Fundamental Discrepancies Between Average-Case Analyses Under Discrete and Continuous Distributions: A Bin Packing Case Study

E.G Coffman, Jr.[1], C. Courcoubetis[1,2], M. R. Garey[1], D. S. Johnson[1]
L. A. McGeoch[3], P. W. Shor[1], R. R. Weber[4], M. Yannakakis[1]

Abstract. We consider the average case behavior of one-dimensional bin packing algorithms in the case where bins have unit capacity and item sizes are chosen according to the "discrete uniform" distribution $U\{j;k\}$, $1 \le j \le k$, where each item size in the set $\{1/k, 2/k, \ldots, j/k\}$ has probability $1/j$ of being chosen. Note that for fixed j,k the distributions $U\{mj;mk\}$ approach the continuous distribution $U(0,j/k]$ as $m \to \infty$, where in $U(0,j/k]$ the item sizes are chosen uniformly from the half-open interval $(0,j/k]$. In this paper, we show that average case behavior can differ substantially under the two types of distributions. We show that for all j,k, $j < k-1$, there exist on-line algorithms that have constant expected waste under $U\{j;k\}$, whereas no on-line algorithm can have less than $\Omega(n^{1/2})$ waste under $U(0,u]$ for any $u \le 1$. Contrariwise, although the First Fit Decreasing (off-line) algorithm has constant expected waste under $U(0,u]$ for all $u \le 1/2$, there are many combinations j,k with $j < k/2$ such that First Fit Decreasing has $\Theta(n)$ expected waste under $U\{j;k\}$. The constant of proportionality is maximized for $j = 6$ and $k = 13$, in which case the expected waste is $n/624$.

1. Introduction.

Suppose you are given items of sizes $1, 2, 3, \ldots, j$, one of each size, and are asked to pack them into bins of capacity k with as little wasted space as possible. For what values of j and k can you pack them perfectly (with no waste)? Clearly the sum of the item sizes must be divisible by k, but what other conditions must be satisfied?

Surprisingly, the divisibility constraint is not only necessary, but suffices. Readers might want to try their hand at proving this. Relatively short proofs exist, but a certain amount of ingenuity is required to find one. The exercise can serve as a warm up for the following more general (and more difficult to prove) theorem:

Theorem 1 (Perfect Packing Theorem). *For positive integers k, j, and r, with $k \ge j$, the list L of rj items, r each of the sizes 1 through j, can be packed perfectly into bins of size k if and only if the sum of the rj item sizes is a multiple of k.*

Our proof of this theorem is by induction. Using relatively simple packing strategies, we first show that we may assume that certain relations hold between j, k, and r (e.g., k and r are odd, $k(r-1)/2r < j < k/2$, and $r < k$). We then perform a complicated analysis to reduce the remaining possibilities for j,k,r to the conjunction of two subproblems: (1) a "smaller" instance of the original problem in which the new bin size k equals the old number of items r, and (2) a directly solvable instance of a related problem in which item sizes go up by 2 rather than one, and bins of two different sizes must be filled. (Full details will appear in the journal version of this paper.)

The question answered by Theorem 1 is more than just an intriguing puzzle in pure combinatorics. Our motivation to work on it came from its relevance to certain fundamental questions about the average-case analysis of algorithms.

Consider the standard bin packing problem, in which one is given a list of items $L = a_1, a_2, \cdots, a_n$, where each a_i has a positive *size* $s(a_i) \le 1$, and is asked to find a *packing* of these items into a minimum number of unit-capacity bins. Clearly we could recast Theorem 1 into this setting. In most real-world applications of bin packing, as in Theorem 1, the item sizes are drawn from some finite set. The usual average-case analysis of bin packing heuristics, however, has assumed that item sizes are chosen according to continuous probability distributions, which by their nature allow an uncountable number of possible item sizes (e.g., see [3,4,7,9,12-14,16-21]). The assumption of a continuous distribution has the advantage of sometimes simplifying the analysis, and has been justified on the grounds that

[1] AT&T Bell Laboratories, Murray Hill, NJ 07974
[2] F.O.R.T.H., Iraklion, Crete, Greece
[3] Amherst College, Amherst, MA 01002
[4] Cambridge University, Cambridge, England

continuous distributions should serve as reasonable approximations for discrete ones. It seems fair, however, to ask whether this is actually true. For example, if we let $U(0,u]$ denote the continuous uniform distribution over the interval $(0,u]$ and $U\{j;k\}$ denote the "discrete uniform" distribution in which each value i/k, $1 \le i \le j$, is chosen with equal probability $1/j$, then the limit of the distributions $U\{mj;mk\}$ as $m \to \infty$ is $U(0,j/k]$. Note, however, that combinatorial questions such as those addressed by Theorem 1 evaporate when one reaches the limit. This suggests that something important (and interesting) may in fact be lost in the transition from discrete to continuous.

In this paper, we shall show that it is not just combinatorial questions that evaporate, but that fundamental aspects of the average-case behavior of classical bin packing algorithms are completely obscured by the continuous approximation.

For example, consider the following two results. If A is an algorithm and L is a list of items, let $A(L)$ denote the number of bins used when A is applied to L, and let $s(L)$ denote the sum of the item sizes in L (an obvious lower bound on the optimal number of bins, since the bin size is 1). Then we have the following contrasting results:

Theorem 2. *For all $u < 1$, if L_n is an n-item list with item sizes drawn independently from $U(0,u]$ and A is any on-line bin packing algorithm, then* $E[A(L_n) - s(L_n)] = \Omega(n^{1/2})$.

Theorem 3. *For any pair j,k with $j \le k-2$, there is an on-line algorithm A that packs the n-item list L_n, having item sizes drawn independently from $U\{j;k\}$, so that* $E[A(L_n) - s(L_n)] = O(1)$.

In order to prove this, we borrow techniques from the proof of Theorem 1 to show that the distribution $U\{j;k\}$ lies in the interior of a specified cone in "packing space," and then apply a geometric theorem of Courcoubetis and Weber [5,6]. The nature of our construction ensures that the running times of the algorithms are polynomial in both n and k. Our proof techniques differ markedly from those associated with the continuous case (just as the results do).

As a second example, let us consider the behavior of the well-known First Fit Decreasing bin packing heuristic (FFD). This is an off-line algorithm, and in the continuous case can do better than the best possible on-line algorithms. This would not seem to be possible for the distributions $U\{j;k\}$, in light of Theorem 3, but FFD's generality, simplicity, and robustness still make it worthy of study.

FFD works as follows: We begin by sorting the items into decreasing order, so that $s(a_1) \ge s(a_2) \ge \cdots$. (This is the operation that requires off-line processing.) We then apply the "FIRST FIT" packing rule, in which the first item is placed in the first bin, and thereafter each item in turn is placed in the lowest indexed (i.e., first) bin that still has room enough for it. It was proved in [3] that for $u \le 1/2$ in the continuous case, the "expected waste" $E[FFD(L_n) - s(L_n)] = O(1)$ (the same bound as in Theorem 3 for the discrete case, and significantly better than the lower bound of $\Omega(n^{1/2})$ for on-line algorithms under the continuous uniform distribution). For $1/2 < u < 1$, FFD still beats the on-line lower bound, but not by quite as much: the expected waste grows as $\Theta(n^{1/3})$. When we turn to discrete uniform distributions, the behavior of FFD becomes significantly more erratic. A first result is the following:

Theorem 4. *(a) For any pair j,k, $j \le k$, if L_n is the n-item list with sizes drawn independently from $U\{j;k\}$, then $E[FFD(L_n)-s(L_n)]$ is either $O(1)$, $\Theta(n^{1/2})$, or $\Theta(n)$.*

(b) Waste $\Theta(n^{1/2})$ occurs whenever $j \in \{k-1,k\}$, and wastes $O(1)$ and $\Theta(n)$ both occur for specific pairs j,k with $j/k < 1/2$ and with $j/k > 1/2$, and with j arbitrarily large.

