

Bin Packing with Discrete Item Sizes, Part II: Tight Bounds on First Fit

E. G. Coffman, Jr.,

Bell Labs, Lucent Technologies
Murray Hill, New Jersey 07974

D. S. Johnson, P. W. Shor

AT&T Labs
Murray Hill, New Jersey 07974

R. R. Weber

Cambridge University
Cambridge, England

ABSTRACT

In the bin packing problem, a list L of n items is to be packed into a sequence of unit capacity bins with the goal of minimizing the number of bins used. First Fit (FF) is one of the most natural on-line algorithms for this problem, based on the simple rule that each successive item is packed into the first bin of the sequence that currently has room for it. We present an average-case analysis of FF in the discrete uniform model: the item sizes are drawn independently and uniformly at random from the set $\{1/k, \dots, (k-1)/k\}$, for some $k > 1$. Let $W^{FF}(L)$ denote the wasted space in the FF packing of L , i.e., the total space still available in the occupied bins. We prove that $E[W^{FF}(L)] = O(\sqrt{nk})$, i.e., there exists a constant $c > 0$ such that $E[W^{FF}(L)] \leq c\sqrt{nk}$ for all n, k sufficiently large. By a complementary lower bound argument, we prove that this result is sharp, unless k is allowed to grow with n at a rate faster than $n^{1/3}$, in which case $E[W^{FF}(L)] = \Theta(n^{2/3})$. Finally, we show that this last result carries over to the continuous uniform model, where item sizes are chosen uniformly from $[0, 1]$. The $O(n^{2/3})$ upper bound for the continuous model is new and solves a problem posed a decade ago. The proofs of many of these results require extensions to the theory of stochastic planar matching.

August 30, 1996

Bin Packing with Discrete Item Sizes, Part II: Tight Bounds on First Fit

E. G. Coffman, Jr.,

Bell Labs, Lucent Technologies
Murray Hill, New Jersey 07974

D. S. Johnson, P. W. Shor

AT&T Labs
Murray Hill, New Jersey 07974

R. R. Weber

Cambridge University
Cambridge, England

1. Introduction

We study the First Fit (FF) packing of a list L of n items into a sequence of initially empty, equal capacity bins. The item sizes are all at most the bin capacity. According to FF, each successive item is packed in the first bin of the sequence that has room for it. FF is an on-line algorithm, in that items are assigned to bins in the order in which they are input, with each assignment depending only on the packing constructed so far and without reference to the sizes or number of remaining unpacked items. The FF packing also has a useful “off-line” characterization, however: it can be constructed by packing the bins one at a time, for each bin repeatedly adding the first as-yet-unpacked item that will fit, until no such items remain. (An easy induction establishes the equivalence of these two formulations.)

For a given distribution of item sizes, we are interested in the following question: As a function of n , what is the expected wasted space (total unused capacity of the occupied bins) in the final packing? In the classical continuous model, the bin size is taken to be 1 for convenience, and the item sizes are independent samples from the uniform distribution on $[0, 1]$. In a discrete version of this model, the item sizes are independent samples from the uniform distribution on $\{1/k, 2/k, \dots, (k-1)/k\}$, for some $k \geq 2$ (trivialities are avoided by disallowing items of size 0 or 1). In this case, results are normally expressed in terms of both n and k . Here, it is more convenient to take the bin size to be k and the item sizes to be uniform on $\{1, 2, \dots, k-1\}$. We measure wasted space in units of the bin size, so the results for the two discrete versions will be the same.

Let $W^{FF}(L)$ denote the wasted space in a First Fit packing of L . For the continuous model, Shor [10] proved that $E[W^{FF}(L)] = \Omega(n^{2/3})$ and $E[W^{FF}(L)] = O(n^{2/3} \log^{1/2} n)$. This paper shows that for any fixed value of k in the discrete model, $E[W^{FF}(L)] = O(\sqrt{n})$, a significantly slower growth rate. This is a corollary of the more general result that $E[W^{FF}(L)] = O(\sqrt{nk})$, i.e., there exists a constant $\alpha > 0$ such that for all n and k sufficiently large, $E[W^{FF}(L)] \leq \alpha\sqrt{nk}$. Using this latter result, we prove that Shor's lower bound for the continuous case is in fact tight, i.e. that for this case $E[W^{FF}(L)] = O(n^{2/3})$ and hence $E[W^{FF}(L)] = \Theta(n^{2/3})$. We also prove a lower bound for the discrete case which shows that the $O(\sqrt{nk})$ upper bound is tight if $k = O(n^{1/3})$. Finally, we prove that if $k = \Omega(n^{1/3})$ then $E[W^{FF}(L)] = \Theta(n^{2/3})$, as in the continuous model.

This paper is the second in a series of papers currently being written based on the results announced in [1] and [6]. The theme of the series is the extension of bin-packing theory to problems in which item sizes are drawn from discrete distributions. The paper in the series most closely related to the current one is [3], which will analyze the behavior of the Best Fit algorithm (BF) under the same distributions we study here. BF is online like FF but places successive items into bins where they fit best, i.e., minimize the resulting leftover space. (A tie is resolved in favor of the lowest indexed bin.) As in the case of continuous uniform distributions, BF slightly outperforms FF, the main result being that $W^{BF}(L)$ is $\Theta(\sqrt{n} \log^{3/4} k)$ when $k = O(n)$ and is $\Theta(\sqrt{n} \log^{3/4} n)$ (the bound for the continuous uniform case) when $k = \Omega(n)$. It does not appear that any on-line algorithm can do significantly better than this. It is easy to see that the expected waste must be at least $\Theta(\sqrt{n})$, and, based on analogies with results for the continuous case in [10] and [11], we expect that the best possible on-line waste growth rate is $\Theta(\sqrt{n \log k})$ when $k = O(\sqrt{n})$ and $\Theta(\sqrt{n \log n})$ (the bound for the continuous uniform case) when $k = \Omega(n)$.

The other papers in the series consider a wider variety of discrete distributions, especially the distributions $U\{j, k\}$, $1 \leq j < k - 1$, where in $U\{j, k\}$ the item sizes are $1/k, 2/k, \dots, j/k$, all equally likely, and the bin size is 1. (In this terminology, the distributions considered in the current paper are the $U\{k - 1, k\}$.) For each such distribution with $j < k - 1$, paper [2] shows that there exists an on-line algorithm whose expected waste is bounded by a constant, independent of n . For many of these distributions, First and Best Fit also appear to have bounded waste (based on simulations). For Best Fit this is proved for several such distributions in [4]. (Further results of this kind can be found in [8].) On the other hand, [4] also proves that

Best Fit's expected waste is $\Theta(n)$ for $U\{8, 11\}$ and $U\{9, 12\}$. Even such off-line algorithms as First and Best Fit Decreasing (FFD and BFD) can have $\Theta(n)$ expected waste for some such distributions, and the behavior of these algorithms is investigated in [5].

The current paper is organized as follows. In Section 2 we introduce a number of preliminary results needed in later sections. Sections 3 and 4 then prove the FF upper and lower bounds, respectively. Section 5 concludes the paper with remarks on extensions and open problems.

2. Preliminaries

Instead of analyzing FF packings of random lists of n items, for fixed n , it is more convenient to analyze FF packings of random lists of N items, where N is Poisson distributed with mean n and independent of item sizes. The two models are called the *fixed- n model* and *Poisson model*, respectively. It is a trivial consequence of the following general lemma that the estimates obtained for the Poisson model also apply to the fixed- n model. Let $A(L_n)$ and $A(L_N)$ be the respective numbers of bins required by an on-line algorithm A in packing lists L_n and L_N in the fixed- n and Poisson models. Similarly, define $W^A(L_n)$ and $W^A(L_N)$ for the wasted space under A in the respective models.

Lemma 2.1. *Assume a general distribution F of item sizes on $[0, 1]$. Then*

$$|E[A(L_n)] - E[A(L_N)]| = O(\sqrt{n}) ,$$

where the hidden multiplicative constant is independent of F .

Remark. Since the expected occupied space in a packing of L_n is the same as that in a packing of L_N , viz. $n/2$ in units of the bin size, Lemma 2.1 also shows that, for the expected wasted space,

$$|E[W^A(L_n)] - E[W^A(L_N)]| = O(\sqrt{n}) .$$

Proof. Consider the A packing of a list of n items, with sizes drawn independently from the distribution F . For j a sample of a random variable J distributed as $N - n$, modify the packing as follows. If $j > 0$, then extend the given A packing by packing j more items with sizes drawn independently from F . If $j < 0$, then remove the last $|j|$ items of the given A packing; this will empty just those bins started by the last $|j|$ items packed. Because A is on-line, the above operation produces an A packing of a random list of $n + J$ items, where $n + J$ is equal in

distribution to N . Moreover, it is easy to see that the numbers of occupied bins added when $j > 0$, and subtracted (emptied) when $j < 0$ are at most $|j|$. The lemma follows from the bound $E|J| = E|N - n| = O(\sqrt{n})$ given by standard estimates of the Poisson distribution. ■

Hereafter, unless stated otherwise, the Poisson model of FF packing is to be assumed. To be consistent with standard formulations of the bin-packing problem, we have chosen to use the fixed- n model in the statements of the main theorems. For simplicity, the above subscript convention for L will be dropped in what follows.

