SIAM J. DISCRETE MATH.
Vol. 18, No. 2, pp. 418–435

ON THE MULTIPLICITY OF PARTS IN A RANDOM COMPOSITION OF A LARGE INTEGER*

PAWEŁ HITCZENKO† AND CARLA D. SAVAGE‡

Abstract. In this paper we study the following question posed by H. S. Wilf: what is, asymptotically as $n \to \infty$, the probability that a randomly chosen part size in a random composition of an integer n has multiplicity m? More specifically, given positive integers n and m, suppose that a composition λ of n is selected uniformly at random and then, out of the set of part sizes in λ, a part size j is chosen uniformly at random. Let $\mathbb{P}(A_n^{(m)})$ be the probability that j has multiplicity m. We show that for fixed m, $\mathbb{P}(A_n^{(m)})$ goes to 0 at the rate $1/\ln n$. A more careful analysis uncovers an unexpected result: $(\ln n)\mathbb{P}(A_n^{(m)})$ does not have a limit but instead oscillates around the value $1/m$ as $n \to \infty$.

This work is a counterpart of a recent paper of Corteel, Pittel, Savage, and Wilf, who studied the same problem in the case of partitions rather than compositions.

Key words. compositions of an integer, random compositions, geometric random variables

AMS subject classifications. 05A16, 60C05

DOI. 10.1137/S0895480199363155

1. Introduction.

In this paper we consider the multiplicity of a randomly chosen part size in a random composition of an integer n. Let us recall that a multiset $\lambda = \{\lambda_1, \dots, \lambda_k\}$ is a *partition* of an integer n if the λ_j are positive integers, called *parts*, such that $\sum \lambda_j = n$. *Compositions* are merely partitions in which the order of parts is significant. Thus, for example, the integer 3 admits three partitions, $\{1, 1, 1\}$, $\{2, 1\}$, and $\{3\}$, and four compositions, namely $(1, 1, 1)$, $(1, 2)$, $(2, 1)$, and (3).

Integer partitions (as deterministic objects) have been studied for quite some time, but Erdös and Lehner [6] were apparently the first to study integer partitions from the probabilistic perspective; namely, they considered the set of all partitions, $P(n)$, of an integer n, as a probability space equipped with the uniform probability measure. Quantities of interest are treated as random variables, and one can study their probabilistic properties, most typically the limiting properties as $n \to \infty$. Erdös and Lehner, for example, considered the limiting distribution of the total number of parts in a partition. Their paper opened a new line of investigation.

Goh and Schmutz [11] obtained the central limit theorem for the number of different part sizes in a random partition; that is, they proved that the number of different part sizes, appropriately normalized, has, approximately, the standard Gaussian distribution. (Several years earlier, Wilf [18] found an asymptotic formula for the expected number of distinct part sizes.) This approach culminated in an important paper by Fristedt [10], who proved that the joint distribution of the multiplicities of

*Received by the editors November 2, 1999; accepted for publication (in revised form) October 8, 2003; published electronically December 9, 2004. Research supported in part by National Science Foundation grant DMS9622772.

http://www.siam.org/journals/sidma/18-2/36315.html

†Department of Mathematics, Drexel University, Philadelphia, PA 19104 (phitczenko@mcs.drexel.edu).

‡Department of Computer Science, North Carolina State University, Raleigh, NC 27695-8206 (savage@cayley.csc.ncsu.edu).

part sizes is that of independent geometric random variables (Y_k), with parameters $(1-p^k)$, conditioned on the event $\{\sum kY_k = n\}$.

Fristedt's work, in turn, opened new possibilities and resulted in further progress in our understanding of the structure of random partitions. A good example is a paper of Pittel [17] substantiating two well-known conjectures concerning integer partitions. Utilizing Fristedt's result, Corteel, Pittel, Savage, and Wilf [4] quite recently provided an answer to the following question. Consider the following two-step sampling procedure: first choose uniformly at random a partition λ of n. Then, out of all different part sizes in λ pick one uniformly at random. What is the asymptotic *unconditional* probability that this part size has a certain specified multiplicity, say, m? For example, partition $\lambda = \{3,2,2,1,1,1\}$ of the number 10 has three different part sizes 1, 2, and 3, and only one of them has multiplicity three, namely 1. Thus, for this particular partition, the probability of choosing a part that has multiplicity three is 1/3. In order to find the unconditional probability of randomly choosing a part of multiplicity three in a randomly chosen partition of 10, one would have to average similar probabilities over all partitions of 10. Corteel, Pittel, Savage, and Wilf showed that in general the probability in question approaches $1/(m(m+1))$ (in particular, the probability that the randomly chosen part size in a random partition is unrepeated approaches 1/2 as $n \to \infty$).

Wilf then asked the same question for random compositions: what is the asymptotic value of the probability that a randomly chosen part size in a random composition of an integer n has multiplicity m? Our aim here is to provide an answer as complete as we can. On the "first level" of precision the answer is simple: for every fixed m this probability approaches zero. One would then like to know the rate of this convergence. We will show that the rate is $1/\ln n$. Specifically, if $A_n^{(m)}$ is the event that a randomly chosen part size in a random composition of n has multiplicity m, then there exist constants $c_1(m) \le c_2(m)$ such that $c_1(m) \le (\ln n)\mathbb{P}(A_n^{(m)}) \le c_2(m)$ for $n \ge 2$.

The next natural step is to find possibly tight bounds on $c_1(m)$ and $c_2(m)$, or to show that the limit $(\ln n)\mathbb{P}(A_n^{(m)})$ exists as $n \to \infty$. This is the place where things become a bit tricky. In order to describe the difficulties let us briefly discuss the argument. Letting $U_n^{(m)}$ and D_n denote the number of parts of multiplicity m and the number of distinct part sizes, respectively, we have $\mathbb{P}(A_n^{(m)}) = \mathbb{E}(U_n^{(m)}/D_n)$. In the case of partitions, Corteel, Pittel, Savage, and Wilf used Fristedt's result to argue that D_n is heavily concentrated around its expectation, and therefore, $\mathbb{P}(A_n^{(m)})$ is asymptotic to the ratio of expectations $\mathbb{E}U_n^{(m)}/\mathbb{E}D_n$ and one needs to find asymptotic values of these two expectations. In the case of compositions, much of the story is the same, with one crucial exception: the expected value of $U_n^{(m)}$ does not have a limit, but exhibits oscillations around $1/(m \ln 2)$. (This phenomenon is not new and was observed in the context of head runs in coin tossing; see, e.g., [3], [12], or [13].) Since, as we will show, the behavior of $(\ln n)\mathbb{P}(A_n^{(m)})$ is governed by the behavior of $\mathbb{E}U_n^{(m)}$, it will follow that $(\ln n)\mathbb{P}(A_n^{(m)})$ oscillates around the value $1/m$ as $n \to \infty$.

The rest of the paper is organized as follows: in the next section we will introduce notation and state our result precisely. In section 3 we will describe the probabilistic set-up. In section 4 we estimate the number of distinct part sizes and show that D_n is heavily concentrated about its expectation. In section 5, we give an estimate for the expected number of parts of given multiplicity. In section 6, we compute bounds on the oscillation.

