

# LIMIT SHAPES OF BUMPING ROUTES IN THE ROBINSON-SCHENSTED CORRESPONDENCE

DAN ROMIK AND PIOTR ŚNIADY

ABSTRACT. We prove a limit shape theorem describing the asymptotic shape of bumping routes when the Robinson-Schensted algorithm is applied to a finite sequence of independent, identically distributed random variables with the uniform distribution  $U[0, 1]$  on the unit interval, followed by an insertion of a deterministic number  $\alpha$ . The bumping route converges after scaling, in the limit as the length of the sequence tends to infinity, to an explicit, deterministic curve depending only on  $\alpha$ . This extends our previous result on the asymptotic determinism of Robinson-Schensted insertion, and answers a question posed by Moore in 2006.

## 1. INTRODUCTION

Let  $\mathfrak{S}_n$  denote the symmetric group of order  $n$ . Recall that the **Robinson-Schensted correspondence** associates with a permutation  $\sigma^{(n)} \in \mathfrak{S}_n$  a pair of standard Young tableaux  $(P_n, Q_n)$  whose common shape  $\lambda$  is a Young diagram of order  $n$ . A fruitful area of study concerns asymptotic properties of the Robinson-Schensted shape  $\lambda$  and the tableaux  $P_n, Q_n$  associated with a *random* permutation  $\sigma^{(n)}$  sampled from the uniform distribution on  $\mathfrak{S}_n$ . The existing results on this subject are too numerous to list here, but some of the important highlights of the theory are the limit shape result of Logan-Shepp [LS77] and Vershik-Kerov [VK77, VK85], which led to the solution of the so-called Ulam-Hammersley problem on the typical length of a longest increasing subsequence in random permutations; and the celebrated Baik-Deift-Johansson theorem [BDJ99a] and its refinements and variants [BDJ99b, BOO00, Joh00] that tied the behavior of longest increasing subsequences in random permutations to the Tracy-Widom distribution and other naturally-occurring stochastic processes from random

---

2010 *Mathematics Subject Classification.* 68Q87 (Primary); 60C05, 05E10, 20C30 (Secondary).

*Key words and phrases.* Robinson-Schensted correspondence, bumping routes, Young tableau, limit shape.

matrix theory. See the book [Rom14] for a survey of many of these developments that also touches on diverse connections to random growth processes, interacting particle systems, representation theory and more.

In this paper we continue this line of investigation by studying the **bumping route** computed during the application of an insertion procedure, which is the fundamental building block of the Robinson-Schensted correspondence. Let us recall briefly the relevant definitions. A **Young diagram**  $\lambda$  of order  $n$  is an integer partition of  $n$ , that is, an array of positive integers  $\lambda_1 \geq \dots \geq \lambda_m \geq 0$  such that  $n = \sum_{j=1}^m \lambda_j$ , represented graphically as a diagram of left-justified square boxes wherein the  $j$ th row contains  $\lambda_j$  boxes. If  $\lambda$  is a Young diagram of order  $n$  and  $x_1, \dots, x_n$  are distinct real numbers, an **increasing tableau** of shape  $\lambda$  and entries given by  $x_1, \dots, x_n$  is a filling of the boxes of  $\lambda$  with the numbers  $x_1, \dots, x_n$  that is increasing along rows and columns. A **standard Young tableau** is such an increasing tableau whose entries are precisely the numbers  $1, \dots, n$ .

Given an increasing tableau  $P$  with entries  $x_1, \dots, x_n$  and a number  $z$  distinct from  $x_1, \dots, x_n$ , the **insertion procedure** applied to  $P$  and  $z$  produces a new increasing tableau  $P \leftarrow z$  with entries  $x_1, \dots, x_n, z$  whose shape  $\lambda^+$  is obtained from  $\lambda$  by the addition of a single box. The new tableau  $P \leftarrow z$  is computed by performing a succession of **bumping steps** whereby the number  $z$  is inserted into the first row of the diagram, displacing an existing entry from the first row; the displaced entry is bumped onto the second row, and in turn bumps an entry of the second row onto the third row; and so on, until finally the entry being bumped settles down in an unoccupied position outside the diagram  $\lambda$ . In each row, the position where the bumping (or settling-down, in the last step) occurs is the leftmost position containing an entry bigger than the incoming number, or a new unoccupied position to the right of all existing entries if no such entry exists. An example is shown in Figure 1.

Define the **bumping route**  $\mathbf{B}_{P,z}$  associated with an insertion procedure performed on the tableau  $P$  with a new input  $z$  to be the sequence of positions where a bumping occurred during the insertion, together with the position of the final box added to the shape. The  $j$ th position in the bumping route is of the form  $(b_{P,z}(j), j)$ , so it is convenient to encode the bumping route using only the  $x$ -coordinates of the positions, which form a monotone nonincreasing sequence of positive integers  $b_{P,z}(1) \geq b_{P,z}(2) \geq \dots \geq b_{P,z}(k_{P,z})$  whose length we denote by  $k_{P,z}$ . For example, the bumping route associated with the insertion step in Figure 1 is  $(4, 3, 2, 2, 2)$ .

