = \sup_{y\le 0} H_{x,\eta}(y) = H_{x,\eta}(0) = 0$. Let us now deal with $x < 2$. First, from (3.31), it can be shown that for $y < -6$,

$$y < S(y) < y + 1,$$

and thus since $x - J'(x)$ is increasing on $(-\infty, 2]$,

$$S(y) - y = S(y) - J'\big(S(y)\big) < y + 1 - J'(y+1) < -\frac{2}{y+1},$$

which further yields

$$y < S(y) < y - \frac{2}{y+1}.$$

Moreover, when $y < -6$,

$$\begin{aligned}\left|H_{x,\eta}(y) - \left(xy - y^2 + J(y) + \frac{\eta y^2}{2}\right)\right| &\le |y|\big|S(y) - y\big| + \big|J\big(S(y)\big) - J(y)\big| \\ &\le 2\left|\frac{y}{y+1}\right| + \big|J'(y+1)\big|\big|S(y) - y\big| \qquad (3.34)\\ &\le 3 + 3 = 6.\end{aligned}$$

Combining (3.34) with (3.30), it follows that for $y < -6$,

$$\left|H_{x,\eta}(y) - \left(xy + \log(-y) - \frac{1-\eta}{2}y^2\right)\right| \le 7. \qquad (3.35)$$

When $\eta = 1$, for any $x \le 0$, $H'_{x,1}(y) < 0$ for $y \le 0$, thus

$$K_1(x) = \lim_{y\to-\infty} H_{x,1}(y) = +\infty.$$

For $0 < x < 2$, since $S(y) - y$ is increasing on $(-\infty, 0]$ with a range of $(0, 2]$, there exists a unique solution, denoted by $T_1(x)$, to $H'_{x,1}(y) = x - S(y) + y = 0$. Note that $y = T_1(x)$ is the maximizer of $H_{x,1}(y)$ and as $x \to 0$, $T_1(x) \to -\infty$, thus there exists $\delta > 0$, such that for $x < \delta$,

$$K_1(x) = \sup_{y \le -6} H_{x,1}(y).$$

Since for $x < 1/6$,

$$\sup_{y \le -6} \big(xy + \log(-y)\big) = -1 - \log x,$$

when combined with (3.35), it follows that for x close enough to 0,

$$\big|K_1(x) - (-\log x)\big| \le 8.$$

When $0 < \eta < 1$, for any $x < 2$, there exists a unique solution, denoted by $T_\eta(x)$, to $H'_{x,\eta}(y) = x - S(y) + \eta y = 0$. Again, $y = T_\eta(x)$ is the maximizer of $H_{x,\eta}(y)$ and as $x \to -\infty$, $T_\eta(x) \to -\infty$. Repeating the arguments of the case $\eta = 1$, gives as $x \to -\infty$,

$$K_\eta(x) \sim \frac{x^2}{2(1-\eta)} + \log\left(-\frac{x}{1-\eta}\right).$$

This last statement (clearly consistent with the case $\eta = 0$), finishes to prove the last assertions of Remark 1.4. The rest of the proof of Theorem 1.4 follows along the lines of the proofs of Lemma 3.1 and of Theorem 1.3, and is therefore left to the interested reader. □

4. Proof of Theorem 1.5 and Theorem 1.6

Left and right concentration inequalities for the largest eigenvalue λ_1^m of an element of the $m \times m$ GUE are respectively given in [2] and [21]. More precisely:

Proposition 4.1. *Let $m \ge 1$ and let $\varepsilon > 0$, then for some absolute constant $C_0 > 0$,*

$$\mathbb{P}\big(\lambda_1^m \ge 2\sqrt{m}(1+\varepsilon)\big) \le C_0 \mathrm{e}^{-m\varepsilon^{3/2}/C_0}. \tag{4.1}$$

Likewise, for some absolute constant $\bar{C}_0 > 0$, and all $m \ge 1$ and $0 < \varepsilon \le 1$,

$$\mathbb{P}\big(\lambda_1^m \le 2\sqrt{m}(1-\varepsilon)\big) \le \bar{C}_0 \mathrm{e}^{-m^2\varepsilon^3/\bar{C}_0}. \tag{4.2}$$

Next, to prove (1.21), assume first that $m\varepsilon^{3/2} \ge 1$. Then for any $0 < \varepsilon < 1$,

$$\begin{aligned}
&\mathbb{P}\left(\frac{V_1(n,m) - n/m}{\sqrt{n/m}} \ge 2\sqrt{m}(1+\varepsilon)\right) \\
&\quad \le \mathbb{P}\left(\sqrt{\frac{m-1}{m}}\frac{\tilde{L}_1(n,m)}{2\sqrt{mn}} \ge 1 + \frac{\varepsilon}{2}\right) \\
&\qquad + \mathbb{P}\left(\sqrt{\frac{m-1}{m}}\frac{|\tilde{V}_1(n,m) - \tilde{L}_1(n,m)|}{2\sqrt{mn}} \ge \frac{\varepsilon}{2}\right).
\end{aligned} \tag{4.3}$$

As before,

$$\sqrt{\frac{m-1}{m}}\frac{\tilde{L}_1(n,m)}{\sqrt{n}} \overset{\mathcal{L}}{=} \lambda_1^{m,0}$$

and

$$\lambda_1^m \overset{\mathcal{L}}{=} \lambda_1^{m,0} + g_m,$$

where g_m is a centered Gaussian random variable with variance $1/m$, independent of $\lambda_1^{m,0}$. So,

$$\begin{aligned}\mathbb{P}\left(\sqrt{\frac{m-1}{m}}\frac{\tilde{L}_1(n,m)}{2\sqrt{mn}} \geq 1+\frac{\varepsilon}{2}\right) &\leq \mathbb{P}\left(\lambda_1^m \geq 2\sqrt{m}\left(1+\frac{\varepsilon}{4}\right)\right) + \mathbb{P}\left(g_m \geq \frac{\sqrt{m}\varepsilon}{2}\right)\\ &\leq C_1 \mathrm{e}^{-m\varepsilon^{3/2}/C_1} + C_1 \mathrm{e}^{-m^2\varepsilon^2/C_1},\end{aligned}$$

for some positive constant C_1. Now from (2.7), (2.8) and (2.9), the second term on the right-hand side of (4.3) is upper-bounded by:

$$\mathbb{P}\left(\frac{|\tilde{V}_1(n,m) - \tilde{L}_1(n,m)|}{2\sqrt{mn}} \geq \frac{\varepsilon}{2}\right) \leq C_2\sqrt{mn}\mathrm{e}^{-\sqrt{n}\varepsilon/C_2 m} + C_2 mn \mathrm{e}^{-n\varepsilon^2/C_2 m}.$$

In order to reach (1.21), we need to show that there exists a positive constant $C(A,\alpha)$, depending only on A and α, such that

$$C(A,\alpha)\mathrm{e}^{-m\varepsilon^{3/2}/C(A,\alpha)} \geq C_1 \mathrm{e}^{-m^2\varepsilon^2/C_1}, \tag{4.4}$$

$$C(A,\alpha)\mathrm{e}^{-m\varepsilon^{3/2}/C(A,\alpha)} \geq C_2\sqrt{mn}\mathrm{e}^{-\sqrt{n}\varepsilon/C_2 m}, \tag{4.5}$$

$$C(A,\alpha)\mathrm{e}^{-m\varepsilon^{3/2}/C(A,\alpha)} \geq C_2 mn \mathrm{e}^{-n\varepsilon^2/C_2 m}. \tag{4.6}$$

