



Asymptotics of Some Plancherel Averages Via Polynomiality Results

Werner Schachinger¹ 

Received: 31 December 2022 / Revised: 28 June 2023 / Accepted: 14 August 2023 /
Published online: 19 September 2023
© The Author(s) 2023

Abstract

Consider Young diagrams of n boxes distributed according to the Plancherel measure. So those diagrams could be the output of the RSK algorithm, when applied to random permutations of the set $\{1, \dots, n\}$. Here we are interested in asymptotics, as $n \rightarrow \infty$, of expectations of certain functions of random Young diagrams, such as the number of bumping steps of the RSK algorithm that leads to that diagram, the side length of its Durfee square, or the logarithm of its probability. We can express these functions in terms of hook lengths or contents of the boxes of the diagram, which opens the door for application of known polynomiality results for Plancherel averages. We thus obtain representations of expectations as binomial convolutions, that can be further analyzed with the help of Rice's integral or Poisson generating functions. Among our results is a very explicit expression for the constant appearing in the almost equipartition property of the Plancherel measure.

Keywords Robinson–Schensted algorithm · Young diagram · Plancherel measure · Durfee square · Asymptotic expansion · Vershik–Kerov conjecture

1 Introduction

We identify Young diagrams (sets consisting of left aligned decreasingly ordered rows of square boxes) with partitions $\lambda = (\lambda_1, \dots, \lambda_k)$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$, and denote $|\lambda| = \lambda_1 + \dots + \lambda_k$. The notation $\lambda \vdash n$ then signifies that λ is a partition of n , i.e., $|\lambda| = n$. We let $Y(\pi) = Y_\lambda := \sum_{\ell=1}^k \lambda_\ell(\ell - 1)$ denote the number of bumping steps of the Robinson–Schensted algorithm (see Figs. 1 and 2) when applied to a permutation π that is mapped to a pair of standard Young tableaux of shape λ . A standard Young tableau is a Young diagram λ filled with numbers $1, \dots, |\lambda|$ in a

✉ Werner Schachinger
Werner.Schachinger@univie.ac.at

¹ Department of Statistics and Operations Research (ISOR), University of Vienna, Vienna, Austria

way such that numbers in each row and each column are increasing. See e.g. [23, Sect. 1.6] or [24, Sect. 3.1] for nice expositions of the algorithm and references to the original articles by Gilbert de Beauregard Robinson, by Craige Eugene Schensted, and by Donald Ervin Knuth, who significantly widened the scope of the algorithm, the abbreviation with reference to all three authors, *RSK algorithm*, now frequently being used also to refer to the original Robinson–Schensted algorithm.

We denote by Y_n the restriction of $Y(\pi)$ to permutations of the set $\{1, 2, \dots, n\}$ chosen uniformly at random. The Young diagrams λ obtained by the RSK algorithm are then distributed according to the n th Plancherel measure, i.e., $\mathbb{P}^{(n)}(\lambda) = \frac{f_\lambda^2}{n!} = \frac{n!}{p_\lambda^2}$, where f_λ is the number of standard Young tableaux of shape λ , satisfying $f_\lambda = \frac{n!}{p_\lambda}$, and where $p_\lambda := \prod_{u \in \lambda} h_u$ denotes the product of the hook lengths of the diagram λ , see [7]. Here the hook length h_u of a particular box u of λ is one more than the number of boxes to the right of u plus the number of boxes below u . Note that Y_λ has also the meaning of $|\lambda|$ times the y -coordinate of the barycenter of the set

$$S_\lambda := \{(i, j) \in \mathbb{Z}^2 : 0 \leq j \leq k - 1, 0 \leq i \leq \lambda_{j+1} - 1\},$$

which is just the set of lower left corners of the boxes of λ in French notation, which addresses boxes by Cartesian coordinates of the first quadrant. Apart from here we always stick to English notation with its matrix style indexing of boxes. Note that X_λ , the x -coordinate of the barycenter of S_λ , is given by $Y_{\lambda'}$, where λ' is the partition

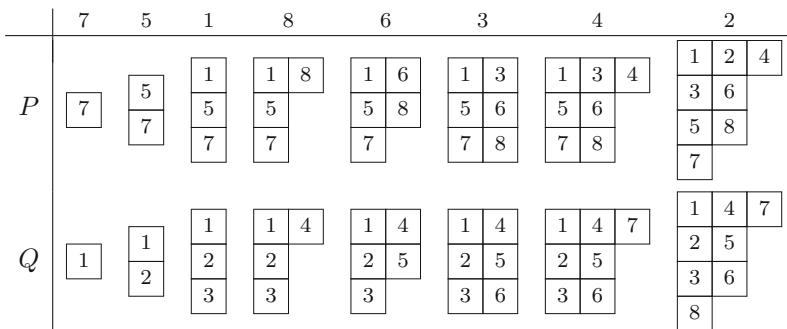


Fig. 1 Tableaux P and Q , as they evolve when subjecting the RSK algorithm to the permutation $\pi = (75186342)$. P is constructed by row insertions of elements of π one by one, while Q is recording the position of boxes as they are added

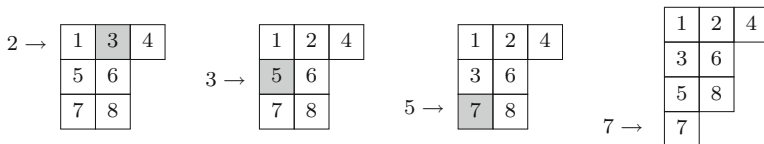


Fig. 2 Detailed insertion of 2 into the second to last tableau P . Inserting 2, 3, and 5 into their (shaded) destination boxes causes bumps of elements down one row, being now in need of insertion themselves. In the last step, 7 is the largest element of row four, so this insertion happens without a pump

conjugate to λ , its parts being defined by $\lambda'_j := |\{i : \lambda_i \geq j\}|$. Stated differently, λ and λ' are mirror images of one another with respect to the main diagonal (upper left to lower right). The sets of hook lengths are therefore the same for λ and λ' , which yields invariance of Plancherel measure under conjugation. Thus X_n and Y_n are identical in distribution. This allows for a representation of $\mathbb{E} Y_n$ and $\text{Var} Y_n$ in terms of $X_n + Y_n$ and $X_n - Y_n$,

$$\begin{aligned} \mathbb{E} Y_n &= \frac{1}{2} \mathbb{E} (X_n + Y_n), \\ \text{Var} Y_n &= \frac{1}{2} (\text{Var} X_n + \text{Var} Y_n) = \frac{1}{4} (\text{Var} (X_n + Y_n) + \text{Var} (X_n - Y_n)). \end{aligned}$$

Note that we can express $X_\lambda - Y_\lambda$, resp. $X_\lambda + Y_\lambda$, in terms of contents $\{c_u : u \in \lambda\}$, resp. hook lengths $\{h_u : u \in \lambda\}$, of the diagram λ :

$$X_\lambda - Y_\lambda = \sum_{u \in \lambda} c_u, \quad X_\lambda + Y_\lambda = \sum_{u \in \lambda} h_u - |\lambda|. \tag{1.1}$$

Here the content c_u of a box $u = (i, j)$ of λ is $j - i$, i.e., the column number of u minus the row number of u , see Fig. 3 for an illustration of hook lengths, contents, and bumping step counts. For a proof of (1.1) note

$$\begin{aligned} Y_\lambda &= \sum_{i=1}^k \lambda_i (i - 1) = \sum_{(i,j) \in \lambda} (i - 1) = \sum_{j=1}^{\lambda_1} \sum_{i=1}^{\lambda'_j} (i - 1) = \sum_{j=1}^{\lambda_1} \sum_{i=1}^{\lambda'_j} (\lambda'_j - i) \\ &= \sum_{(i,j) \in \lambda} (\lambda'_j - i), \end{aligned}$$

and similarly

$$X_\lambda = Y_{\lambda'} = \sum_{(i,j) \in \lambda} (j - 1) = \sum_{(i,j) \in \lambda} (\lambda_i - j),$$

leading to $X_\lambda - Y_\lambda = Y_{\lambda'} - Y_\lambda = \sum_{(i,j) \in \lambda} [(j - 1) - (i - 1)] = \sum_{(i,j) \in \lambda} (j - i) = \sum_{u \in \lambda} c_u$ and $X_\lambda + Y_\lambda + |\lambda| = \sum_{(i,j) \in \lambda} [(\lambda_i - j) + (\lambda'_j - i) + 1] = \sum_{u \in \lambda} h_u$, where $(\lambda_i - j) + (\lambda'_j - i) + 1$ is clearly the hook length of box (i, j) .

Further functions of λ that can be written in terms of hook lengths or contents are $\log p_\lambda = \sum_{u \in \lambda} \log(h_u)$ making its appearance in Sect. 3, and $D(\lambda) = \sum_{u \in \lambda} \delta_{0, c_u}$, the number of boxes of λ on the main diagonal, that we will meet in Sect. 4.

Being able to express some function of λ in terms of the contents or hook lengths of the boxes of λ can allow us to employ the polynomiality results for Plancherel averages derived by Stanley [26].

Theorem 1.1 ([26, Theorems 2.1, 4.3]) *Let $F(x)$ be a formal power series over \mathbb{Q} of bounded degree that is symmetric in the variables $x = (x_1, x_2, \dots)$. Then both*

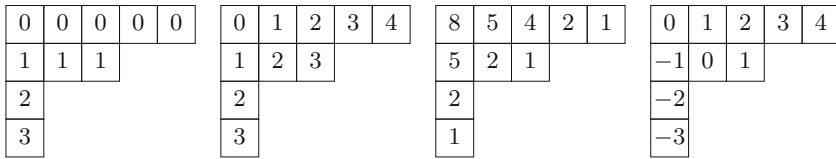


Fig. 3 The partition $(5, 3, 1, 1) \vdash 10$, drawn as Young diagram, filled from left to right with bumping step counts, sums of box coordinates, hook lengths, and contents. The sum of the entries of the second diagram is 10 less than the sum of the hook lengths

averages

$$\frac{1}{n!} \sum_{\lambda \vdash n} f_{\lambda}^2 F(c_u : u \in \lambda) \quad \text{and} \quad \frac{1}{n!} \sum_{\lambda \vdash n} f_{\lambda}^2 F(h_u^2 : u \in \lambda)$$

are polynomial functions of n .

