

# DISTRIBUTIONS OF HOOK LENGTHS IN INTEGER PARTITIONS

MICHAEL GRIFFIN, KEN ONO, AND WEI-LUN TSAI

*In memory of Christine Bessenrodt*

ABSTRACT. Motivated by the many roles that hook lengths play in mathematics, we study the distribution of the number of  $t$ -hooks in the partitions of  $n$ . We prove that the limiting distribution is normal with mean  $\mu_t(n) \sim \frac{\sqrt{6n}}{\pi} - \frac{t}{2}$  and variance  $\sigma_t^2(n) \sim \frac{(\pi^2-6)\sqrt{6n}}{2\pi^3}$ . Furthermore, we prove that the distribution of the number of hook lengths that are multiples of a fixed  $t \geq 4$  in partitions of  $n$  converge to a shifted Gamma distribution with parameter  $k = (t-1)/2$  and scale  $\theta = \sqrt{2/(t-1)}$ .

## 1. INTRODUCTION AND STATEMENT OF RESULTS

The study of the statistical properties of partitions and their Young diagrams is rich with deep results. Works by Pittel [12], Szalay and Turán [14, 15, 16], Temperley [17], and Vershik [18] form a large body of work on questions related to the expected “limiting shapes” of Young diagrams (see the more recent paper by Bogachev [2] for more recent results in the field). In this paper we study the statistical properties of the hook lengths in Young diagrams of integer partitions. In this regard, there is recent work by Mutafchiev [10] concerning the expected hook length of a randomly chosen cell. Here we study a different aspect. To make this precise we first recall some notation.

A *partition*  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$  of  $n$ , denoted  $\lambda \vdash n$ , is a nonincreasing sequence of positive integers that sum to  $n$ . Its *Young diagram* is the left-justified array of boxes where the row lengths are the parts. The *hook*  $H(k, j)$  of the cell in position  $(k, j)$  is the set of cells below or to the right of that cell, including the cell itself, and the *hook length*  $h(k, j) := (\lambda_k - k) + (\lambda'_j - j) + 1$ , is the number of cells in the hook  $H(k, j)$ . Here  $\lambda'_j$  is the number of boxes in column  $j$ , which is the same as the number of parts of the partition that are at least  $j$ .

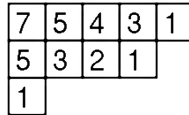


FIGURE 1. Hook lengths for  $\lambda = (5, 4, 1)$

Multisets  $\mathcal{H}(\lambda)$  of partition hook lengths have many roles in combinatorics, number theory, and representation theory (e.g. [1, 8, 13]). For instance, a standard Young tableaux for a partition  $\lambda$  of  $n$  is obtained by writing the numbers 1 through  $n$  in the boxes of the Young diagram so that each column and each row forms an increasing sequence. The Frame-Robinson-Thrall hook length formula

$$d_\lambda = \frac{n!}{\prod_{h \in \mathcal{H}(\lambda)} h}$$

gives the number of standard Young tableaux for  $\lambda$ . This is also the degree of the canonical irreducible representation of the symmetric group  $S_n$  associated to  $\lambda$ . As another important example, we have

---

2020 *Mathematics Subject Classification.* 11P82, 05A17.

*Key words and phrases.* Primary: Partitions, Secondary: Hook lengths.

K.O. thanks the Thomas Jefferson Fund and the NSF (DMS-2002265 and DMS-2055118) for their support.

the famous Nekrasov-Okounkov identity (see (6.12) of [11])<sup>1</sup>

$$(1.1) \quad \sum_{\lambda} q^{|\lambda|} \prod_{h \in \mathcal{H}(\lambda)} \left(1 - \frac{z}{h^2}\right) = \prod_{n=1}^{\infty} (1 - q^n)^{z-1},$$

which arises in combinatorics, mathematical physics and the theory of modular forms.

In this paper, we study the numbers  $Y_t(n)$  which count the  $t$ -hooks (i.e. hooks of length  $t$ ) among all partitions of  $n$ . For fixed  $t$ , we derive the limiting behavior of the sequence  $\{Y_t(n)\}$  for  $n \in \mathbb{N}$ , and we give asymptotics for the accumulation function

$$(1.2) \quad D_t(k; n) := \frac{\#\{\lambda \vdash n \text{ with } \leq k \text{ many hook lengths of size } t\}}{p(n)}.$$

**Theorem 1.1.** *If  $t$  is a fixed positive integer, then the following are true for the sequence  $\{Y_t(n)\}$ .*

- (1) *The sequence is asymptotically normal with mean  $\mu_t(n) \sim \frac{\sqrt{6n}}{\pi} - \frac{t}{2}$  and variance  $\sigma_t^2(n) \sim \frac{(\pi^2-6)\sqrt{6n}}{2\pi^3}$ .*
- (2) *If we let  $k_{t,n}(x) := \mu_t(n) + \sigma_t(n)x$ , then in terms of Gauss's error function  $E(x)$  we have*

$$\lim_{n \rightarrow +\infty} D_t(k_{t,n}(x); n) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy =: E(x).$$

**Remark.** The  $t = 1$  case of Theorem 1.1 recovers a result by Brennan, Knopfmacher and Wagner [3] on the distribution of ascents in partitions, as this number equals the number of size 1-hooks.

**Example.** Theorem 1.1 asserts that the limiting distribution of 2-hooks is a normal distribution with mean  $\mu_2(n) \sim \frac{\sqrt{6n}}{\pi} - 1$  and variance  $\sigma_2^2(n) \sim \frac{(\pi^2-6)\sqrt{6n}}{2\pi^3}$ . For  $n = 5000$ , we find that

$$\sum_{\lambda \vdash 5000} T^{\#\{2 \in \mathcal{H}(\lambda)\}} = 704T + 9211712T^2 + \dots + 1805943379138T^{98} + 2T^{99}.$$

Figure 2 plots  $Y_2(5000)$ .

