SRC Technical Note
1997 - 002

February 21, 1997

Average-Case Analysis of First Fit and Random Fit Bin Packing

Susanne Albers and Michael Mitzenmacher

Systems Research Center
130 Lytton Avenue
Palo Alto, California 94301

http://www.research.digital.com/SRC/

Abstract

We prove that the First Fit bin packing algorithm is stable under the input distribution $U\{k-2, k\}$ for all $k \geq 3$, settling an open question from the recent survey by Coffman, Garey, and Johnson [3]. Our proof generalizes the multi-dimensional Markov chain analysis used by Kenyon, Rabani, and Sinclair to prove that Best Fit is also stable under these distributions [11]. Our proof is motivated by an analysis of Random Fit, a new simple packing algorithm related to First Fit, that is interesting in its own right. We show that Random Fit is stable under the input distributions $U\{k-2, k\}$, as well as present worst-case bounds and some results on distributions $U\{k-1, k\}$ and $U\{k, k\}$ for Random Fit.

1 Introduction

In the one-dimensional bin packing problem, one is given a sequence $a_1, \ldots, a_n \in (0, 1]$ of items to pack into bins of unit capacity so as to minimize the number of bins used. A great deal of literature has focused on this problem, perhaps because, as Coffman, Garey, and Johnson [3] observe in their recent survey on bin packing, "The classical one-dimensional bin packing problem has long served as a proving ground for new approaches to the analysis of approximation algorithms." For example, recently the study of Best Fit bin packing under discrete uniform distributions has led to a novel analysis technique, based on the theory of multi-dimensional Markov chains. In this paper we extend this approach to analyze First Fit and a new bin packing algorithm, called Random Fit, under discrete uniform distributions.

First Fit and Best Fit are two classical algorithms for online bin packing. With First Fit, the bins are indexed in increasing order of their creation. Each item is sequentially placed into the lowest indexed bin into which it will fit, or into a empty bin if no such bin is available. With the Best Fit algorithm, each incoming item is placed into the non-empty bin with smallest residual capacity that can contain it; if no such bin exists, the item is placed in an empty bin. The performance of First Fit and Best Fit in the worst case and uniform average case has been settled for quite some time. In the worst case, the number of bins used by any of these algorithms is at most $\frac{17}{10}$ times the optimum number of bins, as shown by Johnson *et al.* [10]. When item sizes are generated by $U(0, 1)$, the continuous uniform distribution on $(0, 1]$, then the performance measure of interest is the *expected waste*, which is the difference between the number of bins used and the total size of the items packed so far. Shor [16] showed that the expected waste created by First Fit is $\Theta(n^{2/3})$. Shor [16] and Leighton and Shor [13] proved that Best Fit does better, generating expected waste $\Theta(\sqrt{n}\log^{3/4} n)$.

Because of these tight bounds, research on First Fit and Best Fit is now focused on analyzing expected waste when item sizes are generated by discrete uniform distributions. A discrete uniform distribution, denoted by $U\{j, k\}$, $1 \leq j \leq k$, is one where item sizes are chosen uniformly from the set $\{1/k, 2/k, \ldots, j/k\}$. For $U\{k, k\}$, $k > 1$, First Fit and Best Fit achieve expected waste $\Theta(\sqrt{nk})$ and $O(\sqrt{n}\log k)$, respectively, (see Coffman *et al.* [2]). Similar bounds hold for $U\{k-1, k\}$. Of particular interest are distributions for which the algorithms are *stable*. We say that a algorithm is stable under a distribution if the expected waste remains bounded (that is, $O(1)$), even as the number of items n goes to infinity. Coffman *et al.* [2] proved that First Fit is stable under $U\{j, k\}$, when $k \geq j^2$, and Best Fit is stable under $U\{j, k\}$, when $k \geq j(j+3)/2$. Later, Coffman *et al.* [4] introduced a novel method for proving the stability (and instability) of bin packing algorithms based on multi-dimensional Markov chains. Their methodology allowed them to show that $U\{j, k\}$ is stable under Best Fit for several specific pairs of values for j and k. Kenyon *et al.* [11] expanded on this work by proving that Best Fit is stable under the entire family of distributions $U\{k-2, k\}$, using a complex analysis of the underlying Markov chains.

We briefly describe the Markov chain setting used in the results described above. Using the Best Fit algorithm under a discrete uniform distribution, a packing can be represented by the number of bins of each possible residual capacity. The order of the bins is irrelevant. This packing process can therefore be easily represented by a Markov chain, where the state at any time is a vector $s = (s_1, \ldots, s_{k-1})$, and s_i is the number of bins of residual capacity i/k.

The Best Fit algorithm is well suited to the Markov chain approach, because the order of the bins is irrelevant, leading to a simple representation of the packing. In contrast, in the First Fit algorithm, the order of the bins cannot be dismissed. Because of the difficulty of representing the state in the First Fit algorithm, until now these Markov chain techniques have not been successfully applied to the First Fit algorithm.

In this paper, we remedy this problem by demonstrating a Markov chain argument that shows that First Fit is in fact stable under the family of distributions $U\{k-2, k\}$. This result disproves a conjecture made by Coffman *et al*. [3], who state that limited experiments suggest that the expected waste may grow unbounded on $U\{k-2, k\}$ for sufficiently large k. Moreover, it demonstrates that the Markov chain approach may be more generally applicable than previously believed.

Our proof emerges from an analysis of a new bin packing algorithm, called *Random Fit* (RF). Random Fit is a simple randomized variant of First Fit. With Random Fit, each time an item is to be placed in a bin the bins are indexed in an order determined by a permutation chosen independently and uniformly at random. Each item is sequentially placed into the lowest indexed bin into which it will fit, or into a empty bin if no such bin is available.

In Section 2 we introduce Random Fit by analyzing its worst-case behavior. In the following sections we then concentrate on average-case analysis. Random Fit has the advantage that, like Best Fit, a packing can be represented by the number of bins of each possible residual capacity. Therefore, in Section 3, we first generalize the analysis of Best Fit shown in [11] to Random Fit. We prove stability of Random Fit under the input distribution $U\{k-2, k\}$ and derive some related results for $U\{k-1, k\}$ and $U\{k, k\}$. Using ideas developed in Section 3, we proceed to prove stability of First Fit under input distribution $U\{k-2, k\}$ in Section 4.

2 Worst-case analysis of Random Fit

We compare the behavior of Random Fit with an optimal offline algorithm. Recall that with Random Fit, each time an item is to be placed in a bin the bins are indexed in an order determined by a permutation chosen independently and uniformly at random. Each item is sequentially placed into the lowest indexed bin into which it will fit, or into a empty bin if no such bin is available.

Given a sequence $S = (a_1, a_2, \ldots, a_n)$ of items and a bin packing algorithm A, let $A(S)$ denote the number of bins used by A to pack S. In particular, OPT(S) is the number of bins used by an optimal offline algorithm, i.e., it is the minimum number of bins required to pack S.