2. Notation and statement of the result. A *composition* κ of an integer n is an ordered tuple $(\gamma_1,\dots,\gamma_k)$, where $\gamma_1,\dots,\gamma_k$ are positive integers such that $\sum_{i=1}^k \gamma_i = n$. The numbers $\gamma_1,\dots,\gamma_k$ are called *parts*, k is the total number of parts, and the elements of the set $\{\gamma_1,\dots,\gamma_k\}$ are the *part sizes* of κ. For example, $(2,1,2,3,1,1)$ is a composition of the number 10 into six parts with part sizes 1, 2, and 3, where part size 1 has multiplicity 3, 2 has multiplicity 2, and 3 has multiplicity 1. We denote the set of all compositions of n by $C(n)$ and note that $|C(n)| = 2^{n-1}$. For a composition $\kappa = (\gamma_1,\dots,\gamma_k)$ we let $D_n(\kappa)$ denote the number of distinct part sizes and, for fixed integer m, $U_n^{(m)}(\kappa)$ will denote the number of part sizes of κ that have multiplicity m. More formally,

$$D_n(\kappa) = 1 + \sum_{i=2}^{k} I_{\{\gamma_i \neq \gamma_j,\ j=1,\dots,i-1\}},$$

where I_A, the indicator of event A, is 1 if A takes place and 0, otherwise. Similarly,

$$U_n^{(m)}(\kappa) = \sum_{i=1}^{k} I_{B_i},$$

where

$$B_i = \{\gamma_i \neq \gamma_j,\ j < i \text{ and } \operatorname{card}\{\ell > i:\ \gamma_\ell = \gamma_i\} = m-1\}.$$

We equip $C(n)$ with the uniform probability measure $\mathbb{P}$ (i.e., $\mathbb{P}(\kappa) = |C(n)|^{-1} = 2^{-n+1}$ for every $\kappa \in C(n)$), and we will denote the expectation with respect to that measure by $\mathbb{E}$.

Throughout the paper the letter c is reserved for an absolute constant whose value is of no relevance and may change from line to line.

We consider the following experiment. First, a composition is chosen at random. Then, out of all distinct part sizes one is selected uniformly at random. We would like to know what the unconditional probability is that this part size has multiplicity m. We will denote this event by $A_n^{(m)}$. Since for a given composition κ the probability that a randomly chosen part size has multiplicity m is given by the ratio

$$\frac{U_n^{(m)}(\kappa)}{D_n(\kappa)},$$

the unconditional probability that a randomly chosen part size in a random composition has this multiplicity is just the expected value of that ratio. That is,

$$\mathbb{P}(A_n^{(m)}) = \mathbb{E}\frac{U_n^{(m)}}{D_n}.$$

Thus, our goal is to approximate this expectation. Our result is as follows.

THEOREM 1. *Under the above notation we have the following: for a fixed integer m*

$$(\ln n)\mathbb{P}(A_n^{(m)}) = \Theta(1),$$

i.e., there exist two positive constants $c_1(m)$ and $c_2(m)$ such that for all $n \geq 2$,

$$c_1(m) \leq (\ln n)\mathbb{P}(A_n^{(m)}) \leq c_2(m).$$

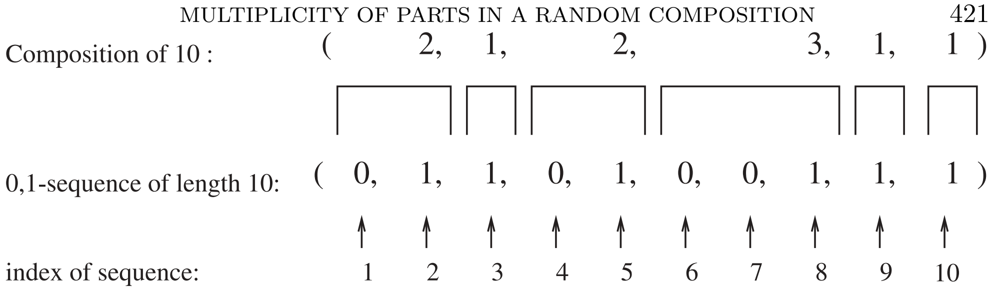

FIG. 1. *Correspondence between compositions of n and $0,1$-sequences of length n which end with* 1.

More precisely, as $n \to \infty$,

$$(\ln n)\mathbb{P}(A_n^{(m)}) = \frac{1}{m} + H^{(m)}(c \ln n) + o(1),$$

where $H^{(m)}$ is a mean-zero function of period 1 *whose Fourier coefficients are given by*

$$\phi_\ell = \frac{1}{m!}\Gamma\left(m - \frac{2\ell\pi i}{\ln 2}\right), \qquad \ell \neq 0.$$

3. Probabilistic set-up. Much of our proof relies on an appropriate interpretation of a composition, found, e.g., in Andrews [2]. This interpretation allows us to connect the study of random compositions to another much investigated topic, namely the study of runs of successes in independent Bernoulli trials (see, for example, Erdös and Rényi [7] or Erdös and Révész [8]). In order to describe this connection we interpret compositions as follows: consider a composition $\kappa = (\gamma_1, \ldots, \gamma_k)$ of n into parts $\gamma_1, \gamma_2, \ldots, \gamma_k$ (for example, $(2,1,2,3,1,1)$ is a composition of the number 10 into six parts $2,1,2,3,1,1$.) Such a composition is associated with a $\{0,1\}$-valued sequence $(x_1, \ldots, x_n)$ in which $x_i = 1$ for $i \in \{\gamma_1, \gamma_1+\gamma_2, \ldots, \gamma_1+\cdots+\gamma_k\}$ and otherwise $x_i = 0$. (Note that this forces $x_n = 1$.) For example, the composition $(2,1,2,3,1,1)$ is associated with the sequence $(0,1,1,0,1,0,0,1,1,1)$, as illustrated in Figure 1. Clearly there is a one-to-one correspondence between compositions of n and $\{0,1\}$-sequences $(x_1, \ldots, x_n)$ with $x_n = 1$.

To say that a composition is chosen at random is to say that the 0's and 1's occur with probability $1/2$ at each of the first $n-1$ positions, and the occurrences at different positions are independent of each other. In other words, the number of 1's in the first $n-1$ positions is a binomial random variable, $\mathrm{Bin}(n-1, 1/2)$, with parameters $n-1$ and $1/2$. With this interpretation the total number of parts is just the number of 1's (including the one in the nth position) and thus it is equidistributed with $1+\mathrm{Bin}(n-1,1/2)$. (This contrasts with the case of "unordered" partitions where the exact distribution of the number of parts is unknown and it took a considerable effort to find a limiting distribution of the total number of parts; see Erdös and Lehner [6].) Furthermore, the numbers $\gamma_1, \ldots, \gamma_k$ can be viewed as "waiting times" for the first, second, ... , and kth appearance of 1 in the associated $\{0,1\}$-sequence $(x_1, \ldots, x_n)$. (In our example, 1 appears in the second, third, fifth, eighth, ninth, and, of course, tenth positions.) It is well known and easy to check that in an infinite sequence of independent Bernoulli trials with the probability of success p, waiting times for successes are independent and identically distributed (i.i.d.) random variables whose common distribution is that of a geometric random variable with parameter p. Since we are considering only $n-1$ trials, this is no longer true. But we have the following fact.

