Some Unexpected Expected Behavior Results for Bin Packing

J. L. Bentley and D. S. Johnson
AT&T Bell Laboratories
Murray Hill, NJ 07974

F. T. Leighton
Massachusetts Institute of Technology
Cambridge, MA 02139

C. C. McGeoch and L. A. McGeoch
Carnegie-Mellon University
Pittsburgh, PA 15213

ABSTRACT

We study the asymptotic expected behavior of the First Fit and First Fit Decreasing bin packing algorithms applied to items chosen uniformly from the interval (0,u], $u \leqslant 1$. Our results indicate that the algorithms perform even better than previously expected.

1. INTRODUCTION

In the standard one-dimensional bin packing problem we are given a sequence $L = (a_1, a_2, ..., a_n)$ of numbers in the interval (0,1], and are asked to pack them into a minimum number of unit-capacity bins (i.e., bins that can contain items totaling at most 1). This NP-hard problem has served as a testbed for new ideas in the analysis of approximation algorithms for well over a decade (e.g., see [3,9]).

One strand of this research has been the study of worst-case behavior, starting in the early 1970's with proofs that the FIRST FIT (FF) heuristic could use as many as 17/10 times the optimal number of bins (but no more) and that the corresponding asymptotic worst-case ratio for FIRST FIT DECREASING (FFD) was 11/9 [9,10]. This line of research has culminated with the results of Lueker and Fernandez de la Vega [6] and Karmarkar and Karp [12], which imply that an asymptotic worst-case ratio of 1 is achievable with a polynomial time algorithm (the number of excess bins is guaranteed to be "little o" of the optimal number of bins).

This paper follows a second and more recent line of research: analysis of *expected* behavior. We shall begin by stating our two main results, and then explain why we find them interesting and (in two senses) unexpected. Both results concern the standard probabilistic model for this problem, in which instances consist of items independently and uniformly distributed in the interval $(0,u]$ for some $u \leqslant 1$. Let L_n^u denote the random variable whose values are lists of n items generated according to this distribution. For algorithm A and list L, let $A(L)$ denote the number of bins in the packing of L generated by A, let $OPT(L)$ denote the number of bins in an optimal packing, and $\Sigma(L)$ denote the obvious lower bound $\Sigma_{a \in L} a \leqslant OPT(L)$. Note the easy result that $E(\Sigma(L_n^u)) = un/2$.

Our results concern the FF and FFD algorithms mentioned above. In both we start with an ordered sequence of initially-empty bins, into which the items are placed, one at a time. In FF, we place a_1 in bin 1, and treat the remaining items in order, placing each in the first bin that still has room enough for it. FFD differs only in that we initially sort the list so that $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n$. Our main results are the following:

Theorem 1. $E\left[FF(L_n^1) - \Sigma(L_n^1)\right] = O(n^{0.8})$.

Theorem 2. For $u \leqslant 1/2$, $E\left[FFD(L_n^u) - \Sigma(L_n^u)\right] = O(1)$.

Theorem 1 says that the expected performance of FF is "asymptotically optimal" in the same sense as was the worst-case performance of the algorithms of Lueker et al.: the expected excess is $o(E[OPT(L_n^1)])$. The result gains importance from the fact that FIRST FIT is an "on-line" algorithm, i.e., it packs items in the order given, without knowledge of the sizes and number of later items. There are many applications in which only on-line algorithms can be used, and for these the sophisticated algorithms of Lueker et al. do not apply; in fact, it has been proven that no on-line algorithm can guarantee an asymptotic worst-case ratio better than 1.536 [2]. Nevertheless, our result says that asymptotic optimality is still possible, *on average*, and can be obtained using a simple and well-known heuristic.

Theorem 1 is *unexpected* because most researchers, including the authors of this paper, had assumed that $E[FF(L_n^1)/\Sigma(L_n^1)]$ approached some constant greater than 1 as $n \rightarrow \infty$. (Simple hand-waving arguments seem to prove this.) In his thesis [9], Johnson conjectured that the ratio was something like 1.07. In [17], Ong, Magazine and Wee estimated it to be 1.056. However, these estimates were based on simulations involving only 200-item and 1000-item lists respectively. The true nature of FF did not become apparent until recent simulations, run on lists with n growing as large as 128,000 [1], raised the possibility that the expected excess was growing sublinearly.

Turning now to off-line algorithms, we first note that the Karmarkar and Karp algorithm provides a *worst case* upper bound of $O(\log^2(OPT(L)))$ on $A(L) - OPT(L)$ [12], which is better than our bound on the *expected* excess of FF (over $\Sigma(L)$). However, if we are considering off-line algorithms, we can consider FFD as a contender in addition to FF. Theorem 2 serves to emphasize just how good a contender FFD can be in practice, despite the fact that it, like FF, is totally out-classed from an asymptotic worst-case point of view. It was already known that for $u = 1$ the expected excess under FFD was $\Theta(n^{1/2})$ [15]. Theorem 2 gives a much stronger result for $u \leqslant 1/2$: here there is a fixed upper bound on the expected excess, no matter how many items there are in the list. It was well-known that FFD performs well in practice, but few expected that it was actually *this* good. This result was again foreshadowed (and inspired) by simulations. Of additional theoretical interest is the fact that Theorem 2 implies an unexpected discontinuity in the behavior of FFD at $u = 1/2$: we can show that for $u > 1/2$ the expected excess is $\Omega(n^{1/3})$.

So far we have argued that our results were "unexpected" because initially they were not believed to be true. A second sense in which Theorems 1 and 2 are unexpected is that, even supposing that they are true, it is surprising that they can be proven, at least given the current state of the art. A standard objection to probabilistic analysis of algorithms is its supposed difficulty, especially when it comes to analyzing algorithms that do anything half-way clever (as do FF and

FFD). Such algorithms start conditioning their data as soon as they make their first step, thus making it difficult to retain the independence assumptions so vital to straightforward probabilistic arguments. Moreover, these algorithms require large amounts of memory, thus making the state spaces far too large for easy Markov chain arguments. (Although the NEXT FIT heuristic, which never has more than one active bin, has been thoroughly analyzed [4,10], the probability distribution for the height of even just the *second* bin in a FF packing is still unknown.)

The key to escaping these difficulties is an old trick for avoiding dependence effects, one that is perhaps not as widely known as it should be: do all the probabilistic arguments in one part of the proof, and in the rest of the proof use only worst-case analysis. In the context of running-time analysis, Knuth has called this the "O of probability" argument [14]. In our context it yields two-part proofs of the following form:

1. [PROBABILISTIC] Using counting arguments, show that all but a negligible proportion of the instances satisfy a set of desirable properties.
2. [WORST-CASE] Show that for all instances satisfying the given properties, the algorithm generates packings whose excess obeys the claimed bounds.

These two parts can occur in either order (and will do so in what follows). Note that, because we use this type of argument, our results are "asymptotic probability 1" results as well as expected behavior results.

Section 2 reviews several easy lemmas from the bin packing folklore. Sections 3 and 4 then sketch the proofs of Theorems 1 and 2, respectively. The concluding section summarizes our experimental findings in more detail so that they can be compared to our analytic results, and mentions some additional expected behavior results that we have proved or are currently in the process of "nailing down" and which, if true, will complete the classification of the behavior of $E[FFD(L_n^u)]$ and $E[OPT(L_n^u)]$ for all $u \in (0,1]$. We also discuss the extent to which our results can be extended to other distributions.

2. GENERAL PURPOSE LEMMAS

Our first lemma is just an easy observation that provides us with a different way of looking at the quantity (the "expected excess") to which both our theorems refer (this version will be used extensively in proving Theorem 2).

Lemma 2.1. For any list L and bin packing algorithm A,

$$A(L) - \Sigma(L) = \sum_{i=1}^{A(L)} (1-B_i)$$

where B_i is the sum of the items in the i^{th} bin of the packing constructed when A is applied to L (and $1-B_i$ is therefore the "empty" or "unused" space in the bin).

The next (trivial) lemma gives us an alternative way of looking at the construction of the FF and FFD packings: instead of scanning the items in the outer loop and the bins in the inner loop, we scan the bins in the outer loop and the items in the inner loop.

Lemma 2.2. For a given list L, the following procedure is an alternative way of constructing the FF packing of L: Pack the bins in turn, starting with Bin 1. To pack Bin i, suppose L_i is the list obtained from L by deleting all items packed in earlier bins (hence $L_1=L$). Bin i is initially empty and hence has a gap of size 1. Repeat the following until the gap in Bin i is smaller than every item remaining in L_i. Find the first item in L_i that is less than the current gap, remove it from L_i, and add it to Bin i.

It is now easy to derive the following lemma, which we will use in proving Theorem 1.

Lemma 2.3. Suppose B is the set of items contained in Bin j of the FF packing of L. Then if we delete the items in B from L and pack the resulting list by FF, the resulting packing will be identical to that obtained by deleting Bin j from the original packing.

Our final lemma says that we can draw conclusions about the expected excess from results that hold "almost surely".

Lemma 2.4. If A is either FF or FFD, and if for a given c, n, u, and f the probability is greater than $1-c/n$ that $A(L_n^u)-\Sigma(L_n^u) \leqslant f(n)$, then the expected value of of the difference is at most $f(n)+c/2$.

Proof. The set of instances where the bound of $f(n)$ does not hold has measure c/n or less. It is an easy observation to show that $A(L) \leqslant |L|$ for all lists L when A is FF or FFD. Thus the contribution of these anomalous instances to the expected difference is at most $(n/2)(c/n)=c/2$. □

3. PROOF OF THEOREM 1

In this section we prove $E[FF(L_n^1)-\Sigma(L_n^1)] = O(n^{4/5})$. We shall actually prove a somewhat stronger result, showing that a restricted version of FF, one that never outperforms FF, obeys this expected performance bound. This approach is analogous to that used by Lueker [15] and Frederickson [7] in proving corresponding results for FFD: to prove results about a sophisticated algorithm A, first show that the desired bounds hold for a less sophisticated (and easier to analyze) algorithm B, and then show that $A(L) \leqslant B(L)$ for all lists L.