The patterns of j,k values for which rates $O(1)$ and $\Theta(n)$ occur are not at all straightforward. See Figure 1, which covers all pairs (j,k) with $j < k/2$ and $k \le 500$, with a dot placed at point (j,k) if the expected waste under $U\{j;k\}$ is $\Theta(n)$. We have omitted pairs with $k-2 \ge j \ge k/2$ since for each such pair the growth rate is the same as that for the pair $(k-j-1,k)$ to within a constant factor. All the unmarked pairs yield $O(1)$ expected waste. The growth rate $\Theta(n^{1/2})$ does not occur for any pair j,k with $k \le 1{,}000$ and $j \le k-2$, and in fact is not known to hold for *any* pair with $j \le k-2$. Our proof techniques do not rule this out, however, and they do show that the expected waste *is* $\Theta(n^{1/2})$ when $j \in \{k-1,k\}$.

The first pair with linear expected waste is (6,13), in which case the precise asymptotic growth rate is $n/624$. Observe in Figure 1 that for a given fixed k, the values of j that yield linear waste are often broken up into separate intervals. The maximum number of separate intervals in this figure is 5 (at $k = 151$), but 2 is a more typical number. For a significant number of k's, the minimum value of j that yields linear waste is about $k/3$; less frequent but similar trends occur along the lines $j = k/4$ and $j = k/5$.

Our proof of Theorem 4 does not yield any closed-form formula that might help to explain these behaviors, but it does provide us with a computational procedure that, given j,k, determines which of the three options applies and, in the case of $\Theta(n)$ waste, computes the constant of proportionality. The basic idea is to use a different sort of continuous approximation. Instead of letting the item *sizes* range continuously over $(0,1]$, we let the *numbers* of items be continuous quantities in that range, viewing the items as part of an infinitely divisible fluid rather than as discrete units. We then algorithmically pack this fluid (starting with equal amounts for each item size) into appropriately fractional numbers of bins. The "answer" for a particular pair j,k depends on how the fluid for items of size $1/k$ ends up being packed. (To prove this latter claim, we combinatori-

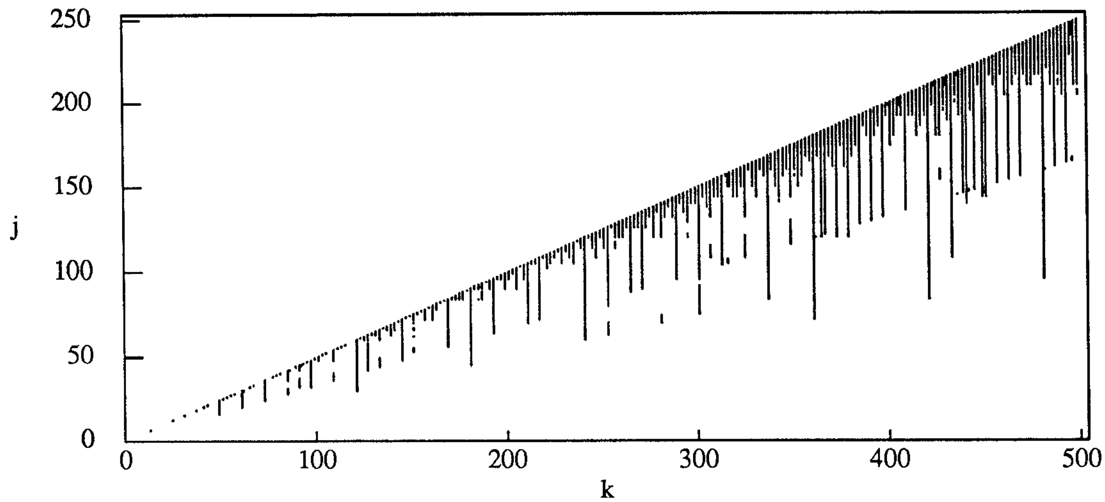

FIGURE 1. Values of $j,k, j < k \le 500$, for which expected waste is $\Theta(n)$ under $U\{j;k\}$.

ally bound the effects of departures from the "perfectly uniform" sample in which all items occur the same number of times.)

When the fluid algorithm reports that the waste for a given pair j,k grows linearly, we can determine the constant of proportionality by looking more closely at the "fractional" packing produced. This only gives us one data point, however, and one might reasonably be interested in more global questions. For instance, how bad can the linear growth rate for waste be under a discrete uniform distribution? We address this latter question by analyzing the performance of the fluid algorithm from a worst-case point of view. The analysis only becomes tractable for large k, so we augment it by actual runs of the fluid algorithm for all pairs j,k with $k \le 1{,}000$, obtaining the following results.

Theorem 5. *(a) For all pairs j,k with $k \ge j \ge 1$, the expected waste for FFD when item sizes are drawn according to $U\{j;k\}$ is no more than* $0.00614\,n/k^{1/2}$.

(b) The worst possible growth rate for expected waste of FFD over all pairs j,k is $n/624 = (.00160...)\,n$, the rate attained by the pair (6,13).

The actual constants revealed by this theorem suggest that the "linear waste" that FFD can create under these distributions may not, after all, be such a dire consequence. If one normalizes by the number of bins used, computing the expected ratio $FFD(L_n)/s(L_n)$, one gets larger, but still quite manageable constants. The worst possible value for this expected ratio is 1.00595..., again attained by the pair (6,13).

FFD is, as we have said, an off-line algorithm, and so may not be applicable in all situations. In this paper we also consider the expected behavior under discrete uniform distributions of such standard on-line algorithms as FIRST FIT (FF) and BEST FIT (BF). (BF differs from the already-described FF in that when we pack a_i, we consider not just the first bin that has room for it, but *all* bins with room, choosing the one into which a_i will fit with the least room left over.)

In the continuous case, the most extensive studies of on-line algorithms have been under the $U(0,1]$ distribution. For instance, it has been shown that any on-line algorithm must have $\Omega((n\log n)^{1/2})$ expected waste under $U(0,1]$, that the expected waste for BF is $\Theta(n^{1/2}\log^{3/4} n)$, and that the expected waste for FF is $\Omega(n^{2/3})$. Once again, different results hold for the analogous discrete uniform distributions, in this case $U\{j;k\}$ for $j \in \{k-1,k\}$. For any fixed k, both BF and FF have expected waste $\Theta(n^{1/2})$ under $U\{k-1;k\}$ and $U\{k;k\}$. Our results actually provide more detail than this, by showing the dependence of the expected waste on k as well as n. (Straightforward adaptions of arguments for the continuous case show that the off-line FFD algorithm yields $\Theta(n^{1/2})$ expected wasted space, independent of k, as does the optimal packing.) The following theorem indicates that, as in the continuous $U(0,1]$ case, there is a penalty for imposing an on-line restriction.