Note that an even more general result is given in [26, Theorem 4.4]. See also [20] for alternative proofs and further generalizations.

The proof of [26, Theorem 2.1] restricts w.l.o.g. to elementary symmetric functions indexed by partitions $\mu = (\mu_1, \dots, \mu_k)$, i.e., to functions $F(\cdot) = e_{\mu}(\cdot) = \prod_{i=1}^k e_{\mu_i}(\cdot)$, where $e_m(x_1, x_2, \dots) = \sum_{i_1 < \dots < i_m} x_{i_1} \dots x_{i_m}$ for $m \geq 1$. As remarked in [26] right below that proof, the resulting polynomial N_{μ} is of degree $|\mu|$ if and only if $|\mu|$ is even and $\mu_1 \leq \frac{|\mu|}{2}$, otherwise, $N_{\mu} = 0$. Here is an immediate application of these degree considerations.

Lemma 1.2 $\text{Var}(X_n - Y_n) = \binom{n}{2}$.

Proof Since $\mathbb{E}(X_n - Y_n) = 0$, we have

$$\text{Var}(X_n - Y_n) = \frac{1}{n!} \sum_{\lambda \vdash n} f_{\lambda}^2 \left(\sum_{u \in \lambda} c_u \right)^2 = \frac{n(n-1)}{2}.$$

Note that here we have $F(c_u : u \in \lambda) = (e_1(c_u : u \in \lambda))^2$, i.e., $\mu = (1, 1)$. The polynomial N_{μ} is therefore of degree 2, and it is completely determined by its values at $n \in \{0, 1, 2\}$, which are $N_{\mu}(0) = N_{\mu}(1) = 0, N_{\mu}(2) = 1$, proving the claim. The result $N_{\mu}(n) = \frac{n(n-1)}{2}$ is also stated as a special case in [26, p. 94]. \square

A workaround is needed for $\text{Var}(X_n + Y_n)$, or even $\mathbb{E}(X_n + Y_n)$, because $X_{\lambda} + Y_{\lambda}$ is not a symmetric function of $\{h_u^2 : u \in \lambda\}$, but only of $\{h_u : u \in \lambda\}$. Finding a series representation $x = \sum_{k \geq 0} a_k p_k(x^2)$ with polynomials p_k , that holds for integers $x \geq 1$, (but need not hold or even converge elsewhere) would allow, interchanging summations, to apply the polynomiality results termwise. If we are lucky—and we are—the polynomials p_k have well known Plancherel averages. Such kind of workaround is employed in this paper to deal with Plancherel averages of several interesting functions of partitions, leading firstly to a representation of the expectation as a binomial

Table 1 Some functions of partitions $\lambda \vdash 4$ and their expectations with respect to Plancherel measure

λ						\mathbb{E}
$\mathbb{P}l^{(4)}(\lambda)$	$\frac{1}{24}$	$\frac{3}{8}$	$\frac{1}{6}$	$\frac{3}{8}$	$\frac{1}{24}$	
$X_\lambda - Y_\lambda$	6	2	0	-2	-6	0
$X_\lambda + Y_\lambda$	6	4	4	4	6	$\frac{25}{6}$
$\log \mathbb{P}l^{(4)}(\lambda)$	$\log \frac{1}{24}$	$\log \frac{3}{8}$	$\log \frac{1}{6}$	$\log \frac{3}{8}$	$\log \frac{1}{24}$	$\frac{8}{3} \log \frac{1}{2} + \frac{\log 3}{2}$
$D(\lambda)$	1	1	2	1	1	$\frac{7}{6}$

convolution, that is free of references to partitions, and can be analyzed using Rice’s integral, or Poisson generating functions. In some cases holonomicity of the sequence of expectations can be inferred from the binomial convolution representation. This then allows for fast computation of many terms, that can be used to numerically confirm error terms, or conduct experiments.

The paper is organized as follows: In Sect. 2 we consider the expected number of bumping steps in the RSK algorithm. In particular, we derive asymptotics for $\mathbb{E}(X_n + Y_n)$, thus refining the result obtained by Romik [22]. In Sect. 3 we consider $\mathbb{E} \log \mathbb{P}l^{(n)}(\lambda)$, with λ distributed according to Plancherel measure. From the first asymptotic terms we obtain a very explicit representation of a constant appearing in an *almost equipartition property* (abbreviated AEP) for Plancherel measure, conjectured in [28], and proven in [4]. In Sect. 4 we derive asymptotics of the expectation of the side length $D(\lambda)$ of the *Durfee square* of λ , i.e., the largest square fitting in the upper left corner of the Young diagram of λ , when partitions λ are distributed according to Plancherel measure, see Table 1. Considering in Sect. 5 more generally lengths of south-east directed cuts through the Young diagram of λ , we enter the realm of a sequence of random curves ψ_λ known to converge uniformly in probability to the *Logan-Shepp-Vershik-Kerov limit shape curve* Ω for $|\lambda| \rightarrow \infty$. For any fixed integer a the sequence with terms $\sqrt{n} \mathbb{E} \psi_\lambda \left(\frac{a}{\sqrt{2n}} \right)$, with expectation computed with respect to $\mathbb{P}l^{(n)}$, turns out to be holonomic. Experiments then strongly hint at convergence of $\mathbb{E} \psi_\lambda \left(\frac{[u\sqrt{2n}]}{\sqrt{2n}} \right) \rightarrow \Omega(u)$, uniformly in u , and reveal that second order terms show interesting fluctuations. However we are only able to prove asymptotic results in the case of fixed a , i.e., in the vicinity of $u = 0$. In Sect. 6 we return to the number Y_n of bumping steps, giving a heuristic argument for $\text{Var } Y_n = \mathcal{O}(n^2)$, based on the limit shape curve.

2 Refined Asymptotics of the Expected Number of Bumping Steps in the RSK Algorithm

Recall that $\mathbb{E}(X_n + Y_n) = 2\mathbb{E} Y_n$ denotes twice the expected number of bumping steps of the RSK algorithm when applied to a random permutation of $\{1, \dots, n\}$. Romik

[22, Eq. (1)] derived the following asymptotic result, $\mathbb{E} Y_n \sim \frac{128}{27\pi^2} n^{\frac{3}{2}}$, and showed $Y_n/\mathbb{E} Y_n \rightarrow 1$ in probability. The sequence of interest starts

$$(\mathbb{E}(X_n + Y_n))_{n=1}^{10} = (0, 1, \frac{7}{3}, \frac{25}{6}, \frac{19}{3}, \frac{44}{5}, \frac{347}{30}, \frac{8181}{560}, \frac{541273}{30240}, \frac{1943453}{90720}).$$

The next theorem leads to a refinement of Romik’s asymptotic equivalent for $\mathbb{E} Y_n$.

Theorem 2.1 (Expected number of bumping steps in the RSK algorithm) *Let $\delta_n := \log n + 2\gamma + 12 \log 2$, with γ denoting Euler’s constant. Then*

$$\begin{aligned} \mathbb{E}(X_n + Y_n) &= \frac{256}{27\pi^2} n^{\frac{3}{2}} - n + \frac{9\delta_n - 77}{9\pi^2} n^{\frac{1}{2}} + \frac{3510\delta_n - 31589}{27648\pi^2} n^{-\frac{1}{2}} \\ &\quad + \frac{5565\delta_n - 62224}{786432\pi^2} n^{-\frac{3}{2}} + \frac{e^8}{2^{12}\pi^{3/2}} \cos\left(8\sqrt{n} + \frac{\pi}{4}\right) n^{-\frac{7}{4}} + \mathcal{O}\left(n^{-\frac{9}{4}}\right). \end{aligned}$$

Proof As $X_n + Y_n$ is not a symmetric polynomial of the multiset $\{h_u^2 : u \in \lambda\}$, but only of the multiset $\{h_u : u \in \lambda\}$, we can not expect $\mathbb{E}(X_n + Y_n)$ to be a polynomial in $|\lambda|$. Indeed, by Romik’s result, $\mathbb{E}(X_n + Y_n) = \Theta(n^{\frac{3}{2}})$ is definitely not a polynomial. However, we can invoke polynomiality results via the following identity. Using

$$p(x, r) := \prod_{i=1}^r (x^2 - i^2), \tag{2.1}$$

the equation

$$x = 1 + \sum_{r=1}^{\infty} \binom{2r}{r} \frac{(-1)^r}{(1 - 2r)(2r + 1)!} p(x, r) \tag{2.2}$$

holds for $x \in \mathbb{N} := \{1, 2, 3, \dots\}$. This will be proved in the Appendix.