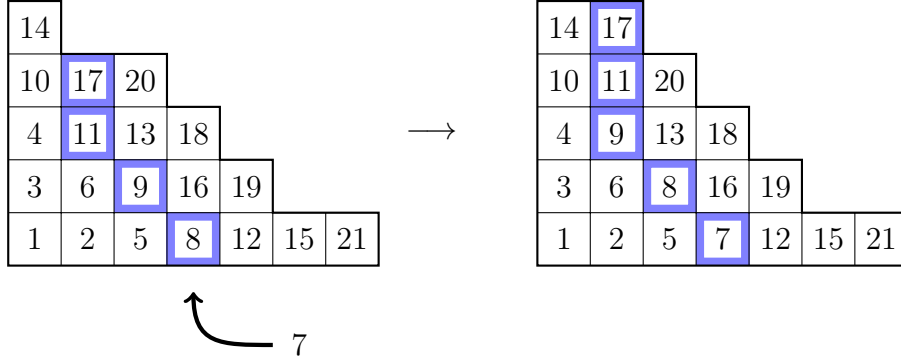


Figure 1. Inserting a number into a tableau results in a cascade of bumping events. The bumping route is the sequence of positions where a bumping occurred.

The **insertion tableau** corresponding to a sequence  $x_1, \dots, x_n$  is defined as the outcome of the iterative application of the insertion procedure

$$P(x_1, \dots, x_n) := \left( \left( \left( \emptyset \leftarrow x_1 \right) \leftarrow x_2 \right) \leftarrow \dots \right) \leftarrow x_n,$$

starting with the empty tableau  $\emptyset$ . The **Robinson-Schensted correspondence** associates with a permutation  $\sigma^{(n)} \in \mathfrak{S}_n$  a pair of standard Young tableaux  $(P_n, Q_n)$  with the first one  $P_n = P(\sigma^{(n)}) = P(\sigma^{(n)}(1), \dots, \sigma^{(n)}(n))$  being the insertion tableau corresponding to the permutation; the definition of the second tableau  $Q_n$ , known as the **recording tableau**, will not be necessary for our purposes. More details on Young tableaux, the Robinson-Schensted correspondence and their properties can be found in several well-known sources such as [Ful97, Knu98, Sta99].

If  $\sigma^{(n)}$  is a uniformly random permutation of order  $n$ , the bumping route computed in the last insertion step performed while calculating the insertion tableau  $P_n$  is equal to  $\mathbf{B}_{P_{n-1}, \sigma^{(n)}(n)}$ , where  $P_{n-1} = P(\sigma^{(n)}(1), \dots, \sigma^{(n)}(n-1))$  denotes the insertion tableau computed from the truncated sequence. The question we wish to address is that of understanding the asymptotic behavior of this bumping route.

Since the computation depends only on the relative order of the numbers  $\sigma^{(n)}(j)$ ,  $1 \leq j \leq n$ , it will be equivalent, and more convenient, to formulate the result in terms of a sequence  $X_1, \dots, X_n$  of independent and identically distributed (i.i.d.) random variables with the uniform distribution  $U[0, 1]$  on the unit interval  $[0, 1]$ , which gives a canonical

way of realizing uniformly random order structures of all orders  $n$  simultaneously. If we denote by  $T_n = P(X_1, \dots, X_n)$  the corresponding insertion tableau, then the bumping route  $\mathbf{B}_{T_{n-1}, X_n}$  is equal in distribution to  $\mathbf{B}_{P_{n-1}, \sigma^{(n)}(n)}$ .

Our result will pertain to an even more general scenario in which the final input  $X_n$  is taken to be an arbitrary (non-random) number  $\alpha$  in the unit interval  $[0, 1]$ . Note that, by obvious monotonicity properties of the insertion procedure, as  $\alpha$  increases from 0 to 1, the bumping route is deformed monotonically between the two extreme cases  $\alpha = 0$  and  $\alpha = 1$ , where in the case  $\alpha = 0$  the bumping route will be the first column of the diagram and an additional new box at the top of the first column, and in the case  $\alpha = 1$  the bumping route consists of a single new box at the end of the first row of the diagram. Note also that the bumping route (except for the last box) is contained in the Young diagram of the tableau  $T_{n-1}$ ; this random Young diagram, whose distribution is known as the **Plancherel measure** of order  $n - 1$ , converges to a well-known limit shape discovered in the celebrated works of Logan-Shepp [LS77] and Vershik-Kerov [VK77, VK85].

Figure 2a shows the bumping routes  $\mathbf{B}_{T_{n-1}, \alpha}$  for various values of  $\alpha$  in a numerical simulation with  $n = 10^4$ . Our goal will be to show that the bumping routes converge after scaling to a family of deterministic limiting curves, which are shown in Figure 2b. As preparation for the precise formulation of this result, let us first define this family of curves. First, define auxiliary functions  $F$ ,  $\Omega$ ,  $u_\alpha$ ,  $v_\alpha$ ,  $x_\alpha$ ,  $y_\alpha$ ,  $\kappa$  by

$$\begin{aligned}
 (1) \quad \Omega(u) &= \frac{2}{\pi} \left( u \sin^{-1} \left( \frac{u}{2} \right) + \sqrt{4 - u^2} \right) && (|u| \leq 2), \\
 (2) \quad F(u) &= \frac{1}{2} + \frac{1}{\pi} \left( \frac{u\sqrt{4 - u^2}}{4} + \sin^{-1} \left( \frac{u}{2} \right) \right) && (|u| \leq 2), \\
 (3) \quad u_\alpha(t) &= \sqrt{t} F^{-1} \left( \frac{\alpha}{t} \right) && (0 \leq \alpha \leq t \leq 1), \\
 (4) \quad v_\alpha(t) &= \sqrt{t} \Omega \left( F^{-1} \left( \frac{\alpha}{t} \right) \right) && (0 \leq \alpha \leq t \leq 1), \\
 (5) \quad x_\alpha(t) &= \frac{v_\alpha(t) + u_\alpha(t)}{2} && (0 \leq \alpha \leq t \leq 1), \\
 (6) \quad y_\alpha(t) &= \frac{v_\alpha(t) - u_\alpha(t)}{2} && (0 \leq \alpha \leq t \leq 1), \\
 (7) \quad \kappa(\alpha) &= y_\alpha(1) = \frac{\Omega(F^{-1}(\alpha)) - F^{-1}(\alpha)}{2} && (0 \leq \alpha \leq 1).
 \end{aligned}$$