We call our restricted version of FF *restricted First Fit (rFF)*. It differs from FF in that, as soon as a bin receives its second item (if it ever does) it is declared "closed" and can receive no more items. Restricted FF is pleasantly free of many of the anomalies that plague FF. For instance, with FF there exist arbitrarily large expamples of lists L and L' such that L' is obtained from L by deleting a single item and yet $FF(L') = (4/3)\cdot FF(L)$ (the details of the construction are left to the enterprising reader). That this and related anomalies cannot occur with rFF is a consequence of the following lemma.

Lemma 3.1. Let L be a list. Then the following two properties hold:

(a) If L^s is formed from L by switching the order of two adjacent items $x<y$ from an ordering in which y precedes x in L to an ordering in which x precedes y in L^s, then $rFF(L) \leqslant rFF(L^s)$.

(b) If L^+ is formed from L by replacing an item x in L by an item $y>x$ in L^+, then $rFF(L) \leqslant rFF(L^+)$.

Proof. The proof is by induction on the length of L and proceeds in tandem for the two properties. See Appendix A1. □

Lemma 3.2. For any lists L and L' such that L' is a sublist of L,

$$rFF(L) \leqslant rFF(L') + |L-L'|.$$

Proof. Let L^* denote the list formed from L by enlarging the elements in $L-L'$ until each has unit size. Since unit-size elements are packed in bins by themselves, $rFF(L^*)=rFF(L')+|L-L'|$. By repeated applications of Lemma 3.1, we also know that $rFF(L) \leqslant rFF(L^*)$. The result follows. □

Lemma 3.3. For any list L, $FF(L) \leqslant rFF(L)$.

Proof. Let $P_{FF}(L)$ and $P_{rFF}(L)$ be the FF and rFF packings of L, respectively. Our argument proceeds by induction on the length of L. In particular, we will find a non-empty set S of items in L that completely occupy a set B_{FF} of bins of $P_{FF}(L)$ and completely occupy a set B_{rFF} of bins in $P_{rFF}(L)$, where $|B_{FF}| \leqslant |B_{rFF}|$. We then remove S from L and use induction and Lemma 2.3 (which clearly applies to rFF as well as FF) to show that

$$\begin{aligned} FF(L) &= |B_{FF}| + FF(L-S) \\ &\leqslant |B_{rFF}| + FF(L-S) \\ &\leqslant |B_{rFF}| + rFF(L-S) \\ &= rFF(L). \end{aligned}$$

The set S is constructed by the following simple augmentation procedure. Initially S, B_{FF}, and B_{rFF} are empty.

1. Put the items contained in the first bin of $P_{rFF}(L)$ in S. Add the first bin of $P_{rFF}(L)$ to B_{rFF}.
2. Add to S those items that share bins in $P_{FF}(L)$ with items of S. Add the corresponding bins of $P_{FF}(L)$ to B_{FF}.

3. Add to S those items that share bins in $P_{rFF}(L)$ with items in S. Add the corresponding bins of $P_{rFF}(L)$ to B_{rFF}.
4. If S has changed, go to step 2.

It is clear that the procedure eventually halts, and it remains only to show that $|B_{FF}| \leqslant |B_{rFF}|$. Observe first that at the end of the first execution of step 2, $|B_{rFF}| = |B_{FF}| = 1$. This is because, by the definitions of FF and rFF, all of the items contained in the first bin of $P_{rFF}(L)$ must also be contained in the first bin of $P_{FF}(L)$. Thereafter, each time a bin is added to B_{rFF} in step 3, at most one item is added to S, since one of the items in the bin must already have been in S and the bin has at most two items. Thus for each bin added to B_{rFF} in step 3, at most one bin can be added to B_{FF} the next time step 2 is executed. Hence when the procedure halts we must have $|B_{FF}| \leqslant |B_{rFF}|$.

We thus can construct the desired non-empty set S, and the lemma is proved. □.

We are now ready to begin our proof of Theorem 1, which we divide into a "worst-case" part (Part I) and a "probabilistic" part (Part II).

In Part I we obtain an upper bound on $FF(L)$ in terms of two parameters of L having to do with the distribution and ordering of its item sizes. First we use Lemma 3.1 to convert L into a list L_d consisting of items with a restricted range of sizes and $rFF(L_d) \geqslant rFF(L)$. We then show how to remove items from L_d to form a sublist L_d' whose rFF packing has an item of size exceeding 1/2 in every bin, so that by Lemma 3.2 $rFF(L_d')$ is no more than $|L_d - L_d'|$ plus the number of items exceeding 1/2 in L_d. Thus, by Lemma 3.3, $FF(L)$ is no more than the sum of these two quantities.

In Part II we use probabilistic arguments to show that, for L_n^1 and sufficiently large n, these quantities are $O(n^{4/5})$ and $n/2 + O(n^{4/5})$, respectively, with probabilities exceeding $1 - 1/n$. Since $\Sigma(L_n^1) = n/2 - O(\sqrt{n \log n})$ with probability exceeding $1 - 1/n$ the desired expected difference bound follows by Lemma 2.4.

Part I: Worst-Case Arguments

Let k be an odd integer such that $k = \lfloor n^{1/5} \rfloor$ or $k = \lfloor n^{1/5} \rfloor + 1$, where $n = |L|$. Partition the interval (0,1] into the k subintervals $(0,1/k], (1/k, 2/k], \ldots, ((k-1)/k, 1]$. Now obtain L_d from L by replacing each item of L in the interval $((i-1)/k, i/k]$ by an item of size i/k for all i, $1 \leqslant i \leqslant k$. By Lemmas 3.1 and 3.3 $FF(L) \leqslant rFF(L) \leqslant rFF(L_d)$, and we henceforth consider only this "discretized" list L_d and the algorithm rFF. Since the unit-size items in L_d are packed in bins by themselves, we shall ignore them in much of what follows.

We now begin removing items from L_d so that we can guarantee that the rFF packing of the resulting list will have an item of size exceeding 1/2 in every bin. The removal process starts with the items of size $(k-2)/k$ and proceeds down the list of item sizes. Begin by removing a sufficiently large initial portion of the $(k-2)/k$-size items so that every initial segment of the resulting list has at least as many $(k-1)/k$-size items as $(k-2)/k$-size items. Next remove a sufficiently large initial portion of the $(k-3)/k$-size items from the resulting list so that every initial segment in the new list has at least as many $(k-2)/k$-size items as $(k-3)/k$-size items. Notice that the new list still has the property that every intitial segment has at least as many $(k-1)/k$-size items as $(k-2)/k$-size items. Continue deleting items in this fashion until a list L_d' is created in which every intial segment has at least as many $(k-i)/k$-size items as $(k-i-1)/k$-size items, $1 \leqslant i \leqslant k-1$. Call this property of L_d' the *k-dominance* property.

Lemma 3.4. If L' consists only of items with sizes of the form i/k and has the k-dominance property, then rFF packs every item of L' with size less than 1/2 in a bin with an item of size exceeding 1/2. (Since k is odd, there are no items of size exactly 1/2 in L'.)

Proof. We shall in fact show that each item of size $x < 1/2$ will be in a bin with an item of size $1-x$ for a perfect match. We proceed by induction on the length of L'. (The claim clearly holds for the empty list.)

Notice that the first items of each size appear in decreasing order in L'. Hence the first $(k-1)/k$-size item precedes all smaller items and starts the first bin that contains an item smaller than 1. Since only $1/k$-size items will fit in this bin, the first such (if any exist) is placed in it as second item. When the first $(k-2)/k$-size item is packed, all previous bins start with larger items, and so it starts a new bin. The first $2/k$-size item (if any such items exist) would not have fit in any of the earlier bins, and so remains the first item in the list that will fit in this bin, and so is placed in it. (If no such $2/k$-size item exists, then no smaller items exist either, so the bin never gets a second item.) In general, when the first $(k-j)/k$-size item, $1 \leqslant j < k/2$, is packed, all previous bins start with larger items and so it starts a new bin and the first j/k-size item remains the first in the list that will fit with it. Thus each first $(k-j)/k$-size item, $1 \leqslant j < k/2$ is paired with the first j/k-size item (if there is one) in a bin of the rFF packing of L' (and if there is none, that $(k-j)/k$-size item remains in a bin all by itself).

By Lemma 2.3, if we remove all these items and their partners from L', we get a new list L'' whose rFF packing looks exactly like that for L' except that the bins in question, each of which contains an item exceeding 1/2 and either a perfect match or nothing, have been deleted. It is easy to see that L'' has the k-dominance property if L' did (the first items in each non-empty size class were deleted and if a class is empty then all classes of smaller items are also empty). Thus by induction all items in L'' that are smaller than 1/2 are packed with matching items larger than 1/2 in the rFF packing of L'', and so the overall rFF packing of L' has the desired properties. □

By Lemma 3.4, it is clear that $rFF(L_d')$ equals the number of items of size greater than 1/2 in L_d'. In order to bound $rFF(L_d)$ we need to obtain bounds on this quantity and on the number of items in $L_d - L_d'$.

Part II: Probabilistic Arguments

Suppose now that our original list L was an instance of the random list L_n^1. The following two lemmas, when combined with the arguments of Part I, suffice to complete the proof of Theorem 1.

Lemma 3.5. With probability exceeding $1 - 1/n$ for sufficiently large n, the number of items of size greater than 1/2 in L_d' is $n/2 + O(n^{4/5})$.