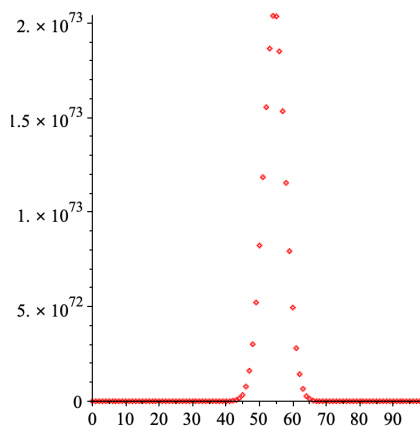


FIGURE 2.  $Y_2(5000)$

Table 1 illustrates the cumulative distribution approximation  $D_2(k_{2,5000}(x); 5000) \approx E(x)$ .

<sup>1</sup>This formula was also obtained by Westbury (see Proposition 6.1 and 6.2 of [19]).

$x$	$D_2(k_{2,5000}(x), 5000)$	$E(x)$	$D_2(k_{2,5000}(x), 5000)/E(x)$
-1.5	0.0658...	0.0668...	0.9849...
$\vdots$	$\vdots$	$\vdots$	$\vdots$
0.0	0.5055...	0.5000...	1.0011...
1.0	0.8246...	0.8413...	0.9802...
2.0	0.9685...	0.9772...	0.9911...

TABLE 1. Asymptotics for the cumulative distribution for  $n = 5000$

We next consider the sequence  $\{\widehat{Y}_t(n)\}$  of distributions of the number of hook lengths in  $t\mathbb{N}$  among the partitions of size  $n$ . This question is motivated by work of Han that extends (1.1) by giving infinite families of modular forms with level structure and cuspidal divisor. If  $\mathcal{H}_t(\lambda)$  is the multiset of hook lengths of  $\lambda$  that are in  $t\mathbb{N}$ , then he proved (see Theorem 1.3 of [7]) that

$$\sum_{\lambda} q^{|\lambda|} \prod_{h \in \mathcal{H}_t(\lambda)} \left( y - \frac{tyz}{h^2} \right) = \prod_{n=1}^{\infty} \frac{(1 - q^{tn})^t}{(1 - (yq^t)^n)^{t-z}(1 - q^n)}.$$

For  $t \geq 4$ , we prove that the limiting distribution is a shifted Gamma distribution with parameter  $k = (t - 1)/2$  and scale  $\theta = \sqrt{2/(t - 1)}$ , and we determine asymptotics for the cumulative distribution

$$(1.3) \quad \widehat{D}_t(k; n) := \frac{\#\{\lambda \vdash n \text{ with } \leq k \text{ many hook lengths in } t\mathbb{N}\}}{p(n)}.$$

Recall (e.g. II.2 of [5]) that a random variable  $X_{k,\theta}$  satisfies the *Gamma distribution with parameter  $k > 0$  and scale  $\theta > 0$*  if its probability distribution function is  $F_{k,\theta}(x) := \frac{1}{\Gamma(k)\theta^k} \cdot x^{k-1} e^{-\frac{x}{\theta}}$ .

**Theorem 1.2.** *If  $t \geq 4$ , then the following are true for the sequence  $\{\widehat{Y}_t(n)\}$ .*

(1) *The sequence satisfies*

$$\widehat{Y}_t(n) \sim \frac{n}{t} - \frac{\sqrt{3(t-1)n}}{\pi t} \cdot X_{\frac{t-1}{2}, \sqrt{\frac{2}{t-1}}},$$

and has mean  $\widehat{\mu}_t(n) \sim \frac{n}{t} - \frac{(t-1)\sqrt{6n}}{2\pi t}$ , mode  $\widehat{m}_t(n) \sim \frac{n}{t} - \frac{(t-3)\sqrt{6n}}{2\pi t}$ , and variance  $\widehat{\sigma}_t^2(n) \sim \frac{3(t-1)n}{\pi^2 t^2}$ .

(2) *If we let  $\widehat{k}_{t,n}(x) := \widehat{\mu}_t(n) + \widehat{\sigma}_t(n)x$ , then in terms of the lower incomplete gamma function we have*

$$\lim_{n \rightarrow +\infty} \widehat{D}_t(\widehat{k}_{t,n}(x); n) = \frac{\gamma\left(\frac{t-1}{2}; \sqrt{\frac{t-1}{2}}x + \frac{t-1}{2}\right)}{\Gamma\left(\frac{t-1}{2}\right)}.$$

**Remark.** The proof of Theorem 1.2, which uses properties of Gamma distributions with  $k > 1$ , does not apply for  $t \in \{2, 3\}$  as  $(t - 1)/2 \leq 1$ . Indeed, the  $\{\widehat{Y}_2(n)\}$  and  $\{\widehat{Y}_3(n)\}$  do not even have continuous limiting distributions. The fact that 100% of  $n$  do not have a 2-core or 3-core partition [6] implies that these distributions are populated with many vanishing terms as illustrated by

$$\sum_{\lambda \vdash 19} T^{\#\mathcal{H}_2(\lambda)} = 300T^9 + 185T^8 + 5T^2.$$

**Example.** Theorem 1.2 gives  $\widehat{Y}_{11}(n) \sim \frac{n}{11} - \frac{\sqrt{30n}}{11\pi} \cdot X_{5, \sqrt{5}}$ , with mean  $\widehat{\mu}_{11}(n) \sim \frac{n}{11} - \frac{5\sqrt{6n}}{11\pi}$  and variance  $\widehat{\sigma}_{11}^2(n) \sim \frac{30n}{121\pi^2}$ . Figure 3 gives  $\widehat{Y}_{11}(1000)$ .