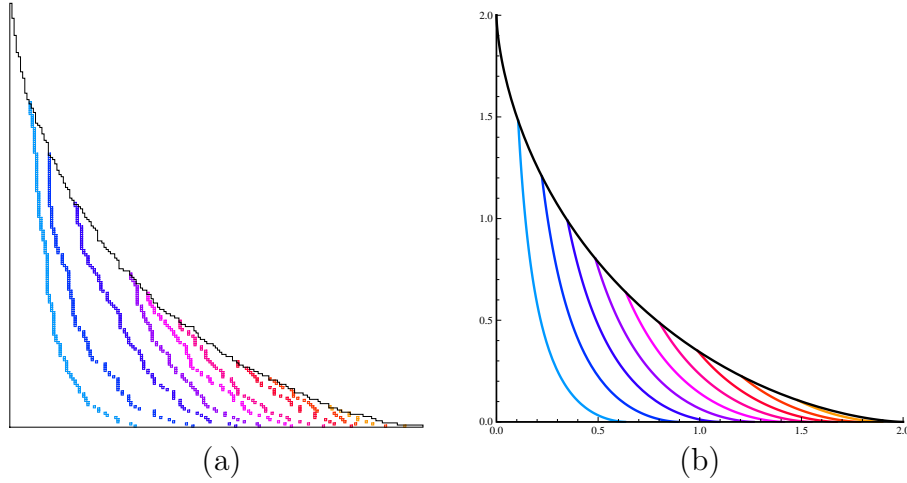


Figure 2. (a) Bumping routes  $\mathbb{B}_{T_{n-1}, \alpha}$  in a simulation with  $n = 10^4$  and the values  $\alpha = 1/10, 2/10, \dots, 9/10$ ; (b) the limiting curves  $(\beta_\alpha(s), s)$  for the same values of  $\alpha$ . The region bounding the limit shapes is the Logan-Shepp-Vershik-Kerov limit shape of Plancherel-random Young diagrams [Rom14].

The **limiting bumping route curves** are now defined as the one-parameter family  $(\beta_\alpha(t))_{0 \leq \alpha < 1}$  of functions, where for each  $\alpha \in [0, 1)$ ,  $\beta_\alpha(\cdot)$  is given by

$$(8) \quad \beta_\alpha(s) = x_\alpha(y_\alpha^{-1}(s)) \quad (0 \leq s \leq \kappa(\alpha)).$$

Our main result is as follows.

**Theorem 1** (Limit shapes of bumping routes). *For each  $0 \leq \alpha < 1$ , the curve  $\beta_\alpha(\cdot)$  describes the limiting bumping route  $\mathbb{B}_{T_{n-1}, \alpha}$ , in the following precise sense: for any  $\epsilon > 0$ , we have that*

$$(9) \quad \mathbb{P} \left( \left| \frac{k_{T_{n-1}, \alpha}}{\sqrt{n}} - \kappa(\alpha) \right| > \epsilon \right) \xrightarrow{n \rightarrow \infty} 0, \text{ and}$$

$$(10) \quad \mathbb{P} \left( \max_{1 \leq m \leq k_{T_{n-1}, \alpha}} \left| \frac{b_{T_{n-1}, \alpha}(m)}{\sqrt{n}} - \beta_\alpha \left( \frac{m}{\sqrt{n}} \wedge \kappa(\alpha) \right) \right| > \epsilon \right) \xrightarrow{n \rightarrow \infty} 0.$$

The problem of understanding the limit shapes of the bumping routes, which our result answers, was posted by C. Moore on his personal web page (along with simulation results similar to our Figure 2a) in 2006 [Moo06].

## 2. PRELIMINARY REMARKS

As a first step towards proving Theorem 1, let us recall some facts from the theory of Plancherel measure, which will immediately prove (9) and also help elucidate the somewhat involved definition of the family of limiting curves  $\beta_\alpha(\cdot)$ . First, when discussing Plancherel-random Young diagrams and Young tableaux it is convenient to use the so-called rotated (also known as “Russian”) coordinate system, related to the standard  $x$ - $y$  coordinates by the linear change of variables

$$(11) \quad \begin{aligned} u &= x - y, \\ v &= x + y. \end{aligned}$$

In this coordinate system, the curve  $v = \Omega(u)$ , where  $\Omega(u)$  is defined in (1), describes the Logan-Shepp-Vershik-Kerov limit shape of Plancherel-random Young diagrams mentioned in the introduction.