Proof. The number of items in L_d' that have size exceeding 1/2 is clearly bounded by the number of items of size greater than $(k-1)/2k$ in L_n^1. By a standard counting argument, this number is at most

$$n\left(1 - \frac{k-1}{2k}\right) + O\left(\sqrt{n \log n}\right) = \frac{n}{2} + \frac{n}{2k} + O\left(\sqrt{n \log n}\right)$$
$$= \frac{n}{2} + O\left(n^{4/5}\right)$$

with probability at least $1 - 1/n$ for sufficiently large n. □

Lemma 3.6. With probability exceeding $1 - 1/n$ for sufficently large n, the number of items in $L_d - L_d'$ is $O(n^{4/5})$.

Proof. Define r_i to be the smallest value such that the removal of the first r_i items of size $(i-1)/k$ from L_d insures that every initial segment of the resulting list contains at least as many items of size i/k as of size $(i-1)/k$. The value of r_i is simply the largest amount that the count of i/k-size items lags behind the count of $(i-1)/k$-size items in any initial segment of L_d. In the procedure described above for constructing L_d' from L_d, the number of items of size $(i-1)/k$ removed from L_d is clearly bounded above by $r_{k-1} + r_{k-2} + \cdots + r_i$ for each i. Hence the total number of items removed from L_d to obtain L_d' obeys the bound

$$|L_d - L_d'| \leqslant r_{k-1}(k-2) + r_{k-2}(k-3) + \cdots + r_3(2) + r_2$$
$$< k \sum_{j=2}^{k-1} r_j$$

In what follows we show that this sum is at most $O(k^{3/2} n^{1/2}) = O(n^{4/5})$ with probability $1 - 1/n$ for sufficently large n, thus completing the proof of the lemma. (For simplicity, we shall not specify the constants involved in our use of the "O"-notation, and shall use the term "with high probability" to mean "with probability exceeding $1-1/n^k$ for sufficiently large n" for some suitable $k \leqslant 5$ depending on the situation. It would be straightforward but tedious to specify these constants precisely and thus convert our sketch to a rigorous proof.)

Standard counting and/or probability arguments reveal that L_d contains $n/k + O(\sqrt{(n/k)}\log n)$ items of size i/k for each $i \leqslant k$ with high probability. For each $i > 1$, the sublist consisting of all size i/k and size $(i-1)/k$ items can be viewed as generating a random walk whose length is $2n/k + O(\sqrt{(n/k)}\log n)$ with high probability. Using the "reflection principle" [5], it is easy to show that the probability that $r_i > d\sqrt{n/k}$, i.e., that there is some initial segment of this sublist whose count of size $(i-1)/k$ items exceeds the count of size i/k items by more than $d\sqrt{n/k}$, is precisely twice the probability that the *total* count of size $(i-1)/k$ items exceeds *total* count of size i/k items by more than $d\sqrt{n/k}$. Again, using standard counting arguments, the probability of the latter event occurring is $O(e^{-\alpha d^2})$ for some constant $\alpha > 0$.

By the preceding facts, we could immediately conclude that $r_i = O(\sqrt{(n/k)}\log n)$ for each i with high probability, but this result only implies that $|L_d - L_d'| = O(k^{3/2}\sqrt{n}\log n)$, which is a factor of $\sqrt{\log n}$ too large and hence yields a bound of $O(n^{4/5}\sqrt{\log n})$ on $E[FF(L_n^1) - \Sigma(L_n^1)]$. To obtain a better bound, we will show in what follows that $r_{k-1}+r_{k-3}+\cdots+r_2 = O(\sqrt{kn})$ with high probability and that $r_{k-2}+r_{k-4}+\cdots+r_3 = O(\sqrt{kn})$ with high probability. This is clearly sufficient to prove that, with high probability, $|L_d - L_d'| \leqslant k\cdot(r_{k-1}+r_{k-2}+\cdots+r_2) = O(k^{3/2}n^{1/2}) = O(n^{4/5})$.

We are forced to consider the r_i in two groups because r_i could be seriously dependent on r_{i-1} for each i. The only effect that r_{i-2} can have on r_i, however, is that it may influence the cardinality of the set consisting of the size i/k and size $(i-1)/k$ items (although not the distribution of the items within that set). Since we have already restricted our attention (with high probability) to the case where this cardinality is $n/k + O(\sqrt{(n/k)}\log n)$, we can take account of this dependency by using the cardinality within this range which yields the largest bounds. In what follows we consider the case of the r_i with i even; the case for odd i is handled similarly.

Consider the probability that $r_{k-1}+r_{k-3}+\cdots+r_2 \geqslant s$. This probability is at most the probability that

$$\left\lfloor r_{k-1}\sqrt{k/n}\right\rfloor + \left\lfloor r_{k-3}\sqrt{k/n}\right\rfloor + \cdots + \left\lfloor r_2\sqrt{k/n}\right\rfloor \geqslant s\sqrt{k/n} - k.$$

We wish to find the minimum value of s for which this probability is small. From before, we know that the probability that $\left\lfloor r_{2i}\sqrt{k/n}\right\rfloor \geqslant d$ for some integer d is $O(e^{-\alpha d^2})$. Let $f_{2i}(x)$ denote a generating function for $\left\lfloor r_{2i}\sqrt{k/n}\right\rfloor$ such that the j^{th} coefficient of $f_{2i}(x)$ is an upper bound on the probability that $\left\lfloor r_{2i}\sqrt{k/n}\right\rfloor$ is greater than j. For sufficiently large $\beta > 0$, the function $f_{2i}(x) = \beta e^x = \sum_{j=0}^{\infty}\beta x^j/j!$ satisfies this requirement since $\beta/j! = \Omega(\beta e^{-j(\ln j - 1)}) > \Theta(e^{-\alpha j^2})$ for sufficiently large β. Let $g(x)$ be a corresponding generating function for $\left\lfloor r_{k-1}\sqrt{k/n}\right\rfloor + \left\lfloor r_{k-3}\sqrt{k/n}\right\rfloor + \cdots + \left\lfloor r_2\sqrt{k/n}\right\rfloor$. Clearly $g(x) = f_{k-1}(x)\cdot f_{k-3}(x)\cdot \cdots \cdot f_2(x) = (\beta e^x)^{(k-1)/2} = \beta^{(k-1)/2}e^{(k-1)x/2}$ is such a function. Hence the probability that

$$\left\lfloor r_{k-1}\sqrt{k/n}\right\rfloor + \left\lfloor r_{k-3}\sqrt{k/n}\right\rfloor + \cdots + \left\lfloor r_2\sqrt{k/n}\right\rfloor \geqslant w$$

is at most

$$\frac{\beta^{(k-1)/2}(\frac{k-1}{2})^w}{w!} \leqslant \frac{\beta^k k^w e^w}{w^w}$$

For $w = 2e\beta k$, this probability is at most

$$\frac{\beta^k k^{2e\beta k} e^{2e\beta k}}{(2e\beta k)^{2e\beta k}} \leqslant 2^{-k} = O\left[\frac{1}{n^2}\right].$$

Hence, if $s\sqrt{k/n} - k \geqslant 2e\beta k$ (and hence if $s = \Omega(\sqrt{kn})$), the probability that $r_{k-1}+r_{k-3}+\cdots+r_2 > s$ is similarly small, which means that $r_{k-1}+r_{k-3}+\cdots+r_2 = O(\sqrt{kn})$ with high probability. By an analogous argument, $r_{k-2}+r_{k-4}+\cdots+r_3$ obeys the same bound with high probability, and the lemma follows. □

This completes the proof of Theorem 1.

4. PROOF OF THEOREM 2

In this section we sketch a proof that, for $u \leqslant 1/2$, $E[FFD(L_n^u) - \Sigma(L_n^u)] = O(1)$. The proof does not involve a surrogate like the rFF algorithm used in the last section, but deals directly with FFD, and is considerably more detailed in its analysis of the packings. We begin with the probabilistic part of the proof, encapsulated in two technical lemmas, each of which allows us to assume that in certain ways our list is close to the perfectly uniform list in which the n items have sizes $u/n, 2u/n, \ldots, nu/n$. Our first lemma is a straightforward application of standard techniques.

Lemma 4.1. For sufficiently large n, the probability that the following properties all hold simultaneously exceeds $1-1/n^2$:

(a) The number of items from L_n^u in the interval $(0, n^{-1/3}]$ is less than $2n^{2/3}/u$ (i.e., less than twice the expected number).

(b) For each k, $\lfloor 1/u \rfloor \leqslant k < n^{1/3}+1$, the number of items from L_n^u in the interval $(1/(k+1), 1/k]$ is less than $2n/(uk(k+1))$ (i.e., less than twice the expected number).

(c) For each k as above, the number of items in the interval $[1/k - 1/n^{1/2}, 1/k]$ is less than $2n^{1/2}/u$ (i.e., less than twice the expected number).

Our second lemma requires some notational preparation. For any integer $J > 0$, the *J-partition* is the division of $(0,u]$ into J consecutive non-overlapping subintervals of length u/J, which we shall call the *blocks* of the partition. Note that the number of items from L_n^u expected to fall into any block is n/J. Call a set of t blocks "bad" if the blocks contain more than $1+\epsilon$ (or less than $1-\epsilon$) times their collective expected number of items tn/J. Intuitively, the next lemma says that, with high probability, the size of the largest bad set of blocks declines exponentially as a function of the expected number of items in a block.

Lemma 4.2. Let $\epsilon = 0.01$ and $\alpha = 1+\epsilon^3$. For sufficiently large n the probability that L_n^u has the following property exceeds $1-1/n^2$:

(d) For all J, $\epsilon^{-4} \leqslant n/J \leqslant 2\log_\alpha n$, and any set of $t = \left\lfloor Je^2/\alpha^{n/J}\right\rfloor$ distinct blocks of the J-partition, the total number of items from L_n^u contained in these blocks lies between $(1-\epsilon)nt/J$ and $(1+\epsilon)nt/J$.