First, since $m\varepsilon^{3/2} \geq 1$, (4.4) is satisfied by choosing $C(A,\alpha) \geq C_1$. Now taking logarithms in (4.5), $C(A,\alpha)$ has to be such that:

$$\log\frac{C_2}{C(A,\alpha)} + \frac{1}{2}\log(mn) \leq m\varepsilon^{3/2}\left(-\frac{1}{C(A,\alpha)} + \frac{\sqrt{n}}{C_2 m^2 \varepsilon^{1/2}}\right). \tag{4.7}$$

Moreover, under the condition $m \leq An^\alpha$, we have:

$$\frac{\sqrt{n}}{C_2 m^2 \varepsilon^{1/2}} \geq \frac{\sqrt{n}}{C_2 m^2} \geq \frac{n^{(1/2)-2\alpha}}{A^2 C_2}.$$

Therefore, if $\alpha < 1/4$, it is enough to choose $C(A,\alpha)$ satisfying

$$\log\frac{\sqrt{A}C_2}{C(A,\alpha)} + \frac{1}{C(A,\alpha)} \leq \frac{n^{(1/2)-2\alpha}}{A^2 C_2} - \frac{1+\alpha}{2}\log n.$$

Since for all integers $n \geq 1$,

$$\frac{n^{1/2-2\alpha}}{A^2C_2} - \frac{1+\alpha}{2}\log n \geq \frac{1+\alpha}{1-4\alpha}\left(1 - \log\frac{A^2C_2(1+\alpha)}{1-4\alpha}\right),$$

we just need to guarantee that

$$\log\frac{\sqrt{A}C_2}{C(A,\alpha)} + \frac{1}{C(A,\alpha)} \leq \frac{1+\alpha}{1-4\alpha}\left(1 - \log\frac{A^2C_2(1+\alpha)}{1-4\alpha}\right). \tag{4.8}$$

But, from our choice of α, $(1+\alpha)/(1-4\alpha) > 1$, so by choosing

$$C(A,\alpha) \geq C\max\bigl(A^{5/2}, 1\bigr)\frac{1+\alpha}{1-4\alpha}\exp\left(\frac{1+\alpha}{1-4\alpha}\right), \tag{4.9}$$

for some large enough absolute constant $C > 0$, (4.8) and (4.5) are satisfied. Finally, by taking logarithms, (4.6) becomes,

$$\log\frac{C_2}{C(A,\alpha)} + \log(mn) \leq m\varepsilon^{3/2}\left(-\frac{1}{C(A,\alpha)} + \frac{n\varepsilon^{1/2}}{C_2m^2}\right). \tag{4.10}$$

From the condition $m \leq An^\alpha$, we just need,

$$\log\frac{AC_2}{C(A,\alpha)} + \frac{1}{C(A,\alpha)} \leq \frac{1}{A^{7/3}C_2}n^{1-7\alpha/3} - (1+\alpha)\log n. \tag{4.11}$$

Now repeating the previous arguments, taking the minimum on the right-hand side of (4.11), it follows that

$$\log\frac{AC_2}{C(A,\alpha)} + \frac{1}{C(A,\alpha)} \leq \frac{1+\alpha}{1-7\alpha/3}\left(1 - \log\frac{A^{7/3}C_2(1+\alpha)}{1-7\alpha/3}\right). \tag{4.12}$$

Again, for $0 < \alpha < 1/4$, $1 < (1+\alpha)/(1-7\alpha/3) < 3$, so as long as

$$C(A,\alpha) \geq C\max\bigl(A^{10/3}, 1\bigr)\frac{1+\alpha}{1-7\alpha/3}\exp\left(\frac{1+\alpha}{1-7\alpha/3}\right), \tag{4.13}$$

for some large enough absolute constant C, then $C(A,\alpha)$ will also satisfy (4.12) and therefore also (4.6).

Combining (4.9) and (4.13), if $m\varepsilon^{3/2} \geq 1$, and $m \leq An^\alpha$, with $\alpha < 1/4$, there exist a positive constant

$$C(A,\alpha) = C\max\bigl(A^{10/3}, 1\bigr)\frac{1+\alpha}{1-4\alpha}\exp\left(\frac{1+\alpha}{1-4\alpha}\right), \tag{4.14}$$

so that (1.21) holds true for all $0 < \varepsilon < 1$. When $m\varepsilon^{3/2} < 1$,

$$C(A,\alpha)\mathrm{e}^{-m\varepsilon^{3/2}/C(A,\alpha)} \geq C\mathrm{e}^{-1/C} \geq 1,$$

as C is large enough, and (1.21) follows. So combining these two cases, there exists $C(A,\alpha)$ as in (4.14), with C large enough, such that (1.21) is satisfied.

Likewise, for the proof of (1.22), first assume that $m^2\varepsilon^3 \geq 1$, and

$$\begin{aligned}
&\mathbb{P}\left(\frac{V_1(n,m)-n/m}{\sqrt{n/m}} \geq 2\sqrt{m}(1-\varepsilon)\right) \\
&\quad \leq \mathbb{P}\left(\sqrt{\frac{m-1}{m}}\frac{\tilde{L}_1(n,m)}{2\sqrt{mn}} \leq 1-\frac{\varepsilon}{2}\right) \\
&\qquad + \mathbb{P}\left(\sqrt{\frac{m-1}{m}}\frac{|\tilde{V}_1(n,m)-\tilde{L}_1(n,m)|}{2\sqrt{mn}} \geq \frac{\varepsilon}{2}\right) \\
&\quad \leq C_1 \mathrm{e}^{-m^2\varepsilon^3/C_1} + C_1 \mathrm{e}^{-m^2\varepsilon^2/C_1} + C_2\sqrt{mn}\,\mathrm{e}^{-\sqrt{n}\varepsilon/C_2 m} + C_2 mn\,\mathrm{e}^{-n\varepsilon^2/C_2 m}.
\end{aligned} \tag{4.15}$$

Repeating previous arguments, and as long as $m \leq An^\alpha$, with $\alpha < 1/6$, there exists a positive constant

$$\bar{C}(A,\alpha) = \bar{C}\max\bigl(A^4, 1\bigr)\frac{1+\alpha}{1-6\alpha}\exp\left(\frac{1+\alpha}{1-6\alpha}\right),$$

so that (1.22) is satisfied. Once more, taking $\bar{C}$ large enough, the case $m^2\varepsilon^3 < 1$ follows, and (1.22) is proved.

The proof for the non-uniform case is similar to the uniform one. For (1.24), assume at first that $k\varepsilon^{3/2} \geq 1$, then

$$\begin{aligned}
&\mathbb{P}\left(\frac{V_1(n,m)-np^m_{\max}}{\sqrt{nkp^m_{\max}}} \geq 2(1+\varepsilon)\right) \\
&\quad \leq \mathbb{P}\left(\frac{V_1(n,m)-V_1'(n,m)}{2\sqrt{nkp^m_{\max}}} \geq \frac{\varepsilon}{3}\right) + \mathbb{P}\left(\sqrt{1-p^m_{\max}}\frac{\tilde{V}_1'(n,m)-F(n,k)}{2\sqrt{nk}} \geq \frac{\varepsilon}{3}\right) \\
&\qquad + \mathbb{P}\left(\sqrt{1-p^m_{\max}}\frac{F(n,k)}{2\sqrt{nk}} \geq 1+\frac{\varepsilon}{3}\right) \\
&\quad = A_1 + A_2 + A_3.
\end{aligned}$$