Now, by [21, Theorem 1], we have

$$\frac{1}{n!} \sum_{\lambda \vdash n} f_{\lambda}^2 \sum_{u \in \lambda} p(h_u, r) = K_r \binom{n}{r + 1}, \tag{2.3}$$

with $K_r = \frac{(2r)!(2r+1)!}{(r+1)!2^r!}$, leading to

$$\begin{aligned} \mathbb{E}(X_n + Y_n) &= \frac{1}{n!} \sum_{\lambda \vdash n} f_{\lambda}^2 \sum_{u \in \lambda} (h_u - 1) \stackrel{(2.2)}{=} \frac{1}{n!} \sum_{\lambda \vdash n} f_{\lambda}^2 \sum_{u \in \lambda} \sum_{r=1}^{\infty} \binom{2r}{r} \frac{(-1)^r}{(1 - 2r)(2r + 1)!} p(h_u, r) \\ &\stackrel{(2.3)}{=} \sum_{r=1}^{\infty} \binom{2r}{r} \frac{(-1)^r}{(1 - 2r)(2r + 1)!} K_r \binom{n}{r + 1} \\ &= \sum_{r=1}^{\infty} \binom{2r}{r}^2 \frac{(-1)^{r+1}}{(2r - 1)(r + 1)(r + 1)!} \binom{n}{r + 1} = \sum_{r=2}^n \binom{2r - 2}{r - 1}^2 \frac{(-1)^r}{(2r^2 - 3r)!} \binom{n}{r} \\ &= (-1)^n \frac{n!}{2\pi i} \oint_{\mathcal{C}} \frac{f(z)}{z(z - 1)(z - 2) \cdots (z - n)} dz, \end{aligned}$$

where C is a contour that encircles integers $2, 3, \dots, n$, but neither any other integers, nor poles of f , which is given by

$$f(z) := \frac{\Gamma^2(2z - 1)}{z^2(2z - 3)\Gamma^5(z)}.$$

For the method of evaluating a large finite difference via the so-called *Rice’s integral* used above see the article [9]. By computing (leading asymptotic terms of) residues of $g_n(z) := f(z) \frac{n!\Gamma(-z)}{\Gamma(n+1-z)}$ at $\pm\frac{3}{2}, 1, \pm\frac{1}{2}$, (note that $g_n(z)$ is analytic for $z \in \{-1, 0\}$) we obtain

$$\begin{aligned} \mathbb{E}(X_n + Y_n) &= \frac{256n^{\frac{3}{2}}}{27\pi^2} \left(1 - \frac{3}{8n} - \frac{7}{128n^2} - \frac{9}{1024n^3} \right) \\ &\quad - n \\ &\quad + \delta_n \frac{n^{\frac{1}{2}}}{\pi^2} \left(1 + \frac{1}{8n} + \frac{1}{128n^2} \right) - \frac{5n^{\frac{1}{2}}}{\pi^2} \left(1 + \frac{1}{8n} - \frac{1}{1920n^2} \right) \\ &\quad + \delta_n \frac{n^{-\frac{1}{2}}}{512\pi^2} \left(1 - \frac{3}{8n} \right) + \frac{n^{-\frac{1}{2}}}{1024\pi^2} \left(1 + \frac{13}{8n} \right) \\ &\quad - \delta_n \frac{n^{-\frac{3}{2}}}{2^{18}\pi^2} + \frac{n^{-\frac{3}{2}}}{2^{14} \cdot 3\pi^2} \\ &\quad + \mathcal{O}\left(n^{-\frac{5}{2}} \log n\right) + \frac{1}{2\pi i} \oint_{C'} f(z) \frac{n!\Gamma(-z)}{\Gamma(n+1-z)} dz, \end{aligned}$$

where we recall $\delta_n = \log n + 2\gamma + 12 \log 2$, and where C' encircles the interval $[-\frac{3}{2}, n]$, but no poles of the integrand outside that interval. This integral is taken care for in the appendix, yielding the contribution involving $\cos(8\sqrt{n} + \frac{\pi}{4})$. Such a term is not completely uncommon, see e.g. the example at the end of [9, Sect. 5]. Slightly rearranging the terms completes the proof. \square

2.1 Holonomicity of the Sequence $(\mathbb{E}(X_n + Y_n) + n)_{n \in \mathbb{N}}$

Defining

$$u_n := \sum_{r=1}^n \binom{2r-2}{r-1}^2 \frac{(-1)^r}{(2r^2-3r)r!} \binom{n}{r},$$

we have $\mathbb{E}(X_n + Y_n) = u_n - n$. The sequence (u_n) satisfies the linear recurrence relation

$$u_{n+4} = \frac{4n^3+23n^2+63n+78}{(n+3)(n+4)^2} u_{n+3} - \frac{(2n-1)(3n+2)}{(n+4)^2} u_{n+2} + \frac{(4n-7)(n+2)}{(n+4)^2} u_{n+1} - \frac{(n+1)(n+2)}{(n+4)^2} u_n.$$

with initial conditions

$$u_0 = 0, \quad u_1 = 1, \quad u_2 = 3, \quad u_3 = \frac{16}{3}.$$

Clearly, the terms $\frac{u_n}{n!}$ comprise a sequence, that is the convolution of two sequences that are obviously holonomic. For said convolution the *gfun* package [25] then easily produces a recursion.

For the Poisson generating function $U(z) := e^{-z} \sum_{n \geq 0} \frac{u_n}{n!} z^n$ we obtain

$$U(z) = -z \sum_{k \geq 0} \binom{2k}{k}^2 \frac{(-z)^k}{(k+1)(2k-1)(k+1)!^2} = z {}_2F_3 \left[\begin{matrix} -\frac{1}{2}, \frac{1}{2} \\ 2, 2, 2 \end{matrix}; -16z \right],$$

a hypergeometric function that may also be used to recover asymptotics of u_n , see [15, Sect. 5.11.2] for asymptotic expansions of generalized hypergeometric functions. Indeed, the leading terms of an asymptotic expansion of $U(z)$, provided by Maple, together with Depoissonization via the saddle point method, yield an alternative proof of Theorem 2.1.

Note that the recurrence relation allows for easily computing millions of terms of the sequence $(\mathbb{E}(X_n + Y_n))_{n \geq 1}$, which can be used to numerically confirm the error term in Theorem 2.1.

3 The Constant Appearing in the AEP for Plancherel Measure

We consider the random variables $Z_n = Z_n(\lambda) := \sum_{u \in \lambda} \log h_u$, where $\lambda \vdash n$ is distributed according to the Plancherel measure, and denote $z_n := \mathbb{E} Z_n$. The sequence starts

$$\begin{aligned} (z_1, \dots, z_5) &= \left(0, \log 2, \frac{\log 2}{3} + \log 3, \frac{17}{6} \log 2 + \frac{\log 3}{4}, \frac{13}{6} \log 2 + \frac{7}{10} \log 3 + \frac{7}{12} \log 5 \right) \\ &\approx (0, 0.6931471806, 1.329661349, 2.238570083, 3.209686276) \end{aligned}$$

The first few asymptotic terms of z_n will lead to a representation of the constant H , conjectured by Vershik and Kerov [28] to exist as the limit in probability of random variables $-\frac{1}{\sqrt{n}} \log \mathbb{P}l^{(n)}(\lambda)$, where $\mathbb{P}l^{(n)}(\lambda) := n! \left(\prod_{u \in \lambda} \frac{1}{h_u} \right)^2$, with $\lambda \vdash n$ again distributed according to the Plancherel measure. A strengthening of the conjecture (convergence in L_p for $p < \infty$) has been proved by Bufetov [4], from which we borrowed above notation, and an expression for H in terms of a threefold integral has been given in [4, Eq. (15)]. We aim here at a less involved representation of H , and at more terms of an asymptotic expansion of $\mathbb{E} [n^{-\frac{1}{2}} (2Z_n - \log n!)]$.

Theorem 3.1 Let $H_n = 1 + \frac{1}{2} + \dots + \frac{1}{n}$ denote the n th harmonic number. Then, as $n \rightarrow \infty$, we have

$$-\frac{\mathbb{E} \log \mathbb{P}^{(n)}(\lambda)}{\sqrt{n}} = H - \left(\frac{13}{24} \log n + \frac{13\gamma}{12} + \log \sqrt{2\pi} + \frac{1}{4} - h'(0) \right) \frac{1}{\sqrt{n}} + o(n^{-\frac{1}{2}}),$$

where

$$H = \frac{16}{3\pi^2}(4\gamma + 1) + \frac{64}{\pi^2} \sum_{\ell \geq 2} \frac{\ell^2}{4\ell^2 - 1} \left(\log \ell - H_\ell + \gamma + \frac{1}{2\ell} \right) \approx 1.87702830628$$

and

$$h'(0) = \sum_{\ell \geq 2} \ell \left(H_\ell - \log \ell - \gamma - \frac{1}{2\ell} + \frac{1}{12\ell^2} \right) \approx 0.001562493.$$

Proof As we will prove in the appendix, the Kronecker delta defined on $\mathbb{N} \times \mathbb{N}$ can be expressed in terms of the polynomials $p(x, r)$ given in (2.1) as follows,

$$\delta_{\ell,n} = \sum_{r=\ell-1}^{\infty} (-1)^{\ell+r+1} \frac{2\ell^2}{(r + \ell + 1)!(r - \ell + 1)!} p(n, r). \tag{3.1}$$

From this we deduce

$$\log n = \sum_{\ell \geq 2} \log \ell \sum_{r \geq \ell-1} \frac{2(-1)^{\ell+r+1} \ell^2}{(r + \ell + 1)!(r - \ell + 1)!} p(n, r) = 2 \sum_{r \geq 2} (-1)^r g(r) p(n, r - 1),$$

for $n \in \mathbb{N}$, where g is given by

$$g(r) = \sum_{\ell=2}^r \frac{(-1)^\ell \ell^2 \log \ell}{\Gamma(r + \ell + 1)\Gamma(r - \ell + 1)}.$$

We want to extend g to a meromorphic function in the right halfplane $\Re r > -1$. Therefore we employ

$$\log \ell = H_\ell - \gamma - \frac{1}{2\ell} + \frac{1}{12\ell^2} + \mathcal{O}(\ell^{-4}),$$

and the identities (all with easy proofs, only the last one is proven in the appendix)

$$\sum_{\ell=2}^r \frac{(-1)^\ell \ell^2}{\Gamma(r + \ell + 1)\Gamma(r - \ell + 1)} = \frac{1}{\Gamma(r)\Gamma(r + 2)} \tag{3.2a}$$

$$\sum_{\ell=2}^r \frac{(-1)^\ell \ell}{\Gamma(r + \ell + 1)\Gamma(r - \ell + 1)} = \frac{3(r - 1)}{2(2r - 1)\Gamma(r)\Gamma(r + 2)} \tag{3.2b}$$

$$\sum_{\ell=2}^r \frac{(-1)^\ell}{\Gamma(r + \ell + 1)\Gamma(r - \ell + 1)} = \frac{r - 1}{2r\Gamma(r)\Gamma(r + 2)} \tag{3.2c}$$

$$\sum_{\ell=2}^r \frac{(-1)^\ell \ell^2 (H_\ell - 1)}{\Gamma(r + \ell + 1)\Gamma(r - \ell + 1)} = \frac{1}{4(r - 1)(2r - 1)\Gamma(r)^2}. \tag{3.2d}$$