Proof. See Appendix A2. □

We are now finished with the probabilistic part of our argument. For the remainder of the proof, we assume that L is an arbitrary list satisfying properties (a) through (d). By Lemmas 4.1 and 4.2 the probability that L_n^u is such a list exceeds $1-1/n$ for sufficiently large n, and so if we can show that $FFD(L) - \Sigma(L) \leqslant C$ for some constant C and all such L, Theorem 2 will follow (by Lemma 2.4). The assumption that $n = |L|$ is "sufficiently large" will be used extensively in what follows (and often implicitly). Precise bounds on how large is "sufficiently large" are not needed for our result (and so, due to space constraints, we will not derive them). We begin by assuming that n is sufficiently large that $\log_\alpha n > 1/\epsilon^4 = 10^8$. The fact that Property (d) is satisfied is then exploited in the following lemma.

Lemma 4.3. For all J, $10^8 \leqslant n/J \leqslant 2\log_\alpha n$,

(1) The number of blocks in the J-partition with less than $(0.99)n/J$ items from L is less than $Je^2/\alpha^{n/J}$.

(2) The number of blocks in the J-partition with more than $(1.01)n/J$ items from L is less than $Je^2/\alpha^{n/J}$.

(3) If the *deficiency* of a block of the J-partition is defined to be 0 if the number N of items from L that it contains is at least $(0.99)n/J$ and to be $(0.99)n/J - N$ otherwise, then the sum of the deficiencies of all the blocks in the J-partition is less than $ne^2/\alpha^{n/J}$.

(4) If the *excess* of a block of the J-partition is defined to be 0 if the number N of items from L that it contains is at most $(1.01)n/J$ and to be $N-(1.01)n/J$ otherwise, then the sum of the excesses of all the blocks in the J-partition is less than $(0.02)ne^2/\alpha^{n/J}$.

Proof. Properties (1) and (2) follow immediately from Property (d) with the value of $\epsilon=0.01$ substituted. Property (3) follows from Property (1) and the fact that the maximum deficiency per block is less than n/J. The argument for (4) is slightly more complicated. Choose the $t=\lfloor Je^2/\alpha^{n/J}\rfloor$ most populated blocks. By Property (2), all the excess must be contained within these blocks. Moreover, it is easy to verify that $t<J/2$ for J in the allowed range, so that, by (1), none of these t blocks can have less than $(0.99)n/J$ items from L. Finally, by Property (d), the total contents of these t blocks is at most $(1.01)tn/J$. In the worst case, all but one of the blocks have exactly $(0.99)n/J$ elements and the other has $(0.02)(t-1)n/J$ excess items. Substituting for t yields (4). □

We are now prepared to look at the FFD packing of L in detail. We shall use the following terminology. An item whose size is in the interval $(1/(k+1),1/k]$ will be called a *k-item.* Viewing the FFD packing as described in Lemma 2.2, we shall call an item *regular* if, when it was assigned to its bin in the packing, it was the first of the unpacked items remaining in L. The first item in a bin is thus by definition regular, and if that first item is a k-item, then at least the first k items in the bin must be regular. A bin whose first item is a k-item will be called a k-bin. A k-bin is *regular* if it contains k k-items; otherwise it is a *borderline* k-bin (note that there can be at most one such bin for each k).

An item that is not regular is called a *fallback* item. The size of the gap in the bin after it receives its last regular item will be called the bin's *(initial) request*, and the size of the gap after a bin receives its first fallback item will be called the bin's *residual request*. To distinguish between items from L that are in a given subinterval of $(0,u]$ and requests that are in the same subinterval, we shall say that items are *in* subintervals and that requests *hit* subintervals. Note that each request must hit to the left of, i.e., be smaller than, the last regular item in the bin generating the request, and, in particular, the initial request for a k-bin must be smaller than $1/(k+1)$.

The next lemma bounds the maximum size of a residual request. By itself it is enough to prove an $O(\log n)$ bound on $E[FFD(L_n^u)-\Sigma(L_n^u)]$, but it is also a vital technical step toward proving our desired constant bound.

Lemma 4.4. For all sufficiently large L satisfying (a) through (d) and all k, $\lfloor 1/u\rfloor\leqslant k<n^{1/3}+1$, (i) the residual request for each k-bin with an initial request exceeding $240ku(\log_\alpha n)/n$ is at most $120ku(\log_\alpha n)/n$ and (ii) the second fallback item (and hence all subsequent fallback items) in each k-bin has size at most $120ku(\log_\alpha n)/n$.

Proof. Suppose not, and consider the situation when the first violation occurred. Let K be the corresponding value of k. First, note that since this is the first violation, it must be a violation of (i) rather than (ii): If an earlier initial request from a k-bin, $k\leqslant K$, was of size exceeding $240ku(\log_\alpha n)/n$, then the residual request (and hence any subsequent fallback items) was at most $120ku(\log_\alpha n)/n$. If the initial request was $240ku(\log_\alpha n)/n$ or less, any second or later fallback item can have been no more than half this size.

Choose a value of J of the form 2^r, r integer, such that $x=n/J$ satisfies $\log_\alpha n<x\leqslant 2\log_\alpha n$. The proof of the lemma proceeds by a series of claims. The first is a simple consequence of Lemma 4.3.

Claim 4.4.1. For the given choice of J, all blocks of the J-partition contain between $(0.99)x$ and $(1.01)x$ items from L.

Proof. By Lemma 4.3, the total number of blocks with contents out of this range is $2Je^2/\alpha^{n/J}=2ne^2/x\alpha^x$. For x in the given range, this is bounded by $2e^2/(\log_\alpha n)$ and hence is less than 1 for sufficiently large n. □

Group the blocks of the J-partition into sequences of $10K$ consecutive blocks, starting with the block whose right endpoint is u and working down. This induces a less-refined partition of $(0,u]$ which we shall call the $(J/10K)$-partition by a slight abuse of notation (since the last block of this partition may be shorter than $10Ku/J$ if J is not an integral multiple of $10K$).

Claim 4.4.2. For each block of the $(J/10K)$-partition whose left endpoint is at least $120Ku(\log_\alpha n)/n$, the total number of hits received by that block by the time (i) was violated is at most $(7.6)Kx$ whereas the number of items from L that it contains is at least $(9.9)Kx$.

Proof. The second assertion follows immediately from Claim 4.4.1. For the first, consider a block $(iu/J,(i+1)u/J]$ of the J-partition containing only k-items for some $k\leqslant K$. Any regular k-bin with all its regular items from this block must have its request in the interval $(1-(kiu/J)-(ku/J),1-(kiu/J)]$, an interval that has width ku/J. These derived intervals, plus the two partial intervals derived from the two blocks of the J-partition that contain the largest k-item and the smallest k-item respectively, partition the interval $(0,1/(k+1)]$ into subintervals, each of whose length is ku/J (except possibly for the endmost subintervals, which can be shorter). A sequence of h of these intervals can contain at most $(1.01)hx/k+2$ regular hits. This is because the sequence of blocks of the J-partition from whence the intervals came each contained at most $(1.01)x$ items from L by Claim 4.4.1 for a total of $(1.01)hx$, these must go k per bin, and there are at most two straddle bins (bins that also contain k-items from neighboring blocks *not* in the sequence).

Let us now turn to the $(J/10K)$-partition and look at a particular block of this partition whose left endpoint exceeds $120Ku(\log_\alpha n)/n$. Its length is $10Ku/J$ and so for each $k\leqslant K$ it intersects no more than $(10K/k)+2$ of the subintervals corresponding to blocks containing k-items. Hence it can receive no more than $(1.01)(x/k)[(10K/k)+2]+2$ regular hits from k bins. In addition, for each $k<K$, it might receive a hit from the borderline k-bin if such a bin exists. It cannot have received any residual hits, since so far all residual hits have by hypothesis been no greater than $120Ku(\log_\alpha n)/n$. Moreover, the smallest k for which there are k-bins is at least $\lfloor 1/u\rfloor\geqslant 2$. Thus the total number of hits it can have received so far is at most

$$10.1Kx\sum_{k=2}^{K}\frac{1}{k^2}+2.02x\sum_{k=2}^{K}\frac{1}{k}+2+(K-2)$$
$$<\ 10.1Kx(.645)+2.02x\ln K+K$$

since $\sum_{k=2}^{\infty}1/k^2=(\pi^2/6)-1$ and $\sum_{k=2}^{K}1/k<\ln K$. If we now note that for $K\geqslant 2$, $\ln K<K/(2.02)$, and that $K<(.05)Kx$ for sufficiently large n (and hence sufficiently large $x=n/J$), we conclude that the total number of hits received is at most $Kx((10.1)(.645)+1.05)<(7.6)Kx$. □

Observe that, while many of the requests that hit a block of the $(J/10K)$-partition may be satisfied by items occurring in that block (which thus become the first fallback items in the bins generating the requests), some may "overflow" the block and have to be satisfied by some item in a later block. (The distance between an initial request and the item that satisfies it is the residual gap in the bin making the request.) The next claim bounds the number of such "overflow" requests.

Claim 4.4.3. For each block of the $(J/10K)$-partition whose left endpoint is at least $120Ku(\log_\alpha n)/n$, the total number of requests that had overflowed out of the block at the time (i) was violated is at most $(7.6)Kx$.