Theorem 1 *a) For every sequence S, $RF(S) \leq 2 \cdot OPT(S) - 1$.*

b) There exist sequences S, with arbitrarily large values of OPT(S), such that with high probability $RF(S) = 2 \cdot OPT(S) - 1$.

Proof: *Part a)* At any time, the sequence of bins used by RF contains at most one bin with residual capacity of at least $\frac{1}{2}$. Thus, for any sequence S, the number of bins used by OPT is at least $\lfloor \frac{1}{2} RF(S) \rfloor + 1$.

Part b) For any integer $n \geq 2$, let S_n be a sequence that contains n *large* items of size $\frac{1}{2}$. In addition, in between any two large items, n^2 *small* items each of size $\frac{1}{2n^3}$ must be inserted. Thus

$$S_n = (\frac{1}{2}, \frac{1}{2n^3}, \ldots, \frac{1}{2n^3}, \frac{1}{2}, \frac{1}{2n^3}, \ldots, \frac{1}{2n^3}, \frac{1}{2}, \frac{1}{2n^3}, \ldots, \frac{1}{2n^3}, \frac{1}{2}).$$

Note that the sum of all the small items is $\frac{1}{2n^3} n^2 (n-1) < \frac{1}{2}$.

Clearly, OPT$(S_n) = \lfloor \frac{n}{2} \rfloor + 1$. We show that with high probability Random Fit uses n bins on this sequence. More precisely, immediately before an insertion of a large item, the probability that a bin holding a large item does not contain a small item is bounded by $(1 - \frac{1}{n})^{n^2} \leq e^{-n}$. Thus, the probability that at any of the n insertions of large items, some open bin having a large item does not contain a small item is bounded by $n^2 e^{-n}$. We conclude that with probability at least $1 - \frac{n^2}{e^n}$, RF needs n bins to pack S_n. ■

While RF has a guaranteed worst-case performance, it does not achieve the same bounds as First Fit and Best Fit. In the worst case, RF is only as good as Next Fit and Worst Fit.

Motivated by recent work [1, 15], we also consider an extension of Random Fit, called *Random-Fit*(d), that is defined for any integer $d \geq 2$. Whenever a new item arrives, RF(d) examines bins in the same way as RF until d bins are found that can hold the item. Among these d bins, the item is inserted into the bin with smallest residual capacity, i.e., the Best Fit rule is applied. If there are only i, $i < d$, open bins that can hold the item, then the item is inserted into one of these i bins, using again the Best Fit strategy. If none of the open bins can hold the item, then the item is inserted into a new bin.

Interestingly, when making the transition from RF to RF(d), the worst-case performance improves.

For any algorithm A, let

$$R_A^\infty = \{r \geq 1 | \text{ for some } N > 0,\ A(S)/OPT(S) \leq r \ \text{ for all } S \text{ with } OPT(S) \geq N\}.$$

Theorem 2 *For every $d \geq 2$, $R_{RF(d)}^\infty \leq \frac{17}{10}$.*

Proof: Follows from a result by Johnson [7, 8] because RF(d) belongs to the class of Almost Any Fit algorithms. ■

3 Average-case analysis of Random Fit

In this section we prove that Random Fit is stable under the input distribution $U\{k-2, k\}$ and derive some related results for $U\{k-1, k\}$ and $U\{k, k\}$.

3.1 Preliminaries

We begin by reviewing some important definitions and lemmas from [11]. For considering the distribution $U\{j, k\}$, rather than have bins of size 1, we shall instead think of having bins of size k and item sizes chosen uniformly from $\{1, \ldots, j\}$. The two notions are clearly equivalent. We shall model the system using $k-1$ *tokens* that move on the non-negative integers. The value of token i at time t, denoted by $s_i(t)$ represents the number of bins with residual capacity i after t items have been placed. The *state* of the system at time t is given by a vector $s(t) = (s_1(t), ..., s_{k-1}(t))$. Initially, $s(0) = (0, \ldots, 0)$, as there are no open bins with residual capacity. We shall often drop the explicit reference on t when the meaning is clear. The *waste* at time t is given by $\sum_{i=1}^{k-1} i s_i(t)$. We wish to show that the expected waste as $t \to \infty$ remains bounded.

We shall divide the tokens into classes. The token i is called *small* if $1 \leq i \leq \lceil \frac{j}{2} \rceil$ and is called *large* if $\lfloor \frac{j}{2} + 2 \rfloor \leq i \leq j$. In the case where j is even, there is also a *middle* token, namely $\lceil \frac{j}{2} \rceil + 1$. For convenience, we shall temporarily restrict ourselves to the case where j is odd, as the case where j is even requires some additional work to handle the middle token. We shall explain the modifications necessary for the case where j is even after the proof of the case where j is odd.

We begin with the following lemma:

Lemma 3 *State s is reachable from the initial state $s(0) = (0, \ldots, 0)$ only if*

1. *For distinct indices i and i' with $i + i' \geq k$, either $s_i = 0$ or $s_{i'} = 0$.*
2. $\sum_{i \text{ not small}} s_i \leq 1$

Proof: By induction; it follows from the fact that we will not open a new bin if an item can be packed in a current bin. ■

It is also not hard to see that all states that satisfy the conditions of Lemma 3 are reachable, and hence we assume hereafter that our state space consists exclusively of all states satisfying the conditions of Lemma 3.

From Lemma 3, if $s_{\lceil \frac{j}{2} \rceil}(t) > 0$, then all large tokens must be 0 at time t. Our proof of stability will rely on this simple feature of the chain. In particular, this feature allows us to focus on the behavior of the small tokens, which is considered in the following lemma:

Lemma 4 *Using Random Fit, the motion of a small token i has the following properties:*

1. *For $i > 1$, the motion of s_i at all positions other than 0 is a random walk on Z_+, such that a positive step is taken with probability at least $\frac{1}{j}$ and a negative step is taken with probability at most $\frac{1}{j} + \frac{s_i}{s_{i-1}+s_i}$.*
2. *The time spent by s_i on each visit to 0 is stochastically dominated by a random variable D with constant expectation and variance (that depend only on j).*

Proof: For the first part, note that, if $s_i > 0$, then s_i increases whenever an element of size $k - i$ enters the system, by Lemma 3. Hence we need only consider negative steps. If an item of size i enters, then s_i may decrease; if an item of size less than i enters, then it is clear that the probability of it landing in a bin of capacity i it at most $\frac{s_i}{s_{i-1}+s_i}$. The result follows.

The second part is almost exactly the same as in Proposition 4 of [11], which we sketch here for completeness. If $s_i = 0$, and $s_{i'} = 0$ for all $i' \geq k - i$, then clearly s_i moves to 1 with probability at least $1/j$. If $s_{i'} = 1$ for some $i' \geq k - i$, however, this is not the case. It suffices to note that if two consecutive items have size $k - i$, then s_i will go to 1 even in this case. One may check that this fact suffices to prove the lemma. ■

3.2 Outline of the proof

We now sketch how we will prove that RF is stable, following the same approach as [11]. We first note that by Lemma 3, the amount of waste from non-small tokens is bounded by a constant. Hence we need only consider the waste due to small tokens, which we denote by $f(t) = \sum_{i=1}^{\lceil j/2 \rceil} i s_i(t)$.