By Euler’s reflection formula, for complex $r \notin \mathbb{Z}$ and for real $\ell \rightarrow \infty$ we have

$$\Gamma(r + \ell + 1)\Gamma(r - \ell + 1) \frac{\sin \pi(\ell - r)}{\pi} = \frac{\Gamma(r + \ell + 1)}{\Gamma(\ell - r)} \sim \ell^{2r+1}.$$

Therefore the following series

$$h(r) := \sum_{\ell \geq 2} (-1)^\ell \ell^2 \frac{\log \ell - H_\ell + \gamma + \frac{1}{2\ell} - \frac{1}{12\ell^2}}{\Gamma(r + \ell + 1)\Gamma(r - \ell + 1)}$$

converges for $\Re r > -1$, and satisfies $h(1) = h(0) = 0$, with $h'(0)$ as given in the theorem. Hence

$$g(z) = h(z) + \frac{1 - \gamma}{\Gamma(z)\Gamma(z + 2)} + \frac{1}{4(z - 1)(2z - 1)\Gamma(z)^2} - \frac{(z - 1)(16z + 1)}{24z(2z - 1)\Gamma(z)\Gamma(z + 2)}$$

is the sought extension, meromorphic for $\Re z > -1$. Now

$$\begin{aligned} z_n &= \frac{1}{n!} \sum_{\lambda \vdash n} f_\lambda^2 \sum_{u \in \lambda} \log h_u = 2 \sum_{r \geq 2} (-1)^r g(r) \frac{1}{n!} \sum_{\lambda \vdash n} f_\lambda^2 \sum_{u \in \lambda} p(h_u, r - 1) \\ &= 2 \sum_{r \geq 2} (-1)^r g(r) K_{r-1} \binom{n}{r} \\ &= (-1)^n \frac{n!}{2\pi i} \oint_C \frac{\phi(z)}{z(z - 1)(z - 2) \dots (z - n)} dz, \end{aligned}$$

where

$$\phi(z) = 2g(z) \frac{\Gamma(2z)\Gamma(2z - 1)}{\Gamma(z + 1)^2\Gamma(z)}.$$

Here C is a contour that encircles integers $2, \dots, n$, but neither any other integers, nor poles of ϕ .

By computing (leading asymptotic terms of) residues of $\Phi_n(z) := \phi(z) \frac{n!\Gamma(-z)}{\Gamma(n+1-z)}$ at $1, \frac{1}{2}$, and 0 , we obtain

Table 2 Some terms of the sequence approaching H

n	$\frac{1}{\sqrt{n}}(2z_n - \log n!)$	n	$\frac{1}{\sqrt{n}}(2z_n - \log n!)$	n	$\frac{1}{\sqrt{n}}(2z_n - \log n!)$
2	0.4901290717	7	0.8208116414	128	1.4880650932
3	0.5008878635	8	0.8690239552	256	1.5781760349
4	0.649543169	16	1.0657023619	512	1.6489336120
5	0.7297992837	32	1.2347905493	1024	1.7039138626
6	0.7726513179	64	1.3748129422	2048	1.7462734777

$$z_n = \frac{n \log n - n}{2} - \frac{1}{4} + \frac{4}{9\pi^2} \left(24\gamma + 7 - 18\pi h\left(\frac{1}{2}\right) \right) n^{\frac{1}{2}} - \frac{\log n}{48} - \frac{13\gamma}{24} + \frac{1}{8} + \frac{h'(0)}{2} + \mathcal{O}(n^{-\frac{1}{2}}) + \frac{1}{2\pi i} \oint_{C'} \Phi_n(z) dz,$$

where C' encircles the interval $[0, n]$, but no poles of the integrand outside that interval. As shown in the appendix, the latter integral is $o(1)$, thus we arrive at

$$\begin{aligned} -\mathbb{E} \frac{\log \mathbb{P}^{(n)}(\lambda)}{\sqrt{n}} &= \mathbb{E} \frac{2Z_n - \log n!}{\sqrt{n}} \\ &= \frac{8}{9\pi^2} \left(24\gamma + 7 - 18\pi h\left(\frac{1}{2}\right) \right) - \frac{13 \log n}{24 \sqrt{n}} \\ &\quad - \left(\frac{13\gamma}{12} + \log \sqrt{2\pi} + \frac{1}{4} - h'(0) \right) \frac{1}{\sqrt{n}} + o(n^{-\frac{1}{2}}). \end{aligned}$$

Finally, for the evaluation of $h(\frac{1}{2})$, we use $\Gamma(\frac{1}{2} + \ell + 1)\Gamma(\frac{1}{2} - \ell + 1) = (-1)^{\ell+1} \frac{\pi}{4} (4\ell^2 - 1)$, as well as $\sum_{\ell \geq 2} (4\ell^2 - 1)^{-1} = \frac{1}{6}$, leading to

$$h\left(\frac{1}{2}\right) = -\frac{4}{\pi} \sum_{\ell \geq 2} \frac{\ell^2}{4\ell^2 - 1} \left(\log \ell - H_\ell + \gamma + \frac{1}{2\ell} \right) + \frac{1}{18\pi},$$

which completes the proof. □

Remark 3.2 Note that the term $2 \sum_{r \geq 2} (-1)^r g(r) K_{r-1} \binom{n}{r}$ can be used to compute z_n for values of n so large that naively generating all partitions $\lambda \vdash n$ is not an option. Of course, care has to be taken, since cancellations will occur in numerical computations because of alternating signs of summands. Table 2 shows that $\frac{1}{\sqrt{n}}(2z_n - \log n!)$ is slowly approaching H from below, with the values obtained in [29, Table 1] from simulations fitting neatly into this pattern. The convergence rate is in good accordance with the error term given in Theorem 3.1. Observe that $\frac{13\gamma}{12} + \log \sqrt{2\pi} + \frac{1}{4} - h'(0) \approx 1.792693$. For $n = 2048$ we then get $H - (\frac{13}{24} \log 2048 + 1.792693) / \sqrt{2048} \approx 1.746154$, which matches the table entry fairly well.

4 The Expected Side Length of the Durfee Square

Here we consider the side length of the Durfee square of partition λ ,

$$D(\lambda) := \max\{i : \lambda_i \geq i\},$$

and we denote the restriction of $D(\lambda)$ to $\lambda \vdash n$ distributed according to Plancherel measure by D_n . With respect to uniform measure, where all partitions $\lambda \vdash n$ are equally likely, the expectation and the most likely value of $D(\lambda)$ have been studied in [5, 6, 18]. Regarding Plancherel measure, it is known since the days of the limit shape theorem (see Theorem 5.1 in the next section) that $\frac{1}{\sqrt{n}}D_n \rightarrow \frac{2}{\pi}$ in probability. Furthermore, we may deduce from [2, Theorem 3.6] convergence in distribution of $\frac{\pi}{\sqrt{\log n}}(D_n - \frac{2}{\pi}\sqrt{n})$ to a standard normal random variable. Should that convergence in distribution be accompanied by convergence of second moments, $\mathbb{E} D_n = \frac{2}{\pi}\sqrt{n} + \mathcal{O}(\sqrt{\log n})$ would follow. We are not aware of a proof of such result, let alone of any results in the literature regarding fine asymptotics of $\mathbb{E} D_n$.

Theorem 4.1 *Let $d_n := \mathbb{E} D_n$. Then, as $n \rightarrow \infty$, we have*

$$d_n = \frac{2}{\pi}\sqrt{n} + \left(\frac{3}{16\pi} - \frac{e^2}{8\pi} \sin(4\sqrt{n}) \right) \frac{1}{\sqrt{n}} + \mathcal{O}(n^{-1}).$$

Proof In terms of contents c_u of a Young diagram λ , we have

$$D(\lambda) = \sum_{u \in \lambda} \delta_{0, c_u}.$$

Define polynomials in terms of the polynomials $p(x, r)$ given in (2.1) via

$$q(x, r) := \prod_{i=0}^{r-1} (x^2 - i^2) = \begin{cases} x^2 p(x, r - 1), & r \geq 1, \\ 1, & r = 0. \end{cases}$$

These also allow for a representation of the Kronecker delta, similar to (3.1),

$$\delta_{\ell, n} = \sum_{r=\ell}^{\infty} (-1)^{\ell+r} \frac{2 - \delta_{0, n}}{(r + \ell)!(r - \ell)!} q(n, r), \tag{4.1}$$

now valid for non-negative integers ℓ, n . By [26, Eq. (7)], see also [8, Theorem A.1], we have

$$\frac{1}{n!} \sum_{\lambda \vdash n} f_{\lambda}^2 \sum_{u \in \lambda} q(c_u, r) = \frac{(2r)!}{(r + 1)!} \binom{n}{r + 1}. \tag{4.2}$$

This leads to the representation

$$d_n = \frac{1}{n!} \sum_{\lambda \vdash n} f_\lambda^2 \sum_{u \in \lambda} \delta_{0,c_u} = \sum_{r \geq 1} \frac{(-1)^{r+1} (2r - 2)!}{(r - 1)! 2^r} \binom{n}{r}. \tag{4.3}$$

For the Poisson generating function $D(z) := e^{-z} \sum_{n \geq 0} \frac{d_n}{n!} z^n$ we obtain

$$D(z) = z \sum_{k \geq 0} \frac{(-z)^k (2k)!}{k!^2 (k + 1)!^2} = z \left(2J_0^2(2\sqrt{z}) - \frac{J_0(2\sqrt{z})J_1(2\sqrt{z})}{\sqrt{z}} + 2J_1^2(2\sqrt{z}) \right),$$

where J_0 and J_1 are Bessel functions of the first kind. We may use $D(z)$ to recover asymptotics of d_n , see [15, Sect. 5.11.4] for asymptotic expansions of Bessel functions. We find

$$D(z) = \frac{2}{\pi} \sqrt{z} - \frac{1}{16\pi \sqrt{z}} - \frac{\sin(4\sqrt{z})}{8\pi \sqrt{z}} + \frac{3 \cos(4\sqrt{z})}{64\pi z} + \mathcal{O}(z^{-\frac{3}{2}}),$$

for $|z| \rightarrow \infty$, $|\arg z| \leq \pi - \delta$ with $\delta > 0$. A uniform bound is furnished by $|D(z)| \leq \cosh(4\sqrt{|z|})$. Evaluating now $d_n = \frac{n!}{2\pi i} \oint_C z^{-n-1} e^z D(z) dz$, with contour $C := \{z \in \mathbb{C} : |z| = n\}$, observing that there is an approximate saddle point at $z = n$, finishes the proof. □

Remark 4.2 The sequence starts $(d_n)_{n=1}^{10} = (1, 1, 1, \frac{7}{6}, \frac{17}{12}, \frac{33}{20}, \frac{109}{60}, \frac{3217}{1680}, \frac{39703}{20160}, \frac{364859}{181440})$, and it again satisfies a linear recurrence relation,

$$d_{n+3} = \frac{3n^2 + 9n + 8}{(n + 2)(n + 3)} d_{n+2} - \frac{3n + 1}{n + 3} d_{n+1} + \frac{n + 1}{n + 3} d_n, \tag{4.4}$$

with initial conditions

$$d_0 = 0, \quad d_1 = 1, \quad d_2 = 1,$$

readily obtained from (4.3) using *gfun*.