Proof. Suppose not and consider the first block to violate the claim. It received at most $(7.6)Kx$ overflow requests from the previous block, has at most $(7.6)Kx$ internal hits, and initially contained at least $(9.9)Kx$ items. If none of the requests that overflowed into the block overflow out of it, then all the requests that overflow out were generated by internal hits and hence there are at most $(7.6)Kx$ of them. If, on the other hand, some request that overflowed in also overflows out, then all the items originally in the block were used to satisfy requests that either overflowed in or were caused by internal hits. There are at most $(15.2)Kx$ such requests and $(9.9)Kx$ such items, so the overflow can be at most $(5.3)Kx$. This contradicts our assumption that the Claim was violated. □

Claim 4.4.4. At the time (i) was violated, each sequence of 4 consecutive blocks of the $(J/10K)$-partition contained entirely in the interval between $120Ku(\log_\alpha n)/n$ and $1/(K+1)$ contained at least one unpacked item.

Proof. If the sequence were completely empty, then all the at least $4\cdot 9.9Kx = 39.6Kx$ items that were initially present in the sequence of blocks by Claim 4.4.2 must have been used to satisfy requests. However, this is impossible, since by Claim 4.4.2 there were at most $4\cdot(7.6)Kx = 30.4Kx$ internal requests and by Claim 4.4.3 at most $7.6Kx$ requests overflowed into the sequence, for a total of $38Kx$. □

Now we can derive a contradiction of our supposition that (i) was violated. Since the length of a block in the $(J/10K)$-partition is $10Ku/J$, there cannot be an empty subinterval of $(120Ku(\log_\alpha n)/n, 1/(K+1)]$ with length $60Ku/J$ or greater that does not contain a sequence of 4 empty blocks from the $(J/10K)$-partition. By Claim 4.4.4, such a sequence must contain at least one item at the time (i) was first violated. Thus no empty subinterval can have length exceeding $60Ku/J$. By our choice of $x = n/J$, this forbidden length is $60Ku/J = 60Kux/n \leqslant 120Ku(\log_\alpha n)/n$. However, our violating residual request exceeds $120Ku(\log_\alpha n)/n$ and arose from a request of at least $240Ku(\log_\alpha n)/n$, thus giving rise to an empty subinterval of length at least $120Ku(\log_\alpha n)/n$ with left endpoint at least $120Ku(\log_\alpha n)/n$. This is a contradiction and the Lemma is proved. □

Let us now make a first attempt at summing the cumulative empty space in the packing, which by Lemma 2.1 equals $FFD(L)-\Sigma(L)$. For the k-bins with $k < n^{1/3}+1$, Lemma 4.4 implies that the residual gap (and hence the empty space) is always bounded by $240ku(\log_\alpha n)/n$. Lemma 4.1(b) implies that for each such k the number of k bins is less than $2n/(uk^3)$, since each such bin except the last requires k k-items. Thus the total contribution of such k-bins to the cumulative empty space in the packing is at most

$$\sum_{k=2}^{n^{1/3}+1} 480(\log_\alpha n)/k^2 \;<\; (480)(.645)(\log_\alpha n) \;=\; O(\log_\alpha n)$$

The following lemma estimates the total contribution from the remaining bins.

Lemma 4.5. The cumulative empty space in k-bins, $k \geqslant n^{1/3}$, is at most $2/u+1$.

Proof. Each such bin (except possibly the last) can have a residual gap of at most $1/n^{1/3}$ and must contain at least $n^{1/3}$ regular items. Since by Lemma 4.1(a) there are at most $2n^{2/3}/u$ such items, there are at most $2n^{1/3}/u$ such bins (except the last). Since the last gap is less than 1, the Lemma follows. □

Thus we have proved that $FFD(L)-\Sigma(L) = O(logn)$. To tighten this bound to $O(1)$, we will need a stronger version of Lemma 4.4. The current version, however, is strong enough to dispose of another class of bins: those with initial gaps of $1/n^{1/2}$ or less.

Lemma 4.6. The cumulative empty space in k-bins, $k < n^{1/3}+1$, whose initial request is bounded by $1/n^{1/2}$ is less than 1 for sufficiently large n.

Proof. Any regular k-bin with an initial request bounded by $1/n^{1/2}$ must have all its k-items larger than $1/k - 1/n^{1/2}$. By Lemma 4.1(c), there are at most $2n^{1/2}/u$ such items, so there can be at most $2n^{1/2}/ku$ such bins, for a cumulative empty space of at most $480(\log_\alpha n)/n^{1/2}$ by Lemma 4.4. Summing over all $k < n^{1/3}+1$ we obtain a bound of $O((\log_\alpha n)/n^{1/6})$, which is less than 1 for sufficiently large n. All that remains are the borderline k-bins, $k < n^{1/3}+1$, with initial gaps less than $1/n^{1/2}$. But these bins by definition end up with $O(1/n^{1/2})$ empty space, for a total of at most $O(1/n^{1/6})$, again less than 1. □

Thus for the remainder of the proof we need only consider k-bins with $k < n^{1/3}+1$ and with initial requests exceeding $1/n^{1/2}$. A key point to observe is that these initial requests are much larger than the size of the maximum residual request from a k-bin, $k < n^{1/3}+1$: By Lemma 4.1 we know that all the latter are bounded by $120ku(\log_\alpha n)/n$, which is $O((\log_\alpha n)/n^{2/3})$ for $k < n^{1/3}+1$ and hence much less than $1/n^{1/2}$ for sufficiently large n.

In order to show that the empty space is on average substantially less than $(\log_\alpha n)/n$ in the remaining bins, we shall have to look at J-partitions for which $x = n/J$ is substantially less than $\log_\alpha n$. For these it is no longer true that all blocks of the J-partition contain between $(0.99)x$ and $(1.01)x$ items, and we will need the full power of Lemma 4.3 to deal with the excesses and deficiencies.

Lemma 4.7. Let k be such that $\lfloor 1/u \rfloor \leqslant k < n^{1/3}+1$ and J be such that $10^8 \leqslant x = n/J \leqslant 2\log_\alpha n$. Define the $(J/10k)$-partition as in Lemma 4.4, with each block corresponding to a sequence of $10k$ blocks from the J-partition. Let R consist of those blocks of the $(J/10k)$-partition that are entirely contained in the interval $(1/n^{1/2} - 60kux/n, u]$ and R' be the subset of R that in addition contains no blocks with right endpoint exceeding $1/(k+1)$. Then the number of maximal sequences of 4 or more consecutive blocks in R' that are empty after the last k-bin is packed is at most $(.65)e^2 n/(kx\alpha^x)$.

Proof. For each block of the J-partition that is in excess, arbitrarily choose $(1.01)x$ of the items in the block to be "normal," with the rest labelled "excess." Any request from a bin containing a regular item labelled "excess" will be called an *excess* hit; all other hits will be called *normal*. Note that, by the argument following Lemma 4.6 and the fact that $60kux/n = O((\log_\alpha n)/n^{2/3})$, none of the blocks in R ever receives a residual hit so long as n is sufficiently large, and hence each such block receives a maximum of $(7.6)kx$ normal hits by the time the last k-bin is packed, by the same argument that was used for the proof of Claim 4.4.2.

For each block B in R, let X_B be the number of excess hits in B plus the deficiency of B. (The deficiency is defined to be 0 if the number N of items in B is at least $9.9kx$; otherwise it is $9.9kx - N$.) Let I_B be the number of requests overflowing into B and O_B be the number of requests overflowing out of B.

Claim 4.7.1.

(a) If B is empty after all the k-bins have been packed, then $O_B \leqslant I_B + X_B - (2.3)kx$;
otherwise $O_B \leqslant (7.6)kx + X_B$.

(b) If $I_B + X_B < (2.3)kx$, then B will still be non-empty when the last k-bin is packed.

Proof. The arguments are analogous to those of Claim 4.4.3. □

Claim 4.7.2. Let S be a maximal sequence of 4 or more consecutive empty blocks from R' as prescribed above, such that S is not the rightmost such sequence, and let C be the set of 5 blocks consisting of the rightmost 4 blocks of S and the (non-empty) block immediately to the right of S. Then $\sum_{B\in C} X_B \geqslant (1.6)kx$.

Proof. Label the bins from left to right as $B[1],\ldots,B[5]$. Since $B[5]$ is non-empty, at most $(7.6)kx + X_{B[5]}$ requests overflow into $B[4]$ by Claim 4.7.1(a). Since $B[2]$ through $B[4]$ are empty, the overflow into $B[1]$ is thus at most $(7.6)kx + \sum_{i=2}^{5} X_{B[i]} - (3)(2.3)kx$. Thus, by Claim 4.7.1(b) and the fact that $B[1]$ is empty after the last k-bin is packed, we must have $\sum_{i=1}^{5} X_{B[i]} \geqslant (1.6)kx$, as desired. □

Claim 4.7.3. Let S be the rightmost maximal sequence of 4 or more empty blocks from R' as prescribed above, and let C be the set of the rightmost 4 blocks of S plus all blocks to the right of S in R. Then $\sum_{B\in C} X_B \geqslant (1.6)kx$.

Proof. If the block to the right of S is non-empty after the last k-bin has been packed, then we can proceed as for Claim 4.7.2. However, if the rightmost block of S is also the rightmost block of R', then the block to the right might even have been empty when the last $(k-1)$-bin was packed. Nevertheless, by an inductive application of Claim 4.7.1(a), it is not difficult to see that the overflow from this block into S can be at most $(7.6)kx$ plus the sum of X_B for all the blocks B to the right of S, and so an argument analogous to that for Claim 4.7.2 still suffices. □

As a consequence of Claims 4.7.2 and 4.7.3, we know that if there are M maximal sequences as specified in the lemma, then the sum of the deficiencies and excesses for all the blocks of the J-partition must be at least $(1.6)kxM$. By parts (3) and (4) of Lemma 4.3, this means that

$$M \leqslant \frac{(1.02)ne^2/\alpha^x}{(1.6)kx} < \frac{(.65)ne^2}{kx\alpha^x}$$

as desired, completing the proof of Lemma 4.7. □

We are now ready to begin summing the empty space in the k-bins with initial requests exceeding $1/n^{1/2}$. To simplify matters, we shall use the following upper bound on empty space. For a k-bin b, let $gap(b)$ be the distance from the initial hit for b to the first item remaining to its left after all k-bins have been packed. Note that this is at least as large as the residual gap for b and hence is an upper bound on the empty space in b at the end of the FFD packing.