Second, the function  $F(u)$  defined in (2) is the cumulative distribution function of the semicircle distribution on  $[-2, 2]$ ; that is, we have

$$F(u) = \frac{1}{2\pi} \int_{-2}^u \sqrt{4 - s^2} ds \quad (|u| \leq 2).$$

Its importance for the present discussion is that, according to one of the main results of our previous paper [RS14], the point

$$(12) \quad (U(\alpha), V(\alpha)) = (F^{-1}(\alpha), \Omega(F^{-1}(\alpha)))$$

is the limiting scaled position (in rotated coordinates) of the new box added to the Robinson-Schensted shape after applying the insertion procedure with the number  $\alpha \in [0, 1]$  to the existing insertion tableau  $T_{n-1}$ . More precisely, when stated using our current terminology, the result [RS14, Theorem 5.1], which we dubbed the “**asymptotic determinism of RSK insertion**,” says that the last position  $(b_\alpha(k_{T_{n-1}, \alpha}), k_{T_{n-1}, \alpha})$  of the bumping route  $\mathbf{B}_{T_{n-1}, \alpha}$  satisfies

$$(13) \quad \frac{1}{\sqrt{n}} (b_\alpha(k_{T_{n-1}, \alpha}) - k_{T_{n-1}, \alpha}, b_\alpha(k_{T_{n-1}, \alpha}) + k_{T_{n-1}, \alpha}) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} (U(\alpha), V(\alpha)).$$

After applying the inverse transformation of (11) to rewrite the result in  $x$ - $y$  coordinates, and noting that

$$\kappa(\alpha) = y_\alpha(1) = \frac{V(\alpha) - U(\alpha)}{2},$$

we get the relation (9), the first claim of Theorem 1.

Next, turn to the pair of functions  $(u_\alpha(t), v_\alpha(t))$  defined in (3)–(4), with the associated pair  $(x_\alpha(t), y_\alpha(t))$  from (5)–(6) representing the same functions in  $x$ - $y$  coordinates. Note that for any fixed  $\alpha$ , the planar curve  $(x_\alpha(t), y_\alpha(t))_{\alpha \leq t \leq 1}$  is a reparametrized version of the curve  $(\beta_\alpha(s), s)_{0 \leq s \leq \kappa(\alpha)}$ , which according to our claim (10) is the limit shape of the bumping route  $\mathbf{B}_{T_{n-1}, \alpha}$  (one needs to note that  $y_\alpha(\cdot)$  is a strictly decreasing function; see Lemma 2 below). It turns out that the parametrization of the curve as  $(x_\alpha(t), y_\alpha(t))$  is the correct one when trying to prove the limit shape result (although the parametrization  $(\beta_\alpha(s), s)$  is the one that answers the original question). To see why, we need to explain the role of the parameter  $t$ . Note that for fixed  $t$ , the points  $(u_\alpha(t), v_\alpha(t))_{0 \leq \alpha \leq 1}$  all lie on the curve  $v = \sqrt{t} \Omega(u/\sqrt{t})$ , which is a copy of the limit shape  $v = \Omega(u)$  scaled down by the factor  $\sqrt{t}$ . The idea is that this scaled-down copy represents the limiting shape of the “ $t$ -sublevel tableau” of  $T_{n-1}$ —that is, the subset of boxes of  $T_{n-1}$  containing an entry  $\leq t$ . We will show below that the point  $(u_\alpha(t), v_\alpha(t))$  (or  $(x_\alpha(t), y_\alpha(t))$ , in the usual coordinate system) corresponds to the limiting position, after scaling, of the point at which the bumping route  $\mathbf{B}_{T_{n-1}, \alpha}$  exits this sublevel tableau. The reason for this is that this exit position relates to the sublevel tableau in roughly the same way that the final bumping route position  $(b_\alpha(k_{T_{n-1}, \alpha}), k_{T_{n-1}, \alpha})$  relates to the entire tableau  $T_{n-1}$ , except that there is a scaling relation that causes the number being inserted to change from  $\alpha$  to  $\alpha/t$ . A more precise formulation of this statement is discussed in the next section, after which we will see that our main result follows without much difficulty by another appeal to the “asymptotic determinism of RSK” theorem. A schematic illustration of the argument described above is shown in Figure 3.

We conclude this section with another small but useful observation.

**Lemma 2.** *The function  $y_\alpha(\cdot)$  is strictly decreasing. In particular, the limit shape functions  $\beta_\alpha(\cdot)$  defined in (8) are well-defined.*

*Proof.* Denote

$$(X(\alpha), Y(\alpha)) = \frac{1}{2}(V(\alpha) + U(\alpha), V(\alpha) - U(\alpha)),$$

where  $U(\alpha), V(\alpha)$  are defined in (12). First, note that  $Y(\cdot)$  is strictly decreasing, since it is the composition of the increasing function  $\alpha \mapsto F^{-1}(\alpha)$  with the function  $u \mapsto \frac{1}{2}(\Omega(u) - u)$ , the latter being decreasing (this can be seen by interpreting this function geometrically, or by differentiating and using the fact that  $\Omega'(u) = \frac{2}{\pi} \sin^{-1}(u/2)$ ). Now, if