Theorem 6. *The expected wasted space under the distributions $U\{k-1;k\}$ and $U\{k;k\}$ is (a) $\Omega((n\log k)^{1/2})$ for any on-line algorithm, (b) $O(n^{1/2}\log k)$ for BF, and (c) $\Theta((nk)^{1/2})$ for FF.*

As with the results of [22] for the continuous case, our proofs model the bin packing process by 2-dimensional matching problems. For (a) and (b), we use discretized versions of techniques from [1,13,22]. The tight bounds for FF in (c) require a new matching formulation, and for the upper bound, a transformation to an equivalent particle system. As an unexpected bonus, we can adapt this upper bound argument to the continuous case and close the long-existing gap between the upper and lower bounds on

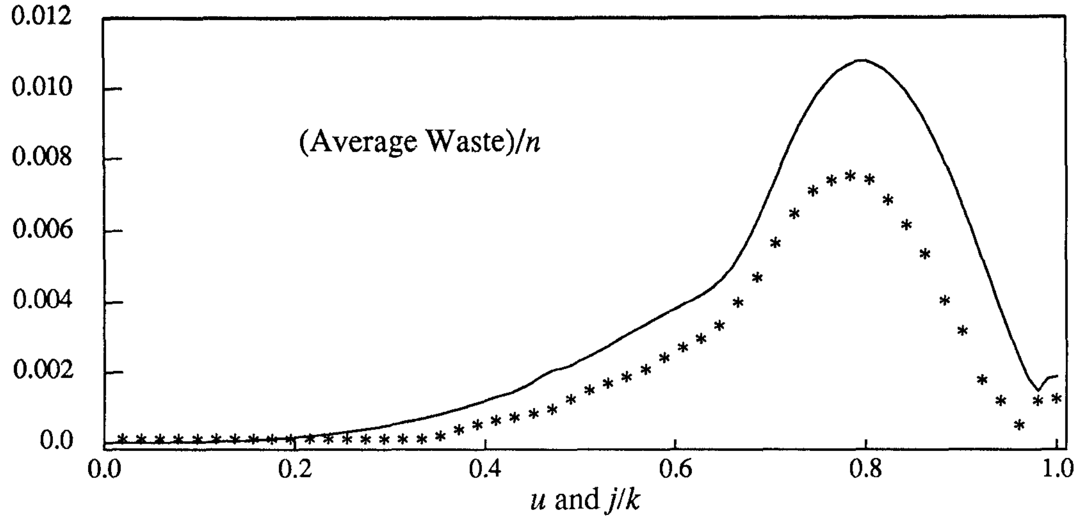

FIGURE 2. Average waste for BEST FIT as a function of the maximum item size under $U(0,u]$ (the smooth line) and $U\{j;51\}$ (the individual points).

expected waste for FF under $U(0,1]$. For five years the best upper bound we have had has been $O(n^{2/3}\log^{1/2} n)$, as opposed to the abovementioned lower bound of $\Omega(n^{2/3})$. We show that if we set $k = n^{1/3}$ in the upper bound argument for (c) above, we can adapt our proof to the continuous case, and obtain the following:

Theorem 7. *The expected waste for FF under* $U(0,1]$ *grows as* $\Theta(n^{2/3})$.

In addition to the results for $j \in \{k-1,k\}$, we also have partial results for smaller j. Basically, we can show that if j is small enough, both FF and BF can do as well as the theoretical on-line algorithms of Theorem 3.

Theorem 8. *If* $k \geq j(j+3)/2$, *BF has* $O(1)$ *expected waste under* $U\{j;k\}$. *If* $k \geq j^2$, *then so does FF.*

This contrasts to the situation in the continuous case, where it is hypothesized that expected waste for both BF and FF grows linearly under $U(0,u]$ for all $u < 1$ [2,8]. The proof of Theorem 8 applies a result of Hajek [11] about stable random walks by showing that certain parameters of the packing process have properly bounded drifts.

Simulations indicate that Theorem 8 is not tight: j can be larger than the bounds mentioned in the theorem and FF and BF will still have expected $O(1)$ waste. For instance, when $k = 13$, FF and BF both appear to yield $O(1)$ expected waste when $j = 6$, even though Theorem 8 only guarantees it for $j \leq 3$. Note the contrast here to our earlier observation that FFD yields linear waste under $U\{6;13\}$. It has long been known that there exist instances where FF outperforms FFD, but these were thought to be pathological. Here we appear to have found a quite reasonable probability distribution under which FF outperforms FFD *on average*!

Constant waste does not continue indefinitely for FF and BF, however. Once j crosses an as-yet-undetermined threshold, linear expected waste appears to ensue. See Figure 2. For each of the distributions $U(0,u]$, $u = .01,.02,\ldots,.99,1.00$, and $U\{j;51\}, j = 1,2,\ldots,51$, 25 sample lists of 2,048,000 items were generated and then packed by BF. The figure displays the average values of $(FF(L)-s(L))/n$ for each distribution. (Experiments with n = 10,000,000 yielded essentially the same ratios, except for $j \in \{50,51\}$, where Theorem 6 says the expected waste is $O(n^{1/2}\log k)$, and for the rightmost values for u, where sublinear factors with large constants of proportionality continue to affect the picture). In terms of the number of *units* of waste, the average for BF under $U\{j;51\}$ was less than one bin for all $j \leq 17$ (i.e., the average packing was optimal!). At j = 18, the average suddenly jumped to over 450 bins. In contrast, the average waste under $U(0,u]$ was a smoothly increasing function of u. On the other hand, once the linear-waste threshold for j had been passed in the discrete case, the dependence on j/k of the constants of proportionality appeared to track the analogous dependence on u in the continuous case (and note that the constants of proportionality were substantially larger than those we saw for FFD).

Other experiments indicate that the thresholds where linear waste begins for FF and BF have decreasing values of j/k as k increases, as suggested by Theorem 8. Thus the curves for $U\{j;k\}$ look more and more like those for $U(0,u]$ as $k \to \infty$. Also as suggested by Theorem 8, the thresholds for FF and BF may differ. For instance, simulations suggest that under $U\{18;55\}$, FF has linear expected waste while the expected waste for BF remains $O(1)$.)

It should be pointed out that no *proof* of linear waste for BF or FF is known for any distribution $U\{j;k\}$ or $U(0,u]$. Finding a proof of linear expected waste is one of the major

Expected Waste under $U(0,u]$ and $U\{j;k\}$				
	$u<1$	$j<k-1$	$u=1$	$j \in \{k-1,k\}$
Optimal Packing	$O(1)$ [3]	$O(1)$ [Thm.3]	$\Theta(n^{1/2})$ [14,16]	$\Theta(n^{1/2})$
Best Poly. Time Alg.	$O(1)$ [3]	$O(1)$ [Thm.3]	$\Theta(n^{1/2})$ [14,16]	$\Theta(n^{1/2})$
Best Poss. On-Line	$\Omega(n^{1/2})$ [Thm.2]	$O(1)$ [Thm.3]	$\Omega((n\log n)^{1/2})$ [22]	$\Omega(n^{1/2})$ [Thm.6]
Best Known On-Line	$O((n\log n)^{1/2})$ [23]	$O(1)$ [Thm.3]	$\Theta((n\log n)^{1/2})$ [23]	$\Theta(n^{1/2})$ [Thm.6]
First Fit Decreasing	$\{O(1),\Theta(n^{1/3})\}$ [3]	$\{O(1),\Theta(n^{1/2}),\Theta(n)\}$ [Thm.4,5]	$\Theta(n^{1/2})$ [14,16]	$\Theta(n^{1/2})$ [Thm.4]
Best Fit	$\Theta(n)$?	$\{O(1),\Theta(n)?, ... (?)\}$ [Thm.8]	$\Theta(n^{1/2}\log^{3/4} n)$ [22]	$\Theta(n^{1/2})$ [Thm.6]
First Fit	$\Theta(n)$?	$\{O(1),\Theta(n)?, ... (?)\}$ [Thm.8]	$\Theta(n^{2/3})$ [Thm.7]	$\Theta(n^{1/2})$ [Thm.6]

TABLE 1. Bounds on expected waste, a simplified overview.

challenges remaining in the area of expected-behavior analysis for bin packing. We suspect that it may well be easier in the discrete case, and at the end of this paper will describe some potentially useful insights that have been revealed by our simulations.