Lemma 4.8. For each integer J such that $10^8 \leqslant x = n/J \leqslant 2\log_\alpha n$, the total number of k-bins b with initial request exceeding $1/n^{1/2}$ and with gap(b) at least $60kux/n$ but no more than $120kux/n$ is at most

$$\min\left\{\frac{40ne^2}{\alpha^x}, \frac{2n}{uk^3}\right\}$$

Proof. The existence of such a k-bin b implies the existence of a sequence of 4 or more empty blocks from the set R' defined in Lemma 4.7. However, each such maximal sequence can account for no more than $(8)(7.6)kx$ normal k-bins b with gap(b) in the stated range. This is because they all must arise from requests that hit to the right of the leftmost 4 empty blocks but not to the right of the right of the leftmost 12 (if the sequence is that long). Multiplying by the total number of maximal sequences of 4 or more empty blocks (from Lemma 4.7), we obtain a bound of $(39.6)ne^2/\alpha^x$ on the number of normal k-bins b with gap(b) in the stated range. Even if all the excess hits gave rise to k-bins with gaps in this range, this could add at most $(0.02)ne^2/\alpha^x$ to the total by part (4) of Lemma 4.3, and since there is only one borderline k-bin, we see that $40ne^2/\alpha^x$ is an upper bound on the total number of k-bins b with gap(b) in the stated range, as claimed.

For the other upper bound, simply observe that by Lemma 4.1(b) there are at most $2n/(uk(k+1))$ k-items, hence at most $2n/(uk^3)$ k-bins in total. □

Lemma 4.9. The total empty space in k-bins with initial request exceeding $1/n^{1/2}$ is $O((\ln k)/k^2)$.

Proof. Let r_0 be the maximum integer r such that $n/2^r \geqslant 10^8$ and let $x_0 = n/2^{r_0}$. For each $h > 0$, let $x_h = 2^h x_0$ and let H be the maximum value of h such that $x_h \leqslant 2\log_\alpha n$. By the proof of Lemma 4.4 (see the paragraph after Claim 4.4.4), none of the bins under consideration has empty space exceeding $60kux_H/n$. We thus can partition the k-bins into those having gaps in the following regions: $(0,60kux_0/n]$, $(60kux_0/n,60kux_1/n]$, $(60kux_1/n,60kux_2/n]$, ..., $(60kux_{H-1}/n,60kux_H/n]$. The first set might contain all the k-bins, but we can use Lemma 4.8 to bound all the other classes, thus obtaining the following overall bound on empty space for the k-bins under consideration:

$$\frac{2n}{uk^3}\cdot\frac{60kux_0}{n} + \sum_{h=0}^{H-1}\left[\min\left\{\frac{40ne^2}{\alpha^{x_h}}, \frac{2n}{uk^3}\right\}\cdot\frac{120kux_h}{n}\right]$$

If we let h' be such that $x_h \leqslant 3\log_\alpha k$ for all $h < h'$ and $x_{h'} > 3\log_\alpha k$, this can be rewritten as

$$\frac{120x_0}{k^2} + \sum_{h=0}^{h'-1}\frac{240ux_h}{k^2} + \sum_{h=h'}^{H-1}\frac{4{,}800e^2kux_h}{\alpha^{x_h}}$$

Taking into account the effect of the bounds on x_h and the fact that the values of x_h form a geometric series, this can be bounded by

$$\frac{240\cdot 10^8}{k^2} + \frac{240u(6\log_\alpha k)}{k^2} + \frac{4{,}800e^2u(6\log_\alpha k)}{k^2} = O\left(\frac{\ln k}{k^2}\right). \ \square$$

To complete the proof of Theorem 2, we need only sum up the last bound over all k, $\lfloor 1/u \rfloor \leqslant k < n^{1/3}+1$. Since $\sum_{k=2}^{\infty}(\ln k)/k^2$ converges, this sum is $O(1)$. Combining this bound with the bounds for the remaining bins determined by Lemmas 4.5 and 4.6, we conclude that $FFD(L) - \Sigma(L) = O(1)$. Hence L_n^u obeys this bound with probability at least $1-1/n$ for n sufficiently large, and, by Lemma 2.1, Theorem 2 is proved. □

5. FURTHER RESULTS

As mentioned in the introduction, Theorem 2 is tight in the sense that, for any $u > 1/2$, $E[FFD(L_n^u) - \Sigma(L_n^u)]$ is *not* $O(1)$; we can prove that it is in fact $\Omega(n^{1/3})$ for all such u. On the other hand, we have a tentative proof that the expected excess for FFD can be no worse than this if $1/2 < u < 1$, and we have a modification to FFD which we believe *will* preserve the bounded nature of the expected excess for all $u < 1$. (This is not possible for $u = 1$ because, as Lueker showed [15], even the *optimal* packing has expected waste $\Theta(n^{1/2})$ when $u = 1$.)

Table I gives our conjectured complete classification of the expected excess over $\Sigma(L_n^u)$ for both $FFD(L_n^u)$ and $OPT(L_n^u)$. The constants implicit in the Θ-notation depend on u. The results in row 1 are from [15], those in row 3 follow from Theorem 2, and we hope to include those in row 2 in the final version of this paper.

	$E[OPT(L_n^u) - \Sigma(L_n^u)]$	$E[FFD(L_n^u) - \Sigma(L_n^u)]$
$u = 1$	$\Theta(n^{1/2})$	$\Theta(n^{1/2})$
$\frac{1}{2} < u < 1$	$\Theta(1)$	$\Theta(n^{1/3})$
$0 < u \leqslant \frac{1}{2}$	$\Theta(1)$	$\Theta(1)$

TABLE I. Expected excess for OPT and FFD.

The analogous classification for the algorithm FF is not yet known for $u < 1$. Simulations, even for lists of length 128,000, do not give a clear picture [1]. They do suggest, however, that the asympotic expected behavior of FF applied to L_n^u is not a monotonic function of u, a phenonenon that has already been observed for NEXT FIT [11]. There may even be values of u for which the expected excess grows linearly with n. As to the actual behavior when $u = 1$, this appears to be even better than indicated by Theorem 1. The simulations suggest that the expected excess actually grows roughly as $n^{0.7}$ rather than as $n^{0.8}$, and even this may be an over-estimate due to insufficient data: we believe that our techniques can be extended to show a bound of $O(n^{2/3})$.

We can also make a more detailed comparison between Theorem 2 and the experimental data. If one keeps track of the numbers involved in the proof of Theorem 2, the actual bound proved on the expected value of $FFD(L_n^{1/2}) - \Sigma(L_n^{1/2})$ is ludicrously large, at least as big as 10^{10}. Moreover this bound only holds for "sufficiently large n," thus leaving open the possibility that even larger bounds might be necessary if we wished our result to hold for all n, including the "small" ones less than, say, e^{100}. Fortunately, the worst-case nature of many of our arguments appears to have led us to a slight over-estimate of the expected excess. The extensive simulations of [1], involving lists with n as large as 128,000, indicate that the actual expected excess is about 0.7. This was the average for each of the values of n tried. Moreover, no excess exceeding 1.3 was ever encountered, and the excess was less than 1 (and hence the packing was optimal) roughly 75 percent of the time. Although our hybrid probabilistic/worst-case proof techniques would be incapable of proving a bound as small as 0.7, the fact that there *is* a constant bound provides confidence in these empirical estimates.

Let us turn now to the question of distributions other than the ones we have mentioned so far. Karmarkar et al. [13] have proved that the bound of $O(n^{1/2})$ on the expected excess for FFD packings holds for *any* distribution of item sizes that is symmetric about 1/2. Our result that the expected excess for FF packings is $O(n^{4/5})$ also extends to all such distributions, by a simple modification of the proof. (Instead of dividing the interval (0,1] into k subintervals of equal length for the purpose of classifying items, divide it into subintervals of equal *probability*. Points in (0,1] that have non-zero probability can be viewed as having infinitesimal width and divided appropriately, with items of that size assigned randomly to the relevant subintervals.)

A related question concerns the expected excess when the items are chosen uniformly from the interval $(b,u]$, for values of $b > 0$. In a recent paper by Lueker [16], this question was partially answered for

OPT. For FFD, the set of intervals $(b,u]$, $0 < b < u$, which yield sublinear expected excess is easy to characterize mathematically: only intervals that are symmetric about 1/2 qualify. Otherwise, the growth is always linear with a constant of proportionality exceeding 1. The interesting problem is thus to determine these constants of proportionality. In a sense, we have solved this problem. We have a simple program which, given b and u, outputs the corresponding constant. As yet, however, we have not run our program on enough b's and u's to detect a meaningful pattern in the results.

We can similarly estimate the expected excess for FFD when items are uniformly distributed among the discrete values $1/k, 2/k, \ldots, j/k$, for any positive integers $j \leqslant k$, but here are unable to give a concise classification of those j and k which yield sublinear expected excess. (One observation we can make is that if $k = m!+1$ for some $m > 4$, then there will be values of $j < k/2$ for which the expected excess will be linear, something that cannot happen with the *continuous* uniform distribution for $(0,j/k]$.)