PROPOSITION 2. *Let* $\Gamma_1, \Gamma_2 \dots$ *be i.i.d. geometric random variables with parameter* $1/2$ *(that is,* $\mathbb{P}(\Gamma_1 = j) = 2^{-j}$, $j = 1, 2 \dots$*) and define*

$$\tau = \inf\{k \geq 1 : \; \Gamma_1 + \Gamma_2 + \cdots + \Gamma_k \geq n\}.$$

Then we have the following: if the set $C(n)$ *of all compositions of an integer* n *is equipped with the uniform probability measure, then the distribution of a randomly chosen composition is given by*

$$\left(\Gamma_1, \Gamma_2, \dots, \Gamma_{\tau-1}, n - \sum_{j=1}^{\tau-1} \Gamma_j\right).$$

4. The number of distinct parts. In this section we will study certain aspects of the behavior of D_n. For the purpose of approximating $\mathbb{P}(A_n^{(m)})$ we will work with the ratio $U_n^{(m)}/D_n$, but it will be clear from our argument, for example, that

$$\frac{\mathbb{E}D_n}{\log_2 n} \to 1 \quad \text{as} \quad n \to \infty.$$

We will proceed in the following fashion: we will establish the existence of two sequences of natural numbers (ℓ_n) and (k_n) which increase to infinity and are asymptotically the same, i.e.,

$$\lim_{n\to\infty} \frac{\ell_n}{k_n} = 1,$$

and such that both probabilities

$$\mathbb{P}(D_n \leq \ell_n), \qquad \mathbb{P}(D_n \geq k_n)$$

tend to zero as $n \to \infty$ at a rate faster than $1/\log_2 n$. This will allow us to replace the D_n in the denominator by either of the sequences (ℓ_n) or (k_n), and then in the next section, we will approximate the expected value of $U_n^{(m)}$. We begin with establishing the existence of (ℓ_n). For a composition $\kappa = (\gamma_1, \dots, \gamma_k)$, let $S_n(\kappa)$ denote the number of consecutive part sizes (starting with size 1) in κ. That is,

$$S_n(\kappa) = \max\{\ell : \; \forall j \leq \ell \; \exists i \leq k : \gamma_i = j\}.$$

Consider for a moment an arbitrary integer ℓ_n. Since $S_n(\kappa) \leq D_n(\kappa)$ we have

$$\mathbb{P}(D_n \leq \ell_n) \leq \mathbb{P}(S_n \leq \ell_n) \leq \mathbb{P}(\exists j \leq \ell_n, \; \forall i < \tau, \; : \; \Gamma_i \neq j), \tag{1}$$

where we purposely ignored the last part $n - \sum_{j=1}^{\tau-1} \Gamma_i$ by writing "$i < \tau$." In order to bound the last probability we first notice that, since τ is equidistributed with the random variable $1 + \text{Bin}(n-1, 1/2)$, we have

$$\mathbb{E}\tau = 1 + (n-1)/2 = (n+1)/2.$$

Moreover, τ is well concentrated around its mean. Namely (see, for example, [1, section A.1]), for every $t > 0$ we have

$$\mathbb{P}(|\tau - \mathbb{E}\tau| \geq t) \leq 2\exp\left\{-\frac{2t^2}{n-1}\right\}.$$

In particular, letting $t_n = \sqrt{\alpha(n-1)\ln n}$, we get

$$\mathbb{P}(|\tau - \mathbb{E}\tau| \geq t_n) \leq 2\exp\{-2\alpha \ln n\} = \frac{2}{n^{2\alpha}}. \tag{2}$$

(The value of α plays a minimal role in the argument, so we will set it to be 1 for the rest of this section; we just want to mention that by increasing this value as necessary we can get arbitrary polynomial rate of convergence to zero of this probability. This will be useful in the next section.) Let $q_n^- = \mathbb{E}\tau - t_n = (n+1)/2 - \sqrt{(n-1)\ln n}$. Then we can bound (1) by

$$\begin{aligned}\mathbb{P}(\exists j \leq \ell_n,\ \forall i < \tau,\ :\ \Gamma_i \neq j) &\leq \mathbb{P}(|\tau - \mathbb{E}\tau| > t_n)\\ &\quad + \mathbb{P}(\{\exists j \leq \ell_n,\ \forall i < \tau,\ :\ \Gamma_i \neq j\} \cap \{|\tau - \mathbb{E}\tau| \leq t_n\}).\end{aligned} \tag{3}$$

From (2), the first probability in the right-hand side (rhs) of (3) goes to 0 at a polynomial rate, so we concentrate on the second. Since $|\tau - \mathbb{E}\tau| \leq t_n$ implies that $\tau \geq q_n^- = (n+1)/2 - o(n)$, we bound the second term in the rhs of (3) by

$$\begin{aligned}&\mathbb{P}(\{\exists j \leq \ell_n,\ \forall i < \tau,\ :\ \Gamma_i \neq j\} \cap \{|\tau - \mathbb{E}\tau| \leq t_n\})\\
&\quad \leq \mathbb{P}\left(\bigcup_{j=1}^{\ell_n}\bigcap_{i=1}^{\tau-1}\{\Gamma_i \neq j\} \cap \{\tau > q_n^-\}\right) \leq \mathbb{P}\left(\bigcup_{j=1}^{\ell_n}\bigcap_{i=1}^{q_n^-}\{\Gamma_i \neq j\}\right)\\
&\quad \leq \sum_{j=1}^{\ell_n}\mathbb{P}\left(\bigcap_{i=1}^{q_n^-}\{\Gamma_i \neq j\}\right) = \sum_{j=1}^{\ell_n}(\mathbb{P}(\Gamma_1 \neq j))^{q_n^-} = \sum_{j=1}^{\ell_n}\left(1 - \frac{1}{2^j}\right)^{q_n^-}\\
&\quad \leq \sum_{j=1}^{\ell_n}\exp\left\{-\frac{q_n^-}{2^j}\right\} = \sum_{j=1}^{\ell_n}\exp\left\{-\frac{q_n^-}{2^{\ell_n}}2^{\ell_n - j}\right\}\\
&\quad \leq \sum_{k=0}^{\infty}\exp\left\{-\frac{q_n^-}{2^{\ell_n}}2^k\right\} \leq \sum_{k=1}^{\infty}\exp\left\{-\frac{q_n^-}{2^{\ell_n}}k\right\} \leq 2\exp\left\{-\frac{q_n^-}{2^{\ell_n}}\right\}\end{aligned}$$

as long as $q_n^-/2^{\ell_n} \geq 1$. Furthermore, the upper bound will go to 0 as $n \to \infty$ if $q_n^-/2^{\ell_n} \to \infty$. For that it is enough to let $\ell_n \sim \log_2(q_n^-/\phi(q_n^-))$, where $q_n^-/\phi(q_n^-) \to \infty$ as $n \to \infty$. For our purpose, the choice $\phi(q_n^-) = \log_2(q_n^-)$ will be convenient. With this choice, we conclude that

$$\mathbb{P}(D_n \leq \ell_n) \leq 2\exp\left\{-\frac{q_n^-}{q_n^-/\ln(q_n^-)}\right\} \leq \frac{2}{q_n^-} = O\left(\frac{1}{n}\right).$$

Using the fact that $0 \leq U_n^{(m)}/D_n \leq 1$, we infer that

$$\mathbb{E}\frac{U_n^{(m)}}{D_n} = \mathbb{E}\frac{U_n^{(m)}}{D_n}I_{D_n \leq \ell_n} + \mathbb{E}\frac{U_n^{(m)}}{D_n}I_{D_n > \ell_n} \leq \mathbb{P}(D_n \leq \ell_n) + \frac{\mathbb{E}U_n^{(m)}}{\ell_n}.$$

As we will see in the next section, $\mathbb{E}U_n^{(m)} = \Theta(1)$, so that the second term in the last sum is dominating.