From (3.22), (3.13) and (3.10),

$$\begin{aligned}
A_1 &\leq \frac{C_1 p^m_{\mathrm{2nd}}\sqrt{n}}{\varepsilon\sqrt{kp^m_{\max}}}, \\
A_2 &\leq C_2 nk\exp\left(-\frac{n\varepsilon^2}{C_2 k}\right) + C_2 k\sqrt{np^m_{\max}}\exp\left(-\frac{\sqrt{np^m_{\max}}\,\varepsilon}{C_2\sqrt{k}}\right), \\
A_3 &\leq \mathbb{P}\left(Z_k \geq \frac{\varepsilon}{3}\right) + \mathbb{P}\left(\lambda_1^k \geq 2\left(1+\frac{\varepsilon}{6}\right)\right) \\
&\leq C_3\exp\left(-\frac{k^2\varepsilon^2}{C_3}\right) + C_3\exp\left(-\frac{k\varepsilon^{3/2}}{C_3}\right).
\end{aligned}$$

In order to reach (1.24), we need to show that there exists a positive constant $C(A,B,\alpha)$, depending only on A, B and α, such that

$$C(A,B,\alpha)\exp\left(-\frac{k\varepsilon^{3/2}}{C(A,B,\alpha)}\right)\geq\frac{C_1p_{\text{2nd}}^m\sqrt{n}}{\varepsilon\sqrt{kp_{\max}^m}},\tag{4.16}$$

$$C(A,B,\alpha)\exp\left(-\frac{k\varepsilon^{3/2}}{C(A,B,\alpha)}\right)\geq C_2nk\exp\left(-\frac{n\varepsilon^2}{C_2k}\right),\tag{4.17}$$

$$C(A,B,\alpha)\exp\left(-\frac{k\varepsilon^{3/2}}{C(A,B,\alpha)}\right)\geq C_2k\sqrt{np_{\max}^m}\exp\left(-\frac{\sqrt{np_{\max}^m}\varepsilon}{C_2\sqrt{k}}\right),\tag{4.18}$$

$$C(A,B,\alpha)\exp\left(-\frac{k\varepsilon^{3/2}}{C(A,B,\alpha)}\right)\geq C_3\exp\left(-\frac{k^2\varepsilon^2}{C_3}\right).\tag{4.19}$$

First, taking logarithms in (4.18), gives:

$$\log\frac{C_2}{C(A,B,\alpha)}+\log k+\frac{1}{2}\log\bigl(np_{\max}^m\bigr)\leq k\varepsilon^{3/2}\left(-\frac{1}{C(A,B,\alpha)}+\frac{\sqrt{np_{\max}^m}}{C_2\sqrt{\varepsilon k^3}}\right).$$

Next,

$$\frac{\sqrt{np_{\max}^m}}{C_2\sqrt{\varepsilon k^3}}\geq\frac{\sqrt{(np_{\max}^m)^{1-3/\alpha}}}{A^{3/2\alpha}C_2},$$

so if $\alpha>3$, then there exists a constant $C(A,B,\alpha)>0$, satisfying (4.18). In fact, here $C(A,B,\alpha)$ just needs to be such that

$$\log\frac{A^{1/\alpha}C_2}{C(A,B,\alpha)}+\frac{1}{C(A,B,\alpha)}\leq\frac{\alpha+2}{\alpha-3}\left(1-\log\frac{A^{3/2\alpha}C_2(\alpha+2)}{\alpha-3}\right),$$

which forces

$$C(A,B,\alpha)\geq C\max\bigl(A^{2/\alpha},1\bigr)\frac{\alpha+2}{\alpha-3}\exp\left(\frac{\alpha+2}{\alpha-3}\right),\tag{4.20}$$

for a large enough absolute constant $C>0$.

Second, taking logarithms in (4.16), gives:

$$\log\frac{C_1}{C(A,B,\alpha)}+\log\left(\frac{p_{\text{2nd}}^m\sqrt{n}}{\sqrt{kp_{\max}^m}}\right)\leq-\frac{k\varepsilon^{3/2}}{C(A,B,\alpha)}+\log\varepsilon.$$

From (1.23) and the assumption $k\varepsilon^{3/2}\geq 1$, in order for (4.16) to hold true, $C(A,B,\alpha)$ needs to satisfy

$$\log\frac{C_1\sqrt{B}}{C(A,B,\alpha)}-\frac{k}{2}\leq-\frac{k}{C(A,B,\alpha)}-\frac{2}{3}\log k,$$

which further forces

$$C(A, B, \alpha) \geq C \max(\sqrt{B}, 1), \tag{4.21}$$

with the absolute constant C large enough.

For (4.17), as we did with (4.6), and under the condition $k^\alpha / p_{\max}^m \leq An$ with $\alpha > 3$, we need:

$$C(A, B, \alpha) \geq C \max\bigl(A^{10/3\alpha}, 1\bigr) \frac{3\alpha + 3}{3\alpha - 7} \exp\left(\frac{3\alpha + 3}{3\alpha - 7}\right), \tag{4.22}$$

with the absolute constant C large enough. Finally, (4.19) is easily satisfied since $k\varepsilon^{3/2} \geq 1$. Moreover, when $k\varepsilon^{3/2} < 1$, then (1.24) holds, given $C > 0$ large enough. Combining (4.20), (4.21) and (4.22), choosing

$$C(A, B, \alpha) = C \max\bigl(A^{10/3\alpha}, 1\bigr) \max(\sqrt{B}, 1) \frac{\alpha + 2}{\alpha - 3} \exp\left(\frac{\alpha + 2}{\alpha - 3}\right),$$

with $C > 0$, some large enough absolute constant, then (1.24) holds under the given conditions. Likewise, we can prove (1.26).

Appendix: Large deviations for the spectrum of the traceless GUE

For any integer $m \geq 2$, let the random matrix $\mathbf{X}$ be an element of the $m \times m$ GUE. Let $(\lambda_1, \lambda_2, \dots, \lambda_m)$ be the spectrum of $\mathbf{X}$, and let

$$(\xi_1, \xi_2, \dots, \xi_m) = \frac{1}{\sqrt{m}} (\lambda_1, \lambda_2, \dots, \lambda_m).$$

The joint probability density of $(\xi_1, \xi_2, \dots, \xi_m)$ is given by

$$\phi_m(\xi_1, \xi_2, \dots, \xi_m) = \frac{1}{Z_m} \exp\left(-\frac{m}{2} \sum_{i=1}^{m} \xi_i^2\right) \prod_{1 \leq i < j \leq m} (\xi_i - \xi_j)^2, \tag{A.1}$$

where

$$Z_m = (2\pi)^{m/2} m^{-m^2/2} \prod_{j=1}^{m} j!, \tag{A.2}$$

see Theorem 2.5.2 in [1] and also Theorem 3.3.1 in [25].