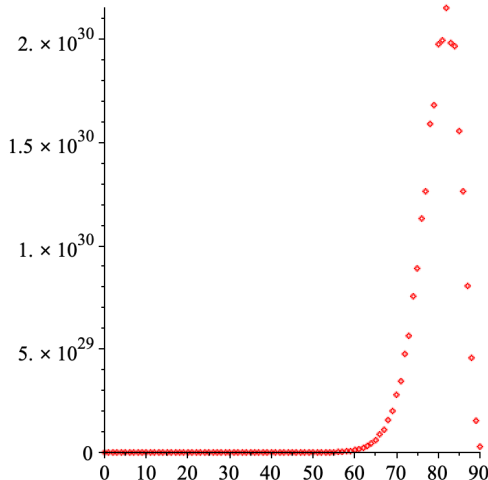
FIGURE 3.  $\widehat{Y}_{11}(1000)$ 

Table 2 illustrates the approximation  $\widehat{D}_{11}(k_{11,1000}(x); 1000) \approx \frac{\gamma(5; \sqrt{5x+5})}{24} =: \widehat{E}_{11}(x)$ .

$x$	$\widehat{D}_{11}(k_{11,1000}(x); 1000)$	$\widehat{E}_{11}(x)$	$\widehat{D}_{11}(k_{11,1000}(x); 1000)/\widehat{E}_{11}(x)$
-1.00	0.1319...	0.1467...	0.8993...
$\vdots$	$\vdots$	$\vdots$	$\vdots$
0.75	0.7410...	0.7954...	0.9315...
1.00	0.8226...	0.8474...	0.9707...
1.25	0.8872...	0.8880...	0.9991...

TABLE 2. Asymptotics for the cumulative distribution for  $n = 1000$ 

This paper is organized as follows. In Section 2 we recall work of Han that offers the relevant enumerative generating functions, and we then determine their asymptotics via the saddle point method, with assistance from the Euler-Maclaurin summation formula. In Section 3 we use these asymptotics to compute the moments of these statistics, which in turn imply Theorems 1.1 and 1.2 thanks to a classical theorem of Curtiss.

## ACKNOWLEDGEMENTS

The authors thank George Andrews, Kathrin Bringmann, Richard Stanley and Ole Warnaar for their valuable correspondence on this project. Finally, they thank the referees for their careful reading of the original submission and for their helpful suggestions.

## 2. NUTS AND BOLTS

We recall work of Han on the enumeration of hook lengths, and we derive important propositions (see Proposition 2.1 and 2.2) that are central to the proof of Theorems 1.1 and 1.2. Han obtained (see Thm. 1.4 and Cor. 5.1 of [7]) the following important generating functions for each fixed positive integer  $t$ :

$$(2.1) \quad G_t(T; q) = \sum_{n=0}^{\infty} P_t(n; T)q^n = \sum_{m,n} p_t(m; n)T^m q^n := \sum_{\lambda} q^{|\lambda|} T^{\#\{t \in \mathcal{H}(\lambda)\}} = \prod_{n=1}^{\infty} \frac{(1 + (T-1)q^{tn})^t}{1 - q^n},$$

$$(2.2) \quad \widehat{G}_t(T; q) = \sum_{n=0}^{\infty} \widehat{P}_t(n; T) q^n = \sum_{m, n} \widehat{p}_t(m; n) T^m q^n := \sum_{\lambda} q^{|\lambda|} T^{\#\mathcal{H}_t(\lambda)} = \prod_{n=1}^{\infty} \frac{(1 - q^{tn})^t}{(1 - (Tq^t)^n)^t (1 - q^n)}.$$

The next two propositions on  $P_t(n; T)$  and  $\widehat{P}_t(n; T)$  are the main results of this section.

**Proposition 2.1.** *Suppose that  $\eta \in (0, 1]$  and  $\eta \leq T \leq \eta^{-1}$ . If  $c(T) := \sqrt{\pi^2/6 - \text{Li}_2(1 - T)}$ , then*

$$P_t(n; T) = \frac{c(T)}{2\sqrt{2}\pi n T^{\frac{t}{2}}} \cdot e^{c(T)(2\sqrt{n} - \frac{1}{\sqrt{n}})} \cdot \left(1 + O_{\eta}(n^{-\frac{1}{7}})\right),$$

where  $\text{Li}_2(z) := -\int_0^z \frac{\log(1-u)}{u} du$  is the dilogarithm function.

The next proposition is more subtle, and pertains to suitable real sequences.

**Proposition 2.2.** *If  $t$  is a positive integer and  $T := \{T_n\}$  is a positive real sequence for which  $T_n = e^{\frac{\alpha(T) + \varepsilon_T(n)}{\sqrt{n}}}$ , where  $\alpha(T)$  is real and  $\varepsilon_T(n) = o_T(1)$ , then*

$$\begin{aligned} \widehat{P}_t(n; T_n) &= \\ & \frac{1}{2^{\frac{7}{4}} 3^{\frac{1}{4}} n} \cdot \sqrt{\frac{1}{\sqrt{6}} + \frac{\alpha(T) + \varepsilon_T(n)}{\pi t}} \left( \frac{\pi t}{\pi t + \sqrt{6}(\alpha(T) + \varepsilon_T(n))} \right)^{\frac{t}{2}} \cdot e^{\pi\sqrt{n}(\sqrt{\frac{2}{3}} + \frac{\alpha(T) + \varepsilon_T(n)}{\pi t})} \cdot (1 + O_T(n^{-\frac{1}{7}})). \end{aligned}$$

**2.1. Proof of Proposition 2.1.** The proof of Proposition 2.1 requires the next lemma.

**Lemma 2.3.** *If  $\eta \in (0, 1]$ , then for  $0 < \alpha < 1$  and  $\eta \leq T \leq \eta^{-1}$  we have*

$$(2.3) \quad \sum_{j=1}^{\infty} \log(1 - e^{-j\alpha}) = -\frac{\pi^2}{6\alpha} - \frac{1}{2} \log\left(\frac{\alpha}{2\pi}\right) + O(\alpha),$$

$$(2.4) \quad \sum_{n=1}^{\infty} \frac{t^2 n(T-1)}{T-1 + e^{tn\alpha}} = -\frac{\text{Li}_2(1-T)}{\alpha^2} + O_{\eta}(1),$$

$$(2.5) \quad \sum_{n=1}^{\infty} \log(1 + (T-1)e^{-tn\alpha}) = -\frac{\text{Li}_2(1-T)}{t\alpha} - \frac{1}{2} \log T + O_{\eta}(\alpha),$$

$$(2.6) \quad \sum_{n=1}^{\infty} \frac{t^3 n^2 e^{-tn\alpha}}{(1 + (T-1)e^{-tn\alpha})^2} = -\frac{2}{\alpha^3} \frac{\text{Li}_2(1-T)}{T-1} + O_{\eta}(\alpha).$$