As for the lower bound, consider a sequence (k_n) which will be specified later. We then have

$$\mathbb{E}\frac{U_n^{(m)}}{D_n} \geq \mathbb{E}\frac{U_n^{(m)}}{D_n} I_{D_n \leq k_n} \geq \frac{1}{k_n}\mathbb{E}U_n^{(m)} I_{D_n \leq k_n} = \frac{1}{k_n}(\mathbb{E}U_n^{(m)} - \mathbb{E}U_n^{(m)} I_{D_n > k_n}).$$

We will choose (k_n) so that the term $\mathbb{E}U_n^{(m)} I_{D_n > k_n}$ will be of lower order than $\mathbb{E}U_n^{(m)}$. Since the latter term will be shown to be bounded away from zero, this means that it suffices to choose (k_n) so that

$$\mathbb{E}U_n^{(m)} I_{D_n > k_n} \to 0 \qquad \text{as} \qquad n \to \infty.$$

Since the number of distinct part sizes is no larger than the largest part size, letting Γ_n^* and Γ_τ^* denote $\max\{\Gamma_1, \ldots, \Gamma_n\}$ and $\max\{\Gamma_1, \ldots, \Gamma_\tau\}$, respectively, we have

$$U_n^{(m)} \leq D_n \leq \Gamma_\tau^* \leq \Gamma_n^*.$$

(The second inequality is valid since the size of the last part is no more than Γ_τ.) It follows that

$$\{D_n > k_n\} \subset \{\Gamma_n^* \geq k_n\},$$

and thus

$$\mathbb{E}U_n^{(m)} I_{D_n > k_n} \leq \mathbb{E}\Gamma_n^* I_{\Gamma_n^* \geq k_n}.$$

To find a choice of (k_n) that would make this latter expectation go to 0 we write

$$\mathbb{E}\Gamma_n^* I_{\Gamma_n^* \geq k_n} = \sum_{t=k_n}^{\infty} t\mathbb{P}(\Gamma_n^* = t) \leq n \sum_{t=k_n}^{\infty} t\mathbb{P}(\Gamma_1 = t) = n \sum_{t=k_n}^{\infty} \frac{t}{2^t} = \frac{n(2+2k_n)}{2^{k_n}}.$$

Choosing $k_n \sim \log_2(n\psi(n))$, we get

$$\mathbb{E}\Gamma_n^* I_{\Gamma_n^* \geq k_n} \leq \frac{2n + 2n\log_2(n\psi(n))}{n\psi(n)},$$

which goes to 0 for $\psi(n) = \log_2^2 n$, for example. Thus one can set $k_n \sim \log_2(n \log_2^2 n)$. With these choices of (ℓ_n) and (k_n) we obtain that

$$\begin{aligned}\mathbb{P}(A_n^{(m)}) = \mathbb{E}\frac{U_n^{(m)}}{D_n} &= \mathbb{E}\frac{U_n^{(m)}}{D_n} I_{\ell_n \leq D_n \leq k_n} + \mathbb{E}\frac{U_n^{(m)}}{D_n} I_{\{\ell_n \leq D_n \leq k_n\}^c} \\ &= \mathbb{E}\frac{U_n^{(m)}}{\log_2 n \pm o(\log n)} + \mathbb{E}\frac{U_n^{(m)}}{D_n} I_{D_n < \ell_n} + \mathbb{E}\frac{U_n^{(m)}}{D_n} I_{D_n > k_n}.\end{aligned}$$

By the choice of (ℓ_n) and (k_n) the last two expectations are bounded above by

$$\begin{aligned}\mathbb{P}(D_n < \ell_n) + \mathbb{P}(D_n > k_n) &\leq 2 \cdot 2^{-q_n^-/2^{\ell_n}} + \mathbb{P}(\Gamma_n^* > k_n) \\ &\leq 2 \cdot 2^{-\phi(q_n^-)} + n\mathbb{P}(\Gamma_1 > k_n) \leq O\left(\frac{1}{n}\right) + \frac{n}{2^{k_n}} \leq O\left(\frac{1}{\log_2^2 n}\right),\end{aligned}$$

and we see that

$$(\ln n)\mathbb{P}(A_n^{(m)}) = \frac{\mathbb{E}U_n^{(m)}}{\log_2 e + o(1)} + o(1), \tag{4}$$

provided that $\mathbb{E}U_n^{(m)} = \Theta(1)$. Thus, the asymptotic behavior of $(\ln n)\mathbb{P}(A_n^{(m)})$ is determined completely by the behavior of $\mathbb{E}U_n^{(m)}$, and to complete the proof we need to estimate $\mathbb{E}U_n^{(m)}$.

5. Parts of multiplicity *m*. In this section we will approximate $\mathbb{E}U_n^{(m)}$. Let $\mathcal{U}^{(m)} = \mathcal{U}^{(m)}(\kappa)$ denote the set of part sizes in κ that have multiplicity m, and let us write $j \in \mathcal{U}^{(m)}$ to indicate that size "j" has multiplicity m. We have

$$\mathbb{E}U_n^{(m)} = \mathbb{E} \sum_{j \le n/m} I_{j\in\mathcal{U}^{(m)}} = \sum_{j\le n/m} \mathbb{P}(j \in \mathcal{U}^{(m)}).$$

Therefore, we need to estimate the sum of $\mathbb{P}(j \in \mathcal{U}^{(m)})$. The degree of difficulty of this approximation increases with the accuracy that one desires to achieve. Furthermore, since, as we will see, $\mathbb{E}U_n^{(m)}$ is an oscillatory function, explicit bounds on $\mathbb{E}U_n^{(m)}$, no matter how tight, cannot be used to show that $(\ln n)\mathbb{P}(A_n^{(m)})$ converges. Thus, one may consider devoting too much attention to an accurate approximation to be a questionable investment. We will present the detailed argument for the fairly precise bound on $\mathbb{E}U_n^{(m)}$, but the reader interested in just the fact that this expectation is $\Theta(1)$ (which is all that is needed to establish (4)) will notice that the argument may be simplified. To make this point more transparent, let $\tilde\Gamma_i(\kappa)$, $i = 1, \dots, \tau(\kappa)$, denote the parts of a composition κ, i.e.,

$$\tilde\Gamma_i(\kappa) = \Gamma_i(\kappa) \quad \text{for} \quad i < \tau(\kappa) \quad \text{and} \quad \tilde\Gamma_{\tau(\kappa)}(\kappa) = n - \sum_{i=1}^{\tau(\kappa)-1} \Gamma_i(\kappa).$$

It is much more convenient to work with Γ's rather than with $\tilde\Gamma$'s, because the last part, $\tilde\Gamma_\tau$, complicates the dependence structure. As a result, a nonnegligible part of our argument is to show that "tildes" can be neglected. This is, of course, not an issue if one is interested merely in a $\Theta(1)$ result; tildes may be dropped since the single part $\tilde\Gamma_\tau$ can be ignored without affecting $U_n^{(m)}$ by more than 1. To estimate $\mathbb{P}(j \in \mathcal{U}^{(m)})$ write

$$\begin{aligned}
\mathbb{P}(j \in \mathcal{U}^{(m)}) =& \mathbb{P}\left(\sum_{i=1}^{\tau} I_{\tilde\Gamma_i = j} = m\right) = \mathbb{P}\left(\{\tilde\Gamma_\tau = j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m-1\right\}\right)\\
&+ \mathbb{P}\left(\{\tilde\Gamma_\tau \ne j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right\}\right)\\
&= \mathbb{P}\left(\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right) + \mathbb{P}\left(\{\tilde\Gamma_\tau = j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m-1\right\}\right)\\
&- \mathbb{P}\left(\{\tilde\Gamma_\tau = j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right\}\right).
\end{aligned} \tag{5}$$

We begin by estimating the first probability in (5), and then we will show that the sums over j of the last two probabilities are negligible. Let $q_n^\pm = (n+1)/2 \pm t_n$. As in the previous section, let $t_n = \sqrt{\alpha(n-1)\ln n}$, but we now choose $\alpha = 2$ so that from (2) we get $n\mathbb{P}(|\tau - \mathbb{E}\tau| \geq t_n) \leq 2/n^3$. To get an upper bound on the first term in (5) write

$$\mathbb{P}\left(\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right) \leq \mathbb{P}\left(\{|\tau - \mathbb{E}\tau| \leq t_n\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right\}\right) + \mathbb{P}(|\tau - \mathbb{E}\tau| > t_n). \tag{6}$$