The proof breaks down into three steps. For the first step, we show that if $s_{\lceil \frac{j}{2} \rceil}(t) > 0$, then the expected change in $f(t)$, or $\mathsf{E}[f(t+1) - f(t)|f(t)]$, is negative. For the second step, we show that if we begin a state where $f(t)$ is large, then for some suitably large T, for almost all of the next T steps $s_{\lceil \frac{j}{2} \rceil} > 0$ with a suitably high probability. Combining these two steps, we find that, whenever $f(t)$ is sufficiently large, the expected change in $f(t)$ is negative over a suitably long interval T. The third step is to this fact and results from the general theory of Markov chains that to show that we may conclude that the expected waste is bounded.

The challenging part of the proof is the second step, where we must show that $s_{\lceil \frac{j}{2} \rceil} > 0$ for most of a suitably large interval. The first step is actually a simple lemma, entirely similar to one given in [11]. However, since the lemma is heavily based on the fact that $j = k - 2$, we present a proof here.

Lemma 5 ([11], Proposition 5) *Suppose that $s_{\lceil \frac{j}{2} \rceil}(t) > 0$. Then $\mathsf{E}[f(t+1) - f(t)|f(t)] = -1/j$.*

Proof: Consider the size i of the item inserted at time $t+1$. If $1 \le i \le \lceil j/2 \rceil$, then the new item is assigned to a bin with remaining capacity l, $i \le l \le \lceil j/2 \rceil$, and f decreases by i. If $\lceil \frac{j}{2} \rceil < i \le j$, then, since $s_{\lceil j/2 \rceil} > 0$, Proposition 3 implies that there is no bin with remaining capacity i. Thus, the incoming item is put into a new bin, i.e., s_{k-i} increases by 1 and f increases by $k - i$. The expected change in f is therefore

$$\frac{1}{j}\left(\sum_{i=1}^{\lceil j/2 \rceil}(-i) + \sum_{i=\lceil j/2 \rceil+1}^{j}(k-i)\right) = \frac{1}{j}\left(\sum_{i=1}^{\lceil j/2 \rceil}(-i) + \sum_{i'=2}^{\lceil j/2 \rceil} i'\right), \tag{1}$$

because $k - j = k - (k-2) = 2$ and, since j is odd, $k - (\lceil j/2 \rceil + 1) = \lceil j/2 \rceil$. It is easy to verify that equation (1) evaluates to $-1/j$. ■

The third step relies on general conditions for a multi-dimensional Markov chain to be ergodic; we cite the appropriate lemma from [11], which is derived from [5].

Lemma 6 ([11], Lemma 6, or [5], Corollary 7.1.3) *Let $\mathcal{M}$ be an irreducible, aperiodic Markov chain with state space $S \subseteq \mathbf{Z}^k$, and b a positive integer. Denote by $p^b_{s\,s'}$ the transition probability from s to s' in $\mathcal{M}^b$, the b-step version of $\mathcal{M}$. Let $\Phi : S \to \mathbf{R}_+$ be a non-negative real-valued function on S which satisfies the following conditions:*

1. *There are constants $C_1, \mu > 0$ such that $\Phi(s) > C_1 ||s||^{\mu}$ for all $s \in S$.*
2. *There is a constant $C_2 > 0$ such that $p^b_{s\,s'} = 0$ whenever $|\Phi(s) - \Phi(s')| > C_2$, for all $s, s' \in S$.*
3. *There is a finite set $B \subset S$ and a constant $\epsilon > 0$ such that $\sum_{s' \in S} p^b_{s\,s'}(\Phi(s') - \Phi(s)) < -\epsilon$ for all $s \in S \backslash B$.*

Then $\mathcal{M}$ is ergodic with stationary distribution π satisfying $\pi(s) < C\mathrm{e}^{-\delta\Phi(s)}$ for all $s \in S$, where C and δ are positive constants.

For the bin-packing problem, we shall use $\Phi(s) = \sum_{i=1}^{\lceil \frac{j}{2} \rceil} i s_i + k - 1 = f + k - 1$, where f is the waste from small tokens. This is an upper bound on the total waste. One may check that the first two conditions of Lemma 6 are satisfied for any choice of b. It remains to find appropriate b, B, and ϵ; this is equivalent to the second step of our proof sketch, which we now focus on.

3.3 Random Fit over long intervals

We now show that, for all but a finite number of starting states, $s_{\lceil \frac{j}{2} \rceil} > 0$ for most of sufficiently large intervals. We shall often compare the behavior of a token with a *random walk* over an interval $[0, R]$. We shall use $p_{\uparrow}(i)$ to denote the probability that a walk at i moves to $i + 1$ in one step. Similarly $p_{\downarrow}(i)$ is the probability that a walk at i moves to $i - 1$ in one step, and $p_{\rightarrow} = 1 - p_{\uparrow}(i) - p_{\downarrow}(i)$ (the *self-loop probability*) is the probability that the walk remains at i when at i. We shall drop the i in cases where $p_{\uparrow}(i)$ is independent of i (except at 0 and R, as $p_{\downarrow}(0)$ and $p_{\uparrow}(R)$ are necessarily 0, and the self-loop probabilities are increased accordingly); this is called the *homogeneous* case. A random walk is *downward biased* if $p_{\uparrow}(i) \leq p_{\downarrow}(i)$ for all i in the range of the walk (except the boundaries).

In order to bound the behavior of the random walks we study, we shall require the following lemma, which is a weak bound derived from Corollary 4.2 of [12]:

Lemma 7 *Let $\lambda_1 < 1$ denote the second largest eigenvalue of the transition matrix for a random walk W on $[0, R]$. Let $\pi(A) = \sum_{a \in A} \pi_a$ be the stationary probability that the walk lies in $A \subset R$, and $W_l(A)$ be the number of steps the walk spends in A during the first l time steps. If the walk starts at 0, then for any integer $l \geq 1$ and $2 \leq \beta < 1/\pi(A)$,*

$$\mathbf{Pr}[W_l(A) \geq \beta\pi(A)l] \leq \frac{\beta}{\sqrt{\pi_0}} \exp\Big(-\pi(A)^2(1 - \lambda_1)l\Big).$$

To use the above lemma we will require the following fact about the eigenvalues:

Lemma 8 *For the random walk on $[0, R]$ with $p_{\uparrow} = p_{\downarrow} = \alpha$, $\lambda_1 \leq 1 - \frac{2\alpha}{R^2}$.*

We start with a preliminary lemma that provides both the first step and the main idea of the proof. In this lemma, and all that follows, we assume that T is at least as large as some constant chosen so that the bounds hold.

