

BIN PACKING WITH DISCRETE ITEM SIZES, PART I: PERFECT PACKING THEOREMS AND THE AVERAGE CASE BEHAVIOR OF OPTIMAL PACKINGS*

E. G. COFFMAN, JR.[†], C. COURCOUBETIS[‡], M. R. GAREY[§], D. S. JOHNSON[¶],
P. W. SHOR[¶], R. R. WEBER^{||}, AND M. YANNAKAKIS[§]

Abstract. We consider the one-dimensional bin packing problem with unit-capacity bins and item sizes chosen according to the discrete uniform distribution $U\{j, k\}$, $1 < j \leq k$, where each item size in $\{1/k, 2/k, \dots, j/k\}$ has probability $1/j$ of being chosen. Note that for fixed j, k as $m \rightarrow \infty$ the discrete distributions $U\{mj, mk\}$ approach the continuous distribution $U(0, j/k]$, where the item sizes are chosen uniformly from the interval $(0, j/k]$. We show that average-case behavior can differ substantially between the two types of distributions. In particular, for all j, k with $j < k - 1$, there exist on-line algorithms that have constant expected wasted space under $U\{j, k\}$, whereas no on-line algorithm has even $o(n^{1/2})$ expected waste under $U(0, u]$ for any $0 < u \leq 1$. Our $U\{j, k\}$ result is an application of a general theorem of Courcoubetis and Weber [C. Courcoubetis and R.R. Weber, *Probab. Engrg. Inform. Sci.*, 4 (1990), pp. 447–460] that covers all discrete distributions. Under each such distribution, the optimal expected waste for a random list of n items must be either $\Theta(n)$, $\Theta(n^{1/2})$, or $O(1)$, depending on whether certain “perfect” packings exist. The perfect packing theorem needed for the $U\{j, k\}$ distributions is an intriguing result of independent combinatorial interest, and its proof is a cornerstone of the paper.

Key words. bin packing, on-line, average-case analysis, approximation algorithms

AMS subject classifications. 68W25, 68W40

PII. S0895480197325936

1. Introduction. Suppose one is given items of sizes $1, 2, 3, \dots, j$, one of each size, and is asked to pack them into bins of capacity k with as little wasted space as possible, i.e., one is asked to find a least cardinality partition (packing) of the set of items such that the sizes of the items in each block (bin) sum to at most k . For what values of j and k can the set be packed perfectly (i.e., so that the sizes of the items in each block sum to exactly k)? 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 sufficient. Readers might want to try their hand at proving this. Relatively short proofs exist, as illustrated in the next section, but a certain ingenuity is required to find one. The exercise serves as a warm-up for the following more general and more difficult theorem, also proved in the next section, in which there are r copies of each size, for some $r \geq 1$.

THEOREM 1 (perfect packing theorem). *For positive integers k, j , and r , with $k \geq j$, one can perfectly pack a list L consisting of rj items, r each of sizes 1 through j , into bins of size k if and only if the sum of the rj item sizes is a multiple of k .*

*Received by the editors August 11, 1997; accepted for publication (in revised form) October 14, 1999; published electronically May 26, 2000.

<http://www.siam.org/journals/sidma/13-3/32593.html>

[†]New Jersey Institute of Technology, Newark, NJ 07102 (coffman@homer.njit.edu). The research of this author was completed while at Bell Labs.

[‡]Department of Computer Science, Athens University of Economics and Business, Athens, Greece (courcou@csi.forth.gr).

[§]Bell Labs, Murray Hill, NJ 07974 (mihalis@research.bell-labs.com). M.R. Garey is now retired.

[¶]AT&T Labs–Research, Florham Park, NJ 07932 (dsj@research.att.com, shor@research.att.com).

^{||}Statistical Laboratory, University of Cambridge, Cambridge CB2 1SB, UK (R.R.Weber@statslab.cam.ac.uk).

In set-theoretic terms, the question answered by Theorem 1 is an intriguing puzzle in pure combinatorics. But our motivation to work on it came from its relevance to certain fundamental questions about the average-case analysis of algorithms. In particular, consider the standard bin packing problem, in which one is given a list of items $L = (a_1, a_2, \dots, a_n)$, where each a_i has a positive size $s_i \leq 1$, and is asked to find a packing of these items into a minimum number of unit-capacity bins. Clearly, Theorem 1 can be cast in these terms by a simple rescaling.

In most real-world applications of bin packing, as in Theorem 1, the item sizes are drawn from some finite set. However, the usual average-case analysis of bin packing heuristics has assumed that item sizes are chosen according to continuous probability distributions, which by their nature allow an uncountable number of possible item sizes (see [1, 8], for example). The assumption of a continuous distribution has the advantage of sometimes simplifying the analysis, and has been justified on the grounds that continuous distributions should serve as reasonable approximations for discrete ones. But there are reasons to ask whether this is actually true. For example, let $U(0, u]$ denote the continuous uniform distribution over the interval $(0, u]$ and let $U\{j, k\}$ denote the discrete uniform distribution on the set $\{1/k, \dots, j/k\}$. Then the limit of the distributions $U\{mj, mk\}$, as $m \rightarrow \infty$ is $U(0, j/k]$, but in the limit combinatorial questions such as those addressed by Theorem 1 evaporate. This suggests that something important (and interesting) may in fact be lost in the transition from discrete to continuous models. The results in this paper illustrate that, indeed, fundamental aspects of the average-case behavior of classical bin packing algorithms are obscured by the continuous models.

To describe these results, we need the following notation. If A is an algorithm and L is a list of items, then $A(L)$ is the number of bins used when A is applied to L , and $s(L)$ is the sum of the item sizes in L . The waste in the packing of L by A is denoted by $W^A(L) = A(L) - s(L)$. Note that this equals the sum of the gaps in the bins used by A 's packing of L , where the gap in a bin containing items of total size σ is $1 - \sigma$. Let $\text{OPT}(L)$ denote the length of an optimal packing; OPT also denotes a corresponding optimal off-line packing algorithm. In what follows, L_n is a list of n items whose sizes are independent samples from a given distribution. To avoid trivialities, we disallow the item size k , and hence ignore the distribution $U\{k, k\}$. Further, we will ignore the (deterministic) distribution $U\{1, k\}$. Thus, only the distributions $U\{j, k\}$, $1 < j < k$, will be of interest.

This paper is the first in a series of articles covering the results announced in [2, 6], all dealing with classical bin packing under discrete distributions and showing qualitative differences between algorithmic behavior under such discrete distributions and their continuous analogues. The specific problems addressed by the series are summarized in Table 1, and are shown along with the continuous analogues for contrast. The present paper, Part I of the series, proves the results marked by a bullet (\bullet) in the table. In this table $\Omega(f(n))$ has the Knuthian sense of “greater than $cf(n)$ for some $c > 0$ and all sufficiently large n ” and $\underline{\Omega}(f(n))$ has the weaker sense of Hardy and Littlewood’s “not $o(f(n))$,” i.e. “greater than $cf(n)$ for some $c > 0$ and infinitely many n .” We say $g(n)$ is $\Theta(f(n))$ (respectively, $\underline{\Theta}(f(n))$) when $g(n)$ is both $O(f(n))$ and $\Omega(f(n))$ (respectively, $\underline{\Omega}(f(n))$). A subscript of u or k indicates that the hidden multiplicative constants depend on u or k and so we have asymptotics in n that hold for fixed u or k . Where more than one growth rate is possible, we list all that have been observed for distributions of the stated type, using “...” to indicate when other possibilities have not been ruled out. In the case of general discrete distributions it is NP-hard to determine which possibilities apply under first fit decreasing (FFD),

TABLE 1
Results for expected waste.

	$U(0, 1]$	Ref	$U\{j, k\}, j = k - 1, k$	Ref
OPT	$\Theta(n^{1/2})$	[17, 16]	$\Theta(n^{1/2})$	[8]
FFD, BFD	$\Theta(n^{1/2})$	[17, 16]	$\Theta(n^{1/2})$	[8]
FF	$\Theta(n^{2/3})$	[7]	$\Theta(n^{1/2}k^{1/2}), k = O(n^{1/3})$	[7]
			$\Theta(n^{2/3}), k = \Omega(n^{1/3})$	[7]
BF	$\Theta(n^{1/2} \log^{3/4} n)$	[19]	$\Theta(n^{1/2} \log^{3/4} k), k = O(n)$	[5]
			$\Theta(n^{1/2} \log^{3/4} n), k = \Omega(n)$	[5]
Best on-line	$\underline{\Theta}(n^{1/2} \log^{1/2} n)$	[19, 20]	$\underline{\Omega}_k(n^{1/2})^\dagger$	[-]

	$U(0, u], u < 1$	Ref	$U\{j, k\}, 1 \leq j \leq k - 2$	Ref
OPT	$\Theta_u(1)$	[1]	$\Theta_k(1)$	[•]
FFD	$\Theta_u(1), u \leq 1/2$	[1, 13]	$\Theta_k(1), \Theta_k(n^{1/2})^\ddagger, \Theta_k(n)^\S$	[3]
	$\Theta_u(n^{1/3}), u > 1/2$	[1]		
FF	$\Theta_u(n)^*$	[2]	$\Theta_k(1), \Theta_k(n^{1/2})^\ddagger, \Theta_k(n)^*$	[6, 4]
BF	$\Theta_u(n)^*$	[2]	$\Theta_k(1), \Theta_k(n^{1/2})^\ddagger, \Theta_k(n)$	[6, 4, 15]
Best on-line	$\underline{\Omega}_u(n^{1/2})$	[•]	$\Theta_k(1)$	[•]

	General continuous	Ref	General discrete	Ref
OPT	$\Theta_u(1), \Theta_u(n^{1/2}),$ $\Theta_u(n), \dots$	[1, 17, 16]	$O_k(1), \Theta_k(n^{1/2}), \Theta_k(n)$	[9]
FFD	$\Theta_u(1), \Theta_u(n^{1/3}),$ $\Theta_u(n^{1/2}), \Theta_u(n), \dots$	[1, 17, 16]	$O_k(1), \Theta_k(n^{1/2}), \Theta_k(n)^\S$	[3]
Best on-line	$\underline{\Theta}(n^{1/2} \log^{1/2} n),$ $\Theta_u(n), \dots$	[19, 20]	$O_k(1), \Theta_k(n^{1/2}), \Theta_k(n)$	[9]

§ These results also hold for BFD.