As in [11] we express instances L in terms of the sample paths of a Poisson process in two dimensions. Figure 1 gives an example. We describe the method first with k odd. The

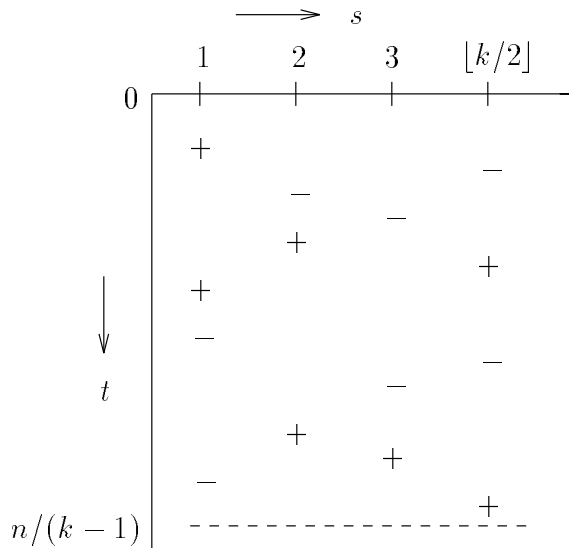


Figure 1: L as the superposition of Poisson processes.

discrete horizontal dimension consists of $\frac{k-1}{2}$ columns indexed by s . The continuous vertical dimension is a time axis with the time t starting at 0 at the top of the figure and increasing as one goes downward (to be consistent with the literature). In each column, sample paths of two independent Poisson point processes are laid out, each at rate 1, one generating points labeled with a $+$ and the other generating points labeled with a $-$. The list L is constructed top-down from the points that appear in the interval $[0, n/(k-1)]$; a $-$ in column s becomes an item of size $s < k/2$ in L , and a $+$ in column s becomes an item of size $k - s > k/2$ in L . The superposition of all $k - 1$ processes, two per column, gives a Poisson process at rate $(k - 1)$. Thus, the number N of points in L is Poisson distributed with mean $(k - 1) \cdot \frac{n}{(k-1)} = n$, as

desired. Each successive point of the superposition process is equally likely to be a + or -, and it is equally likely to be in any one of the $\frac{k-1}{2}$ columns. Thus, successive items in L have sizes independent and uniformly distributed on $\{1, \dots, k-1\}$, again as desired.

If k is even, the independent Poisson processes of +'s and -'s are laid out in $k/2$ columns; the processes in the first $\frac{k}{2} - 1$ columns are at rate 1 as before, but in column $k/2$ the + and - processes each have rate $1/2$. The list is constructed as before, but note that both +'s and -'s in column $k/2$ become items of size $k/2 = k - k/2$. It is easy to verify that the construction again gives lists of N independent item sizes uniformly distributed over $\{1, \dots, k-1\}$ with N Poisson distributed and with $E[N] = n$.

Our bounds on expected wasted space will invariably be expressed in terms of algorithms matching +'s and -'s in the above two-dimensional instances. These matching algorithms all satisfy the constraint that the sizes of the items represented by a matched + and - must sum to at most k . Thus, since matching a + in column s to a - in column s' corresponds to matching items with respective sizes $k-s$ and s' , we must have $k-s+s' \leq k$ and hence $s \geq s'$. Graphically then, our matching constraint means that, for each matched pair, the + must be to the right of or in the same column as the -.

For matching algorithm A , $M^A(L)$ denotes the set of pairs of matched points in L , or equivalently, the set of edges (straight-line segments) connecting the paired +'s and -'s. Let $U^A(L)$ count the number of points left unmatched, i.e., $U^A(L) = N - 2|M^A(L)|$. The connection between matching and packing lies in the fact that a matching algorithm A corresponds in an obvious way to a packing algorithm A ; in the A packing of L , each pair in $M^A(L)$ is packed into a single bin, as is each unmatched item counted by $U^A(L)$. The two interpretations of A yield the following simple result.

Lemma 2.2. *For the matching/packing algorithm A and any symmetric distribution of item sizes on $\{1, \dots, k-1\}$ (i.e., any distribution for which the probability of size i equals that of size $k-i$, $1 \leq i < k/2$), the expected wasted space under A for a random list L satisfies*

$$E[W^A(L)] = \frac{1}{2}E[U^A(L)] \tag{2.1}$$

$$E[W^A(L)] \geq \frac{1}{k} \sum_{1 \leq i \leq |M^A(L)|} E[H_i^A], \tag{2.2}$$

where H_i^A is the horizontal component of the i^{th} edge in $M^A(L)$.

Proof. The number of occupied bins in the A packing of L is $A(L) = (N - U^A(L))/2 + U^A(L)$, so that $E[A(L)] = \frac{1}{2}E[N] + \frac{1}{2}E[U^A(L)]$. We have N items of average size $k/2$, so in units of the bin size, $\frac{1}{2}E[N]$ is the expected occupied space. Then (2.1) gives the expected wasted space.

The horizontal component of an edge in $M^A(L)$ is the wasted space in the bin containing the pair of items connected by the edge; dividing by k gives the wasted space in units of the bin size. Summing the expected value over all edges yields the lower bound in (2.2). ■

The analysis of matching algorithms will often reduce to the analysis of a process

$$W_0 = 0 , \tag{2.3}$$

$$W_j = (W_{j-1} + \xi_j)^+ , \quad j \geq 1 ,$$

where $x^+ = \max(0, x)$, and where the ξ_j are the successive steps of a random walk in \mathbb{R}_+ . Figure 2 shows an example. The sequence $\{W_j\}$ is a Lindley process and can be viewed as the queueing process induced by the random walk $\{S_j\}$, with $S_j = \sum_{1 \leq i \leq j} \xi_i$ (Feller [7], Sect. VI.9). In the applications of this paper, the i.i.d. random variables ξ_i satisfy

$$E[\xi_i] = 0, \quad \sigma^2 < \infty , \tag{2.4}$$

where σ^2 denotes the variance of the ξ_i 's.

Lemma 2.3. *Under the assumptions in (2.4), $E[W_j] = O(\sqrt{j})$, where the hidden multiplicative constant depends only on σ .*

Proof. Solving (2.3), one can show by induction that $W_j \stackrel{d}{=} \max\{0, S_1, \dots, S_j\}$ (Feller [7], p. 197), where $\stackrel{d}{=}$ denotes equality in distribution and is justified here by the fact that the ξ_i are independent and identically distributed. Then we can write

$$\begin{aligned} E[W_j] &= \int_0^\infty Pr\{\max\{0, S_1, \dots, S_j\} > x\} dx \\ &\leq \sqrt{j} + \int_{\sqrt{j}}^\infty Pr\{\max\{0, S_1, \dots, S_j\} > x\} dx . \end{aligned} \tag{2.5}$$

Since the ξ_i have zero means, we have by Kolmogorov's inequality (Feller [7], p. 235, eq. (8.3)) $Pr\{\max\{0, S_1, \dots, S_j\} > x\} \leq j\sigma^2/x^2$ whereupon substitution into (2.5) proves the lemma. ■

Let Q_j denote the number of times that $W_i = 0$ in the sequence W_1, \dots, W_j . The FF lower-bound argument will need the following lower bound on $E[Q_j]$.

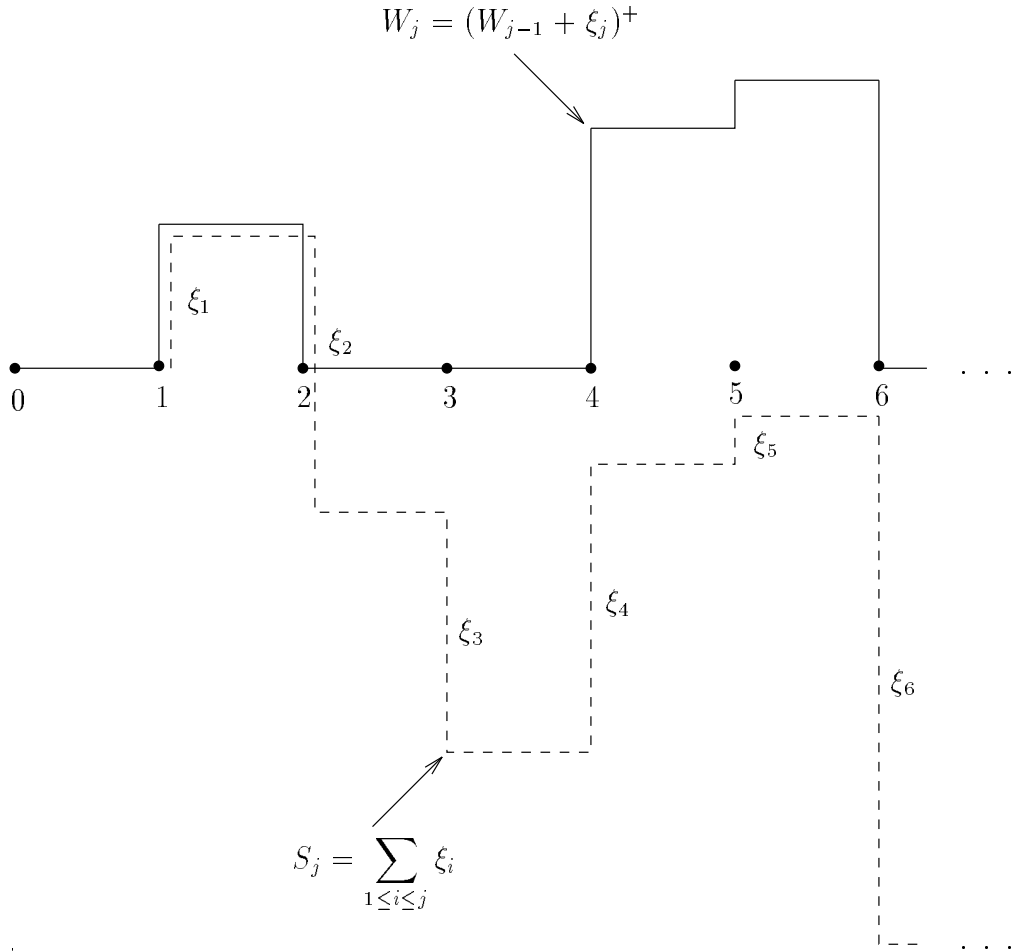


Figure 2: $\{W_j\}$ and $\{S_j\}$ illustrated as step functions.