As a general guide to the contents of this paper, Table 1 lists the results known for the continuous uniform case, and contrasts them with the new results we have obtained for the discrete uniform case. Note that this table is just an overview, and so does not reflect all the details of our results as described above (such as dependencies on particular values of u or j,k). The remainder of this paper will sketch the proofs of several of the above theorems, with emphasis on the novel proof techniques needed for the discrete case. Section 2 considers the question of the best possible expected behavior for on-line algorithms under discrete and continuous distributions. First we sketch a proof of Theorem 2 for the continuous case. We then describe the geometric results of Courcoubetis and Weber [5,6], and show how to apply them using results about perfect packings to prove Theorem 3 for the discrete case. In Section 3 we consider the classical off-line algorithms FFD and BFD. We describe the "fluid algorithm" proof technique in more detail and use it to prove Theorem 4. We also sketch the proof of Theorem 5. Section 4 then briefly covers our results for the classical on-line algorithms FF and BF (Theorems 6, 7, and 8). We conclude in Section 5 with a discussion of how well our results and techniques might extend to other forms of discrete distributions.

2. Best Possible On-Line Algorithms.

In this section we consider the best possible expected behavior for on-line bin packing algorithms under $U\{j;k\}$ and $U(0,u]$ distributions. For all cases, better performance bounds are attainable in the discrete case than are mathematically possible under the analogous continuous distribution. We first describe the limits on behavior under $U(0,u]$ distributions. We should first point out that our description of these results in Section 1 (and Table 1) used "Ω" notation in the sense of Hardy and Littlewood [10], rather than of Knuth [15]. That is, we used "$\Omega(f(n))$" to mean "not $o(f(n))$," rather than "greater than $cf(n)$ for some $c > 0$ and *all* sufficiently large n."

In the case where $u = 1$, this is the best we can hope for. In [22], Shor showed that if L_n is an n-item list with item sizes drawn independently from $U(0,1]$ and A is any on-line bin packing algorithm, then $E[A(L_n) - s(L_n)] = \Omega(n^{1/2}\log^{1/2} n)$, but only in the Hardy and Littlewood sense. The stronger sense is not possible, since there is, for instance, an on-line algorithm whose waste is proportional to $n^{1/2}$ whenever n is a power of 2. (This follows from the fact that an expected difference of $O(n^{1/2})$ *is* possible for an algorithm that is on-line except that it knows n in advance [14,16].)

Theorem 2 of this paper, which covers the case when $u < 1$, suffers from a similar technical restriction. A more detailed statement of that result goes as follows.

Theorem 2*. *Suppose that $u < 1$, L_n is an n-item list with item sizes drawn independently from $U(0,u]$, and A is any on-line bin packing algorithm that does not know n in advance. Then $E[A(L_n) - s(L_n)]$ is not $o(n^{1/2})$.*

Proof. Let $w(t)$ denote the amount of empty space in partially filled bins after t items have been packed. We shall show that the average value of $w(t)$, $1 \le t \le n$ is $\Omega(n^{1/2}u^3)$ with high probability (in the "for all sufficiently large n" sense of Ω). The theorem follows.

Let $\upsilon(t)$ denote the number of non-empty bins that are filled to less than $1-u^2/8$ after the first t items have been packed. Suppose that $\upsilon(t) \ge n^{1/2}$ at least $u/8$ of the time. Then we have $w(t) \ge n^{1/2}u^2/8$ for at least a fraction $u/8$ of the time. This means that $w(t)$ must average at least $n^{1/2}u^3/64$.

On the other hand, suppose that for at least a fraction $1-u/8$ of the time, we have $\upsilon(t) \le n^{1/2}$. Now, consider the last item that we put into a bin. At time n, with high probability there must be at least $nu/2-o(n)$ bins, because the sum of the item sizes is at least that. Therefore, there must have been at least $nu/2-o(n)$ items that were the last item put into some bin. We will show that many of these items must have left large gaps. At time t, when we are about to pack item a_{t+1}, there are only $\upsilon(t)$ bins into which one can put an item larger than $u^2/8$. Therefore, if a_{t+1} is to leave a gap of less than δ in its bin, it either must have size less than $u^2/8$ or its size must be within δ of the empty space in one of these $\upsilon(t)$ bins with gaps larger than $u^2/8$. The probability of this is at most $[u^2/8+\delta\upsilon(t)]/u$. By choosing $\delta = u^2 n^{-1/2}/8$, we make this probability equal to $u/4$. This means that the fraction of items that can fill a bin to within $\delta = u^2 n^{-1/2}/8$ is with high probability at most $3u/8-o(1)$, because $\upsilon(t)$ is large (greater than $n^{1/2}$) at most $u/8$ of the time, and if $\upsilon(t)$ is small, the probability being able to fill a bin to within δ is at most $u/4$. As we have already noted, however, there are with high probability at least $nu/2-o(n)$ items that were the last item put into a bin. Thus, with high probability, there are at least $nu/8-o(n)$ bins with more than δ empty space, giving $w(n) \geq nu\delta/8 = n^{1/2}u^3/64$.

To complete our proof, we need to show that, given our assumption about $\upsilon(t)$, $w(t)$ is large on average as well as just for $t = n$. Fortunately, the argument just given can be adapted to show that the wasted space at time n' for $n/2<n'<n$ is large, because if $\upsilon(t)$ is small for $1-u/8$ fraction of the interval $[0,n]$, then for any time $n'>n/2$, $\upsilon(t)$ will be small for some slightly smaller fraction of the interval $[0,n']$. The bounds on $w(n')$ thus will go through with only a change in the hidden constants, which will suffice to prove the theorem. □

It should be noted that the above proof relies heavily on the fact that the distribution is continuous, since this is the reason why the union of $n^{1/2}$ intervals of size δ cannot cover the full probability space. Our discrete distributions $U\{j;k\}$, however, do not have this failing, and so it is possible to obtain significantly better average-case behavior for them. The key tool for realizing this improved performance is a result of Courcoubetis and Weber [5,6] that casts bin packing with discrete distributions into geometric terms.

Suppose D is a discrete distribution over a set S of d distinct item sizes, $s_1, s_2, \ldots, s_d$, and let $B > 0$. Any packing of items of these sizes into bins of size B can be viewed as a non-negative integer vector $(c_1, \ldots, c_d)$, where $\Sigma_{i=1}^{d} c_i s_i \leq B$. Of particular interest are those vectors which give rise to a sum of *exactly* B, which we shall call *perfect packing configurations*. For instance, if $S = \{1,2,3\}$ and $B = 7$, one such configuration would be $(1,0,2)$. Let $P_{S,B}$ denote the set of all perfect packing configurations for a given S and B. Let $\Lambda_{S,B}$ be the convex cone in $\mathbf{R}^d$ spanned by all nonnegative linear combinations of configurations in $P_{S,B}$.

Theorem (Courcoubetis and Weber [5]). *Let B be a bin size, D be a discrete distribution with item sizes in the finite set $S = \{s_1, \ldots, s_d\}$, and p_i be the probability of item size s_i under D, $1 \leq i \leq d$. Let L_n be a random n-item list with item sizes drawn independently according to D. Then*

(a) If $(p_1, \ldots, p_d)$ lies in the interior of $\Lambda_{S,B}$, there exists an on-line bin packing algorithm A such that $E[A(L_n) - s(L_n)] = O(1)$.