*Proof.* For  $f \in C^{j+1}([a, b])$  and  $a, b \in \mathbb{Z}$ , Euler-Maclaurin summation (e.g. Thm. 2.1.9 of [9]) gives

$$\sum_{a < n \leq b} f(n) = \int_a^b f(x) dx + \sum_{r=0}^j \frac{(-1)^{r+1}}{(r+1)!} \left( f^{(r)}(b) - f^{(r)}(a) \right) B_{r+1} + \frac{(-1)^j}{(j+1)!} \int_a^b B_{j+1}(x - [x]) f^{(j+1)}(x) dx,$$

where  $B_r(x)$  is the  $r$ th Bernoulli polynomial and  $B_r := B_r(0)$ . Letting  $a = 0$  and  $j = 0$  gives

$$\begin{aligned} \sum_{n=1}^b \frac{t^2 n(T-1)}{T-1 + e^{tn\alpha}} &= \int_0^b \frac{t^2(T-1)x}{T-1 + e^{t\alpha x}} dx + \frac{t^2(T-1)b}{2(T-1 + e^{tab})} \\ & \quad + \int_0^b B_1(x - [x]) \frac{t^2(T-1)(T-1 + e^{t\alpha x}) - t^3\alpha(T-1)x e^{t\alpha x}}{(T-1 + e^{t\alpha x})^2} dx \\ &= \frac{1}{\alpha^2} \left[ \text{Li}_2((1-T)e^{-tab}) - \text{Li}_2(1-T) - t\alpha \log\left((T-1)e^{-tab} + 1\right) \right] + O_{\eta}\left(\frac{b^3}{e^{tab}}\right). \end{aligned}$$

To obtain (2.4), we let  $b \rightarrow \infty$  and find that

$$\sum_{n=1}^{\infty} \frac{t^2 n(T-1)}{T-1 + e^{tn\alpha}} = -\frac{\text{Li}_2(1-T)}{\alpha^2} + O_{\eta}(1).$$

Applying Euler-Maclaurin summation proves the other cases *mutatis mutandis*.  $\square$

*Proof of Proposition 2.1.* We first note that (2.1) implies that

$$(2.7) \quad P_t(n; T) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (ze^{ix})^{-n} G_t(T; ze^{ix}) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{gt(T; ze^{ix})} dx,$$

where  $g_t(T; w) := \text{Log}(w^{-n} G_t(T; w))$  for  $0 < |w| < 1$ . To apply the saddle point method, we must determine  $z = e^{-\alpha}$  for  $\alpha > 0$ , such that  $g'_t(T; z) = 0$  (throughout we consider the derivative with respect to the second parameter). By (2.1), this is equivalent to

$$\sum_{j=1}^{\infty} \frac{t^2 j(T-1)}{T-1 + e^{tj\alpha}} + \sum_{j=1}^{\infty} \frac{j}{e^{j\alpha} - 1} = n.$$

By combining (2.4) with

$$(2.8) \quad \sum_{j=1}^{\infty} \frac{j}{e^{j\alpha} - 1} = \frac{\pi^2}{6\alpha^2} - \frac{1}{2\alpha} + O(1),$$

which holds for  $0 < \alpha < 1$ , we find that

$$(2.9) \quad \alpha = c(T) \cdot n^{-\frac{1}{2}} - \frac{1}{4}n^{-1} + O_{\eta}(n^{-\frac{3}{2}}).$$

We now estimate  $g_t(T; z)$ ,  $g''_t(T; z)$ , and  $g'''_t(T; z)$ . Plugging  $z = e^{-\alpha}$  into  $g_t(T; z)$ , we obtain

$$g_t(T; z) = t \sum_{j=1}^{\infty} \log(1 + (T-1)e^{-tj\alpha}) - \sum_{j=1}^{\infty} \log(1 - e^{-j\alpha}) + n\alpha.$$

Therefore, (2.3), (2.5) and (2.9) gives

$$(2.10) \quad g_t(T; z) = 2c(T)\sqrt{n} + \frac{1}{2} \log\left(\frac{c(T)}{2\pi T\sqrt{n}}\right) + O_{\eta}\left(n^{-\frac{1}{2}}\right).$$

Similarly, by using (2.4) and (2.6) we obtain

$$(2.11) \quad \sum_{j=1}^{\infty} \frac{j^2 e^{-j\alpha}}{(1 - e^{-j\alpha})^2} = \frac{\pi^2}{3\alpha^3} - \frac{1}{2\alpha^2} + O(\alpha),$$

which implies that

$$(2.12) \quad \begin{aligned} g''_t(T; z) &= \left[ n + \sum_{j=1}^{\infty} \frac{t^3 j^2 e^{-tj\alpha}}{(1 + (T-1)e^{-tj\alpha})^2} - \sum_{j=1}^{\infty} \frac{t^2 j(T-1)}{T-1 + e^{tj\alpha}} + \sum_{j=1}^{\infty} \frac{j^2 e^{-j\alpha}}{(1 - e^{-j\alpha})^2} - \sum_{j=1}^{\infty} \frac{j}{e^{j\alpha} - 1} \right] e^{2\alpha} \\ &= e^{2c(T)n^{-\frac{1}{2}} + O_{\eta}(n^{-1})} \left( \frac{2}{c(T)} n^{\frac{3}{2}} + O_{\eta}(n) \right). \end{aligned}$$