The second probability in the rhs of (6) is $O(1/n^4)$ and for the first one we have

$$\begin{aligned}
&\mathbb{P}\left(\left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right\} \cap \{|\tau - \mathbb{E}\tau| \leq t_n\}\right)\\
&\quad = \mathbb{P}\Bigg(\bigcup_{1\leq i_1<\cdots<i_m\leq\tau}\Bigg(\bigcap_{\ell=1}^{m}\{\Gamma_{i_\ell}=j\} \cap \bigcap_{\substack{i=1\\ i\neq i_1,\dots,i_m}}^{\tau-1}\{\Gamma_i \neq j\}\Bigg) \cap \{|\tau - \mathbb{E}\tau| \leq t_n\}\Bigg)\\
&\quad \leq \mathbb{P}\Bigg(\bigcup_{1\leq i_1<\cdots<i_m\leq q_n^+}\Bigg(\bigcap_{\ell=1}^{m}\{\Gamma_{i_\ell}=j\} \cap \bigcap_{\substack{i=1\\ i\neq i_1,\dots,i_m}}^{q_n^- -1}\{\Gamma_i \neq j\}\Bigg) \cap \{|\tau - \mathbb{E}\tau| \leq t_n\}\Bigg)\\
&\quad \leq \mathbb{P}\Bigg(\bigcup_{1\leq i_1<\cdots<i_m\leq q_n^+}\Bigg(\bigcap_{\ell=1}^{m}\{\Gamma_{i_\ell}=j\} \cap \bigcap_{\substack{i=1\\ i\neq i_1,\dots,i_m}}^{q_n^- -1}\{\Gamma_i \neq j\}\Bigg)\Bigg)\\
&\quad \leq \binom{q_n^+}{m}\frac{1}{2^{jm}}\left(1 - \frac{1}{2^j}\right)^{q_n^- -1-m}.
\end{aligned}$$

Similarly, to get a lower bound for the first term of (5) we have

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right) \geq \mathbb{P}\left(\left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right\} \cap \{|\tau - \mathbb{E}\tau| \leq t_n\}\right)\\
&\quad \geq \mathbb{P}\Bigg(\bigcup_{1\leq i_1<\cdots<i_m\leq q_n^-}\Bigg(\bigcap_{\ell=1}^{m}\{\Gamma_{i_\ell}=j\} \cap \bigcap_{\substack{i=1\\ i\neq i_1,\dots,i_m}}^{q_n^+}\{\Gamma_i \neq j\}\Bigg) \cap \{|\tau - \mathbb{E}\tau| \leq t_n\}\Bigg)\\
&\quad \geq \mathbb{P}\Bigg(\bigcup_{1\leq i_1<\cdots<i_m\leq q_n^-}\Bigg(\bigcap_{\ell=1}^{m}\{\Gamma_{i_\ell}=j\} \cap \bigcap_{\substack{i=1\\ i\neq i_1,\dots,i_m}}^{q_n^+}\{\Gamma_i \neq j\}\Bigg)\Bigg) - \mathbb{P}(|\tau - \mathbb{E}\tau| > t_n)\\
&\quad \geq \binom{q_n^-}{m}\frac{1}{2^{jm}}\left(1 - \frac{1}{2^j}\right)^{q_n^+ -m} - O\left(\frac{1}{n^4}\right).
\end{aligned}$$

It remains to bound the sum over j of the terms

$$\mathbb{P}\left(\{\tilde{\Gamma}_\tau = j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right\}\right) \tag{7}$$

and

$$\mathbb{P}\left(\{\tilde{\Gamma}_\tau = j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m-1\right\}\right) \tag{8}$$

in (5) and to show that they are negligible compared to the sum over j of the first term in (5). Since

$$\{\tilde{\Gamma}_\tau = j\} \subset \{\tilde{\Gamma}_\tau \geq j\} \subset \{\Gamma_\tau \geq j\},$$

for the probability in (7) we have

$$\begin{aligned}
\mathbb{P}\left(\{\tilde{\Gamma}_\tau = j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right\}\right) &\leq \mathbb{P}\left(\{\Gamma_\tau \geq j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right\}\right)\\
&\leq \mathbb{P}\left(\{|\tau - \mathbb{E}\tau| \leq t_n\} \cap \{\Gamma_\tau \geq j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i=j} = m\right\}\right) + O\left(\frac{1}{n^4}\right)\\
&\leq \sum_{k=q_n^-}^{q_n^+} \mathbb{P}\left(\{\tau = k\} \cap \{\Gamma_k \geq j\} \cap \left\{\sum_{i=1}^{k-1} I_{\Gamma_i=j} = m\right\}\right) + O\left(\frac{1}{n^4}\right)\\
&\leq \sum_{k=q_n^-}^{q_n^+} \mathbb{P}\left(\{\Gamma_k \geq j\} \cap \left\{\sum_{i=1}^{k-1} I_{\Gamma_i=j} = m\right\}\right) + O\left(\frac{1}{n^4}\right)\\
&\leq \sum_{k=q_n^-}^{q_n^+} \frac{1}{2^{j-1}} \binom{k-1}{m} \frac{1}{2^{jm}} \left(1 - \frac{1}{2^j}\right)^{k-1-m} + O\left(\frac{1}{n^4}\right)\\
&\leq c\sqrt{n \ln n} \binom{q_n^+}{m} \frac{1}{2^{j(m+1)}} \left(1 - \frac{1}{2^j}\right)^{q_n^- - m} + O\left(\frac{1}{n^4}\right)
\end{aligned}$$

by the definition of q_n^+ and q_n^-. Thus, summing up over j, we get

$$\begin{aligned}
c\sqrt{n \ln n} &\binom{q_n^+}{m} \sum_{j=1}^{n} \frac{1}{2^{j(m+1)}} \left(1 - \frac{1}{2^j}\right)^{q_n^- - m} + O\left(\frac{1}{n^3}\right)\\
&\leq c\sqrt{n \ln n} \binom{q_n^+}{m} \int_1^\infty \frac{1}{2^{(m+1)x}} \left(1 - \frac{1}{2^x}\right)^{q_n^- - m} dx + O\left(\frac{1}{n^3}\right)\\
&\leq c\sqrt{n \ln n} \binom{q_n^+}{m} \int_0^1 u^m (1-u)^{q_n^- - m} du + O\left(\frac{1}{n^3}\right)\\
&= c\sqrt{n \ln n} \binom{q_n^+}{m} \frac{\Gamma(m+1)\Gamma(q_n^- - m + 1)}{\Gamma(q_n^- + 2)} + O\left(\frac{1}{n^3}\right)\\
&= \Theta\left(\sqrt{\frac{\ln n}{n}}\right).
\end{aligned}$$

For the second probability (8) the argument is essentially the same:

$$\begin{aligned}
\mathbb{P}\left(\{\tilde{\Gamma}_\tau = j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i = j} = m-1\right\}\right) &\le \mathbb{P}\left(\{\Gamma_\tau \ge j\} \cap \left\{\sum_{i=1}^{\tau-1} I_{\Gamma_i = j} = m-1\right\}\right)\\
&\le \sum_{k=q_n^-}^{q_n^+} \mathbb{P}\left(\{\tau = k\} \cap \{\Gamma_k \ge j\} \cap \left\{\sum_{i=1}^{k-1} I_{\Gamma_i = j} = m-1\right\}\right) + O\left(\frac{1}{n^4}\right)\\
&\le \sum_{k=q_n^-}^{q_n^+} \frac{1}{2^{j-1}} \binom{k-1}{m-1} \frac{1}{2^{j(m-1)}} \left(1 - \frac{1}{2^j}\right)^{k-1-m+1} + O\left(\frac{1}{n^4}\right)\\
&\le c\sqrt{n \ln n} \binom{q_n^+}{m-1} \frac{1}{2^{jm}} \left(1 - \frac{1}{2^j}\right)^{q_n^- - m} + O\left(\frac{1}{n^4}\right),
\end{aligned}$$

and in the same fashion as before we see that the sum over j of these terms does not exceed $\Theta(\sqrt{(\ln n)/n})$.