(b) If $(p_1, \ldots, p_d)$ lies on the boundary of $\Lambda_{S,B}$, then $E[OPT(L_n) - s(L_n)] = \Theta(n^{1/2})$ and there exists an on-line bin packing algorithm A such that $E[A(L_n) - s(L_n)] = \Theta(n^{1/2})$.

(c) If $(p_1, \ldots, p_d)$ lies outside of $\Lambda_{S,B}$, then $E[OPT(L_n) - s(L_n)] = \Theta(n)$.

The algorithms of cases (a) and (b) run in polynomial time. We shall sketch how they work below, after we have used the above result to prove the following extended version of Theorem 3.

Theorem 3*. *For any pair j,k of positive integers with $j \leq k$, let L_n be an n-item list L_n, having item sizes drawn independently from $U\{j;k\}$. Then*

(a) If $j < k-1$ there exists a polynomial-time on-line algorithm A such that $E[A(L_n) - s(L_n)] = O(1)$.

(b) If $j \in \{k-1,k\}$, then $E[OPT(L_n) - s(L_n)] = \Theta(n^{1/2})$ and there exists a polynomial-time on-line bin packing algorithm A such that $E[A(L_n) - s(L_n)] = \Theta(n^{1/2})$.

Proof. For this result we can simplify our notation by normalizing so that item sizes range from 1 to j and bins have capacity k; this will only affect the constants of proportionality. Let us first observe that the lower bound on $E[OPT(L_n) - s(L_n)]$ in (b) follows easily from the fact that in any packing of items of sizes 1 through $k-1$ into bins of size k, the waste is at least the excess of the number of items of size $k-1$ over those of size 1. Thus we need only concern ourselves with the on-line algorithm claims of (a) and (b). For these, by the Courcoubetis-Weber theorem, all we need show is that the j-dimensional vector $(1/j, 1/j, \ldots, 1/j)$ (or equivalently the vector $\bar{e} = (1,1,\ldots,1)$) is in the appropriate cone (strictly inside when $j < k-1$). To prove that $\bar{e}$ is in the cone, it suffices to show the following:

Lemma 2.1. *If $1 \leq j \leq k$, then for some integer $r > 0$, the set of rj items consisting of r copies each of items of size $1,2,\ldots,j$ can be packed perfectly into bins of size k.*

To prove in addition that $\bar{e}$ is strictly inside the cone when $j < k-1$, it suffices to show the following:

Lemma 2.2. *For each i,j,k with $i \leq j < k-1$, there exist positive integers r_1 and r_2 such that the set of $r_1 j + r_2$ items, consisting of r_1 items of sizes $1,2,\ldots,j$ and r_2 additional items of size i, can be packed perfectly into bins of size k.*

Lemma 2.1 follows immediately from the Perfect Packing Theorem (Theorem 1), which says that such packings exist for *all* $r > 0$ with $rj(j+1)/2$ a multiple of k. Space limitations prevent us from presenting a complete proof of Theorem 1. As we shall see, however, Lemma 2.1 also follows from Lemma 2.2, except in the case where $j \in \{k-1,k\}$. In this latter case we may assume that $j = k-1$, since items of size k clearly pack bins of size k perfectly all by themselves. There are then two cases to consider, depending on the parity of k. If k is odd, we can simply take $r = 1$ and pair up the item of size $k-i$ with the item of size i, $1 \le i < k/2$. If k is even, we take $r = 2$, pair up items as before, and then pair up the two remaining items of size $k/2$.

To see that Lemma 2.2 implies Lemma 2.1 when $j < k-1$, let $r_1(i)$ and $r_2(i)$, $1 \le i \le j$, be the constants whose existence is guaranteed by Lemma 2.2 for the triple i,j,k, and let $s = r_2(1) r_2(2) \cdots r_2(j)$. Then for each i, $sr_1(i)/r_2(i)$ items of size $1,2,...,j$ plus s additional items of size i will pack perfectly into bins of size k. The set of items formed by summing the above over i from 1 to j will pack perfectly into bins of size k, and this set will contain an equal number of items of each size. Thus it suffices to prove Lemma 2.2.

The proof is by induction on k. It is clearly true for $k=2$. Suppose Lemma 2.2 (and hence Lemma 2.1) holds for bins of size less than k. In the case $k \ge i+j$, the induction hypothesis implies that there exist r and m such that r items each of the sizes $1, \ldots, j$ will perfectly pack into m bins of size $k-i$. In this case we simply take $r_1 = r$ and $r_2 = m$. Use m items of size i to create m bins with remaining space of $k-i$ and then perfectly pack these with r items of each of the sizes $1, \ldots, j$. On the other hand, for $k < i+j$, Lemma 2.1 implies that there are r and m such that r items of each of the sizes $1, \ldots, k-j-1$ perfectly pack in m bins of size $k-i$. Take $r_1 = 2r$ (to ensure it is even) and completely fill bins with pairs of items of sizes $(j,k-j), (j-1,k-j+1), \ldots, \lceil k/2 \rceil, \lfloor k/2 \rfloor)$. Take $r_2 = 2m$ and use $2m$ items of size i to create $2m$ bins with remaining wasted-space of $k-i$. These can be completed filled by the remaining items, $2r$ of each of the sizes $1, \ldots, k-j-1$. Note that it is here we use $j < k-1$. This completes the inductive step and the proof of Lemma 2.2. As explained above, Theorem 3* follows. □

The on-line algorithm that achieves the performance of Theorem 3* in case (a) can be described as follows. (The algorithm for case (b) does not require the full Courcoubetis-Weber machinery and is left as an exercise to the reader.) Let $c_{l,i}$ denote the number of times that an item of size i is used in packing bin type l. We choose a set of c_l's that span a cone having $(1, \ldots, 1)$ in its interior. These may correspond to the packings generated in the proofs of lemmas 2.1 and 2.2. The time required to identify these packings is polynomial in j and k. We imagine that there is a separate packing facility for each type of bin c_l, and that each arriving item is routed to a packing facility, where it finds space in a partially full bin or starts a new bin of that type. As a function of the contents of partially full bins at facility l, let $e_{l,i}$ be the number of different item sizes amongst those used in bins of type l for which there are presently fewer spaces at facility l than there are for item size i. Note $e_{l,i} = 0$ for i such that $c_{l,i} = 0$. The routing probabilities are determined by finding an $\varepsilon > 0$ such that the following linear program has a feasible solution:

$$\sum_l \alpha_l (c_{l,i} + \varepsilon e_{l,i}) = 1/j, \quad i = 1, \ldots, j.$$

$$\alpha_l > 0, \quad i = 1, \ldots, j.$$

Such an ε exists because $(1,1,...,1)$ is in the interior of the cone spanned by the c_l's; a precise value for ε can be computed directly from the distance of $(1,1,...,1)$ from the boundary of the cone.

When an item of size i is to be packed, it is randomly routed for packing at facility l with probability

$$\alpha_l (c_{l,i} + \varepsilon e_{l,i}) / (1/j).$$

The impact of this construction is to ensure that facility l receives items of size i at a faster rate than items of size i' whenever there are more spaces for items of size i than i' in partially full bins of type l. For the details of why this suffices, see [5].

Note that, as described, the above algorithm is randomizing. It is possible to construct an algorithm that is not randomizing, but the analysis that shows it gives $O(1)$ wasted space is considerably more complicated [6]. It is also possible to design an on-line algorithm that does not know j and k in advance but determines them by sampling and still yields $O(1)$ waste for all distributions $U\{j;k\}$ with $j < k-1$ [5].

