The Annals of Applied Probability
2019, Vol. 29, No. 3, 1311–1355
https://doi.org/10.1214/17-AAP1351
© Institute of Mathematical Statistics, 2019

THE LENGTH OF THE LONGEST COMMON SUBSEQUENCE OF TWO INDEPENDENT MALLOWS PERMUTATIONS[1]

BY KE JIN

University of Delaware

The Mallows measure is a probability measure on S_n where the probability of a permutation π is proportional to $q^{l(\pi)}$ with $q > 0$ being a parameter and $l(\pi)$ the number of inversions in π. We prove a weak law of large numbers for the length of the longest common subsequences of two independent permutations drawn from the Mallows measure, when q is a function of n and $n(1-q)$ has limit in $\mathbb{R}$ as $n \to \infty$.

1. Introduction.

1.1. *Background.* The longest common subsequence (LCS) problem is a classical problem which has application in fields such as molecular biology [see, e.g., Pevzner (2000)], data comparison and software version control. Most previous works on the LCS problem are focused on the case when the strings are generated uniformly at random from a given alphabet. Notably, Chvátal and Sankoff (1975) proved that the expected length of the LCS of two random k-ary sequences of length n when normalized by n converges to a constant γ_k. Since then, various endeavors [Dancík (1994), Dančík and Paterson (1995), Deken (1979), Lueker (2009)] have been made to determine the value of γ_k. The exact values of γ_k are still unknown. The known lower and upper bounds Lueker (2009) for γ_2 are

$$0.788071 < \gamma_2 < 0.826280.$$

In contrast to the LCS of two random strings, the LCS of two permutations is well connected to the longest increasing subsequence (LIS) problem [cf. Proposition 3.1 in Houdré and Işlak (2014)]. This can be seen from the following two facts:

- For any $\pi \in S_n$, the length of the LCS of π and the identity in S_n is equal to the length of the LIS of π.
- For any $\pi, \tau \in S_n$, the length of the LCS of π and τ is equal to the length of the LCS of $\tau^{-1} \circ \pi$ and the identity in S_n.

Received July 2017.
[1]Supported by NSF Grant DMS-1261010 and Sloan Research Fellowship.
MSC2010 subject classifications. 60F05, 60B15, 05A05.
Key words and phrases. Longest common subsequence, longest increasing subsequence, Mallows measure.

From the above two properties, it is easily seen that, if π, τ are independent and either π or τ is uniformly distributed on S_n the length of the LCS of π and τ has the same distribution as the length of the LIS of a uniformly random permutation. The length of the LIS of a uniformly random permutation has been well studied with major contributions from Hammersley (1972), Logan and Shepp (1977), Kerov and Vershik (1977) and culminating with the groundbreaking work of Baik, Deift and Johansson (1999) who prove that, under proper scaling, the length of the LIS converges to the Tracy–Widom distribution. Therefore, the length of the LCS of two permutations is only of interest when both permutations are nonuniformly distributed. In this paper, we study the length of the LCS of two independent permutations drawn from the Mallows measure.

DEFINITION 1.1. Given $\pi \in S_n$, the inversion set of π is defined by

$$\mathrm{Inv}(\pi) := \{(i, j) : 1 \le i < j \le n \text{ and } \pi(i) > \pi(j)\},$$

and the inversion number of π, denoted by $l(\pi)$, is defined to be the cardinality of $\mathrm{Inv}(\pi)$.

The Mallows measure on S_n is introduced by Mallows in Mallows (1957). For $q > 0$, the (n, q)—Mallows measure on S_n is given by

$$\mu_{n,q}(\pi) := \frac{q^{l(\pi)}}{Z_{n,q}},$$

where $Z_{n,q}$ is the normalizing constant. In other words, under the Mallows measure with parameter $q > 0$, the probability of a permutation π is proportional to $q^{l(\pi)}$. The Mallows measure has been used in modeling ranked and partially ranked data [see, e.g., Critchlow (1985), Fligner and Verducci (1993), Marden (1995)].

DEFINITION 1.2. For any $\pi, \tau \in S_n$, define the length of the longest common subsequence of π and τ as follows:

$$\mathrm{LCS}(\pi, \tau) := \max\bigl(m : \exists i_1 < \cdots < i_m \text{ and } j_1 < \cdots < j_m \text{ such that } \pi(i_k) = \tau(j_k) \text{ for all } k \in [m]\bigr).$$

Given the close connection of the LCS of two permutations and the LIS problem, to prove our results, we are able to make use of the techniques developed in Bhatnagar and Peled (2015), Mueller and Starr (2013) in which weak laws of large numbers of the length of the LIS of permutation under the Mallows measure have been proven for different regimes of q.

1.2. *Results*. Before stating the main theorem, we introduce the following lemma proved in Jin (2017), which shows the convergence of the empirical measure of a collection of random points defined by two independent Mallows permutations.

LEMMA 1.3. *Suppose that* $\{q_n\}_{n=1}^{\infty}$ *and* $\{q'_n\}_{n=1}^{\infty}$ *are two sequences such that* $\lim_{n\to\infty} n(1-q_n)=\beta$ *and* $\lim_{n\to\infty} n(1-q'_n)=\gamma$, *with* $\beta,\gamma\in\mathbb{R}$. *Let* $\mathbb{P}_n$ *denote the probability measure on* $S_n\times S_n$ *such that* $\mathbb{P}_n((\pi,\tau))=\mu_{n,q_n}(\pi)\cdot\mu_{n,q'_n}(\tau)$, *that is*, $\mathbb{P}_n$ *is the product measure of* μ_{n,q_n} *and* μ_{n,q'_n}. *For any* $R=(x_1,x_2]\times(y_1,y_2]\subset[0,1]\times[0,1]$, *we have*

$$\lim_{n\to\infty}\mathbb{P}_n\left(\left|\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_R\left(\frac{\pi(i)}{n},\frac{\tau(i)}{n}\right)-\int_R\rho(x,y)\,dx\,dy\right|>\varepsilon\right)=0, \tag{1}$$

for any $\varepsilon>0$, *with*

$$\rho(x,y):=\int_0^1 u(x,t,\beta)\cdot u(t,y,\gamma)\,dt, \tag{2}$$

where

$$u(x,y,\beta):=\frac{(\beta/2)\sinh(\beta/2)}{(e^{\beta/4}\cosh(\beta[x-y]/2)-e^{-\beta/4}\cosh(\beta[x+y-1]/2))^2}, \tag{3}$$

for $\beta\neq0$, *and* $u(x,y,0):=1$.

The density $u(x,y,\beta)$ in (3), obtained by Starr (2009), is the limiting distribution of the empirical measure induced by Mallows permutation when the parameters q_n satisfy that $\lim_{n\to\infty} n(1-q_n)=\beta$. The limiting distribution of the points of a random permutation is known as a *permuton* [cf. Hoppen et al. (2013)] and has recently been studied in the context of finding the limiting distribution of permutation statistics such as cycle lengths [Mukherjee et al. (2016)] and certain limit shapes of permutations with fixed pattern densities [Kenyon et al. (2015)].

The main result of this paper is a weak law of large numbers of the LCS of two permutations drawn independently from the Mallows measure. The first observation, which is proved in Corollary 2.4, is that the length of LCS of two permutations π and τ is equal to the length of the longest increasing points in the collection of points:

$$z(\pi^{-1},\tau^{-1}):=\left\{\left(\frac{\pi^{-1}(i)}{n},\frac{\tau^{-1}(i)}{n}\right)\right\}_{i\in[n]}.$$

At a high level, our proof follows the approach of Deuschel and Zeitouni (1995) who showed weak laws for the LIS of i.i.d. points drawn according to some density in a box. They partition the box into a grid of smaller boxes whose size is chosen to be such that the distribution of points within them is close to uniform. The weak law for the LIS of uniformly random permutations [Kerov and Vershik (1977)] can be applied to points in these boxes to estimate the number of increasing points in the neighborhood of any increasing path. In our case, this approach fails because the points in the box are no longer i.i.d.

Indeed, in a prior work, Mueller and Starr (2013) applied Deuschel and Zeitouni's approach to show a weak law for the LIS of a Mallows permutation, where due to properties of the Mallows measure, the permutation induced by the points in a smaller box is also Mallows distributed. They coupled the distribution of points to two i.i.d. point processes to overcome this problem. In our case, this does not seem to be applicable directly, since the induced permutation by the points in a box is no longer Mallows or the product of independent Mallows permutations. We follow a different approach. We prove a combinatorial fact using the properties of the weak Bruhat order to say that the distribution of the LIS of points in a small box can be stochastically bounded between the LIS and the LDS of a Mallows permutation restricted to a certain fixed set of indices. In their work, Mueller and Starr derived estimates on the LIS of a Mallows permutation in a small box; however, we cannot use these estimates directly because of the restriction to an arbitrary set of indices. To overcome this, we generalize their estimates to the LIS of a Mallows permutation restricted to an arbitrary set of indices. Our argument recovers their result for small enough β (which is the relevant case) and gives a slightly more streamlined proof.

Specifically, we establish two results to apply the approach above. The first result, deduced from Lemma 1.3, is that the number of points $z(\pi^{-1}, \tau^{-1})$ contained in any fixed rectangle, when divided by the size of the permutation, converges in probability to a constant. The second result, proved in Lemma 4.4, is that the length of the longest increasing points in $z(\pi^{-1}, \tau^{-1})$ within a small box R is close to the size of the LIS in the uniform case, that is, it is approximately $2\sqrt{|z(\pi^{-1}, \tau^{-1}) \cap R|}$. The main theorem is the following.

THEOREM 1.4. *Let $B^1_{\nearrow}$ denote the set of nondecreasing, C^1_b functions $\phi : [0, 1] \to [0, 1]$, with $\phi(0) = 0$ and $\phi(1) = 1$, where C^1_b denotes the set of functions which have bounded and continuous first-order derivative. Define function $J : B^1_{\nearrow} \to \mathbb{R}$,*

$$J(\phi) := \int_0^1 \sqrt{\dot{\phi}(x)\rho\bigl(x, \phi(x)\bigr)}\, dx \quad \text{and} \quad \bar{J} := \sup_{\phi \in B^1_{\nearrow}} J(\phi),$$

where $\rho(x, y)$ is the density defined in (2) and $\dot{\phi}$ denotes the derivative of ϕ. Under the same conditions as in Lemma 1.3, for any $\varepsilon > 0$, we have

$$\lim_{n\to\infty} \mathbb{P}_n\left(\left|\frac{\mathrm{LCS}(\pi, \tau)}{\sqrt{n}} - 2\bar{J}\right| < \varepsilon\right) = 1. \tag{4}$$

Finally, we derive the limiting constant in the special case when $\beta = \gamma$.

COROLLARY 1.5. *Suppose that* $\{q_n\}_{n=1}^{\infty}$ *and* $\{q'_n\}_{n=1}^{\infty}$ *are two sequences such that* $\lim_{n\to\infty} n(1-q_n) = \lim_{n\to\infty} n(1-q'_n) = \beta$ *with* $\beta \neq 0$. *Then the constant* $\bar{J}$ *in Theorem* 1.4 *is given by*

$$\bar{J} = \sqrt{\frac{\beta}{6\sinh(\beta/2)}} \cdot \int_0^1 \sqrt{\cosh(\beta/2) + 2\cosh\bigl(\beta[2x-1]/2\bigr)}\, dx.$$

2. Reducing the LCS problem to the LIS problem.

DEFINITION 2.1. Given a set of points in $\mathbb{R}^2$: $\boldsymbol{z} = \{z_1, z_2, \dots, z_n\}$, where $z_i = (x_i, y_i) \in \mathbb{R}^2$, we say that $(z_{i_1}, z_{i_2}, \dots, z_{i_m})$ is an increasing subsequence if

$$x_{i_j} < x_{i_{j+1}}, \qquad y_{i_j} < y_{i_{j+1}}, \qquad j = 1, \dots, m-1.$$

In the above, we do not require $i_j < i_{j+1}$. Let $\mathrm{LIS}(\boldsymbol{z})$ denote the length of the longest increasing subsequence of $\boldsymbol{z}$.

DEFINITION 2.2. Given $\boldsymbol{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$, $\boldsymbol{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$, we say that $((a_{i_1}, b_{i_1}), (a_{i_2}, b_{i_2}), \dots, (a_{i_m}, b_{i_m}))$ is an increasing subsequence between $\boldsymbol{a}$ and $\boldsymbol{b}$ if

$$a_{i_j} < a_{i_{j+1}}, \qquad b_{i_j} < b_{i_{j+1}}, \qquad j = 1, \dots, m-1.$$

In the above, we do not require $i_j < i_{j+1}$. Let $\mathrm{LIS}(\boldsymbol{a}, \boldsymbol{b})$ denote the length of the longest increasing subsequence between $\boldsymbol{a}$ and $\boldsymbol{b}$. Let $\mathrm{LIS}(\boldsymbol{a}) := \mathrm{LIS}(id, \boldsymbol{a})$, $\mathrm{LDS}(\boldsymbol{a}) := \mathrm{LIS}(id^r, \boldsymbol{a})$, where $id = (1, 2, \dots, n)$ denotes the identity in S_n and $id^r = (n, \dots, 1)$ denotes the reversal of identity in S_n. Hence $\mathrm{LIS}(\boldsymbol{a})$ is the length of the longest increasing subsequence of $\boldsymbol{a}$ and $\mathrm{LDS}(\boldsymbol{a})$ is the length of the longest decreasing subsequence of $\boldsymbol{a}$.

Note that Definition 2.2 allows us to define $\mathrm{LIS}(\pi, \tau)$, the length of the longest increasing subsequence of two permutations, by regarding π and τ as vectors in $\mathbb{Z}^n$. We show that $\mathrm{LCS}(\pi, \tau) = \mathrm{LIS}(\pi^{-1}, \tau^{-1})$, which allows us to reduce the LCS problem to an LIS problem.

LEMMA 2.3. *Given* $\pi, \tau \in S_n$, *we have*

$$\mathrm{LCS}(\pi, \tau) = \mathrm{LCS}(\sigma \circ \pi, \sigma \circ \tau), \qquad \mathrm{LIS}(\pi, \tau) = \mathrm{LIS}(\pi \circ \sigma, \tau \circ \sigma),$$

for any $\sigma \in S_n$.

PROOF. Suppose $(a_1, a_2, \dots, a_m)$ is a common subsequence of π and τ, then $(\sigma(a_1), \sigma(a_2), \dots, \sigma(a_m))$ is a common subsequence of $\sigma \circ \pi$ and $\sigma \circ \tau$. Hence

$$\mathrm{LCS}(\pi, \tau) \le \mathrm{LCS}(\sigma \circ \pi, \sigma \circ \tau) \le \mathrm{LCS}\bigl(\sigma^{-1} \circ \sigma \circ \pi, \sigma^{-1} \circ \sigma \circ \tau\bigr) = \mathrm{LCS}(\pi, \tau).$$

Similarly, suppose $((\pi(i_1), \tau(i_1)), (\pi(i_2), \tau(i_2)), \ldots, (\pi(i_m), \tau(i_m)))$ is an increasing subsequence between π and τ, then $((\pi \circ \sigma(i'_1), \tau \circ \sigma(i'_1)), (\pi \circ \sigma(i'_2), \tau \circ \sigma(i'_2)), \ldots, (\pi \circ \sigma(i'_m), \tau \circ \sigma(i'_m)))$ is an increasing subsequence between $\pi \circ \sigma$ and $\tau \circ \sigma$, where $i'_k = \sigma^{-1}(i_k)$ for $k \in [m]$. Hence

$$\mathrm{LIS}(\pi, \tau) \leq \mathrm{LIS}(\pi \circ \sigma, \tau \circ \sigma) \leq \mathrm{LIS}\bigl(\pi \circ \sigma \circ \sigma^{1}, \tau \circ \sigma \circ \sigma^{-1}\bigr) = \mathrm{LIS}(\pi, \tau). \quad \square$$

COROLLARY 2.4. *For any* $\pi, \tau \in S_n$, $\mathrm{LCS}(\pi, \tau) = \mathrm{LIS}(\pi^{-1}, \tau^{-1})$.

PROOF. By the previous lemma, we have

$$\mathrm{LCS}(\pi, \tau) = \mathrm{LCS}\bigl(id, \pi^{-1} \circ \tau\bigr) = \mathrm{LIS}\bigl(id, \pi^{-1} \circ \tau\bigr) = \mathrm{LIS}\bigl(\tau^{-1}, \pi^{-1}\bigr).$$

In the second equality, we use the following trivial fact:

$$\mathrm{LCS}(id, \pi) = \mathrm{LIS}(\pi) = \mathrm{LIS}(id, \pi). \quad \square$$

3. Weak Bruhat order. Before introducing the weak Bruhat order, we make the following definition.

DEFINITION 3.1. Given $\pi \in S_n$ and $\boldsymbol{a} = (a_1, a_2, \ldots, a_k)$, where $a_i \in [n]$ and $a_1 < a_2 < \cdots < a_k$, let $\pi(\boldsymbol{a}) = (\pi(a_1), \pi(a_2), \ldots, \pi(a_k))$. Let $\pi_{\boldsymbol{a}} \in S_k$ denote the permutation induced by $\pi(\boldsymbol{a})$, that is, $\pi_{\boldsymbol{a}}(i) = j$ if $\pi(a_i)$ is the jth smallest term in $\pi(\boldsymbol{a})$.

Lemma 4.4 says that the LIS of the points $\{(\frac{\pi(i)}{n}, \frac{\tau(i)}{n})\}_{i \in [n]}$ that fall in a small box is close to the uniform case. Let $\boldsymbol{a} = (a_1, \ldots, a_k)$ be an increasing sequence of indices with $a_i \in [n]$. To prove Lemma 4.4, we will show that there exists a coupling of permutations (X, Y, X', X''), where X, X' and X'' are distributed according to $\mu_{n,q}$ and Y is independent of X with an arbitrary distribution on S_n. Under this coupling, $\mathrm{LIS}(X_{\boldsymbol{a}}, Y_{\boldsymbol{a}})$ will be bounded by $\mathrm{LIS}(X'_{\boldsymbol{a}})$ and $\mathrm{LDS}(X''_{\boldsymbol{a}})$. The main tool that we use to construct the coupling is the weak Bruhat order on S_n.

Recall that for a permutation $\pi \in S_n$, $l(\pi)$ denotes the number of inversions of π and $\mathrm{Inv}(\pi)$ denotes the set of inversions of π as defined in Definition 1.1. Let (i, j) denote the transposition in S_n and $s_i := (i, i+1)$ the adjacent transposition in S_n.

DEFINITION 3.2. The left weak Bruhat order $(S_n, \leq_L)$ is defined as the transitive closure of the relations

$$\pi \leq_L \tau \qquad \text{if } \tau = s_i \circ \pi \text{ and } l(\tau) = l(\pi) + 1.$$

We are multiplying permutations right-to-left. For instance, $s_2 \circ 2413 = 3412$. The *right weak Bruhat order* $(S_n, \leq_R)$ is defined in the same way except that the covering relationship is given by $\tau = \pi \circ s_i$ and $l(\tau) = l(\pi) + 1$.

One characterization of the left weak order is the following proposition [cf. Abello (1991)]. We provide its proof here for the completeness of the paper.