Lemma 9 *For sufficiently large T, if $s_i > T^4$ over the time interval $[0, T]$, then $s_{i+1} \geq T^{1/16}$ for all but at most $T^{15/16}$ steps with probability at least $1 - \frac{2}{T^2}$.*

Proof: By Lemma 4, the behavior of the token s_{i+1} at any point on the interval $[0, T]$ can be related to a random walk over the positive integers, where $p_{\uparrow}(i) \geq 1/j$ and $p_{\downarrow}(i) \leq \frac{1}{j} + \frac{s_i}{s_i + s_{i+1}}$ (except at $i = 0$).

Furthermore, the probability that $s_{i+1} \geq T^{1/16}$ for all but at most $T^{15/16}$ steps, which we shall hereafter call z, is clearly minimized if we start s_{i+1} at 0. This information is sufficient to prove that $z \geq \frac{2}{T^2}$ directly; however, we suggest an easier approach.

We first note that, since we are comparing the behavior of s_{i+1} to a specific random walk, z can only increase if we restrict the walk (or, equivalently, the token s_{i+1}) to the interval $[0, T^{1/4} - 1]$. Bounding the walk in this way will simplify the analysis. Also, for convenience, we also temporarily ignore the problem of the waiting time when $s_{i+1} = 0$ as described in Lemma 6.

We now split each step, or item arrival, into two phases. In phase one, a random permutation order is chosen for the open bins. In phase two, an item size is chosen from the distribution $U\{j, k\}$, and this item is placed according to the RF rule.

By breaking each step up in this manner, we see that whenever the permutation chosen in phase one has a bin with remaining capacity i ahead of all bins of remaining capacity $i + 1$, then for phase two, the worst possible case is that s_{i+1} behaves like an unbiased random walk, with $p_\uparrow = p_\downarrow = 1/j$. (Note that it is possible that $p_\downarrow \leq 1/j$, but we maximize the time that $s_{i+1} \geq T^{1/16}$ by assuming that $p_\downarrow = 1/j$.) In the alternate case where a bin with remaining capacity $i + 1$ lies ahead of all bins of capacity i in phase one, we may again overestimate z by assuming that $p_\uparrow = 0$ and $p_\downarrow = 1$ in phase two. As we now show, by splitting each step into two phases in this way, we have essentially reduced the problem to an unbiased walk.

We note that, over the interval $[0, T]$ we have enforced the restrictions $s_{i+1} \leq T^{1/4}$ and $s_i \geq T^4$. Hence, with probability at least $\frac{1}{T^2}$, for *no* steps in this interval do we place a bin of capacity $i + 1$ ahead of all bins of capacity i in phase one. We call this event $\mathcal{E}$. Conditioned on $\mathcal{E}$, s_{i+1} behaves like an unbiased random walk on $[0, T^{1/4} - 1]$ over the entire interval. In particular, the stationary distribution is uniform, so $\pi_i = T^{-1/4}$ for all i. Let Z be the number of steps for which $s_{i+1} \leq T^{1/16}$. From Lemmas 7 and 8, we find that for sufficiently large T,

$$\begin{aligned} \mathbf{Pr}[Z \geq T^{15/16} \mid \mathcal{E}] &\leq T^{1/8} \cdot T^{1/8} \exp\Big(\frac{-2T^{1/8}}{j}\Big) && (2) \\ &\leq \frac{1}{T^2}. && (3) \end{aligned}$$

Using a union bound on probabilities now yields the lemma.

To handle the discrepancy when the walk is at 0, we note that we can explicitly bound the total number of steps at 0 with sufficiently high probability using part 2 of Lemma 4. The bounds given by equations (2) and (3) can also be tightened so that for sufficiently large T, the lemma as stated holds. ■

We have shown that if s_i is extremely large over a sufficiently long interval, then s_{i+1} is also be large over most of the interval with high probability. Our actual goal is to show that if any s_i is extremely large (for $i \leq \lceil \frac{j}{2} \rceil$), then $s_{\lceil \frac{j}{2} \rceil} > 0$ over most of the interval. Hence we will require an inductive, but slightly weaker, version of Lemma 9.

One problem in generalizing Lemma 9 is that if s_i is large only for most, and not all, of the steps, then there are several steps where we cannot explicitly say how s_{i+1} behaves. Moreover, these steps may affect the behavior of s_{i+1} at any point. We avoid the problem by introducing an adversary model, generalizing a similar argument from [11]. This adversary model allows us to consider the worst possible case for the steps where s_i is smaller than we need.

We consider how an adversary can affect a homogeneous downward biased random walk on $[0, R]$. The goal of the adversary is to keep the random walk at or below some level l, $l \geq 2$, for as many steps as possible. The adversary may control a fixed number of steps. In a controlled step, the adversary may specify any probability distribution on the legal moves from the current state; the step of the walk is then made according to that distribution. In all the other steps, the process behaves like a homogeneous downward biased random walk.

In the following, given an adversary strategy A, let $p_A(y, i, n, m, l)$ denote the probability that a homogeneous downward biased random walk of n steps on the interval $[0, R]$ starting at i with y controlled steps used according to A, spends at least m steps at or below l.

Lemma 10 *For all non-negative integers y, i, n, m and l, with $l < R$ and $i < R-1$, the exists an adversary strategy A_0*

(a) that never uses a controlled step when the walk is below l

(b) that always uses a controlled step as soon as possible when the walk is at or above $l+1$ to push the walk downwards

such that $p_{A_0}(y, i, n, m, l) \geq p_A(y, i, n, m, l)$ for all adversaries A.

Proof: The case where $l = 0$, the walk is unbiased, and the self-loop probability is 0 corresponds to what is proven in [11, Lemma 7]; we extend the argument to this more general case. We use induction on n. We first note that any adversary that uses a downward move when the walk is below l can be replaced by one that does not. This follows by a simple coupling argument. Compare the strategy where the adversary uses a downward move below l to one where the adversary waits until the walk is at l by coupling all random moves; the second strategy will be at the same height or below the first after the downward move. (It will end up below only if the walk reaches 0.) Thus we have shown that there is an optimal adversary strategy that satisfies condition (a).

We now concentrate on adversary strategies that use their moves at or above $l+1$. Let D^yR denote the strategy A_1 which uses the y adversary-controlled steps as soon as possible when the walk is at or above $l+1$, and then follows the random walk. Let RD^y denote the strategy A_2 that begins with a random step, and then uses the adversary-controlled steps as soon as possible when the walk is at or above $l+1$. Let $p_{A_1}(y, i, n, m, l)$ be the probability of the event that the walk is at or below l for at least m of the next n steps after starting at i when adversary strategy $A_1 = D^yR$ is used. Similarly, let $p_{A_2}(y, i, n, m, l)$ be the probability of the event that the walk is at or below l for at least m of the next n steps after starting at i when adversary strategy $A_2 = RD^y$. We claim

$$p_{A_1}(y, i, n, m, l) \geq p_{A_2}(y, i, n, m, l), \tag{4}$$

and by induction this suffices to prove that there is an optimal strategy satisfying condition (b).