Let $(\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0)$ be the spectrum of $\mathbf{X} - \mathrm{tr}(\mathbf{X})/m$, an element of the $m \times m$ traceless GUE, and again, let

$$\bigl(\xi_1^0, \xi_2^0, \dots, \xi_m^0\bigr) = \frac{1}{\sqrt{m}} \bigl(\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0\bigr).$$

The joint distribution function of $(\xi_1^0, \xi_2^0, \dots, \xi_m^0)$ is given by

$$\begin{aligned}\mathbb{P}\big(\xi_1^0 \le s_1, \xi_2^0 \le s_2, \dots, \xi_m^0 \le s_m\big) \\ = \sqrt{2\pi} \int_{\mathcal{L}(s_1,\dots,s_m)} \phi_m(x_1, x_2, \dots, x_m)\, \mathrm{d}x_1 \cdots \mathrm{d}x_{m-1},\end{aligned} \tag{A.3}$$

where

$$\mathcal{L}(s_1, \dots, s_m) := \left\{ x = (x_1, \dots, x_m) \in \mathbb{R}^m : \sum_{i=1}^m x_i = 0, \text{ and } x_i \le s_i, \right. \\ \left. \text{for each } i = 1, \dots, m \right\}.$$

Let $(\xi_1^m, \xi_2^m, \dots, \xi_m^m)$ be the non-increasing rearrangement of $(\xi_1, \xi_2, \dots, \xi_m)$, and let $(\xi_1^{m,0}, \xi_2^{m,0}, \dots, \xi_m^{m,0})$ be the non-increasing rearrangement of $(\xi_1^0, \xi_2^0, \dots, \xi_m^0)$, then, for example, see [17],

$$\big(\xi_1^m, \xi_2^m, \dots, \xi_m^m\big) \stackrel{\mathcal{L}}{=} \big(\xi_1^{m,0}, \xi_2^{m,0}, \dots, \xi_m^{m,0}\big) + g_m \mathbf{e}_m, \tag{A.4}$$

where g_m is a centered Gaussian random variable with variance $1/m^2$, independent of the vector $(\xi_1^{m,0}, \xi_2^{m,0}, \dots, \xi_m^{m,0})$, and where $\mathbf{e}_m = (1, 1, \dots, 1)$.

As shown in [8], the law of the spectral measure $\hat{\mu}^m = \frac{1}{m} \sum_{i=1}^m \delta_{\xi_i}$ satisfies a large deviation principle on the set $\mathcal{P}(\mathbb{R})$ of probability measures on $\mathbb{R}$, and with good rate function I, in the scale m^2. Moreover, I is given by

$$I(\mu) = \frac{1}{2} \int x^2 \mu(\mathrm{d}x) - \iint \log|x - y| \mu(\mathrm{d}x)\mu(\mathrm{d}y) - \frac{3}{4}, \tag{A.5}$$

and its unique minimizer is the semicircular probability measure

$$\sigma(\mathrm{d}x) = \frac{1}{2\pi} \mathbf{1}_{|x|\le 2} \sqrt{4 - x^2}\, \mathrm{d}x.$$

Based on this LDP for $\hat{\mu}^m$, the LDP for the largest (or rth largest) eigenvalue of the GOE with an explicit rate function is obtained in [7] and [3] (see also [19] for generalizations). Following the approach and the techniques developed there, and taking into account (A.4), we get a multidimensional LDP for the first r eigenvalues of the traceless GUE:

Theorem A.1. *Let $r \in \mathbb{N}$, on $\mathcal{L}^r := \{(x_1, x_2, \dots, x_r) \in \mathbb{R}^r : x_1 \ge x_2 \ge \cdots \ge x_r\}$, $(\xi_1^{m,0}, \xi_2^{m,0}, \dots, \xi_r^{m,0})$ satisfies a LDP with speed m and a good rate function*

$$I_r(x_1, x_2, \dots, x_r) = \begin{cases} 2\sum_{i=1}^r \int_2^{x_i} \sqrt{(z/2)^2 - 1}\, \mathrm{d}z, & \text{if } x_1 \ge x_2 \ge \cdots \ge x_r \ge 2, \\ +\infty, & \text{otherwise.} \end{cases}$$

Proof. Let

$$Q_m(\mathrm{d}\xi_1, \mathrm{d}\xi_2, \dots, \mathrm{d}\xi_m) = \frac{1}{Z_m} \exp\left(-\frac{m}{2}\sum_{i=1}^m \xi_i^2\right) \prod_{1\le i<j\le m} (\xi_i - \xi_j)^2 \prod_{i=1}^m \mathrm{d}\xi_i.$$

From [7], $(\xi_1^m, \xi_2^m, \dots, \xi_r^m)$ satisfies a LDP with speed m and rate function I_r on $\mathcal{L}^r$. To prove the validity of the same results for $(\xi_1^{m,0}, \xi_2^{m,0}, \dots, \xi_r^{m,0})$, it is enough to show that

$$\limsup_{m\to\infty} \frac{1}{m} \log Q_m\big(\xi_r^{m,0} \le x\big) = -\infty, \tag{A.6}$$

for any $x < 2$, and since $I_r(x_1, x_2, \dots, x_r)$ is continuous, increasing in each individual variable, on $\mathcal{L}^r \cap [2,\infty)^r$,

$$\lim_{m\to\infty} \frac{1}{m} \log Q_m\big(\xi_1^{m,0} \ge x_1, \dots, \xi_r^{m,0} \ge x_r\big) = -2\sum_{i=1}^r \int_2^{x_i} \sqrt{(z/2)^2 - 1}\,\mathrm{d}z, \tag{A.7}$$

for all $x_1 \ge x_2 \ge \cdots \ge x_r \ge 2$.

First, for $x < 2$, let $\delta = 2 - x$, so

$$\begin{aligned} Q_m\big(\xi_r^{m,0} \le x\big) &\le Q_m\big(\xi_r^{m,0} + g_m \le x + \delta/2\big) + \mathbb{P}(g_m \ge \delta/2) \\ &= Q_m\big(\xi_r^m \le x + \delta/2\big) + \mathbb{P}(g_m \ge \delta/2). \end{aligned} \tag{A.8}$$

Since,

$$\mathbb{P}(g_m \ge \delta) \sim \frac{1}{\sqrt{2\pi} m\delta} \mathrm{e}^{-m^2\delta^2/2}, \qquad \text{as } m \to \infty, \tag{A.9}$$

(A.6) follows. For (A.7), fix $x_1 \ge x_2 \ge \cdots \ge x_r \ge 2$, for any $0 < \varepsilon < x_r$, we have

$$\begin{aligned} &\limsup_{m\to\infty} \frac{1}{m} \log Q_m\big(\xi_1^{m,0} \ge x_1, \dots, \xi_r^{m,0} \ge x_r\big) \\ &\quad \le \limsup_{m\to\infty} \frac{1}{m} \log\big(Q_m\big(\xi_1^m \ge x_1 - \varepsilon, \dots, \xi_r^m \ge x_r - \varepsilon\big) + \mathbb{P}(g_m \ge \varepsilon)\big). \end{aligned}$$

Moreover,

$$Q_m\big(\xi_1^m \ge x_1 - \varepsilon, \dots, \xi_r^m \ge x_r - \varepsilon\big) = \exp\big\{-m\big(I_r(x_1 - \varepsilon, x_2 - \varepsilon, \dots, x_r - \varepsilon) + \mathrm{o}(1)\big)\big\},$$

where o(1) goes to 0 as m goes to infinity. So for fixed $0 < \varepsilon < x_r$,

$$\frac{\mathbb{P}(g_m \ge \varepsilon)}{Q_m(\xi_1^m \ge x_1 - \varepsilon, \dots, \xi_r^m \ge x_r - \varepsilon)} \to 0, \qquad m \to \infty,$$

hence,

$$\limsup_{m\to\infty} \frac{1}{m} \log Q_m\big(\xi_1^{m,0} \geq x_1, \dots, \xi_r^{m,0} \geq x_r\big) \leq -I_r(x_1 - \varepsilon, x_2 - \varepsilon, \dots, x_r - \varepsilon).$$

Likewise,

$$\liminf_{m\to\infty} \frac{1}{m} \log Q_m\big(\xi_1^{m,0} \geq x_1, \dots, \xi_r^{m,0} \geq x_r\big) \geq -I_r(x_1 + \varepsilon, x_2 + \varepsilon, \dots, x_r + \varepsilon),$$

and letting $\varepsilon \to 0$, the continuity of the rate function leads to (A.7). □