† Easily shown to be $\underline{\Omega}(n^{1/2})$; $O(n^{1/2} \log^{3/4} k)$ by the results for Best Fit; conjectured to be $\underline{\Theta}(n^{1/2} \log^{1/2} k), k = O(n)$ and $\underline{\Theta}(n^{1/2} \log^{1/2} n), k = \Omega(n)$.

‡ Not ruled out by theorems, but no occurrences known either. For FFD and $j \leq k - 2$, $\Theta_k(n^{1/2})$ is known not to occur for any $k \leq 10,000$ [3].

* Conjecture supported by simulation studies.

best fit decreasing (BFD), and the best on-line algorithm, although exponential-time decision algorithms exist [9, 4].

Our primary focus in this paper is on the distributions $U\{j, k\}$ with $1 < j \leq k - 2$. The marked results for these distributions are both contained in the following theorem, which is proved in section 3.

THEOREM 2. *For any $j, k, 1 < j \leq k - 2$, there exists an on-line algorithm A with running time bounded by a polynomial in n and k such that if L_n has item sizes generated according to $U\{j, k\}$ then $E[W^A(L_n)] = \Theta(1)$.*

The corresponding result in the continuous model, which is proved in section 5, shows substantially worse behavior.

THEOREM 3. *If L_n has item sizes generated according to $U(0, u]$ for $0 < u < 1$, and A is any on-line algorithm, then there exists a constant $c > 0$ such that $E[W^A(L_n)] > cn^{1/2}$ for infinitely many n .*

We should note that the comparison between discrete and continuous cases is much closer in the case of $U\{k - 1, k\}$, which corresponds to $U(0, 1]$. In fact, for unconstrained (off-line) optimal packings, one obtains the same $\Theta(n^{1/2})$ expected waste results in both cases. The proof of this fact in the discrete case can be assembled from standard techniques in the continuous theory, for example by observing that the algorithm MATCH and its analysis in [8, p. 100] carry over to the discrete case. Note that this bound is independent of k . There does appear to be a dependence on k

in the on-line case, but one can obtain $O_k(n^{1/2})$ behavior in a quite straightforward manner. Simply keep a pool of $\lceil k/2 \rceil$ bin types, one each for each matching pair of item sizes $(i, k - i)$, $1 \leq i \leq k/2$ and one for item size $k/2$ if k is even. When an item of size i arrives, we look to see if there is a partially full bin of its type containing an item of size $k - i$ and if so, add the new item to the bin, filling it. Otherwise a new bin of the given type is started. It is easy to show that this procedure has $O_k(n^{1/2})$ expected waste.

The plan for the remaining papers in this series is as follows. Part II of the series [7] proves the first fit (FF) results under $U\{k - 1, k\}$, and in so doing, the $O(n^{2/3})$ upper bound for $U(0, 1]$. Part III [3] will cover the results on FFD and BFD under discrete distributions (both uniform and general). Parts IV [4] and V [5] will present the results for best fit (BF) under $U\{j, k\}$, $1 < j \leq k - 2$ and $U\{k - 1, k\}$, respectively (see also [15] which generalizes the results of [6, 4] for the case $U\{k - 2, k\}$).

The current paper is organized as follows. The proof of the perfect packing theorem (Theorem 1) appears in section 2. In addition to being of independent combinatorial interest, this result contributes to the proof of Theorem 2, which we present in sections 3 and 4. In section 3, we describe a general result of Courcoubetis and Weber [9] which implies that for any discrete item-size distribution, $E[W^{\text{OPT}}(L_n)]$ must be one of $\Theta(n)$, $\Theta(n^{1/2})$, or $O(1)$, and no matter which case applies, there is an on-line algorithm A whose expected waste obeys the same bounds. Which case applies depends on the existence of certain perfect packings, and here is where Theorem 1 comes in. The algorithms provided by [9] for $U\{j, k\}$, $1 < j \leq k - 2$, are randomized algorithms, and in section 4 we present corresponding deterministic algorithms and give the somewhat more difficult proof that they also have $O(1)$ expected waste. Theorem 3, this paper's contribution to the theory of continuous distributions is proved in section 5.

We conclude in section 6 with a brief discussion of the significance of Theorem 2 and possible extensions to it. As stated, the theorem provides algorithms that are tailored to the particular distribution $U\{j, k\}$ in question. We can, however, provide a single algorithm A with running time bounded by a polynomial in n and k that yields $E[W^A(L_n)] = \Theta(1)$ for any such distribution, even if j and k are not known in advance and have to be discovered on-line. We also discuss the possibility of extending Theorem 2 to more general classes of discrete distributions and mention some more recent developments. In particular, in [10, 11] a simple new deterministic on-line algorithm is introduced that supersedes the algorithms presented in this paper, in that it has $O(1)$ expected waste for *every* discrete distribution that has $E[W^{\text{OPT}}(L_n)] = O(1)$, and has a running time $O(nk)$ for all $U\{j, k\}$ distributions.

2. The perfect packing theorem. We begin our proof of Theorem 1 with three lemmas that list a number of special instances that lead to perfect packing. The first lemma takes care of the special case, $r = 1$.

LEMMA 4. *Suppose m, j , and k are positive integers such that $j \leq k$ and $mk = j(j + 1)/2$. Then the set of j items, one each of sizes $1, \dots, j$, perfectly packs into m bins of size k .*

Proof. The proof is by induction. Pick j and k and assume the theorem is true for all pairs that are smaller in lexicographic order than (k, j) . The theorem is clearly true for $k \leq 2$ or $j \leq 2$, so assume $k, j > 2$.

If $j > k/2$ then we can start by perfectly packing bins with pairs of items $(j - i, k - j + i)$, $0 \leq i < j - k/2$, after which the remaining items are those of sizes $1, \dots, k - j - 1$, plus the item of size $k/2$ if k is even. Since the sum of the sizes of

the items that have been packed at this point is a multiple of k , the sum of the sizes of remaining items is also a multiple of k . If k is odd, the unpacked items are an instance of $(k, k - j - 1)$, with $k - j - 1 < j$ and the induction hypothesis applies. If k is even then $k/2$ divides $j(j + 1)/2$ and all remaining items are no larger than $k/2$. Thus the items $1, \dots, k - j - 1$ form an instance of $(k/2, k - j - 1)$ and by the induction hypothesis can be perfectly packed into bins of size $k/2$. These half-bins and the item of size $k/2$ can then be combined into bins of size k .

Now suppose $j \leq k/2$. If k is even then we have an instance of $(k/2, j)$ and the induction hypothesis applies. If k is odd, first note that $k/2 \geq j$ and $j > 2$ implies $k > j + 1$, which together with $mk = j(j + 1)/2$ implies $j > 2m$. Thus we can construct m pairs of items each of total size $k' = 2j - 2m + 1$ by combining $j - i$ with $k' - j + i$, $0 \leq i \leq m - 1$. If we place one pair in each of our m bins, we now have m bins with gaps of size $k - k' = k - 2j + 2m - 1$ and items of sizes $1, \dots, j - 2m$. Because $mk = j(j + 1)/2$, the sum of these item sizes must be $m(k - k')$, and so an application of the induction hypothesis to the instance $(k - k', j - 2m)$ completes the proof. \square

LEMMA 5. Consider $r > 1$ sets, the i th of which consists of j items of consecutive sizes, $\ell_i + 1, \dots, \ell_i + j$, for some $\ell_i \geq 0$. Suppose either (a) r is even or (b) j is odd. Then these rj items perfectly pack into j bins of size equal to the sum of the average item sizes in the r groups, i.e., $r(j + 1)/2 + \sum_{i=1}^r \ell_i$.

Proof. The lemma will follow if we can show that for $\ell_i = 0$, $1 \leq i \leq r$, and bins of size $r(j + 1)/2$ it is possible to pack perfectly the items into the j bins in such a manner that each bin contains exactly one item from each of the r sets.