Lemma 2.4. *Under the assumptions (2.4), $E[Q_j] = \Omega(\sqrt{j})$.*

Proof. The *descending ladder epochs* of the random walk $\{S_i\}$ are those index values $i \geq 0$ where the position of the walk is lower than at any preceding epoch (see Fig. 2). It is easily verified that Q_j is equal in distribution to the number of descending ladder epochs encountered by S_1, \dots, S_j (Feller [7], p. 196). Moreover, under (2.4) this quantity has the same distribution as in the classical symmetric random walk with step sizes ± 1 (Feller [7], p. 396, Corollary 1). In the latter random walk the probability that, for any $\alpha > 0$, the position after j steps is less than $-\alpha\sqrt{j}$ is bounded away from 0 for all j sufficiently large. The lemma follows at once from the trivial observation that the classical random walk must encounter at least r descending ladder epochs in reaching a final position of $-r$. ■

Although estimates more precise than Lemmas 2.3 and 2.4 are possible, they will not be needed. A further advantage of their present form is that the following generalization is trivial to prove. The details are left to the interested reader.

Corollary 2.1. *Let the process $\{W_i\}$ have an initial state $W_0 = O(\sqrt{j})$, and consider J steps of the process, where J is Poisson distributed with mean j . Then Lemmas 2.3 and 2.4 still hold, i.e., $E[W_J] = O(\sqrt{j})$ and $E[Q_J] = \Omega(\sqrt{j})$.*

3. The FF Upper Bound

We begin with an analysis of an algorithm that majorizes FF. This approach mimics that in [10], where the problem with continuous item sizes is analyzed. We define the *matching First Fit* (MFF) algorithm as follows, in terms of the item labels introduced in Fig. 1. As in that figure, let s and t denote respectively a horizontal size coordinate (column index) and a vertical time coordinate. Recall that the size dimension is “folded”; a point with size coordinate s represents an item of size s if the point is a $-$, but an item of size $k - s$ if the point is a $+$. The first item (the one with smallest t coordinate) is packed in the first bin. The bin is then closed if the item was a $-$; otherwise, it remains open for another item. Thereafter, if the next item to be packed is a $+$, it starts a new bin, which remains open. If the next item is a $-$, it is packed in the first open bin, if any, in which it fits; if no such open bin exists, then the $-$ starts a new bin. In either case, the bin receiving the $-$ is then closed. Note that if k is odd, then MFF is the same as FF, except that it closes a bin whenever it receives a $-$. When k is even, there is a further difference between MFF and FF: MFF rejects opportunities to pack two items of size $k/2$ into the same bin unless the first is a $+$ and the second is a $-$.

MFF has the following useful monotonicity property, which FF does not share.

Lemma 3.1. (Shor [10]) *Suppose L' is obtained from L by the removal of one or more items, leaving the ordering of items unchanged. Then $MFF(L') \leq MFF(L)$.*

Shor [10] proves this result for the continuous case, but his arguments carry over directly to our discrete model; the details are left to the interested reader. Shor also proves that $W^{MFF}(L) \geq W^{FF}(L)$ for all L in the continuous case. Unfortunately, this result holds in our discrete model only for k odd, where the special case of items of size $k/2$ does not arise. However, we need only the average-case majorization, as given in the next result, which holds for all k .

Lemma 3.2. *Let the items of L be independent with a general distribution on $\{1, \dots, k-1\}$ for any $k > 1$. Then $E[W^{FF}(L)] \leq E[W^{MFF}(L)]$.*

Proof. We have $W^{MFF}(L) \geq W^{FF}(L)$ for all L when k is odd by the arguments of [10] for the continuous case, which we omit. Thus, the lemma holds trivially for k odd. Assume that k is even for the remainder of the proof.

Consider the $FF(L)$ packing, i.e., the FF packing of list L . The nonempty bins are of 5 types, to be denoted $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$, $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$, $\left[\begin{smallmatrix} - \\ + \end{smallmatrix} \right]$, $\left[\begin{smallmatrix} - \\ - \end{smallmatrix} \right]$, $\left[\begin{smallmatrix} - \\ - \end{smallmatrix} \right]$. A type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ bin contains only a $+$. A type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ bin contains two $+$'s of size $k/2$ and a type- $\left[\begin{smallmatrix} - \\ - \end{smallmatrix} \right]$ bin contains two $-$'s of size $k/2$. Type- $\left[\begin{smallmatrix} - \\ + \end{smallmatrix} \right]$ bins include all those bins that start with a $+$ except the type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ and type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ bins. Similarly, type- $\left[\begin{smallmatrix} - \\ - \end{smallmatrix} \right]$ bins include all those bins that start with a $-$ except the type- $\left[\begin{smallmatrix} - \\ - \end{smallmatrix} \right]$ bins. Let ν^+ and ν^- be the numbers of type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ and type- $\left[\begin{smallmatrix} - \\ - \end{smallmatrix} \right]$ bins, respectively.

We claim that, if a type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ bin is removed from the $FF(L)$ packing, then an $FF(L')$ packing remains, where L' is obtained from L by deleting the items, say X and X' , that were in the type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ bin removed, and by retaining the order of the remaining items in L . Let X come before X' in L , so X' is the first item of size $k/2$ following X in L ; and let B_i be the bin containing X, X' in the $FF(L)$ packing. To verify the claim, note first that all items that came before X in L were packed in B_1, \dots, B_{i-1} , so they will be identically packed in B_1, \dots, B_{i-1} of the $FF(L')$ packing. Any $-$ following X in L fits with X , so all $-$'s between X and X' in L must have been packed in B_1, \dots, B_{i-1} ; then these items will also appear in the $FF(L')$ packing just as they did in the $FF(L)$ packing. All $+$'s between X and X' have sizes $> k/2$ and had to be packed in B_{i+1}, B_{i+2}, \dots , for otherwise, X could have been packed into a bin $B_j, j < i$. Thus, these $+$'s appear in B_i, B_{i+1}, \dots in the $FF(L')$ packing in the same sequence as before. When X' was packed in B_i , B_i became full. Thus, if an item following X' in L appeared in B_j in the $FF(L)$ packing, it will appear in B_j in the $FF(L')$ packing if $j < i$, and in B_{j-1} if $j > i$. The claim follows.

Now remove all type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ bins from the $FF(L)$ packing to obtain the $FF(L_1)$ packing, where L_1 is obtained from L by removing all items that appeared in type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ bins in the $FF(L)$ packing. Next, remove from L_1 all items except those that were either items packed first in any bin or are $-$'s that were packed second in bins of type- $\left[\begin{smallmatrix} - \\ + \end{smallmatrix} \right]$. We are left with type- $\left[\begin{smallmatrix} + \\ + \end{smallmatrix} \right]$ bins containing only a $+$, as before; type- $\left[\begin{smallmatrix} - \\ + \end{smallmatrix} \right]$ bins containing only a $+$ and $-$ with the $+$ packed first; and type- $\left[\begin{smallmatrix} - \\ - \end{smallmatrix} \right]$ bins containing only a $-$. Let L_2 be the list of remaining items, and call the above packing of L_2 the *reduced* packing. We claim that the reduced packing is

an $MFF(L_2)$ packing.

To verify the claim, note that each bin type in the reduced packing is a valid MFF bin type, by the definitions of L_2 and MFF. Suppose the first i bins of the reduced packing are an MFF packing of the items contained in these bins, and consider where MFF would pack the one or two items in the $(i + 1)^{\text{st}}$ bin of the reduced packing. By the definition of MFF, the only open bins among the first i are type- $\left[+\right]$ bins. But these were also type- $\left[+\right]$ bins in the $FF(L)$ and $FF(L_1)$ packings. Since the items in the $(i + 1)^{\text{st}}$ bin of the reduced packing did not fit in these type- $\left[+\right]$ bins of the $FF(L_1)$ packing, they would be packed by MFF into an $(i + 1)^{\text{st}}$ bin just as they appear in the reduced packing. A simple induction thus establishes the claim.

Now add back to L_2 the ν^- -'s of size $k/2$ that were removed from type- $\left[-\right]$ bins in the $FF(L_1)$ packing, preserving the order of items in L . Our final claim is that the MFF packing of the new list L_3 consists of the bins of the $MFF(L_2)$ packing plus ν^- new type- $\left[-\right]$ bins; the new type- $\left[-\right]$ bins will be interspersed among the bins of the $MFF(L_2)$ packing according to the positions of their items in L_3 , but the ordering of the bins of the $MFF(L_2)$ packing will be preserved. This claim is proved by the same reasoning as before. Suppose just one - of size $k/2$ is returned to L_2 . When that item comes to be packed by MFF, it can not fit in any open bin, because such a bin would have to be a type- $\left[+\right]$ bin which also existed in the $FF(L)$ and $FF(L_1)$ packings. Then MFF packs the new - of size $k/2$ in a new bin, which it then closes. An easy induction on ν^- completes the argument.