Finally, we might ask about the expected behavior of the related Best Fit and Best Fit Decreasing algorithms, which have the same worst-case behavior as FF and FFD [9,10]. (In Best Fit, each item x is placed in the bin which, among all those bins with sufficient room, has the smallest gap.) Our proof techniques do not obviously extend, but similar results seem to hold for these algorithms. Our preliminary simulations indicate that for L_n^u, $u \leqslant 1/2$, the average unused space for BFD is the same as that for FFD, and that, for L_n^1, the average unused space for BF may grow slightly slower than that for FF, the empirical estimate of the growth rate being roughly $O(n^{0.6})$.

REFERENCES

1. J. L. BENTLEY, D. S. JOHNSON, T. LEIGHTON, AND C. C. MCGEOCH, "An experimental study of bin packing," *in Proc. 21st Ann. Allerton Conf. on Communication, Control, and Computing*, University of Illinois, Urbana, Ill., 1983, 51-60.
2. D. J. BROWN, "A lower bound for on-line one-dimensional bin packing algorithms," Report No. R-864, Coordinated Science Laboratory, University of Illinois, Urbana, IL, 1979.
3. E. G. COFFMAN, JR, M. R. GAREY, AND D. S. JOHNSON, "Approximation algorithms for bin-packing — An updated survey," to appear (copies are available from the authors).
4. E. G. COFFMAN, JR, M. HOFRI, K. SO, AND A. C. YAO, "A stochastic model of bin packing," *Information and Control* **44** (1980), 105-115.
5. W. FELLER, *An Introduction to Probability Theory and its Applications, Volume I*, John Wiley & Sons, New York, 1968, 72-73.
6. W. FERNANDEZ DE LA VEGA AND G. S. LUEKER, "Bin packing can be solved within $1+\epsilon$ in linear time," *Combinatorica* **1** (1981), 312-320.
7. G. N. FREDERICKSON, "Probabilistic analysis for simple one- and two-dimensional bin packing algorithms," *Inform. Process. Lett.* **11** (1980), 156-161.
8. M. R. GAREY AND D. S. JOHNSON, "Approximation algorithms for bin packing problems — A survey," *in Analysis and Design of Algorithms in Combinatorial Optimization*, G. Ausiello and M. Lucertini (eds.), Springer-Verlag, New York, 1981, 147-172.
9. D. S. JOHNSON, "Near-Optimal Bin Packing Algorithms," Doctoral Dissertation, Department of Mathematics, Massachussetts Institute of Technology, Cambridge, MA, 1973.
10. D. S. JOHNSON, A. DEMERS, J. D. ULLMAN, M. R. GAREY, AND R. L. GRAHAM, "Worst-case performance bounds for simple one-dimensional packing algorithms," *SIAM J. Comput.* **3** (1974), 299-325.
11. N. KARMARKAR, "Probabilistic analysis of some bin-packing alogorithms," *in* "Proceedings 23rd Ann. Symp. on Foundations of Computer Science," IEEE Computer Society, Los Angeles, Calif., 1982, 107-111.
12. N. KARMARKAR AND R. M. KARP, "An efficient approximation scheme for the one-dimensional bin packing problem," *in* "Proceedings 23rd Ann. Symp. on Foundations of Computer Science," IEEE Computer Society, Los Angeles, Calif., 1982, 312-320.
13. N. KARMARKAR, R. M. KARP, G. S. LUEKER, AND F. MURGOLO, "Bin packing with symmetric or decreasing distributions," (tentative title), in preparation.
14. D. E. KNUTH, Lecture delivered in the course Computer Science 255, Stanford University, Spring Quarter, 1980.
15. G. S. LUEKER, "An average-case analysis of bin packing with uniformly distributed item sizes," Report No. 181, Department of Information and Computer Science, University of California, Irvine, CA, 1982.
16. G. S. LUEKER, "Bin packing with items uniformly distributed over intervals $[a,b]$," *in* "Proceedings 24th Ann. Symp. on Foundations of Computer Science," IEEE Computer Society, Los Angeles, Calif., 1983, 289-297.
17. H. L. ONG, M. J. MAGAZINE, AND T. S. WEE, "Probabilistic analysis of bin packing heuristics," *Operations Res.*, to appear.

Appendix A1: Proof of Lemma 3.1

The proof is by induction and proceeds in tandem for the two conclusions of the Lemma. The basis is trivially established for $|L| = 1$ and we henceforth assume that both conclusions hold for all lists of length less than $|L|$. The proof is argued in two parts, one for each of the two conclusions.

Part I: If L^s is obtained from L by reversing the order of two adjacent items $x < y$ from largest first to smallest first, then $rFF(L) \leqslant rFF(L^s)$

Proof. Suppose $rFF(L) > rFF(L^s)$. Let B be the contents of the first bin in of the packing $P_{rFF}(L)$ of L by rFF, and let B' be the contents of the first bin of $P_{rFF}(L^s)$.

Without loss of generality we may assume that x appears in B'. If not, then neither does y, which follows it in L^s and is larger, nor do either of them appear in B. Hence we must have $B = B'$ and, by Lemma 2.3, a shorter counterexample could be obtained by deleting B from both L and L^s, violating the induction hypothesis.

We can also assume that y appears in B. If not, then the first item in L must be some item z different from both x and y. This z must be the first item in both B and B', and since x went with it in B', the only item that could have prevented x from going with it in B would be the only item ahead of x in L but not ahead of x in L^s, i.e., y. Since we are supposing that y does not appear in B, this means that x does appear there. Hence again $B = B'$, and again we can obtain a smaller counterexample, contrary to the induction hypothesis.

There are now two subcases to consider, depending on whether y is the first item in L. Let us first consider the case when it is not, and so some other item z is first in L and L^s and hence first in both B and B'. Let $L_1 = L - \{z,y\}$ and $L_1^+ = L^s - \{z,x\}$. Note that L_1^+ differs from L_1 only in that an element x has been replaced by a larger element y. Thus, by the induction hypothesis that the second conclusion of the Lemma holds for shorter lists, we know that $rFF(L_1) \leqslant rFF(L_1^+)$. However, by Lemma 2.3, $rFF(L) = rFF(L_1)+1$ and $rFF(L^s) = rFF(L_1^+)+1$, and hence $rFF(L) \leqslant rFF(L^s)$, contrary to assumption.

The second subcase is the one in which y *is* the first item in L, and hence is the first item in B while x is the first item in B'. If $x+y \leqslant 1$, then $B = B' = \{x,y\}$ and the two packings are identical except for the ordering of the items in their first bins, contradicting the assumption that one packing used more bins. If $x+y > 1$, then the bottom item in the second bin of $P_{rFF}(L)$ is x and the bottom item in the second bin of $P_{rFF}(L^s)$ is y. Let z be the first item in L such that $x+z \leqslant 1$. If no such item exists, then x and y remain as singletons in both packings, which are hence identical up to a reordering of their first two bins, again contradicting the assumption than one packing used more bins. If $y+z \leqslant 1$ we have the same situation as in the first subcase, and can obtain a smaller counterexample by removing $\{z,y\}$ from L and $\{z,x\}$ from L^s, thus violating the induction hypothesis. If $y+z > 1$, then z is paired with x in both packings, and so the packings are identical except that the order of the first two bins is reversed, contradicting the assumption that one packing used more bins. □

Part II: If L^+ is obtained from L by replacing an item x by a larger item y, then $rFF(L) \leqslant rFF(L^+)$

Proof. Suppose $rFF(L) > rFF(L^+)$. Let B be the first bin of $P_{rFF}(L)$ and let B' be the first bin of $P_{rFF}(L^+)$. As in Part I we may assume that x appears in B, as otherwise we would have $B = B'$ and a shorter counterexample could be derived by deleting B from the two lists. Our argument is once again divided into subcases depending on whether x is the first or second element in B.

Suppose that x is the second element in B. We may assume that y is not the second element of B', since if it were the two packings would be identical except for the contents of their first bins. Let z be the first item in B (and hence in B' also). First assume that some item w is paired with with z in B'. Note that x must precede w in L as both items fit with z and yet x was preferred. Thus y precedes w in L^+, and so y and all items between y and w exceed $1-z$ and hence must exceed w. Hence we can repeatedly apply the induction hypothesis to obtain the following series of inequalities:

$$\begin{aligned} rFF(L) &= 1 + rFF(L - \{x,z\}) \\ &\leqslant 1 + rFF(L - \{x,z\} \textit{ with } w \textit{ moved forward to } x\textit{'s old position}) \\ &= 1 + rFF(L^+ - \{y,z\} \textit{ with } w \textit{ in } y\textit{'s old position}) \\ &\leqslant 1 + rFF(L^+ - \{y,z\} \textit{ with } w \textit{ in } y\textit{'s old position and increased to } y\textit{'s size}) \\ &= 1 + rFF(L^+ - \{w,z\}) = rFF(L^+), \end{aligned}$$

thus violating our assumption that $rFF(L) > rFF(L^+)$.

Now suppose no item is paired with z in B', but that there is an item v paired with y in some bin of $P_{rFF}(L^+)$. In this case we use the following applications of the induction hypothesis:

$$\begin{aligned} rFF(L) &= 1 + rFF(L - \{x,z\}) \\ &\leqslant 1 + rFF(L - \{x,z\} \textit{ with } v \textit{ increased to } 1) \\ &= 1 + rFF(L^+ - \{y,z\} \textit{ with } v \textit{ increased to } 1) \\ &= 2 + rFF(L^+ - \{y,z,v\}) = rFF(L^+), \end{aligned}$$

contrary to our assumption that $rFF(L) > rFF(L^+)$.

Finally, suppose that no item is paired with either z or y in $P_{rFF}(L^+)$. Then the two packings are the same except that $P_{rFF}(L^+)$ has two singleton bins (containing z and y) whereas $P_{rFF}(L)$ had a single bin containing both x and z. This implies that the latter packing has fewer bins than the former, again a contradiction. This concludes the subcase where x is the second item in B.