If r is even then we simply take two of the sets and pack the i th largest item in one set with the i th smallest item in the other set, i.e., as the pair $(i, j - i + 1)$, $i = 1, \dots, j$. This fills j bins to level $j + 1$. By repeating this $r/2$ times we fill j bins of size $r(j + 1)/2$.

If r and j are both odd then an extra step is required. The idea is first to pack items in triples, one item from each of three sets, such that the sum of each triple is the same. It is easiest to appreciate the construction by considering an example, say $j = 9$. The triples, which each sum to 15, are given in the columns below.

1	2	3	4	5	6	7	8	9
6	7	8	9	1	2	3	4	5
8	6	4	2	9	7	5	3	1

In general, the triples are $(i, i + (j + 1)/2, j + 1 - 2i)$, $i = 1, \dots, (j - 1)/2$, and $(i, i - (j - 1)/2, 2j - 2i + 1)$, $i = (j + 1)/2, \dots, j$. The result of packing these one per bin is to fill all j bins to level $3(j + 1)/2$. The number of remaining sets is even and the remaining spaces in the j bins are equal. Thus, the procedure for case (a) can be applied to complete the packing in each bin. \square

The following lemma provides part of the induction step used in the proof of Theorem 1.

LEMMA 6. Consider a quadruple (k, j, r, m) of positive integers such that $k \geq j$ and $mk = rj(j + 1)/2$. Then there exists a perfect packing of r copies of $1, \dots, j$ into m bins of size k if there exists a perfect packing for each lexicographically smaller quadruple of this form, and if any one of the following holds:

- (a) $j \geq k/2$.
- (b) r does not divide k .
- (c) k or r is even.

(d) $j \leq (r - 1)k/2r$.

Proof. First, using the arguments of Lemma 4, we demonstrate how to reduce the problem to a smaller instance if (a) holds. If $j \geq k/2$ and k is odd, then we can pack bins with pairs $(j - i, k - j + i)$, $i = 0, \dots, j - (k + 1)/2$. The remaining items, which are of sizes $1, \dots, k - j - 1$, define the smaller instance $(k, k - j - 1, r, m')$, where $m' = m - r(2j + 1 - k)/2$. If $j > k/2$ and k is even, then we can pack bins in the same way, $i = 0, \dots, j - 1 - k/2$. The remaining items, which are of sizes $1, \dots, k - j - 1$ and $k/2$, can be packed into bins of size $k/2$, by the induction hypothesis that there exists a perfect packing for $(k/2, k - j - 1, r, m')$, where $m' = 2m - r(2j - k)$.

If (b) holds then r and m must have a common factor $p > 1$ and the problem reduces to the instance $(k, j, r/p, m/p)$.

Now suppose neither (a) nor (b) holds but (c) does. If k is even then $k/2$ divides $rj(j + 1)/2$. Since (a) doesn't hold, $j < k/2$. Thus the problem reduces to a smaller instance in which the bin size is $k/2$. If r is even, then since (b) does not hold, k is divisible by r and so must be even too. Thus the same argument applies.

Finally, for case (d), assume that (a), (b), and (c) do not hold, i.e., $j < k/2$, r divides k and k and r are both odd. Let $r_1 = (r + 1)/2$, $k_1 = kr_1/r$ and $r_2 = (r - 1)/2$, $k_2 = kr_2/r$. Note that $r_1 + r_2 = r$ and $k_1 + k_2 = k$. The fact that r is odd implies that r_1 and r_2 are integers. The fact that r divides k implies that k_1 and k_2 are integers, with $mk_1 = r_1j(j + 1)/2$ and $mk_2 = r_2j(j + 1)/2$. Since by assumption $j < k/2$, we have $k_1 \geq j$, and hence by hypothesis for the instance (k_1, j, r_1, m) , we can pack r_1 copies of $1, \dots, j$ into m bins of size k_1 . Similarly, if also $j \leq k_2$ then we can pack r_2 copies of $1, \dots, j$ into m bins of size k_2 . Since $k_1 + k_2 = k$ we can combine pairs of bins of sizes k_1 and k_2 into bins of size k . Thus there is a reduction to smaller instances if $j \leq k_2 = (r - 1)k/2r$, i.e., if (d) holds. \square

Proof of the perfect packing theorem. Instances for which the theorem is to be proved are described by the quadruples of Lemma 6. Notice that it would be enough to specify the triple (k, j, r) ; however, it is helpful to mention m explicitly. The proof of the theorem is by induction on (k, j, r) under lexicographical ordering. By Lemma 4 it is true for $r = 1$. Assume all quadruples that are smaller than (k, j, r, m) can be perfectly packed and $r > 1$. We show there exists a perfect packing of r copies of $1, \dots, j$ into m bins of size k . By Lemma 6, we need only consider the case when k and r are odd, r divides k and $(r - 1)k/2r < j < k/2$. Note that in this case $(r - 1)k/2r$ is an integer and $k/2r$ is 0.5 more than an integer. We show below that we can perfectly pack all the items of sizes from $j + 1 - (r - 1)k/2r$ through j into bins of size k . (Note that the lower bound on this range is greater than 1 because of the above lower bound on j .) The theorem then follows because the remaining items form a smaller quadruple, so by the induction hypothesis they can be perfectly packed into bins of size k .

To follow the construction below, the reader may find it helpful to consider a specific example. Consider the quadruple $(k, j, r, m) = (165, 77, 5, 91)$. Note that k and r are odd, r divides k , and j lies between $(r - 1)k/2r = 66$ and $k/2 = 82.5$. We show below how to perfectly pack all items of sizes $12, \dots, 77$. The remaining items form the smaller quadruple $(165, 11, 5, 2)$.

To pack all items of sizes from $j + 1 - (r - 1)k/2r$ through j , we divide the range of item sizes into intervals, i.e., sets of consecutive integers. Each interval is symmetric about a multiple of $k/2r$ and has one of two lengths depending on whether the interval is symmetric about an odd or even multiple of $k/2r$. To form the intervals, we first take the largest interval that is symmetric about $(r - 1)k/2r$; this is the interval

$[(r-1)k/r - j, j]$. Note that this interval does not include $(r-2)k/2r$ since $j < k/2 = rk/2r$. Next we take the largest interval that can be formed from the remaining items that is symmetric about $(r-2)k/2r$, obtaining the interval $[j - k/r + 1, (r-1)k/r - j - 1]$. Continuing in this fashion and taking intervals symmetric about further multiples of $k/2r$, we end up with intervals of two kinds. First, there are $(r-1)/2$ intervals centered on even multiples of $k/2r$, with the interval centered on $(r-1-2i)k/2r$ being $[(r-1-i)k/r - j, j - ik/r]$, where i ranges from 0 to $(r-3)/2$. Second, there are an equal number of interval centered on odd multiples of $k/2r$, with the interval centered on $(r-2i)k/2r$ being $[j - ik/r + 1, (r-i)k/r - j - 1]$, where i ranges from 1 to $(r-1)/2$. Note that the smallest endpoint is $j - ik/r + 1$ for $i = (r-1)/2$, which equals $j + 1 - (r-1)k/2r$ as claimed above.

For the numerical example above, there are two intervals of each type. Intervals of the first type are $[22, 44]$ and $[55, 77]$; they are of length 23 and symmetric about 33 and 66. Intervals of the second type are $[12, 21]$ and $[45, 54]$; they are of length 10 and symmetric about 16.5 and 49.5. In general, intervals of the first type have odd length $2j - (r-1)k/r + 1$ and are symmetric about an even multiple of $k/2r$. Intervals of the second type have even length $k - 2j - 1$ and are symmetric about an odd multiple of $k/2r$. Our plan is to use Lemma 5 to perfectly pack into bins of size k those items whose sizes lie in intervals of the same type.

We begin by considering all those intervals of the first type. These have odd lengths and they are symmetric about points ik/r , $i = 1, \dots, (r-1)/2$. There are r items of each size in each of these intervals. Our strategy is to partition these intervals into groups that satisfy the hypotheses of Lemma 5 (b). That is, we arrange for the midpoints of the intervals within each group to sum to k . Since the midpoints correspond to the average item sizes for the corresponding intervals, and the number of items in the intervals is odd, Lemma 5 (b) implies that we can perfectly pack the items in the intervals of each group. Constructing these groups is a bin packing problem in which the midpoints of the intervals take on the role of item sizes. In what follows we write ‘items’ in single quotes when speaking of the midpoints of intervals, possibly normalized, and viewing them as items to be perfectly packed in bins of some required size. In considering intervals of the first type, it is as though we had r ‘items’ of each of the sizes ik/r , $i = 1, \dots, (r-1)/2$, and wished to pack them into bins of size k . After a normalization that multiplies each item size by r/k , this is equivalent to the problem of packing r ‘items’ of each of the sizes $1, \dots, (r-1)/2$ into bins of size r . That is, we have a smaller version (k', j', r', m') of our packing problem, with $j' = (r-1)/2$, $k' = r' = r$ and $m' = r'j'(j'+1)/2k'$. But by the induction hypothesis this means that the desired packing can be achieved. In the example, it is as though we had 5 ‘items’ of sizes 33 and 66 that are to be packed in bins of size 165. Normalizing by a factor of $1/33$, this is equivalent to the problem instance $(5, 2, 5, 3)$.