For any $\mu \in \mathcal{P}(\mathbb{R})$, construct a discrete approximation via

$$x_i^m = \inf\left\{x \in \mathbb{R} : \mu\big((-\infty, x]\big) \geq \frac{i}{m+1}\right\}, \qquad 1 \leq i \leq m, \tag{A.10}$$

and $\mu^m = \sum_{i=1}^m \delta_{x_i^m}/m$ (note that the choice of the length $1/(m+1)$ of the intervals rather that $1/m$ is only made in order to ensure that x_m^m is finite). Using these discrete constructions, set:

$$\mathcal{X} = \left\{\mu \in \mathcal{P}(\mathbb{R}) : \frac{1}{\sqrt{m}} \sum_{i=1}^m x_i^m \to 0, \text{ as } m \to \infty\right\} \tag{A.11}$$

and

$$\mathcal{P}_0(\mathbb{R}) = \left\{\mu \in \mathcal{P}(\mathbb{R}) : \int x\mu(\mathrm{d}x) = 0\right\}. \tag{A.12}$$

Since the condition in (A.11) ensures that μ has mean zero, it is clear that $\mathcal{X}$ is a proper subset of $\mathcal{P}_0(\mathbb{R})$. With the above, and the arguments and results in [8], the large deviation principle for the spectral measure of the traceless GUE follow:

Theorem A.2. *The spectral measure $\hat{\mu}_0^m = \sum_{i=1}^m \delta_{\xi_i^0}/m$ satisfies a large deviation principle on $\mathcal{X}$ in the scale m^2 and with the good rate function I.*

Proof. Since this proof closely follows [8], it is just sketched here. Write the density of the eigenvalues as:

$$\begin{aligned} &Q_m\big(\mathrm{d}\xi_1^0, \mathrm{d}\xi_2^0, \dots, \mathrm{d}\xi_m^0\big) \\ &\quad = \frac{\sqrt{2\pi}}{Z_m} \exp\left(-m^2 \iint_{x\neq y} f(x,y)\hat{\mu}_0^m(\mathrm{d}x)\hat{\mu}_0^m(\mathrm{d}y)\right) \prod_{i=1}^m \mathrm{e}^{-\xi_i^{0\,2}/2}\,\mathrm{d}\xi_1^0 \cdots \mathrm{d}\xi_{m-1}^0, \end{aligned}$$

where $\xi_m^0 = -\sum_{i=1}^{m-1} \xi_i^0$ and

$$f(x,y) = \tfrac{1}{4}\big(x^2 + y^2\big) - \log|x - y|.$$

Let $\bar{Q}_m$ be the non-normalized positive measure $\bar{Q}_m = Z_m Q_m/\sqrt{2\pi}$. Via Stirling's formula,

$$\lim_{m\to\infty} \frac{1}{m^2} \log \frac{\sqrt{2\pi}}{Z_m} = \frac{1}{2} - \int_0^1 x \log x \,\mathrm{d}x = \frac{3}{4}, \tag{A.13}$$

so if under $\bar{Q}_m$, $\hat{\mu}_0^m$ satisfies a large deviation with rate function

$$J(\mu) = \iint f(x, y)\mu(\mathrm{d}x)\mu(\mathrm{d}y), \tag{A.14}$$

then combined with (A.13), this will lead to the statement of the theorem.

First, observe that for any Borel subset $A \subset \mathcal{X}$, any $N \in \mathbb{R}^+$,

$$\limsup_{m\to\infty} \frac{1}{m^2} \log\bigl(\bar{Q}_m\bigl(\hat{\mu}_0^m \in A\bigr)\bigr) \le -\inf_{\mu\in A}\left(\iint \bigl(f(x, y) \wedge N\bigr)\mu(\mathrm{d}x)\mu(\mathrm{d}y)\right). \tag{A.15}$$

Moreover, from arguments as in [8], the sequence $(\hat{\mu}_0^m)_{m\in\mathbb{N}}$ is exponentially tight under $\bar{Q}_m$ on $\mathcal{X}$. So we just need to prove that $(\hat{\mu}_0^m)_{m\in\mathbb{N}}$ satisfies a weak large deviation principle with rate function $J(\mu)$ under the measure $\bar{Q}_m$. The upper bound is clear. Indeed, $\mu \to \iint (f(x, y) \wedge N)\mu(\mathrm{d}x)\mu(\mathrm{d}y)$ is continuous on $\mathcal{X}$, therefore (A.15) implies that for any $\mu \in \mathcal{X}$,

$$\limsup_{\delta\to 0}\limsup_{m\to\infty} \frac{1}{m^2} \log\bigl(\bar{Q}_m\bigl(\hat{\mu}_0^m \in B(\mu, \delta)\bigr)\bigr) \le -\iint \bigl(f(x, y) \wedge N\bigr)\mu(\mathrm{d}x)\mu(\mathrm{d}y),$$

where $B(\mu, \delta)$ is the open ball of center μ and radius δ, with respect to the distance given by

$$d(\mu_1, \mu_2) = \sup_{g\in \mathrm{Lip_b}(1)}\left|\int g\,\mathrm{d}\mu_1 - \int g\,\mathrm{d}\mu_2\right|, \qquad \mu_1, \mu_2 \in \mathcal{X},$$

where for some fixed $b > 0$,

$$\mathrm{Lip_b}(1) = \bigl\{g : \mathbb{R} \to \mathbb{R} : \|g\|_{\mathrm{Lip}} \le 1, \|g\|_\infty \le b\bigr\},$$

are bounded Lipschitz functions. Then, by monotone convergence,

$$\limsup_{\delta\to 0}\limsup_{m\to\infty} \frac{1}{m^2} \log\bigl(\bar{Q}_m\bigl(\hat{\mu}_0^m \in B(\mu, \delta)\bigr)\bigr) \le -\iint f(x, y)\mu(\mathrm{d}x)\mu(\mathrm{d}y), \tag{A.16}$$

finishing the proof of the upper bound.

To prove the lower bound, let $\nu \in \mathcal{X}$. Since $I(\nu) = +\infty$ if ν has an atom, assume without loss of generality that ν is atomless. As in (A.10), let $\nu^m = \sum_{i=1}^m \delta_{x_i^m}/m$. Now, as $m \to \infty$, ν^m converges weakly, with probability one, towards ν. Hence, for any $\delta > 0$ and m large enough, setting $\Delta_m := \{\xi_1^0 \le \xi_2^0 \le \cdots \le \xi_m^0\}$,

$$\begin{aligned}
&\bar{Q}_m\bigl(\hat{\mu}_0^m \in B(\nu, \delta)\bigr) \\
&\quad \ge \bar{Q}_m\left(\left\{\max_{1\le i\le m-1}\bigl|\xi_i^0 - x_i^m\bigr| < \frac{\delta}{2\sqrt{m}}\right\} \cap \Delta_m\right)
\end{aligned}$$

$$\geq \int_{\mathcal{T}(\xi_1,\ldots,\xi_m)} \exp\left(-\frac{m}{2}\sum_{i=1}^{m}(\xi_i + x_i^m)^2\right) \prod_{1\leq i<j\leq m} \left|\xi_i - \xi_j + x_i^m - x_j^m\right|^2 \prod_{i=1}^{m-1} \mathrm{d}\xi_i$$

$$\geq \prod_{i+1<j} \left|x_i^m - x_j^m\right|^2 \times \prod_{i=1}^{m-1}\left|x_{i+1}^m - x_i^m\right| \exp\left\{-\frac{m}{2}\sum_{i=1}^{m-1}\left(\left|x_i^m\right| + \frac{\delta}{\sqrt{m}}\right)^2\right\}$$

$$\times \left|x_m^m - x_{m-1}^m\right| \exp\left(-m\left(\sum_{i=1}^{m-1} x_i^m\right)^2 - m^2\delta^2\right) \int_{\mathcal{T}(\xi_1,\ldots,\xi_m)} \prod_{i=1}^{m-2} |\xi_{i+1} - \xi_i| \prod_{i=1}^{m-1} \mathrm{d}\xi_i, \quad \text{(A.17)}$$

where

$$\mathcal{T}(\xi_1,\ldots,\xi_m) := \left\{\max_{1\leq i\leq m-1} |\xi_i| < \frac{\delta}{2\sqrt{m}}, \xi_1 \leq \xi_2 \leq \cdots \leq \xi_m, \sum_{i=1}^{m} \xi_i + \sum_{i=1}^{m} x_i^m = 0\right\}.$$