By the same argument, thanks to (2.9), we find that

$$(2.13) \quad g'''_t(T; z) = O_{\eta} \left( \sum_{j=1}^{\infty} \frac{j^3 e^{-j\alpha}}{(1 - e^{j\alpha})^4} \right) = O_{\eta}(\alpha^{-4}) = O_{\eta}(n^2).$$

To complete the proof, we now let  $P_t(n; T) = I + II$ , where

$$I := \frac{1}{2\pi} \int_{|x| \leq n^{-5/7}} e^{g_t(T; ze^{ix})} dx \quad \text{and} \quad II := \frac{1}{2\pi} \int_{|x| > n^{-5/7}} e^{g_t(T; ze^{ix})} dx.$$

To estimate  $I$ , we use the Taylor expansion of  $g_t(T; w)$  centered at the saddle point  $z = e^{-\alpha}$

$$g_t(T; w) = g_t(T; z) + \frac{g_t''(T; z)(w - z)^2}{2} + O_\eta(g_t'''(T; z))(w - z)^3.$$

Since  $|x| \leq n^{-5/7}$ , estimate (2.9) gives

$$w - z = ze^{ix} - z = e^{-\alpha} (ix + O(x^2)) = \left(1 + O_\eta\left(n^{-\frac{1}{2}}\right)\right) \left(ix + O\left(n^{-\frac{10}{7}}\right)\right) = ix + O_\eta\left(n^{-\frac{17}{14}}\right).$$

Therefore, we obtain

$$(2.14) \quad g_t(T; w) = g_t(T; z) - \frac{g_t''(T; z)(x)^2}{2} + O_\eta\left(n^{-\frac{1}{7}}\right).$$

Combining (2.10), (2.12), (2.13), and (2.14), we obtain the main term asymptotic

$$(2.15) \quad \begin{aligned} I &= \frac{e^{g_t(T; z)}}{2\pi} \left[ \int_{-\infty}^{\infty} e^{-\frac{g_t''(T; z)x^2}{2}} dx - \int_{|x| > n^{-5/7}} e^{-\frac{g_t''(T; z)x^2}{2}} dx \right] \cdot \left(1 + O_\eta\left(n^{-\frac{1}{7}}\right)\right) \\ &= \frac{c(T)}{2\sqrt{2\pi n} T^{\frac{1}{2}}} \cdot e^{c(T)\left(2\sqrt{n} - \frac{1}{\sqrt{n}}\right)} \cdot \left(1 + O_\eta\left(n^{-\frac{1}{7}}\right)\right). \end{aligned}$$

To estimate the tail error term  $II$ , we estimate  $\frac{G_t(T; ze^{ix})}{G_t(T; z)}$  using

$$e^{g_t(T; ze^{ix})} = e^{g_t(T; z)} \frac{G_t(T; ze^{ix})}{G_t(T; z)}.$$

Since  $T > 0$ , letting  $w = ze^{ix}$  gives

$$(2.16) \quad \begin{aligned} \left| \frac{G_t(T; w)}{G_t(T; z)} \right|^2 &\leq \prod_{j=1}^{\infty} \text{Max} \left\{ 1, \left| \frac{1 + (T-1)w^j}{1 + (T-1)z^j} \right|^2 \right\} \left| \frac{1 - z^j}{1 - w^j} \right|^2 \\ &\leq \prod_{j=1}^{\infty} \text{Max} \left\{ 1, \left( 1 + \frac{2z^j(1-T)(1 - \cos(xj))}{(1 - z^j)^2} \right) \right\} \left( 1 + \frac{2z^j(1 - \cos(xj))}{(1 - z^j)^2} \right)^{-1} \\ &\leq \prod_{\sqrt{n} \leq j \leq 2\sqrt{n}} \text{Max} \left\{ 1, \left( 1 + \frac{2z^j(1-T)(1 - \cos(xj))}{(1 - z^j)^2} \right) \right\} \left( 1 + \frac{2z^j(1 - \cos(xj))}{(1 - z^j)^2} \right)^{-1}. \end{aligned}$$

To reduce to the finite product in the last line, we used the fact that for all  $j \geq 1$  we have

$$\text{Max} \left\{ 1, \left( 1 + \frac{2z^j(1-T)(1 - \cos(xj))}{(1 - z^j)^2} \right) \right\} \left( 1 + \frac{2z^j(1 - \cos(xj))}{(1 - z^j)^2} \right)^{-1} \leq 1.$$

We consider two cases (i.e.  $T > 1$  and  $T \leq 1$ ) to estimate (2.16). If  $T > 1$  and  $j \in [\sqrt{n}, 2\sqrt{n}]$ , then by (2.9) we have  $2z^j/(1 - z^j)^2 \leq c_\eta$ , for some  $c_\eta > 0$ . This implies that

$$(2.17) \quad \left| \frac{G_t(T; w)}{G_t(T; z)} \right|^2 \leq \prod_{\sqrt{n} \leq j \leq 2\sqrt{n}} (1 + c_\eta(1 - \cos(xj)))^{-1}.$$

If  $T \leq 1$ , then we have

$$\text{Max} \left\{ 1, \left( 1 + \frac{2z^j(1-T)(1 - \cos(xj))}{(1 - z^j)^2} \right) \right\} = 1 + \frac{2z^j(1-T)(1 - \cos(xj))}{(1 - z^j)^2}.$$

Moreover, we have  $2z^j/(1-z^j)^2 \leq c_\eta$ . A similar calculation also shows that (2.17) still holds for  $T \leq 1$  by choosing a suitable  $c_\eta > 0$ .