We now observe that for $q = q_n \sim n/2$ the sum

$$\binom{q}{m} \sum_{j=1}^{\infty} 2^{-jm} (1 - 2^{-j})^{q-m}$$

is easily seen to be $\Theta(1)$ (it suffices to compare it to the integral $\int_1^\infty 2^{-mx}(1 - 2^{-x})^{q-m} dx$). Therefore, since q_n^+ and q_n^- are asymptotically the same, m is fixed, and

$$\sum_{j>n} 2^{-jm} (1 - 2^{-j})^{q_n^\pm - m} \le \sum_{j>n} 2^{-j} = \frac{1}{2^n},$$

we conclude that

$$\mathbb{E}U_n^{(m)} \sim \binom{q}{m} \sum_{j=1}^{\infty} 2^{-jm} (1 - 2^{-j})^{q-m}. \tag{9}$$

A more detailed analysis reveals a quite interesting and unexpected phenomenon. The rhs in (9) does not have a limit, but exhibits oscillations about $1/(m \ln 2)$. To see this, one approach is as follows (for convenience we will replace $q - m$ with q in (9)—this does not affect asymptotics): expanding $(1 - 2^{-j})^q$ using the binomial formula, and summing over j, gives

$$\sum_{j=1}^{\infty} 2^{-jm} (1 - 2^{-j})^q = \sum_{k=0}^{q} (-1)^k \binom{q}{k} \frac{1}{2^{k+m} - 1}.$$

Alternating sums of this type appear surprisingly often in the analysis of certain algorithms and can be approximated using methods of complex analysis. Since the standard method, attributed to Rice, has been described recently in several papers, we will not reproduce the details here. Rather, we refer to [16, section 5.2.2], [9], [14], or [15] for some examples of applications and illustration of the method. In particular, these last two papers explicitly treat the asymptotics of the sum

$$\sum_{k=0}^{q} (-1)^k \binom{q}{k} \frac{1}{2^{k+m} - 1}.$$

We would like to indicate an alternative approach to approximating (9) shown to the first author by Bennett Eisenberg and Gilbert Stengle [5]. Although it seems less general than the Rice method, in the case of our sum it gives a more elementary and direct proof of the asymptotics. Consider a sequence (q_s) such that for some $1 \le \beta < 2$, $q_s/2^s \to \beta = 2^x$, $0 \le x < 1$. Then, for s large, replacing q_s with 2^{x+s} and j with $s+r$ we get

$$\binom{q_s}{m} \sum_{j=1}^{\infty} \frac{1}{2^{jm}} \left(1 - \frac{1}{2^j}\right)^{q_s} \sim \frac{2^{m(x+s)}}{m!} \sum_{r=-s+1}^{\infty} 2^{-sm} 2^{-rm} \left(1 - \frac{1}{2^s 2^r}\right)^{2^x 2^s}$$
$$\sim \frac{1}{m!} \sum_{r=-\infty}^{\infty} 2^{m(x-r)} e^{-2^{x-r}},$$

where the "legality" of passing to the limits is easily checked (see [5]). The latter expression defines a 1-periodic function, and its Fourier coefficients are easily found:

$$\begin{aligned} \phi_\ell =& \frac{1}{m!} \int_0^1 \sum_{r=-\infty}^{\infty} 2^{m(x-r)} e^{-2^{x-r}} e^{-2\pi i \ell x} dx \\ =& \frac{1}{m!} \int_{-\infty}^{\infty} 2^{mx} e^{-2^x} e^{-2\pi i \ell x} dx, \end{aligned}$$

which, upon substitution $u = 2^x$, becomes

$$\phi_\ell = \frac{1}{m! \ln 2} \int_0^\infty u^{m-1-2\pi i \ell / \ln 2} e^{-u} du = \frac{1}{m! \ln 2} \Gamma\left(m - \frac{2\pi i \ell}{\ln 2}\right).$$

Note that $\phi_0 = 1/(m \ln 2)$, so if we let

$$H^{(m)}(x) = \frac{\ln 2}{m!} \sum_{r=-\infty}^{\infty} 2^{m(x-r)} e^{-2^{x-r}} - \frac{1}{m},$$

then $H^{(m)}$ satisfies the conditions of Theorem 1. Combining this with (4) completes the proof of the theorem.

6. Bounding the oscillation. The sum below is used in section 5 of the paper to approximate $\mathbb{E}U_n^{(m)}$:

$$\binom{q}{m} \sum_{j=1}^{\infty} 2^{-jm} (1 - 2^{-j})^{q-m}, \tag{10}$$

where $q = \lfloor (n/2) \rfloor$. The data displayed in Figures 2, 3, and 4 indicate that $1/(m \ln 2)$ is a reasonable approximation to the actual value $\mathbb{E}U_n^{(m)}$ for small m and that the sum (10) is a good approximation to $\mathbb{E}U_n^{(m)}$ as n gets large.

As noted in the previous section, the sum (10) oscillates about $1/(m \ln 2)$. We would like to note that the oscillation is not an artifact of our approximation. The data show that the actual value of $\mathbb{E}U_n^{(m)}$ does itself oscillate about $1/(m \ln 2)$. This is illustrated in Figures 5, 6, and 7 for $m = 1, 5, 10$, respectively. These plots use successively coarser scales and show how the amplitude of the oscillation of $\mathbb{E}U_n^{(m)}$ about $1/(m \ln n)$ increases as m increases.

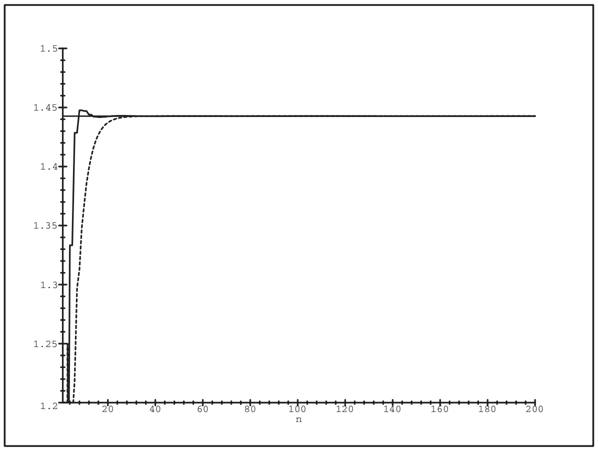

FIG. 2. *Comparison of* $\mathbb{E}U_n^{(1)}$ *(dotted) the approximating sum* (10) *and* $1/(\ln 2)$ *(bold).*

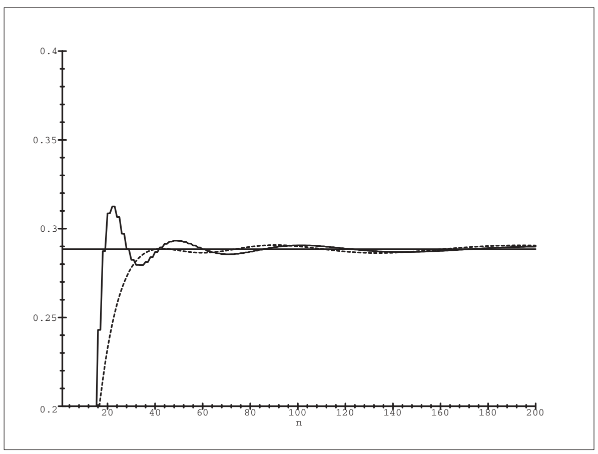

FIG. 3. *Comparison of* $\mathbb{E}U_n^{(5)}$ *(dotted), the approximating sum* (10), *and* $1/(5\ln 2)$ *(bold).*