PROPOSITION 3.3.

$$\pi \leq_L \tau \quad \textit{if and only if} \quad \operatorname{Inv}(\pi) \subseteq \operatorname{Inv}(\tau).$$

PROOF. Suppose τ covers π, that is, $s_i \circ \pi = \tau$ and $l(\pi) + 1 = l(\tau)$. It is easy to see that $\operatorname{Inv}(\tau) = \operatorname{Inv}(\pi) \cup \{(\pi^{-1}(i), \pi^{-1}(i+1))\}$. For arbitrary π and τ, $\pi \leq_L \tau$ implies that there exists a sequence of permutations $\{\sigma_0, \dots, \sigma_k\}$ such that σ_{i+1} covers σ_i and $\pi = \sigma_0 \leq_L \cdots \leq_L \sigma_k = \tau$. Hence $\pi \leq_L \tau$ implies $\operatorname{Inv}(\pi) \subseteq \operatorname{Inv}(\tau)$. On the other hand, given $\operatorname{Inv}(\pi) \subseteq \operatorname{Inv}(\tau)$, to show $\pi \leq_L \tau$ it suffices to show that there exists an adjacent transposition s_i such that $\operatorname{Inv}(\pi) \subseteq \operatorname{Inv}(s_i \circ \tau) \subset \operatorname{Inv}(\tau)$. Let k be the smallest i such that $\pi^{-1}(i) \neq \tau^{-1}(i)$. Let $j = \pi^{-1}(k)$ and $h = \tau(j)$. Since $h > k \geq 1$, define $j' = \tau^{-1}(h-1)$. By the choice of k, we have $\pi(j') > k$. It follows that $j < j'$, since otherwise $(j', j) \in \operatorname{Inv}(\pi)$ and $(j', j) \notin \operatorname{Inv}(\tau)$. Therefore, we have $\operatorname{Inv}(\pi) \subseteq \operatorname{Inv}(s_{h-1} \circ \tau) \subset \operatorname{Inv}(\tau)$. $\square$

LEMMA 3.4. *Given $\pi, \tau \in S_k$ with $\pi \leq_L \tau$, for any $n \geq k$, $0 < q \leq 1$ and increasing indices $\boldsymbol{a} = (a_1, a_2, \dots, a_k)$ with $a_i \in [n]$, there exists a coupling (X, Y) such that $X \sim \mu_{n,q}$, $Y \sim \mu_{n,q}$ and*

$$\operatorname{LIS}(X_{\boldsymbol{a}}, \pi) \geq \operatorname{LIS}(Y_{\boldsymbol{a}}, \tau).$$

PROOF. First, we claim that it suffices to show the case when τ covers π in $(S_k, \leq_L)$, that is, $l(\tau) = l(\pi) + 1$ and $\tau = s_i \circ \pi$ for some $i \in [k-1]$. The claim can be shown by induction on the Kendall's tau distance of π and τ, that is, the minimum number of adjacent transpositions multiplied to π from the left to get τ. Suppose we have $\pi \leq_L \sigma \leq_L \tau$ in S_k with $l(\pi) < l(\sigma) < l(\tau)$. By the induction hypothesis, there exist two couplings (X, Y) and (Y', Z), which are not necessarily defined in the same probability space, such that X, Y, Y', Z have the same marginal distribution $\mu_{n,q}$ and

$$\operatorname{LIS}(X_{\boldsymbol{a}}, \pi) \geq \operatorname{LIS}(Y_{\boldsymbol{a}}, \sigma), \qquad \operatorname{LIS}(Y'_{\boldsymbol{a}}, \sigma) \geq \operatorname{LIS}(Z_{\boldsymbol{a}}, \tau). \tag{5}$$

We can construct a new coupling (X', Z') as follows:

(1) Sample a permutation $\xi \in S_n$ according to the distribution $\mu_{n,q}$.

(2) Sample X' according to the induced distribution on S_n by the first coupling (X, Y) conditioned on $Y = \xi$.

(3) Sample Z' according to the induced distribution on S_n by the second coupling (Y', Z) conditioned on $Y' = \xi$.

By the law of total probability, it is easily seen that $X' \sim \mu_{n,q}$ and $Z' \sim \mu_{n,q}$. Also, regardless of which permutation ξ being sampled in the first step, by (5), we have

$$\operatorname{LIS}(X'_{\boldsymbol{a}}, \pi) \geq \operatorname{LIS}(\xi_{\boldsymbol{a}}, \sigma) \geq \operatorname{LIS}(Z'_{\boldsymbol{a}}, \tau).$$

In the remainder of the proof, we assume $\tau = s_i \circ \pi$ and $l(\tau) = l(\pi) + 1$. Note that, for any $\sigma \in S_n$,

(6) $$\sigma \circ (i,j) = (\sigma(i), \sigma(j)) \circ \sigma, \qquad \sigma_{\boldsymbol{a}} \circ (i,j) = (\sigma \circ (a_i, a_j))_{\boldsymbol{a}}.$$

Let $r = a_{\pi^{-1}(i)}$ and $t = a_{\pi^{-1}(i+1)}$. Since $l(\tau) = l(\pi) + 1$, we have $\pi^{-1}(i) < \pi^{-1}(i+1)$; thus $r < t$. Let $A := \{\{\sigma, \sigma \circ (r,t)\} : \sigma \in S_n \text{ and } \sigma(r) < \sigma(t)\}$. Clearly, A is a partition of S_n. Then we construct the coupling (X, Y) as follows:

(1) Choose a set in A according to measure $\mu_{n,q}$, that is, the set $\{\sigma, \sigma \circ (r,t)\}$ is chosen with probability $\mu_{n,q}(\{\sigma, \sigma \circ (r,t)\})$.

(2) Suppose the set $\{\sigma, \sigma \circ (r,t)\}$, with $\sigma(r) < \sigma(t)$, is chosen in the first step. Flip a coin with probability of heads being

$$p = \frac{q^{l(\sigma)} - q^{l(\sigma \circ (r,t))}}{q^{l(\sigma)} + q^{l(\sigma \circ (r,t))}}.$$

Note that the probability of heads p is nonnegative because we have $0 < q \le 1$ and the following fact:

$$i < j \quad \text{and} \quad \sigma(i) < \sigma(j) \quad \Rightarrow \quad l(\sigma) < l(\sigma \circ (i,j)) \qquad \forall \sigma \in S_n.$$

(3) If the outcome is heads, then we set $X = Y = \sigma$.

(4) If the outcome is tails, then with equal probability, we set either $X = \sigma$, $Y = \sigma \circ (r,t)$ or $X = \sigma \circ (r,t)$, $Y = \sigma$.

It can be verified that (X, Y) thus defined has the correct marginal distribution $\mu_{n,q}$. In the following, we show that

(7) $$\mathrm{LIS}(X_{\boldsymbol{a}} \circ \pi^{-1}) \ge \mathrm{LIS}(Y_{\boldsymbol{a}} \circ \tau^{-1}).$$

Then the lemma follows by Lemma 2.3 because

$$\mathrm{LIS}(X_{\boldsymbol{a}} \circ \pi^{-1}) = \mathrm{LIS}(X_{\boldsymbol{a}} \circ \pi^{-1}, id) = \mathrm{LIS}(X_{\boldsymbol{a}}, \pi),$$
$$\mathrm{LIS}(Y_{\boldsymbol{a}} \circ \tau^{-1}) = \mathrm{LIS}(Y_{\boldsymbol{a}} \circ \tau^{-1}, id) = \mathrm{LIS}(Y_{\boldsymbol{a}}, \tau).$$

Suppose the set $\{\sigma, \sigma \circ (r,t)\}$, with $\sigma(r) < \sigma(t)$, is chosen in the first step. If the outcome in the second step is tails, we verify that $X_{\boldsymbol{a}} \circ \pi^{-1} = Y_{\boldsymbol{a}} \circ \tau^{-1}$. When $X = \sigma$, $Y = \sigma \circ (r,t)$, by (6), we have

$$\begin{aligned} X_{\boldsymbol{a}} \circ \pi^{-1} &= \sigma_{\boldsymbol{a}} \circ \pi^{-1}, \\ Y_{\boldsymbol{a}} \circ \tau^{-1} &= (\sigma \circ (r,t))_{\boldsymbol{a}} \circ \pi^{-1} \circ s_i \\ &= (\sigma \circ (r,t))_{\boldsymbol{a}} \circ (\pi^{-1}(i), \pi^{-1}(i+1)) \circ \pi^{-1} \\ &= (\sigma \circ (r,t) \circ (r,t))_{\boldsymbol{a}} \circ \pi^{-1} \\ &= \sigma_{\boldsymbol{a}} \circ \pi^{-1}. \end{aligned}$$

When $X=\sigma\circ(r,t)$, $Y=\sigma$, again by (6), we have

$$\begin{aligned}X_{\boldsymbol{a}}\circ\pi^{-1}&=\big(\sigma\circ(r,t)\big)_{\boldsymbol{a}}\circ\pi^{-1}\\&=\sigma_{\boldsymbol{a}}\circ\big(\pi^{-1}(i),\pi^{-1}(i+1)\big)\circ\pi^{-1}\\&=\sigma_{\boldsymbol{a}}\circ\pi^{-1}\circ s_i,\\Y_{\boldsymbol{a}}\circ\tau^{-1}&=\sigma_{\boldsymbol{a}}\circ\pi^{-1}\circ s_i.\end{aligned}$$

If the outcome in the second step is heads, we have

$$X_{\boldsymbol{a}}\circ\pi^{-1}=\sigma_{\boldsymbol{a}}\circ\pi^{-1}\quad\text{and}\quad Y_{\boldsymbol{a}}\circ\tau^{-1}=\sigma_{\boldsymbol{a}}\circ\pi^{-1}\circ s_i.$$

Since $\sigma(r)<\sigma(t)$, that is, $\sigma(a_{\pi^{-1}(i)})<\sigma(a_{\pi^{-1}(i+1)})$, we have $\sigma_{\boldsymbol{a}}\circ\pi^{-1}(i)<\sigma_{\boldsymbol{a}}\circ\pi^{-1}(i+1)$. Hence $Y_{\boldsymbol{a}}\circ\tau^{-1}$ covers $X_{\boldsymbol{a}}\circ\pi^{-1}$ in $(S_k,\leq_R)$. (7) follows. □

REMARK. A special case of Lemma 3.4 is when $k=n$, in which the only choice for $\boldsymbol{a}$ is the vector $(1,2,\dots,n)$ whence $X_{\boldsymbol{a}}=X$, $Y_{\boldsymbol{a}}=Y$.

In Lemma 3.6, we prove a similar result for the case when $q\geq 1$, using the following property of Mallows permutations [cf. Lemma 2.2 in Bhatnagar and Peled (2015)].

PROPOSITION 3.5. *For any $n\geq 1$ and $q>0$, if $\pi\sim\mu_{n,q}$ then $\pi^r\sim\mu_{n,1/q}$ and $\pi^{-1}\sim\mu_{n,q}$.*

LEMMA 3.6. *Given $\pi,\tau\in S_k$ with $\pi\leq_L\tau$, for any $n\geq k$, $q\geq 1$ and increasing indices $\boldsymbol{a}=(a_1,a_2,\dots,a_k)$ with $a_i\in[n]$, there exists a coupling (X,Y) such that $X\sim\mu_{n,q}$, $Y\sim\mu_{n,q}$ and*

$$\operatorname{LIS}(X_{\boldsymbol{a}},\pi)\leq\operatorname{LIS}(Y_{\boldsymbol{a}},\tau).$$

PROOF. Given $\pi\in S_n$, recall that π^r denote the reversal of π. For any $\pi\in S_n$, we have $\operatorname{Inv}(\pi^r)=\{(i,j):1\leq i<j\leq n \text{ and } (n+1-j,n+1-i)\notin\operatorname{Inv}(\pi)\}$. Hence $\pi\leq_L\tau$ implies $\tau^r\leq_L\pi^r$. By Lemma 3.4, there exists a coupling (U,V) such that $U\sim\mu_{n,1/q}$, $V\sim\mu_{n,1/q}$ and

$$\operatorname{LIS}\big(U_{\boldsymbol{a}'},\pi^r\big)\leq\operatorname{LIS}\big(V_{\boldsymbol{a}'},\tau^r\big),$$

where $\boldsymbol{a}'=(a_1',\dots,a_k')$ with $a_i'=n+1-a_{k+1-i}$. Define $(X,Y):=(U^r,V^r)$. By Proposition 3.5, $X\sim\mu_{n,q}$, $Y\sim\mu_{n,q}$. Moreover, we have

$$\begin{aligned}\operatorname{LIS}(X_{\boldsymbol{a}},\pi)&=\operatorname{LIS}\big((X_{\boldsymbol{a}})^r,\pi^r\big)=\operatorname{LIS}\big((X^r)_{\boldsymbol{a}'},\pi^r\big)=\operatorname{LIS}\big(U_{\boldsymbol{a}'},\pi^r\big)\\&\leq\operatorname{LIS}\big(V_{\boldsymbol{a}'},\tau^r\big)=\operatorname{LIS}\big((Y^r)_{\boldsymbol{a}'},\tau^r\big)=\operatorname{LIS}\big((Y_{\boldsymbol{a}})^r,\tau^r\big)\\&=\operatorname{LIS}(Y_{\boldsymbol{a}},\tau).\end{aligned}$$

□

LEMMA 3.7. *Given increasing indices* $\boldsymbol{a} = (a_1, a_2, \dots, a_k)$ *with* $a_i \in [n]$, *for any* $0 < q \leq 1$ *and any distribution* ν *on* S_k, *there exists a coupling* (X, Y, Z) *such that the following holds*:

(a) *X and Y are independent.*
(b) *$X \sim \mu_{n,q}$, $Y \sim \nu$ and $Z \sim \mu_{n,q}$.*
(c) $\mathrm{LIS}(X_{\boldsymbol{a}}, Y) \leq \mathrm{LIS}(Z_{\boldsymbol{a}})$.

PROOF. Let id_k denote the identity in S_k. By the definition of weak Bruhat order, for any $\xi \in S_k$, we have $id_k \leq_L \xi$. Hence, given $\xi \in S_k$, by Lemma 3.4, there exists a coupling (U, V) such that $U \sim \mu_{n,q}$, $V \sim \mu_{n,q}$ and $\mathrm{LIS}(U_{\boldsymbol{a}}, \xi) \leq \mathrm{LIS}(V_{\boldsymbol{a}}, id_k) = \mathrm{LIS}(V_{\boldsymbol{a}})$. Then we construct the coupling (X, Y, Z) as follows:

- Sample Y according to the distribution ν.
- Conditioned on $Y = \xi$, (X, Z) has the same distribution as (U, V) defined above.

First, we point out that X and Y are independent. Since whatever value Y takes, the conditional distribution of X is $\mu_{n,q}$. Moreover, it can be seen that X, Y and Z thus defined have the correct marginal distributions. Finally, (c) holds by the construction of the coupling. □

We can prove a similar result for the case when $q \geq 1$.

LEMMA 3.8. *Given* $\boldsymbol{a} = (a_1, a_2, \dots, a_k)$, *where* $a_1 < \cdots < a_k$ *and* $a_i \in [n]$, *for any* $q \geq 1$ *and any distribution* ν *on* S_k, *there exists a coupling* (X, Y, Z) *such that the following holds*:

(a) *X and Y are independent.*
(b) *$X \sim \mu_{n,q}$, $Y \sim \nu$ and $Z \sim \mu_{n,q}$.*
(c) $\mathrm{LIS}(X_{\boldsymbol{a}}, Y) \geq \mathrm{LIS}(Z_{\boldsymbol{a}})$.

PROOF. The lemma follows by the same argument as in the proof of Lemma 3.7 except that here we use Lemma 3.6 instead of Lemma 3.4. □

LEMMA 3.9. *Given* $\boldsymbol{a} = (a_1, a_2, \dots, a_k)$, *where* $a_1 < \cdots < a_k$ *and* $a_i \in [n]$. *Define* $\bar{\boldsymbol{a}} := (n + 1 - a_k, n + 1 - a_{k-1}, \dots, n + 1 - a_1)$. *For any* $0 < q \leq 1$ *and any distribution* ν *on* S_k, *there exists a coupling* (X, Y, Z) *such that the following holds*:

(a) *X and Y are independent.*
(b) *$X \sim \mu_{n,q}$, $Y \sim \nu$ and $Z \sim \mu_{n,1/q}$.*
(c) $\mathrm{LIS}(X_{\boldsymbol{a}}, Y) \geq \mathrm{LIS}(Z_{\bar{\boldsymbol{a}}})$.

PROOF. Recall that π^r denotes the reversal of π. If $\pi \sim \nu$, we use ν^r to denote the distribution of π^r. Clearly, $\nu = (\nu^r)^r$. By Lemma 3.8, there exists a coupling (U, V, Z) such that:

- U and V are independent.
- $U \sim \mu_{n,1/q}$, $V \sim \nu^r$ and $Z \sim \mu_{n,1/q}$.
- $\mathrm{LIS}(U_{\bar{\boldsymbol{a}}}, V) \geq \mathrm{LIS}(Z_{\bar{\boldsymbol{a}}})$.

Define $X := U^r$ and $Y := V^r$. We have

$$\begin{aligned}
\mathrm{LIS}(U_{\bar{\boldsymbol{a}}}, V) &= \mathrm{LIS}\big(\big\{\big(U_{\bar{\boldsymbol{a}}}(i), V(i)\big)\big\}_{i\in[k]}\big) \\
&= \mathrm{LIS}\big(\big\{\big((U_{\bar{\boldsymbol{a}}})^r(i), V^r(i)\big)\big\}_{i\in[k]}\big) \\
&= \mathrm{LIS}\big(\big\{\big((U^r)_{\boldsymbol{a}}(i), V^r(i)\big)\big\}_{i\in[k]}\big) \\
&= \mathrm{LIS}\big(\big\{\big(X_{\boldsymbol{a}}(i), Y(i)\big)\big\}_{i\in[k]}\big) \\
&= \mathrm{LIS}(X_{\boldsymbol{a}}, Y). \qquad \square
\end{aligned}$$

4. Proof of Theorem 1.4. We start this section by introducing the following lemma which is analogous to Corollary 4.3 in Mueller and Starr (2013). That result shows that the LIS of a Mallows distributed permutation scaled by $n^{-1/2}$ can be bounded within the interval $(2e^{-|\beta|/2}, 2e^{|\beta|/2})$. We postpone the proof of Lemma 4.2 to the end of this paper.

DEFINITION 4.1. For any positive integer n and $m \in [n]$, define

$$Q(n,m) := \big\{(b_1, b_2, \dots, b_m) : b_i \in [n] \text{ and } b_i < b_{i+1} \text{ for all } i\big\}.$$

LEMMA 4.2. *Suppose that $\{q_n\}_{n=1}^\infty$ is a sequence such that $q_n > 0$ and $\lim_{n\to\infty} n(1-q_n) = \beta$ with $|\beta| < \ln 2$. For any sequence $\{k_n\}_{n=1}^\infty$ such that $k_n \in [n]$ and $\lim_{n\to\infty} k_n = \infty$, we have*

$$\lim_{n\to\infty} \max_{\boldsymbol{b}\in Q(n,k_n)} \mu_{n,q_n}\bigg(\pi \in S_n : \frac{\mathrm{LIS}(\pi_{\boldsymbol{b}})}{\sqrt{k_n}} \notin \big(2e^{\frac{-|\beta|}{2}} - \varepsilon, 2e^{\frac{|\beta|}{2}} + \varepsilon\big)\bigg) = 0$$

for any $\varepsilon > 0$.

4.1. *The scale of* $\mathrm{LIS}(\pi, \tau)$ *within a rectangle.* We introduce the following way to sample a permutation according to $\mu_{n,q}$ which will be used in the proofs. Given $\boldsymbol{c} = (c_1, c_2, \dots, c_m)$, where $c_i \in \mathbb{Z}^+$ and $\sum_{i=1}^m c_i = n$, define

$$d_0 := 0, \qquad d_k := \sum_{i=1}^k c_i \qquad \forall k \in [m],$$

$$A(\boldsymbol{c}) := \big\{(A_1, A_2, \dots, A_m) : \{A_i\}_{i\in[m]} \text{ is a partition of } [n], |A_i| = c_i\big\}.$$

Given $(A_1, \dots, A_m) \in A(\boldsymbol{c})$, define the inversion number of $(A_1, \dots, A_m)$ as follows:

$$\begin{aligned}
&l\big((A_1, \dots, A_m)\big) \\
&\quad := \big|\big\{(x, y) : x < y \text{ and there exists } i > j \text{ such that } x \in A_i, y \in A_j\big\}\big|.
\end{aligned}$$

Let $\boldsymbol{a}_i$ be the vector which consists of the numbers in A_i in increasing order. There exists a bijection $f_{\boldsymbol{c}}$ between S_n and $A(\boldsymbol{c}) \times S_{c_1} \times S_{c_2} \times \cdots \times S_{c_m}$ such that, for any $\pi \in S_n$, $f_{\boldsymbol{c}}(\pi) = ((A_1, A_2, \ldots, A_m), \tau_1, \tau_2, \ldots, \tau_m)$ if and only if

$$\{\pi(j) : j \in A_i\} = \{d_{i-1}+1, d_{i-1}+2, \ldots, d_i\}, \qquad \pi_{\boldsymbol{a}_i} = \tau_i \qquad \forall i \in [m].$$

In other words, set A_i consists of those indices j such that $\pi(j) \in [d_{i-1}+1, d_i]$ and τ_i denotes the relative ordering of $\{d_{i-1}+1, \ldots, d_i\}$ in π. For example, given $(A_1, A_2, A_3) = (\{1, 5, 6\}, \{2, 4, 9\}, \{3, 7, 8\})$ and $\tau_1 = (1, 3, 2)$, $\tau_2 = (2, 3, 1)$, $\tau_3 = (3, 2, 1)$, the corresponding permutation π under the bijection $f_{\boldsymbol{c}}$ is $(1, 5, 9, 6, 3, 2, 8, 7, 4)$. From the definition above, it is not hard to see that the following relation holds:

$$l(\pi) = l\big((A_1, A_2, \ldots, A_m)\big) + \sum_{i=1}^{m} l(\tau_i). \tag{8}$$

Define the random variable $X_{\boldsymbol{c}}$ which takes value in $A(\boldsymbol{c})$ such that

$$\mathbb{P}\big(X_{\boldsymbol{c}} = (A_1, A_2, \ldots, A_m)\big) \propto q^{l((A_1, A_2, \ldots, A_m))}.$$

Independent of $X_{\boldsymbol{c}}$, let $Y_1, Y_2, \ldots, Y_m$ be independent random variables such that, for any $i \in [m]$, $Y_i \sim \mu_{c_i,q}$. Define $Z := f_{\boldsymbol{c}}^{-1}(X_{\boldsymbol{c}}, Y_1, Y_2, \ldots, Y_m)$. By (8), we have $Z \sim \mu_{n,q}$, since

$$\mathbb{P}(Z = \pi) \propto q^{l(\pi)}.$$

As our last step in preparation for the proof of Lemma 4.4, we introduce the following elementary result in analysis.