3. First Fit Decreasing.

In this section we discuss the expected behavior of the off-line First Fit Decreasing algorithm (FFD) under $U\{j;k\}$ distributions, providing proofs or proof sketches for the key results listed in Theorem 4. The contrast of these results with those for the analogous continuous distributions has already been pointed out in Section 1.

The key tool we have developed for treating these algorithms is what we shall call the *Fluid FFD* technique. The easiest way to explain this technique is to see it in action for a particular distribution, $U\{6;13\}$. To simplify our discussion, we shall again multiply all item- and bin-sizes by k, and so consider ourselves to be packing items of size $1,2,...,6$ into bins of size 13. Suppose we had a sample list that was *perfectly uniform*, i.e., which contained $n = 6m$ items, m of each size. Suppose further that we can take m to be any arbitrarily large number, and hence can consider it to be divisible by any integer we choose. (This is like viewing the set of items as an infinitely divisible fluid; hence the name for the technique.) Once sorted, such a list would consist of m items of size 6, followed by m items of size 5, etc. Let us now consider how FF would pack such a list:

Size-6 Items: These m items would go two per bin into $m/2$ bins, leaving $m/2$ gaps of size 1.

Size-5 Items: These m items would go two per bin into $m/2$ bins, leaving $m/2$ gaps of size 3.

Size-4 Items: These m items would go three per bin into $m/3$ bins, leaving $m/3$ additional gaps of size 1, for a total of $5m/6$.

Size-3 Items: The first $m/2$ of these items would go into the gaps of size 3 created by size-5 items. The remaining $m/2$ would go four per bin into $m/8$ bins, leaving $m/8$ additional gaps of size 1, for a total of $23m/24$.

Size-2 Items: These m items would go six per bin into $m/6$ bins, leaving $m/6$ additional gaps of size 1, for a total of $9m/8$.

Size-1 Items: These would fill m of the $9m/8$ gaps of size 1 in previous bins, leaving $m/8$ gaps unfilled.

Note that we conclude that $FFD(L) - s(L) = m/8 = n/48$, i.e., the waste is linear in the number of items. (The normalized waste, with bin size reduced from 13 to our standard of 1, is $n/624$.)

It should be clear how to generalize the Fluid FFD procedure to any other pair $j,k, j \le k$, assuming bins are of size k and there are an infinitely divisible number m of size-i items, $1 \le i \le j$. Three qualitatively different types of outcomes are possible:

Type 1. When the size-1 items come to be packed, the sum of the gap sizes exceeds m, and so not all the gaps can be filled (as was the case for the pair 6,13).

Type 2. When the size-1 items come to be packed, the sum of the gap sizes is less than m, and so all gaps will be filled and a non-zero number of bins will be constructed consisting solely of size-1 items. This happens, for instance, for the pair 6,11.

Type 3. When the size-1 items come to be packed, the total sum of gap sizes was precisely m, so that these are filled with 1-items and no 1-items are left over. This happens, for instance, for the pair 12,13.

We claim that these answers tell us all we need to know to characterize $E[FFD(L_n)-s(L_n)]$ as a function of n, for L_n drawn from $U\{j;k\}$. This despite the fact that real samples L_n would not be perfectly uniform, and in practice n tends not to be infinitely divisible. To be specific, we shall prove the following more detailed version of Theorem 4(a):

Theorem 4a*. *(a) For any pair j,k, $j \le k$, if L_n is the n-item list with sizes drawn independently from $U\{j;k\}$, then $E[FFD(L_n)-s(L_n)]$ is $\Theta(n)$ if the output of Fluid FFD is of Type 1 for j and k, is $O(1)$ if the output is of Type 2, and is $\Theta(n^{1/2})$ if the output is of Type 3.*

Proof. Let us examine the worst-case effect on an FFD packing of changing an arbitrary list by one item so that it looks a little more perfectly uniform. Suppose that in a set of n items, say I, the number of items of sizes $1,\ldots,j$ are $(m_1,\ldots,m_j)$. Consider an i such that $m_i > 0$ and let J be the set of $n-1$ items obtained by deleting an item of size i from I. We shall compare the FFD packing of J against that obtained for I. Following the packing of all items of sizes i or more, the packing for J will leave one bin less full. Now consider the packing of items of next smallest size. Some of these, but certainly not more than $k-1$, might go in the bin that is less full. This means that some other bins, which would otherwise have received items of this size, will be less full. There are at most $k-1$ such. So following the packing of these items there are at most k bins whose contents differ from those that would have been obtained by packing I.

The argument continues in this fashion. Suppose we are about to pack items of size $l < i$ and there are N bins whose contents differ from what they would have been if we had been packing I. Any such bin that is less full might take up to $k-1$ more items of size l, so there might be up to $k-1$ other bins that receive fewer items of size l than would have been the case if we had been packing I. Any bin that is more full might take up to $k-1$ fewer items of size l, so there might be up to $k-1$ other bins that receive more items of size l than would have been the case if we had been packing I. Since items of size l are identical, those bins whose contents start this phase no differently and receive exactly the same number of items of size l as they would have received if we had been packing I end this phase of packing holding exactly the same contents as if we had been packing I. Thus at the end of this phase there are no more than kN bins whose contents differ from what they would have been if we had been packing I. It follows that there are no more than k^j bins whose contents differ for packings of I and J. A similar argument holds if we imagine that J is constructed from I by adding an item.

Thus for any two size distributions $(m_1,\ldots,m_j)$ and $(m'_1,\ldots,m'_j)$ we can bound the difference in wasted space by $2k^{j+1}\sum_i |m_i - m'_i|$. Taking the expected value over $m_1,\ldots,m_j$, when these have been generated by $U\{j;k\}$, the expected difference in wasted space from that which would be obtained if there were equal numbers of items of each size is bounded by $2k^{j+1}E(\Sigma_i |[n/j]-m_i|)$. Since m_i has the binomial distribution $B(n,n/j)$ this expectation is $O(n^{1/2})$. Note also that the actual difference is $O(n^{1/2}\log n)$ with probability $1 - 1/n$.

The above analysis goes through even if we consider only the packing of the items of size greater than 1. Thus we can conclude that after we have packed all items larger than 1, the sum of the gapsizes will differ from that under Fluid FFD by $O(n^{1/2})$ in expectation, and by $O(n^{1/2}\log n)$ with probability $1 - 1/n$. The fact that n is not infinitely divisible, and may not even be divisible by j, only adds a constant factor to this difference.)

If the Fluid FFD outcome is of Type 1, this means that for large enough n the total gap size at this point remains less than cn/j for some $c < 1$ with probability $1 - 1/n$, whereas the number of size-1 items is with the same probability larger than $n/j - dn^{1/2}\log n$ for some constant d, and hence is almost surely larger than cn/j for sufficiently large n. Thus with probability $1 - 1/n$ the only bin with a gap will be the last one created, which will contain only size-1 items. Normalizing to unit-capacity bins, this means that $FFD(L_n) - s(L_n)$ is almost surely less than 1. With probability $1/n$ there may be some additional gaps, but these can contribute at most an additional 1 to the (normalized) expected value since there are at most n bins, so that $E[FFD(L_n) - s(L_n)] < 2$, a somewhat stronger result than claimed above.