By the previous claim and the definition of L_1 , we have

$$MFF(L_3) = MFF(L_1) + \nu^- = FF(L_1) + \nu^- = FF(L) - \nu^+ + \nu^- .$$

By Lemma 3.1, $MFF(L_3) \leq MFF(L)$, so

$$E[FF(L)] \leq E[MFF(L)] + E[\nu^+] - E[\nu^-] .$$

But FF packings are determined solely by item sizes, not labels. Thus, since items of size $k/2$ are equally likely to be labeled + or -, we get $E[\nu^+] = E[\nu^-]$, and hence, $E[FF(L)] \leq E[MFF(L)]$. The $FF(L)$ and $MFF(L)$ packings have the same occupied space, so $E[W^{FF}(L)] \leq E[W^{MFF}(L)]$, and the lemma is proved. \blacksquare

The preliminaries to the upper bound proof conclude with combinatorial properties of MFF viewed as a matching algorithm. In the instance L , scan the -'s top-down, matching each to

the highest unmatched $+$, if any, above and to the right of the $-$ (this also includes $+$'s directly above the $-$). Figure 3 shows an example. It is easy to see that, in the final matching, two

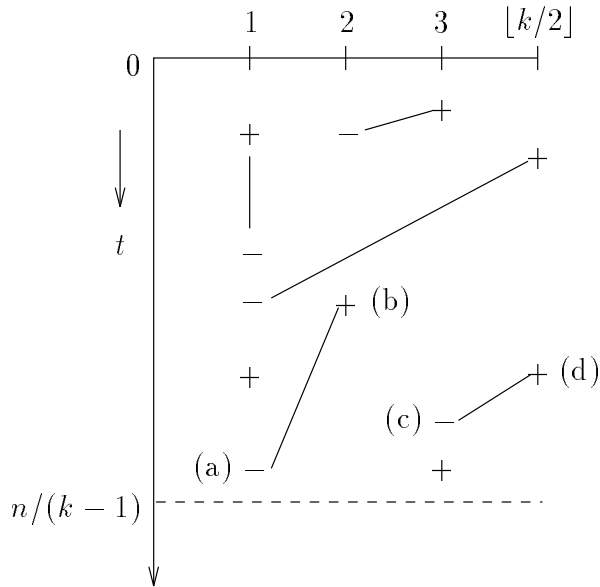


Figure 3: An MFF matching. Points a, b, c , and d illustrate the weak FF property.

items are matched if and only if they are packed in the same bin by MFF. Lemma 2.2 then applies to MFF as a matching algorithm, so by Lemma 3.2,

$$E[W^{FF}(L)] \leq E[W^{MFF}(L)] = \frac{1}{2}E[U^{MFF}(L)]. \quad (3.1)$$

Note that MFF matchings are in the class of *up-right* matchings, i.e., matchings in which each edge goes up and to the right from the $-$ to the $+$ (these edges include those going straight up or directly to the right). Observe also that MFF matchings satisfy the property that, if (a, b) and (c, d) are any two $(-, +)$ edges with time coordinates satisfying $t_a > t_c > t_d > t_b$, then their size coordinates satisfy $s_b < s_c$. (See Fig. 3 for an example.) This property holds because if $s_c \leq s_b$, then MFF would have matched c to b instead of d . We shall be introducing up-right matchings with the following weaker property implied by $s_b < s_c$; such matchings will be easier to analyze.

The Weak FF property: *If (a, b) and (c, d) are $(-, +)$ edges with time coordinates satisfying $t_a > t_c > t_d > t_b$, then their size coordinates satisfy $s_a \leq s_d$.*

In the upper-bound proof below, we analyze an algorithm that generates matchings with the weak FF property. The following result shows that the expected number of points left

unmatched by this algorithm will give an upper bound on the expected number left unmatched in MFF matchings. The proof in [10] for the continuous case carries over directly to the discrete case.

Lemma 3.3. (Shor [10]) *For any instance L , the MFF matching has a cardinality at least that of any up-right matching with the weak FF property.*

With these preliminaries, we are now ready for the upper-bound theorem.

Theorem 3.1. *Let L be a list of n items with sizes drawn independently and uniformly at random from $\{1, \dots, k-1\}$. Then $E[W^{FF}(L)] = O(\sqrt{nk})$.*

Proof. We prove the result for $MFF(L)$ under the Poisson model with parameter n ; the theorem will then follow from (3.1) and Lemma 2.1. First, we prove that the argument can be restricted to odd k . Consider some even k and a random instance $L^{(k)}$ for this bin size. We produce as follows a random instance $L^{(2k-1)}$ for bin size $2k-1$. Each of the first $\frac{k}{2}-1$ columns is expanded into two columns; as shown in Fig. 4, column 1 of $L^{(k)}$ is expanded into columns 1 and 2 of $L^{(2k-1)}$, column 2 of $L^{(k)}$ is expanded into columns 3 and 4 of $L^{(2k-1)}$, and so on, with column $\frac{k}{2}$ of $L^{(k)}$ becoming column $k-1$ of $L^{(2k-1)}$. For each $s = 1, \dots, \frac{k}{2}-1$, the $+$'s and $-$'s in column s of $L^{(k)}$ are each assigned by an independent toss of a fair coin to either column $2s-1$ or column $2s$ of $L^{(2k-1)}$, retaining the same time coordinate in $L^{(2k-1)}$. By the properties of Poisson processes, it is easy to see that this filtering of the original processes does indeed produce a random instance for bin size $2k-1$. (Note the harmless rescaling of the latter instance in terms of $k' = 2k-1$; the standard rate of the Poisson $+$ and $-$ processes has been halved, but the length of the standard interval has been doubled, since $n/(k-1) = 2n/(k'-1)$.)

Now construct a matching of $L^{(k)}$ such that two points of $L^{(k)}$ are matched if and only if they are matched in the MFF matching of $L^{(2k-1)}$ (see Fig. 4). Let A denote this procedure so that $M^A(L^{(k)})$ is the matching constructed. It is easy to verify that the edges of $M^A(L^{(k)})$ must be up-right (since the edges of $M^{MFF}(L^{(2k-1)})$ are) and that, although $M^A(L^{(k)})$ need not be an MFF matching, it retains the weak FF property of such matchings. Then by Lemma 3.3, $U^{MFF}(L^{(k)}) \leq U^A(L^{(k)}) = U^{MFF}(L^{(2k-1)})$, and since $L^{(k)}$ and $L^{(2k-1)}$ are random instances for bin sizes k and $2k-1$, we have $E[U^{MFF}(L^{(k)})] \leq E[U^{MFF}(L^{(2k-1)})]$. We conclude from (3.1) that, if the $O(\sqrt{nk})$ expected waste bound holds for k odd, then it must hold for k even.

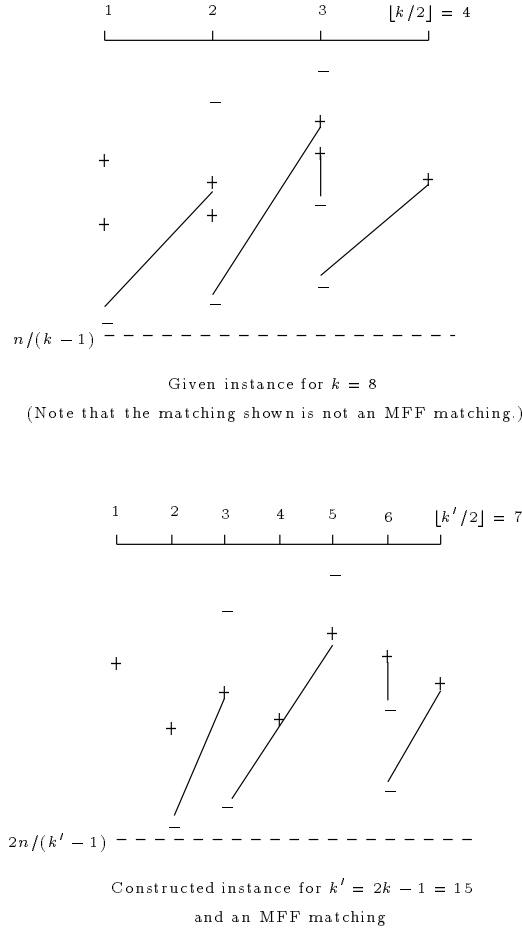


Figure 4: Converting an instance for k to one for $k' = 2k - 1$.

In the remainder of the proof, assume k is odd. Convenient notation will be $k_* = \lfloor k/2 \rfloor$ for the number of columns in Fig. 1, and $n_* = n/(k-1)$, for the length of the time interval. We take n_* to be an integer, an assumption that can affect only hidden multiplicative constants.

Next, we define an algorithm that is easier to analyze than MFF. The algorithm is defined as follows in terms of a grid or lattice superimposed on the instance L . As shown in Fig. 5, place n_* equally spaced grid points on each of the columns, $1, \dots, k_*$. Note that $n_*k_* = n/2$. In the first step, scan the $-$'s top-down, matching each to a highest available grid point, if any, above and to the right of the $-$, resolving ties by choosing the right-most such point. Available means simply that the grid point has not already been matched.