The last term on the right-hand side of (A.17) can be lower-bounded by changing variables: $\xi_1 = x_1$ and $\xi_i - \xi_{i-1} = x_i$, $2 \leq i \leq m-1$. Next, let

$$\mathcal{R}(x_1,\ldots,x_{m-1}) := \left\{-\frac{\delta}{2\sqrt{m}} \leq x_1 \leq -\frac{\delta}{4\sqrt{m}}, 0 \leq x_i \leq -\frac{\delta}{4m^2}, 2 \leq i \leq m-1\right\}.$$

Recalling that, $\sum_{i=1}^m x_i^m/m \to 0$,

$$\begin{aligned}\int_{\mathcal{T}(\xi_1,\ldots,\xi_m)} \prod_{i=1}^{m-2} |\xi_{i+1} - \xi_i| \prod_{i=1}^{m-1} \mathrm{d}\xi_i &\geq \int_{\mathcal{R}(x_1,\ldots,x_{m-1})} \prod_{i=2}^{m-1} |x_i| \prod_{i=1}^{m-1} \mathrm{d}x_i \\ &\geq \frac{\delta}{4\sqrt{m}}\left(\frac{1}{2}\left(\frac{\delta}{4m^2}\right)^2\right)^{m-2}.\end{aligned} \quad \text{(A.18)}$$

Hence,

$$\begin{aligned}&\bar{Q}_m\left(\hat{\mu}_0^m \in B(\nu,\delta)\right) \\ &\quad\geq \prod_{i+1<j} \left|x_i^m - x_j^m\right|^2 \prod_{i=1}^{m-1}\left|x_{i+1}^m - x_i^m\right| \exp\left(-\frac{m}{2}\sum_{i=1}^{m}\left(x_i^m\right)^2\right) \\ &\qquad\times \left|x_m^m - x_{m-1}^m\right| \frac{\delta}{4\sqrt{m}}\left(\frac{1}{2}\left(\frac{\delta}{4m^2}\right)^2\right)^{m-2} \exp\left(-\sqrt{m}\delta\sum_{i=1}^{m}\left|x_i^m\right| - \delta^2\right).\end{aligned} \quad \text{(A.19)}$$

Now by arguments as in [8],

$$\liminf_{\delta\to 0}\liminf_{m\to\infty}\frac{1}{m^2}\log\bigl(\bar{Q}_m\bigl(\hat{\mu}_0^m\in B(\nu,\delta)\bigr)\bigr)\geq -\iint f(x,y)\nu(\mathrm{d}x)\nu(\mathrm{d}y). \tag{A.20}$$

Combining (A.16) and (A.20), establishes the weak large deviation principle, finishing the proof of the theorem. □

We are now ready to give the large deviations for $\xi_1^{m,0}$ when on the left of its mean. To do so, let us introduce some notations: Let $\mathcal{M}((-\infty,x])$ be the set of all probability measures on $(-\infty,x]$, $x\in\mathbb{R}$, let $\mathcal{M}_{\mathcal{X}}((-\infty,x])=\mathcal{M}((-\infty,x])\cap\mathcal{X}$, and let $\mathcal{M}_0((-\infty,x])=\mathcal{M}((-\infty,x])\cap\mathcal{P}_0(\mathbb{R})$. Since $\{\xi_1^{m,0}\leq x\}=\{\hat{\mu}_0^m\in\mathcal{M}_{\mathcal{X}}((-\infty,x])\}$, then for any $x\leq 2$,

$$\lim_{m\to\infty}\frac{1}{m^2}\log\mathbb{P}\bigl(\xi_1^{m,0}\leq x\bigr)=-\inf_{\mu\in\mathcal{M}_{\mathcal{X}}((-\infty,x])}I(\mu). \tag{A.21}$$

For each $x\in\mathbb{R}$, let

$$K(x)=\inf_{\mu\in\mathcal{M}_0((-\infty,x])}I(\mu). \tag{A.22}$$

When $x\geq 2$, the semicircular law σ is both in $\mathcal{M}_{\mathcal{X}}((-\infty,x])$ and $\mathcal{M}_0((-\infty,x])$, and so $\inf_{\mu\in\mathcal{M}_{\mathcal{X}}((-\infty,x])}I(\mu)=K(x)=I(\sigma)=0$. Moreover, when $x\leq 0$, and since both $\mathcal{M}_{\mathcal{X}}((-\infty,x])$ and $\mathcal{M}_0((-\infty,x])$ are empty, $\inf_{\mu\in\mathcal{M}_{\mathcal{X}}((-\infty,x])}I(\mu)=K(x)=I(\sigma)=+\infty$.

When $0<x\leq 2$, and from arguments as in [18], it is next shown that K is continuous. Indeed, for any $y<0$ and $0<x\leq 2$, let

$$J_\mu(y,x)=\frac{1}{2}\int_y^x u^2\mu(\mathrm{d}u)-\int_y^x\int_y^x\log|u-t|\mu(\mathrm{d}u)\mu(\mathrm{d}t)-\frac{3}{4}, \tag{A.23}$$

and let ν_x be the minimizer of $I(\mu)$ on $\mathcal{M}_0((-\infty,x])$. Then, for any $0<\varepsilon<x$,

$$K(x)\leq K(x-\varepsilon)\leq\frac{J_{\nu_x}(y_\varepsilon,x-\varepsilon)}{\nu_x^2([y_\varepsilon,x-\varepsilon])}, \tag{A.24}$$

where y_ε is the value for which

$$\int_{y_\varepsilon}^{x-\varepsilon}t\,\mathrm{d}\nu_x(t)=0.$$

Since, as $\varepsilon\to 0$, the right-hand side of (A.24) converges to $K(x)$, K is left continuous.

To show the right continuity, note that by a simple change of variables,

$$K(x)=\inf_{\mu\in\mathcal{M}_0((-\infty,x+\varepsilon])}J_\mu^\varepsilon(x),$$

where

$$J_\mu^\varepsilon(x)=\frac{1}{2}\int_{-\infty}^{x+\varepsilon}(u-\varepsilon)^2\mu(\mathrm{d}u)-\int_{-\infty}^{x+\varepsilon}\int_{-\infty}^{x+\varepsilon}\log|u-t|\mu(\mathrm{d}u)\mu(\mathrm{d}t)-\frac{3}{4}.$$

Therefore,

$$0 \le K(x) - K(x+\varepsilon) \le J^{\varepsilon}_{\nu_{x+\varepsilon}}(x) - K(x+\varepsilon) = \frac{\varepsilon^2}{2},$$

and the right continuity of K follows. Likewise, $\inf_{\mu \in \mathcal{M}_{\mathcal{X}}((-\infty,x])} I(\mu)$ is right-continuous with respect to x.