If the Fluid FFD outcome is of Type 2, an analogous argument leads to the conclusion that $FFD(L_n) - s(L_n)$ is almost surely $\Theta(n)$. A simpler argument suffices if the Fluid FFD outcome is of Type 3. In this case we use the fact that the expected difference in total gap size between L_n and the perfectly uniform list is $O(n^{1/2})$ even after the size-1 items have been packed. Since the expected total gap size for the latter is $O(1)$, even when we take into account the fact that n might not be infinitely divisible, the result follows. This concludes the proof of Theorem 4a*. □

In the light of Theorem 4a*, Theorem 4(b) reduces to the following:

Theorem 4b*. *For all j,k with $j \in \{k-1,k\}$, the output of Fluid FFD is of Type 3, while outputs of Types 1 and 2 occur for arbitrarily high values of j with j/k both above and below $1/2$.*

Proof. When $j \in \{k-1,k\}$, Fluid FFD places the size-k items (if any) in bins by themselves, and thereafter pairs up the remaining items, size $k-j$ with size j, running out of items (and gaps) just as the last size-1 item goes in the last bin containing an item of size $k-1$. We thus have the claimed Type 3 outcome.

To show that Type 1 can occur for arbitrarily large j and j/k either greater than or less than $1/2$, consider the pairs j,k with $k = 1260t+1$, $j \in \{420t, 840t\}$ and arbitrarily large $t \geq 1$. We first note that if $j = 840t$, Fluid FFD will pair each of the items of size $k/2 + a$, $1/2 \leq a \leq 210t - 1/2$, with an item of size $k/2 - a$, thus forming perfectly packed bins. We will then be in the initial state for the $420t, 1260t$ pair, so we need only consider this latter pair. The key to our argument is to observe that when we go to pack the items of size $420t$, $315t$, $252t$, $210t$, and $180t$, in each case there will be no gaps left into which these items will fit, so that they will go three, four, five, six, and seven items per bin, respectively, creating a total of $m/3 + m/4 + m/5 + m/6 + m/7 > m$ gaps of size 1, yielding a Type 1 result. (The value 1260 is chosen so that no gaps precisely equal to the item sizes in question will be created using larger items.)

As to Type 2 results, we can actually prove the following somewhat stronger result than claimed above, one that yields the same conclusion for FFD as Theorem 8, mentioned in the Introduction, yields for FF.

Lemma 3.1. *For all pairs j,k with $j < k^{1/2}$ or $j > k - k^{1/2}$, Fluid FFD gives a Type 2 outcome and hence for the uniform distributions $U\{j;k\}$ based on those pairs, $E[FFD(L_n) - s(L_n)] = O(1)$.*

The proof is straightforward. As before, we need only consider the case where $j < k^{1/2}$, as the other case reduces to it. Note that for each j', $1 < j' \leq j$, the size-j' items go at least $k^{1/2}$ per bin, creating at most $m/k^{1/2}$ gaps of some specified size g. A simple induction can thus be used to show that, when we come to pack the items of size g, no gaps of size greater than g will remain, and the total number of gaps of size g itself will be at most $(m/k^{1/2})(k^{1/2} - 1) < m$, meaning that all those gaps will be filled perfectly, with some items of size g left over. Since this holds true for $g = 1$, the outcome is by definition of Type 2. This completes the proof of Theorem 4b*. □

In light of the results of our running Fluid FFD on all pairs j,k with $k \leq 1{,}000$ and $j \leq k/2$, we know that Type 2 outcomes can actually occur for much larger j than those indicated in Lemma 3.1, or any straightforward strengthening of it. Moreover, Type 1 outcomes occur for many more pairs than those specified above. A key question is how bad Type 1 behavior can be, and it is this question that is addressed by Theorem 5, with the aid of some additional information that we can obtain from Fluid FFD.

Observe that in the Type 1 case, we can actually determine the constant of proportionality on the $\Theta(n)$ waste from the result of running Fluid FFD. If the total gap size left after packing the size-1 items is cm, which normalizes to cm/k, then the (normalized and asymptotic) expected value of $E[FFD(L_n) - s(L_n)]$ is cn/jk ($n/624$ in the $6,13$ case, where $c = 1/8$). Noting that the normalized value of $s(L_n)$ is almost surely $(n/j)(j(j+1)/2)/k + o(n) = n(j+1)/2k + o(n)$, we can further conclude that $\lim_{n\to\infty} E[FFD(L_n)/s(L_n)] = 1 + 2k(cn/jk)/(n(j+1)) = 1 + 2c/(j(j+1))$. This is perhaps a more interesting measure in this case, since we know by Theorem 3 that for all n, $s(L_n) - OPT(L_n)$ is almost surely less than some fixed constant, meaning that this limiting ratio is also $\lim_{n\to\infty} E[FFD(L_n)/OPT(L_n)]$. (The value is 1.00595... in the case of $6,13$.)

Our results for Fluid FFD with $k \leq 1{,}000$ suggest that the $U\{6;13\}$ distribution might be the worst possible for FFD under both metrics, and this indeed turns out to be the case, as already indicated in the statement of Theorem 5 and the paragraph following in Section 1. The proof is based on a worst-case analysis of Fluid FFD, and relies on the assumption, justified by our computer runs, that $k > 1{,}000$. It involves two separate arguments, both somewhat complicated. The first, covering the waste in bins that start with items larger than $k^{1/2}$, has something of the flavor of the analysis of FFD in [3] for the continuous case, although edge effects that previously could be ignored must now be taken into account. The argument for bins that start with items smaller than $k^{1/2}$ must confront distinctly new issues. Space limitations unfortunately prevent us from saying more here.

We conclude this section by noting that all the arguments presented here about FFD apply equally well to Best Fit Decreasing (BFD), the algorithm that applies Best Fit rather than First Fit to the presorted list. Moreover, for $k \le 1{,}000$, $j \le k/2$, Fluid BFD always has output of the same Type as does Fluid FFD, although on 48 occasions where both had Type 1 outputs, BFD ended up with a slightly smaller constant of proportionality on its $\Theta(n)$ expected waste. The biggest relative difference occurred for the pair 422,847, where the constant of proportionality for BFD was 4.821×10^{-5} while that for FFD was 4.860×10^{-5}, roughly 0.8% worse (admittedly a mostly theoretical distinction).

4. First Fit and Best Fit.

Given the rapidly approaching page limit, we do not have space here to elaborate on what was said in Section 1 about the proofs of the results listed in Theorem 6 for the case of distributions $U\{j;k\}$, $j \in \{k-1, k\}$. (These adapted and improved on known techniques for the continuous case.) We shall instead turn to the case of $j < k-1$, where techniques quite different from those used in the continuous case are needed. Let us concentrate on the proof of the second part of Theorem 8, which addressed the behavior of First Fit. (The Theorem 8 result for Best Fit follows by a slight tightening of the argument.)

Theorem 8b. *If* $k \ge j^2$, $E[FF(L_n) - s(L_n)] = O(1)$.

Proof. For simplicity, let us once again normalize so that the bin size in k and the item sizes range from 1 to j. Consider an infinite Markov process in which the states are the possible sequences of partially filled bins in a packing. (Note that these "states" ignore the position and number of completely full bins in the packing.) Let $n(t)$, $t \ge 1$ be the index of the item that first pushes the total size of items in the tth bin over the threshold $k-j$, and let $x(t)$ be the state in existence at that time. Let $f(x(t))$ be the number of partially full bins in state $x(t)$. We shall prove that $E[f(x(t))] = O(1)$, which implies the theorem. The proof uses a criterion due to Hajek [11] which states that for a stochastic process $x(t)$ if (1) the increment $f(x(t+1)) - f(x(t))$ is bounded independently of $x(t)$ by a random variable whose distribution has an exponentially bounded tail, and (2) there exists $\delta > 0$ such that for all y outside a finite set S the drift $E[f(x(t+1)) - f(x(t)) : x(t) = y]$ is less than $-\delta$, then the stochastic process $f(x)$ is stable in the sense that there exists a $B < \infty$, such that $E[f(x(t))] < B$, for all t. (Note that a random variable X has exponentially bounded tails if there is a $\omega > 0$ such that $P(|X| > a) \le e^{-\omega a}$ for all a.)