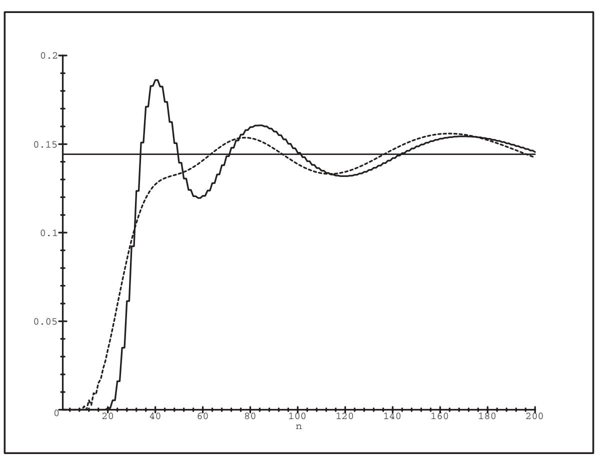

FIG. 4. *Comparison of* $\mathbb{E}U_n^{(10)}$ *(dotted) the approximating sum* (10) *and* $1/(10\ln 2)$ *(bold).*

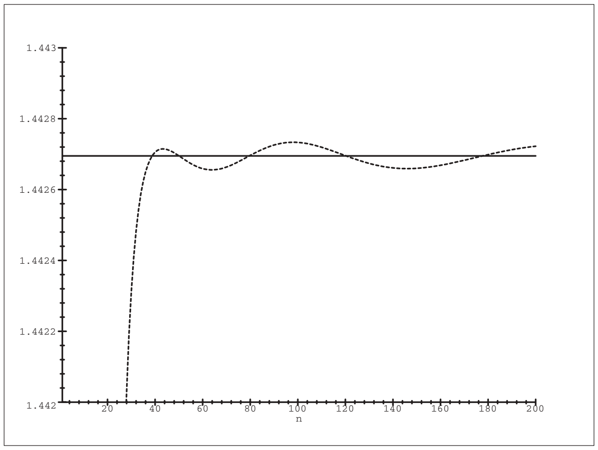

FIG. 5. *The oscillation of* $\mathbb{E}U_n^{(1)}$ *about* $1/(\ln 2)$.

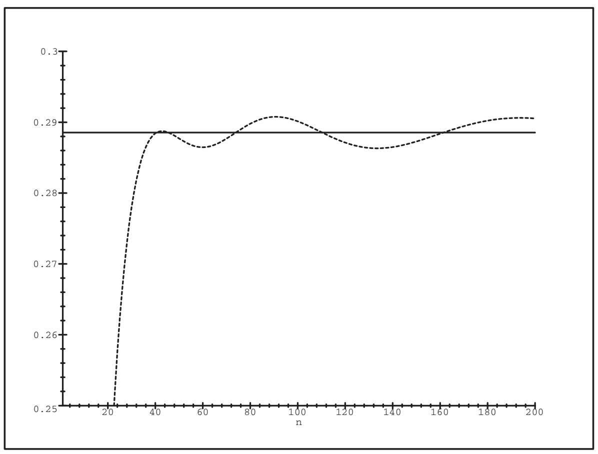

FIG. 6. *The oscillation of* $\mathbb{E}U_n^{(5)}$ *about* $1/(5\ln 2)$.

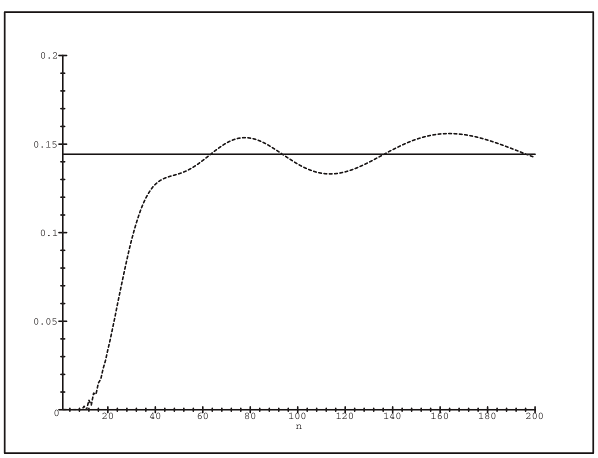

FIG. 7. *The oscillation of* $\mathbb{E}U_n^{(10)}$ *about* $1/(10\ln 2)$.

In order to bound the amplitude of the oscillation, note that the coefficients of Fourier expansion of the function $H^{(m)}/\ln 2$ (which is asymptotic to $\mathbb{E}U_n^{(m)} - 1/(m\ln 2)$) satisfy

$$\sum_{\ell\neq 0}|\phi_\ell| = \frac{1}{m!\ln 2}\sum_{\ell\neq 0}\left|\Gamma\left(m-\frac{2\pi i\ell}{\ln 2}\right)\right|,$$

and therefore,

$$\limsup_n \left|\mathbb{E}U_n^{(m)} - \frac{1}{m\ln 2}\right| \le \left|\frac{H^{(m)}(c\ln n)}{\ln 2}\right| \le \frac{1}{m!\ln 2}\sum_{\ell\neq 0}\left|\Gamma\left(m-\frac{2\pi i\ell}{\ln 2}\right)\right|.$$

Using the properties of gamma function,

$$|\Gamma(it)| = \sqrt{\frac{\pi}{t\sinh(\pi t)}} \quad \text{and} \quad \Gamma(z+1) = z\Gamma(z),$$

and letting $\rho = 2\pi/\ln 2$, we get a bound on the oscillation:

$$\frac{1}{m!\ln 2}\sum_{\ell\neq 0}\left|\Gamma(m-\rho\ell i)\right| \le \frac{2}{m!\ln 2}\sum_{\ell=1}^{\infty}\left(\prod_{k=0}^{m-1}|k-\rho\ell i|\right)\sqrt{\frac{\pi}{\rho\ell\sinh(\pi\rho\ell)}}$$

$$= \frac{\sqrt{2}}{m!\sqrt{\ln 2}}\sum_{\ell=1}^{\infty}\left(\prod_{k=0}^{m-1}\sqrt{k^2+\rho^2\ell^2}\right)\frac{1}{\sqrt{\ell\sinh(\pi\rho\ell)}}. \tag{11}$$

For small m, this bound is very good, but as illustrated in Table 1, as m increases it becomes increasingly weaker. In fact, for m exceeding the value 52, it becomes useless, as this bound on the amplitude exceeds the mean value, $1/(m\ln 2)$. (Thus, from this bound on the oscillation, one could not even conclude that the quantity (10) is positive.) A more detailed analysis of the nature of these fluctuations is perhaps an interesting question but we do not pursue it much further in this paper. One thing worth pointing out is that oscillations of $\mathbb{E}U_n^{(m)}$ are highly nonsymmetric around $1/(m\ln 2)$. On one hand, considering q of the form $m2^p$ and replacing the sum over j by its largest term (corresponding to $j = p$) we find that

$$\limsup_n \mathbb{E}U_n^{(m)} \ge \binom{q}{m}\left(\frac{m}{q}\right)^m\left(1-\frac{m}{q}\right)^{q-m} \sim \frac{m^m}{m!}e^{-m} \ge \frac{e^{-1/(12m)}}{\sqrt{2\pi m}}$$