We divide the range of  $x$  into two cases  $n^{-5/7} \leq |x| \leq \frac{\pi}{2\sqrt{n}}$ , and  $\frac{\pi}{2\sqrt{n}} \leq |x| \leq \pi$ . For the first case, we can use the inequality  $1 - \cos(xj) \geq \frac{2}{\pi^2}(xj)^2$  to estimate (2.17), giving

$$(2.18) \quad \left| \frac{G_t(T; w)}{G_t(T; z)} \right|^2 \leq \prod_{\sqrt{n} \leq j \leq 2\sqrt{n}} \left( 1 + \frac{2c_\eta}{\pi^2}(xj)^2 \right)^{-1} \ll e^{-c_\eta(x^2 n^{\frac{3}{2}})} \ll e^{-c_\eta n^{\frac{1}{4}}}.$$

In the case where  $\frac{\pi}{2\sqrt{n}} \leq |x| \leq \pi$ , we count the  $j \in [\sqrt{n}, 2\sqrt{n}]$  for which there is an  $\ell \in \mathbb{Z}$  with  $-n^{-\frac{1}{12}} + 2\ell\pi \leq xj \leq n^{-\frac{1}{12}} + 2\ell\pi$ . The total number of such  $j$  is  $\gg n^{1/2} + O(n^{5/12})$ . Hence, we have

$$(2.19) \quad \left| \frac{G_t(T; w)}{G_t(T; z)} \right|^2 \leq \left( 1 + c_\eta(1 - \cos(n^{-\frac{1}{12}})) \right)^{-(n^{\frac{1}{2}} + O(n^{\frac{5}{12}}))} \ll e^{-c_\eta n^{\frac{1}{4}}}.$$

By combining (2.18) and (2.19), we obtain the upper bound for the tail

$$II \ll \frac{1}{2\pi} \int_{|x| > n^{-5/7}} e^{g_t(T; z)} \left| \frac{G_t(T; ze^{ix})}{G_t(T; z)} \right| dx \ll_\eta e^{-\frac{c(T)}{\sqrt{n}} - \frac{c_\eta}{2} n^{\frac{1}{4}}}.$$

As  $P_t(n; T) = I + II$ , the proposition follows from this inequality and (2.15).  $\square$

**2.2. Proof of Proposition 2.2.** For each positive integer  $n$ , (2.2) implies that

$$(2.20) \quad \widehat{P}_t(n; T_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (ze^{ix})^{-n} \widehat{G}_t(T_n; ze^{ix}) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\widehat{g}_t(T_n; ze^{ix})} dx,$$

where  $\widehat{g}_t(T_n; w) := \text{Log}(w^{-n} \widehat{G}_t(T_n; w))$  for  $0 < |w| < 1$ . We aim to locate the saddle point  $z = e^{-\beta_n}$ , with  $\beta_n > 0$ . To this end, we solve

$$-\sum_{j=1}^{\infty} \frac{t^2 j}{e^{tj\beta_n} - 1} + \sum_{j=1}^{\infty} \frac{t^2 j T_n^j}{e^{tj\beta_n} - T_n^j} + \sum_{j=1}^{\infty} \frac{j}{e^{j\beta_n} - 1} = n.$$

By (2.8) and the definition of  $\alpha(T)$  and  $\varepsilon_T(n)$ , we obtain we find that

$$(2.21) \quad \beta_n = \left( \frac{\pi}{\sqrt{6}} + \frac{\alpha(T) + \varepsilon_T(n)}{t} \right) \cdot n^{-\frac{1}{2}} + O_T(n^{-1}).$$

Since we have  $\varepsilon_T(n) = o_T(1)$ , it follows that

$$(2.22) \quad \beta_n = \left( \frac{\pi}{\sqrt{6}} + \frac{\alpha(T)}{t} \right) \cdot n^{-\frac{1}{2}} + o_T(n^{-\frac{1}{2}}).$$

We now estimate  $\widehat{g}_t(T_n; z)$ ,  $\widehat{g}_t''(T_n; z)$ , and  $\widehat{g}_t'''(T_n; z)$ . Plugging  $z = e^{-\beta_n}$  into  $\widehat{g}_t(T_n; z)$ , we obtain

$$\widehat{g}_t(T_n; z) = t \sum_{j=1}^{\infty} \log(1 - e^{-tj\beta_n}) - t \sum_{j=1}^{\infty} \log(1 - T_n^j e^{-tj\beta_n}) - \sum_{j=1}^{\infty} \log(1 - e^{-j\beta_n}) + n\beta_n.$$

Applying (2.3) to all three terms gives

$$(2.23) \quad \widehat{g}_t(T_n; z) = \frac{t\pi^2}{6(t\beta_n - \log T_n)} + \frac{1}{2} \log\left(\frac{\beta_n}{2\pi}\right) + \frac{t}{2} \log\left(\frac{t\beta_n - \log T_n}{t\beta_n}\right) + n\beta_n + O_T(\beta_n).$$



Similarly, by using (2.8) and (2.11) we obtain

$$(2.24) \quad \widehat{g}_t''(T_n; z) = \left[ n + \frac{\pi^2 t^3}{3(\beta_n t - \log T_n)^3} + \frac{t-1}{2\beta_n^2} - \frac{(t^3 + \pi^2 t^2)}{2(\beta_n t - \log T_n)^2} + \frac{1-t}{2\beta_n} + \frac{t^2}{2(\beta_n t - \log T_n)} + O_T(\beta_n) \right] e^{2\beta_n}.$$

By the same argument, thanks to (2.22), we find that

$$(2.25) \quad \widehat{g}_t'''(T_n; z) = O_T((\beta_n t - \log T_n)^{-4}) = O_T(\beta_n^{-4}).$$

Arguing as in the proof of Proposition 2.1 with (2.22), (2.23), (2.24), and (2.25), we obtain

$$\begin{aligned} \widehat{P}_t(n; T_n) &= \frac{e^{\widehat{g}_t(T_n; z)}}{2\pi} \cdot \int_{-\infty}^{\infty} e^{-\frac{\widehat{g}_t''(T_n; z)x^2}{2}} dx \cdot (1 + O_T(n^{-\frac{1}{7}})) = \frac{e^{\widehat{g}_t(T_n; z)}}{2\pi} \cdot \sqrt{\frac{2\pi}{|\widehat{g}_t''(T_n; z)|}} \cdot \left(1 + O_T\left(n^{-\frac{1}{7}}\right)\right) \\ &= \frac{1}{2^{\frac{7}{4}} 3^{\frac{1}{4}} n} \cdot \sqrt{\frac{1}{\sqrt{6}} + \frac{\gamma(T) + \alpha_T(n)}{\pi t}} \left( \frac{\pi t}{\pi t + \sqrt{6}(\alpha(T) + \varepsilon_T(n))} \right)^{\frac{t}{2}} \cdot e^{\pi\sqrt{n}\left(\sqrt{\frac{2}{3}} + \frac{\alpha(T) + \varepsilon_T(n)}{\pi t}\right)} \cdot \left(1 + O_T\left(n^{-\frac{1}{7}}\right)\right). \end{aligned}$$

This completes the proof of the proposition.