5 Expected Fluctuations Around the Limit Shape Curve

Let us introduce the *limit shape curve*

$$\Omega(u) = \begin{cases} \frac{2}{\pi} \left(u \arcsin \frac{u}{\sqrt{2}} + \sqrt{2 - u^2} \right), & \text{if } |u| \leq \sqrt{2}, \\ |u|, & \text{if } |u| > \sqrt{2}. \end{cases} \tag{5.1}$$

The lower right boundary of the Young diagram of a partition $\lambda \vdash n$, scaled to have unit area, rotated together with parts of positive x -axis and negative y -axis by 135° ,

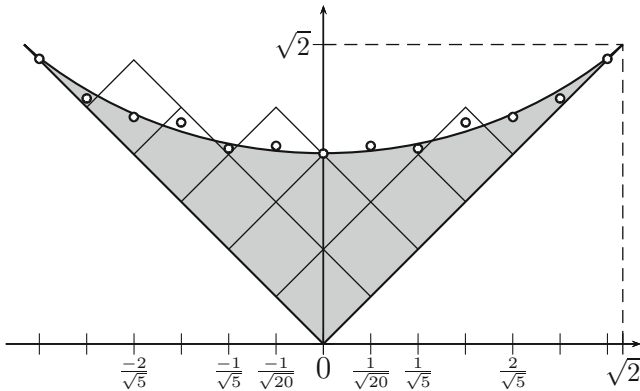


Fig. 4 The Logan–Shepp–Vershik–Kerov limit shape curve Ω (upper boundary of the grey region). Superimposed is the partition $(5, 3, 1, 1) \vdash 10$, properly scaled and rotated by 135° . The dots have coordinates $\left(\frac{a}{\sqrt{20}}, \frac{1}{\sqrt{5}}(\omega_{a,10} + \frac{|a|}{2})\right)$, for $a \in \{-6, \dots, 6\}$

gives rise to a piecewise linear function ψ_λ , also defined on \mathbb{R} . When λ is distributed according to Plancherel measure, the random functions ψ_λ approach the limit shape curve, as $n \rightarrow \infty$, in a sense that is made precise in the following result by Vershik and Kerov [27] and Logan and Shepp [14], which we present following closely [23, Theorem 1.22].

Theorem 5.1 (Limit shape theorem for Plancherel-random partitions) *For all $\varepsilon > 0$, we have $\mathbb{P}(\|\psi_\lambda - \Omega\|_\infty > \varepsilon) \rightarrow 0$ as $n \rightarrow \infty$, i.e., the random functions ψ_λ converge to Ω in probability in the norm $\|\cdot\|_\infty$.*

A discretized version of ψ_λ , defined on the set $\{\frac{a}{\sqrt{2n}} : a \in \mathbb{Z}\}$, can be expressed in terms of contents c_u of λ via $\Psi_\lambda(a) := \sum_{u \in \lambda} \delta_{-a, c_u}$.

Indeed, the set $\left\{\left(\frac{a}{\sqrt{2n}}, \frac{2}{\sqrt{2n}}(\Psi_\lambda(a) + \frac{|a|}{2})\right) : a \in \mathbb{Z}\right\}$ is a subset of the graph of ψ_λ containing, among others, all the points where the slope of ψ_λ changes from 1 to -1 or back. For example, if $\lambda = (5, 3, 1, 1)$, then $(\Psi_\lambda(a))_{a=-4}^4 = (1, 1, 1, 2, 2, 1, 1, 1, 0)$. Define now a related function, $\Phi_\lambda(a) := \frac{1}{2}(\Psi_\lambda(a) + \Psi_\lambda(-a))$, i.e.,

$$\Phi_\lambda(a) = \begin{cases} D(\lambda), & a = 0, \\ \frac{1}{2} \sum_{u \in \lambda} \delta_{|a|, |c_u|}, & \text{else.} \end{cases}$$

This symmetrised function is used because it can be expressed in terms of Kronecker deltas restricted to pairs of nonnegative integers, thus allowing to use the representation (4.1). Next, let $\omega_{a,n} := \mathbb{E} \Phi_\lambda(a)$, with $\lambda \vdash n$ distributed according to the Plancherel measure, and define a sequence of functions

$$\tilde{\Omega}_n(u) := \sqrt{\frac{2}{n}} \left(\omega_{\lfloor \sqrt{2nu} \rfloor, n} + \frac{1}{2} |\lfloor \sqrt{2nu} \rfloor| \right),$$

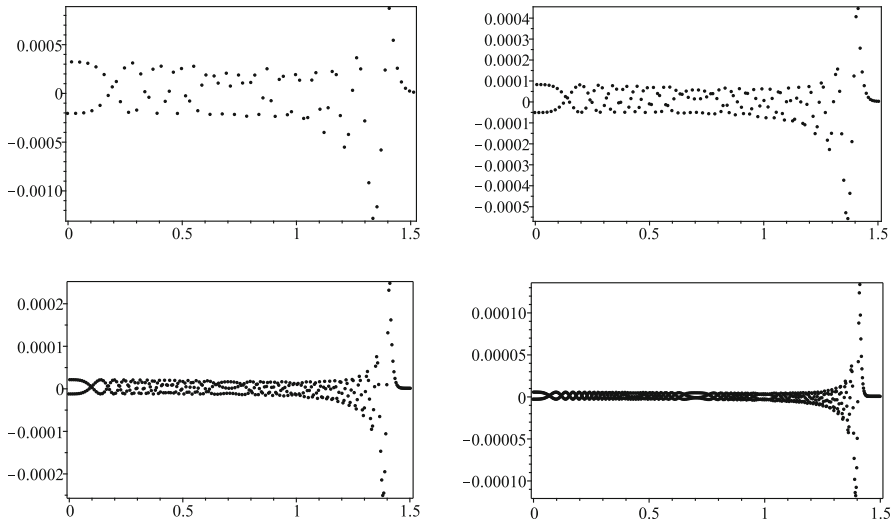


Fig. 5 Plots of $\tilde{\Omega}_n(u) - \Omega(u)$ for roughly quadrupling values of $n \in \{1573, 6230, 24798, 98943\}$ from top left to bottom right

that one would expect to converge to $\Omega(u)$, although such convergence is not implied by Theorem 5.1. By [2, Theorem 3.6] we have convergence in distribution of $\frac{\pi\sqrt{n}}{\sqrt{2\log n}} \left[\sqrt{\frac{2}{n}} \left(\Phi_\lambda(\lfloor \sqrt{2nu} \rfloor) + \frac{1}{2} |\lfloor \sqrt{2nu} \rfloor| \right) - \Omega(u) \right]$ to a standard normal random variable in case that $|u| < \sqrt{2}$. Should there also be convergence of second moments, $\tilde{\Omega}_n(u) = \Omega(u) + \mathcal{O}\left(\sqrt{\frac{\log n}{n}}\right)$ would follow for $|u| < \sqrt{2}$. See Fig. 4 for the limit shape curve, and, scaled to unit area, a superimposed partition of 10, and values $\tilde{\Omega}_{10}(u)$ for $u \in \{\frac{a}{\sqrt{20}} : -6 \leq a \leq 6\}$. There is a seeming coincidence on the y-axes, yet $\Omega(0) = \frac{2\sqrt{2}}{\pi} \approx .900316$, $\frac{\omega_{0,10}}{\sqrt{5}} = \frac{364859}{181440\sqrt{5}} \approx .899305$, and the ordinate of the upper corner of the rotated Young diagram, $\frac{2}{\sqrt{5}} \approx .894427$, are all different.

An alternating sum representation of $\omega_{a,n}$, building upon (4.1), is the following

$$\omega_{a,n} = \sum_{r>a} \frac{(-1)^{r+a+1}(2r-2)!}{(r-1+a)!(r-1-a)!r!} \binom{n}{r}, \tag{5.2}$$

which again gives rise to a linear recurrence relation (obtained using *gfun*)

$$\begin{aligned} &(n+4)(n+a+3)(n-a+3)\omega_{a,n+4} \\ &= [4n^3 + 32n^2 + (86 - 2a^2)n + 78 - 7a^2]\omega_{a,n+3} \\ &\quad - (n+3)(6n^2 + 22n + 20 - a^2)\omega_{a,n+2} \\ &\quad + (n+1)(n+2)(n+3)(4\omega_{a,n+1} - \omega_{a,n}), \end{aligned}$$

holding for $n \geq a - 2$, with initial conditions

$$\omega_{a,n} = 0, \text{ for } \min(0, a - 2) \leq n \leq a, \quad \omega_{a,a+1} = \frac{1}{(a + 1)!}.$$

Note that there is a common factor $(n + 1)$ in the recurrence relation, when $a = 2$. Note also that setting $a = 0$ yields a recurrence relation with both order and degree one larger than the one given in (4.4).

