Modeling of bumping routes in the RSK algorithm

Vasilii Duzhin[1], Artem Kuzmin[1], Nikolay Vassiliev[1,2]

[1]Saint Petersburg Electrotechnical University "LETI"

[2]Saint Petersburg Department of Steklov Mathematical Institute of Russian Academy of Sciences

Polynomial Computer Algebra
April 21, 2023

Robinson-Schensted-Knuth (RSK) correspondence

The Robinson-Schensted-Knuth correspondence is a bijection between a set of sequences of elements of a linearly ordered set and a set of pairs of Young tableaux of the same shape: insertion tableau P and recording tableau Q.

Robinson-Schensted-Knuth (RSK) correspondence

The Robinson-Schensted-Knuth correspondence is a bijection between a set of sequences of elements of a linearly ordered set and a set of pairs of Young tableaux of the same shape: insertion tableau P and recording tableau Q.

We consider the sequences consisting of random real numbers uniformly distributed in the interval [0, 1].

An example

0.570, 0.563, 0.771, 0.729, 0.025, 0.064, 0.556, 0.698, 0.528, 0.346, 0.010, 0.241, 0.569, 0.455

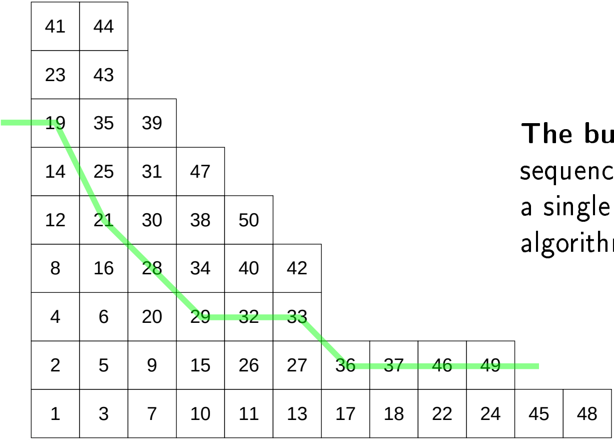
P :

0.455					
0.241	0.569	0.698			
0.064	0.346	0.528	0.729	0.771	
0.010	0.025	0.171	0.556	0.563	0.570

Q :

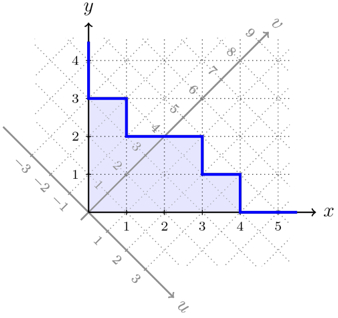
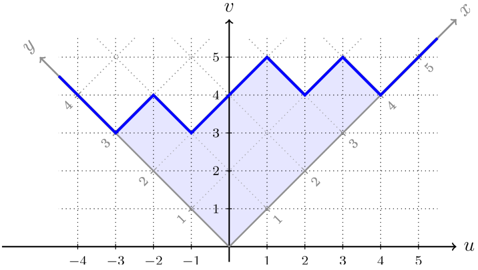
9					
4	8	14			
2	5	7	13	15	
1	3	6	10	11	12

Bumping routes



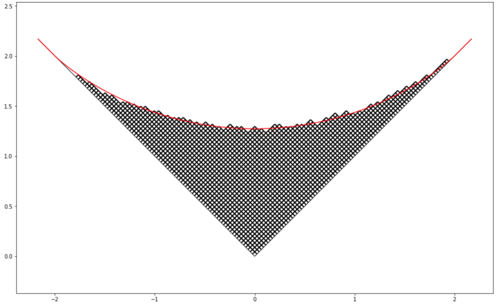
The bumping route is a sequence of boxes moved during a single iteration of the RSK algorithm in P tableau.