LEMMA 4.3. *Suppose $\{B_i\}_{i=1}^{\infty}$ is a partition of $\mathbb{N}$, that is, $\bigcup_{i=1}^{\infty} B_i = \mathbb{N}$ and $B_i \cap B_j = \varnothing, \forall i \neq j$. Moreover, each B_i is a finite nonempty set. Given a sequence $\{x_i\}_{i=1}^{\infty}$, if $\lim_{n\to\infty} x_{b_n} = a$, for any sequence $\{b_i\}_{i=1}^{\infty}$ with $b_i \in B_i$, then we have $\lim_{n\to\infty} x_n = a$.*

PROOF. We prove the lemma by contradiction. Suppose $\lim_{n\to\infty} x_n = a$ does not hold. Then there exists $\varepsilon > 0$ and a subsequence $\{x_{n_j}\}_{j=1}^{\infty}$ such that $x_{n_j} \notin (a-\varepsilon, a+\varepsilon)$ for all j. We can construct a sequence $\{b_i\}_{i=1}^{\infty}$ with $b_i \in B_i$, such that $x_{b_i} \notin (a-\varepsilon, a+\varepsilon)$ infinitely often. Specifically, we define the sequence $\{b_i\}_{i=1}^{\infty}$ as follows. For each i, if there exists an $n_j \in B_i$, let $b_i = n_j$; otherwise, let b_i be the smallest number in B_i. Thus, we get the contradiction. □

For any $\pi, \tau \in S_n$, define $z(\pi, \tau) := \{(\frac{\pi(i)}{n}, \frac{\tau(i)}{n})\}_{i\in[n]}$. Given a rectangle $R \subset [0, 1] \times [0, 1]$, let $l_R(\pi, \tau)$ denote the length of the longest increasing subsequence of $z(\pi, \tau)$ within R. The following lemma addresses the size of the LIS in in a small rectangle and this result will be the most crucial building block in the proof of Theorem 1.4.

LEMMA 4.4. *Let* $R = (x_1, x_2] \times (y_1, y_2] \subset [0,1] \times [0,1]$. *Under the same conditions as in Lemma* 1.3, *if* $\Delta x|\beta| < \ln 2$, *we have*

$$\lim_{n\to\infty} \mathbb{P}_n\left(\frac{l_R(\pi,\tau)}{\sqrt{n\rho(R)}} \in \left(2e^{-\Delta x|\beta|/2} - \varepsilon, 2e^{\Delta x|\beta|/2} + \varepsilon\right)\right) = 1, \tag{9}$$

for any $\varepsilon > 0$, *where* $\rho(R) := \iint_R \rho(x,y)\,dx\,dy$ *and* $\Delta x := x_2 - x_1$.

PROOF. To simplify the proof, we divide the lemma into the following three cases:

Case 1: $\beta > 0$ or $\beta = 0$ and $q_n \le 1$ when n is sufficiently large.
Case 2: $\beta < 0$ or $\beta = 0$ and $q_n \ge 1$ when n is sufficiently large.
Case 3: $\beta = 0$.

First, Case 3 follows from Case 1 and Case 2 because if $\lim_{n\to\infty} n(1-q_n) = 0$, we can divide the sequence $\{q_n\}_{n=1}^\infty$ into two disjoint subsequences such that one of them falls into Case 1 and the other falls into Case 2.

Next, we argue that Case 2 follows from Case 1. If $\pi \sim \mu_{n,q}$, by Proposition 3.5, we have $\pi^r \sim \mu_{n,1/q}$. Since $\lim_{n\to\infty} n(1-q_n) = \beta \in \mathbb{R}$, we have $\lim_{n\to\infty} q_n = 1$. Hence

$$\lim_{n\to\infty} n(1 - 1/q_n) = \lim_{n\to\infty} n(q_n - 1)/q_n = -\beta.$$

Therefore, Case 2 follows from Case 1 by considering the reversal of π and τ in (9). Specifically, if $\pi \sim \mu_{n,q_n}$ and $\tau \sim \mu_{n,q'_n}$, after reversing, we have $\pi^r \sim \mu_{n,1/q_n}$ and $\tau^r \sim \mu_{n,1/q'_n}$ and the n points induced by π and τ do not change, that is, $z(\pi,\tau) = z(\pi^r, \tau^r)$.

To prove Case 1, in the following, we assume $x_1, y_1 > 0$ and $x_2, y_2 < 1$. The proofs for the cases when $x_1 = 0$ or $y_1 = 0$ or $x_2 = 1$ or $y_2 = 1$ are similar. Let $x_3 = y_3 = 1$. Given $n \in \mathbb{N}$, we will sample (π, τ) according to $\mathbb{P}_n$ by the method introduced before Lemma 4.3. Define

$$\begin{aligned} d_{n,i} &:= \lfloor nx_i \rfloor, & c_{n,i} &:= d_{n,i} - d_{n,i-1} & &\text{for } i = 1,2,3,\\ d'_{n,i} &:= \lfloor ny_i \rfloor, & c'_{n,i} &:= d'_{n,i} - d'_{n,i-1} & &\text{for } i = 1,2,3, \end{aligned}$$

where we assume that $d_{n,0} = d'_{n,0} = 0$. Then it is trivial that

$$\begin{aligned} d_{n,i} &= \left|\left\{ j \in [n] : \frac{j}{n} \in (0, x_i]\right\}\right|, & c_{n,2} &= \left|\left\{ j \in [n] : \frac{j}{n} \in (x_1, x_2]\right\}\right|,\\ d'_{n,i} &= \left|\left\{ j \in [n] : \frac{j}{n} \in (0, y_i]\right\}\right|, & c'_{n,2} &= \left|\left\{ j \in [n] : \frac{j}{n} \in (y_1, y_2]\right\}\right|. \end{aligned}$$

Since $\lim_{n\to\infty} \frac{\lfloor nx \rfloor}{n} = x, \forall x \in \mathbb{R}$, it follows that $\lim_{n\to\infty} \frac{d_{n,i}}{n} = x_i$. Hence

$$\lim_{n\to\infty} \frac{c_{n,2}}{n} = x_2 - x_1 = \Delta x. \tag{10}$$

Next, for any nonnegative integer i, define $B_i := \{n \in \mathbb{N} : c_{n,2} = i\}$. Clearly, $\{B_i\}_{i=0}^{\infty}$ thus defined is a partition of $\mathbb{N}$ and we show that each B_i is a nonempty finite set. Since, by (10), $\lim_{n\to\infty} c_{n,2} = \infty$, we conclude that each B_i is a finite set. From the definition of $d_{n,i}$, it is easily seen that the sequence $\{d_{n,1}\}$ is nondecreasing and the increment of consecutive terms is either 0 or 1. The same is true for the sequence $\{d_{n,2}\}$. Hence we have

$$|c_{n+1,2} - c_{n,2}| = \big|d_{n+1,2} - d_{n,2} - (d_{n+1,1}, -d_{n,1})\big| \le 1.$$

Since $c_{1,2} \in B_0$ and $\lim_{n\to\infty} c_{n,2} = \infty$, the inequality above guarantees that each B_i is nonempty. Next, define $\boldsymbol{c}_n := (c_{n,1}, c_{n,2}, c_{n,3})$ and $\boldsymbol{c}'_n := (c'_{n,1}, c'_{n,2}, c'_{n,3})$. Define $X_{\boldsymbol{c}_n}$ which takes values in $A(\boldsymbol{c}_n)$ such that

$$\mathbb{P}\big(X_{\boldsymbol{c}_n} = (A_1, A_2, A_3)\big) \propto q_n^{l((A_1,A_2,A_3))} \qquad \forall (A_1, A_2, A_3) \in A(\boldsymbol{c}_n).$$

Independently, define three independent random variables $Y_{n,1}, Y_{n,2}, Y_{n,3}$ such that $Y_{n,i} \sim \mu_{c_{n,i},q_n}$. Independent of all the variables defined above, define $X_{\boldsymbol{c}'_n}$ and $Y'_{n,1}, Y'_{n,2}, Y'_{n,3}$ in the same fashion. That is, $X_{\boldsymbol{c}'_n}$ takes value in $A(\boldsymbol{c}'_n)$ with

$$\mathbb{P}\big(X_{\boldsymbol{c}'_n} = \big(A'_1, A'_2, A'_3\big)\big) \propto \big(q'_n\big)^{l((A'_1,A'_2,A'_3))} \qquad \forall \big(A'_1, A'_2, A'_3\big) \in A\big(\boldsymbol{c}'_n\big)$$

and $Y'_{n,1}, Y'_{n,2}, Y'_{n,3}$ are three independent random variables with $Y'_{n,i} \sim \mu_{c'_{n,i},q'_n}$. Define

$$\pi := f_{\boldsymbol{c}_n}^{-1}(X_{\boldsymbol{c}_n}, Y_{n,1}, Y_{n,2}, Y_{n,3}), \qquad \tau := f_{\boldsymbol{c}'_n}^{-1}\big(X_{\boldsymbol{c}'_n}, Y'_{n,1}, Y'_{n,2}, Y'_{n,3}\big).$$

From the discussion before Lemma 4.3, it follows that (π, τ) thus defined has distribution $\mathbb{P}_n$. Moreover, given $X_{\boldsymbol{c}_n} = (A_1, A_2, A_3)$ and $X_{\boldsymbol{c}'_n} = (A'_1, A'_2, A'_3)$, we have

$$A_2 = \left\{i \in [n] : \frac{\pi(i)}{n} \in (x_1, x_2]\right\}, \qquad A'_2 = \left\{i \in [n] : \frac{\tau(i)}{n} \in (y_1, y_2]\right\}.$$

Hence we have

$$A_2 \cap A'_2 = \left\{i \in [n] : \left(\frac{\pi(i)}{n}, \frac{\tau(i)}{n}\right) \in R\right\}. \tag{11}$$

Define $M = |z(\pi, \tau) \cap R|$, that is, M denotes the number of points $\{(\frac{\pi(i)}{n}, \frac{\tau(i)}{n})\}_{i=1}^{n}$ within R. Then, by (11), we have $M = |A_2 \cap A'_2|$. Hence M only depends on the values of $X_{\boldsymbol{c}_n}$ and $X_{\boldsymbol{c}'_n}$ and is independent of $\bigcup_{i\in[3]}\{Y_{n,i}, Y'_{n,i}\}$. Next, note that, conditioning on $X_{\boldsymbol{c}_n} = (A_1, A_2, A_3)$ and $X_{\boldsymbol{c}'_n} = (A'_1, A'_2, A'_3)$, $l_R(\pi, \tau)$ is determined by $Y_{n,2}$ and $Y'_{n,2}$. To see this, we first define a new function I as follows, given any finite set $A \subset \mathbb{Z}$ and any $a \in A$, define $I(A, a) := k$ if a is the kth smallest number in A. Suppose $A_2 \cap A'_2 = \{a_j\}_{j\in[M]}$ with $a_1 < a_2 < \cdots < a_M$. Define $\boldsymbol{b} \in Q(c_{n,2}, M)$ and $\boldsymbol{b}' \in Q(c'_{n,2}, M)$ by

$$\begin{aligned} \boldsymbol{b} &:= \big(I(A_2, a_1), I(A_2, a_2), \ldots, I(A_2, a_M)\big), \\ \boldsymbol{b}' &:= \big(I\big(A'_2, a_1\big), I\big(A'_2, a_2\big), \ldots, I\big(A'_2, a_M\big)\big). \end{aligned} \tag{12}$$

Note that $\boldsymbol{b}$ and $\boldsymbol{b}'$ are determined by A_2 and A_2'. Then we have

$$l_R(\pi,\tau) = \mathrm{LIS}\big((Y_{n,2})_{\boldsymbol{b}}, (Y_{n,2}')_{\boldsymbol{b}'}\big). \tag{13}$$

Indeed, conditioning on $X_{\boldsymbol{c}_n} = (A_1, A_2, A_3)$, we know that $\{\pi(i) : i \in A_2\} = \{d_{n,1}+1, d_{n,1}+2, \dots, d_{n,2}\}$. And the value of $Y_{n,2}$ determines the relative ordering of $\pi(i)$ for those $i \in A_2$. Similarly, the value of $Y_{n,2}'$ determines the relative ordering of $\tau(i)$ for those $i \in A_2'$.

Now we are in the position to prove (9) for Case 1. From the discussion above and Lemma 4.3, it suffices to show that, for any sequence $\{s_n\}_{n=1}^{\infty}$ with $s_n \in B_n$, that is, when $c_{s_n,2} = n$, we have

$$\lim_{n\to\infty} \mathbb{P}_{s_n}\left(\frac{l_R(\pi,\tau)}{\sqrt{s_n\rho(R)}} \in \big(2e^{-\Delta x\beta/2} - \varepsilon, 2e^{\Delta x\beta/2} + \varepsilon\big)\right) = 1, \tag{14}$$

for any $\varepsilon > 0$. Note that by the definition of $\mathbb{P}_{s_n}$ in Lemma 1.3, π and τ above are of size s_n with $\pi \sim \mu_{s_n,q_{s_n}}$, $\tau \sim \mu_{s_n,q_{s_n}'}$.

We separate the proof of (14) into two parts. Specifically, we need to show that, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_{s_n}\left(\frac{l_R(\pi,\tau)}{\sqrt{s_n\rho(R)}} < 2e^{\Delta x\beta/2} + \varepsilon\right) = 1, \tag{15}$$

and

$$\lim_{n\to\infty} \mathbb{P}_{s_n}\left(\frac{l_R(\pi,\tau)}{\sqrt{s_n\rho(R)}} > 2e^{-\Delta x\beta/2} - \varepsilon\right) = 1. \tag{16}$$

Since $\{s_n\}_{n\ge1}$ is a subsequence of $\{i\}_{i\ge0}$, $\lim_{n\to\infty} s_n = \infty$. Hence, by (10) and the fact that $c_{s_n,2} = n$, we get

$$\lim_{n\to\infty}\frac{n}{s_n} = \lim_{n\to\infty}\frac{c_{s_n,2}}{s_n} = \Delta x.$$

Thus,

$$\lim_{n\to\infty} n(1-q_{s_n}) = \lim_{n\to\infty}\frac{n}{s_n}s_n(1-q_{s_n}) = \Delta x\beta < \ln 2. \tag{17}$$

To prove (15), for any $\varepsilon > 0$, we can choose $\varepsilon_1 > 0$ sufficiently small such that

$$(1-\varepsilon_1)\big(2e^{\Delta x\beta/2} + \varepsilon\big) > 2e^{\Delta x\beta/2}. \tag{18}$$

For this fixed ε_1, we can choose $\delta > 0$ such that

$$\sqrt{\frac{\rho(R)}{\rho(R)+\delta}} > 1 - \varepsilon_1. \tag{19}$$

Given $n \in \mathbb{N}$, define $k_n = \lfloor s_n(\rho(R)+\delta)\rfloor$. Clearly, we have $\lim_{n\to\infty} k_n = \infty$. Hence, by Lemma 4.2, (17) and (18), there exists $N_1 > 0$ such that, for any $n > N_1$, we have

$$\min_{\boldsymbol{b}\in Q(n,k_n)} \mu_{n,q_{s_n}}\left(\eta \in S_n : \frac{\mathrm{LIS}(\eta_{\boldsymbol{b}})}{\sqrt{k_n}} < (1-\varepsilon_1)\big(2e^{\Delta x\beta/2} + \varepsilon\big)\right) > 1 - \varepsilon. \tag{20}$$

Given $\boldsymbol{b} \in Q(n, k_n)$, for any $\boldsymbol{b}'$ which is a subsequence of $\boldsymbol{b}$, we have $\mathrm{LIS}(\eta_{\boldsymbol{b}}) \geq \mathrm{LIS}(\eta_{\boldsymbol{b}'})$. Thus we can make (20) stronger as follows:

$$\min_{\boldsymbol{b}\in\bar{Q}(n,k_n)} \mu_{n,q_{s_n}}\left(\eta \in S_n : \frac{\mathrm{LIS}(\eta_{\boldsymbol{b}})}{\sqrt{k_n}} < (1-\varepsilon_1)\big(2e^{\Delta x\beta/2}+\varepsilon\big)\right) > 1-\varepsilon, \tag{21}$$

where $\bar{Q}(n, k_n) = \bigcup_{i\in[k_n]} Q(n, i)$. Since $\lim_{n\to\infty} s_n = \infty$, we have

$$\lim_{n\to\infty} s_n(1-q_{s_n}) = \beta \quad \text{and} \quad \lim_{n\to\infty} s_n\big(1-q'_{s_n}\big) = \gamma. \tag{22}$$

Hence, by Lemma 1.3, there exists $N_2 > 0$ such that, for any $n > N_2$, we have

$$\mathbb{P}_{s_n}\left(\frac{|z(\pi,\tau)\cap R|}{s_n} \leq \rho(R)+\delta\right) > 1-\varepsilon. \tag{23}$$

In the following, let $E_n(A_2, A_2')$ denote the event that the second entries of $X_{\boldsymbol{c}_{s_n}}$ and $X_{\boldsymbol{c}'_{s_n}}$ are A_2 and A_2', respectively. Let $\mathbb{P}$ denote the probability space on which $(X_{\boldsymbol{c}_{s_n}}, Y_{s_n,1}, Y_{s_n,2}, Y_{s_n,3})$ and $(X_{\boldsymbol{c}'_{s_n}}, Y'_{s_n,1}, Y'_{s_n,2}, Y'_{s_n,3})$ are defined. Then, for any $n > \max(N_1, N_2)$, we have

$$\begin{aligned}
&\mathbb{P}_{s_n}\left(\frac{l_R(\pi,\tau)}{\sqrt{s_n\rho(R)}} < 2e^{\Delta x\beta/2}+\varepsilon\right)\\
&\quad\geq \sum_{|A_2\cap A_2'|\leq k_n} \mathbb{P}\left(\frac{l_R(\pi,\tau)}{\sqrt{s_n\rho(R)}} < 2e^{\Delta x\beta/2}+\varepsilon \,|\, E_n\big(A_2,A_2'\big)\right)\times \mathbb{P}\big(E_n\big(A_2,A_2'\big)\big)\\
&\quad= \sum_{|A_2\cap A_2'|\leq k_n} \mathbb{P}\left(\frac{\mathrm{LIS}((Y_{s_n,2})_{\boldsymbol{b}}, (Y'_{s_n,2})_{\boldsymbol{b}'})}{\sqrt{s_n\rho(R)}} < 2e^{\Delta x\beta/2}+\varepsilon\right)\times \mathbb{P}\big(E_n\big(A_2,A_2'\big)\big)\\
&\quad\geq \sum_{|A_2\cap A_2'|\leq k_n} \mu_{n,q_{s_n}}\left(\frac{\mathrm{LIS}(\eta_{\boldsymbol{b}})}{\sqrt{s_n\rho(R)}} < 2e^{\Delta x\beta/2}+\varepsilon\right)\times \mathbb{P}\big(E_n\big(A_2,A_2'\big)\big)\\
&\quad= \sum_{|A_2\cap A_2'|\leq k_n} \mu_{n,q_{s_n}}\left(\frac{\mathrm{LIS}(\eta_{\boldsymbol{b}})}{\sqrt{s_n(\rho(R)+\delta)}} < \frac{\sqrt{\rho(R)}}{\sqrt{\rho(R)+\delta}}\big(2e^{\Delta x\beta/2}+\varepsilon\big)\right)\\
&\qquad\times \mathbb{P}\big(E_n\big(A_2,A_2'\big)\big)\\
&\quad\geq \sum_{|A_2\cap A_2'|\leq k_n} \mu_{n,q_{s_n}}\left(\frac{\mathrm{LIS}(\eta_{\boldsymbol{b}})}{\sqrt{k_n}} < (1-\varepsilon_1)\big(2e^{\Delta x\beta/2}+\varepsilon\big)\right)\times \mathbb{P}\big(E_n\big(A_2,A_2'\big)\big)\\
&\quad\geq (1-\varepsilon)\times \sum_{|A_2\cap A_2'|\leq k_n} \mathbb{P}\big(E_n\big(A_2,A_2'\big)\big)\\
&\quad= (1-\varepsilon)\times \mathbb{P}_{s_n}\big(\big|z(\pi,\tau)\cap R\big| \leq k_n\big)\\
&\quad= (1-\varepsilon)\times \mathbb{P}_{s_n}\big(\big|z(\pi,\tau)\cap R\big| \leq s_n\big(\rho(R)+\delta\big)\big)\\
&\quad> (1-\varepsilon)^2.
\end{aligned}$$

The first equality follows by (13) and the independence of $(X_{\boldsymbol{c}_{s_n}}, X_{\boldsymbol{c}'_{s_n}})$ and $(Y_{s_n,2}, Y'_{s_n,2})$. Note that $\boldsymbol{b}$ and $\boldsymbol{b}'$ are determined by A_2 and A'_2 as in (12). The second inequality follows by Lemma 3.7, since $Y_{s_n,2}$ and $Y'_{s_n,2}$ are independent with $Y_{s_n,2} \sim \mu_{n,q_{s_n}}$. The third inequality follows by (19) and the fact that $k_n = \lfloor s_n(\rho(R) + \delta)\rfloor \le s_n(\rho(R) + \delta)$. The fourth inequality follows by (21) and the fact that the dimension of $\boldsymbol{b}$ equals to $|A_2 \cap A'_2|$. The last inequality follows by (23). Hence (15) follows.