We must now pack items whose sizes lie in intervals of the second type. These intervals are of even length, symmetric about the points $ik/2r$, for i odd and $i = 1, \dots, r-2$. Again, there are r items of each size in these intervals. As above, we exhibit a reduction to a smaller perfect packing problem. After multiplying item sizes by $2r/k$ the problem is equivalent to perfectly packing r copies of ‘items’ of sizes $1, 3, 5, \dots, r-2$ into bins of size $2r$. For the example, this is 5 copies of ‘items’ of sizes 1 and 3, to be perfectly packed into 2 bins of size 10. Unfortunately, if the sum of the ‘item’ sizes is an odd multiple of r the ‘items’ cannot be perfectly packed into bins of size $2r$. For this reason, and also because it is convenient to do so even when the sum of the ‘item’ sizes is a multiple of $2r$, we consider perfect packings into bins of sizes r and $2r$. Assume for the moment that r copies of ‘items’ of sizes $1, 3, 5, \dots, r-2$

can be perfectly packed into bins of sizes r and $2r$. If they are packed entirely into bins of size $2r$, then the number of ‘items’ in each bin must be even (as all ‘item’ sizes are odd), and so Lemma 5 applies and implies that the original items can be perfectly packed into bins of size k . On the other hand, suppose a bin of size r is required. The set of ‘items’ that are packed into a bin of size r corresponds to a set of intervals whose midpoints sum to $k/2$. Recall that the intervals are of even length. We divide each such interval into its first half and its second half, obtaining twice as many intervals, whose midpoints now sum to k . Now we can again use Lemma 5 to construct the perfect packing.

The final step in the proof is to show that we can indeed perfectly pack r copies of each of the item sizes $1, 3, 5, \dots, r - 2$ into bins of sizes r and $2r$. We shall use a different packing depending upon whether $r = 4\alpha + 1$ or $r = 4\alpha + 3$.

For the case $r = 4\alpha + 1$, the ‘items’ are perfectly packed by the following simple procedure. We begin by packing one bin with $(1, 1, r - 2)$, one bin with $(2i + 1, 2i + 1, r - 2i - 2, r - 2i)$, for each $i = 1, \dots, \alpha - 1$, and one bin with $(2i - 1, 2i + 1, r - 2i, r - 2i)$, for each $i = 1, \dots, \alpha$, (noting that in the final case, when $i = \alpha$, we get three ‘items’ of size $2i + 1 = r - 2i$). This packs four ‘items’ of each size larger than 1, and three ‘items’ of size 1. We can apply this packing α times, leaving us with one ‘item’ of each size larger than 1 and $\alpha + 1$ ‘items’ of size 1. Then we pack one bin with $(1, 2i - 1, r - 2i)$ for each $i = 1, \dots, \alpha$. This uses up all the remaining ‘items’ (where $\alpha + 1$ items of size 1 are used because there are two ‘items’ of size 1 when $i = 1$). For the numerical example, in which $\alpha = 1$, this construction says that we should pack 5 copies of 1 and 3 into bins of size 5 and 10 by first packing one bin with $(1, 1, 3)$ and then one bin with $(1, 3, 3, 3)$. This leaves one ‘item’ of size 3 and two of size 1. These perfectly pack into a bin of size 5.

When $r = 4\alpha + 3$ the procedure is very similar to that above. We begin by packing one bin with $(1, 1, r - 2)$, one bin with $(2i + 1, 2i + 1, r - 2i - 2, r - 2i)$, for each $i = 1, \dots, \alpha$, and one bin with $(2i - 1, 2i + 1, r - 2i, r - 2i)$, for each $i = 1, \dots, \alpha$. As before, this packs four ‘items’ of each size larger than 1, and three ‘items’ of size 1. We apply this packing α times, leaving us with three ‘items’ of each size larger than 1, and $\alpha + 3$ ‘items’ of size 1. Then we pack one bin with $(2i - 1, 2i + 1, r - 2i, r - 2i)$, for each $i = 1, \dots, \alpha$. This leaves us with $\alpha + 2$ ‘items’ of size 1, two ‘items’ of size $2\alpha + 1$, and one ‘item’ of each other size. Finally, as before, we pack one bin with $(1, 2i - 1, r - 2i)$, for each $i = 1, \dots, \alpha + 1$, which uses up all remaining items. \square

3. Proof of Theorem 2 using randomized algorithms. Recall the theorem statement: For any distribution $U\{j, k\}$, with $1 < j \leq k - 2$, there is an on-line algorithm A with running time bounded by a polynomial in n and k , and with $E[W^A(L_n)] = \Theta(1)$. In this section we show that this is true if we are willing to consider randomized algorithms.

We rely on the following general result of Courcoubetis and Weber [9]. Let $\vec{p} = (p_1, \dots, p_d)$ be a discrete distribution on the set $S = \{s_1, \dots, s_d\}$; p_i is the probability of item size s_i . Any packing of items with sizes from S into a bin of size k can be viewed as a nonnegative integer vector $\vec{c} = (c_1, \dots, c_d)$, where $\sum_{i=1}^d c_i s_i \leq k$. Of particular interest are those vectors that give rise to a sum of exactly k , which we shall call *perfect packing configurations*. For instance, if $S = \{1, 2, 3\}$ and $k = 7$, one such configuration would be $(1, 0, 2)$. Let $\mathbb{P}_{S,k}$ denote the set of all perfect packing configurations for a given S and k . Let $\Lambda_{S,k}$ be the convex cone in \mathbb{R}^d spanned by all nonnegative linear combinations of configurations in $\mathbb{P}_{S,k}$.

THEOREM (Courcoubetis and Weber [9]).

(a) If \vec{p} lies in the interior of $\Lambda_{S,k}$, then there is a polynomial-time randomized on-line bin packing algorithm A with $E[W^A(L_n)] = O(1)$.

(b) If \vec{p} lies on the boundary of $\Lambda_{S,k}$, then $E[W^{OPT}(L_n)] = \Theta(n^{1/2})$ and there is a polynomial-time randomized on-line bin packing algorithm A with $E[W^A(L_n)] = \Theta(n^{1/2})$.

(c) If \vec{p} lies outside of $\Lambda_{S,k}$, then $E[W^{OPT}(L_n)] = \Theta(n)$.

Case (a) is the relevant one here. Case (b) holds for the distributions $U\{k-1, k\}$, but we already have shown in the introduction how the conclusion of (b) can be proved by more direct means for those distributions. We explain below how the algorithm implicit in case (a) works, but first we use that case to prove the version of Theorem 2 involving randomized algorithms. In general it is NP-hard to determine which of the three cases applies to a given distribution (as can be proved by a straightforward transformation from the PARTITION problem [14]). However, for the distributions $U\{j, k\}$ we can use the following lemma, which we shall prove using the perfect packing theorem.

LEMMA 7. For each i, j, k with $i \leq j < k - 1$, there exist positive integers $r_i, s_i, m_i < 2k^2$ such that the set of $r_i j + s_i$ items consisting of $r_i + s_i$ items of size i together with r_i items of each of the other $j - 1$ sizes can be packed perfectly into m_i bins of size k .

Note that this lemma implies that the j -dimensional vector $\vec{e} = (1, 1, \dots, 1)$ is strictly inside the appropriate cone when $S = \{1, 2, \dots, j\}$, $j < k - 1$. This is because \vec{e} is in the interior of the cone spanned by vectors of the form $(r_i, \dots, r_i, r_i + s_i, r_i, \dots, r_i)$, $i = 1, \dots, j$, and those vectors are sums of perfect packing configurations by Lemma 7. The proof of Theorem 2 thus follows from case (a) of the Courcoubetis and Weber theorem.

Proof of Lemma 7. We make use of the perfect packing theorem. There are two cases. If $k \geq i + j$, we simply set $r_i = k - i$ and $s_i = m_i = j(j + 1)/2$. Note that the total size of r_i items each of the sizes $1, \dots, j$ equals $r_i j(j + 1)/2$, so by the perfect packing theorem, we can perfectly pack them into $j(j + 1)/2 = m_i$ bins of size $r_i = k - i$. The remaining $s_i = m_i$ items of size i can then go one per bin to fill these bins up to size precisely k .

On the other hand, suppose $k < i + j$. Now things are a bit more complicated. We have $r_i = 2(k - i)$, $s_i = (k - j)(k - j - 1)$, and $m_i = s_i + r_i(2j - k + 1)/2$. By the perfect packing theorem r_i items each of the sizes $1, \dots, k - j - 1$ perfectly pack in s_i bins of size $k - i$. (Such items exist because by assumption $j < k - 1$.) We then add the additional s_i items of size i to these bins, one per bin, to bring each bin up to size k . There remain r_i items each of sizes $k - j$ through j , for a total of $r_i(j - (k - j - 1)) = 2(m_i - s_i)$ items. These can be used to completely fill the remaining $m_i - s_i$ bins with pairs of items of sizes $(j, k - j), (j - 1, k - j + 1), \dots, (\lceil k/2 \rceil, \lfloor k/2 \rfloor)$. Note that if k is even, the last bin type contains two items of size $k/2$, but we have an even number of such items by our choice of $r_i = 2(k - i)$, so this presents no difficulty.