We first present two useful propositions.

Proposition 11 $p_{A_1}(y, l, n, m, l) \geq p_{A_1}(y, l, n, m+1, l)$

Proposition 12 $p_{A_1}(y, l-1, n, m, l) \geq p_{A_1}(y+1, l, n+1, m+1, l)$

Proposition 11 is easy to verify. We prove Proposition 12. Let W_{l-1} be the walk that starts at $l-1$ and follows strategy $D^y R$; similarly let W_l be the walk that starts at l and follows strategy $D^{y+1} R$. Let T_{l-1} be the time when W_{l-1} first makes the transition $(l-1) \to l$ and let T_l be the time when W_l first makes the transition $l \to (l+1)$. Clearly, $T_{l-1} = T_l$ in distribution. We only have to consider the event that $T_{l-1} = T_l \leq n+1$ and $T_{l-1} = T_l \geq m$. Then, the remainder of W_{l-1} is a walk starting at l that follows strategy $D^y R$ and must be at or below l for at least $m - T_{l-1}$ of the next $n - T_{l-1}$ steps. In the case of W_l, the adversary first pushes the walk down to l and the remainder is also a walk that starts at l, follows strategy $D^y R$ and must be at or below l for at least $m+1-T_l = m+1-T_{l-1}$ of the next $n+1-(T_l+1) = n-T_l$ steps. Using Proposition 11 and taking again into account that $T_{l-1} = T_l$ in distribution, we conclude the probability of the first walk is not smaller than that of the second walk, i.e., $p_{A_1}(y, l-1, n, m, l) \geq p_{A_1}(y+1, l, n+1, m+1, l)$.

We return to the proof of inequality (4). If $i \leq l$, both strategies A_1 and A_2 start the same and we are done by induction. If $i > y + l$, both strategies give the same distribution after $y+1$ steps, so the two quantities $p_{A_1}(y, i, n, m, l)$ and $p_{A_2}(y, i, n, m, l)$ are equal. The interesting case is when $l < i \leq y + l$. In this case, strategy A_1 forces the walk from i down to l using $i - l$ controlled steps. Thus,

$$p_{A_1}(y, i, n, m, l) = p_{A_1}(y', l, n', m-1, l),$$

where $n' = n - i + l$ and $y' = y - i + l$. Also

$$\begin{aligned} p_{A_2}(y, i, n, m, l) &= p_{\uparrow} \cdot p_{A_1}(y'-1, l, n'-2, m-1, l) + p_{\downarrow} \cdot p_{A_1}(y'+1, l, n', m-1, l) \\ &\quad + p_{\rightarrow} \cdot p_{A_1}(y', l, n'-1, m-1, l) \end{aligned}$$

and

$$\begin{aligned} p_{A_1}(y', l, n', m-1, l) &= p_{\uparrow} \cdot p_{A_1}(y'-1, l, n'-2, m-2, l) + p_{\downarrow} \cdot p_{A_1}(y', l-1, n'-1, m-2, l) \\ &\quad + p_{\rightarrow} \cdot p_{A_1}(y', l, n'-1, m-2, l). \end{aligned}$$

Using Proposition 11, we have $p_{A_1}(y'-1, l, n'-2, m-2, l) \geq p_{A_1}(y'-1, l, n'-2, m-1, l)$ and $p_{A_1}(y', l, n'-1, m-2, l) \geq p_{A_1}(y', l, n'-1, m-1, l)$. Thus,

$$p_{A_1}(y, i, n, m, l) - p_{A_2}(y, i, n, m, l) \geq p_{\downarrow}(p_{A_1}(y', l-1, n'-1, m-2, l) - p_{A_1}(y'+1, l, n', m-1, l)).$$

Proposition 12 implies that the last term in non-negative. ■

Lemma 13 *Suppose, over a period of T steps, $s_{i-1} \geq T^{\alpha}$ over all but $T^{1-\alpha}$ steps for some $\alpha \leq 1/16$. Then $s_i \geq T^{\alpha/16}$ for all but at most $T^{1-\alpha/16}$ steps with probability at least $1 - 3T^{-\alpha/4}$.*

Proof: As in Lemma 9, we may, without loss of generality, restrict s_i to the interval $[0, T^{\alpha/4} - 1]$. Then s_i behaves like a slightly biased random walk on all but the $T^{1-\alpha}$ steps for which s_{i-1} lies below T^{α}. Rather than consider the biased walk, however, we use the same technique as in Lemma 9 to reduce the problem to an unbiased random walk by splitting each step into two phases. We give the adversary control on all steps in which a bin with remaining capacity i lies ahead of all bins with capacity $i-1$ after the first phase. On any step for which $s_{i-1} \geq T^{\alpha}$ and $s_i \leq T^{\alpha/4}$, the probability that a bin with remaining capacity i lies ahead of all bins with capacity $i-1$ after the first phase is at most $\frac{1}{T^{3\alpha/4}}$. Hence, the expected number of such steps is at most $T^{1-3\alpha/4}$, and by Markov's inequality, the number of such steps is at most $T^{1-\alpha}$ with probability at least $T^{-\alpha/4}$. Let $\mathcal{E}$ be the event that there are no more than $T^{1-\alpha}$ such steps.

Conditioned on $\mathcal{E}$, the adversary controls at most $2T^{1-\alpha}$ steps: $T^{1-\alpha}$ from the above paragraph, and $T^{1-\alpha}$ from the steps where $s_{i-1} < T^{\alpha}$. On all other steps the walk behaves like an unbiased random walk with $p_{\uparrow} = p_{\downarrow} = 1/j$. (Again, this is not quite true when $s_i = 0$, but this small discrepancy can be easily handled explicitly as described in Lemma 9; for convenience we dismiss the problem here.) We use this to bound the probability that s_i lies below $T^{\alpha/16}$ for more than $T^{1-\alpha/16}$ steps.

We first consider the moves controlled by the adversary. In the worst case, s_i begins at 0. By Lemma 10, there exists an optimal adversary strategy A_0 that uses a controlled step whenever s_i reaches $T^{\alpha/16} - 1$ or $T^{\alpha/16}$. Hence, to overestimate the effect of the adversary, we assume the following: the adversary uses its moves whenever s_i reaches $T^{\alpha/16}$; the adversary's move returns the walk to $s_i = 0$; and all steps until the adversary's moves are used count as steps where $s_i \in [0, T^{\alpha/16} - 1]$. These assumptions can only increase the time until the adversary's moves are used. The expected time for s_i to reach $T^{\alpha/16}$ from 0 is $cT^{\alpha/8}$ for some constant c. Thus the expected number of steps until A has used all of its moves it bounded by $cT^{1-7\alpha/8}$. Let Z_1 be the number of steps until the A_0 uses all of its moves. Then by Markov's inequality

$$\mathbf{Pr}\left[Z_1 \geq \frac{T^{1-\alpha/16}}{2} | \mathcal{E}\right] \leq 2cT^{-13\alpha/16} \leq T^{-\alpha/4}$$

for sufficiently large T.