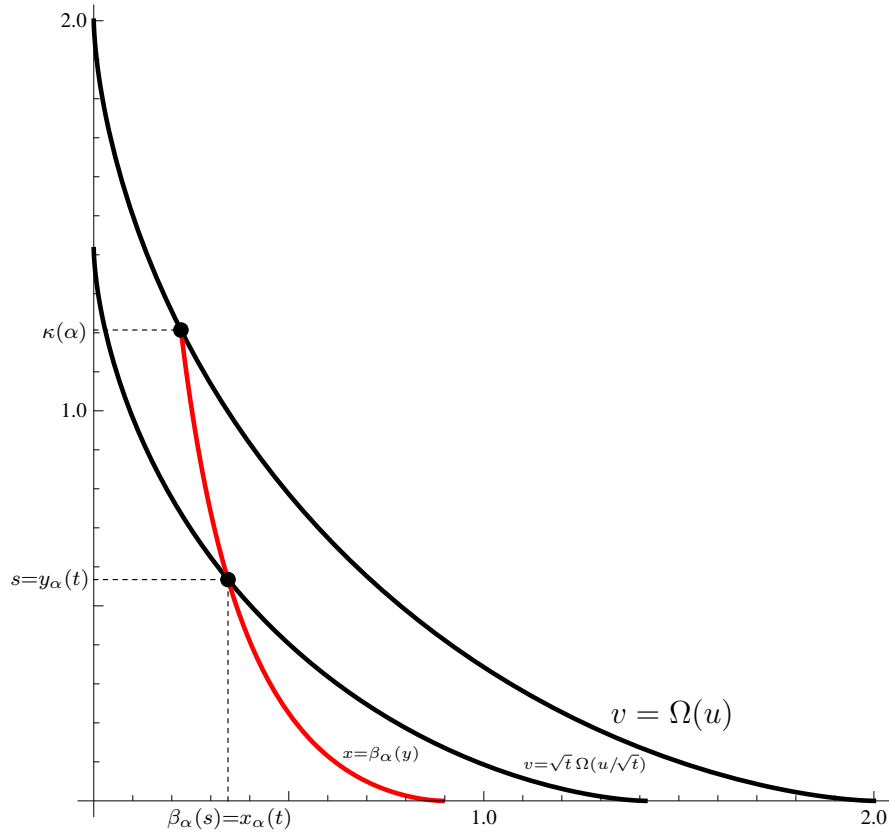


Figure 3. The meaning of the parameter  $t$ : for a fixed value of  $\alpha$ , the intersection  $(u_\alpha(t), v_\alpha(t))$  of the asymptotic bumping curve with the scaled-down copy  $v = \sqrt{t}\Omega(u/\sqrt{t})$  of the Logan-Shepp-Vershik-Kerov limit shape is computed by applying the asymptotic determinism theorem to the  $t$ -sublevel tableau; when the numbers in the sublevel tableau are scaled to the range  $[0, 1]$ , the number  $\alpha$  being inserted is scaled to  $\alpha/t$ .

$t < t'$  are numbers in  $[\alpha, 1]$  then

$$y_\alpha(t) = \sqrt{t}Y(\alpha/t) < \sqrt{t'}Y(\alpha/t) < \sqrt{t'}Y(\alpha/t') = y_\alpha(t'),$$

proving the claim.  $\square$

### 3. PLANCHEREL MEASURE AND SUBLEVEL TABLEAUX

Recall that the distribution of the insertion tableau  $P_n$  associated via the Robinson-Schensted correspondence with a uniformly random permutation  $\sigma^{(n)}$  in  $\mathfrak{S}_n$  is also called the Plancherel measure of order



$n$  (this refers to the measure on standard Young tableaux; the distribution of the shape of this tableau is the Plancherel measure on Young diagrams of order  $n$  that was mentioned in the introduction). By the remarks made in the introduction, a tableau  $P_n$  with this distribution arises by taking the tableau  $T_n$  and “standardizing” it by replacing each entry  $X_k$  with the ranking of  $X_k$  in the list  $X_1, \dots, X_n$  (i.e., the number  $\sigma(k)$  such that  $X_k = X_{\sigma(k)}^{(n)}$ , where  $X_1^{(n)} \leq \dots \leq X_n^{(n)}$  are the order statistics of  $X_1, \dots, X_n$ ). Note also that the vector  $(X_1^{(n)}, \dots, X_n^{(n)})$  of order statistics is independent of the tableau  $P_n$  and is distributed uniformly on the simplex

$$\Delta_n = \{(x_1, \dots, x_n) : 0 \leq x_1 \leq \dots \leq x_n \leq 1\}.$$

It follows that, conversely, if we start with a random standard Young tableau  $P_n$  distributed according to the Plancherel measure of order  $n$  and a random vector  $(W_1, \dots, W_n)$  that is independent of  $P_n$  and is distributed uniformly on  $\Delta_n$ , then the tableau  $T'_n$  obtained by replacing each entry  $p$  of  $P_n$  by  $W_p$  is equal in distribution to  $T_n$ .

We will now apply these observations to prove a simple lemma about sublevel tableaux. As mentioned above, for any  $0 < t \leq 1$ , the  $t$ -**sublevel tableau** of  $T_{n-1}$ , which we denote by  $T_{n-1}^{(t)}$ , is the subtableau of  $T_{n-1}$  consisting of those boxes with entries  $\leq t$ . An essential fact that makes our main result possible is a self-similarity property that says that  $T_{n-1}^{(t)}$  is distributed roughly as a scaled version of  $T_k$  for  $k \approx tn$ . Since the number of entries in  $T_{n-1}^{(t)}$  is itself random, we need to condition on its value to make this statement precise. The details are as follows.