It is easy to verify that in both cases the values of r_i, s_i, m_i are all less than $2k^2$. \square

We now describe how the algorithms implicit in the Courcoubetis and Weber theorem [9] work for our distributions. Each bin can be labeled with a “type” that specifies the configuration of items it contains. Given a distribution $U\{j, k\}$, we choose a set of perfect packing configurations $\{\vec{c}_q : 1 \leq q \leq Q\}$ that spans a cone having $(1, 1, \dots, 1)$ in its interior, say those configurations generated in the proof of Lemma 7.

Let $c_{q,i}$ denote the number of items of size i used in configuration \vec{c}_q . (Note that by the lemma, we may assume that $Q \leq \sum_{i=1}^j m_i < 2jk^2 = O(k^3)$.) We imagine that there is a separate packing facility for each configuration \vec{c}_q , and that each arriving item is routed to one such packing facility, where it finds space in a partially full bin of that type or starts a new bin of that type. Let $e_{q,i}$ denote the number of different item sizes amongst those used in \vec{c}_q for which there are presently fewer spaces at facility q than there are for item size i . Note that $e_{q,i} = 0$ for i such that $c_{q,i} = 0$. The routing probabilities are determined by finding an $\epsilon > 0$ such that the following linear program has a feasible solution:

$$\sum_{q=1}^Q \alpha_q (c_{q,i} + \epsilon e_{q,i}) = 1/j, \quad i = 1, \dots, j.$$

$$\alpha_q > 0, \quad 1 \leq q \leq Q.$$

Such an ϵ exists because $(1, 1, \dots, 1)$ is in the interior of the cone spanned by the c_q 's; a precise value for ϵ can be computed directly from the distance of $(1, 1, \dots, 1)$ from the boundary of the cone, which itself can be derived from the proof of Lemma 7. Moreover, it is not difficult to see that this distance is such that the number of bits of precision needed to describe ϵ is $O(\log k)$. Thus the linear program can be constructed and solved in time polynomial in k , or in constant time for fixed k .

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

$$\alpha_q (c_{q,i} + \epsilon e_{q,i}) / (1/j).$$

The impact of this construction is to ensure that facility q 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 q . For the details of why this suffices, see [9].

4. A deterministic algorithm. The algorithm described at the end of the previous section is randomized but relatively straightforward. Our deterministic algorithm is a bit more complicated and works on slightly different principles. (Designing a new algorithm from scratch turns out to be easier than attempting to derandomize the above algorithm directly.) We shall use the following consequence of Lemma 7.

LEMMA 8. *For each i, j, k with $i \leq j < k - 1$, there exist positive integers $r'_i, s'_i, m'_i < 4k^5$, with $s'_i \leq r'_i$, such that the set of $r'_i j - s'_i$ items consisting of $r'_i - s'_i$ items of size i together with r'_i items of each of the $j - 1$ other sizes can be packed perfectly into m'_i bins of size k .*

Proof. Let $r_i, s_i, m_i > 0$ be as given in the conclusion of Lemma 7. This means that for $1 \leq i \leq j$, the set R_i consisting of r_i items each of sizes $1, 2, \dots, j$ together with an additional s_i items of size i can be packed perfectly into m_i bins. Note that by the proof of Lemma 7, we may assume that $s_i = (k - j)(k - j - 1)$ if $i > k - j$ and otherwise $s_i = j(j + 1)/2$. Denote the first quantity by \bar{s} and the second by \underline{s} .

First let us suppose $j \leq k/2$, in which case there is no i for which $i > k - j$. We can thus construct the set desired for i, j, k by merging together one copy each of the sets R_h for all $h, 1 \leq h \leq j$ except i . This set has the parameters $r'_i = \sum_{q \neq i} r_h + \underline{s}$ and $s'_i = \underline{s}$, and packs perfectly into $m'_i = \sum_{q \neq i} m_h$ bins by Lemma 7.

The situation is a bit more complicated if $j > k/2$ and s_i can take on both values \underline{s} and \bar{s} . Now we construct our desired set by merging together \bar{s} copies of each R_h ,

$1 \leq h \leq k - j$, with \underline{s} copies of each R_h , $k - j < h \leq j$, and then deleting one copy of R_i . This set has the parameters $r'_i = \sum_{h=1}^{k-j} r_h \bar{s} + \sum_{h=k-j+1}^j r_h \underline{s} + \underline{s} \bar{s} - r_i$ and $s'_i = \underline{s}$ or \bar{s} depending on whether or not $i \leq k - j$. It packs perfectly into $m'_i = \sum_{h=1}^{k-j} m_h \bar{s} + \sum_{h=k-j+1}^j m_h \underline{s} - m_i$ bins by Lemma 7.

The fact that in both cases $r'_i, s'_i, m'_i < 4k^5$ can be derived from the above arguments and the proof of Lemma 7. \square

We now begin the description of our deterministic algorithm for $U\{j, k\}$. As in the randomized algorithm of the previous section, we shall build our packing based on the fixed set of $Q < 2k^3$ perfect packing configurations $\{\bar{c}_q : q = 1, \dots, Q\}$ derived from Lemma 7 for j and k . (Note that these are the only bin configurations needed in the proof of Lemma 8, since that proof constructs its packings based on the packings of Lemma 7.) Every time we start a new bin in our packing, we permanently assign one of these configurations to it, as in the previous algorithm. The bin remains open until it is completely packed in the manner required by its assigned configuration. During certain phases of the algorithm we may also open some empty bins and assign them configurations, with the bins losing those assignments (and open status) if they remain empty at the end of the phase.

If $\bar{c} = (c_1, c_2, \dots, c_j)$ is one of our perfect packing configurations, let $(c_1 : a_1, c_2 : a_2, \dots, c_j : a_j)$ be a labeled configuration denoting a bin with configuration \bar{c} that contains $a_i \leq c_i$ items of size i , $1 \leq i \leq j$. As far as the packing process is concerned, the "state" of the system can at any time be described by giving the current number of open bins of each possible labeled configuration.

Let I_n be the state after the n th item has been packed; I_0 denotes the empty state in which there are no partially packed bins. We define two functions of I_n that are of use in the description and subsequent analysis of our algorithm. Suppose there are x_i as-yet-unfilled spaces for items of size i in the open bins. Let $X(I_n) = \sum_{i=1}^j x_i$ denote the total number of items that are required to complete the packing of the open bins of state I_n . Let $T(I_n) = j \sum_{i=1}^j r'_i \lfloor x_i / s'_i \rfloor$, where r'_i and s'_i are the numbers given in Lemma 8. Note that by Lemma 8, we can assign configurations to additional (empty) bins in such a manner that if there were exactly $T(I_n)/j$ items of each of the sizes $1, \dots, j$ to be packed then these could be used to perfectly pack these newly configured empty bins, with $s'_i \lfloor x_i / s'_i \rfloor$ items of size i left over, $1 \leq i \leq j$: For each i we would use $(r'_i - s'_i) \lfloor x_i / s'_i \rfloor$ items of size i and $r'_i \lfloor x_i / s'_i \rfloor$ items of each of the other sizes to perfectly pack $m'_i \lfloor x_i / s'_i \rfloor$ empty bins. The $s'_i \lfloor x_i / s'_i \rfloor$ surplus items of size i could be used to fill all but $u_i = x_i - s'_i \lfloor x_i / s'_i \rfloor$ of the spaces for items of size i that already existed in our packing of the first n items. Note that $0 \leq u_i < s'_i$. Let $R_{min} = \min\{r'_i / s'_i : 1 \leq i \leq j\}$, and note that $R_{min} \geq 1$ since by Lemma 8, $s'_i \leq r'_i$ for all i . It is easy to see that by definition of $T(I_n)$,

$$(4.1) \quad jR_{min}X(I_n) - j \sum_{i=1}^j r'_i \leq T(I_n) < 4k^6 X(I_n).$$

The algorithm proceeds in phases, depending on the value of $X(I_n)$ at the end of the previous phase. The final state of one phase will also be viewed as the initial state of the following phase. We denote the indices of the final states of the phases by n_h , $h \geq 0$, with $n_0 = 0$. The sequence of n_h 's is defined inductively as follows. Suppose there are more than n_h items so the current phase is not the last.

If $X(I_{n_h}) < L$, where L is specified below, we take $n_{h+1} = n_h + 1$, that is, the next phase involves only the packing of the next item. This item is placed in the

first bin in the list having a space for it or, if necessary, a new bin. For the purpose of defining “first,” we assume the open bins are ordered lexicographically according to their labeled configuration, with ties broken according to the order in which the bins were opened. (Any definition of “lexicographic” will do, although presumably we would want all the empty labeled configurations to come after all the partially full ones.) When no open bin has space for the item, the new bin where it is placed is assigned the lexicographically first configuration that includes an item of the given size.