In the second step, the $+$'s are matched to grid points in a complementary fashion. Scanning top-down, match each $+$ to the highest available grid point below and to the left of the $+$, with ties decided in favor of the left-most such point. In this case, available means not already

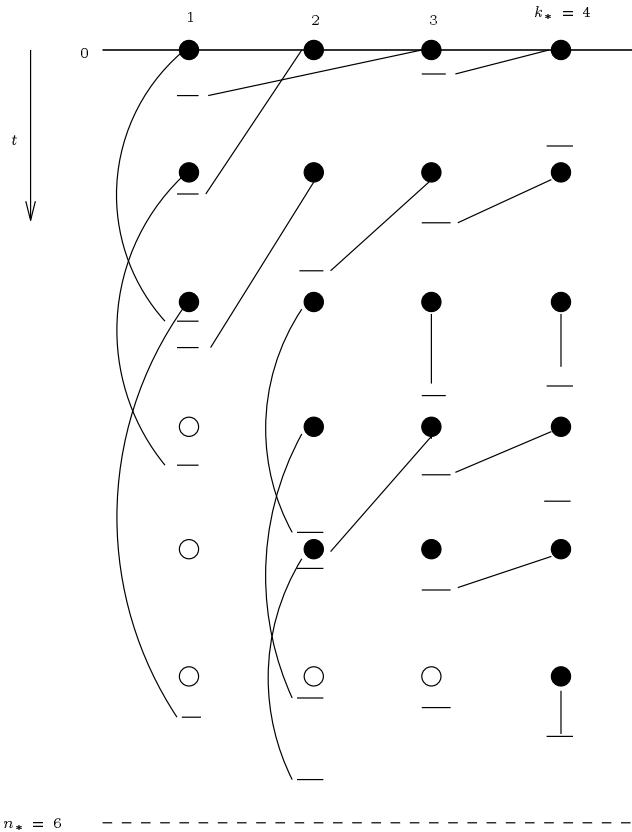


Figure 5: Matching $-$'s to grid points ($+$'s not shown).

matched to a $+$. Finally, the third step constructs an up-right matching of $-$'s to $+$'s by matching a $-$ with a $+$ if and only if they are matched to the same grid point. Denote this final matching by $M = M(L)$ and let $U = U(L)$ be the number of points left unmatched by M . In what follows, the matchings of $-$'s and $+$'s to grid points in the first two steps will be called the $-$ and $+$ grid matchings, respectively.

Lemma 3.4. *The matching M has the weak FF property.*

Proof. To violate the lemma we need $(-, +)$ edges (a, b) and (c, d) in M with $t_a > t_c > t_d > t_b$, but $s_d < s_a$, as shown in Fig. 6. Let g_{ab} and g_{cd} be the grid points coupling a, b and c, d , respectively, according to the grid matchings producing M . Figure 6 shows the three cases where g_{cd} is above, below, and at the same level as g_{ab} . (Note that there are two possibilities each for the relative positions of g_{ab} , g_{cd} and their associated edges. The argument is the same no matter which possibility holds and only one is shown for each case in Fig. 6.) The first case would violate the $+$ grid matching, since b would be matched to g_{cd} as the highest

available below-left grid point. This would also apply in the third, equal-height case, because of the tie-breaking rule. Finally, the second case would violate the $-$ grid matching rule, since c would be matched to g_{ab} as the highest available up-right grid point. ■

In the remaining, probabilistic part of the proof we show that $E[U] = O(\sqrt{nk})$; this bound together with (3.1) and Lemmas 2.1, 3.3, 3.4 proves the theorem. To prove the bound, we analyze the $+$ and $-$ grid matchings individually. The analyses of these two matchings will be rather different, although they eventually reduce to the analysis of very similar interacting particle systems. The reason for the difference is that the $-$ and $+$ grid matchings are not entirely symmetric. To obtain symmetry, we could have required $+$'s to be matched to grid points in a *bottom-up* scan, matching each $+$ to the *lowest* available grid point below and to the left of the $+$, with ties resolved in favor of the left-most such point. But the final matching M would not then have had the weak FF property, as the reader can easily verify.

Let N_+ and N_- be the numbers of $+$'s and $-$'s in L , respectively. In the construction of M , let $U^{(g)}$ be the number of grid points that are not matched to both a $+$ and a $-$. Then $n/2 - U^{(g)}$ is the number of grid points matched to both a $+$ and $-$, and therefore also the number of $+$'s and $-$'s matched in M . So

$$U = N_+ - (n/2 - U^{(g)}) + N_- - (n/2 - U^{(g)}) .$$

We conclude that

$$E[U] = 2E[U^{(g)}] , \tag{3.2}$$

since the means of N_+ and N_- are both $n/2$.

Now define $U_+^{(g)}$ and $U_-^{(g)}$ as the numbers of grid points left unmatched in the $+$ and $-$ grid matchings, respectively. We have $U^{(g)} \leq U_+^{(g)} + U_-^{(g)}$, since the grid points counted by $U_+^{(g)}$ and $U_-^{(g)}$ need not be the same. Then by (3.2),

$$E[U] \leq 2E[U_+^{(g)}] + 2E[U_-^{(g)}] . \tag{3.3}$$

Letting U_+ and U_- be the respective numbers of $+$ and $-$ points left unmatched in the $+$ and $-$ grid matchings, we note also that

$$E[U_+] = E[N_+ - (n/2 - U_+^{(g)})] = E[U_+^{(g)}] \tag{3.4}$$

Below, we show first that $E[U_-^{(g)}] = O(\sqrt{nk})$ and then that $E[U_+] = O(\sqrt{nk})$. By (3.3) and (3.4), the theorem will then be proved.

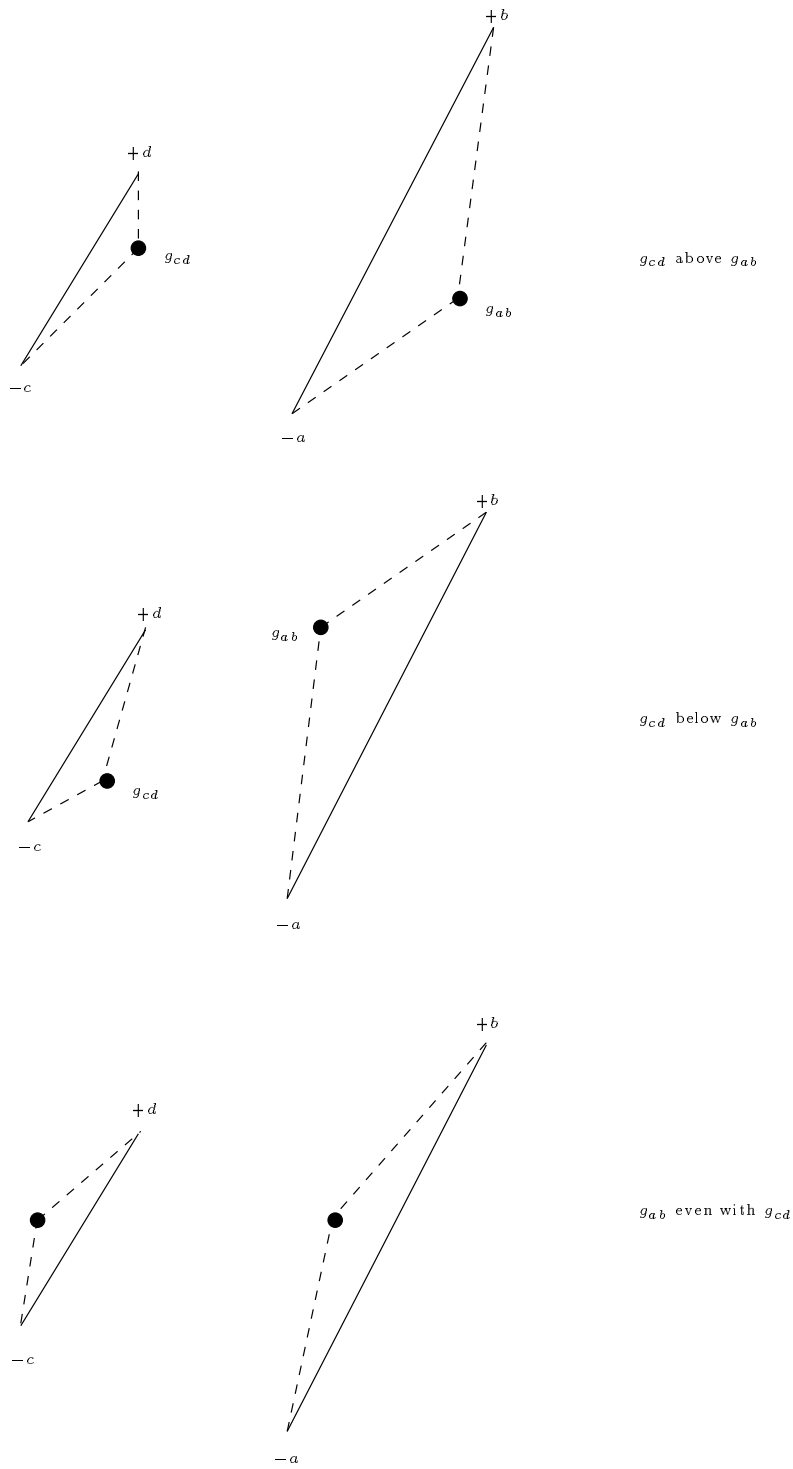


Figure 6: Illustrating Lemma 3.4. (Recall that time increases downward.)