The coordinate systems

The French notation	The Russian notation (Vershik-Kerov coordinates)
	

The Plancherel measure

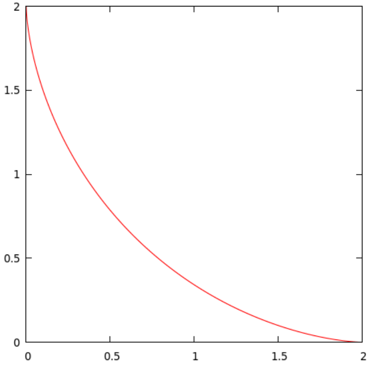
RSK defines a correspondence between a **uniform** measure on sequences and a **Plancherel** measure on pairs of Young tableaux of the same shape. In the limit, the shape of such tableaux is described by the Vershik-Kerov-Logan-Shepp curve [Vershik Kerov 85, Logan Shepp 77]:



$$\frac{2}{\pi}\left(u \arcsin u + \sqrt{1-u^2}\right)$$

The Plancherel measure

RSK defines a correspondence between a **uniform** measure on sequences and a **Plancherel** measure on pairs of Young tableaux of the same shape. In the limit, the shape of such tableaux is described by the Vershik-Kerov-Logan-Shepp curve [Vershik Kerov 85, Logan Shepp 77]:



$$\frac{2}{\pi}\left(u \arcsin u + \sqrt{1-u^2}\right)$$

Limit shape of bumping routes

In [Romik, Śniady '16], formulae for computing the limit shapes of bumping routes are presented:

$$(1) \qquad \Omega(u) = \frac{2}{\pi}\left(u \arcsin\frac{u}{2} + \sqrt{4-u^2}\right) \qquad (|u| \le 2),$$

$$(2) \qquad F(u) = \frac{1}{2} + \frac{1}{\pi}\left(\frac{u\sqrt{4-u^2}}{4} + \arcsin\frac{u}{2}\right) \qquad (|u| \le 2),$$

$$(3) \qquad u_\alpha(t) = \sqrt{t} \cdot F^{-1}\left(\frac{\alpha}{t}\right) \qquad (0 \le \alpha \le t \le 1),$$

$$(4) \qquad v_\alpha(t) = \sqrt{t} \cdot \Omega\left(F^{-1}\left(\frac{\alpha}{t}\right)\right) \qquad (0 \le \alpha \le t \le 1),$$

$$(5) \qquad y_\alpha(t) = \frac{v_\alpha(t) + u_\alpha(t)}{2} \qquad (0 \le \alpha \le t \le 1),$$

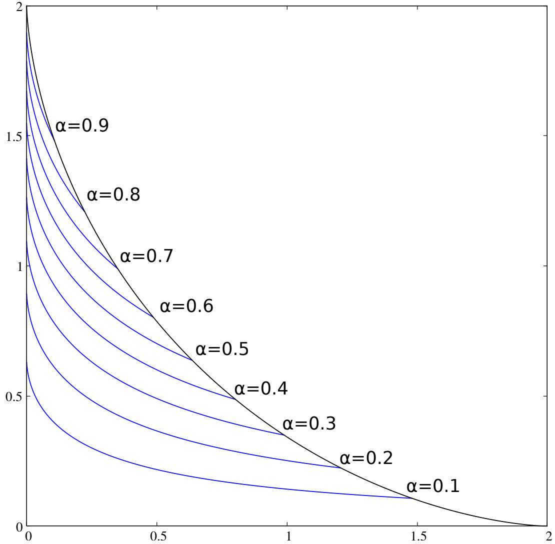
$$(6) \qquad x_\alpha(t) = \frac{v_\alpha(t) - u_\alpha(t)}{2} \qquad (0 \le \alpha \le t \le 1),$$

$$(7) \qquad k(\alpha) = x_\alpha(1) = \frac{\Omega\big(F^{-1}(\alpha)\big) - F^{-1}(\alpha)}{2} \qquad (0 \le \alpha \le 1).$$

The problem

How the bumping routes converge to their limit curves with increasing size of Young tableaux (i.e. with increasing length of input sequence)?