The proof of (16) follows in a similar way as the proof of (15). First, by (17) and the fact that $\lim_{n\to\infty} q_n = 1$, we have

$$\lim_{n\to\infty} n(1 - 1/q_{s_n}) = \lim_{n\to\infty} \frac{n(q_{s_n} - 1)}{q_{s_n}} = -\Delta x\beta > -\ln 2. \tag{24}$$

For any $\varepsilon > 0$, we can choose $\varepsilon_1 > 0$ sufficiently small such that

$$(1 + \varepsilon_1)\big(2e^{-\Delta x\beta/2} - \varepsilon\big) < 2e^{-\Delta x\beta/2}. \tag{25}$$

For this fixed ε_1, we can choose $\delta > 0$ such that

$$\sqrt{\frac{\rho(R)}{\rho(R) - \delta}} < 1 + \varepsilon_1. \tag{26}$$

Given $n \in \mathbb{N}$, define $k'_n = \lceil s_n(\rho(R) - \delta)\rceil$. Clearly, we have $\lim_{n\to\infty} k'_n = \infty$. Moreover, under conditions of Case 1, $1/q_n \ge 1$ for sufficiently large n. Hence, by Lemma 4.2, (24) and (25), there exist $N_3 > 0$ such that, for any $n > N_3$, we have

$$\min_{\boldsymbol{b}\in Q(n,k'_n)} \mu_{n,1/q_{s_n}}\left(\eta \in S_n : \frac{\mathrm{LIS}(\eta_{\boldsymbol{b}})}{\sqrt{k'_n}} > (1 + \varepsilon_1)\big(2e^{-\Delta x\beta/2} - \varepsilon\big)\right) > 1 - \varepsilon. \tag{27}$$

Given $\boldsymbol{b} \in Q(n, k'_n)$, for any $\boldsymbol{b}'$ such that $\boldsymbol{b}$ is a subsequence of $\boldsymbol{b}'$, we have $\mathrm{LIS}(\eta_{\boldsymbol{b}}) \le \mathrm{LIS}(\eta_{\boldsymbol{b}'})$. Thus we can make (27) stronger as follows:

$$\min_{\boldsymbol{b}\in \hat{Q}(n,k'_n)} \mu_{n,1/q_{s_n}}\left(\eta \in S_n : \frac{\mathrm{LIS}(\eta_{\boldsymbol{b}})}{\sqrt{k'_n}} > (1 + \varepsilon_1)\big(2e^{-\Delta x\beta/2} - \varepsilon\big)\right) > 1 - \varepsilon, \tag{28}$$

where $\hat{Q}(n, k'_n) = \bigcup_{k'_n \le i \le n} Q(n, i)$. By (22) and Lemma 1.3, there exists $N_4 > 0$ such that, for any $n > N_4$, we have

$$\mathbb{P}_{s_n}\left(\frac{|z(\pi, \tau) \cap R|}{s_n} \ge \rho(R) - \delta\right) > 1 - \varepsilon. \tag{29}$$

Then, assuming the notation defined in the proof of (15), for any $n > \max(N_3, N_4)$, we have

$$\begin{aligned}
&\mathbb{P}_{s_n}\left(\frac{l_R(\pi, \tau)}{\sqrt{s_n\rho(R)}} > 2e^{-\Delta x\beta/2} - \varepsilon\right) \\
&\quad\ge \sum_{|A_2\cap A'_2|\ge k_n} \mathbb{P}\left(\frac{l_R(\pi, \tau)}{\sqrt{s_n\rho(R)}} > 2e^{-\Delta x\beta/2} - \varepsilon \,|\, E_n(A_2, A'_2)\right) \times \mathbb{P}\big(E_n(A_2, A'_2)\big)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{|A_2 \cap A_2'| \geq k_n'} \mathbb{P}\left(\frac{\mathrm{LIS}((Y_{s_n,2})_{\boldsymbol{b}}, (Y'_{s_n,2})_{\boldsymbol{b}'})}{\sqrt{s_n \rho(R)}} > 2e^{-\Delta x \beta/2} - \varepsilon\right) \times \mathbb{P}\big(E_n(A_2, A_2')\big) \\
&\geq \sum_{|A_2 \cap A_2'| \geq k_n'} \mu_{n,1/q_{s_n}}\left(\frac{\mathrm{LIS}(\eta_{\bar{\boldsymbol{b}}})}{\sqrt{s_n \rho(R)}} > 2e^{-\Delta x \beta/2} - \varepsilon\right) \times \mathbb{P}\big(E_n(A_2, A_2')\big) \\
&= \sum_{|A_2 \cap A_2'| \geq k_n'} \mu_{n,1/q_{s_n}}\left(\frac{\mathrm{LIS}(\eta_{\bar{\boldsymbol{b}}})}{\sqrt{s_n(\rho(R) - \delta)}} > \frac{\sqrt{\rho(R)}}{\sqrt{\rho(R) - \delta}}\big(2e^{-\Delta x \beta/2} - \varepsilon\big)\right) \\
&\quad \times \mathbb{P}\big(E_n(A_2, A_2')\big) \\
&\geq \sum_{|A_2 \cap A_2'| \geq k_n'} \mu_{n,1/q_{s_n}}\left(\frac{\mathrm{LIS}(\eta_{\bar{\boldsymbol{b}}})}{\sqrt{k_n'}} > (1 + \varepsilon_1)\big(2e^{-\Delta x \beta/2} - \varepsilon\big)\right) \\
&\quad \times \mathbb{P}\big(E_n(A_2, A_2')\big) \\
&\geq (1 - \varepsilon) \times \sum_{|A_2 \cap A_2'| \geq k_n'} \mathbb{P}\big(E_n(A_2, A_2')\big) \\
&= (1 - \varepsilon) \times \mathbb{P}_{s_n}\big(\big|z(\pi, \tau) \cap R\big| \geq k_n'\big) \\
&= (1 - \varepsilon) \times \mathbb{P}_{s_n}\big(\big|z(\pi, \tau) \cap R\big| \geq s_n\big(\rho(R) - \delta\big)\big) \\
&> (1 - \varepsilon)^2.
\end{aligned}
$$

The first equality follows by (13) and the independence of $(X_{\boldsymbol{c}_{s_n}}, X_{\boldsymbol{c}'_{s_n}})$ and $(Y_{s_n,2}, Y'_{s_n,2})$. The second inequality follows by Lemma 3.9, since $Y_{s_n,2}$ and $Y'_{s_n,2}$ are independent with $Y_{s_n,2} \sim \mu_{n,q_{s_n}}$. The third inequality follows by (26) and the fact that $k_n' = \lceil s_n(\rho(R) - \delta) \rceil \geq s_n(\rho(R) - \delta)$. The fourth inequality follows by (28) and the fact that $\bar{\boldsymbol{b}}$ has the same dimension as of $\boldsymbol{b}$ which equals to $|A_2 \cap A_2'|$. The last inequality follows by (29). Hence, (16) follows and this completes the proof of Lemma 4.4. □

4.2. *Deuschel and Zeitouni's approach.* The following lemma establishes a certain degree of smoothness of the densities u and ρ defined in Lemma 1.3.

LEMMA 4.5. *The density functions $u(x, y, \beta)$ defined in (3) and $\rho(x, y)$ defined in (2) satisfy the following*:

(a) $e^{-|\beta|} \leq u(x, y, \beta) \leq e^{|\beta|}$, $e^{-|\beta| - |\gamma|} \leq \rho(x, y) \leq e^{|\beta| + |\gamma|}$,
(b) $u(x, y, \beta) \in C_b^1$, $\rho(x, y) \in C_b^1$,
(c) $\max(|\frac{\partial u}{\partial x}|, |\frac{\partial u}{\partial y}|) \leq |\beta| e^{|\beta|}$,
(d) $\max(|\frac{\partial \rho}{\partial x}|, |\frac{\partial \rho}{\partial y}|) \leq (|\beta| + |\gamma|) e^{|\beta| + |\gamma|}$,

where $(x, y) \in [0, 1] \times [0, 1]$.

PROOF. First, we show that $e^{-|\beta|} \le u(x, y, \beta) \le e^{|\beta|}$ for any $0 \le x, y \le 1$. Here, we assume $\beta > 0$. The proof for the case when $\beta < 0$ is similar. By (3), we have

$$
\begin{aligned}
u(x, y, \beta) &= \frac{(\beta/2)\sinh(\beta/2)}{(e^{\beta/4}\cosh(\beta[x-y]/2) - e^{-\beta/4}\cosh(\beta[x+y-1]/2))^2} \\
&= \frac{\beta(e^{\beta}-1)}{(2e^{\beta/2}\cosh(\beta[x-y]/2) - 2\cosh(\beta[x+y-1]/2))^2}.
\end{aligned} \tag{30}
$$

Since $-1 \le x - y \le 1$ and $-1 \le x + y - 1 \le 1$, we have

$$2e^{\beta/2} \le 2e^{\beta/2}\cosh\bigl(\beta[x-y]/2\bigr) \le e^{\beta} + 1, \tag{31}$$

$$2 \le 2\cosh\bigl(\beta[x+y-1]/2\bigr) \le e^{\beta/2} + e^{-\beta/2}. \tag{32}$$

Since $e^{\beta/2} + e^{-\beta/2} < 2e^{\beta/2}$, from (31) and (32), we have

$$e^{\beta/2} - e^{-\beta/2} \le 2e^{\beta/2}\cosh\bigl(\beta[x-y]/2\bigr) - 2\cosh\bigl(\beta[x+y-1]/2\bigr) \le e^{\beta} - 1. \tag{33}$$

By (30) and (33), it follows that

$$\frac{\beta}{e^{\beta}-1} \le u(x, y, \beta) \le \frac{\beta(e^{\beta}-1)}{(e^{\beta/2} - e^{-\beta/2})^2}. \tag{34}$$

It is easily verified that

$$\frac{\beta}{e^{\beta}-1} \ge e^{-\beta} \quad \Longleftrightarrow \quad e^{-\beta} \ge 1 - \beta, \tag{35}$$

$$\frac{\beta(e^{\beta}-1)}{(e^{\beta/2} - e^{-\beta/2})^2} \le e^{\beta} \quad \Longleftrightarrow \quad \bigl(e^{\beta}-1\bigr)\bigl(e^{\beta} - 1 - \beta\bigr) \ge 0. \tag{36}$$

By the inequality $e^x \ge 1 + x$, the right-hand side of (35) and (36) hold. It follows from (34) and the left-hand side of (35) and (36) that

$$e^{-\beta} \le u(x, y, \beta) \le e^{\beta} \qquad \forall 0 \le x, y \le 1.$$

By the definition of $\rho(x, y)$, it follows trivially that

$$e^{-|\beta|-|\gamma|} \le \rho(x, y) \le e^{|\beta|+|\gamma|} \qquad \forall 0 \le x, y \le 1.$$

In Starr (2009), he shows that $\frac{\partial^2 \ln u(x,y,\beta)}{\partial x\, \partial y} = 2\beta u(x, y, \beta)$. Thus

$$\int_0^x u(t, y, \beta)\, dt = \frac{1}{2\beta}\left(\frac{\partial \ln u(x, y, \beta)}{\partial y} - \frac{\partial \ln u(0, y, \beta)}{\partial y}\right). \tag{37}$$

By direct calculation, we have $u(1, y, \beta) = \frac{\beta e^{\beta y}}{e^{\beta}-1}$, $u(0, y, \beta) = \frac{\beta e^{-\beta y}}{1-e^{-\beta}}$. Therefore, we get $\frac{\partial \ln u(1,y,\beta)}{\partial y} = \beta$ and $\frac{\partial \ln u(0,y,\beta)}{\partial y} = -\beta$. By (37), it follows that

$$\frac{\partial u(x, y, \beta)}{\partial y} = 2\beta u(x, y, \beta)\left(\int_0^x u(t, y, \beta)\, dt - \frac{1}{2}\right), \tag{38}$$

and

$$(39) \qquad \int_0^x u(t, y, \beta)\, dt \le \int_0^1 u(t, y, \beta)\, dt = 1.$$

From (38) and (39), we get

$$(40) \qquad \left|\frac{\partial u}{\partial y}\right| \le |\beta| u(x, y, \beta) \le |\beta| e^{|\beta|}.$$

Since $u(x, y, \beta)$ is uniformly continuous on $[0, 1] \times [0, 1]$, $\int_0^x u(t, y, \beta)\, dt$ is also continuous on $[0, 1] \times [0, 1]$. Hence, by (38), $\frac{\partial u}{\partial y}$ is bounded and continuous on $[0, 1] \times [0, 1]$. A similar argument can be made for $\frac{\partial u}{\partial x}$. Thus we have shown that $u(x, y, \beta) \in C_b^1$ and

$$\max\left(\left|\frac{\partial u}{\partial x}\right|, \left|\frac{\partial u}{\partial y}\right|\right) \le |\beta| e^{|\beta|}.$$

Next, since $|\frac{\partial u(x,t,\beta)}{\partial x} \cdot u(t, y, \gamma)| \le |\beta| e^{|\beta|+|\gamma|}$ for any $0 \le x, y, t \le 1$, by the dominated convergence theorem, we have

$$(41) \quad \frac{\partial \rho(x, y)}{\partial x} = \frac{\partial}{\partial x}\left(\int_0^1 u(x, t, \beta) u(t, y, \gamma)\, dt\right) = \int_0^1 \frac{\partial u(x, t, \beta)}{\partial x} u(t, y, \gamma)\, dt.$$

Hence $|\frac{\partial \rho}{\partial x}| \le |\beta| e^{|\beta|+|\gamma|}$. Moreover, $\frac{\partial u(x,t,\beta)}{\partial x} \cdot u(t, y, \gamma)$ as a function of x, y, t is uniformly continuous on $[0, 1] \times [0, 1] \times [0, 1]$. Thus, by (41), $\frac{\partial \rho}{\partial x}$ is continuous on $[0, 1] \times [0, 1]$. By a similar argument, it can be shown that $\frac{\partial \rho}{\partial y}$ is continuous on $[0, 1] \times [0, 1]$, and $|\frac{\partial \rho}{\partial y}| \le |\gamma| e^{|\beta|+|\gamma|}$. Therefore, $\rho(x, y) \in C_b^1$ and

$$\max\left(\left|\frac{\partial \rho}{\partial x}\right|, \left|\frac{\partial \rho}{\partial y}\right|\right) \le (|\beta| + |\gamma|) e^{|\beta|+|\gamma|}.$$

□

The next lemma shows that for any nondecreasing curve in the unit square, in a strip of small width around it, with probability going to 1, there exists an increasing subsequence whose length can be bounded from below. The proof of Lemma 4.7 uses similar arguments as in the proof of Lemma 8 in Deuschel and Zeitouni (1995). Before stating the lemma, we need the following notation.

DEFINITION 4.6. Let $B_\nearrow$ be the set of nondecreasing, right continuous functions $\phi : [0, 1] \to [0, 1]$. For $\phi \in B_\nearrow$, we have $\phi(x) = \int_0^x \dot{\phi}(t)\, dt + \phi_s(x)$, where ϕ_s is singular and has a zero derivative almost everywhere. Let $\rho(x, y)$ be the density defined in (2). Define function $J : B_\nearrow \to \mathbb{R}$,

$$J(\phi) := \int_0^1 \sqrt{\dot{\phi}(x) \rho\big(x, \phi(x)\big)}\, dx \quad \text{and} \quad \bar{J} := \sup_{\phi \in B_\nearrow} J(\phi).$$

REMARK. By Theorems 3 and 4 in Deuschel and Zeitouni (1995), it follows from Lemma 4.5(a) and (b), that

$$\sup_{\phi \in B_{\nearrow}} J(\phi) = \sup_{\phi \in B^1_{\nearrow}} J(\phi),$$

where $B^1_{\nearrow}$ is defined in Theorem 1.4. Hence we use the same notation $\bar{J}$ to denote the supremum over $B_{\nearrow}$.

Given a function $\phi(x)$ and any $\delta > 0$, we say that a point (x, y) is in the δ neighborhood of ϕ if $\phi(x) - \delta < y < \phi(x) + \delta$.