As was done for d_n , asymptotics via Poisson generating functions (which can again be expressed in terms of Bessel functions) can be obtained also for $\omega_{a,n}$ for fixed integer $a > 0$:

$$\omega_{a,n} + \frac{a}{2} = \frac{2}{\pi} \sqrt{n} + \left(\frac{4a^2 + 3}{16\pi} - (-1)^a \frac{e^2}{8\pi} \sin(4\sqrt{n}) \right) \frac{1}{\sqrt{n}} + \mathcal{O}(n^{-1}). \tag{5.3}$$

In order to obtain asymptotics of $\omega_{a,n}$ for n and a simultaneously approaching ∞ , which would be needed for asymptotics of $\tilde{\Omega}_n(u)$, one could use the parametrization $n = 2\kappa^2 \in \mathbb{N}$, $a = \lfloor 2\kappa u \rfloor$, and consider

$$\frac{\omega_{\lfloor 2\kappa u \rfloor, 2\kappa^2}}{\kappa} = \frac{1}{2\pi i} \oint_{C''} \frac{1}{\kappa} \frac{(-1)^{\lfloor 2\kappa u \rfloor + 1} \Gamma(-z) \Gamma(2z - 1)}{\Gamma(z + \lfloor 2\kappa u \rfloor) \Gamma(z - \lfloor 2\kappa u \rfloor) \Gamma(z + 1)} \frac{(2\kappa^2)! \Gamma(-z)}{\Gamma(2\kappa^2 + 1 - z)} dz,$$

implied by (5.2), where C'' is a contour that encircles integers $1, \dots, 2\kappa^2$, but neither any other integers, nor poles of the integrand. Outside C'' , the integrand has poles at $\frac{1}{2}$, at 0, and at all negative half-integers. For fixed u it turns out that each residue contributes to the leading (constant) term of the asymptotics in the limit $\kappa \rightarrow \infty$, with the sum of those contributions converging, but for fixed κ the sum of residues does not converge. Balancing those two limiting processes (taking more and more residues into account, letting $\kappa \rightarrow \infty$) and at the same time bounding the integral over a sequence of correspondingly deformed contours appears to be intricate, so unfortunately we have not been able to prove $\tilde{\Omega}_n(u) \rightarrow \Omega(u)$ for $u \neq 0$.

Using holonomicity of $(\omega_{a,n})_{n \geq 0}$ to generate many terms of that sequence for many values of a , we obtain the plots in Figs. 5 and 6. The values for n in Fig. 5 and in the second plot in Fig. 6 have been chosen to satisfy $\sin(4\sqrt{n}) \approx 1$ in order to give maximal weight to the term $(-1)^a$ present in (5.3) and thus ensure better comparability of the plots in the vicinity of 0. The value of n in the first plot of Fig. 6 satisfies $\sin(4\sqrt{n}) \approx 0$. We conclude this section with some (non-rigorous) observations based on these plots.

- (a) Regarding convergence of the sequence $(\tilde{\Omega}_n(\cdot))_{n \geq 1}$ to $\Omega(\cdot)$, the plots in Fig. 5 suggest that for any $\epsilon > 0$ we have $\max_{0 \leq u \leq \sqrt{2} - \epsilon} |\tilde{\Omega}_n(u) - \Omega(u)| = \mathcal{O}(\frac{1}{n})$. Near $\sqrt{2}$ we only seem to have $\max_{|u - \sqrt{2}| \leq \epsilon} |\tilde{\Omega}_n(u) - \Omega(u)| = \mathcal{O}(\frac{1}{\sqrt{n}})$, with this slower convergence in agreement with the convergence rate stated in [2, Eq. (3.4)] for convergence of $\tilde{\Omega}_n(\sqrt{2}) \rightarrow \Omega(\sqrt{2})$ in probability, building upon results from [1, 3, 12, 19] addressing the *problem of the longest increasing subsequence*, see also [23].

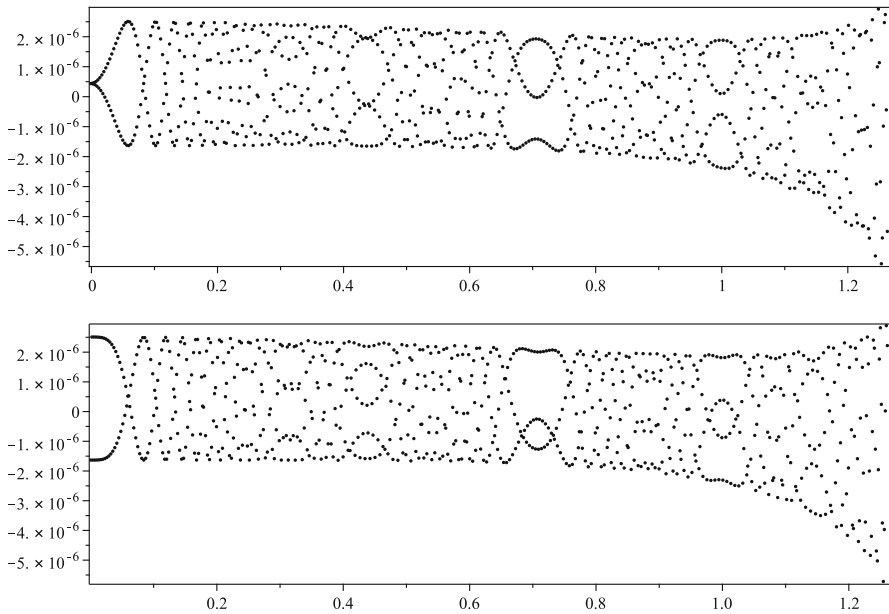


Fig. 6 Plots of $\tilde{\Omega}_n(u) - \Omega(u)$ for $n = 200415$ (upper plot) and for $n = 200767$ (lower plot), with u varying in the discrete set $\{\frac{a}{\sqrt{2n}} : 0 \leq a \leq 800\}$

(b) A common feature of all plots in Figs. 5 and 6 is the presence of subintervals of “smooth” behavior surrounded by regions of more “irregular” behavior. In order to enforce “smooth” dependence of $\omega_{a,n}$ on a , one would, in the light of (5.3), restrict to odd (or to even) a . However, this would only work for $0 \leq a \leq \alpha_n$ with $\alpha_n = o(\sqrt{n})$. The location of the first “peak” to the right of 0 seems to suggest, that $\alpha_n = \Theta(n^{\frac{1}{4}})$ may hold. For larger a it is no longer useful to distinguish between even and odd, instead one should consider $\omega_{a,n}$ evaluated at a belonging to other arithmetic progressions: Near $\frac{a}{\sqrt{2n}} = \sqrt{2} \cos \frac{\pi}{3} \approx 0.707$ the way to go would be to consider $(\omega_{a+3k,n})_k$, whereas near $\frac{a}{\sqrt{2n}} = \sqrt{2} \cos \frac{\pi}{4} = 1$ it would be $(\omega_{a+4k,n})_k$. Every fifth term should be taken near $\sqrt{2} \cos \frac{\pi}{5} \approx 1.144$ and $\sqrt{2} \cos \frac{2\pi}{5} \approx 0.437$. We expect this pattern to continue, with regions of smoothness near $\sqrt{2} \cos \frac{\ell\pi}{m}$ for $1 \leq \ell < \frac{m}{2}$, and ℓ, m coprime. For larger m these regions will become noticeable only if n gets large enough, and those regions will shrink with n further increasing, making room for yet other regions to pop up.

6 A Heuristic Upper Bound for the Variance of the Number of Bumping Steps in the RSK Algorithm

Let $L_n := X_n + Y_n$, $\ell_n := \mathbb{E}(X_n + Y_n)$, and $v_n := \text{Var}(X_n + Y_n)$. We now give a heuristic derivation of ℓ_n , and an upper bound for v_n based on [13]. Let $\Omega(x)$ be the function defined in (5.1), describing the limit shape of normalized Young diagrams

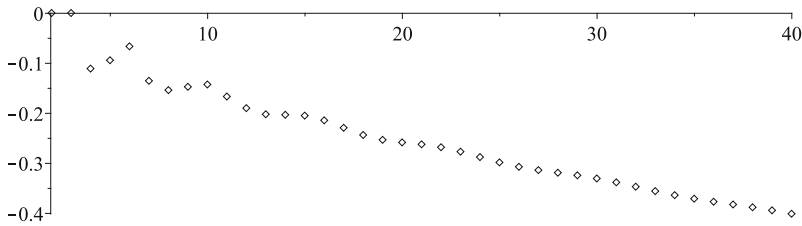


Fig. 7 A plot of $\text{Cov}(L_n - L_{n-1}, L_{n-1})$ for $n \in \{2, \dots, 40\}$

with respect to Plancherel measure. Denote $s(x) := \frac{1}{\pi}\sqrt{2-x^2}$, the density of the semicircle distribution with support $[-\sqrt{2}, \sqrt{2}]$. As shown in [13], this is also the limiting density of the random abscissa of a newly inserted box into a scaled and rotated Young diagram that closely resembles the limit shape curve, when new insertions are made according to the Plancherel growth process, that ensures that at each stage of the process the Young diagram is distributed according to Plancherel measure, see also [23, Sect. 1.19]. This leads to

$$\ell_n - \ell_{n-1} \sim \sqrt{2n} \int_{-\sqrt{2}}^{\sqrt{2}} \Omega(x)s(x)dx = \frac{128}{9\pi^2} \sqrt{n},$$

and thus $\ell_n \sim \frac{256}{27\pi^2} n^{\frac{3}{2}}$. Moreover, assuming independence of L_{n-1} and $L_n - L_{n-1}$,

$$v_n - v_{n-1} \sim 2n \int_{-\sqrt{2}}^{\sqrt{2}} \left(\Omega(x) - \frac{64\sqrt{2}}{9\pi^2} \right)^2 s(x)dx = \frac{54\pi^4 + 2835\pi^2 - 32768}{162\pi^4} n,$$

and thus

$$v_n \sim \frac{54\pi^4 + 2835\pi^2 - 32768}{324\pi^4} n^2 \approx 0.01496867061 n^2.$$

Numerically we have e.g. $\frac{v_{50}}{50^2} \approx 0.01216526413$.