Now suppose x is the first item in B. Then x was the first item in L and so y is the first item in L^+ and hence the first item in B'. If x is the only item in B, then since y is bigger, it must be the only item in B', and so the two packings are the same except for the contents of their first bins, contradicting our assumption that one packing used more bins. Hence let z be the mate of x in B. Now suppose y has mate w in $P_{rFF}(L^+)$. We may assume that $w \neq z$, for otherwise the two packings would be identical except for their first bins. Since $x + w < y + w \leqslant 1$, w would have gone in B had not z preceded it in L. The fact that z, which precedes w in L^+, did not go in B' means that z and every item between z and w must be bigger than $1-y$ and hence bigger than w. We thus can use the induction hypotheses to derive the following contradiction:

$$\begin{aligned} rFF(L) &= 1 + rFF(L - \{x,z\}) \\ &\leqslant 1 + rFF(L - \{x,z\} \textit{ with } w \textit{ in } z\textit{'s old position}) \\ &\leqslant 1 + rFF(L - \{x,z\} \textit{ with } w \textit{ in } z\textit{'s old position and increased to } z\textit{'s size}) \\ &= 1 + rFF(L - \{x,w\}) \\ &= 1 + rFF(L^+ - \{y,w\}) = rFF(L^+). \end{aligned}$$

If y has no mate in $P_{rFF}(L^+)$, suppose z has a mate v. We then can use the induction hypotheses to derive the following contradiction:

$$\begin{aligned} rFF(L) &= 1 + rFF(L - \{x,z\}) \\ &\leqslant 1 + rFF(L - \{x,z\} \textit{ with } v \textit{ increased to } 1) \\ &= 1 + rFF(L^+ - \{y,z\} \textit{ with } v \textit{ increased to } 1) \\ &= 2 + rFF(L^+ - \{y,z,v\}) = rFF(L^+). \end{aligned}$$

Finally, suppose neither y nor z has a mate in $P_{rFF}(L^+)$. Then the two packings are the same except that y and z use up two bins in $P_{rFF}(L^+)$ and x and z use up just one in $P_{rFF}(L)$, meaning that the latter uses fewer bins than the former, again a contradiction, the final one needed for this proof. □

Appendix A2: Proof of Lemma 4.2

Throughout this proof we will have frequent recourse to a set of standard approximations, proved by straightforward application of the definitions, Stirling's formula, or the appropriate Taylor Series expansions, and summarized below as "Facts" (a) through (e).

(a) For integers $a > b > 0$, $\binom{a}{b} \leqslant \frac{a^b}{b!}$.

(b) If $z \geqslant 1$, $z! \geqslant z^z/e^z$.

(c) For all real z, $e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots \geqslant 1 + z$.

(d) If $0 \leqslant z \leqslant 1$, $1 + z \geqslant e^{z/2}$.

(e) If $-0.1 < z < 0.1$, then
$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots \geqslant z - \frac{2z^2}{3}.$$

We now begin our proof. Let ϵ, α, J, and $t = \left\lfloor Je^2/\alpha^{n/J} \right\rfloor$ be as in the statement of the Lemma to be proved, and let $x = n/J$, which means that t can be rewritten as $\left\lfloor ne^2/(x\alpha^x) \right\rfloor$. We may assume that $t \geqslant 1$, for if $t = 0$, the lemma holds trivially.

Claim 1. The probability P that there exist t blocks of the J-partition of $(0,u]$ that contain precisely $s = \lfloor (1+\epsilon)xt \rfloor$ items from L_n^u is less than $1/n^4$ for sufficently large n, as is the probability that there exist t blocks that contain precisely $s = \lceil (1-\epsilon)xt \rceil$ items.

Proof. To simplify our arguments, we shall substitute $(1+\epsilon)xt$ and $(1-\epsilon)xt$ for the two quantities in the Claim's statement. Since by the Lemma's hypothesis $x \geqslant 10^8$ and by assumption $t \geqslant 1$, the substituted values can be off by a factor of at most 1.00000001, and the reader may readily verify that this is not enough to invalidate our conclusions. Hence the number of items we are concerned with can be written as $s = (1+\delta)xt$, where δ is either ϵ or $-\epsilon$.

The probability that t blocks contain precisely s items is simply

$$P = \frac{\binom{J}{t}\binom{n}{s} t^s (J-t)^{n-s}}{J^n} \leqslant \frac{J^t n^s t^s (1-t/J)^{n-s}}{t!s!J^s}$$

Substituting for J and applying Facts (b) and (c) we obtain

$$P \leqslant \frac{n^t x^s e^t e^s t^s e^{-tx(n-s)/n}}{x^t t^t s^s} = \frac{n^t (xt)^s e^{t+s-tx+tsx/n}}{(xt)^t s^s}.$$

Substituting for s we obtain

$$P \leqslant \frac{n^t e^{t+\delta xt+(1+\delta)(xt)^2/n}}{(xt)^t (1+\delta)^{(1+\delta)xt}} = \left[\frac{ne}{xt} \cdot \frac{e^{\delta x+(1+\delta)x^2 t/n}}{(1+\delta)^{(1+\delta)x}}\right]^t.$$

We now substitute for all except the outermost occurrence of t. In inner occurrences of t except the first, an upper bound will suffice, so we use $ne^2/(x\alpha^x)$. In the first occurrence, where t is in the denominator, we observe that since $t \geqslant 1$ by assumption, t must be at least half the quantity whose floor it equals, i.e., at least $ne^2/(2x\alpha^x)$. We thus obtain

$$P \leqslant \left(\frac{2\alpha^x}{e} \cdot \frac{e^{\delta x + (1+\delta) x e^2/\alpha^x}}{(1+\delta)^{(1+\delta)x}} \right)^t = \frac{2^t}{e^t} \cdot \left(\frac{\alpha e^{\delta + (1+\delta) e^2/\alpha^x}}{(1+\delta)^{1+\delta}} \right)^{xt}$$

We now observe that since $x \geqslant \epsilon^{-4}$, $\alpha = 1+\epsilon^3$, and $\delta \leqslant \epsilon = .01$,

$$\frac{(1+\delta)e^2}{\alpha^x} \leqslant \frac{(1+\delta)e^2}{(1+\epsilon^3)^{1/\epsilon^4}} \leqslant \frac{e^3}{(e^{\epsilon^3/2})^{1/\epsilon^4}} = \frac{e^3}{e^{1/(2\epsilon)}}$$

by Fact (d). For $\epsilon \leqslant .01$, this is clearly less than ϵ^3. Applying this to the numerator, along with the fact that $\alpha = 1+\epsilon^3 \leqslant e^{\epsilon^3}$ by Fact (c), and applying Fact (e) to the denominator, we obtain

$$P \leqslant \frac{2^t}{e^t} \cdot \left(\frac{e^{\delta + 2\epsilon^3}}{e^{(\delta - 2\delta^2/3)(1+\delta)}} \right)^{xt} = \frac{2^t}{e^t} \cdot \left(e^{-\frac{\delta^2}{3} + 2\epsilon^3 + 2\delta^3/3} \right)^{xt}$$

$$< (0.75)^t e^{-\epsilon^2 xt/4}$$

since $2/e < 0.75$ and $\delta^2 = \epsilon^2$, and since $\epsilon \leqslant .01$ implies that $2\epsilon^3 \pm 2\epsilon^3/3 < \epsilon^2/12$.

There are now two cases to consider. First, suppose $t \geqslant \sqrt{n}$. Then $P < (0.75)^t \leqslant (0.75)^{\sqrt{n}}$ which is clearly less than $1/n^4$ for sufficiently large n.

On the other hand, suppose $t < \sqrt{n}$. Since, as we have already observed, $t > ne^2/(2x\alpha^x)$, it follows that $x\alpha^x \geqslant e^2\sqrt{n}/2$ and hence $\ln x + x \ln \alpha \geqslant (\ln n)/2$. Using the fact that, by hypothesis, $x \leqslant 2\log_\alpha n$ and hence $\ln x \leqslant 2 \ln \ln n$ for sufficently large n, we thus obtain, for sufficiently large n,

$$x \geqslant \frac{(\ln n)/2 - 2\ln\ln n}{\ln \alpha} \geqslant \frac{\ln n}{3 \ln \alpha} \geqslant \frac{\ln n}{3\epsilon^3}$$

by Fact (c) and the fact that $\alpha = 1+\epsilon^3$. Since $t \geqslant 1$, this means that

$$P < e^{-\epsilon^2 xt/4} \leqslant e^{(-\ln n)/(12\epsilon)} = n^{-1/(12\epsilon)}$$

which is certainly less than $1/n^4$ for $\epsilon = .01$.

This completes the proof of the Claim. □

Claim 1 holds only for two particular values of s, the endpoints of the interval $[\lceil (1-\epsilon)xt \rceil, \lfloor (1+\epsilon)xt \rfloor]$. However, it is easy to verify that the probability that there exist t blocks of the J-partition that contain precisely s items from L_n^u is maximal at $s = xt$, the expected number of items that any given set of t blocks contains, and declines monotonically as s increases above xt and as s decreases below xt. (To verify this, consider the ratio of the probabilities for two consecutive values of s.) Thus the $1/n^4$ probability bound proved in the Claim also holds for all values of s outside the interval $[(1-\epsilon)xt, (1+\epsilon)xt]$.

Since there are at most n possible values of s (and hence at most n possible values of s outside this interval), this means that for each $J = n/x$, the probability that there is a set of t blocks containing more than $(1+\epsilon)xt$ or less than $(1-\epsilon)xt$ items from L_n^u is at most $n(1/n^4) = 1/n^3$. Since there are at most $n/(2\log_\alpha n)$ possible values of J that can yield $n/x \leqslant 2\log_\alpha n$ as required, this means that the overall probability is at most $n(1/n^3) = 1/n^2$ and the Lemma is proved. □