After the adversary steps are used, the number of steps that s_i spends in the interval $I = [0, T^{\alpha/16} - 1]$ is stochastically dominated by that of an unbiased random walk U on $[0, T^{\alpha/4}]$ that runs for T steps and begins at 0. Let Z_2 be the number of steps U spends in I. As in the proof of Lemma 9, the equilibrium distribution of U is uniform over $[0, T^{\alpha/4} - 1]$. Thus $\pi(I) = T^{-3\alpha/16}$. Using Lemmas 7 and 8 we obtain

$$\begin{aligned}\mathbf{Pr}\left[Z_2 \geq \frac{T^{1-\alpha/16}}{2}\right] &\leq \frac{T^{\alpha/8} \cdot T^{\alpha/8}}{2} \exp\left(\frac{-T^{-3\alpha/8} \cdot T^{-\alpha/2} \cdot T}{j}\right) \\ &\leq T^{-\alpha/4}\end{aligned}$$

for sufficiently large T.

Taking a union bound, we find that the probability that $Z_1 + Z_2 \geq T^{1-\alpha/16}$ is at least $1 - 3T^{-\alpha/4}$, which proves the lemma. ■

We are now ready to prove the main theorem:

Theorem 14 *Random Fit is stable under the distribution $U\{k-2, k\}$ for all $k \geq 3$.*

Proof: As in our previous calculations we first assume that k is odd. As in Theorem 1 of [11], it suffices to consider the drift of $f(s)$ over a suitably large interval T, and show that it is negative for all but a finite number of states. The excluded set of states G will be

$$G = \{s \in S : \forall i, s_i \leq T^4\},$$

where T will be determined. Consider any starting state outside of this set G. Applying Lemma 9 and then Lemma 13 inductively, we find that with probability at least $1 - (c_1/T^{\epsilon_1})$, $s_{\lceil \frac{j}{2} \rceil} > 0$ over all but T^{ϵ_2} of the steps, for some constants c_1 and $\epsilon_1, \epsilon_2 < 1$ dependent only on j. Let $\mathcal{A}$ be the event that $s_{\lceil \frac{j}{2} \rceil} > 0$ over all but T^{ϵ_2} of the steps. As the expected value of f decreases by $1/j$ whenever $s_{\lceil \frac{j}{2} \rceil} > 0$ by Lemma 5, and it increases by at most j otherwise,

$$\begin{aligned} \mathsf{E}[f(T) - f(0) | f(0)] &\leq \mathsf{E}[f(T) - f(0) | f(0) \wedge \mathcal{A}] + (1 - \mathbf{Pr}[\mathcal{A}])\, \mathsf{E}[f(T) - f(0) | f(0) \wedge \neg \mathcal{A}] \quad (5) \\ &\leq \left[-\frac{1}{j} \left(T - T^{\epsilon_2}\right) + jT^{\epsilon_2} \right] + c_1 T^{1-\epsilon_1} j. \quad (6) \end{aligned}$$

By choosing T sufficiently large, we may make this expression smaller than $-\delta$ for some constant δ. This suffices to prove the theorem, by Lemma 6.

If k is even, then there is middle token $s_{\lceil j/2 \rceil + 1}$. If $s_{\lceil j/2 \rceil + 1} = 0$, everything is exactly as in the case where k is odd. If $s_{\lceil j/2 \rceil + 1} > 0$, then by Lemma 3 $s_{\lceil j/2 \rceil + 1} = 1$ and no bins with larger capacity are open. We consider the time steps when $s_{\lceil j/2 \rceil + 1} = 1$. In these steps f might increase because a small item may be inserted in the bin of capacity $\lceil j/2 \rceil + 1$. Lemmas 9 and 13, which apply when k is even, give that with probability at least $1 - (c_1/T^{\epsilon_1})$, $s_{\lceil \frac{j}{2} \rceil} > T^{1-\epsilon_2}$ over all but T^{ϵ_2} of the steps, for some constants c_1 and $\epsilon_1, \epsilon_2 < 1$ dependent only on j. Hence it should be a very rare event for a small item to be placed into a bin of capacity $\lceil j/2 \rceil + 1$.

In fact, in exactly the same manner as shown in Lemma 5, one may show the following:

Proposition 15 *Suppose that k is even and $s_{\lceil j/2 \rceil} > Z$. Then $\mathsf{E}[f(t+1) - f(t) | f(t)] \leq -1/j + 2/Z$.*

We conclude that in this case

$$\begin{aligned} \mathsf{E}[f(T) - f(0) | f(0)] &\leq \mathsf{E}[f(T) - f(0) | f(0) \wedge \mathcal{A}] + (1 - \mathbf{Pr}[\mathcal{A}])\, \mathsf{E}[f(T) - f(0) | f(0) \wedge \neg \mathcal{A}] \quad (7) \\ &\leq \left[\left(-\frac{1}{j} + \frac{2}{T^{1-\epsilon_2}} \right) \left(T - T^{\epsilon_2}\right) + jT^{\epsilon_2} \right] + c_1 T^{1-\epsilon_1} j. \quad (8) \end{aligned}$$

This expression can also be bounded by $-\delta$ if T is chosen large enough. ■

One may check that from the inductive use of Lemma 13, the ϵ_2 in Theorem 14 is exponential in j, and hence our bounds on the expected waste is doubly exponential in j. It is an interesting question whether better bounds are possible.

It is also worthwhile to note the following:

Theorem 16 *Random Fit(d) for $d \geq 2$ is stable under the distribution $U\{k-2, k\}$ for all $k \geq 3$.*

The proof is entirely similar to that for Random Fit. Simulations suggest that as d increases, the behavior of Random Fit(d) rapidly approaches that of Best Fit, as one might expect.

Theorem 17 *Random Fit and Random Fit(d), for $d \geq 2$, have expected waste $o(n)$ under the distributions $U\{k-1, k\}$ and $U\{k, k\}$, for all $k \geq 3$.*

Proof: We only consider the distribution $U\{k-1, k\}$, as the waste under the distribution $U\{k, k\}$ is entirely similar. Under this distribution, the statement corresponding to Lemma 5 is that if $s_{\lceil \frac{j}{2} \rceil}(t) > 0$, then $\mathsf{E}[f(t+1) - f(t) | f(t)] = 0$. Using the same notation as in the proof of Theorem 14 we obtain

$$\mathsf{E}[f(T) - f(0) | f(0)] \leq jT^{\epsilon_2} + c_1 T^{1-\epsilon_1} j$$

for some constants c_1 and $\epsilon_1, \epsilon_2 < 1$ dependent only on j. Hence, once the expected waste reaches a certain constant, its expected growth is sublinear, proving the theorem. ■

Whether tighter bounds, more like those known for Best Fit and First Fit, are possible for Random Fit under these distributions remains an open question.