**Lemma 3.** *Let  $0 < t \leq 1$  and  $0 \leq k \leq n - 1$ .*

- *The number  $|T_{n-1}^{(t)}|$  of boxes in  $T_{n-1}^{(t)}$  satisfies*

$$(14) \quad \frac{|T_{n-1}^{(t)}|}{n-1} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} t.$$

- *Conditioned on the event  $|T_{n-1}^{(t)}| = k$ , the rescaled sublevel tableau  $\frac{1}{t}T_{n-1}^{(t)}$  (where the rescaling means that each entry of  $T_{n-1}^{(t)}$  is divided by  $t$ ) is equal in distribution to  $T_k$ .*

*Proof.* Recall that the entries of  $T_{n-1}$  are the i.i.d. uniform random numbers  $X_1, \dots, X_{n-1}$ , so the entries of the sublevel tableau  $T_{n-1}^{(t)}$  consist of the subset of the numbers  $X_1, \dots, X_{n-1}$  which are  $\leq t$ . It follows

that the random variable  $Z_n = |T_{n-1}^{(t)}|$  is given by

$$Z_n = \sum_{j=1}^{n-1} 1_{\{X_j \leq t\}},$$

so the first claim follows from the law of large numbers.

Denote by  $(Y_1, \dots, Y_{Z_n})$  the sequence (of random length  $Z_n$ ) of those  $X_j$ 's (for  $1 \leq j \leq n-1$ ) for which  $X_j \leq t$ , in the order in which they appear. From elementary probability theory, it is easy to see that, conditioned on the event  $\{Z_n = k\}$ , the random variables  $Y_1, \dots, Y_k$  are independent and uniformly distributed in  $[0, t]$ . But now observe that (still conditioning on  $\{Z_n = k\}$ ) the sublevel tableau  $T_{n-1}^{(t)}$  can be constructed as follows:

**Standardize:**

Replace  $T_{n-1}$  by a standard Young tableau  $P_{n-1}$ , where each entry of  $P_{n-1}$  is the ranking of the corresponding entry of  $T_{n-1}$  in the list  $X_1, \dots, X_{n-1}$ .

**Sublevel:**

Take the  $k$ -sublevel tableau  $P_{n-1}^{(k)}$  of  $P_{n-1}$ .

**Destandardize:**

Replace each entry  $p$  of  $P_{n-1}^{(k)}$  by the  $p$ th order statistic  $Y_p^{(k)}$  of the sequence  $Y_1, \dots, Y_k$ .

By the remarks made at the beginning of this section,  $P_{n-1}$  is a Plancherel-random standard Young tableau of order  $n-1$ . We now use the elementary fact that the Plancherel measures are a consistent family of probability measures, in the sense that for any  $k \leq m$ , the  $k$ -sublevel tableau of a Plancherel-random tableau  $Q_m$  of order  $m$  is a Plancherel-random tableau of order  $k$ . (The case  $k = m-1$  of this claim corresponds to the simple operation of removing the maximal entry of a Plancherel-random tableau; this clearly implies the general case by induction, and the fact that the claim is true in this case is a version of a well-known property of the Plancherel measures, mentioned for example in [Rom14, Lemma 1.25].) So, the tableau  $P_{n-1}^{(k)}$  is distributed according to the Plancherel measure of order  $k$ . Finally, since (still conditioning on the event  $\{Z_n = k\}$  as before) the vector of order statistics  $(Y_1^{(k)}, \dots, Y_k^{(k)}) = (X_1^{(n)}, \dots, X_k^{(n)})$  is independent of  $P_{n-1}$  (and hence also of  $P_{n-1}^{(k)}$ ) and is distributed like  $t$  times a random vector distributed uniformly in  $\Delta_k$ , again by the remarks made above we have that  $T_{n-1}^{(t)}$  is (conditionally on  $\{Z_n = k\}$ ) equal in distribution to  $t \cdot T_k$ .  $\square$

## 4. FINISHING THE PROOF

To prove (10), we first reparametrize the bumping route  $\mathbf{B}_{T_{n-1},\alpha}$  according to the parameter  $t$  associated with the sublevel tableaux. For each  $0 \leq \alpha < 1$ , this reparametrized bumping route will now be a random function  $\Phi_{n,\alpha} : [\alpha, 1] \rightarrow \mathbb{N} \times \mathbb{N}$  defined by

$$(15) \quad \Phi_{n,\alpha}(t) = (b_{T_{n-1},\alpha}(m), m)$$

where for each  $t$  we denote by  $m$  the minimal number for which  $(b_{T_{n-1},\alpha}(m), m)$  lies outside the sublevel tableau  $T_{n-1}^{(t)}$ . Note that almost surely we have that

$$\begin{aligned} \Phi_{n,\alpha}(\alpha) &= (b_{T_{n-1},\alpha}(1), 1), \\ \Phi_{n,\alpha}(1) &= (b_{T_{n-1},\alpha}(k_{T_{n-1},\alpha}), k_{T_{n-1},\alpha}), \end{aligned}$$

and the range of  $\Phi_{n,\alpha}$  consists of the entire bumping route  $\mathbf{B}_{T_{n-1},\alpha}$ .