LEMMA 4.7. *Under the same conditions as in Theorem 1.4, for any $\phi \in B^1_{\nearrow}$ and any $\delta, \varepsilon > 0$, define the event*

$$E_n := \Big\{(\pi, \tau) \in S_n \times S_n : \exists \textit{ an increasing subsequence of } \Big\{\Big(\frac{\pi(i)}{n}, \frac{\tau(i)}{n}\Big)\Big\}_{i\in[n]}$$
$$\textit{which is wholly contained in the } \delta \textit{ neighborhood of } \phi(\cdot)$$
$$\textit{and the length of which is greater than } 2J(\phi)(1-\varepsilon)\sqrt{n}\Big\}.$$

Then

$$\lim_{n\to\infty} \mathbb{P}_n(E_n) = 1.$$

PROOF. Given $\delta, \varepsilon > 0$, fix an integer K. Let $\Delta x := 1/K$. Let $x_i := i\Delta x$ and $y_i := \phi(x_i)$ for $i \in [K]$. Let $x_0 := 0$, $y_0 := 0$. Define the rectangles $R_i := [x_{i-1}, x_i] \times [y_{i-1}, y_i]$ for $i \in [K]$. Since ϕ is in C^1_b, for any $0 < \delta' < 1$, we can choose K large enough such that

$$\max_i (y_i - y_{i-1}) < \delta, \qquad e^{-\Delta x|\beta|/2} > 1 - \delta', \qquad \Delta x|\beta| < \ln 2 \tag{42}$$

$$\max_i \max_{x,y\in R_i} \max\Big(\frac{\rho(x,y)}{\rho(x_i,y_i)}, \frac{\rho(x_i,y_i)}{\rho(x,y)}\Big) < \frac{1}{1-\delta'} \tag{43}$$

and

$$\sum_{i=1}^K \sqrt{\rho(x_i,y_i)(y_i - y_{i-1})\Delta x} > (1-\delta')J(\phi). \tag{44}$$

(43) follows from the uniform continuity of $\rho(x, y)$ on $[0, 1] \times [0, 1]$ and the fact that $\rho(x, y)$ is bounded away from 0, which is proved in Lemma 4.5(a). (44) fol-

lows since

$$\lim_{K\to\infty}\sum_{i=1}^{K}\sqrt{\rho(x_i,y_i)(y_i-y_{i-1})\Delta x}$$
$$=\lim_{K\to\infty}\sum_{i=1}^{K}\sqrt{\rho(x_i,y_i)\frac{y_i-y_{i-1}}{x_i-x_{i-1}}}\Delta x$$
$$=J(\phi),$$

where the last equality follows from the definition of Riemann integral, the mean value theorem and the fact that $\phi\in C_b^1$. Next, for any $i\in[K]$, define $\rho(R_i):=\iint_{R_i}\rho(x,y)\,dx\,dy$. By (43), we have

$$\frac{\rho(R_i)}{1-\delta'}>\rho(x_i,y_i)(y_i-y_{i-1})\Delta x.$$

Hence, for any $i\in[K]$, we have

$$\frac{l_{R_i}(\pi,\tau)}{2\sqrt{n\rho(x_i,y_i)(y_i-y_{i-1})\Delta x}}\geq\frac{l_{R_i}(\pi,\tau)\sqrt{1-\delta'}}{2\sqrt{n\rho(R_i)}}. \tag{45}$$

By fixing the ε in Lemma 4.4 to be $2\delta'$, we have

$$\lim_{n\to\infty}\mathbb{P}_n\left(\frac{l_{R_i}(\pi,\tau)}{\sqrt{n\rho(R_i)}}>2e^{-\Delta x|\beta|/2}-2\delta'\right)=1. \tag{46}$$

Moreover,

$$\begin{aligned}\mathbb{P}_n&\left(\frac{l_{R_i}(\pi,\tau)}{2\sqrt{n\rho(x_i,y_i)(y_i-y_{i-1})\Delta x}}>(1-2\delta')\sqrt{1-\delta'}\right)\\&\geq\mathbb{P}_n\left(\frac{l_{R_i}(\pi,\tau)}{2\sqrt{n\rho(R_i)}}>1-2\delta'\right)\\&\geq\mathbb{P}_n\left(\frac{l_{R_i}(\pi,\tau)}{\sqrt{n\rho(R_i)}}>2e^{-\Delta x|\beta|/2}-2\delta'\right).\end{aligned} \tag{47}$$

The first inequality follows by (45), and the second inequality follows by (42), since

$$2e^{-\Delta x|\beta|/2}-2\delta'>2(1-\delta')-2\delta'=2(1-2\delta').$$

Hence, by (46) and (47), we get

$$\lim_{n\to\infty}\mathbb{P}_n\left(\frac{l_{R_i}(\pi,\tau)}{2\sqrt{n\rho(x_i,y_i)(y_i-y_{i-1})\Delta x}}>(1-2\delta')\sqrt{1-\delta'}\right)=1, \tag{48}$$

for any $i\in[K]$. Note that by concatenating the increasing subsequences of $\{(\frac{\pi(i)}{n},\frac{\tau(i)}{n})\}_{i\in[n]}$ in each R_i we get a increasing subsequence in $[0,1]\times[0,1]$

which is wholly contained in a δ neighborhood of ϕ. Combining (44) and (48), it follows that, with probability converging to 1 as $n \to \infty$, there exists an increasing subsequence of $\{(\frac{\pi(i)}{n}, \frac{\tau(i)}{n})\}_{i\in[n]}$ in a δ neighborhood of ϕ whose length is at least

$$\sum_{i=1}^{K} 2\sqrt{n}(1-2\delta')\sqrt{1-\delta'}\sqrt{\rho(x_i, y_i)(y_i - y_{i-1})\Delta x} > 2\sqrt{n}(1-2\delta')(1-\delta')^{\frac{3}{2}} J(\phi).$$

The lemma follows since we can choose δ' small enough in the first place such that $(1-2\delta')(1-\delta')^{\frac{3}{2}} > 1-\varepsilon$. □

DEFINITION 4.8. Given $K, L \in \mathbb{N}$, define

$$\boldsymbol{B}_{KL} := \{(b_0, b_1, \dots, b_K) \in \mathbb{Z}^{K+1} : 0 = b_0 \le b_1 \le \cdots \le b_K = KL - 1\}.$$

DEFINITION 4.9. Given $K, L \in \mathbb{N}$ and $\boldsymbol{b} = (b_0, b_1, \dots, b_K) \in \boldsymbol{B}_{KL}$, for any $i \in [K]$, define the rectangle $R_i := ((i-1)\Delta x, i\Delta x] \times (b_{i-1}\Delta y, (b_i + 1)\Delta y]$, where $\Delta x := \frac{1}{K}$ and $\Delta y := \frac{1}{KL}$. Let $M_i := \sup_{(x,y)\in R_i} \rho(x, y)$ and $m_i := \inf_{(x,y)\in R_i} \rho(x, y)$. Define

$$J_{\boldsymbol{b}}^{K,L} := \sum_{i=1}^{K} \sqrt{M_i(b_i - b_{i-1} + 1)\Delta x \Delta y}.$$

LEMMA 4.10.

$$\varlimsup_{\substack{K\to\infty\\ L\to\infty}} \max_{\boldsymbol{b}\in\boldsymbol{B}_{KL}} J_{\boldsymbol{b}}^{K,L} \le \bar{J},$$

where $\bar{J}$ is defined in Definition 4.6.

PROOF. Let M be an upper bound of $\rho(x, y)$. In the context of Definition 4.9, let $\phi_{\boldsymbol{b}}(x)$ be the piecewise linear function on $[0, 1]$ such that $\phi_{\boldsymbol{b}}(i\Delta x) = b_i\Delta y$, $i = 0, 1, \dots, K$. From the two definitions above, we have

$$\begin{aligned} J(\phi_{\boldsymbol{b}}) &= \int_0^1 \sqrt{\dot{\phi}_{\boldsymbol{b}}(x)\rho(x, \phi_{\boldsymbol{b}}(x))}\,dx \\ &= \sum_{i=1}^{K} \int_{(i-1)\Delta x}^{i\Delta x} \sqrt{\dot{\phi}_{\boldsymbol{b}}(x)\rho(x, \phi_{\boldsymbol{b}}(x))}\,dx \\ &= \sum_{i=1}^{K} \int_{(i-1)\Delta x}^{i\Delta x} \sqrt{\frac{(b_i - b_{i-1})\Delta y}{\Delta x} \cdot \rho(x, \phi_{\boldsymbol{b}}(x))}\,dx \\ &\ge \sum_{i=1}^{K} \int_{(i-1)\Delta x}^{i\Delta x} \sqrt{\frac{(b_i - b_{i-1})\Delta y}{\Delta x} \cdot m_i}\,dx \end{aligned} \tag{49}$$

$$
\begin{aligned}
&= \sum_{i=1}^{K} \sqrt{m_i (b_i - b_{i-1}) \Delta x \Delta y} \\
&\geq \sum_{i=1}^{K} \sqrt{M_i (b_i - b_{i-1}) \Delta x \Delta y} - \sum_{i=1}^{K} \sqrt{(M_i - m_i)(b_i - b_{i-1}) \Delta x \Delta y},
\end{aligned}
$$

where the last inequality follows since $\sqrt{a} + \sqrt{b} \geq \sqrt{a+b}$ for any $a, b \geq 0$. Moreover,

$$
\begin{aligned}
&\sum_{i=1}^{K} \sqrt{M_i (b_i - b_{i-1}) \Delta x \Delta y} \\
&\quad = J_{\boldsymbol{b}}^{K,L} - \sum_{i=1}^{K} \left(\sqrt{M_i (b_i - b_{i-1} + 1) \Delta x \Delta y} - \sqrt{M_i (b_i - b_{i-1}) \Delta x \Delta y}\right) \\
&\quad = J_{\boldsymbol{b}}^{K,L} - \sum_{i=1}^{K} \frac{M_i \Delta x \Delta y}{\sqrt{M_i (b_i - b_{i-1} + 1) \Delta x \Delta y} + \sqrt{M_i (b_i - b_{i-1}) \Delta x \Delta y}} \\
&\quad \geq J_{\boldsymbol{b}}^{K,L} - \sum_{i=1}^{K} \frac{M_i \Delta x \Delta y}{\sqrt{M_i (b_i - b_{i-1} + 1) \Delta x \Delta y}} \qquad (50) \\
&\quad \geq J_{\boldsymbol{b}}^{K,L} - \sum_{i=1}^{K} \frac{M_i \Delta x \Delta y}{\sqrt{M_i \Delta x \Delta y}} \\
&\quad \geq J_{\boldsymbol{b}}^{K,L} - \sqrt{M} \sum_{i=1}^{K} \sqrt{\Delta x \Delta y} \\
&\quad = J_{\boldsymbol{b}}^{K,L} - \sqrt{\frac{M}{L}}.
\end{aligned}
$$

Next, define

$$
\begin{aligned}
D_1(\boldsymbol{b}) &:= \{ i \in [K] : (b_i - b_{i-1} + 1) \Delta y \leq \sqrt[3]{\Delta x} \}, \\
D_2(\boldsymbol{b}) &:= \{ i \in [K] : (b_i - b_{i-1} + 1) \Delta y > \sqrt[3]{\Delta x} \}.
\end{aligned}
$$

For $i \in D_1(\boldsymbol{b})$, the height of R_i is no greater than $\sqrt[3]{\Delta x}$, and for $i \in D_2(\boldsymbol{b})$, the height of R_i is greater than $\sqrt[3]{\Delta x}$. To bound the cardinality of $D_2(\boldsymbol{b})$, we have

$$
\begin{aligned}
|D_2(\boldsymbol{b})| \sqrt[3]{\Delta x} &\leq \sum_{i \in D_2(\boldsymbol{b})} (b_i - b_{i-1} + 1) \Delta y \\
&\leq \sum_{i \in D_2(\boldsymbol{b})} (b_i - b_{i-1}) \Delta y + |D_2(\boldsymbol{b})| \Delta y
\end{aligned}
$$

$$\begin{aligned}
(51) \qquad & \leq \sum_{i=1}^{K}(b_i - b_{i-1})\Delta y + K\Delta y \\
& \leq 1 + \frac{1}{L} \\
& \leq 2.
\end{aligned}$$

Given $\varepsilon > 0$, by the uniform continuity of $\rho(x, y)$ on $[0, 1] \times [0, 1]$, there exists $K_0 > 0$ such that, for any $K > K_0$ and any $i \in D_1(\boldsymbol{b})$, we have $M_i - m_i < \varepsilon^2$. We can also choose K_0 sufficiently large such that, for any $K > K_0$,

$$2\sqrt{M}(\Delta x)^{\frac{1}{6}} < \varepsilon. \tag{52}$$

Thus, for any $K > K_0$, we have

$$\begin{aligned}
& \sum_{i=1}^{K}\sqrt{(M_i - m_i)(b_i - b_{i-1})\Delta x \Delta y} \\
& \quad \leq \sum_{i \in D_1(\boldsymbol{b})} \sqrt{\varepsilon^2(b_i - b_{i-1})\Delta x \Delta y} + \sum_{i \in D_2(\boldsymbol{b})} \sqrt{M(b_i - b_{i-1})\Delta x \Delta y} \\
& \quad \leq \varepsilon \sum_{i=1}^{K}\sqrt{(b_i - b_{i-1})\Delta x \Delta y} + \sum_{i \in D_2(\boldsymbol{b})} \sqrt{M \Delta x} \qquad (53) \\
& \quad \leq \varepsilon \sqrt{\sum_{i=1}^{K}\Delta x}\sqrt{\sum_{i=1}^{K}(b_i - b_{i-1})\Delta y} + 2\sqrt{M}(\Delta x)^{\frac{1}{6}} \\
& \quad < \varepsilon + \varepsilon,
\end{aligned}$$

where the second to last inequality follows by the Cauchy–Schwarz inequality and (51). Let $L_0 := \lceil \frac{M}{\varepsilon^2} \rceil$. By combining (49), (50) and (53), we get, for any $K > K_0$, $L > L_0$ and any $\boldsymbol{b}$,

$$J_{\boldsymbol{b}}^{K,L} \leq J(\phi_{\boldsymbol{b}}) + \sqrt{\frac{M}{L}} \leq J(\phi_{\boldsymbol{b}}) + 3\varepsilon \leq \bar{J} + 3\varepsilon,$$

where the last inequality follows from the fact that $\phi_{\boldsymbol{b}} \in B_{\nearrow}$ and Definition 4.6. □

DEFINITION 4.11. In the context of Definition 4.9, we call a sequence of points $(z_1, \dots, z_m)$ with $z_i = (x_i, y_i)$ a $\boldsymbol{b}$-*increasing sequence* if the following two conditions are satisfied:

(a) $(z_1, \dots, z_m)$ is an increasing sequence, that is, $x_i < x_{i+1}$ and $y_i < y_{i+1}$ for all $i \in [m-1]$.

(b) Every point in the sequence is contained in some rectangle R_j with $j \in [K]$. In other words, $(j-1)\Delta x < x_i \le j\Delta x$ implies $b_{j-1}\Delta y < y_i \le (b_j + 1)\Delta y$.

Given a collection of points $\boldsymbol{z} = \{z_i\}_{i\in[n]}$, let $\mathrm{LIS}_{\boldsymbol{b}}(\boldsymbol{z})$ denote the length of the longest $\boldsymbol{b}$-increasing subsequence of $\boldsymbol{z}$. That is,

$$\begin{aligned}\mathrm{LIS}_{\boldsymbol{b}}(z) := \max\bigl\{m : \exists (i_1, i_2, \dots, i_m) \\ \text{such that } (z_{i_1}, z_{i_2}, \dots, z_{i_m}) \text{ is a } \boldsymbol{b}\text{-increasing sequence}\bigr\}.\end{aligned}$$

Note that we do not require $i_j < i_{j+1}$ above.

LEMMA 4.12. *Under the same conditions as in Lemma* 1.3, *for any* $\delta > 0$, *there exist* K_0, L_0 *such that*, *for any* $K > K_0$, $L > L_0$ *and any* $\boldsymbol{b} = (b_0, b_1, \dots, b_K) \in \boldsymbol{B}_{KL}$:

$$\lim_{n\to\infty} \mathbb{P}_n\bigl(\mathrm{LIS}_{\boldsymbol{b}}\bigl(\boldsymbol{z}(\pi,\tau)\bigr) > 2\sqrt{n}(\bar{J}+\delta)\bigr) = 0, \tag{54}$$

where $\boldsymbol{z}(\pi,\tau) := \{(\frac{\pi(i)}{n}, \frac{\tau(i)}{n})\}_{i\in[n]}$.

PROOF. Given $\delta > 0$, by Lemma 4.10, there exist $K_1, L_1 > 0$ such that, for any $K > K_1$, $L > L_1$ and any $\boldsymbol{b} = (b_0, b_1, \dots, b_K) \in \boldsymbol{B}_{KL}$, we have

$$J_{\boldsymbol{b}}^{K,L} < \bar{J} + \frac{\delta}{2}.$$

Then we get

$$\mathbb{P}_n\bigl(\mathrm{LIS}_{\boldsymbol{b}}\bigl(\boldsymbol{z}(\pi,\tau)\bigr) > 2\sqrt{n}(\bar{J}+\delta)\bigr) \le \mathbb{P}_n\bigl(\mathrm{LIS}_{\boldsymbol{b}}\bigl(\boldsymbol{z}(\pi,\tau)\bigr) > 2\sqrt{n}\bigl(J_{\boldsymbol{b}}^{K,L}+\delta/2\bigr)\bigr).$$

Hence, to show (54), it suffices to show that there exists K_2, L_2 such that, for any $K > K_2$, $L > L_2$ and any $\boldsymbol{b}$,

$$\lim_{n\to\infty} \mathbb{P}_n\bigl(\mathrm{LIS}_{\boldsymbol{b}}\bigl(\boldsymbol{z}(\pi,\tau)\bigr) > 2\sqrt{n}\bigl(J_{\boldsymbol{b}}^{K,L}+\delta/2\bigr)\bigr) = 0. \tag{55}$$

Given $K, L > 0$, whose values are to be determined, and any $\boldsymbol{b} \in \boldsymbol{B}_{KL}$, we inherit all the notation introduced in Definition 4.9. Let $l_{R_i}(\pi,\tau)$ denote the length of the longest increasing subsequence of $\boldsymbol{z}(\pi,\tau)$ wholly contained in the rectangle R_i. For any $i \in [K]$, define

$$E_i(\boldsymbol{b}) := \bigl\{(\pi,\tau) : l_{R_i}(\pi,\tau) \ge 2\sqrt{n}\bigl(\sqrt{M_i(b_i - b_{i-1}+1)\Delta x\Delta y} + \delta\Delta x/2\bigr)\bigr\}.$$

Since $\mathrm{LIS}_{\boldsymbol{b}}(\boldsymbol{z}(\pi,\tau)) \le \sum_{i=1}^K l_{R_i}(\pi,\tau)$, we get

$$\bigl\{\mathrm{LIS}_{\boldsymbol{b}}\bigl(\boldsymbol{z}(\pi,\tau)\bigr) > 2\sqrt{n}\bigl(J_{\boldsymbol{b}}^{K,L}+\delta/2\bigr)\bigr\} \subset \bigcup_{i\in[K]} E_i(\boldsymbol{b}).$$

Hence, to show (55), it suffices to show

$$\lim_{n\to\infty} \mathbb{P}_n\bigl(E_i(\boldsymbol{b})\bigr) = 0 \qquad \forall i \in [K]. \tag{56}$$

Let $M := \sup_{0\le x,y\le 1}\rho(x,y)$. Since $e^{\Delta x|\beta|/2}-1=\Theta(\Delta x)$, there exists $K_2>0$ such that, for any $K>K_2$, we have

$$e^{\Delta x|\beta|/2} < 1+\frac{\delta\sqrt{\Delta x}}{2\sqrt{M}} \quad \text{and} \quad \Delta x|\beta|<\ln 2. \tag{57}$$

Moreover, for any $i\in[K]$,

$$\begin{aligned}
&\mathbb{P}_n(E_i(\boldsymbol{b}))\\
&\le \mathbb{P}_n\left(l_{R_i}(\pi,\tau)\ge 2\sqrt{n}\sqrt{M_i(b_i-b_{i-1}+1)\Delta x\Delta y}\left(1+\frac{\delta\Delta x}{2\sqrt{M\Delta x}}\right)\right)\\
&\le \mathbb{P}_n\left(l_{R_i}(\pi,\tau)\ge 2\sqrt{n\rho(R_i)}\left(1+\frac{\delta\sqrt{\Delta x}}{2\sqrt{M}}\right)\right).
\end{aligned} \tag{58}$$

The first inequality follows since $(b_i-b_{i-1}+1)\Delta y\le 1$ and $M_i\le M$. The second inequality follows since

$$M_i(b_i-b_{i-1}+1)\Delta x\Delta y \ge \int_{R_i}\rho(x,y)\,dx\,dy=\rho(R_i).$$

Hence, combining (57), (58) and Lemma 4.4, we get, for any $K>K_2$, $L>0$ and any $\boldsymbol{b}$,

$$\lim_{n\to\infty}\mathbb{P}_n(E_i(\boldsymbol{b}))=0 \qquad \forall i\in[K].$$

Thus, (56) as well as the lemma follow. □

PROOF OF THEOREM 1.4. By Proposition 3.5, if $\pi\sim\mu_{n,q}$, π^{-1} has the same distribution $\mu_{n,q}$. Hence, if $(\pi,\tau)\sim\mu_{n,q}\times\mu_{n,q'}$, (π^{-1},τ^{-1}) has the same distribution $\mu_{n,q}\times\mu_{n,q'}$. Note that $\mathrm{LIS}(\pi,\tau)=\mathrm{LIS}(z(\pi,\tau))$. Thus, by Corollary 2.4, to prove Theorem 1.4, it suffices to show

$$\lim_{n\to\infty}\mathbb{P}_n\left(\left|\frac{\mathrm{LIS}(z(\pi,\tau))}{\sqrt{n}}-2\bar{J}\right|<\varepsilon\right)=1, \tag{59}$$

for any $\varepsilon>0$. By Lemma 4.7 and the definition of $\bar{J}$, we have

$$\lim_{n\to\infty}\mathbb{P}_n\left(\frac{\mathrm{LIS}(z(\pi,\tau))}{\sqrt{n}}>2\bar{J}-\varepsilon\right)=1. \tag{60}$$

To show the upper bound in (59), note that, for any $K,L>0$ and any increasing sequence of points $\{(x_j,y_j)\}_{j\in[n]}$ with $0<x_j,y_j\le 1$, there exists a choice of $\boldsymbol{b}'=(b'_0,b'_1,\dots,b'_K)$ such that $\{(x_j,y_j)\}_{j\in[n]}$ is a $\boldsymbol{b}'$ - increasing sequence. Specifically, we can define $\boldsymbol{b}'$ as follows. Let $\Delta x:=\frac{1}{K}$, $\Delta y:=\frac{1}{KL}$:

- Define $b'_0:=0$, $b'_K:=KL-1$.
- For $i\in[K-1]$, define $b'_i:=\lfloor\max\{y_j:(i-1)\Delta x<x_j\le i\Delta x\}\cdot KL\rfloor$.