Indeed, L_{n-1} and $L_n - L_{n-1}$ seem to be negatively correlated. The sequence of covariances $(\text{Cov}(L_n - L_{n-1}, L_{n-1}))_{n \geq 2}$ starts $(0, 0, -\frac{1}{9}, -\frac{17}{180}, -\frac{1}{15}, -\frac{61}{450}, -\frac{863}{5600}, \dots)$, staying negative up to $n = 40$ with roughly linear growth, see Fig. 7.

So it seems that in the light of Lemma 1.2 one can safely guess that for the number Y_n of bumping steps $\text{Var } Y_n = \Theta(n^2)$ holds. It would be desirable to have a proof for that, and also know at least the leading asymptotic term of $\text{Var } Y_n$.

7 Conclusion

In this paper we have obtained asymptotics of expectations of certain statistics of Plancherel distributed Young diagrams. That these statistics could be expressed in terms of hook lengths and contents of the boxes of such diagram was essential,

as it allowed us to invoke polynomiality results for Plancherel averages, leading to representations of expectations as binomial convolutions, that make for easier asymptotic treatment. We hope that this approach will help to analyse further statistics of Plancherel distributed Young diagrams. Now polynomiality results have also been found for measures different from Plancherel (such as the Jack deformation of Plancherel measure, see [20]), or for subclasses of Plancherel distributed Young diagrams, such as strict partitions (see [11, 16, 17]). In case that appropriate substitutes for (2.3) or (4.2) are at hand, it is reasonable to believe that certain statistics in these settings could also be analysed along the lines of this paper.

Acknowledgements We would like to thank two anonymous referees, whose suggestions led to substantial improvements of the paper.

Funding Open access funding provided by University of Vienna.

Data Availability Not applicable.

Declarations

Conflict of interest The author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

8. Appendix

8.1 Proof of Equation (2.2)

We use $p(n, r) = \frac{(2r+1)!}{n} \binom{n+r}{2r+1}$ and rewrite (2.2) as

$$n^2 = \sum_{r \geq 0} \frac{(-1)^r}{1-2r} \binom{2r}{r} \binom{n+r}{2r+1} =: S_n$$

Denoting by Δ the forward difference operator, we will show that $\Delta^3 S$ is the zero sequence, which together with $S_1 = 1$, $\Delta S_1 = 3$, $\Delta^2 S_1 = 2$ yields $S_n = n^2$ for $n \in \mathbb{N}$. Now

$$\begin{aligned} \Delta^3 S_n &= \sum_{r \geq 1} \frac{(-1)^r}{1-2r} \binom{2r}{r} \binom{n+r}{2r-2} = \sum_{r \geq 1} (-1)^{r+1} \frac{2}{r} \binom{2r-2}{r-1} \binom{n+r}{2r-2} \\ &= 2 \sum_{r \geq 1} \frac{(-1)^{r+1}}{n+2} \binom{n+r}{r-1} \binom{n+2}{r} = \frac{2}{n+2} \sum_{r \geq 0} (-1)^r \binom{n+r+1}{n+1} \binom{n+2}{r+1} = 0, \end{aligned}$$

where the last equality follows from [10, Eq. (5.24)]. □

8.2 Proof of Equation (3.1)

The equation is easily checked for $\ell > n$, since then also $r \geq n$ and thus $p(n, r) = 0$. For $1 \leq \ell \leq n$ we have

$$\begin{aligned} & \sum_{r=\ell-1}^{n-1} \frac{2\ell^2(-1)^{r+\ell+1}}{(r+\ell+1)!(r-\ell+1)!} \frac{(2r+1)!}{n} \binom{n+r}{2r+1} \\ &= \frac{2\ell^2(-1)^{\ell+1}}{n(n+\ell)} \sum_{r=\ell-1}^{n-1} (-1)^r \binom{n+r}{n+\ell-1} \binom{n+\ell}{r+\ell+1} \\ &= \frac{2\ell^2(-1)^{\ell+1}}{n(n+\ell)} (-1)^{n+1} \delta_{\ell,n} = \delta_{\ell,n}, \end{aligned}$$

treating the case $\ell = n$ directly, and using [10, Eq. (5.24)] again for $n > \ell$. □

8.3 Proof of Equation (4.1)

We use $q(n, r) = \frac{n(2r)!}{n+r} \binom{n+r}{2r}$ for $n > 0$, and $q(0, 0) = 1$. The equation is easily checked for $n = 0$, and for $\ell > n$, since then also $r > n$ and thus $q(n, r) = 0$. For $n > 0, 0 \leq \ell \leq n$ we have

$$\begin{aligned} & \sum_{r=\ell}^n \frac{2(-1)^{r+\ell}}{(r+\ell)!(r-\ell)!} \frac{n(2r)!}{n+r} \binom{n+r}{2r} = (-1)^\ell \frac{2n}{n+\ell} \sum_{r=\ell}^n (-1)^r \binom{n+r-1}{n+\ell-1} \binom{n+\ell}{r+\ell} \\ &= \frac{2n(-1)^\ell}{n+\ell} (-1)^n \delta_{\ell,n} = \delta_{\ell,n}, \end{aligned}$$

treating the case $\ell = n$ directly, and using [10, Eq. (5.24)] again for $n > \ell$. □

8.4 Proof of Equation (3.2d)

Using $H_\ell - 1 = \sum_{k=2}^\ell \frac{1}{k}$, and interchanging summation, we obtain, using [10, Eq. (5.16)] at several places,

$$\begin{aligned} \sum_{\ell=2}^r \frac{(-1)^\ell \ell^2 (H_\ell - 1)}{(r+\ell)!(r-\ell)!} &= \frac{1}{(2r)!} \sum_{k=2}^r \frac{1}{k} \sum_{\ell=k}^r (-1)^\ell [r^2 - (r+\ell)(r-\ell)] \binom{2r}{r-\ell} \\ &= \frac{1}{(2r)!} \sum_{k=2}^r \frac{1}{k} \sum_{\ell=k}^r (-1)^\ell \left[r^2 \binom{2r}{r-\ell} - 2r(2r-1) \binom{2r-2}{r-\ell-1} \right] \\ &= \frac{1}{(2r)!} \sum_{k=2}^r \frac{1}{k} (-1)^k \left[r^2 \binom{2r-1}{r-k} - 2r(2r-1) \binom{2r-3}{r-k-1} \right] \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{(2r)!} \sum_{k=2}^r \frac{(-1)^k r}{r-1} (k-1) \binom{2r-1}{r-k} \\
 &= \frac{1}{(2r)!} \sum_{k=2}^r \frac{(-1)^k r}{r-1} \left[(r-1) \binom{2r-1}{r-k} - (2r-1) \binom{2r-2}{r-k-1} \right] \\
 &= \frac{1}{(2r)!} \frac{r}{r-1} \left[(r-1) \binom{2r-2}{r} - (2r-1) \binom{2r-3}{r} \right] \\
 &= \frac{1}{(2r)!} \frac{r}{r-1} \frac{r}{r-2} \binom{2r-3}{r} = \frac{1}{4(r-1)(2r-1)(r-1)!^2}.
 \end{aligned}$$

□

8.5 Saddle Point Evaluation of the Integral $I_n := \frac{1}{2\pi i} \oint_{C'} f(z) \frac{n! \Gamma(-z)}{\Gamma(n+1-z)} dz$

This integral appears in the proof of Theorem 2.1, see Sect. 2 for relevant notation. Putting $n = m^2$, the integrand may be rewritten as

$$G(z) := \frac{\pi \Gamma(m^2 + 1) \Gamma^2(2z - 1)}{\sin(\pi z) \Gamma(m^2 + 1 - z) (2z - 3) \Gamma^3(z + 1) \Gamma^3(z)}.$$

Denoting by ψ the digamma function, we have

$$\begin{aligned}
 \frac{G'(z)}{G(z)} &= \psi(m^2 + 1 - z) + 4\psi(2z - 1) - 3\psi(z + 1) - 3\psi(z) - \pi \cot \pi z - \frac{2}{2z - 3} \\
 &\sim \log \frac{(z - m^2 - 1)(2z - 1)^4}{(z + 1)^3 z^3} - \frac{1}{z - \frac{1}{2}} + \frac{3}{2z + 2} + \frac{3}{2z} - \frac{1}{z - \frac{3}{2}},
 \end{aligned}$$

with error terms $\mathcal{O}(m^{-2}) + \mathcal{O}(z^{-2})$, holding for $m \rightarrow +\infty, |z| \rightarrow \infty$, subject to $z = o(m^2)$ and $\delta \leq |\arg z| \leq \pi - \delta$ for some $\delta > 0$.