Proof of $E[U_-^{(g)}] = \mathbf{O}(\sqrt{\mathbf{nk}})$: Consider the $-$ grid matching process illustrated in Fig. 5. Let $m_i(t)$ denote the number of unmatched grid points at time t in column i , $1 \leq i \leq k_*$, with vertical coordinates in $[0, t]$, $0 \leq t \leq n_*$. Then

$$E[U_-^{(g)}] = \sum_{i=1}^{k_*} E[m_i(n_*)]. \quad (3.5)$$

From the definition of the matching rule, we see that when a grid point is matched, all grid points up and to its right must already be matched. It follows easily that

$$t \geq m_1(t) \geq m_2(t) \geq \cdots \geq m_{k_*}(t) \geq 0, \quad 0 < t \leq n_* . \quad (3.6)$$

Note that the joint process $\mathbf{m}(t) = (m_1(t), \dots, m_{k_*}(t))$ is Markovian with a continuous time parameter. The $m_i(t)$ may be viewed as the positions of k_* particles in the interacting-particle process $\mathbf{m}(t)$. The interactions are determined by (3.6), i.e., the position of the i^{th} particle acts as a barrier to the motion of the $(i-1)^{\text{st}}$ particle, $1 < i \leq k_*$. At each integer time up to n_* , a new row of unmatched grid points is introduced, so each m_i increases by 1 at integer times. When a $-$ is encountered, say at time t in column i , then the particles do not move if $m_i(t) = m_{i+1}(t) = \cdots = m_{k_*}(t) = 0$; otherwise, $m_j(t)$ decreases by one, where j is the smallest index, $i \leq j \leq k_*$, such that $m_j(t^-) > m_{j+1}(t^-)$, with the conventions that t^- is the time infinitesimally prior to t and that $m_{k_*+1}(t) = 0$, $t \geq 0$. The former event means that the $-$ is left unmatched, and the latter event corresponds to matching the $-$ to the highest available grid point in column j .

Because the sequences of $-$'s in the k_* columns are independent, the behavior of $\mathbf{m}(t)$ suggests that of the order statistics of k_* independent particle processes. Indeed, let $\tilde{m}(t)$, $\tilde{m}(0) = 0$, be a process stochastically the same as one of the $m_i(t)$ in isolation, i.e., $\tilde{m}(t)$ increases by 1 at integer times and decreases by 1 whenever $\tilde{m}(t^-) > 0$ and a $-$ appears at time t . A comparison of transition probabilities shows that, if $\tilde{m}_i(t)$, $1 \leq i \leq k_*$, denote k_* independent copies of $\tilde{m}(t)$ with $\tilde{m}_i(0) = 0$, $1 \leq i \leq k_*$, then $(m_1(t), \dots, m_{k_*}(t))$ is equal in distribution to $(\tilde{m}_{(1)}(t), \dots, \tilde{m}_{(k_*)}(t))$, where $\tilde{m}_{(i)}(t)$ denotes the i^{th} largest of the $\tilde{m}_i(t)$. But then, by (3.5),

$$\begin{aligned} E[U_-^{(g)}] &= \sum_{i=1}^{k_*} E[\tilde{m}_{(i)}(n_*)] = \sum_{i=1}^{k_*} E[\tilde{m}_i(n_*)] \\ &= k_* E[\tilde{m}(n_*)], \end{aligned} \quad (3.7)$$

so our problem has reduced to the analysis of the one-dimensional process $\tilde{m}(t)$.

Let $\{W_j\}_{j \geq 0}$ be the Markov chain embedded in $\tilde{m}(t)$ just before integer times, with $W_0 = 0$. This chain is a random walk with a barrier at the origin and transitions $W_{j-1} \rightarrow W_j$, $j \geq 1$, that balance an increase of one at time $j - 1$ with a decrease of one for every $-$ that appears in $(j - 1, j)$ at times t with $\tilde{m}(t^-) > 0$. Then

$$W_j = (W_{j-1} + \xi_j)^+, \quad j = 1, 2, \dots, \quad (3.8)$$

where the ξ_j are independent and equal in distribution to $1 - \zeta_j$, with the ζ_j being independent Poisson distributed random variables having unit means and variances. Since the ξ_j satisfy (2.4) we obtain from Lemma 2.3,

$$E[W_{n_*}] = O(\sqrt{n_*}). \quad (3.9)$$

Now $\tilde{m}(n_*) = W_{n_*} + 1$ almost surely, so (3.7) and (3.9) along with $k_*\sqrt{n_*} = \sqrt{nk_*/2} = \frac{1}{2}\sqrt{n(k-1)}$ give the desired result, $E[U_-^{(g)}] = O(\sqrt{nk})$.

Proof of $E[U_+] = O(\sqrt{nk})$: The $+$ grid matching process is illustrated in Fig. 7. As can be

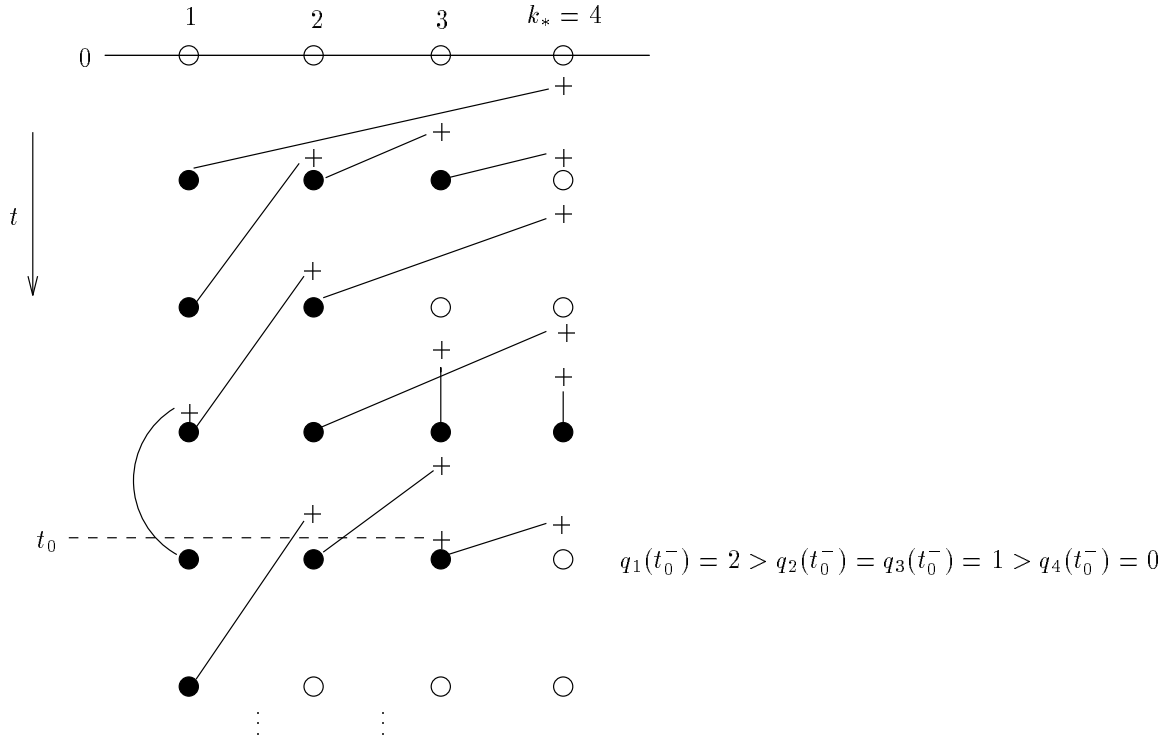


Figure 7: The $+$ matching just before a $+$ at t_0 is matched.

seen, the matched grid points no longer have the simple structure of Fig. 5; the grid points

of any column are not necessarily matched in a contiguous sequence starting from the top. However, we do have the following useful property. At time t in the matching process, let $q_i(t) \geq 0$ denote the number of grid points below t in column i that have already been matched to $+$'s encountered in $[0, t]$. An easy induction establishes that

$$q_1(t) \geq q_2(t) \geq \cdots \geq q_{k_*}(t) \geq 0 ,$$

and that in column i at time t the matched grid points with vertical coordinates exceeding t form a contiguous sequence for each i , $1 \leq i \leq k_*$. Figure 7 gives an illustration.

Define a process $\tilde{q}(t)$, $\tilde{q}(0) = 0$, on a single column having a Poisson pattern of $+$'s at rate 1; $\tilde{q}(t)$ decreases by 1 at integer times t when $\tilde{q}(t^-) > 0$ and increases by 1 whenever a $+$ is encountered. Note the symmetry between $\tilde{q}(t)$ and the process $\tilde{m}(t)$ of the earlier analysis. Let $\tilde{q}_1(t), \dots, \tilde{q}_{k_*}(t)$ be k_* independent copies of $\tilde{q}(t)$ with $\tilde{q}_i(0) = 0$, $1 \leq i \leq k_*$. Using the earlier arguments, the key observation now is that $(q_1(t), \dots, q_{k_*}(t))$ is equal in distribution to $(\tilde{q}_{(1)}(t), \dots, \tilde{q}_{(k_*)}(t))$.