If, on the other hand, $X(I_{n_h}) \geq L$ then $n_{h+1} = n_h + T(I_{n_h})$. The phase proceeds as follows: Initially, a set of $m \equiv \sum_i m'_i \lfloor x_i/s'_i \rfloor < 2k^5 X(I_{n_h})$ empty bins are opened with configurations assigned in the manner described above. Thereafter, as each item arrives during the phase, it is packed in the first bin that has a space for an item of that size, with “first” defined as above. If during the phase we encounter more than $u_i + T(I_{n_h})/j$ items of some size i then it is necessary to open additional bins for items of size i . Each time such an item arrives, the new bin is assigned the lexicographically first configuration that includes items of size i . The phase terminates after $T(I_{n_h})$ items have been packed or after the last item has been packed, whichever comes first. At the end of the phase, any still-empty bin that has a configuration assigned to it is stripped of that assignment and returned to the pool of unopened empty bins.

This completes our description of the algorithm (except for specifying the constant L , which we shall do below). We now begin our analysis. Our first goal is to show that X and X^2 are well behaved. Lemma 9 bounds the expected change in these functions during a phase.

LEMMA 9. For all $h \geq 0$,

$$(4.2) \quad X(I_{n_{h+1}}) \leq X(I_{n_h}) + k - 1 \quad \text{if } X(I_{n_h}) < L$$

$$(4.3) \quad E[X(I_{n_{h+1}})^2 | X(I_{n_h})] \leq 100k^{18} X(I_{n_h}) \quad \text{if } X(I_{n_h}) \geq L.$$

Proof. When $X(I_{n_h}) < L$ the phase lasts for the packing of a single item. In the worst case, this item starts a new bin. Since no configuration can contain more than k items, this means that $X(I_{n_{h+1}})$ has at most $k - 1$ more items, so (4.2) holds.

Now assume $X(I_{n_h}) \geq L$, in which case the phase lasts for $T \equiv T(I_{n_h}) \leq 4k^6 X(I_{n_h})$ items by (4.1). The number N_i of items of size i amongst the T items that are packed during this interval has a binomial distribution with mean T/j and variance $(T/j)(1 - 1/j)$. If these T items contained exactly T/j items of each size then, using the scheme described above, we would have $X(I_{n_{h+1}}) = \sum_i u_i$. Any shortfall or surplus of N_i over T/j can produce at most $|N_i - T/j|$ extra partially full bins. Since by Lemma 8, $u_i < s'_i < 4k^5$, $1 \leq i \leq j$,

$$X(I_{n_{h+1}}) \leq \sum_{i=1}^j (u_i + (k - 1)|N_i - T/j|) < 4jk^5 + k \sum_i |N_i - T/j|.$$

Squaring the above, taking expected values, and using (4.1), knowledge of the mean

and variance of N_j , and Cauchy–Schwarz inequalities, we derive (4.3) from

$$\begin{aligned}
 E[X(I_{n_{h+1}})^2] &< 16k^{12} + 8k^7 E \left[\sum_i |N_i - T/j| \right] + k^2 E \left[\left(\sum_i |N_i - T/j| \right)^2 \right] \\
 &\leq 16k^{12} + 8k^7 \sum_i \sqrt{E(N_i - T/j)^2} + k^2 E \left[j \sum_i |N_i - T/j|^2 \right] \\
 &\leq 16k^{12} + 8k^7 j \sqrt{(T/j)(1 - 1/j)} + k^2 j^2 (T/j)(1 - 1/j) \\
 &\leq 16k^{12} + 8k^{7.5} \sqrt{T} + k^3 T \\
 &< 25k^{12} T \\
 &\leq 100k^{18} X(I_{n_h}). \quad \square
 \end{aligned}$$

The analysis which follows begins by considering the Markov chain $\{I_{n_h}\}$, restricted to the set of states Z reachable from the empty state $z = \phi$ by a sequence of phases of the algorithm — note that in the chain $\{I_{n_h}\}$ a *step* corresponds to an entire phase of the algorithm. It is easily verified that $\{I_{n_h}\}$ restricted to Z is irreducible, since any state can eventually be converted to ϕ by an arrival sequence having positive probability. We may also assume $\{I_{n_h}\}$ is aperiodic. To ensure this, we need only to augment the set $\{\bar{c}_q : q = 1, \dots, Q\}$ of perfect packing configurations used by our algorithm to include the following two configurations, assuming they are not already present: (a) the configuration consisting of k items of size 1 and (b) the configuration consisting of one size-2 item and $k - 2$ items of size 1. (The latter configuration is possible since we are already assuming that $j > 1$.) Now whenever we are in empty state ϕ there will be ways of returning to it in either k or $k - 1$ steps, depending on the next items to arrive. This will prevent ϕ and every other state from being periodic.

Let $P_y^{(t)}(z)$ be the t -step transition probability from state y to state z . We have the following convergence result.

LEMMA 10. *For any j, k , and $L \geq 200k^{18}$, $P_\phi(z)$ converges to a stationary distribution $\{\pi(z)\}$ geometrically quickly in the X^2 norm, i.e., there exists constants $K > 0$ and $\rho > 1$ such that for all $x \in Z$,*

$$(4.4) \quad \left| \sum_{z \in Z} X(z)^2 P_x^{(h)}(z) - \sum_{z \in Z} X(z)^2 \pi(z) \right| \leq K \rho^{-h}, \quad h \geq 1.$$

Proof. We use the following result on geometric ergodicity, a specialization of Theorem 15.0.1 of [18] to countable chains, which uses the function $\mathbf{1}_C(x) = 1$ if $x \in C$ and 0 otherwise.

THEOREM (Meyn and Tweedie [18]). *Suppose $\{\Phi_k\}$ is an irreducible, aperiodic Markov chain on a countable state space Z , and let $P_y^{(t)}(z)$ be the t -step transition probability from state y to state z for $\{\Phi_k\}$. If there is a function $V : Z \rightarrow [1, \infty)$, constants $b < \infty$, $\beta > 0$, and a finite subset $C \subset Z$ such that for all $x \in Z$,*

$$\sum_{z \in Z} P_x(z) V(z) \leq (1 - \beta) V(x) + b \mathbf{1}_C(x),$$

then $P_\phi^{(h)}(z)$ converges to a stationary distribution $\pi(z)$ geometrically quickly in the V norm in the sense that there exist constants $K > 0$, $\rho > 1$ such that for all $x \in Z$,

$$\sup_{|f| \leq V} \left| \sum_{z \in Z} f(z) P_x^{(h)}(z) - \sum_{z \in Z} f(z) \pi(z) \right| \leq K \rho^{-h}.$$

To apply the theorem, suppose $L \geq 200k^{18}$. Then by (4.2) and (4.3)

(4.5)

$$X(I_{n_{h+1}})^2 \leq (1/2)X(I_{n_h})^2 + (1/2)L^2 + 2(k-1)L + (k-1)^2 \quad \text{if } X(I_{n_h}) < L,$$

(4.6)

$$E[X(I_{n_{h+1}})^2 | X(I_{n_h})] \leq 100k^{18}(X(I_{n_h})/L)X(I_{n_h}) \leq (1/2)X(I_{n_h})^2 \quad \text{if } X(I_{n_h}) \geq L.$$

Let $V(z) = X(z)^2$ for all $z \in Z$ except ϕ , for which we take $V(\phi) = 1$, and let $C = \{z : X(z) < L\}$. By (4.5)–(4.6) we can then take $\beta = 1/2$ and $b = L^2/2 + 2(k-1)L + (k-1)^2$. \square

Now note that if we needed only to prove that the expected waste of our algorithm was bounded at the times n_h at which phases end, we would be done. This is because the waste in any state z is bounded by $X(z)$ (and so by $X(z)^2$) and Lemma 10 tells us that the Markov chain that considers only the states at the ends of phases has an equilibrium distribution and that the expected value of $X(z)^2$ under that equilibrium distribution (i.e., $\sum_{z \in Z} X(z)^2 \pi(z)$) is a finite constant; moreover, $E[X(I_{n_h})^2]$ converges geometrically to this constant. Unfortunately, we need to prove that $E[X(I_n)]$ is bounded by a fixed constant for *all* sufficiently large n , not just those n at which phases end, and this latter result is not implied by Lemma 10. So further argument is needed.

Let us now construct a Markov chain that changes state after each item is packed, not after each phase. Note that $\{I_n\}$ is not Markov since we need to retain more information in a state than simply the current packing state, given that the number of steps to the end of a phase depends not only on the current packing state but also on the packing state that existed at the beginning of the phase and the number of steps taken since then. For a given beginning-of-phase state $z \in Z$, let $t(z)$ denote the length of the phase beginning at z . That is, $t(z) = 1$ if $X(z) < L$ and $t(z) = T(z)$ otherwise. The set of states of our Markov chain, which we denote by $\{I'_n\}$, is $Z \cup W$, where

$$W = \left\{ w(z, \xi, i) : z \in Z, \xi \in \{1, 2, \dots, j\}^{t(z)}, 1 \leq i < t(z) \right\}.$$

The state $w(z, \xi, i)$ corresponds to the packing state (the ordered set of partially empty and empty bins to which configurations have been assigned) that would be obtained from z after the beginning-of-phase assignment of configurations to empty bins has been made and after items of sizes $\xi_1, \xi_2, \dots, \xi_i$ have been packed (in that order), where $\xi = (\xi_1, \xi_2, \dots, \xi_{t(z)})$. The norm $X(w(z, \xi, i))$ is simply the number of empty slots in this derived set of bins.