It can be easily verified that with $\boldsymbol{b}'$ thus defined, every point (x_j, y_j) is in some rectangle R_i, where R_i is defined in Definition 4.9. Hence we get

$$\begin{aligned}
&\mathbb{P}_n\biggl(\frac{\mathrm{LIS}(z(\pi,\tau))}{\sqrt{n}} > 2\bar{J}+\varepsilon\biggr) \\
&\qquad = \mathbb{P}_n\Bigl(\max_{\boldsymbol{b}\in\boldsymbol{B}_{KL}}\bigl(\mathrm{LIS}_{\boldsymbol{b}}\bigl(z(\pi,\tau)\bigr)\bigr) > \sqrt{n}(2\bar{J}+\varepsilon)\Bigr) \qquad (61)\\
&\qquad \le \sum_{\boldsymbol{b}\in\boldsymbol{B}_{KL}} \mathbb{P}_n\bigl(\mathrm{LIS}_{\boldsymbol{b}}\bigl(z(\pi,\tau)\bigr) > \sqrt{n}(2\bar{J}+\varepsilon)\bigr).
\end{aligned}$$

By Lemma 4.12, we can choose K, L sufficiently large such that, for any $\boldsymbol{b}\in\boldsymbol{B}_{KL}$,

$$\lim_{n\to\infty}\mathbb{P}_n\bigl(\mathrm{LIS}_{\boldsymbol{b}}\bigl(z(\pi,\tau)\bigr) > \sqrt{n}(2\bar{J}+\varepsilon)\bigr) = 0.$$

Hence, by (61) and the fact that the number of different choices of $\boldsymbol{b}$ is bounded above by $(KL)^K$, we have

$$\lim_{n\to\infty}\mathbb{P}_n\biggl(\frac{\mathrm{LIS}(z(\pi,\tau))}{\sqrt{n}} > 2\bar{J}+\varepsilon\biggr) = 0, \qquad (62)$$

and (59) follows from (60) and (62). □

4.3. *Solving $\bar{J}$ when $\beta=\gamma$.* The following lemma lets us solve for the supremum $\bar{J}$ when the underlying density $\rho(x,y)$ satisfies $\rho(\frac{x+y}{2},\frac{x+y}{2}) \ge \rho(x,y)$.

LEMMA 4.13. *Given a density $\rho(x,y)$ on $[0,1]\times[0,1]$ such that $\rho(x,y)$ is C_b^1 and $c<\rho(x,y)<C$ for some $C,c>0$, if $\rho(x,y)\le\rho(\frac{x+y}{2},\frac{x+y}{2})$ for any $0\le x,y\le 1$, then we have*

$$\bar{J} = \int_0^1 \sqrt{\rho(x,x)}\,dx,$$

that is, the supremum of $J(\phi)$ on $B_{\nearrow}$ is attained for $\phi(x)=x$.

PROOF. By the remark following Definition 4.6, it suffices to show that, for any $\phi\in B_{\nearrow}^1$, we have

$$J(\phi) \le \int_0^1 \sqrt{\rho(x,x)}\,dx. \qquad (63)$$

Define $g_\phi(x) := x+\phi(x)$. Since $\dot{\phi}(x)\ge 0$, we have $\dot{g_\phi}(x)\ge 1$. Next, we reparameterize $\phi(x)$ as follows:

$$t := \frac{g_\phi(x)}{2} = \frac{x+\phi(x)}{2}. \qquad (64)$$

Thus, we have $x = g_\phi^{-1}(2t)$ and $\phi(x) = 2t - x = 2t - g_\phi^{-1}(2t)$ where $t \in [0, 1]$. Moreover, since $g_\phi(x)$ is strictly increasing, x is strictly increasing as a function of t. Hence we have

$$\rho\big(x, \phi(x)\big) = \rho\big(g_\phi^{-1}(2t), 2t - g_\phi^{-1}(2t)\big) \le \rho(t, t), \tag{65}$$

where the last inequality follows since $\rho(x, y) \le \rho(\frac{x+y}{2}, \frac{x+y}{2})$. Next, by taking the derivative with respect to t on both sides of (64), we have

$$1 = \frac{1}{2}\left(\frac{dx}{dt} + \dot{\phi}(x)\frac{dx}{dt}\right). \tag{66}$$

By multiplying $2\frac{dx}{dt}$ on both sides of (66), we get

$$\dot{\phi}(x)\left(\frac{dx}{dt}\right)^2 = 2\frac{dx}{dt} - \left(\frac{dx}{dt}\right)^2 \le 1. \tag{67}$$

Hence, by (65) and (67), we have

$$\begin{aligned} J(\phi) &= \int_0^1 \sqrt{\dot{\phi}(x)\rho\big(x, \phi(x)\big)}\, dx \\ &\le \int_0^1 \sqrt{\dot{\phi}(x)\rho(t, t)} \cdot \frac{dx}{dt}\, dt \\ &= \int_0^1 \sqrt{\rho(t, t)\dot{\phi}(x)\left(\frac{dx}{dt}\right)^2}\, dt \\ &\le \int_0^1 \sqrt{\rho(t, t)}\, dt. \end{aligned}$$

Therefore, $\bar{J}$ is attained for $\phi(x) = x$. □

PROOF OF COROLLARY 1.5. Note that in the special case where $\beta = \gamma$, the density $\rho(x, y)$ in (2) is given by

$$\rho(x, y) := \int_0^1 u(x, t, \beta) \cdot u(t, y, \beta)\, dt. \tag{68}$$

In this case, we will show that $\rho(x, y) \le \rho(\frac{x+y}{2}, \frac{x+y}{2})$ for any $0 \le x, y \le 1$. Hence, by Lemma 4.5 and Lemma 4.13, $\bar{J}$ defined in Theorem 1.4 is attained when $\phi(x) = x$. In fact, by direct calculation, it can be shown that

$$u(x, t, \beta) \cdot u(t, y, \beta) \le u\left(\frac{x+y}{2}, t, \beta\right) \cdot u\left(t, \frac{x+y}{2}, \beta\right), \tag{69}$$

for any $0 \le x, y, t \le 1$. By the definition of $u(x, y, \beta)$, we have

$$
\begin{aligned}
&u(x,t,\beta)\cdot u(t,y,\beta)\\
&\quad= \frac{(\beta/2)\sinh(\beta/2)}{(e^{\beta/4}\cosh(\beta[x-t]/2)-e^{-\beta/4}\cosh(\beta[x+t-1]/2))^2}\\
&\qquad\times \frac{(\beta/2)\sinh(\beta/2)}{(e^{\beta/4}\cosh(\beta[t-y]/2)-e^{-\beta/4}\cosh(\beta[t+y-1]/2))^2}\\
&\quad= \frac{\beta(e^{\beta}-1)}{(2e^{\beta/2}\cosh(\beta[x-t]/2)-2\cosh(\beta[x+t-1]/2))^2}\\
&\qquad\times \frac{\beta(e^{\beta}-1)}{(2e^{\beta/2}\cosh(\beta[t-y]/2)-2\cosh(\beta[t+y-1]/2))^2}.
\end{aligned}
\tag{70}
$$

Considering the term inside the square of the denominator, by using the hyperbolic trigonometric identities,

$$
\begin{aligned}
\cosh(x)\cosh(y) &= \big(\cosh(x+y)+\cosh(x-y)\big)/2,\\
\cosh(x+y) &= \cosh(x)\cosh(y)+\sinh(x)\sinh(y),\\
\cosh(x-y) &= \cosh(x)\cosh(y)-\sinh(x)\sinh(y),
\end{aligned}
$$

we get

$$
\begin{aligned}
&\big(2e^{\beta/2}\cosh\big(\beta[x-t]/2\big)-2\cosh\big(\beta[x+t-1]/2\big)\big)\\
&\quad\times\big(2e^{\beta/2}\cosh\big(\beta[t-y]/2\big)-2\cosh\big(\beta[t+y-1]/2\big)\big)\\
&\quad= 2e^{\beta}\big(\cosh\big(\beta[x-y]/2\big)+\cosh\big(\beta[x+y-2t]/2\big)\big)\\
&\qquad-2e^{\beta/2}\big(\cosh\big(\beta[x+y-1]/2\big)+\cosh\big(\beta[x-y-2t+1]/2\big)\big)\\
&\qquad-2e^{\beta/2}\big(\cosh\big(\beta[x-y+2t-1]/2\big)+\cosh\big(\beta[x+y-1]/2\big)\big)\\
&\qquad+2\big(\cosh\big(\beta[x+y+2t-2]/2\big)+\cosh\big(\beta[x-y]/2\big)\big)\\
&\quad= S_t^- + S_t^+,
\end{aligned}
\tag{71}
$$

where S_t^- denotes the sum of those terms in the above equation containing the term $x-y$ and S_t^+ denotes the sum of those which contain the term $x+y$. After further simplification using the identities above, we have

$$
S_t^- = 2\cosh\big(\beta[x-y]/2\big)\big(e^{\beta}-2e^{\beta/2}\cosh\big(\beta[2t-1]/2\big)+1\big).
\tag{72}
$$

It is easily seen that the minimum of $e^{\beta}-2e^{\beta/2}\cosh(\beta[2t-1]/2)+1$ for $0 \le t \le 1$ is attained when $t = 0, 1$, and the minimum is 0. Hence, for any $t \in [0, 1]$, S_t^- is minimized when $x = y$. Thus to prove (69), it suffices to show that $S_t^+ \ge 0$, since $S_t^- + S_t^+$ is the term inside the square of the denominator of (70). After simplification, we have

$$\begin{aligned} S_t^+ = {} & 2e^{\beta}\big(\cosh(\beta[x+y-1]/2)\cosh(\beta[2t-1]/2) \\ & - \sinh(\beta[x+y-1]/2)\sinh(\beta[2t-1]/2)\big) \\ & - 4e^{\beta/2}\cosh(\beta[x+y-1]/2) \\ & + 2\big(\cosh(\beta[x+y-1]/2)\cosh(\beta[2t-1]/2) \\ & + \sinh(\beta[x+y-1]/2)\sinh(\beta[2t-1]/2)\big). \end{aligned} \tag{73}$$

Next, we make change of variables. Define $r := e^{\beta(x+y-1)/2}$, $s := e^{\beta(2t-1)/2}$. Then, from (73), we have

$$\begin{aligned} S_t^+ &= \frac{e^{\beta}}{2}\left(\left(r+\frac{1}{r}\right)\left(s+\frac{1}{s}\right) - \left(r-\frac{1}{r}\right)\left(s-\frac{1}{s}\right)\right) - 2e^{\beta/2}\left(r+\frac{1}{r}\right) \\ &\quad + \frac{1}{2}\left(\left(r+\frac{1}{r}\right)\left(s+\frac{1}{s}\right) + \left(r-\frac{1}{r}\right)\left(s-\frac{1}{s}\right)\right) \\ &= e^{\beta}\left(\frac{r}{s}+\frac{s}{r}\right) - 2e^{\beta/2}\left(r+\frac{1}{r}\right) + \left(rs+\frac{1}{rs}\right) \\ &= \left(\frac{e^{\beta}r}{s} + rs - 2e^{\beta/2}r\right) + \left(\frac{e^{\beta}s}{r} + \frac{1}{rs} - \frac{2e^{\beta/2}}{r}\right) \\ &\geq 0, \end{aligned} \tag{74}$$

where the last inequality follows since $x + y \geq 2\sqrt{xy}$ for any $x, y \geq 0$. We complete the proof of Corollary 1.5 by showing

$$\int_0^1 u(x,t,\beta)\cdot u(t,x,\beta)\,dt = \frac{\beta(\cosh(\beta/2) + 2\cosh(\beta[2x-1]/2))}{6\sinh(\beta/2)}, \tag{75}$$

for $0 \leq x \leq 1$.

By the same change of variables as above, since $y = x$, let $r := e^{\beta(2x-1)/2}$, $s := e^{\beta(2t-1)/2}$. Then we have

$$\frac{dt}{ds} = \frac{1}{\frac{ds}{dt}} = \frac{1}{s\beta}. \tag{76}$$

By (72), we have

$$S_t^- = 2\left(e^{\beta} - e^{\beta/2}\left(s+\frac{1}{s}\right) + 1\right). \tag{77}$$

Then, by (74) and (77), it can be easily verified that

$$rs(S_t^+ + S_t^-) = \left(e^{\beta/2}(r+s) - (rs+1)\right)^2. \tag{78}$$

Hence we have

$$\begin{aligned}
&\int_0^1 u(x,t,\beta)\cdot u(t,x,\beta)\,dt \\
&\quad = \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{\beta^2(e^\beta-1)^2}{(S_t^+ + S_t^-)^2}\frac{1}{s\beta}\,ds \\
&\quad = \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{\beta(e^\beta-1)^2 r^2 s}{(rs(S_t^+ + S_t^-))^2}\,ds \\
&\quad = \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{\beta(e^\beta-1)^2 r^2 s}{(e^{\beta/2}(r+s)-(rs+1))^4}\,ds \\
&\quad = \beta\big(e^\beta-1\big)^2 r^2 \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{s}{((e^{\beta/2}-r)s+e^{\beta/2}r-1)^4}\,ds \\
&\quad = \beta\big(e^\beta-1\big)^2 e^{\beta(2x-1)} \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{s}{(e^{\beta/2}(1-e^{\beta(x-1)})s+e^{\beta x}-1)^4}\,ds.
\end{aligned} \tag{79}$$

The first equality above follows from (70), (71), (76) and change of variables. The third equality follows from (78). Then we make another change of variable by defining

$$w := \frac{e^{\beta/2}(1-e^{\beta(x-1)})s+e^{\beta x}-1}{e^\beta-1},$$

from which we have

$$\frac{ds}{dw} = \frac{e^\beta-1}{e^{\beta/2}(1-e^{\beta(x-1)})} \quad \text{and} \quad w = \begin{cases} 1 & \text{when } s = e^{\beta/2}, \\ e^{\beta(x-1)} & \text{when } s = e^{-\beta/2}. \end{cases}$$

Hence, by (79), we have

$$\begin{aligned}
&\int_0^1 u(x,t,\beta)\cdot u(t,x,\beta)\,dt \\
&\quad = \frac{\beta e^{2\beta(x-1)}}{(e^\beta-1)(1-e^{\beta(x-1)})^2}\int_{e^{\beta(x-1)}}^1 \frac{(e^\beta-1)w-e^{\beta x}+1}{w^4}\,dw \\
&\quad = \frac{\beta e^{2\beta(x-1)}}{(e^\beta-1)(1-e^{\beta(x-1)})^2}\left(\frac{1-e^\beta}{2w^2}+\frac{e^{\beta x}-1}{3w^3}\right)\Bigg|_{e^{\beta(x-1)}}^1 \\
&\quad = \frac{\beta(1+e^\beta+2e^{\beta x}+2e^{-\beta(x-1)})}{6(e^\beta-1)} \\
&\quad = \frac{\beta(\cosh(\beta/2)+2\cosh(\beta[2x-1]/2))}{6\sinh(\beta/2)}.
\end{aligned}$$

□

5. Proof of Lemma 4.2. To prove Lemma 4.2, we use the same techniques developed in the proof of Corollary 4.3 in Mueller and Starr (2013), in which the authors constructed a coupling of two point processes. A point process is a random, locally finite, nonnegative integer valued measure. Let $\mathcal{X}_k$ denote the set of all Borel measures ξ on $\mathbb{R}^k$ such that $\xi(A) \in \{0, 1, 2, \dots\}$ for any bounded Borel set A in $\mathbb{R}^k$. Then a point process on $\mathbb{R}^k$ is a random variable which takes value in $\mathcal{X}_k$.

Suppose μ, ν are two measures on $\mathbb{R}^k$. We say $\mu \le \nu$ if $\mu(A) \le \nu(A)$ for any $A \in \mathcal{B}(\mathbb{R}^k)$.

LEMMA 5.1. *Suppose $\hat{\alpha}$ and α are two probability measures on $[0, 1]$ with density $f(x)$, $g(x)$, respectively. If, for any $x \in [0, 1]$, $f(x) \ge p \cdot g(x)$ for some $0 < p < 1$, then there exist random variables X, Y and B_p such that the following hold*:

- *X is $\hat{\alpha}$-distributed, Y is α-distributed and B_p is Bernoulli distributed with $\mathbb{P}(B_p = 1) = p$.*
- *B_p and Y are independent.*
- *Define two point processes η, ξ on $[0, 1]$ as follows*:

$$\xi(A) := \mathbb{1}_A(X) \quad \text{and} \quad \eta(A) := B_p \cdot \mathbb{1}_A(Y) \qquad \forall A \in \mathcal{B}([0, 1]).$$

Then we have $\eta \le \xi$.

PROOF. Let Y, Y' and B_p be independent random variables defined on the same probability space such that Y is α-distributed, B_p is Bernoulli distributed with $\mathbb{P}(B_p = 1) = p$ and the density of the distribution of Y' is $\frac{f(x)-p\cdot g(x)}{1-p}$. Define $X := B_p Y + (1 - B_p)Y'$. To see that X thus defined is $\hat{\alpha}$-distributed, we have

$$\mathbb{P}(X \in A) = p \int_A g(x)\, dx + (1 - p) \int_A \frac{f(x) - p \cdot g(x)}{1 - p}\, dx = \int_A f(x)\, dx,$$

for any $A \in \mathcal{B}([0, 1])$. Finally, the two point processes ξ and η thus defined satisfy $\eta \le \xi$, since for any $A \in \mathcal{B}([0, 1])$, when $B_p = 1$, we have $\xi(A) = \eta(A)$, and, when $B_p = 0$, we have $\eta(A) = 0$. □

LEMMA 5.2. *Suppose $\hat{\alpha}$ and α are two probability measures on $[0, 1]$ with density $f(x)$, $g(x)$, respectively. If $(1 - \theta_1)g(x) \le f(x) \le (1 + \theta_2)g(x)$ for some $\theta_1, \theta_2 \ge 0$ with $\theta_1 + \theta_2 < 1$, then there exist random variables X, Y, Z and B_θ such that the following hold*:

- *X is $\hat{\alpha}$-distributed, Y and Z are α-distributed and B_θ is Bernoulli distributed with $\mathbb{P}(B_\theta = 1) = \theta$, where $\theta = \theta_1 + \theta_2$.*
- *B_θ, Y and Z are independent.*

- *Define two point processes* ξ, ζ *on* $[0, 1]$ *as follows*:

$$\xi(A) := \mathbb{1}_A(X) \quad \textit{and} \quad \zeta(A) := \mathbb{1}_A(Y) + B_\theta \cdot \mathbb{1}_A(Z) \qquad \forall A \in \mathcal{B}([0,1]).$$

Then we have $\xi \le \zeta$.

PROOF. Let Y, Z and B_θ be independent random variables defined on the same probability space such that Y, Z is α-distributed, B_θ is Bernoulli distributed with $\mathbb{P}(B_\theta = 1) = \theta$. We define a new random variable X as follows. Conditioned on $Y = y$ and $Z = z$:

- If $B_\theta = 0$, define $X = y$.
- If $B_\theta = 1$, flip a coin W with probability of heads being $\frac{f(z)-(1-\theta_1)g(z)}{\theta \cdot g(z)}$. If the result is heads, define $X = z$, else define $X = y$.

Note that, without loss of generality, here we may assume $g(z) > 0$, since $\mathbb{P}(g(Z) = 0) = 0$. It is straightforward that the two point processes ξ and ζ thus defined satisfy $\xi \le \zeta$. We complete the proof by verifying that X thus defined has distribution $f(x)$.

For any $A \in \mathcal{B}([0, 1])$, the event $\{X \in A\}$ can be partitioned into three parts: $\{B_\theta = 0, Y \in A\}$, $\{B_\theta = 1, W \text{ is heads}, Z \in A\}$ and $\{B_\theta = 1, W \text{ is tails}, Y \in A\}$. We have

$$\mathbb{P}(\{B_\theta = 0, Y \in A\}) = (1-\theta)\int_A g(x)\,dx = (1-\theta)\alpha(A),$$

$$\begin{aligned}\mathbb{P}(\{B_\theta = 1, W \text{ is heads}, Z \in A\}) &= \theta \int_A \frac{f(z) - (1-\theta_1)g(z)}{\theta \cdot g(z)} g(z)\,dz \\ &= \int_A f(z)\,dz - (1-\theta_1)\alpha(A),\end{aligned}$$

$$\begin{aligned}&\mathbb{P}(\{B_\theta = 1, W \text{ is tails}, Y \in A\}) \\ &\quad = \theta \int_A g(y)\,dy \int_0^1 \left(1 - \frac{f(z) - (1-\theta_1)g(z)}{\theta \cdot g(z)}\right) g(z)\,dz \\ &\quad = \alpha(A) \int_0^1 (1+\theta_2)g(z) - f(z)\,dz \\ &\quad = \alpha(A)\theta_2.\end{aligned}$$

Here, we evaluate the last two probabilities by conditioning on the value of Z. Summing up the three probabilities, we get

$$\mathbb{P}(\{X \in A\}) = \int_A f(z)\,dz. \qquad \square$$

Next, we define a triangular array of random variables in $[0, 1]$.