The limit shapes of bumping routes for different α



Computer experiments

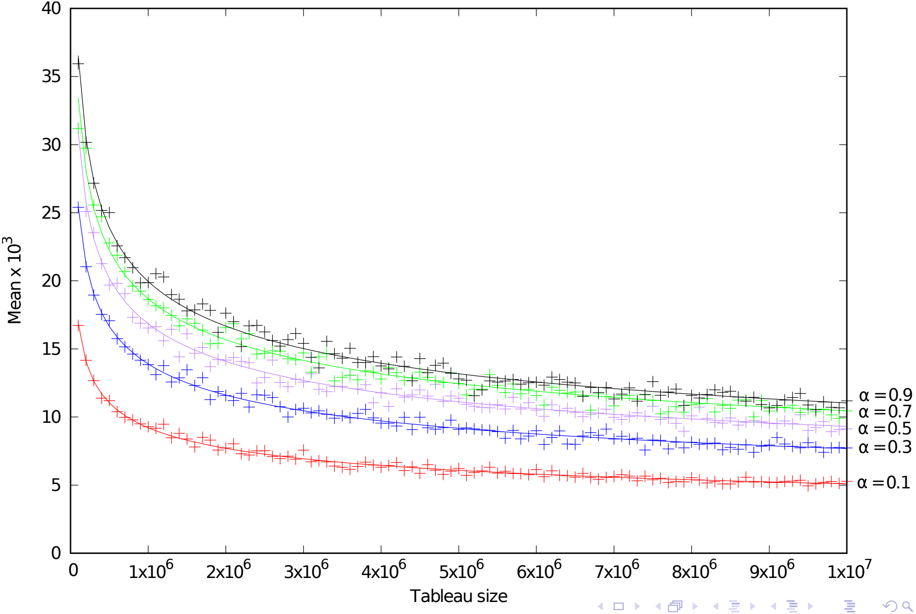
- We have constructed Young tableaux with **sizes** n $= 10^5..10^7$ boxes with a step of 10^5.
- For each tableau size $n < 10^6$, **1000** tableaux and the corresponding routes were built. For larger sizes, **300** tableaux were built for each size.
- The following values of the element α were fed to the **input** of the RSK algorithm: $\alpha = 0.1, 0.3, 0.5, 0.7, 0.9$.
- For a fixed input value α and tableau size n, we calculate the **root-mean-square distance** S between the coordinates (x, y) of the boxes of the constructed bumping route and the corresponding coordinates (x*, y*) of the limit curve.
- For each of the k constructed bumping routes, the mean and variance of S were calculated.

Approximation of the function of distance between a bumping route and its limit shape

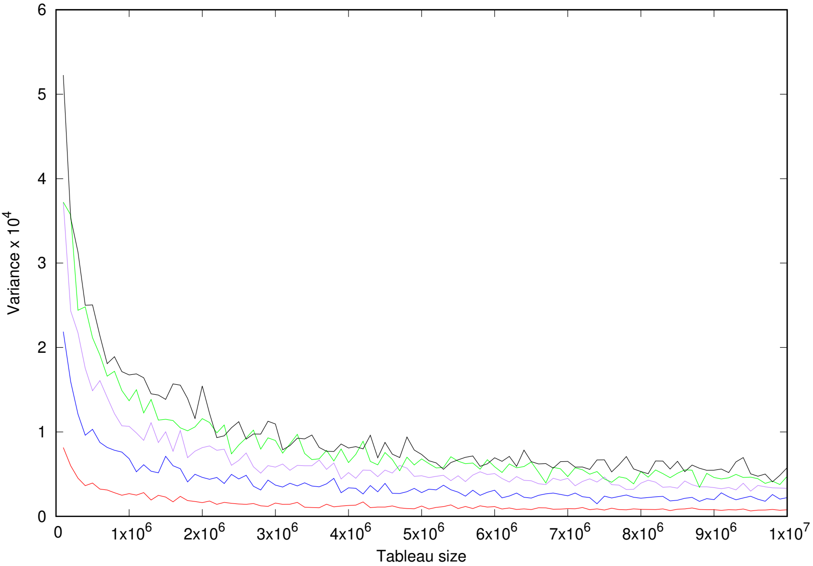
The computer experiments show that the distances between the bumping routes and the limit shapes are well approximated by the formula

$$f(n) = a \cdot n^{-\frac{1}{4}} + b \cdot n^{-\frac{1}{2}}$$

The distances of the bumping routes from their limit shapes



The distances of the bumping routes from their limit shapes



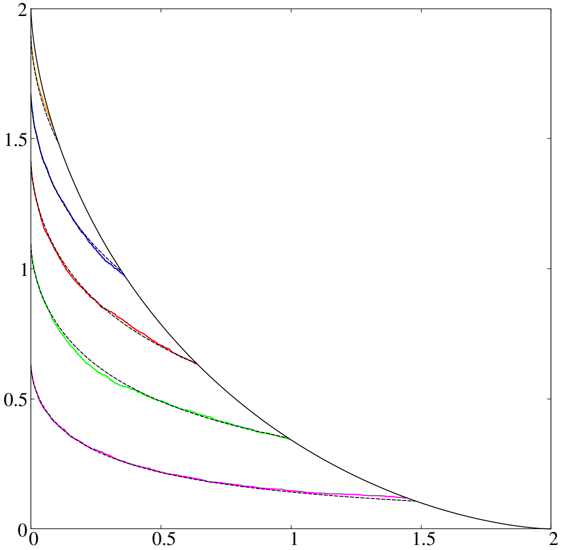
Maximum deviations of the approximation curves from the experimental bumping routes