The transition function p for $\{I'_n\}$ is defined as follows: For states $z \in Z$, $p_z(w(z', \xi, i))$ is nonzero only if $z' = z$ and $i = 1$, with all $j^{t(z)}$ such states being equally likely. We have $p_{w(z, \xi, i)}(w(z, \xi, i+1)) = 1$ for states $w(z, \xi, i) \in W$ with $i < t(z) - 1$, and 0 for all other states, i.e., the transition is deterministic. Finally, the transition from $w(z, \xi, t(z) - 1)$ is also deterministic, with the state to which the transition is made being the state $z' \in Z$ that would result from the phase starting at z if the $t(z)$ items to arrive had the sizes and order specified by ξ . Note that $\{I'_n\}$ inherits from $\{I_{n_h}\}$ the properties of being irreducible and aperiodic.

LEMMA 11. *Let π be the equilibrium distribution for $\{I_{n_h}\}$, and let $E_\pi[t]$ denote the expected value of $t(z)$ under π , i.e., $\sum_{z \in Z} \pi(z)t(z)$. Then $\{I'_n\}$ has an equilibrium*

distribution π' where

$$\begin{aligned} \pi'(z) &= \frac{\pi(z)}{E_\pi[t]}, \quad z \in Z, \\ \pi'(w(z, \xi, i)) &= \frac{\pi(z)}{j^{t(z)} E_\pi[t]}, \quad z \in Z, \xi \in \{1, 2, \dots, j\}^{t(z)}, 1 \leq i < t(z). \end{aligned}$$

Proof. Note that $E_\pi[t]$ is a well-defined constant, since $E_\pi[X] \equiv \sum_{z \in Z} \pi(z)X(z)$ is bounded by Lemma 10, and $t(z) \leq 4k^6 X(z)$ by (4.1). Thus π' is well defined. To prove Lemma 11 we use a standard result from [12, p. 393].

THEOREM (Feller [12]). *Suppose we are given an aperiodic, irreducible Markov chain on a countable state space S with transition function p , and there exists a function $\mu : S \rightarrow [0, 1]$ such that (a) $\sum_{s \in S} \mu(s) = 1$ and (b) for all states $s' \in S$, $\mu(s') = \sum_{s \in S} \mu(s)p_s(s')$. If we let s_n denote the state after n steps, then we have for all states $s, s' \in S$,*

$$\lim_{n \rightarrow \infty} P[s_n = s' | s_0 = s] = \mu(s').$$

Given this theorem, we prove Lemma 11 by showing that (a) and (b) hold for π' . For (a), we have

$$\begin{aligned} \sum_{z \in Z} \pi'(z) + \sum_{w \in W} \pi'(w) &= \frac{\sum_{z \in Z} \pi(z)}{E_\pi[t]} + \sum_{z \in Z} \sum_{i=1}^{t(z)-1} \sum_{\xi \in \{1, 2, \dots, j\}^{t(z)}} \frac{\pi(z)}{j^{t(z)} E_\pi[t]} \\ &= \sum_{z \in Z} \frac{\pi(z)}{E_\pi[t]} + \sum_{z \in Z} \frac{\pi(z)(t(z) - 1)}{E_\pi[t]} \\ &= \sum_{z \in Z} \frac{\pi(z)t(z)}{E_\pi[t]} = 1. \end{aligned}$$

Property (b) holds for all states $w(z, \xi, i) \in W$ by design, and for states $z \in Z$ by the fact that π is the equilibrium distribution for $\{I_{n_h}\}$. Thus the above theorem of Feller applies and the proof of Lemma 11 is complete. \square

There are now only two more steps to completing a proof that our deterministic algorithm has bounded expected waste. We must simply verify the following two claims:

$$(4.7) \quad E_{\pi'}[X] \equiv \sum_{x \in Z \cup W} \pi'(x)X(x) < \infty,$$

$$(4.8) \quad \lim_{n \rightarrow \infty} \sum_{x \in Z \cup W} P[I'_n = x] \cdot X(x) = E_{\pi'}[X].$$

Claim (4.8) follows from (4.7) and Theorem 14.0.1 of [18], which when specialized to countable chains goes as follows.

THEOREM (Meyn and Tweedie [18]). *Suppose $\{\Phi_k\}$ is an irreducible, aperiodic Markov chain on a countable state space Z having equilibrium distribution π . Let π_n be the distribution of states at step n for some fixed initial configuration, and suppose $f : Z \rightarrow [0, \infty]$. Then if $E_\pi[f(z)] < \infty$, we have*

$$\lim_{n \rightarrow \infty} E_{\pi_n}[f(z)] = E_\pi[f(z)].$$

For (4.7) we first need to bound $X(x)$ for $x \in W$. Since $X(x)$ can increase by no more than $k - 1$ each time an item is packed, we conclude by (4.1) that

$$(4.9) \quad X(w(z, \xi, i)) \leq X(z) + (k - 1)t(z) < 4k^7 X(z).$$

Thus

$$\begin{aligned} \sum_{x \in Z \cup W} \pi'(x)X(x) &\leq \sum_{z \in Z} \frac{\pi(z)X(z)}{E_\pi[t]} + \sum_{z \in Z} \sum_{i=1}^{t(z)-1} \sum_{\xi \in \{1,2,\dots,j\}^{t(z)}} \frac{\pi(z)X(w(z, \xi, i))}{j^{t(z)} E_\pi[t]} \\ &\leq \frac{1}{E_\pi[t]} \sum_{z \in Z} \pi(z)t(z)4k^7 X(z) \\ &\leq \frac{1}{E_\pi[t]} \sum_{z \in Z} \pi(z)16k^{13} X(z)^2 = \frac{16k^{13}}{E_\pi[t]} \sum_{z \in Z} \pi(z)X(z)^2. \end{aligned}$$

Lemma 10 says that $\sum_{z \in Z} \pi(z)X(z)^2$ is bounded. Hence (4.7) holds.

5. Proof of Theorem 3. Recall the theorem statement: If L_n has item sizes generated according to $U(0, u]$ for $0 < u < 1$, and A is any on-line algorithm, then there exists a constant $c > 0$ such that $E[W^A(L_n)] > cn^{1/2}$ for infinitely many n .

Proof. Let $w(t)$ denote the amount of empty space in partially filled bins after t items have been packed. We show that for any $n > 0$ the expected value of the average of $w(1), \dots, w(n)$ is $\Omega(n^{1/2}u^3)$. This implies that $E[w(n)]$ must be $\underline{\Omega}(n^{1/2}u^3)$, i.e., not $o(n^{1/2})$.

Consider packing item a_{t+1} . Let $v(t)$ denote the number of nonempty bins that have a gap of at least $u^2/8$ after the first t items have been packed. There are at most $v(t)$ 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, either it must have size less than $u^2/8$ or its size must be within δ of the empty space in one of these $v(t)$ bins with gaps larger than $u^2/8$. The probability of this is at most $[u^2/8 + \delta v(t)]/u$. By choosing $\delta = u^2 n^{-1/2}/8$, conditioning on whether $v(t)$ is greater or less than $n^{1/2}$ and noting that the size of a_{t+1} is distributed as $U(0, u]$ independent of $v(t)$, we have

$$P(a_{t+1} \text{ leaves gap} < \delta) \leq P(v(t) \geq n^{1/2}) + u/4.$$

Now

$$\begin{aligned} E[w(t)] &\geq \delta \sum_{s=0}^{t-1} P(a_{s+1} \text{ is last in a bin and leaves gap} \geq \delta) \\ &= \delta \sum_{s=0}^{t-1} [P(a_{s+1} \text{ is last in a bin}) \\ &\quad - P(a_{s+1} \text{ is last in a bin and leaves gap} < \delta)] \\ &\geq \delta \sum_{s=0}^{t-1} [P(a_{s+1} \text{ is last in a bin}) - P(a_{s+1} \text{ leaves gap} < \delta)] \\ &\geq \delta \sum_{s=0}^{t-1} \left[P(a_{s+1} \text{ is last in a bin}) - \sum_{s=0}^{t-1} P(v(s) \geq n^{1/2}) - \sum_{s=0}^{t-1} u/4 \right]. \end{aligned}$$

Let S_t be the sum of the first t item sizes and note that S_t is a lower bound on the number of bins and hence on the number of items that are the last item in a bin. We thus have

$$E \left[\sum_{s=0}^{t-1} P(a_{s+1} \text{ is last in a bin}) \right] \geq E[S_t] = tu/2.$$

Using the fact that $\delta = u^2 n^{-1/2}/8$, we then have

$$E[w(t)] \geq (u^2 n^{-1/2}/8) \left[tu/4 - \sum_{s=0}^{t-1} P(v(s) \geq n^{1/2}) \right].$$

If $\sum_{s=0}^{n-1} P(v(s) \geq n^{1/2}) \leq nu/24$, we have for all $t \geq n/2$,

$$E[w(t)] \geq (u^2 n^{-1/2}/8)[nu/8 - nu/24] = u^3 n^{1/2}/96.$$

This implies

$$E \left[\frac{1}{n} \sum_{t=1}^n w(t) \right] \geq u^3 n^{1/2}/192.$$

On the other hand, if $\sum_{s=0}^{n-1} P(v(s) \geq n^{1/2}) \geq nu/24$ then

$$\begin{aligned} E \left[\frac{1}{n} \sum_{t=1}^n w(t) \right] &\geq \frac{1}{n} \sum_{t=1}^n P(v(t) \geq n^{1/2}) n^{1/2} (u^2/8) \\ &\geq n^{1/2} (u^2/8) (u/24) = u^3 n^{1/2}/192. \end{aligned}$$

These imply that $E[w(n)]$ is $\Omega(n^{1/2})$. \square

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 $U\{j, k\}$ do not have this failing and for this reason we can obtain significantly better average-case behavior for them.