DEFINITION 5.3. Suppose that $\{q_n\}_{n=1}^{\infty}$ is a sequence such that $q_n > 0$. For any $n \in \mathbb{N}$, we define the random vector $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$ as follows. Let $\{Y_i\}_{i=1}^n$ be i.i.d. uniform random variables on $[0, 1]$. Let $\{Y_{(i)}\}_{i=1}^n$ be the order statistics of $\{Y_i\}_{i=1}^n$. Independently, let π be a μ_{n,q_n}-distributed random variable on S_n. We define $Y_i^{(n)} := Y_{(\pi(i))}$ for all $i \in [n]$.

In the remainder of this paper, we use $(Y_1^{(n)}, \dots, Y_n^{(n)})$ specifically to denote the random vector defined as above. Next, we define the function Φ which maps vectors in $\mathbb{R}^n$ or n points in $\mathbb{R}^2$ to the induced permutation in S_n.

DEFINITION 5.4. Suppose $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ is a vector in $\mathbb{R}^n$ such that all its entries are distinct. Let $\Phi(\boldsymbol{x})$ denote the permutation in S_n such that, for any $i \in [n]$, $\Phi(\boldsymbol{x})(i) = j$ if x_i is the jth smallest entry in $\boldsymbol{x}$. Similarly, suppose $\boldsymbol{z} = \{(x_i, y_i)\}_{i=1}^n$ are n points in $\mathbb{R}^2$ such that they share no x coordinate nor any y coordinate. Let $\Phi(\boldsymbol{z})$ denote the permutation in S_n such that, for any $i \in [n]$, $\Phi(\boldsymbol{z})(i) = j$ if there exists $k \in [n]$, such that x_k is the ith smallest term in $\{x_i\}_{i=1}^n$ and y_k is the jth smallest term in $\{y_i\}_{i=1}^n$.

REMARK. From the above definitions, it can be easily seen that:

(a) For any $x_1 < \dots < x_n$ and $\boldsymbol{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$, we have $\Phi(\boldsymbol{y}) = \Phi(\{(x_i, y_i)\}_{i=1}^n)$.

(b) For any $\boldsymbol{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ and increasing indices $\boldsymbol{b} = (b_1, \dots, b_m)$, we have $\Phi(\boldsymbol{y})_{\boldsymbol{b}} = \Phi((y_{b_1}, \dots, y_{b_m}))$.

(c) $\Phi((Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)}))$ is μ_{n,q_n}-distributed.

Let D_n be the set of vectors in $[0, 1]^n$ which contain (at least two) identical entries. It is not hard to show that the density function of $(Y_1^{(n)}, \dots, Y_n^{(n)})$ is the following:

$$f_n(\boldsymbol{y}) = \mu_{n,q_n}\big(\Phi(\boldsymbol{y})\big) \cdot n! \qquad \text{for all } \boldsymbol{y} \in [0, 1]^n \setminus D_n.$$

Since $\{Y_i^{(n)}\}_{i=1}^n = \{Y_i\}_{i=1}^n$ are n i.i.d. uniform samples from $[0, 1]$, we have $\mathbb{P}((Y_1^{(n)}, \dots, Y_n^{(n)}) \in D_n) = 0$. Intuitively, for any $0 \le y_1 < \dots < y_n \le 1$, there are $n!$ ways to choose the vector $(Y_1, \dots, Y_n)$ such that $\{Y_i\}_{i=1}^n = \{y_i\}_{i=1}^n$. Moreover, conditioned on $\{Y_i\}_{i=1}^n = \{y_i\}_{i=1}^n$, the probability of $(Y_1^{(n)}, \dots, Y_n^{(n)}) = (y_{\pi(1)}, \dots, y_{\pi(n)})$ is $\mu_{n,q_n}(\pi)$. Since the measure of D_n is zero, when $\boldsymbol{y} \in D_n$, we can define $f_n(\boldsymbol{y})$ to be an arbitrary value.

LEMMA 5.5. *Given $i \in [n]$ and a vector $(y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n) \in [0, 1]^{n-1} \setminus D_{n-1}$, let $\hat{\alpha}$ denote the distribution of $Y_i^{(n)}$ conditioned on the event*

$\{Y_j^{(n)} = y_j$ for all $j \in [n] \setminus \{i\}\}$. Then $\hat{\alpha}$ has density $f(y)$ on $[0, 1]$ such that, for any $y, y' \in [0, 1] \setminus \{y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n\}$, we have

$$f(y) \geq \min\left(q_n^n, \frac{1}{q_n^n}\right), \qquad f(y) - f(y') \leq \max\left(q_n^n, \frac{1}{q_n^n}\right) - 1.$$

PROOF. Since $(Y_1^{(n)}, \dots, Y_n^{(n)})$ has density $f_n(\boldsymbol{y}) = \mu_{n,q_n}(\Phi(\boldsymbol{y})) \cdot n!$ on $[0, 1]^n \setminus D_n$, the density $f(y)$ of $\hat{\alpha}$ is given by

$$f(y) = \frac{\mu_{n,q_n}(\Phi((y_1, \dots, y_{i-1}, y, y_{i+1}, \dots, y_n)))}{\int_0^1 \mu_{n,q_n}(\Phi((y_1, \dots, y_{i-1}, t, y_{i+1}, \dots, y_n)))\,dt},$$

for any $y \in [0, 1]/\{y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n\}$. It can be seen from the definition that $f(y)$ is a simple function which takes at most n different values. Let M and m denote the maximum and minimum of $f(y)$, respectively. Then we have $M \geq 1$ and $0 < m \leq 1$. Moreover, for any $y, y' \in [0, 1]$, let $\boldsymbol{y} := (y_1, \dots, y_{i-1}, y, y_{i+1}, \dots, y_n)$ and $\boldsymbol{y}' := (y_1, \dots, y_{i-1}, y', y_{i+1}, \dots, y_n)$. We have

$$\left|l\big(\Phi(\boldsymbol{y})\big) - l\big(\Phi(\boldsymbol{y}')\big)\right| \leq n - 1.$$

That is, if $\boldsymbol{y}$ and $\boldsymbol{y}'$ differ at one entry, the number of inversions of the induced permutations differ at most by $n - 1$. Hence, assuming $q_n \geq 1$, for any $y, y' \in [0, 1]$, we have

$$\frac{1}{q_n^{n-1}} \leq \frac{f(y)}{f(y')} \leq q_n^{n-1}.$$

Choose y' such that $f(y') = M$, we have $f(y) \geq M/q_n^{n-1} \geq 1/q_n^n$. For the second part, we choose y, y' such that $f(y) = M$ and $f(y') = m$. Then we have $M/m - 1 \leq q_n^{n-1} - 1 \leq q_n^n - 1$. Thus, $M - m \leq q_n^n - 1$, since $0 < m \leq 1$. The argument for the case when $0 < q_n < 1$ is similar. □

LEMMA 5.6. *Given $n \in \mathbb{N}$ and $q_n > 0$, for any $m \leq n$ and any increasing indices $\boldsymbol{b} = (b_1, \dots, b_m)$, there exists a random vector $(V_1, \dots, V_n) \in [0, 1]^n$ and $2m$ independent random variables $\{U_i\}_{i=1}^m \cup \{B_i\}_{i=1}^m$ such that $(V_1, \dots, V_n)$ has the same distribution as $(Y_1^{(n)}, \dots, Y_n^{(n)})$, each U_i is uniformly distributed on $[0, 1]$ and each B_i is a Bernoulli random variable with $\mathbb{P}(B_i = 1) = \min(q_n^n, 1/q_n^n)$. Moreover, if we define two point processes as follows:*

$$\xi_{\boldsymbol{b}}^{(n)}(A) := \sum_{i=1}^m \mathbb{1}_A\big((i, V_{b_i})\big),$$

$$\eta_m(A) := \sum_{i=1}^m B_i \cdot \mathbb{1}_A\big((i, U_i)\big) \qquad \forall A \in \mathcal{B}\big(\mathbb{N} \times [0, 1]\big),$$

we have $\eta_m \leq \xi_{\boldsymbol{b}}^{(n)}$ almost surely.

PROOF. Given n, m and $\boldsymbol{b}$, let $(Y_1^{(n)}, \dots, Y_n^{(n)})$ be as defined in Definition 5.3 and, independently, define $2m$ independent random variables $\{U_i\}_{i=1}^m \cup \{B_i\}_{i=1}^m$ such that each U_i is uniformly distributed on $[0, 1]$ and each B_i is a Bernoulli random variable with $\mathbb{P}(B_i = 1) = \min(q_n^n, 1/q_n^n)$. We define the random vector $(V_1, \dots, V_n)$ as follows:

- Sample the random vector $(Y_1^{(n)}, \dots, Y_n^{(n)})$, say, we get $(Y_1^{(n)}, \dots, Y_n^{(n)}) = (y_1, \dots, y_n)$.
- For $j \in [n] \setminus \{b_i\}_{i=1}^m$, let $V_j := y_j$.
- For each $i \in [m]$, we resample $Y_{b_i}^{(n)}$ one by one, conditioned on the current value of other $Y_j^{(n)}$. Let y'_{b_i} denote the new value of $Y_{b_i}^{(n)}$ after the resampling and define $V_{b_i} := y'_{b_i}$. Specifically, for each $i \in [m]$, we sample a value y'_{b_i} according to the distribution of $Y_{b_i}^{(n)}$, conditioned on the event

$$\{Y_{b_j}^{(n)} = y'_{b_j} \text{ for } \forall j < i \text{ and } Y_k^{(n)} = y_k \text{ for } \forall k \in [n] \setminus \{b_j\}_{j \in [i]}\}.$$

- In each resampling step, say, resampling $Y_{b_i}^{(n)}$, let $\hat{\alpha}$ denote the above conditional distribution of $Y_{b_i}^{(n)}$. By Lemma 5.5, we know that that $\hat{\alpha}$ has density $f(y)$ with $f(y) \ge \min(q_n^n, 1/q_n^n)$. Hence we can couple this resampling procedure with variables U_i and B_i in the same fashion as in the proof of Lemma 5.1, with α in that lemma being the uniform measure on $[0, 1]$. Thus, we have $\mathbb{1}_A((i, V_{b_i})) \ge B_i \cdot \mathbb{1}_A((i, U_i))$ a. s. for any $A \in \mathcal{B}(\mathbb{N} \times [0, 1])$.

It can be easily seen from the above procedure that $(V_1, \dots, V_n)$ thus defined has the same distribution as $(Y_1^{(n)}, \dots, Y_n^{(n)})$, and

$$\eta_m(A) = \sum_{i=1}^m B_i \cdot \mathbb{1}_A\big((i, U_i)\big) \le \sum_{i=1}^m \mathbb{1}_A\big((i, V_{b_i})\big) = \xi_{\boldsymbol{b}}^{(n)}(A) \qquad \text{a.s.}$$

for any $A \in \mathcal{B}(\mathbb{N} \times [0, 1])$. □

LEMMA 5.7. *Given $n \in \mathbb{N}$ and $q_n > 0$ such that* $\max(q_n^n, 1/q_n^n) < 2$, *for any $m \le n$ and any increasing indices $\boldsymbol{b} = (b_1, \dots, b_m)$, there exists a random vector $(V_1, \dots, V_n) \in [0, 1]^n$ and $3m$ independent random variables $\{U_i, U'_i, B_i\}_{i=1}^m$ such that $(V_1, \dots, V_n)$ has the same distribution as the vector $(Y_1^{(n)}, \dots, Y_n^{(n)})$, each U_i, U'_i are uniformly distributed on* $[0, 1]$ *and each B_i is a Bernoulli random variable with* $\mathbb{P}(B_i = 1) = \max(q_n^n, 1/q_n^n) - 1$. *Moreover, if we define two point* *processes as follows*:

$$\xi_{\boldsymbol{b}}^{(n)}(A) := \sum_{i=1}^{m} \mathbb{1}_A\big((i, V_{b_i})\big) \qquad \forall A \in \mathcal{B}\big(\mathbb{N} \times [0,1]\big),$$

$$\zeta_m(A) := \sum_{i=1}^{m} \mathbb{1}_A\big((i, U_i')\big) + B_i \cdot \mathbb{1}_A\big((i, U_i)\big) \qquad \forall A \in \mathcal{B}\big(\mathbb{N} \times [0,1]\big)$$

we have $\xi_{\boldsymbol{b}}^{(n)} \le \zeta_m$ *almost surely.*

PROOF. The proof of this lemma is similar to the proof of Lemma 5.6. Given n, m and $\boldsymbol{b}$, define $(Y_1^{(n)}, \dots, Y_n^{(n)})$ as in Definition 5.3 and, independently, define $3m$ independent random variables $\{U_i, U_i', B_i\}_{i=1}^m$ such that each U_i, U_i' are uniformly distributed on $[0, 1]$ and each B_i is a Bernoulli random variable with $\mathbb{P}(B_i = 1) = \max(q_n^n, 1/q_n^n) - 1$. Then we define the random vector $(V_1, \dots, V_n)$ by the same steps as in the proof of Lemma 5.6, except that, in each resampling step, we couple the resampling of $Y_{b_i}^{(n)}$ with the variables U_i, U_i' and B_i in the same way as in the proof of Lemma 5.2, with α in that lemma being the uniform measure on $[0, 1]$. Note that the second inequality in Lemma 5.5 ensures that the conditions in Lemma 5.2 are met. Specifically, in each resampling step, let $f(y)$ denote the density of the conditional distribution of $Y_{b_i}^{(n)}$. Let M, m be the maximum and minimum of $f(y)$, respectively. Define $\theta_1 := 1 - m$ and $\theta_2 := M - 1$. Hence $1 - \theta_1 \le f(y) \le 1 + \theta_2$ almost surely and $\theta_1 + \theta_2 = M - m \le \max(q_n^n, 1/q_n^n) - 1 < 1$. □

Recall that $\mathcal{X}_2$ denotes the set of all Borel measures ξ on $\mathbb{R}^2$ such that $\xi(A) \in \{0, 1, 2, \dots\}$ for any bounded Borel set A in $\mathbb{R}^2$.

DEFINITION 5.8. For any $\xi \in \mathcal{X}_2$, we define the LIS of ξ as follows:

$$\begin{aligned} \mathrm{LIS}(\xi) := \max\big\{k : \exists (x_1, y_1), (x_2, y_2), \dots, (x_k, y_k) \in \mathbb{R}^2 \text{ such that} \\ \xi\big(\{(x_i, y_i)\}\big) \ge 1, \forall i \in [k] \text{ and } (x_i - x_j)(y_i - y_j) > 0, \forall i \ne j\big\}. \end{aligned}$$

It is easily seen that the function $\mathrm{LIS}(\cdot)$ is nondecreasing on $\mathcal{X}_2$ in the sense that, if $\xi, \zeta \in \mathcal{X}_2$ with $\xi \le \zeta$, we have $\mathrm{LIS}(\xi) \le \mathrm{LIS}(\zeta)$. Moreover, for any n points $\{(x_i, y_i)\}_{i=1}^n$ in $\mathbb{R}^2$ such that $x_i \ne x_j$ and $y_i \ne y_j$ for all $i \ne j$, define the integer-valued measure ξ as follows:

$$\xi(A) := \sum_{i=1}^{n} \mathbb{1}_A\big((x_i, y_i)\big) \qquad \forall A \in \mathcal{B}\big(\mathbb{R}^2\big).$$

Then we have $\mathrm{LIS}(\xi) = \mathrm{LIS}(\{(x_i, y_i)\}_{i=1}^n)$, where the latter one is defined in Definition 2.1.

LEMMA 5.9. *Let* $(V_1, \dots, V_n)$ *be a random vector which has the same distribution as* $(Y_1^{(n)}, \dots, Y_n^{(n)})$. *For any* $m \le n$ *and any increasing indices* $\boldsymbol{b} = (b_1, \dots, b_m)$, *define the point process* $\xi_{\boldsymbol{b}}^{(n)}$ *as in the previous two lemmas, that is,*

$$\xi_{\boldsymbol{b}}^{(n)}(A) := \sum_{i=1}^{m} \mathbb{1}_A\big((i, V_{b_i})\big) \qquad \forall A \in \mathcal{B}\big(\mathbb{N} \times [0, 1]\big).$$

Then $\mathrm{LIS}(\xi_{\boldsymbol{b}}^{(n)})$ *and* $\mathrm{LIS}(\pi_{\boldsymbol{b}})$ *have the same distribution, where* $\pi \sim \mu_{n,q_n}$.

PROOF. By the remarks after Definition 5.4, we have

$$\Phi\big(\{(i, V_{b_i})\}_{i=1}^m\big) = \Phi\big((V_{b_1}, V_{b_2}, \dots, V_{b_m})\big) = \Phi\big((V_1, V_2, \dots, V_n)\big)_{\boldsymbol{b}},$$

where $\Phi((V_1, V_2, \dots, V_n))$ in the last term has the distribution μ_{n,q_n}. The lemma follows by the fact that

$$\mathrm{LIS}\big(\xi_{\boldsymbol{b}}^{(n)}\big) = \mathrm{LIS}\big(\{(i, V_{b_i})\}_{i=1}^m\big) = \mathrm{LIS}\big(\Phi\big(\{(i, V_{b_i})\}_{i=1}^m\big)\big). \qquad \square$$

Now we are in the position to prove Lemma 4.2. In the following, we use λ_n to denote the uniform measure on S_n.