**Theorem 4.** *For any  $0 \leq \alpha < 1$  and  $\epsilon > 0$  we have*

$$(16) \quad \mathbb{P} \left( \max_{\alpha \leq t \leq 1} \left\| \frac{\Phi_{n,\alpha}(t)}{\sqrt{n}} - (x_\alpha(t), y_\alpha(t)) \right\| > \epsilon \right) \xrightarrow{n \rightarrow \infty} 0.$$

*Proof.* Fix  $\alpha \in [0, 1)$ . First, we prove the weaker statement that for any  $\epsilon > 0$  and  $t \in [\alpha, 1]$  we have

$$(17) \quad \mathbb{P} \left( \left\| \frac{\Phi_{n,\alpha}(t)}{\sqrt{n}} - (x_\alpha(t), y_\alpha(t)) \right\| > \epsilon \right) \xrightarrow{n \rightarrow \infty} 0.$$

Denote  $Z_n = |T_{n-1}^{(t)}|$  as before, and let  $\delta > 0$  be some small number (depending on  $\epsilon$ ) whose value will be fixed shortly. We have

$$\begin{aligned} (18) \quad & \mathbb{P} \left( \left\| \frac{\Phi_{n,\alpha}(t)}{\sqrt{n}} - (x_\alpha(t), y_\alpha(t)) \right\| > \epsilon \right) \\ & \leq \mathbb{P} \left( \left| \frac{Z_n}{n-1} - t \right| > \delta \right) \\ & \quad + \sum_{\left| \frac{k}{n-1} - t \right| \leq \delta} \mathbb{P}(Z_n = k) \mathbb{P} \left( \left\| \frac{\Phi_{n,\alpha}(t)}{\sqrt{n}} - (x_\alpha(t), y_\alpha(t)) \right\| > \epsilon \mid Z_n = k \right) \\ & \leq \mathbb{P} \left( \left| \frac{Z_n}{n-1} - t \right| > \delta \right) \\ & \quad + \max_{\left| \frac{k}{n-1} - t \right| \leq \delta} \mathbb{P} \left( \left\| \frac{\Phi_{n,\alpha}(t)}{\sqrt{n}} - (x_\alpha(t), y_\alpha(t)) \right\| > \epsilon \mid Z_n = k \right). \end{aligned}$$

In the last expression, the first term tends to 0 as  $n \rightarrow \infty$ , by (14). Let  $k = k(n)$  be the value for which the maximum of the second term

is attained. Note that the second claim of Lemma 3 implies that the conditional probability in the second term can be replaced by its unconditional counterpart

$$(19) \quad \mathbb{P} \left( \left\| \frac{\Phi_{k+1, \alpha/t}(1)}{\sqrt{n}} - (x_\alpha(t), y_\alpha(t)) \right\| > \epsilon \right) \\ = \mathbb{P} \left( \left\| \sqrt{\frac{k}{n}} \frac{\Phi_{k+1, \alpha/t}(1)}{\sqrt{k}} - \sqrt{t} (x_{\alpha/t}(1), y_{\alpha/t}(1)) \right\| > \epsilon \right).$$

If we had the precise equality  $k = tn$ , it would immediately follow from (13) that this probability tends to 0 as  $n$  (and therefore also  $k$ ) tends to  $\infty$ . As it is, such an equality does not hold; however, we restricted  $k$  to a range such that

$$t - \delta \leq \liminf_{n \rightarrow \infty} \frac{k}{n} \leq \limsup_{n \rightarrow \infty} \frac{k}{n} \leq t + \delta.$$

This is good enough, since it is easy to check that if  $\delta$  is taken (as a function of  $\epsilon$ ) to be a small enough positive number, then the right-hand side of (19) can be bounded from above by

$$(20) \quad \mathbb{P} \left( \left\| \frac{\Phi_{k+1, \alpha/t}(1)}{\sqrt{k}} - (x_{\alpha/t}(1), y_{\alpha/t}(1)) \right\| > \frac{\epsilon}{2} \right) \\ + \mathbb{P} \left( \left\| \frac{\Phi_{k+1, \alpha/t}(1)}{\sqrt{k}} \right\| > 3\sqrt{2} \right).$$

The first probability tends to 0 as  $n \rightarrow \infty$  by (13). The second probability is bounded by the probability that a Plancherel-random Young diagram of order  $k$  has a row or column of length  $\geq 3\sqrt{k}$ ; it is well-known that this probability decreases to 0 at a rate that is exponential in  $\sqrt{k}$  (see [Rom14, Lemma 1.5]). Thus, combining these observations with (18), (19) and the bound (20) proves (17).

Finally, to finish the proof we need to show that (17) implies (16). This is a standard argument: first, (17) clearly implies a version of (16) in which the maximum is taken over finitely many values  $\alpha \leq t_1 < \dots < t_p \leq 1$  of  $t$ . Second, since both the functions  $t \mapsto \Phi_{n, \alpha}(t)$  and  $t \mapsto (x_\alpha(t), y_\alpha(t))$  have the property that their  $x$ -coordinate is weakly decreasing and their  $y$ -coordinate is weakly increasing, and since  $t \mapsto (x_\alpha(t), y_\alpha(t))$  is continuous, knowing that the bound

$$\left\| \frac{\Phi_{n, \alpha}(t)}{\sqrt{n}} - (x_\alpha(t), y_\alpha(t)) \right\| > \epsilon$$

holds for all values of  $t$  in a finite set that is sufficiently dense in  $[\alpha, 1]$  ensures that the same inequality (with  $\epsilon$  replaced by, say,  $2\epsilon$ ) will hold for all  $t \in [\alpha, 1]$ . The details are easy and are left to the reader.  $\square$