Next we need a result which, when combined with (A.21), gives

$$\lim_{m\to\infty} \frac{1}{m^2} \log \mathbb{P}\big(\xi_1^{m,0} \le x\big) = -K(x), \tag{A.25}$$

for any $x \le 2$. This is the purpose of our next lemma whose statement as well as proof benefited from Ionel Popescu help.

Lemma A.1. *For any $x \in \mathbb{R}$,*

$$\inf_{\mu \in \mathcal{M}_{\mathcal{X}}((-\infty,x])} I(\mu) = K(x). \tag{A.26}$$

Proof. For $x \ge 2$, both sides of (A.26) are equal to zero and so we just need to consider the case $x < 2$. First, since $\mathcal{X}$ is a proper subset of $\mathcal{P}_0(\mathbb{R})$,

$$K(x) \le \inf_{\mu \in \mathcal{M}_{\mathcal{X}}((-\infty,x])} I(\mu). \tag{A.27}$$

Next, let us show that

$$K(x) \ge \inf_{\mu \in \mathcal{M}_{\mathcal{X}}((-\infty,x])} I(\mu). \tag{A.28}$$

By Theorem 1.10 and Theorem 1.11 of Chapter IV of [28], there exists a unique probability measure, μ_0, minimizing $I(\mu)$, for all $\mu \in \mathcal{M}_0((-\infty,x])$, and its support is an interval, $[a,b]$, with $b \le x$. Since μ_0 is atomless, its distribution function F is continuous, increasing with $F(a) = 0$ and $F(b) = 1$. Moreover, since μ_0 has mean zero, $\int_0^1 F^{-1}(x)\,\mathrm{d}x = 0$, where F^{-1}, the inverse of F, is continuous and increasing on [0, 1], with $F^{-1}(0) = a$ and $F^{-1}(1) = b$.

Now for any integer $n \ge 2$, construct an approximation to F^{-1} as follows: For $i/n \le x \le (i+1)/n$, let

$$G_n^+(x) = \begin{cases} n\left(F^{-1}\left(\frac{i+2}{n}\right)\left(x - \frac{i}{n}\right) + F^{-1}\left(\frac{i+1}{n}\right)\left(\frac{i+1}{n} - x\right)\right), & \text{if } 0 \le i \le n-2, \\ b + x - \frac{i}{n}, & \text{if } i = n-1, \end{cases}$$

and let,

$$G_n^-(x) = \begin{cases} n\left(F^{-1}\left(\frac{i}{n}\right)\left(x - \frac{i}{n}\right) + F^{-1}\left(\frac{i-1}{n}\right)\left(\frac{i+1}{n} - x\right)\right), & \text{if } 1 \le i \le n-1, \\ a + x - \frac{i+1}{n}, & \text{if } i = 0. \end{cases}$$

From this construction, $\int_0^1 G_n^+(x)\,\mathrm{d}x > 0$ and $\int_0^1 G_n^-(x)\,\mathrm{d}x < 0$. Next, let

$$\gamma_n^+ = \frac{-\int_0^1 G_n^-(x)\,\mathrm{d}x}{\int_0^1 G_n^+(x)\,\mathrm{d}x - \int_0^1 G_n^-(x)\,\mathrm{d}x}, \qquad \gamma_n^- = \frac{\int_0^1 G_n^+(x)\,\mathrm{d}x}{\int_0^1 G_n^+(x)\,\mathrm{d}x - \int_0^1 G_n^-(x)\,\mathrm{d}x},$$

and let

$$G_n(x) = \gamma_n^+ G_n^+(x) + \gamma_n^- G_n^-(x).$$

Then,

$$\int_0^1 G_n(x)\,\mathrm{d}x = 0,$$

and since G_n is piece-wise linear, it is Lipschitz. Let μ_n be the probability measure whose distribution function is G_n^{-1}. The Lipschitz continuity of G_n yields that $\mu_n \in \mathcal{X}$, for any $n \geq 2$. From its very construction, μ_n is supported on $[a - 1/n, b + 1/n]$, and μ_n converges to μ_0 weakly, as $n \to \infty$, and thus

$$\lim_{n\to\infty} \int x^2 \mu_n(\mathrm{d}x) = \int x^2 \mu_0(\mathrm{d}x). \tag{A.29}$$

For the second term on the right-hand side of (A.5),

$$\iint \log|x-y|\mu(\mathrm{d}x)\mu(\mathrm{d}y) = 2\iint_{x<y} \log(y-x)\mu(\mathrm{d}x)\mu(\mathrm{d}y), \tag{A.30}$$

let

$$\frac{1}{n^2}\sum_{i<j} \log\left(F^{-1}\left(\frac{j+1}{n}\right) - F^{-1}\left(\frac{i}{n}\right)\right) + \frac{1}{2n^2}\sum_{i=0}^{n-1} \log\left(F^{-1}\left(\frac{i+1}{n}\right) - F^{-1}\left(\frac{i}{n}\right)\right) \tag{A.31}$$

and

$$\frac{1}{n^2}\sum_{i<j} \log\left(G_n\left(\frac{j+1}{n}\right) - G_n\left(\frac{i}{n}\right)\right) + \frac{1}{2n^2}\sum_{i=0}^{n-1} \log\left(G_n\left(\frac{i+1}{n}\right) - G_n\left(\frac{i}{n}\right)\right), \tag{A.32}$$

be respectively Riemann sums approximations of $\iint_{x<y} \log(y-x)\mu_0(\mathrm{d}x)\mu_0(\mathrm{d}y)$ and $\iint_{x<y} \log(y-x)\mu_n(\mathrm{d}x)\mu_n(\mathrm{d}y)$. For any $0 \leq i \leq j \leq n-1$,

$$\begin{aligned} &\log\left(G_n\left(\frac{j+1}{n}\right) - G_n\left(\frac{i}{n}\right)\right) \\ &\quad \geq \gamma_n^+ \log\left(G_n^+\left(\frac{j+1}{n}\right) - G_n^+\left(\frac{i}{n}\right)\right) + \gamma_n^- \log\left(G_n^-\left(\frac{j+1}{n}\right) - G_n^-\left(\frac{i}{n}\right)\right), \end{aligned} \tag{A.33}$$

and moreover, for any $1 \le i \le j \le n-2$,

$$\log\left(G_n\left(\frac{j+1}{n}\right) - G_n\left(\frac{i}{n}\right)\right) \ge \gamma_n^+ \log\left(F^{-1}\left(\frac{j+2}{n}\right) - F^{-1}\left(\frac{i+1}{n}\right)\right) + \gamma_n^- \log\left(F^{-1}\left(\frac{j}{n}\right) - F^{-1}\left(\frac{i-1}{n}\right)\right). \tag{A.34}$$

If $\iint_{x<y} \log(y-x)\mu_0(\mathrm{d}x)\mu_0(\mathrm{d}y) = -\infty$, (A.28) is trivially true, so let us assume that this last integral is finite. Moreover, since $\gamma_n^+ + \gamma_n^- = 1$,

$$\liminf_{n\to\infty}\left(-\iint \log|x-y|\mu_n(\mathrm{d}x)\mu_n(\mathrm{d}y)\right) \le -\iint \log|x-y|\mu_0(\mathrm{d}x)\mu_0(\mathrm{d}y), \tag{A.35}$$

and combining (A.29) and (A.35),

$$\liminf_{n\to\infty} I(\mu_n) \le I(\mu_0).$$

Since μ_n is supported on $[a-1/n, b+1/n]$ and from the right continuity (in x) of $\inf_{\mu\in\mathcal{M}_{\mathcal{X}}((-\infty,x])} I(\mu)$,

$$K(x) \ge \inf_{\mu\in\mathcal{M}_{\mathcal{X}}((-\infty,x])} I(\mu),$$

which finishes the proof. □

To finish this Appendix, the large deviations for the first r eigenvalues of the traceless GUE, when at least one of them is on the left of the asymptotic mean, is established:

Corollary A.1. *For $x_r \le x_{r-1} \le \cdots \le x_1$, and $x_r \le 2$,*

$$\lim_{m\to\infty} \frac{1}{m^2} \log \mathbb{P}\left(\xi_1^{m,0} \le x_1, \ldots, \xi_r^{m,0} \le x_r\right) = -K(x_r).$$

Proof. Next, let $(\xi_1^{m,0}, \xi_2^{m,0}, \ldots, \xi_m^{m,0})$ be the non-increasing rearrangement of $(\xi_1^0, \xi_2^0, \ldots, \xi_m^0)$, and set

$$L := \mathbb{P}\left(\xi_1^{m,0} \le x_1, \ldots, \xi_r^{m,0} \le x_r\right),$$
$$M := \mathbb{P}\left(\xi_1^0 \le x_1, \ldots, \xi_r^0 \le x_r, \xi_{r+1}^0 \le x_r, \ldots, \xi_m^0 \le x_r\right).$$

Then,

$$M \le L \le \frac{m!}{(m-r+1)!(r-1)!} B \le m^r M, \tag{A.36}$$

and therefore,

$$\lim_{m\to\infty}\frac{1}{m^2}\log L=\lim_{m\to\infty}\frac{1}{m^2}\log M. \tag{A.37}$$

Changing variables:

$$\xi_i^0-(x_i-x_r)=\eta_i, \qquad \text{for } 1\le i\le r-1,$$
$$\xi_i^0=\eta_i, \qquad \text{for } r\le i\le m,$$

and so,

$$M=\mathbb{P}(\eta_i\le x_r, 1\le i\le m).$$

Considering the two measures $\sum_{i=1}^m \delta_{\xi_i^0}/m$ and $\sum_{i=1}^m \delta_{\eta_i}/m$, for any bounded Lipschitz function g (with $\|g\|_{\mathrm{Lip}}\le 1$), then as $m\to\infty$,

$$\frac{1}{m}\left|\sum_{i=1}^m g(\xi_i^0)-\sum_{i=1}^m g(\eta_i)\right|\le\frac{1}{m}\sum_{i=1}^m|\xi_i^0-\eta_i|\longrightarrow 0.$$

Therefore, $\sum_{i=1}^m \delta_{\xi_i^0}/m$ and $\sum_{i=1}^m \delta_{\eta_i}/m$ are exponentially equivalent, and Theorem A.2 also applies to the latter (see Theorem 4.2.13 in [12]). So from (A.25), it follows that

$$\lim_{m\to\infty}\frac{1}{m^2}\log M=-K(x_r),$$

and (A.37) finishes the proof. □

Acknowledgements

Many thanks to Satya Majumdar for suggesting that the methods of [26] should provide a closed form expression for the rate function obtained in Theorem 1.2, to Ionel Popescu for his help with Lemma A.1 and its proof, and to a referee for detailed comments. Christian Houdré's research was supported in part by the Simons Foundation Grant, #246283.

References

[1] Anderson, G.W., Guionnet, A. and Zeitouni, O. (2010). *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge: Cambridge Univ. Press. MR2760897

[2] Aubrun, G. (2005). A sharp small deviation inequality for the largest eigenvalue of a random matrix. In *Séminaire de Probabilités XXXVIII*. *Lecture Notes in Math*. **1857** 320–337. Berlin: Springer. MR2126983

[3] Auffinger, A., Ben Arous, G. and Černý, J. (2013). Random matrices and complexity of spin glasses. *Comm. Pure Appl. Math*. **66** 165–201. MR2999295

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

[5] Baryshnikov, Yu. (2001). GUEs and queues. *Probab. Theory Related Fields* **119** 256–274. MR1818248

[6] Benaych-Georges, F. and Houdré, C. (2013). A note on GUE minors, maximal Brownian functionals and longest increasing subsequences. Available at arXiv:1312.3301.

[7] Ben Arous, G., Dembo, A. and Guionnet, A. (2001). Aging of spherical spin glasses. *Probab. Theory Related Fields* **120** 1–67. MR1856194

[8] Ben Arous, G. and Guionnet, A. (1997). Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields* **108** 517–542. MR1465640

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

[10] 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

[11] Dean, D.S. and Majumdar, S.N. (2006). Large deviations of extreme eigenvalues of random matrices. *Phys. Rev. Lett.* **97** 160201, 4. MR2274338

[12] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. New York: Springer. MR1619036

[13] Deuschel, J.-D. and Zeitouni, O. (1999). On increasing subsequences of I.I.D. samples. *Combin. Probab. Comput.* **8** 247–263. MR1702546

[14] 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

[15] Houdré, C. and Litherland, T. (2011). On the limiting shape of Young diagrams associated with Markov random words. Available at arXiv:1110.4570.

[16] 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. Beachwood, OH: IMS. MR2797948

[17] Houdré, C. and Xu, H. (2013). On the limiting shape of Young diagrams associated with inhomogeneous random words. In *High Dimensional Probability VI*: *The Banff Volume*. *Progress in Probability* **66** 277–302. Basel: Birkhäuser.

[18] Ibrahim, J.-P. (2011). Large deviations for directed percolation on a thin rectangle. *ESAIM Probab. Stat.* **15** 217–232. MR2870513

[19] Johansson, K. (2000). Shape fluctuations and random matrices. *Comm. Math. Phys.* **209** 437–476. MR1737991

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

[21] Ledoux, M. and Rider, B. (2010). Small deviations for beta ensembles. *Electron. J. Probab.* **15** 1319–1343. MR2678393

[22] Lifshits, M. (2000). Lecture notes on strong approximation. *Pub. IRMA Lille* **53** 1–25.

[23] Löwe, M. and Merkl, F. (2001). Moderate deviations for longest increasing subsequences: The upper tail. *Comm. Pure Appl. Math.* **54** 1488–1520. MR1852980

[24] Löwe, M., Merkl, F. and Rolles, S. (2002). Moderate deviations for longest increasing subsequences: The lower tail. *J. Theoret. Probab.* **15** 1031–1047. MR1937784

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

[26] Nadal, C., Majumdar, S.N. and Vergassola, M. (2011). Statistical distribution of quantum entanglement for a random bipartite state. *J. Stat. Phys.* **142** 403–438. MR2764133

[27] Rockafellar, R.T. (1970). *Convex Analysis*. *Princeton Mathematical Series* **28**. Princeton, NJ: Princeton Univ. Press. MR0274683

[28] Saff, E.B. and Totik, V. (1997). *Logarithmic Potentials with External Fields*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **316**. Berlin: Springer. Appendix B by Thomas Bloom. MR1485778

[29] Seppäläinen, T. (1998). Large deviations for increasing sequences on the plane. *Probab. Theory Related Fields* **112** 221–244. MR1653841

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

[31] 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

[32] Tracy, C.A. and Widom, H. (2005). Matrix kernels for the Gaussian orthogonal and symplectic ensembles. *Ann*. *Inst*. *Fourier* (*Grenoble*) **55** 2197–2207. MR2187952

[33] Tricomi, F.G. (1957). *Integral Equations*. *Pure and Applied Mathematics* **V**. London: Interscience Publishers. MR0094665

Received December 2011 and revised December 2013