PROOF OF LEMMA 4.2. The lemma can be divided into two parts. For the first part, we show that, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \max_{\boldsymbol{b}\in Q(n,k_n)} \mu_{n,q_n}\left(\pi \in S_n : \frac{\mathrm{LIS}(\pi_{\boldsymbol{b}})}{\sqrt{k_n}} \le 2e^{\frac{-|\beta|}{2}} - \varepsilon\right) = 0. \tag{80}$$

Given $n > 0$, for any $\boldsymbol{b} \in Q(n, k_n)$, by Lemma 5.9, $\mathrm{LIS}(\xi_{\boldsymbol{b}}^{(n)})$ and $\mathrm{LIS}(\pi_{\boldsymbol{b}})$ have the same distribution, where $\xi_{\boldsymbol{b}}^{(n)}$ is the point process defined in that lemma. Moreover, by Lemma 5.6, there exists a point process η_{k_n} such that $\eta_{k_n} \le \xi_{\boldsymbol{b}}^{(n)}$ almost surely and η_{k_n} is defined by

$$\eta_{k_n}(A) := \sum_{i=1}^{k_n} B_{n,i} \cdot \mathbb{1}_A\big((i, U_i)\big) \qquad \forall A \in \mathcal{B}\big(\mathbb{N} \times [0, 1]\big), \tag{81}$$

where $\{U_i\}_{i=1}^{k_n} \cup \{B_{n,i}\}_{i=1}^{k_n}$ are $2k_n$ independent random variables with each U_i being uniformly distributed on $[0, 1]$ and each $B_{n,i}$ being a Bernoulli random variable with $\mathbb{P}(B_{n,i} = 1) = \min(q_n^n, 1/q_n^n)$. Hence, by the monotonicity of $\mathrm{LIS}(\cdot)$ on $\mathcal{X}_2$, we have

$$\begin{aligned}\mu_{n,q_n}\left(\pi \in S_n : \frac{\mathrm{LIS}(\pi_{\boldsymbol{b}})}{\sqrt{k_n}} \le 2e^{\frac{-|\beta|}{2}} - \varepsilon\right) &= \mathbb{P}\left(\frac{\mathrm{LIS}(\xi_{\boldsymbol{b}}^{(n)})}{\sqrt{k_n}} \le 2e^{\frac{-|\beta|}{2}} - \varepsilon\right) \\ &\le \mathbb{P}\left(\frac{\mathrm{LIS}(\eta_{k_n})}{\sqrt{k_n}} \le 2e^{\frac{-|\beta|}{2}} - \varepsilon\right).\end{aligned}$$

We complete the proof of (80) by showing that

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\mathrm{LIS}(\eta_{k_n})}{\sqrt{k_n}} > 2e^{\frac{-|\beta|}{2}} - \varepsilon\right) = 1, \tag{82}$$

for any $\varepsilon > 0$. First, we show that

$$\lim_{n\to\infty} \min\left(q_n^n, 1/q_n^n\right) = e^{-|\beta|}. \tag{83}$$

Assuming $0 < q_n \le 1$, since $\lim_{n\to\infty} n(1-q_n) = \beta$ and $\lim_{n\to\infty} \frac{\ln q_n}{q_n - 1} = 1$, we have

$$\lim_{n\to\infty} q_n^n = \lim_{n\to\infty} e^{n \ln q_n} = \lim_{n\to\infty} e^{n(q_n-1)} = e^{-|\beta|}.$$

The case $q_n > 1$ can be shown similarly. Hence, by (83), for any $\varepsilon_1 > 0$, there exists $N_1 > 0$ such that, for any $n > N_1$, we have $\min(q_n^n, 1/q_n^n) > e^{-|\beta|-\varepsilon_1}$. Thus, by the law of large numbers and the fact that $\lim_{n\to\infty} k_n = \infty$, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\sum_{i=1}^{k_n} B_{n,i} > k_n e^{-|\beta|-\varepsilon_1}\right) = 1. \tag{84}$$

Given $\boldsymbol{U} = (U_1, \dots, U_{k_n})$ and $\boldsymbol{B} = (B_{n,1}, \dots, B_{n,k_n})$, let $\Lambda(\boldsymbol{U}, \boldsymbol{B})$ denote the set of points in $\mathbb{R}^2$ defined by

$$\Lambda(\boldsymbol{U}, \boldsymbol{B}) := \left\{(i, U_i) : i \in [k_n] \text{ and } B_{n,i} = 1\right\}.$$

By the definition of η_{k_n} and Definition 5.8, we have

$$\mathrm{LIS}(\eta_{k_n}) = \mathrm{LIS}\left(\Lambda(\boldsymbol{U}, \boldsymbol{B})\right).$$

Moreover, conditioned on $\sum_{i=1}^{k_n} B_{n,i} = m$, by the independence of $\boldsymbol{U}$ and $\boldsymbol{B}$, it is easily seen that $\mathrm{LIS}(\Lambda(\boldsymbol{U}, \boldsymbol{B}))$ has the same distribution as $\mathrm{LIS}(\pi)$ with $\pi \sim \lambda_m$. For any $0 < \varepsilon_2, \varepsilon_3 < 1$, by the result of Kerov and Vershik (1977), there exists $M > 0$ such that, for any $m > M$,

$$\lambda_m\left(\frac{\mathrm{LIS}(\pi)}{\sqrt{m}} > 2 - \varepsilon_2\right) > 1 - \varepsilon_3. \tag{85}$$

Since $\lim_{n\to\infty} k_n = \infty$ and (84), there exists $N > N_1$ such that, for any $n > N$, we have

$$k_n e^{-|\beta|-\varepsilon_1} > M \quad \text{and} \quad \mathbb{P}\left(\sum_{i=1}^{k_n} B_{n,i} > k_n e^{-|\beta|-\varepsilon_1}\right) > 1 - \varepsilon_3.$$

Let E_m denote the event $\{\sum_{i=1}^{k_n} B_{n,i} = m\}$ and $s := \lfloor k_n e^{-|\beta|-\varepsilon_1} \rfloor + 1$. For any $n > N$, we have

$$\begin{aligned}
&\mathbb{P}\big(\mathrm{LIS}(\eta_{k_n}) > (2-\varepsilon_2)\sqrt{k_n e^{-|\beta|-\varepsilon_1}}\big) \\
&\quad\geq \sum_{m=s}^{k_n} \mathbb{P}\big(\mathrm{LIS}(\eta_{k_n}) > (2-\varepsilon_2)\sqrt{k_n e^{-|\beta|-\varepsilon_1}} | E_m\big) \cdot \mathbb{P}(E_m) \\
&\quad\geq \sum_{m=s}^{k_n} \mathbb{P}\big(\mathrm{LIS}(\eta_{k_n}) > (2-\varepsilon_2)\sqrt{m} | E_m\big) \cdot \mathbb{P}(E_m) \\
&\quad= \sum_{m=s}^{k_n} \lambda_m\big(\mathrm{LIS}(\pi) > (2-\varepsilon_2)\sqrt{m}\big) \cdot \mathbb{P}(E_m) \\
&\quad> (1-\varepsilon_3) \sum_{m=s}^{k_n} \mathbb{P}(E_m) \\
&\quad= (1-\varepsilon_3)\mathbb{P}\left(\sum_{i=1}^{k_n} B_{n,i} > k_n e^{-|\beta|-\varepsilon_1}\right) \\
&\quad> (1-\varepsilon_3)^2.
\end{aligned}$$

The second inequality above follows since $m \geq s \geq k_n e^{-|\beta|-\varepsilon_1}$. The third inequality follows from (85) and the fact that $m \geq k_n e^{-|\beta|-\varepsilon_1} > M$. Therefore, we have shown that $\lim_{n\to\infty} \mathbb{P}(\mathrm{LIS}(\eta_{k_n}) > (2-\varepsilon_2)\sqrt{k_n e^{-|\beta|-\varepsilon_1}}) = 1$, and (82) follows from the fact that, by choosing ε_1 and ε_2 small enough, $(2-\varepsilon_2)\sqrt{e^{-|\beta|-\varepsilon_1}}$ can be arbitrarily close to $2e^{\frac{-|\beta|}{2}}$.

For the second part, we need to show that, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \max_{\boldsymbol{b}\in Q(n,k_n)} \mu_{n,q_n}\left(\pi \in S_n : \frac{\mathrm{LIS}(\pi_{\boldsymbol{b}})}{\sqrt{k_n}} \geq 2e^{\frac{|\beta|}{2}} + \varepsilon\right) = 0. \tag{86}$$

Similar to the proof of (83), we can show that

$$\lim_{n\to\infty} \max\big(q_n^n, 1/q_n^n\big) = e^{|\beta|} < 2. \tag{87}$$

The last inequality follows since $|\beta| < \ln 2$. Thus, for any $0 < \varepsilon_1 < \ln 2 - |\beta|$, there exists $N_1 > 0$ such that, for all $n > N_1$, we have

$$\max\big(q_n^n, 1/q_n^n\big) < e^{|\beta|+\varepsilon_1} < 2.$$

Given $n > N_1$, for any $\boldsymbol{b} \in Q(n, k_n)$, by Lemma 5.9, $\mathrm{LIS}(\xi_{\boldsymbol{b}}^{(n)})$ and $\mathrm{LIS}(\pi_{\boldsymbol{b}})$ have the same distribution, where $\xi_{\boldsymbol{b}}^{(n)}$ is the point process defined in that lemma. Moreover, by Lemma 5.7, there exists a point process ζ_{k_n} such that $\xi_{\boldsymbol{b}}^{(n)} \leq \zeta_{k_n}$ almost surely and ζ_{k_n} is defined by

$$\zeta_{k_n}(A) := \sum_{i=1}^{k_n} \mathbb{1}_A\big((i, U_i')\big) + B_{n,i} \cdot \mathbb{1}_A\big((i, U_i)\big) \qquad \forall A \in \mathcal{B}(\mathbb{N} \times [0,1]), \tag{88}$$

where $\{U_i\}_{i=1}^{k_n} \cup \{U_i'\}_{i=1}^{k_n} \cup \{B_{n,i}\}_{i=1}^{k_n}$ are $3k_n$ independent random variables with each U_i, U_i' being uniformly distributed on $[0,1]$ and each $B_{n,i}$ being a Bernoulli random variable with $\mathbb{P}(B_{n,i} = 1) = \max(q_n^n, 1/q_n^n) - 1$. Hence, by the monotonicity of $\mathrm{LIS}(\cdot)$ on $\mathcal{X}_2$, we have

$$\mu_{n,q_n}\bigg(\pi \in S_n : \frac{\mathrm{LIS}(\pi_{\boldsymbol{b}})}{\sqrt{k_n}} \geq 2e^{\frac{|\beta|}{2}} + \varepsilon\bigg) \leq \mathbb{P}\bigg(\frac{\mathrm{LIS}(\zeta_{k_n})}{\sqrt{k_n}} \geq 2e^{\frac{|\beta|}{2}} + \varepsilon\bigg).$$

We complete the proof of (86) as well as Lemma 4.2 by showing that, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\bigg(\frac{\mathrm{LIS}(\zeta_{k_n})}{\sqrt{k_n}} < 2e^{\frac{|\beta|}{2}} + \varepsilon\bigg) = 1. \tag{89}$$

Since, for all $n > N_1$, we have $\mathbb{P}(B_{n,i} = 1) = \max(q_n^n, 1/q_n^n) - 1 < e^{|\beta|+\varepsilon_1} - 1$, by the law of large numbers, we get

$$\lim_{n\to\infty} \mathbb{P}\Bigg(\sum_{i=1}^{k_n} B_{n,i} < k_n\big(e^{|\beta|+\varepsilon_1} - 1\big)\Bigg) = 1. \tag{90}$$

Given $\boldsymbol{U}' = (U_1', \dots, U_{k_n}')$, $\boldsymbol{U} = (U_1, \dots, U_{k_n})$ and $\boldsymbol{B} = (B_{n,1}, \dots, B_{n,k_n})$, let $\Lambda(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B})$ denote the set of points in $\mathbb{R}^2$ defined by

$$\Lambda\big(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B}\big) := \big\{(i, U_i) : i \in [k_n] \text{ and } B_{n,i} = 1\big\} \cup \big\{\big(i, U_i'\big) : i \in [k_n]\big\}.$$

By the definition of ζ_{k_n} and Definition 5.8 we have

$$\mathrm{LIS}(\zeta_{k_n}) = \mathrm{LIS}\big(\Lambda(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B})\big). \tag{91}$$

Based on $\boldsymbol{U}'$, $\boldsymbol{U}$ and $\boldsymbol{B}$, define another set of points in $\mathbb{R}^2$ as follows:

$$\Lambda^+\big(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B}\big) := \big\{(i + 1/2, U_i) : i \in [k_n] \text{ and } B_{n,i} = 1\big\} \cup \big\{\big(i, U_i'\big) : i \in [k_n]\big\}.$$

Then we have

$$\mathrm{LIS}\big(\Lambda\big(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B}\big)\big) \leq \mathrm{LIS}\big(\Lambda^+\big(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B}\big)\big). \tag{92}$$

Since, by Definition 2.1, no two points with the same x coordinate can be both within an increasing subsequence, by increasing the x coordinates of those points in $\Lambda(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B})$ which reside on the same vertical line as other points by 1/2, the relative ordering of the shifted point with other points does not change, except the one which has the same x coordinate when unshifted. Combining (91) and (92), we have

$$\mathrm{LIS}(\zeta_{k_n}) \leq \mathrm{LIS}\big(\Lambda^+\big(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B}\big)\big). \tag{93}$$

Moreover, conditioned on $\sum_{i=1}^{k_n} B_{n,i} = m$, by independence of $\boldsymbol{U}', \boldsymbol{U}$ and $\boldsymbol{B}$, it is easily seen that $\mathrm{LIS}(\Lambda^+(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B}))$ has the same distribution as $\mathrm{LIS}(\pi)$ with $\pi \sim \lambda_{k_n+m}$. For any $0 < \varepsilon_2, \varepsilon_3 < 1$, by the result of Kerov and Vershik (1977) again, there exists $M > 0$ such that, for any $k > M$,

$$\lambda_k\left(\frac{\mathrm{LIS}(\pi)}{\sqrt{k}} < 2 + \varepsilon_2\right) > 1 - \varepsilon_3.$$

Since $\lim_{n\to\infty} k_n = \infty$ and (90), there exists $N > N_1$ such that, for any $n > N$, we have

$$k_n > M \quad \text{and} \quad \mathbb{P}\left(\sum_{i=1}^{k_n} B_{n,i} < k_n\big(e^{|\beta|+\varepsilon_1} - 1\big)\right) > 1 - \varepsilon_3.$$

Let $s := \lceil k_n(e^{|\beta|+\varepsilon_1} - 1)\rceil - 1$. Recall that E_m denotes the event $\{\sum_{i=1}^{k_n} B_{n,i} = m\}$. For any $n > N$, we have

$$\begin{aligned}
&\mathbb{P}\big(\mathrm{LIS}(\zeta_{k_n}) < (2+\varepsilon_2)\sqrt{k_n e^{|\beta|+\varepsilon_1}}\big) \\
&\quad \geq \sum_{m=0}^{s} \mathbb{P}\big(\mathrm{LIS}(\zeta_{k_n}) < (2+\varepsilon_2)\sqrt{k_n e^{|\beta|+\varepsilon_1}} | E_m\big) \cdot \mathbb{P}(E_m) \\
&\quad \geq \sum_{m=0}^{s} \mathbb{P}\big(\mathrm{LIS}(\zeta_{k_n}) < (2+\varepsilon_2)\sqrt{k_n + m} | E_m\big) \cdot \mathbb{P}(E_m) \\
&\quad \geq \sum_{m=0}^{s} \mathbb{P}\big(\mathrm{LIS}\big(\Lambda^+\big(\boldsymbol{U}', \boldsymbol{U}, \boldsymbol{B}\big)\big) < (2+\varepsilon_2)\sqrt{k_n + m} | E_m\big) \cdot \mathbb{P}(E_m) \\
&\quad = \sum_{m=0}^{s} \lambda_{k_n+m}\big(\mathrm{LIS}(\pi) < (2+\varepsilon_2)\sqrt{k_n + m}\big) \cdot \mathbb{P}(E_m) \\
&\quad > (1-\varepsilon_3) \sum_{m=0}^{s} \mathbb{P}(E_m) \\
&\quad = (1-\varepsilon_3)\mathbb{P}\left(\sum_{i=1}^{k_n} B_{n,i} < k_n\big(e^{|\beta|+\varepsilon_1} - 1\big)\right) \\
&\quad > (1-\varepsilon_3)^2.
\end{aligned}$$

The second inequality follows because

$$k_n + m \leq k_n + s < k_n + k_n\big(e^{|\beta|+\varepsilon_1} - 1\big) = k_n e^{|\beta|+\varepsilon_1},$$

and the third inequality follows from (93). Therefore, we have shown that $\lim_{n\to\infty} \mathbb{P}(\mathrm{LIS}(\zeta_{k_n}) < (2+\varepsilon_2)\sqrt{k_n e^{|\beta|+\varepsilon_1}}) = 1$ and (89) follows from the fact that, by choosing ε_1 and ε_2 small enough, $(2+\varepsilon_2)\sqrt{e^{|\beta|+\varepsilon_1}}$ can be arbitrarily close to $2e^{\frac{|\beta|}{2}}$. □

6. Discussion and open questions.

1. Consider the partially ordered set $(S_n, \le_L)$; we conjecture the following stochastic dominance of Mallows measure: for any $0 < q < q'$, we have $\mu_{n,q} \preceq \mu_{n,q'}$, that is, $\mu_{n,q}$ is stochastically dominated by $\mu_{n,q'}$. By Strassen's theorem [cf. Lindvall et al. (1999)], the conjecture is equivalent to the following statement: there exists a coupling (X, Y) with $X \sim \mu_{n,q}$ and $Y \sim \mu_{n,q'}$ such that $X \le_L Y$.

2. In the proof of Corollary 1.5, we show that $\bar{J}$ defined in Theorem 1.4 is attained when $\phi(x) = x$ given that $\lim_{n\to\infty} n(1 - q_n) = \lim_{n\to\infty} n(1 - q_n') = \beta$. In fact, for any $\beta \in \mathbb{R}$, if $\gamma = 0, \beta, \pm\infty$, taking $\phi(x)$ to be the diagonal of the unit square gives the supremum of the following variational problem:

$$\sup_{\phi \in B^1_{\nearrow}} \int_0^1 \sqrt{\dot{\phi}(x)\rho\big(x, \phi(x)\big)}\, dx.$$

Note that, when $\gamma = \pm\infty$, we extend the definition of $\rho(x, y, \beta, \gamma)$ as follows (we explicitly add β, γ as the argument of the density ρ.):

$$\rho(x, y, \beta, \pm\infty) := \lim_{\gamma\to\pm\infty} \rho(x, y, \beta, \gamma) = \lim_{\gamma\to\pm\infty} \int_0^1 u(x, t, \beta) \cdot u(t, y, \gamma)\, dt.$$

In fact, it is not hard to show that the above limits exist with

$$\rho(x, y, \beta, \infty) = u(x, y, \beta) \quad \text{and} \quad \rho(x, y, \beta, -\infty) = u(x, y, -\beta).$$

It is unknown to us whether $\phi(x) = x$ solves the above variational problem for arbitrary $\beta, \gamma \in \mathbb{R}$.

Acknowledgments. I am grateful to my supervisor, Nayantara Bhatnagar, for her helpful advice and suggestions during the research as well as her guidance in the completion of this paper. I also thank an anonymous referee for pointing out an incorrect proof in the first draft of this paper as well as giving me detailed typographical and grammatical suggestions.

REFERENCES

ABELLO, J. (1991). The weak Bruhat order of S_Σ consistent sets, and Catalan numbers. *SIAM J. Discrete Math.* **4** 1–16. MR1090284

BAIK, J., DEIFT, P. and JOHANSSON, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** 1119–1178.

BHATNAGAR, N. and PELED, R. (2015). Lengths of monotone subsequences in a Mallows permutation. *Probab. Theory Related Fields* **161** 719–780.

CHVÁTAL, V. and SANKOFF, D. (1975). Longest common subsequences of two random sequences. *J. Appl. Probab.* **12** 306–315. MR0405531

CRITCHLOW, D. E. (1985). *Metric Methods for Analyzing Partially Ranked Data*. *Lecture Notes in Statistics* **34**. Springer, Berlin. MR0818986

DANČÍK, V. (1994). Expected length of longest common subsequences. Ph.D. dissertation, Univ. Warwick.

DANČÍK, V. and PATERSON, M. (1995). Upper bounds for the expected length of a longest common subsequence of two binary sequences. *Random Structures Algorithms* **6** 449–458. MR1368846

DEKEN, J. G. (1979). Some limit results for longest common subsequences. *Discrete Math*. **26** 17–31.

DEUSCHEL, J.-D. and ZEITOUNI, O. (1995). Limiting curves for i.i.d. records. *Ann*. *Probab.* **23** 852–878. MR1334175

FLIGNER, M. A. and VERDUCCI, J. S. (1993). *Probability Models and Statistical Analyses for Ranking Data*. *Lecture Notes in Statistics* **80**. Springer, Berlin.

HAMMERSLEY, J. (1972). A few seedlings of research. In *Proc*. *of the Sixth Berkeley Symp*. *Math*. *Statist*. *and Probability*, *Vol*. 1 345–394. Univ. California Press, Berkeley, CA. MR0405665

HOPPEN, C., KOHAYAKAWA, Y., MOREIRA, C. G., RÁTH, B. and SAMPAIO, R. M. (2013). Limits of permutation sequences. *J*. *Combin*. *Theory Ser. B* **103** 93–113.

HOUDRÉ, C. and IŞLAK, Ü. (2014). A central limit theorem for the length of the longest common subsequence in random words. Available at: arXiv:1408.1559.

JIN, K. (2017). The limit of the empirical measure of the product of two independent Mallows permutations. Available at: arXiv:1702.00140.

KENYON, R., KRAL, D., RADIN, C. and WINKLER, P. (2015). Permutations with fixed pattern densities. Available at: arXiv:1506.02340.

KEROV, S. and VERSHIK, A. (1977). Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux. *Sov. Math*. *Dokl*. **18** 527–531.

LINDVALL, T. et al. (1999). On Strassen's theorem on stochastic domination. *Electron*. *Commun*. *Probab.* **4** 51–59.

LOGAN, B. F. and SHEPP, L. A. (1977). A variational problem for random Young tableaux. *Adv. Math*. **26** 206–222.

LUEKER, G. S. (2009). Improved bounds on the average length of longest common subsequences. *J. ACM* **56** 17.

MALLOWS, C. L. (1957). Non-null ranking models. I. *Biometrika* **44** 114–130. MR0087267

MARDEN, J. I. (1995). *Analyzing and Modeling Rank Data*. *Monographs on Statistics and Applied Probability* **64**. Chapman & Hall, London. MR1346107

MUELLER, C. and STARR, S. (2013). The length of the longest increasing subsequence of a random Mallows permutation. *J*. *Theoret*. *Probab.* **26** 514–540. MR3055815

MUKHERJEE, S. et al. (2016). Fixed points and cycle structure of random permutations. *Electron*. *J*. *Probab.* **21**.

PEVZNER, P. (2000). *Computational Molecular Biology*: *An Algorithmic Approach*. MIT Press, Cambridge, MA.

STARR, S. (2009). Thermodynamic limit for the Mallows model on S_n. Available at: arXiv:0904.0696.

DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF DELAWARE
15 ORCHARD RD
NEWARK, DELAWARE 19711
USA
E-MAIL: kejin@udel.edu