*Proof of (10).* It is now easy to derive (10) from (16). The idea is that the relation between  $t$  and  $m$  in (15) can be inverted, expressing the  $m$ th point  $(b_{T_{n-1},\alpha}(m), m)$  of the bumping route as  $\Phi_{n,\alpha}(t(m))$  where  $t(m)$  is the minimal value  $t \geq \alpha$  for which the  $y$ -coordinate of  $\Phi_{n,\alpha}(t)$  is equal to  $m$ . (Note that  $t(\cdot)$  also depends on  $n$  and  $\alpha$ , but for convenience we leave this dependence implicit in our notation.) Expressing (16) in terms of  $t(m)$  gives the convergence in probability

$$\max_{1 \leq m \leq k_{T_{n-1},\alpha}} \left\| \frac{(b_{T_{n-1},\alpha}(m), m)}{\sqrt{n}} - (x_\alpha(t(m)), y_\alpha(t(m))) \right\| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0,$$

which can be broken down into two separate convergence relations,

$$(21) \quad \max_{1 \leq m \leq k_{T_{n-1},\alpha}} \left| y_\alpha(t(m)) - \frac{m}{\sqrt{n}} \right| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0,$$

$$(22) \quad \max_{1 \leq m \leq k_{T_{n-1},\alpha}} \left| \frac{b_{T_{n-1},\alpha}(m)}{\sqrt{n}} - x_\alpha(t(m)) \right| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

Now observe that  $z \mapsto y_\alpha^{-1}(z \wedge \kappa(\alpha))$  is a continuous function on  $[0, 2\kappa(\alpha)]$ . Combining this with (21) and the fact that (by (9))  $k_{T_{n-1},\alpha}/\sqrt{n} \leq 2\kappa(\alpha)$  with asymptotically high probability, we see that

$$\max_{1 \leq m \leq k_{T_{n-1},\alpha}} \left| t(m) - y_\alpha^{-1} \left( \frac{m}{\sqrt{n}} \wedge \kappa(\alpha) \right) \right| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

Finally, this relation, together with (22) and the fact that  $x_\alpha(\cdot)$  is continuous, implies that

$$\max_{1 \leq m \leq k_{T_{n-1},\alpha}} \left| \frac{b_{T_{n-1},\alpha}(m)}{\sqrt{n}} - x_\alpha \left( y_\alpha^{-1} \left( \frac{m}{\sqrt{n}} \wedge \kappa(\alpha) \right) \right) \right| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0,$$

which is exactly (10).  $\square$

#### ACKNOWLEDGMENTS

P.Š.'s research has been supported by a grant number SN 101/1-1 from *Deutsche Forschungsgemeinschaft*. Dan Romik's work was supported by the National Science Foundation under grant DMS-0955584 and by grant #228524 from the Simons Foundation.

## REFERENCES

- [BDJ99a] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12:1119–1178, 1999.
- [BDJ99b] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the second row of a Young diagram under Plancherel measure. *Geom. Funct. Anal.*, 10:702–731, 1999.
- [BOO00] A. Borodin, A. Okounkov, and G. Olshanski. Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.*, 13:491–515, 2000.
- [Ful97] W. Fulton. *Young Tableaux: With Applications to Representation theory and Geometry*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997.
- [Joh00] K. Johansson. Shape fluctuations and random matrices. *Commun. Math. Phys.*, 209:437–476, 2000.
- [Knu98] D. E. Knuth. *The Art of Computer Science, Vol. 3: Sorting and Searching, 2nd Ed.* Addison-Wesley, 1998.
- [LS77] B. F. Logan and L. A. Shepp. A variational problem for random Young tableaux. *Adv. Math.*, 26(2):206–222, 1977.
- [Moo06] C. Moore. Flows in Young diagrams, 2006. Online resource, <http://tuvalu.santafe.edu/~moore/gallery.html>.
- [Rom14] D. Romik. *The Surprising Mathematics of Longest Increasing Subsequences*. Cambridge University Press, 2014. To appear; available at <http://www.math.ucdavis.edu/~romik/book>.
- [RŚ14] D. Romik and P. Śniady. Jeu de taquin dynamics on infinite Young tableaux and second class particles. *Ann. Probab.*, 2014. To appear.
- [Sta99] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.
- [VK77] A. M. Vershik and S. V. Kerov. Asymptotics of the Plancherel measure of the symmetric group and the limit form of Young tableaux. *Soviet Math. Dokl.*, 18:527–531, 1977.
- [VK85] A. M. Vershik and S. V. Kerov. Asymptotic of the largest and the typical dimensions of irreducible representations of a symmetric group. *Functional Anal. Appl.*, 19(1):21–31, 1985.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, DAVIS, ONE SHIELDS AVENUE, DAVIS, CA 95616, USA

*E-mail address:* romik@math.ucdavis.edu

ZENTRUM MATHEMATIK, M5, TECHNISCHE UNIVERSITÄT MÜNCHEN, BOLTZMANNSTRASSE 3, 85748 GARCHING, GERMANY

INSTYTUT MATEMATYCZNY, POLSKA AKADEMIA NAUK, UL. ŚNIADECKICH 8, 00-956 WARSZAWA, POLAND

INSTYTUT MATEMATYCZNY, UNIWERSYTET WROCŁAWSKI, PL. GRUNWALDZKI 2/4, 50-384 WROCŁAW, POLAND

*E-mail address:* piotr.sniady@tum.de, piotr.sniady@math.uni.wroc.pl