4 Analysis of First Fit under distribution $U\{k-2, k\}$

We now consider how to modify the proof of RF on the distribution $U\{k-2, k\}$ to work for First Fit. Again we focus on the case where k is odd; the case where k is even requires some minor additional work, as for Random Fit, which we omit here.

One way of thinking about the difficulty in extending the results from RF to FF is to consider the dependence among the steps. In RF, at each step we have an independent random ordering assigned to the bins, while in FF, the orders of the bins at different steps are clearly dependent. In particular, the order of the bins at each step depends on the initial state, over which we have negligible control. The work of this section will focus on finding ways to circumvent effect of these dependencies so that we can apply the same ideas that we used in Section 3.

Let us consider an initial state, given at time $t = 0$. In order to avoid problems caused by the order of bins in the initial state, we focus on bins that are created after time 0. In fact, we are even more restrictive: let a *single i bin at time t* be a bin created after time 0 that has remaining capacity i and contains only one item, and denote the number of single i bins by $u_i(t)$. Instead of the vector s we considered previously, we shall primarily the vector $u = (u_1, \ldots, u_{\lceil \frac{j}{2} \rceil})$. The following important points about u make it useful:

- If $u_{\lceil \frac{j}{2} \rceil} > 0$, then $s_{\lceil \frac{j}{2} \rceil} > 0$ also. Hence, proving $u_{\lceil \frac{j}{2} \rceil} > 0$ over most of the steps is sufficient.
- Regardless of the initial state, $(u_1, \ldots, u_{\lceil \frac{j}{2} \rceil}) = (0, \ldots, 0)$ at time 0.

To see how the considering u makes things easier, let us prove a lemma similar to Lemma 9 for First Fit.

Lemma 18 *Suppose $s_i(0) \geq T$. Then when $u_{i+1} > 0$, u_{i+1} behaves like a random walk with probability at least $1/j$ of increasing at each step and probability at most $1/j$ of decreasing at each step. Also, the time spent by u_{i+1} on each visit to 0 is stochastically dominated by a random variable D with constant expectation (that depends only on j). In particular, $u_{i+1} \geq T^{1/16}$ for all but at most $T^{15/16}$ steps with probability at least $1 - \frac{1}{T^2}$.*

Proof: Since $s_i(0) \geq T$, over the next T steps there is always a bin with remaining capacity i ahead of all single bins with remaining capacity $i+1$ created after time 0. This implies that u_{i+1} can decrease only when an item of size $i+1$ arrives, and hence decreases with probability at most $1/j$ at each step. When $u_{i+1} > 0$, then u_{i+1} increases whenever an item of size $k-i-1$ arrives, and hence it increases with probability at least $1/j$. The case where $u_{i+1} = 0$ is special, and is handled as in Lemma 4. The final result, that $u_{i+1} \geq T^{1/16}$ most of the time, now follows using an argument similar to Lemma 9. ■

As in the proof for RF, we now want to extend the above lemma inductively. Similar to the RF case, we would like to say that a bin of size i lies ahead of all single $i+1$ bins most of the time, whenever the number of single $i+1$ bins is sufficiently small. In Lemma 13, we accomplished this by splitting each step into two substeps, with the first substep re-ordering the bins randomly. We do not have this luxury for the FF case. However, it seems intuitive that the bins should be "almost" randomly distributed at each step. This point is made explicit in the following lemma:

Lemma 19 *Let $\mathcal{E}$ be the event that a single i bin at time t lies ahead of all single $i+1$ bins. Let $z_t^{b,c} = \mathbf{Pr}\{\mathcal{E} | u_i(t) = b, u_{i+1}(t) = c\}$. Then $z_t^{b,c} \leq \frac{b}{b+c}$.*

Proof: Consider any sequence $a = a_1, a_2, \ldots, a_t$ of t items that ends with a single $i+1$ bin ahead of all single i bins with $u_i(t) = b$ and $u_{i+1}(t) = c$. We center on the steps where the single i and $i+1$ bins were created. We first claim that if a single i bin was created at step g and a single $i+1$ bin was created at step h, then switching the entering items at steps g and h switches the order of these two bins, but has no other effect on the algorithm. This can easily be proven by induction for all bins behind the first single $i+1$ bin, since there is no way a second item could have been placed in any of these bins. The only difficult case is that of the first single $i+1$ bin, call it B. The reason that B is a special case is that it is possible that since B is the frontmost single bin, it may be that a second item could have been placed in it if we change its capacity. However, since switching the appropriate steps g and h would only lower the capacity of B, it is clear that if B has not obtained a second item in the original sequence, it cannot in the modified sequence as well.

We now divide the sequences into equivalence classes. For a sequence a, let $Y_t^i(a)$ be the set of times at which the single i bins at time t were created. Two sequences a and a' are equivalent if $Y_t^i(a) \cup Y_t^{i+1}(a) = Y_t^i(a') \cup Y_t^{i+1}(a')$ and $u_i(t) = b$, $u_{i+1}(t) = c$ for both sequences.

Take any sequence a with a single $i+1$ bin ahead of all single i bins at time t. From the above paragraph, permuting the times when a single $i+1$ bin and a single i bin were created yields equivalent sequences. Hence, by taking all ways of splitting $Y_t^i(a) \cup Y_t^{i+1}(a)$ into two groups of size b and c, and using this division to determine when single i and $i+1$ bins are created, we find that every sequence a has at least $\binom{b+c}{b}$ sequences in its equivalence class. Since the probability a and any of these other $\binom{b+c}{b}$ sequences occurring are equal, it is straightforward to show combinatorially that there are at least b/c times as many sequences with a single i bin ahead of all single $i+1$ bins as there are with a single $i+1$ bin ahead of all single i bin. Hence $z_t^{b,c} \leq \frac{b}{b+c}$. ■

Lemma 19 suggests that the behavior of FF should not be worse than RF, with the understanding that the u_i now play the role of the s_i. As in the case of RF, we would like to say the small tokens u_i behave like a unbiased random walk over most of the steps. This leads us to the prove a variant of Lemma 13 in this setting, which is phrased slightly differently in order to appropriately handle the conditioning.

Lemma 20 *Suppose, over a period of T steps, $u_{i-1} \geq T^\alpha$ over all but at most $T^{1-\alpha}$ steps for some $\alpha \leq 1/16$ with probability at least $1/2$. Then, conditioned on $u_{i-1} \geq T^\alpha$ over all but at most $T^{1-\alpha}$ steps, $u_i \geq T^{\alpha/16}$ for all but at most $T^{1-\alpha/16}$ steps with probability at least $1 - 4T^{-\alpha/4}$.*

Proof: As in Lemma 13, we must bound the number of steps for which the behavior of u_i is not that of an unbiased random walk, and then apply an adversary argument. Also as in Lemma 13, we will restrict our consideration to the behavior of u_i to the interval $[0, T^{\alpha/4} - 1]$. (This can be interpreted as though if $u_i \geq T^{\alpha/4}$, we may assume that a single bin of size $i+1$ lies ahead of all bins of size i, which is a conservative assumption.)