Two approximate saddle points of $G(z)$ are $\zeta := 6 + 4mi$ and $\bar{\zeta} = 6 - 4mi$. Indeed, $\frac{d}{dz} \log G(\zeta) = \frac{G'(\zeta)}{G(\zeta)} = \mathcal{O}(\frac{1}{m^2})$, and $\frac{d^2}{dz^2} \log G(\zeta) = \frac{i}{2m} + \mathcal{O}(\frac{1}{m^2})$, which suggests a contour directed north-west in the point ζ : Let $z = \zeta + e^{i\frac{\pi}{4}}u$ and observe

$$G(z) = G(\zeta) \exp \left(\mathcal{O}(\frac{u}{m^2}) + \frac{i}{4m} (e^{i\frac{\pi}{4}}u)^2 + \mathcal{O}(\frac{u^2}{m^2}) \right) = G(\zeta) e^{-\frac{u^2}{4m}} \left(1 + \mathcal{O}(\frac{u+u^2}{m^2}) \right).$$

Also note that a cumbersome evaluation results in

$$G(\zeta) = \frac{-ie^{8im+8}}{2^{13}\pi m^4} \left(1 + \mathcal{O}\left(\frac{1}{m}\right) \right).$$

Define the counter-clockwise oriented contour C' as the polygon connecting the points

$$\begin{aligned}
 z_0 &:= -\frac{5}{2} + \varepsilon, & z_1 &:= \zeta - m - mi, & z_2 &:= \zeta + m + mi, & z_3 &:= n + 1 + mi, \\
 z_6 &:= \bar{\zeta} - m + mi, & z_5 &:= \bar{\zeta} + m - mi, & z_4 &:= n + 1 - mi,
 \end{aligned}$$

with $\varepsilon > 0$ small, and with segment c_i connecting z_i and z_{i+1} for $0 \leq i \leq 5$, and c_6 connecting z_6 and z_0 . It turns out that the integrals along c_0 and c_6 are of order $\mathcal{O}(m^{-5+2\varepsilon})$, and c_i , for $2 \leq i \leq 4$, make even smaller contributions. Moreover, the combined contribution of c_1 and c_5 is $-2\sqrt{\frac{m}{\pi}} \Im(e^{i\frac{\pi}{4}} G(\zeta)(1 + \mathcal{O}(\frac{1}{m})))$, which, up to error terms of order $\mathcal{O}(m^{-\frac{9}{2}})$, simplifies to

$$-2\sqrt{\frac{m}{\pi}} \Im(e^{i\frac{\pi}{4}} G(\zeta)) = \frac{e^8}{2^{12}\pi^{\frac{3}{2}}m^{\frac{7}{2}}} \Im(ie^{\frac{\pi i}{4}+8im}) = \frac{e^8}{2^{12}\pi^{\frac{3}{2}}m^{\frac{7}{2}}} \cos\left(\frac{\pi}{4} + 8m\right).$$

8.6 Bounding the Integral $J_n := \frac{1}{2\pi i} \oint_{C'} \Phi_n(z) dz$

This integral appears in the proof of Theorem 3.1, see Sect. 3 for relevant notation. Let C' be the boundary of the rectangle with corners $-\frac{1}{2} \pm i4em, n + \frac{1}{2} \pm i4em$, and $m = \sqrt{n}$.

Observe $(-1)^\ell \ell^2 (\log \ell - H_\ell + \gamma + \frac{1}{2\ell} - \frac{1}{12\ell^2}) = \mathcal{O}(\ell^{-2})$, and, abbreviating $\rho = r + \frac{1}{2}$,

$$\left| \frac{\Gamma^2(r+1)}{\Gamma(r+\ell+1)\Gamma(r-\ell+1)} \right| = \left| \frac{r(r-1)\cdots(r-\ell+1)}{(r+1)\cdots(r+\ell)} \right| = \left| \prod_{k=1}^{\ell} \frac{\rho - (k - \frac{1}{2})}{\rho + (k - \frac{1}{2})} \right| \leq 1$$

for $\Re \rho \geq 0$, i.e., for $\Re r \geq -\frac{1}{2}$, therefore $\Gamma^2(r+1)h(r) = \mathcal{O}(1)$ for $\Re r \geq -\frac{1}{2}$, which leads to $\Gamma^2(z+1)g(z) = \mathcal{O}(1)$ for $\Re z \geq -\frac{1}{2}, |z-w| \geq \frac{1}{2}$ for $w \in \{0, \frac{1}{2}, 1\}$. Hence

$$\Phi_n(z) = \mathcal{O}\left(\frac{\Gamma(2z)\Gamma(2z-1)}{\Gamma^4(z+1)\Gamma(z)} \frac{n!\Gamma(-z)}{\Gamma(n+1-z)}\right) = \mathcal{O}\left(\frac{2^{4z}}{z^4\Gamma^2(z+1)\sin\pi z} \frac{\Gamma(n+1)}{\Gamma(n+1-z)}\right),$$

by the reflection and duplication formulas. For $z = -\frac{1}{2} + it$, with $t = \mathcal{O}(\sqrt{n})$, we have $|\Gamma^2(z+1)\sin\pi z| = \pi$ and $\left|\frac{\Gamma(n+1)}{\Gamma(n+1-z)}\right| = \mathcal{O}(n^{-\frac{1}{2}})$, yielding a contribution $\mathcal{O}(n^{-\frac{1}{2}})$ from the integral over the left segment of C' . The contributions from the two horizontal segments is $\mathcal{O}(n^{-\frac{3}{2}})$, while the right segment makes an exponentially small contribution. All this can be seen from the estimates

$$\frac{1}{\Gamma^2(z+1)\sin\pi z} = \mathcal{O}\left(\left(\frac{\sigma^2 + t^2}{e^2}\right)^{-\sigma - \frac{1}{2}}\right),$$

holding for $z = \sigma + it$ with $\sigma \geq -\frac{1}{2}$ and $|z-w| \geq \frac{1}{2}$ for $w \in \mathbb{Z}$, and

$$\left|\frac{2^{4z}\Gamma(n+1)}{\Gamma(n+1-z)}\right| = \mathcal{O}\left((16n)^\sigma \exp\left(\frac{2t^2}{n-\sigma+t}\right)\right),$$

holding for $z = \sigma + it$ with $-\frac{1}{2} \leq \sigma \leq n + \frac{1}{2}$ and $t = \mathcal{O}(\sqrt{n})$.

References

1. Baik, J., Deift, P., Johansson, K.: On the distribution of the length of the longest increasing subsequence of random permutations. *J. Am. Math. Soc.* **12**, 1119–1178 (1999)
2. Bogachev, L.V., Su, Z.: Gaussian fluctuations of Young diagrams under the Plancherel measure. *Proc. R. Soc. Lond. Ser. A* **463**, 1069–1080 (2007)
3. Borodin, A., Okounkov, A., Olshanski, G.: Asymptotics of Plancherel measures for symmetric groups. *J. Am. Math. Soc.* **13**, 481–515 (2000)
4. Bufetov, A.I.: On the Vershik-Kerov conjecture concerning the Shannon-McMillan-Breiman theorem for the Plancherel family of measures on the space of Young diagrams. *Geom. Funct. Anal.* **22**, 938–975 (2012)
5. Canfield, E.R., Corteel, S., Savage, C.D.: Durfee polynomials. *Electron. J. Combin.* **5**, 32 (1998)
6. Canfield, E.R.: From recursions to asymptotics: Durfee and dilogarithmic deductions. *Adv. Appl. Math.* **34**, 768–797 (2005)
7. Frame, J.S., Robinson, G. de B., Thrall, R.M.: The hook graphs of the symmetric groups. *Can. J. Math.* **6**, 316–324 (1954)
8. Fujii, S., Kanno, H., Moriyama, S.: Instanton calculus and chiral one-point functions in supersymmetric gauge theories. *Adv. Theor. Math. Phys.* **12**, 1401–1428 (2008)
9. Flajolet, P., Sedgewick, R.: Mellin transforms and asymptotics: finite differences and Rice’s integrals. *Theor. Comput. Sci.* **144**, 101–124 (1995)
10. Graham, R.L., Knuth, D.E., Patashnik, O.: *Concrete mathematics: a foundation for computer science*. Addison-Wesley, Boston (1994)
11. Han, G.-N., Xiong, H.: Polynomiality of Plancherel averages of hook-content summations for strict, doubled distinct and self-conjugate partitions. *J. Combin. Theory Ser. A* **168**, 50–83 (2019)
12. Johansson, K.: Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. Math.* **153**, 259–296 (2001)
13. Kerov, S.: A differential model for the growth of Young diagrams. In: *Proceedings of the St. Petersburg Mathematical Society*, Vol. IV, Amer. Math. Soc. Transl. Ser. 2, vol. 188, pp. 111–130. American Mathematical Society, Providence (1999)
14. Logan, B.F., Shepp, L.A.: A variational problem for random Young tableaux. *Adv. Math.* **26**, 206–222 (1977)
15. Luke, Y.L.: *The special functions and their approximations*. Academic Press, Cambridge (1969)
16. Matsumoto, S.: Polynomiality of shifted Plancherel averages and content evaluations. *Ann. Math. Blaise Pascal* **24**, 55–82 (2017)
17. Matsumoto, S., Śniady, P.: Random strict partitions and random shifted tableaux. *Sel. Math. New Ser.* **26**, 10 (2020)
18. Mutafchiev, L.R.: On the size of the Durfee square of a random integer partition. *J. Comput. Appl. Math.* **142**, 173–184 (2002)
19. Okounkov, A.: Random matrices and random permutations. *Int. Math. Res. Not.* **20**, 1043–1095 (2000)
20. Olshanski, G.: Plancherel averages: remarks on a paper by Stanley. *Electron. J. Combin.* **17**, 43 (2010)
21. Panova, G.: Polynomiality of some hook-length statistics. *Ramanujan J.* **27**, 349–356 (2012)
22. Romik, D.: The number of steps in the Robinson-Schensted algorithm. *Funct. Anal. Appl.* **39**, 152–155 (2005)
23. Romik, D.: *The surprising mathematics of longest increasing subsequences*. Cambridge University Press, New York (2015)
24. Sagan, B.E.: *The symmetric group*. Springer, New York (2001)
25. Salvy, B., Zimmermann, P.: Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Trans. Math. Softw.* **20**(2), 163–177 (1994)
26. Stanley, R.P.: Some combinatorial properties of hook lengths, contents, and parts of partitions. *Ramanujan J.* **23**, 91–105 (2010)
27. Vershik, A.M., Kerov, S.V.: Asymptotics of the Plancherel measure of the symmetric group and the limiting shape of Young tableaux. *Soviet Math. Dokl.* **18**, 527–531 (1977)

28. Vershik, A.M., Kerov, S.V.: Asymptotic behavior of the maximum and generic dimensions of irreducible representations of the symmetric group. *Funktional. Anal. i Prilozhen.* **19**, 25–36 (1985)
29. Vershik, A., Pavlov, D.: Numerical experiments in problems of asymptotic representation theory. *J. Math. Sci.* **168**, 351–361 (2010)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.