To make use of this fact, extend the lattice below n_* , and extend the matching process so that *all* $+$'s are matched to grid points. Then at time n_* the $+$'s matched to grid points below n_* are just those $+$'s left unmatched by the original $+$ grid matching process. Thus,

$$E[U_+] = \sum_{i=1}^{k_*} E[q_i(n_*)] = \sum_{i=1}^{k_*} E[\tilde{q}_{(i)}(n_*)] = k_* E[\tilde{q}(n_*)] . \quad (3.10)$$

As before, to obtain $E[\tilde{q}(n_*)]$ we analyze a Markov chain $\{\widehat{W}_j\}_{0 \leq j \leq n_*}$ embedded in $\tilde{q}(t)$ at instants just before integer times, with $\widehat{W}_0 = 0$. In analogy with (3.8), we get

$$\widehat{W}_j = (\widehat{W}_{j-1} + \widehat{\xi}_j)^+, \quad j = 1, 2, \dots , \quad (3.11)$$

where $\widehat{\xi}_j$ has zero mean and unit variance. Lemma 2.3 again gives $E[\widehat{W}_{n_*}] = O(\sqrt{n_*})$ and hence $E[\tilde{q}(n_*)] = O(\sqrt{n_*})$. By (3.10), we obtain $E[U_+] = O(\sqrt{nk})$ as before, thus completing the proof of the theorem. ■

Recall that, in the continuous problem, the bin size is 1 and the item sizes are drawn from the uniform distribution on $[0, 1]$. Together with appropriate transformations between discrete and continuous problem instances, Theorem 3.1 provides the basis for proving a tight upper bound for the continuous case.

Theorem 3.2. *Let n item sizes be drawn independently from the uniform distribution on $[0, 1]$. Then*

$$E[FF(L)] = \Theta(n^{2/3}) .$$

Proof. The lower bound $E[FF(L)] = \Omega(n^{2/3})$ was proved in [10], so we need only prove $E[FF(L)] = O(n^{2/3})$. As in Theorem 3.1, the proof reduces to the analysis of an upright matching problem. Again, we prove the result for the Poisson model, and then apply Lemma 2.1.

In analogy with the discrete case, consider an instance L in two dimensions as shown in Fig. 8, where the N points are those of a Poisson pattern with intensity n that fall in the rectangle $[0, \frac{1}{2}] \times [0, 1]$. The first step of the proof transforms this instance into one for a

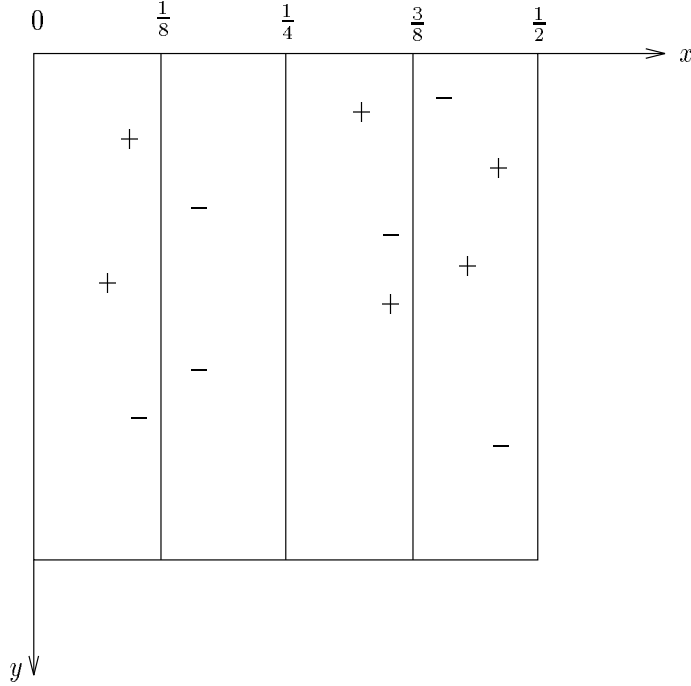


Figure 8: A continuous instance L .

discrete uniform distribution on item sizes, so that we can apply Theorem 3.1. Let k_n , $n \geq 1$, be a sequence of odd integers such that $k_n = \Theta(n^{1/3})$. For given n , construct verticals at $x = \frac{i}{2(k_*+1)}$, $0 \leq i \leq k_* + 1 = \lfloor k_n/2 \rfloor + 1$, as shown in Fig. 8. Move each $+$ horizontally to the nearest vertical on its left, and move each $-$ horizontally to the nearest vertical on its right.

Consider the resulting discrete instance L' restricted to the interior verticals at $x = \frac{i}{2(k_*+1)}$, $1 \leq i \leq k_*$. The second step inserts the edges of an MFF matching for L' . Figure 9 gives an illustration. Recall that this set of edges has the weak FF property. Let U' be the number of points in L' unmatched by the edges.

The number N' of points in L' has the expected value $E[N'] = E[N] - E[Z]$, where Z

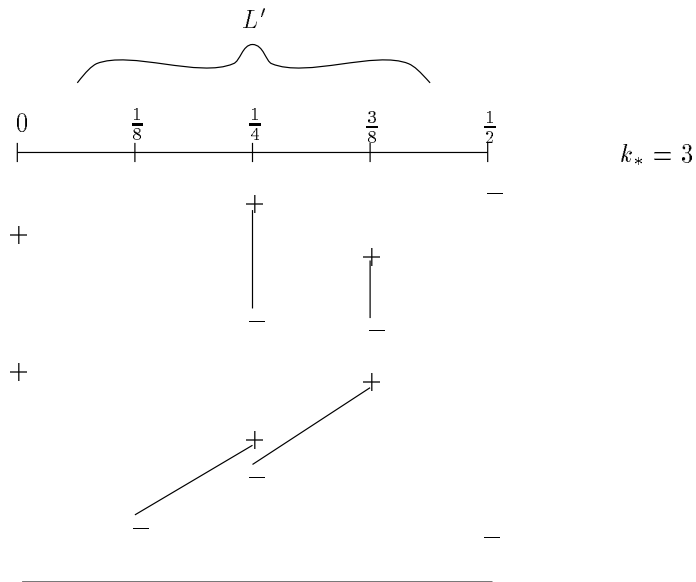


Figure 9: A matching for L' after shifting the points of L in Fig. 8.

counts all of the boundary points on the verticals at $x = 0$ and $x = 1$. By our choice of k , the numbers of $+$'s and $-$'s on each of the $k_* + 1$ verticals are Poisson distributed with mean $\frac{1}{2}E[N]/(k_* + 1) = \Theta(n^{2/3})$. Then $E[Z] = \Theta(n^{2/3})$ and $E[N'] = \Theta(n)$. Thus, by Theorem 3.1 the expected number of unmatched points in L' is $E[U'] = O(\sqrt{nk})$ and hence $E[U'] = O(n^{2/3})$.

The third and final step shifts all of the $+$'s and $-$'s, including those on the boundaries, back to their original positions, extending the edges between paired points as necessary to preserve the matching. Trivially, since the $+$'s move right and the $-$'s move left in this step, the edges remain up-right. Moreover, it is easy to see that the weak FF property of the edges is preserved under the shift of the $+$'s and $-$'s. Thus, by Lemma 3.3 applied to the continuous case [10], the expected number of points left unmatched by the MFF matching of the original instance L is at most $E[U^{MFF}(L)] \leq E[U'] + E[Z]$. Since $E[U']$ and $E[Z]$ are both $O(n^{2/3})$, we conclude that $E[U^{MFF}(L)] = O(n^{2/3})$. (Note that we needed the choice $k = \Theta(n^{1/3})$ so as to balance $E[U']$ and $E[Z]$, i.e., to have both bounded by $O(n^{2/3})$.) Then (3.1) applied to the continuous case [10] proves the theorem. \blacksquare

Returning to the discrete uniform distribution, we remark that if $k = k_n$ is taken as a function of n in Theorem 3.1, then the $O(\sqrt{nk})$ upper bound is sharp only if k_n grows sufficiently slowly. As might be expected, if the discretization is fine enough relative to n , then the asymptotic result for the discrete case is the same as for the continuous case. In particular,

we have the following corollary to Theorem 3.2.

Corollary 3.1. *Let the item sizes in L have a discrete uniform distribution with $k = \Omega(n^{1/3})$.*

Then

$$E[W^{FF}(L)] = O(n^{2/3}) .$$

Proof. The proof technique is similar to that in Theorem 3.2. Consider only odd values of k ; even k can be handled as in Theorem 3.1. Suppose we are given a random instance L for the continuous problem with intensity $n(k+1)/(k-1)$. Construct $(k-1)/2$ equally spaced interior verticals through $[0, 1/2] \times [0, 1]$ as shown in Fig. 10. In each of the $\frac{k+1}{2}$ vertical strips move $-$'s to the nearest vertical on the left and $+$'s to the nearest vertical on the right leaving time coordinates unchanged. Now look upon the $(k-1)/2$ interior verticals as an instance $L^{(k)}$ for bin size k . A column in $L^{(k)}$ receives $+$'s from the vertical strip to its left and $-$'s from the vertical strip to its right; the total number of these points is equal in distribution to the total number in a vertical strip, and hence is Poisson distributed with mean $\frac{n(k+1)}{k-1} \cdot \frac{1}{(k+1)/2} = \frac{2n}{k-1}$. Thus, the instance $L^{(k)}$ is a random instance of N items for bin size k , where N is Poisson distributed with mean n .