To bound the number of steps the adversary controls, then, we bound the number of steps X that satisfy the following conditions:

- $u_{i-1} \geq T^{\alpha}$.
- $u_i \leq T^{\alpha/4} - 1$.
- A single i bin lies ahead of all single $i-1$ bins.

The value of X, in addition to the number of steps for which $u_{i-1} < T^{\alpha}$, bounds the number of steps where the adversary controls the walk; on all other steps, we either have that $u_i \geq T^{\alpha/4}$ or a single $i-1$ bin lies in front of all single i bins, and so u_i behaves (at worst) as an unbiased random walk with $p_{\uparrow} = p_{\downarrow} = 1/j$. (As usual, we ignore the discrepancy at $u_i = 0$.)

Let y_t be the probability that on the tth step the above conditions hold. Then

$$\begin{aligned} \mathsf{E}[X] = \mathsf{E}[\sum_{t=0}^{T-1} y_t] &= \sum_{t=0}^{T-1} \mathsf{E}[y_t] \\ &\leq \sum_{t=0}^{T-1} \frac{T^{\alpha/4}}{T^{\alpha} + T^{\alpha/4}} \\ &< T^{1-3\alpha/4}. \end{aligned}$$

Although it would seem this is enough to bound the number of adversary steps, we must be careful. Let $\mathcal{E}$ be the event that $u_{i-1} \geq T^{\alpha}$ over all but $T^{1-\alpha}$ steps. The expected number of additional adversary steps from single $i-1$ bins being frontmost is not $\mathsf{E}[X]$, but $\mathsf{E}[X|\mathcal{E}]$. From the hypothesis of the lemma that $\mathbf{Pr}(\mathcal{E}) \geq 1/2$, however, we must have $\mathsf{E}[X|\mathcal{E}] \leq 2T^{1-3\alpha/4}$. Using Markov's law, we have

$$\mathbf{Pr}(\{X|\mathcal{E}\} \geq T^{1-\alpha}) \leq 2T^{-\alpha/4}.$$

Hence, conditioned on $\mathcal{E}$, the number of steps the adversary controls is at most $2T^{1-\alpha}$ with a probability at least $1 - 2T^{-\alpha/4}$. The rest of the proof now proceeds as in Lemma 13. ■

We are now ready to prove the main theorem:

Theorem 21 *First Fit is stable under the distribution $U\{k-2, k\}$ for all $k \geq 3$.*

Proof: As in Theorem 14, it suffices to consider the drift of $f(s)$ over a suitably large interval T, and show that it is negative for all but a finite number of states. The excluded set of states G will be

$$G = \{s \in S : \forall i, s_i \leq T\},$$

for some suitably large T. We now apply Lemma 18 and Lemma 20 to obtain a bound on $\mathsf{E}[f(T) - f(0)|f(0)]$ similar to that in Theorem 14.

We would then like to apply Lemma 6; however, technically we cannot do so, as the state space of the underlying Markov chain is not embedded in a fixed dimensional space. Similar results, however, can be applied in this setting, once we have shown that the expected change in the waste f is negative for a suitably large T. For example, [14, Theorem 13.0.1] can be used to show that the chain is ergodic, and [6, Theorem 3.1] implies that in the stationary distribution, the distribution of the waste has an exponentially decreasing tail. ■

5 Conclusions

We have demonstrated that the First Fit bin packing algorithm is stable on the distribution $U\{k-2, k\}$. We believe that our result demonstrates that the Markov chain approach may be useful, even in situations where the natural description of a problem does not have a convenient state space. Our analysis made use of insight gained from a novel packing algorithm, Random Fit, which appears interesting in its own right.

An open question is to tighten the bounds developed in this paper. For both First Fit and Random Fit, our bounds for the expected waste are doubly exponential in j. Simulations suggest that the expected waste for First Fit may only be exponential in j [9]. Unfortunately, the simulations for Random Fit seem to suggest that the expected waste for Random Fit may indeed be doubly exponential in j, in which case it seems that another approach may be necessary to achieve better bounds for First Fit.

References

[1] Y. Azar, A. Broder, A. Karlin and E. Upfal. Balanced allocations. In *Proc. 26 Annual ACM Symposium on Theory of Computing*, pages 593–602, 1994.

[2] E.G. Coffman, 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: A bin packing case study. In *Proc. 23th Annual ACM Symposium on Theory of Computing*, pages 230–240, 1991.

[3] E.G. Coffman Jr., M.R. Garey and D.S. Johnson. Approximation algorithms for bin packing: A survey. In *Approximation Algorithms for NP-hard Problems*, edited by D. Hochbaum, PWS, 1996.

[4] 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 *Proc. 25th Annual ACM Symposium on Theory of Computing*, pages 412–421, 1993.

[5] G. Fayolle, V.A. Malyeshev and M.V. Menshikov. *Topics in Constructive Theory of Countable Markov chains*. Cambridge University Press, 1995.

[6] B. Hajek. Hitting-time and occupation-time bounds implied by drift analysis with applications. *Advances in Applied Probability*, 14:502–525, 1982.

[7] D.S. Johnson. *Near-Optimal Bin Packing Algorithms*. PhD Thesis, Massachusetts Institute of Technology, Department of Mathematics, 1973.

[8] D.S. Johnson. Fast algorithms for bin packing. *Journal of Computer and System Sciences*, 8:272–314, 1974.

[9] D.S. Johnson. Private communication, 1996.

[10] D.S. Johnson, A. Demers, J.D. Ullman, M.R. Garey and R.L. Graham. Worst case bounds for simple one-dimensional packing algorithms. *SIAM Journal on Computing*, 3:299–325, 1974.

[11] C. Kenyon, A. Sinclair, and Y. Rabani. Biased Random Walks, Lyapunov Functions, and Stochastic Analysis of Best Fit Bin Packing. In *Proc. 7th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 351–358, 1995. Submitted to *Journal of Algorithms*.

[12] N. Kahale. Large Deviation Bounds for Markov Chains. DIMACS Technical Report 94-39.

[13] T. Leighton and P. Shor. Tight bounds for minimax grid matching with applications to average case analysis of algorithms. *Combinatorica* 9:161–187, 1989.

[14] S. Meyn and R. Tweedie. *Markov Chains and Stochastic Stability*, Springer-Verlag, 1996.

[15] M. Mitzenmacher. Load balancing and density dependent jump Markov processes. In *Proc. 37th IEEE Annual Symposium on Foundations of Computer Science*, 1996.

[16] P.W. Shor. The average-case analysis of some on-line algorithms for bin packing. *Combinatorica*, 6:179–200, 1986.