### 3. PROOFS OF THEOREMS 1.1 AND 1.2

We prove Theorems 1.1 and 1.2 using the method of moments, where the crucial device is the following classical theorem of Curtiss.

**Theorem 3.1** (Theorem 2 of [4]). *Let  $\{X_n\}$  be a sequence of real random variables. Then define the corresponding moment generating function*

$$M_{X_n}(r) := \int_{-\infty}^{\infty} e^{rx} dF_n(x),$$

where  $F_n(x)$  is the cumulative distribution function associated with  $X_n$ . If the sequence  $\{M_{X_n}(r)\}$  converges pointwise on a neighborhood of  $r = 0$ , then  $\{X_n\}$  converges in distribution.

*Proof of Theorem 1.1.* For each  $n \geq 1$ , we consider the  $r$ th power moment

$$(3.1) \quad M(Y_t(n); r) := \frac{1}{p(n)} \sum_{m=0}^{\infty} p_t(m; n) \cdot e^{\frac{(m - \mu_t(n))r}{\sigma_t(n)}}.$$

By Curtiss's Theorem, combined with the theory of normal distributions, it suffices to show that

$$(3.2) \quad \lim_{n \rightarrow +\infty} M(Y_t(n); r) = e^{\frac{r^2}{2}}.$$

By evaluating  $P_t(n; T)$  at  $T = 1$  (i.e.  $P_t(n; 1) = p(n)$ ) and  $e^{\frac{r}{\sigma_t(n)}}$ , we have

$$M(Y_t(n); r) = \frac{P_t(n; e^{\frac{r}{\sigma_t(n)}})}{p(n)} \cdot e^{-\frac{\mu_t(n)}{\sigma_t(n)} r}.$$

Proposition 2.1 gives

$$(3.3) \quad M(Y_t(n); r) = \frac{c(e^{\frac{r}{\sigma_t(n)}}) \cdot \left(1 + O_n\left(n^{-\frac{1}{7}}\right)\right)}{c(1) \cdot \left(1 + O\left(n^{-\frac{1}{7}}\right)\right)} \cdot e^{-\frac{t}{2\sigma_t(n)} r - \frac{\mu_t(n)}{\sigma_t(n)} r + (2n^{\frac{1}{2}} - n^{-\frac{1}{2}}) \cdot (c(e^{\frac{r}{\sigma_t(n)}}) - c(1))}.$$

Since  $e^{\frac{r}{\sigma_t(n)}} > 0$  and  $e^{\frac{r}{\sigma_t(n)}} \rightarrow 1$ , as  $n \rightarrow \infty$ , the implied constant can be chosen to be independent of  $\eta$ . By direct calculation of the dilogarithm function, we find that  $c(1) = \pi/\sqrt{6}$ , and

$$c(e^{\frac{r}{\sigma_t(n)}}) = \frac{\pi}{\sqrt{6}} + \sqrt{\frac{3}{2}} \frac{1}{\pi} \left( \frac{r}{\sigma_t(n)} \right) + \sqrt{\frac{3}{2}} \frac{(\pi^2 - 6)}{4\pi^3} \left( \frac{r^2}{\sigma_t^2(n)} \right) + O\left( \frac{r^3}{\sigma_t^3(n)} \right).$$

Therefore, (3.3) becomes

$$\begin{aligned} M(Y_t(n); r) &= \left( 1 + O_r(n^{-\frac{1}{7}}) \right) \cdot e^{\left( -\frac{t}{2} - \mu_t(n) + \frac{\sqrt{6n}}{\pi} \right) \frac{r}{\sigma_t(n)} + \sqrt{6n} \left( \frac{\pi^2 - 6}{4\pi^3} \right) \left( \frac{r^2}{\sigma_t^2(n)} \right) + O_r(n^{-\frac{3}{4}})} \\ &= \left( 1 + O_r(n^{-\frac{1}{7}}) \right) \cdot e^{\frac{r^2}{2} + o_r(1)}. \end{aligned}$$

Letting  $n \rightarrow +\infty$ , we obtain (3.2) confirming Theorem 1.1.  $\square$

*Proof of Theorem 1.2.* To prove Theorem 1.2, we recall that if  $k > 1$  and  $r < 1/\theta$ , then the moment generating function for the random variable  $X_{k,\theta}$  is (for example, see II.2 of [5])

$$M(X_{k,\theta}; r) = \frac{1}{(1 - \theta r)^k}.$$

This distribution has mean  $\mu_{k,\theta} = k\theta$ , mode  $\text{mo}_{k,\theta} = (k-1)\theta$ , and variance  $\sigma_{k,\theta}^2 = k\theta^2$ . If  $a$  and  $b$  are real, then the shifted Gamma distribution  $aX_{k,\theta} + b$  has moment generating function

$$M(aX_{k,\theta} + b; r) = e^{br} \cdot M(X_{k,\theta}, ar) = \frac{e^{br}}{(1 - \theta ar)^k},$$

and has mean  $ak\theta + b$ , mode  $a(k-1)\theta + b$ , and variance  $a^2k\theta^2$ . We compare  $\widehat{Y}_t(n)$  with  $aX_{k,\theta} + b$ , where  $(k, \theta) := \left( \frac{t-1}{2}, \sqrt{\frac{2}{t-1}} \right)$ , and  $a := -1$  and  $b := \sqrt{2(t-1)}/2$ . Therefore, we assume that  $(t-1)/2 > 1$ , which is equivalent to  $t \geq 4$ .