It is trivial that $f(x(t+1)) - f(x(t)) \le 1$, so part (1) of Hajek's criteria is satisfied. The idea which leads to a verification of (2) is as follows. Suppose that in state y, there are two or more partially filled bins, other than the last. As each subsequent item arrives (at least until all bins but the $t+1$st are completely filled), there is a probability of at least $1/j$ that some partially full bin becomes completely filled by the packing of the arriving item into that bin. (This holds for both BF and FF.) All we need to do is to take k sufficiently large to ensure that the expected number of partially full bins to become full between $n(t)$ and $n(t+1)$ is greater than $1+\delta$ for some $\delta > 0$. This will give $E[f(x(t+1)) - f(x(t)) : x(t) = y] < -\delta$. A sufficiently large k is $k = j^2$, since this ensures that at least j items will be packed between $n(t)$ and $n(t+1)$, and that there is a positive probability that there will be $j+1$ or more such items (the expected value is roughly $2j$). To obtain Hajek's property (2), all we thus have to do is take B to be the set of states with no more than 2 partially full bins. □

5. Conclusion.

In this paper we have analyzed the average-case behavior of one-dimensional bin packing algorithms under discrete uniform distributions, and found marked differences between this behavior and that under the corresponding continuous uniform distributions. For the most part, we have had to use significantly different proof techniques to handle the discrete case. A natural question is whether we can extend these techniques to bounded domain discrete distributions other than the $U\{j;k\}$.

For many of our techniques, the answer is yes. The results of Courcoubetis and Weber [5,6] apply to arbitrary discrete distributions, although the details of the packing theorems needed to apply them and determine the correct answer will vary (and in some cases the packing problem that needs solution will be NP-complete). One can conclude in general, however, that the *optimal* waste has the same three possible growth rates $\Theta(n)$, $\Theta(n^{1/2})$, and $O(1)$. The only correspondingly general results for the continuous case are those due to Rhee and Talagrand [21], and do not paint such a complete picture. They show only that the expected waste for optimal packings must either be $\Theta(n)$ or $O((n \log n)^{1/2})$, whereas we know that for certain continuous distributions the optimal expected waste can be $O(1)$.

The proof that the "Fluid FFD" procedure can be used to determine the expected waste for FFD applies to arbitrary discrete distributions, at least assuming that the item probabilities are commensurable, thus again yielding the same three possibilities for the growth rate of expected waste. The proof that First and Best Fit give constant expected waste when k is sufficiently large with respect to j can be extended to any distribution in which the item sizes i, $1 \le i \le j$, have different but non-zero probabilities p_i, and $k \ge j/min_i\{p_i\}$.

In our opinion the most interesting open problem in the average-case analysis of bin packing algorithms concerns the situations in which First Fit and Best Fit appear to yield linear expected waste, as indicated in Figure 2. To date we have been unable to prove that the waste must be $\Omega(n)$ in any of those situations (discrete or continuous) where it appears to occur. The simplest situation for which experiments indicate linear waste is the $U\{8;11\}$ distribution. This is perhaps the most likely candidate for a proof, and we are currently pursuing several promising approaches. The process by which the linear waste is created seems clear from experiments: it is all concentrated in bins that receive items of total size 10. These appear to be created at a rate of about $0.126n$, whereas the items of size 1 that are the only candidates for filling the gaps in these bins only arrive at rate $0.125n$. (A similar process appears to go on for all $U\{k-3,k\}$ distributions.) Surprisingly, although First and Best Fit have $\Theta(n^{1/2})$ waste under $U\{k-1,k\}$ and appear to have linear waste under $U\{k-3,k\}$, we conjecture based on experiments that the waste is $O(1)$ under $U\{k-2,k\}$. (The constant bound here, however, appears to be much larger than the bound we observed when $j \le k^{1/2}$, which amounted to the contents of a single bin.)

REFERENCES

1. M. Ajtai, J. Komlós, and G. Tusnády, "On optimal matchings," *Combinatorica* **4** (1983), 259-264.
2. J. L. Bentley, D. S. Johnson, F. 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, IL, 1983, 51-60.
3. J. L. Bentley, D. S. Johnson, F. T. Leighton, C. C. McGeoch, and L. A. McGeoch, "Some unexpected expected behavior results for bin packing," *in* "Proceedings 16th Ann. ACM Symp. on Theory of Computing," Association for Computing Machinery, New York, 1984, 279-288.
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. C. Courcoubetis and R. R. Weber, "Necessary and sufficient conditions for the stability of a bin-packing system," *J. Appl. Prob.* **23** (1986), 989-999.
6. C. Courcoubetis and R. R. Weber, "Stability of flexible manufacturing systems," *Operations Res.*, to appear.
7. J. Csirik, J. B. G. Frenk, A. Frieze, G. Galambos, and A. H. G. Rinnooy Kan, "A probabilistic analysis of the next fit decreasing bin packing heuristic," *Operations Res. Lett.* **5** (1986), 233-236.
8. J. Csirik and D. S. Johnson, "Bounded space on-line bin packing: Best is better than first," *in Proceedings 2nd Ann. ACM-SIAM Symp. on Disc. Algorithms*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1991, 309-319.
9. G. N. Frederickson, "Probabilistic analysis for simple one- and two-dimensional bin packing algorithms," *Inform. Process. Lett.* **11** (1980), 156-161.
10. G. H. Hardy and J. E. Littlewood, "Some problems of Diophantine approximation," *Acta Mathematica* **37** (1914), 155-238.
11. B. Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications," *Adv. Appl. Prob.* **14** (1982), 502-525.
12. N. Karmarkar, "Probabilistic analysis of some bin-packing algorithms," *in* "Proceedings 23rd Ann. Symp. on Foundations of Computer Science," IEEE Computer Society, Los Angeles, Calif., 1982, 107-111.
13. R. M. Karp, M. Luby, and A. Marchetti-Spaccamela, "Probabilistic analysis of multidimensional bin packing problems," *in* "Proceedings 16th Ann. ACM Symp. on Theory of Computing," Association for Computing Machinery, New York, 1984, 289-298.
14. W. Knödel, "A bin packing algorithm with complexity $O(n\log n)$ and performance 1 in the stochastic limit," in "Proc. 10th Symp. on Mathematical Foundations of Computer Science," J. Gruska and M. Chytil (eds.), *Lecture Notes in Computer Science 118*, Springer-Verlag, 1981, 369-378.
15. D. E. Knuth, "Big omicron and big omega and big theta," *SIGACT News* **8:2** (1976), 18-24.
16. 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.
17. 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.
18. P. Ramanan, "Average-case analysis of the smart next fit algorithm," *Inform. Process. Lett.* **31** (1989), 221-225.
19. P. Ramanan and K. Tsuga, "Average-case analysis of the modified harmonic algorithm," *Algorithmica* **4** (1989), 519-534.
20. W. T. Rhee and M. Talagrand, "Optimal bin packing with items of random sizes," *Math. Oper. Res.* **13** (1988), 140-151.
21. W. T. Rhee and M. Talagrand, "Optimal bin packing with items of random sizes II," *SIAM J. Comput.* **18** (1989), 139-151.
22. P. W. Shor, "The average case analysis of some on-line algorithms for bin packing," *Combinatorica* **6** (1986), 179-200.
23. P. W. Shor, "How to pack better than best fit: An improved on-line bin packing algorithm," in preparation.