α	a	b	max.dev.	α	a	b	max.dev.
0.1	0.244	0.191	0.000217	0.5	0.296	0.475	0.000230
0.15	0.270	0.184	0.000239	0.55	0.289	0.597	0.000211
0.2	0.285	0.213	0.000349	0.6	0.281	0.622	0.000233
0.25	0.296	0.236	0.000301	0.65	0.266	0.826	0.000294
0.3	0.302	0.264	0.000209	0.7	0.253	0.914	0.000216
0.35	0.303	0.325	0.000249	0.75	0.233	1.136	0.000270
0.4	0.305	0.319	0.000202	0.8	0.210	1.457	0.000228
0.45	0.304	0.331	0.000298	0.85	0.179	1.925	0.000167

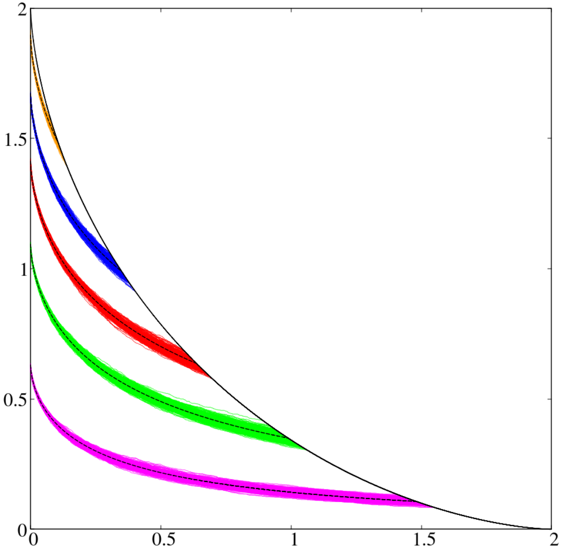
The coordinates of the ends of the bumping routes

- In our computer experiments, we have studied the coordinates distributions of the ends of the bumping routes.
- This experiment uses only end points of bumping routes instead of their entire trajectories as the previous one.
- Also we estimated the values of the parameters of the function approximating these distributions.

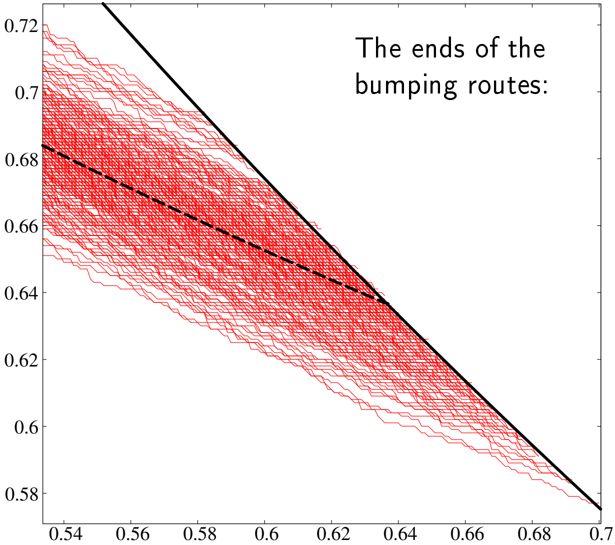
A limit curve and a bumping route for each α