To apply Curtiss's theorem, we compute the moment generating function as in (3.1), with the claimed mean  $\widehat{\mu}_t(n) \sim n/t - (t-1)\sqrt{6n}/2\pi t$ , and variance  $\widehat{\sigma}_t(n) \sim \sqrt{3(t-1)n}/\pi t$ . Applying Proposition 2.2 with  $\alpha(T) := \pi tr/\sqrt{3(t-1)}$  and  $T_n := e^{\frac{\alpha(T)}{\sqrt{n}}}$ , we find that

$$\begin{aligned} M(\widehat{Y}_t(n); r) &= \frac{(2^{\frac{7}{4}} 3^{\frac{1}{4}} n)^{-1} \cdot \sqrt{\frac{1}{\sqrt{6}} + \frac{r}{\sqrt{3(t-1)}}} \cdot \left( 1 + \sqrt{\frac{2}{t-1}} r \right)^{-\frac{t}{2}} \cdot \left( 1 + O_r(n^{-\frac{1}{7}}) \right)}{(4\sqrt{3}n)^{-1} \cdot \left( 1 + O(n^{-\frac{1}{7}}) \right)} \cdot e^{\frac{n}{i\widehat{\sigma}_t(n)} r - \frac{\widehat{\mu}_t(n)}{\widehat{\sigma}_t(n)} r} \\ &= \frac{e^{\frac{\sqrt{2(t-1)}}{2} r}}{\left( 1 + \sqrt{\frac{2}{t-1}} r \right)^{\frac{t-1}{2}}} \cdot \left( 1 + O_r(n^{-\frac{1}{7}}) \right). \end{aligned}$$

Therefore, Curtiss's theorem gives  $\widehat{Y}_t(n) \sim a\widehat{\sigma}_t(n)X_{k,\theta} + b\widehat{\sigma}_t(n) + \widehat{\mu}_t(n)$ , as well as the claimed mean, mode and variance. To obtain claim (2), we recall that if  $k > 1$ , then the Gamma distribution  $X_{k,\theta}$  has cumulative distribution function (e.g. II.2 of [5])  $D_{k,\theta}(x) = \gamma(k; \frac{x}{\theta})/\Gamma(k)$ , where  $\gamma(\alpha; x)$  is the lower incomplete Gamma function.  $\square$

## REFERENCES

- [1] G. E. Andrews and K. Eriksson, *Integer partitions*, Cambridge Univ. Press, 2nd ed., Cambridge, 2004.
- [2] L. V. Bogachev, *Unified derivation of the limit shape for multiplicative ensembles of random integer partitions with equiweighted parts*, Random Struct. Alg. **47** (2015), 227-266.
- [3] C. Brennan, A. Knopfmacher, and S. Wagner, *The distribution of ascents of size  $d$  or more in partitions of  $n$* , Combinatorics, Probability and Computing **17** (2008), 495-509.

- [4] J. Curtiss, *A note on the theory of moment generating functions*, Ann. Math. Statist. **13** (1942), 430-433.
- [5] W. Feller, *An introduction to probability theory and its applications*, Vol. 2, 2nd ed., John Wiley & Sons, Inc. (1971).
- [6] A. Granville and K. Ono, *Defect zero  $p$ -blocks for finite simple groups*, Trans. Amer. Math. Soc. **348** (1996), 331-347.
- [7] G.-N. Han, *The Nekrasov-Okounkov hook length formula: refinement, elementary proof, extension and applications*, Ann. Inst. Fourier (Grenoble) **60** (2010), 1–29.
- [8] G. James and A. Kerber, *Representation theory of the symmetric group*, Cambridge Univ. Press, 1984.
- [9] M. R. Murty, *Problems in Analytic Number Theory*, 2nd ed., Graduate Texts in Mathematics, Vol 206, Readings in Mathematics, Springer, New York, 2008.
- [10] L. Mutafchiev, *The limiting distribution of the hook length of a randomly chosen cell in a random Young diagram*, arXiv:1906.07169, (2019).
- [11] N. Nekrasov and A. Okounkov, *Seiberg-Witten theory and random partitions*, The unity of mathematics, Vol 244, Prog. Math., Birkhauser, Boston, 2006, 525–596.
- [12] B. Pittel, *On a likely shape of the random Ferrers diagram*, Adv. Appl. Math. **18** (1997), 432-488.
- [13] B. E. Sagan, *The symmetric group: Representations, combinatorial algorithms, and symmetric functions*, Springer-Verlag, New York, 2001.
- [14] M. Szalay and P. Turán, *On some problems of the statistical theory of partitions with applications to characters of the symmetric group, I*, Acta Math. Acad. Sci. Hungar. **29** (1977), 361-379.
- [15] M. Szalay and P. Turán, *On some problems of the statistical theory of partitions with applications to characters of the symmetric group, II*, Acta Math. Acad. Sci. Hungar. **29** (1977), 381-392.
- [16] M. Szalay and P. Turán, *On some problems of the statistical theory of partitions with applications to characters of the symmetric group, III*, Acta Math. Acad. Sci. Hungar. **32** (1978), 129-155.
- [17] H. N. V. Temperley, *Statistical mechanics and the partition numbers II. The form of crystal surfaces*, Math. Proc. Cambridge Philos. Soc. **48** (1952), 683-697.
- [18] A. M. Vershik, *Statistical mechanics of combinatorial partitions and their limit shapes*, Funct. Anal. Appl. **30** (1996), 90-105.
- [19] B. Westbury, *Universal characters from the Macdonald identities*, Adv. Math. **202** (2006), 50–63.

DEPARTMENT OF MATHEMATICS, BRIGHAM YOUNG UNIVERSITY, PROVO, UT 84602  
 Email address: mjgriffin@math.byu.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE, VA 22904  
 Email address: ko5wk@virginia.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE, VA 22904  
 Email address: wt8zj@virginia.edu