6. Concluding remarks. In this paper we have proved a combinatorial theorem about perfect packings and used it to derive results about the expected behavior of online bin packing algorithms under the discrete uniform distributions $U\{j, k\}$, $1 < j \leq k - 2$, in which item sizes are restricted to the finite set $\{i/k : 1 \leq i \leq j\}$, each item size being equally likely. These results, which imply that for any such distribution there exist on-line algorithms with bounded expected waste, contrast distinctly with the result proved in section 5 about the corresponding continuous uniform distributions $U[0, j/k]$, for which expected waste must grow as $n^{1/2}$. We presented both randomized algorithms and deterministic ones, with the price for the deterministic ones being an apparent increase in both running time and expected waste. It should be noted, however, that in both cases the upper bounds we have proved on expected waste are quite large, and even though more careful arguments might enable us to lower the estimated bounds substantially, it is not clear that the “constant” waste provided by our algorithms would in practice be better than the slowly growing waste that other heuristics might provide, especially for large values of k .

One additional weakness of our results as proved in sections 3 and 4 is that they require a distinct algorithm for each j, k , rather than yielding a single algorithm that works and provides bounded expected waste for all $U\{j, k\}$, $1 < j \leq k - 2$. It is not difficult to see, however, that our underlying lemmas and their proofs can be used to provide general on-line algorithms (both randomized and deterministic) with running times of the form $O(np(k))$, where $p(k)$ is a polynomial in k that bounds the time to

pack an individual item. The parameters k and j can either be given as part of the input or, better yet, derived empirically by observing the first-arriving items.

Given j and k , the general algorithms determine the required set of perfect packing configurations $\{\vec{c}_h : 1 \leq h \leq H\}$ by implementing the proof of Lemma 7. The required values of r_i, s_i, m_i and (in the deterministic case) r'_i, s'_i, m'_i can be determined using the arguments in the proofs of that lemma and Lemma 8. It is then a simple exercise in data structures (which we leave to the reader) to verify that the packing operations for both the randomized and deterministic algorithms can be performed in worst-case polynomial time (in k), independent of n .

Note, however, that these “generalized” algorithms are still quite limited, in that they only know how to pack items generated according to the set of probability distributions $U\{j, k\}$. They may not even construct legal packings for other discrete distributions, even though many such distributions also have on-line algorithms that yield constant expected waste. This follows from the theorem of Courcoubetis and Weber [9] cited in section 3, which relates the existence of such algorithms for a distribution \vec{p} to the existence of appropriate perfect packing configurations. Although as we have remarked it is NP-complete to tell for an arbitrary distribution which case of the Courcoubetis–Weber theorem applies, it is easy to construct individual distributions other than the $U\{j, k\}$ that satisfy the hypotheses of this theorem, and it is likely that there are other general classes of interesting distributions for which the needed perfect packing configurations can be proved to exist by appropriate analogues of our perfect packing theorem. We leave the investigation of such questions to future research.

One useful tool may be the recent result of [10, 11] that for discrete distributions with integer item sizes one can determine which case of the Courcoubetis–Weber theorem applies in pseudopolynomial time (time bounded by a polynomial in the bin capacity B , rather than polynomial in $\log B$ as is required by the definition of polynomial time). For the integer item size case, these references also show that complicated algorithms such as those presented here will not be necessary when dealing with distributions for which $E[W^{\text{OPT}}(L_n)]$ is sublinear. In [10] it is shown that a simple deterministic $O(nB)$ -time on-line variant on the *Sum-of-Squares* algorithm of [11] has $O(1)$ expected waste for all distributions with $E[W^{\text{OPT}}(L_n)] = O(1)$ and $\Theta(\sqrt{n})$ expected waste for all distributions with $E[W^{\text{OPT}}(L_n)] = \Theta(\sqrt{n})$.

Acknowledgment. We are grateful to a referee for many helpful comments, including a major simplification to the proof of (4.7).

REFERENCES

- [1] 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 of the 16th Annual ACM Symposium on Theory of Computing, ACM, New York, 1984, pp. 279–288.
- [2] E. G. COFFMAN, JR., C. COURCOUBETIS, M. R. GAREY, D. S. JOHNSON, L. A. MCGEOCH, P. W. SHOR, R. R. WEBER, AND M. YANNAKAKIS, *Fundamental discrepancies between average-case analyses under discrete and continuous distributions*, in Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, ACM, New York, 1991, pp. 230–240.
- [3] E. G. COFFMAN, JR., D. S. JOHNSON, L. A. MCGEOCH, P. W. SHOR, AND R. R. WEBER, *Bin Packing with Discrete Item Sizes, Part III: Average Case Behavior of FFD and BFD*, in preparation.
- [4] E. G. COFFMAN, JR., D. S. JOHNSON, P. W. SHOR, AND R. R. WEBER, *Bin Packing with Discrete Item Sizes, Part IV: Markov Chains, Computer Proofs, and Average-Case Analysis of Best Fit Bin Packing*, in preparation.

- [5] E. G. COFFMAN, JR., D. S. JOHNSON, P. W. SHOR, AND R. R. WEBER, *Bin Packing with Discrete Item Sizes, Part V: Tight Bounds on Best Fit*, in preparation.
- [6] E. G. COFFMAN, JR., D. S. JOHNSON, P. W. SHOR, AND R. R. WEBER, *Markov chains, computer proofs, and average-case analysis of best fit bin packing*, in Proceedings of the 25th Annual ACM Symposium on Theory of Computing, ACM, New York, 1993, pp. 412–421.
- [7] E. G. COFFMAN, JR., D. S. JOHNSON, P. W. SHOR, AND R. R. WEBER, *Bin packing with discrete item sizes, part II: Tight bounds on first fit*, Random Structures Algorithms, 10 (1997), pp. 69–101.
- [8] E. G. COFFMAN, JR., AND G. S. LUEKER, *An Introduction to the Probabilistic Analysis of Packing and Partitioning Algorithms*, Wiley, New York, 1991.
- [9] C. COURCOUBETIS AND R. R. WEBER, *Stability of on-line bin packing with random arrivals and long-run average constraints*, Probab. Engrg. Inform. Sci., 4 (1990), pp. 447–460.
- [10] J. CSIRIK, D. S. JOHNSON, C. KENYON, J. B. ORLIN, P. W. SHOR, AND R. R. WEBER, *On the sum-of-squares algorithm for bin packing*, in Proceedings of the 32nd ACM Symposium on Theory of Computing, ACM, New York, 2000, to appear.
- [11] J. CSIRIK, D. S. JOHNSON, C. KENYON, P. W. SHOR, AND R. R. WEBER, *A self organizing bin packing heuristic*, in Algorithm Engineering and Experimentation, M. Goodrich and C. C. McGeoch, eds., Lecture Notes in Comput. Sci. 1619, Springer-Verlag, Berlin, 1999, pp. 246–265.
- [12] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd ed., Wiley, New York, 1968.
- [13] S. FLOYD AND R. M. KARP, *FFD bin packing for item sizes with distributions on $[0, 1/2]$* , Algorithmica, 6 (1991), pp. 222–240.
- [14] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability: A Guide to the Theory of NP-completeness*, W. H. Freeman, New York, 1979.
- [15] C. KENYON, Y. RABANI, AND A. SINCLAIR, *Biased random walks, Lyapunov functions, and stochastic analysis of best fit bin packing*, in Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, Atlanta, GA, 1996, pp. 351–358.
- [16] W. KNÖDEL, *A bin packing algorithm with complexity $o(n \log n)$ in the stochastic limit*, in Proceedings of the 10th Symposium on Mathematical Foundations of Computer Science, J. Gruska and M. Chytil, eds., Lecture Notes in Comput. Sci. 118, Springer-Verlag, Berlin, 1981, pp. 369–378.
- [17] G. S. LUEKER, *An Average-Case Analysis of Bin Packing with Uniformly Distributed Item Sizes*, Tech. Report 181, Dept. of Information and Computer Science, University of California, Irvine, CA, 1982.
- [18] S. P. MEYN AND R. L. TWEEDIE, *Markov Chains and Stochastic Stability*, Springer-Verlag, Berlin, 1993.
- [19] P. W. SHOR, *The average case analysis of some on-line algorithms for bin packing*, Combinatorica, 6 (1986), pp. 179–200.
- [20] P. W. SHOR, *How to pack better than Best Fit: Tight bounds for average-case on-line bin packing*, in Proceedings of the 32nd Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamitos, CA, 1991, pp. 752–759.