by Stirling's formula. On the other hand, we have

$$\liminf_n \mathbb{E}U_n^{(m)} \ge \frac{e^{-2m}}{m\ln 2}.$$

To see this, let

$$f(x) = \binom{q}{m}2^{-mx}(1-2^{-x})^{q-m},$$

so that

$$\mathbb{E}U_n^{(m)} \sim \sum_{j=1}^{\infty} f(j).$$

TABLE 1
Comparison of the bound (11) *with the mean value as* m *increases.*

m	The bound (11) on oscillation	$1/(m \ln 2)$
1	.00001426024765	1.442695041
2	.00006502473820	.7213475205
3	.0002012028112	.4808983470
4	.0004802854938	.3606737603
5	.0009517428766	.2885390082
6	.001642131452	.2404491735
7	.002550173579	.2060992916
8	.003650969724	.1803368801
9	.004904708738	.1602994490
10	.006265585898	.1442695041
⋮	⋮	⋮
50	.02756514237	.02885390082
51	.02757480454	.02828813806
52	.02757887758	.02774413540
53	.02757781675	.02722066115
54	.02757203860	.02671657484
55	.02756192443	.02623081892
56	.02754782372	.02576241145
57	.02753005703	.02531043932
58	.02750891844	.02487405243
59	.02748467827	.02445245832
60	.02745758499	.02404491735
⋮	⋮	⋮
100	.02546701322	.01442695041
150	.02295420798	.009617966940
200	.02098570878	.007213475205
250	.01942401432	.005770780164
300	.01813918847	.004808983470
350	.01704834346	.004121985831
400	.01610016200	.003606737603

Since

$$f'(x) = \binom{q}{m} \frac{\ln 2}{2^{mx}} \left(1 - \frac{1}{2^x}\right)^{q-m-1} \left(\frac{q}{2^x} - m\right),$$

f is increasing for $x < x_0$ and decreasing for $x > x_0$, where $x_0 = \log_2(q/m)$. Therefore, letting $k_0 = [\log_2(q/m)]$ be the integer part of x_0, we see that

$$\begin{aligned}\sum_{j=1}^{\infty} f(j) = \sum_{j=1}^{k_0} f(j) + \sum_{j=k_0+1}^{\infty} f(j) &\geq \int_0^{k_0} f(x)dx + \int_{k_0+1}^{\infty} f(x)dx \\ = \int_0^{\infty} f(x)dx - \int_{k_0}^{k_0+1} f(x)dx.\end{aligned}$$

The first integral upon substitution $u = 2^{-x}$ is easily seen to be equal to $1/(m \ln 2)$, while for the second one, letting $\delta = x_0 - k_0$ and then $x = x_0 + t$, we get

$$\begin{aligned}\int_{k_0}^{k_0+1} f(x)dx &= \int_{x_0-\delta}^{x_0+1-\delta} f(x)dx = \int_{-\delta}^{1-\delta} f(x_0+t)dt \\ &= \binom{q}{m} \int_{-\delta}^{1-\delta} \frac{1}{2^{mx_0}} \frac{1}{2^{mt}} \left(1 - \frac{1}{2^{x_0+t}}\right)^{q-m} dt \\ &= \binom{q}{m} \left(\frac{m}{q}\right)^m \int_{-\delta}^{1-\delta} \frac{1}{2^{mt}} \left(1 - \frac{m}{2^t q}\right)^{q-m} dt,\end{aligned}$$

which upon substitution $u = 2^{-t}$ becomes

$$\binom{q}{m} \left(\frac{m}{q}\right)^m \frac{1}{\ln 2} \int_{2^{\delta-1}}^{2^\delta} u^{m-1} \left(1 - \frac{mu}{q}\right)^{q-m} du.$$

Using $\binom{q}{m} \le q^m/m!$ and letting $q \to \infty$ we see that the latter expression is bounded above by

$$\frac{m^m}{m! \ln 2} \int_{2^{\delta-1}}^{2^\delta} u^{m-1} e^{-mu} du \le \frac{1}{m! \ln 2} \int_0^{m2^\delta} u^{m-1} e^{-u} du,$$

and by a successive partial integration, and because $0 \le \delta \le 1$, we see that the last integral is no more than

$$\frac{1}{m! \ln 2}(m-1)!(1 - e^{-m2^\delta}) \le \frac{1}{m \ln 2}(1 - e^{-2m}).$$

Thus,

$$\int_{k_0}^{k_0+1} f(x)dx \le \frac{1}{m \ln 2}(1 - e^{-2m}),$$

that is,

$$\mathbb{E} U_n^{(m)} \ge \frac{e^{-2m}}{m \ln 2},$$

and the argument is completed.

Acknowledgment. We would like to express our gratitude to Herbert Wilf for bringing to our attention the paper of Kirschenhofer and the work of Knuth on the alternating sums that are used in our analysis. Part of the research of the first author was carried out while he was visiting the Department of Mathematics of Lehigh University in the spring semester of 1999. He would like to thank the department for its hospitality.

REFERENCES

[1] N. ALON, AND J. H. SPENCER, *The Probabilistic Method*, 2nd ed., Wiley-Interscience, New York, 2000.
[2] G. E. ANDREWS, *The Theory of Partitions*, Addison–Wesley, Reading, MA, 1976.
[3] D. W. BOYD, *Losing Runs in Bernoulli Trials*, unpublished manuscript, 1972.
[4] S. CORTEEL, B. PITTEL, C. D. SAVAGE, AND H. S. WILF, *On the multiplicity of parts in a random partition*, Random Structures Algorithms, 14 (1999), pp. 185–197.

[5] B. Eisenberg, G. Stengle, and G. Strang, *The asymptotic probability of a tie for first place*, Ann. Appl. Probab., 3 (1993), pp. 731–745.

[6] P. Erdös and J. Lehner, *The distribution of the number of summands in the partitions of positive integer*, Duke Math. J., 8 (1941), pp. 335–345.

[7] P. Erdös and A. Rényi, *On a new law of large numbers*, J. Anal. Math., 23 (1970), pp. 103–111.

[8] P. Erdös and P. Révész, *On the length of the longest head–run*, in Topics in Information Theory, Colloq. Math. Soc. Janos Bolyai 16, North-Holland, Amsterdam, 1975, pp. 219–228.

[9] P. Flajolet and R. Sedgewick, *Mellin transform and asymptotics: Finite differences and Rice's integrals*, Theoret. Comput. Sci., 144 (1995), pp. 101–124.

[10] B. Fristedt, *The structure of random partitions of large integers*, Trans. Amer. Math. Soc., 337 (1993), pp. 703–735.

[11] W. M. Y. Goh and E. Schmutz, *The number of distinct part sizes in a random integer partition*, J. Combin. Theory Ser. A, 69 (1995), pp. 149–158.

[12] L. Gordon, M. F. Schilling, and M. S. Waterman, *An extreme value theory for long head runs*, Probab. Theory Related Fields, 72 (1986), pp. 279–287.

[13] L. J. Guibas and A. M. Odlyzko, *Long repetitive patterns in random sequences*, Z. Wahrsch. Verv. Gebiete, 53 (1980), pp. 241–262.

[14] P. Kirschenhofer, *A note on alternating sums*, Electron. J. Combin., 3 (1996), Paper #R7.

[15] P. Kirschenhofer and H. Prodinger, *The number of winners in a discrete geometrically distributed sample*, Ann. Appl. Probab., 6 (1996), pp. 687–694.

[16] D. E. Knuth, *The Art of Computer Programming.* Vol. III. *Sorting and Searching*, Addison-Wesley, Reading, MA, 1973.

[17] B. Pittel, *Confirming two conjectures about the integer partitions*, J. Combin. Theory Ser. A, 88 (1999), pp. 123–135.

[18] H. S. Wilf, *Three problems in combinatorial asymptotics*, J. Combin. Theory Ser. A, 35 (1983), pp. 199–207.