A limit curve and 300 bumping routes for each α

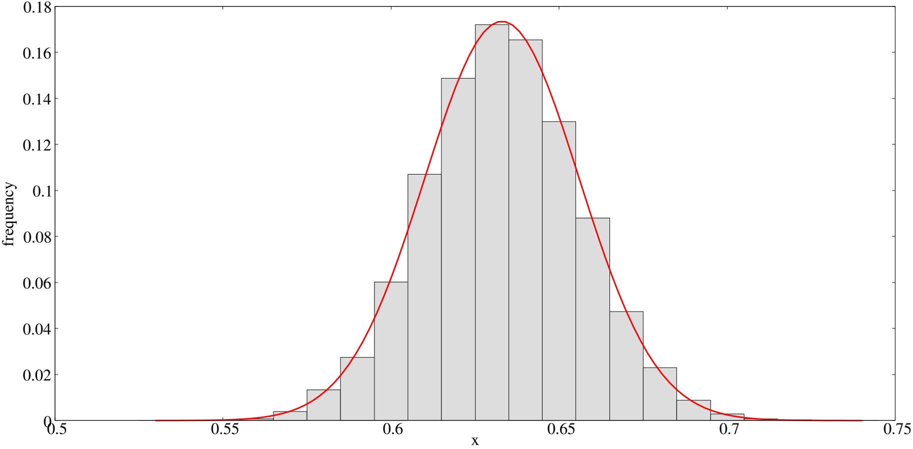


A limit curve and 300 bumping routes for α=0.5

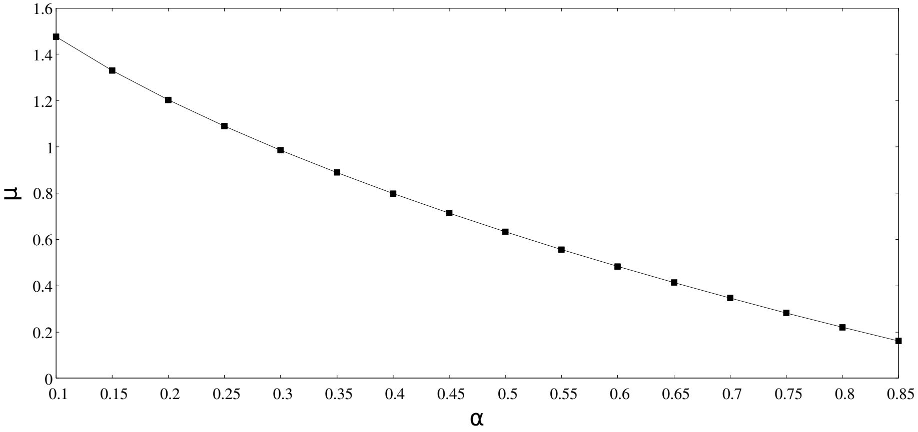


The distribution of the ends of the bumping routes

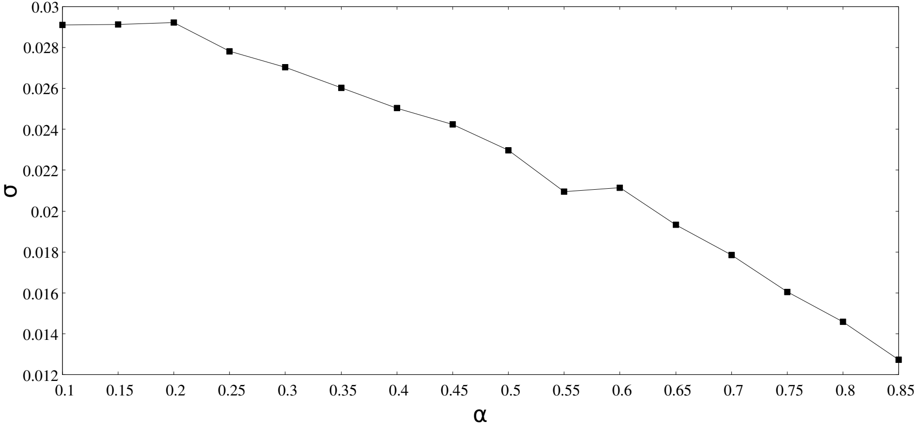
The histogram is well approximated by a Gaussian distribution:



Dependence of the mean of the Gaussian for different α



Dependence of the st. dev. of the Gaussian for different α



Conclusion

- Our computer experiments show **good agreement** with the theoretical results obtained in [Romik, Śniady '16];
- The **distance** between the bumping routes and the limit shapes is well-approximated by the **formula**

$$f(n) = a \cdot n^{-\frac{1}{4}} + b \cdot n^{-\frac{1}{2}}$$

- **Convergence** of the bumping routes to their limit shapes turns out to be rather **slow** with the principal term proportional to $n^{-\frac{1}{4}}$;
- The distributions of the **ends of bumping routes** are well-approximated by a **Gaussian** distribution;
- In the future we plan to extend our research to the **strict** Young tableaux case.

Thanks for your attention!