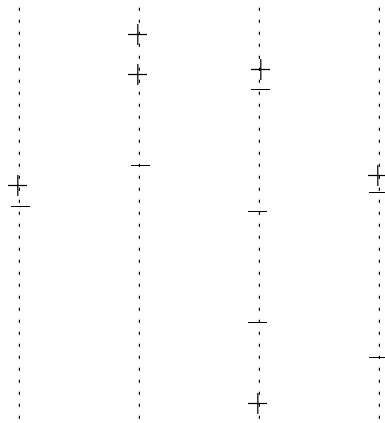
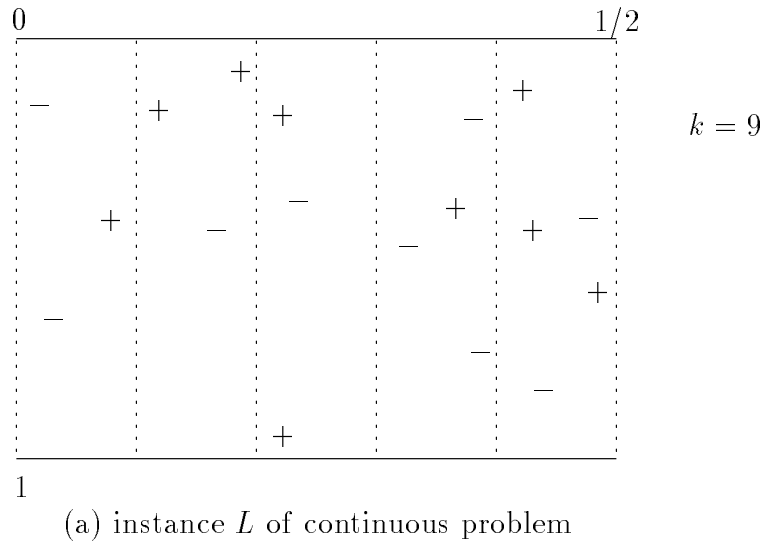
Next, construct a matching of points in $L^{(k)}$ by matching two points if and only if they are matched in the MFF matching $M^{MFF}(L)$ of the original continuous instance. Let A denote this procedure so that $M^A(L^{(k)})$ denotes the matching of the points in $L^{(k)}$. By the construction process ($+$'s move right and $-$'s move left), the edges of $M^A(L^{(k)})$ must be up-right since the edges of M^{MFF} are. Moreover, although $M^A(L^{(k)})$ need not be an MFF matching, it retains the weak FF property of such matchings. Points of the continuous instance that were moved to the left or right boundary verticals, and hence outside $L^{(k)}$, may have been matched in $M^{MFF}(L)$ (and not necessarily with each other). Thus, if Z counts the points moved to boundary verticals, then

$$|M^A(L^{(k)})| \geq |M^{MFF}(L)| - Z .$$

Now substitute $|M^A(L^{(k)})| = \frac{1}{2}[N - U^A(L^{(k)})]$ and $|M^{MFF}(L)| = \frac{1}{2}[N' - U^{MFF}(L)]$ where N and N' are the numbers of points in $L^{(k)}$ and L , respectively. Then since $N' \geq N$, we obtain $U^A(L^{(k)}) \leq U^{MFF}(L) + 2Z$, so by Lemma 3.3,

$$U^{MFF}(L^{(k)}) \leq U^{MFF}(L) + 2Z .$$

It remains to observe that, since $E[N'] = \Theta(n)$, we have $E[U^{MFF}(L)] = O(n^{2/3})$ by Theorem 3.2, and since $k = \Omega(n^{1/3})$, we have $E[2Z] = \frac{n(k+1)}{k-1} \cdot \frac{2}{k+1} = \frac{2n}{k-1} = O(n^{2/3})$. Thus,



(b) instance $L^{(k)}$

Figure 10: Transformation for Corollary 3.1.

$E[U^{MFF}(L^{(k)})] = O(n^{2/3})$ so (3.1) completes the proof. ■

4. The FF Lower Bound

For the discrete uniform distribution, the following lower bound shows that the estimates in Theorem 3.1 and Corollary 3.1 are smallest possible asymptotic upper bounds.

Theorem 4.1. *Let L be a list of n items with sizes drawn independently and uniformly at random from $\{1, \dots, k-1\}$.*

(i) *If $k = O(n^{1/3})$, then $E[W^{FF}(L)] = \Omega(\sqrt{nk})$.*

(ii) *If $k = \Omega(n^{1/3})$, then $E[W^{FF}(L)] = \Omega(n^{2/3})$, as in the continuous case.*

Proof. As in the upper-bound proof, the result is proved for the Poisson model, whereupon Lemma 2.1 completes the proof. We start by proving part (i).

The proof begins with a sequence of reductions in Lemmas 4.1–4.3. Ultimately, we show that the $\Omega(\sqrt{nk})$ lower bound holds if the number of points left unmatched by a certain planar matching algorithm to be defined has the same asymptotic lower bound. The remainder of the proof then proves the latter lower bound. Lemmas 4.1 and 4.2 below adapt very similar results in [10].

Consider an FF packing of L and focus on just those bins having items in $(k/3, k/2)$ and $(k/2, 2k/3)$; these are simply the items in $(k/3, 2k/3)$ if k is odd, but if k is even, the items of size $k/2$ are excluded. There are at most two such items per bin. Call the items in $(k/3, k/2)$ *s-items* and those in $(k/2, 2k/3)$ *b-items* (s and b stand for smaller and bigger, respectively). In terms only of s - and b -items there are just 5 types of bins; these are denoted $\begin{bmatrix} b \\ b \end{bmatrix}$, $\begin{bmatrix} s \\ b \end{bmatrix}$, $\begin{bmatrix} s \\ s \end{bmatrix}$, $\begin{bmatrix} b \\ s \end{bmatrix}$, $\begin{bmatrix} s \\ s \end{bmatrix}$, with $|b|$, $|s|$, $|s|$, $|b|$, $|s|$ giving the numbers of the respective types. The letters in this notation stand for the item types, and the vertical positions denote the order of packing, i.e., the item on the bottom was packed before the item, if any, on top.

The number of item sizes in each of $(k/3, k/2)$ and $(k/2, 2k/3)$ is approximately $n/6$. The precise formula is of no interest; we need only the fact that it is asymptotically $n/6 + O(1)$. Let $\beta_k n/6$ denote the expected numbers of s and b items so that $\beta_k = 1 + O(1/k)$.

Our first reduction is given by

Lemma 4.1. *If*

$$E \left| \begin{bmatrix} s \\ b \end{bmatrix} \right| = \beta_k \frac{n}{6} - \Omega(\sqrt{nk}), \quad (4.1)$$

then $E[W^{FF}(L)] = \Omega(\sqrt{nk})$.

Proof. We first establish three relations among the quantities $|_b|$, $|_s|$, $|_s^s|$, $|_s^b|$, $|_b^s|$. The first simply breaks down the expected number of b -items by bin type,

$$E \left| \begin{smallmatrix} s \\ b \end{smallmatrix} \right| + E \left| \begin{smallmatrix} | \\ b \end{smallmatrix} \right| + E \left| \begin{smallmatrix} b \\ s \end{smallmatrix} \right| = \beta_k \frac{n}{6}. \quad (4.2)$$

Equating the left-hand side of (4.2) to the expected number of s -items, $E \left| \begin{smallmatrix} | \\ s \end{smallmatrix} \right| + 2E \left| \begin{smallmatrix} s \\ s \end{smallmatrix} \right| + E \left| \begin{smallmatrix} b \\ s \end{smallmatrix} \right| + E \left| \begin{smallmatrix} s \\ b \end{smallmatrix} \right|$, gives the second relation

$$E \left| \begin{smallmatrix} | \\ b \end{smallmatrix} \right| = 2E \left| \begin{smallmatrix} s \\ s \end{smallmatrix} \right| + E \left| \begin{smallmatrix} | \\ s \end{smallmatrix} \right|. \quad (4.3)$$

The third relation is given by

$$E \left| \begin{smallmatrix} b \\ s \end{smallmatrix} \right| \leq E \left| \begin{smallmatrix} s \\ s \end{smallmatrix} \right|. \quad (4.4)$$

To prove (4.4), it suffices to show that at any time in the packing process, the probability that the next item creates a bin of type $\left[\begin{smallmatrix} b \\ s \end{smallmatrix} \right]$ is at most the probability that it creates a bin of type $\left[\begin{smallmatrix} s \\ s \end{smallmatrix} \right]$. Neither type bin can be created unless the current packing contains a bin B of type $\left[\begin{smallmatrix} | \\ s \end{smallmatrix} \right]$. Furthermore, by definition of the FF rule, B must be the only such bin in the current packing. Let a be the level of B . Then, again by the FF rule, B must also be the only bin filled to a level in $[a, k - a]$. Therefore, if the next item has a size in $[a, k - a]$, we can pack it directly into B without attempting to pack it elsewhere. But given that the next item has a size in $[a, k - a]$, it is as likely to be an s -item as a b -item. If the next item has a size that is not in that range, then either it is a b -item too large to fit in any partially filled bin, or it is an s -item. In neither case can it create a type $\left[\begin{smallmatrix} b \\ s \end{smallmatrix} \right]$ bin. Thus, the next item is at least as likely to create an $\left[\begin{smallmatrix} s \\ s \end{smallmatrix} \right]$ -type bin as a $\left[\begin{smallmatrix} b \\ s \end{smallmatrix} \right]$ -type bin. Then (4.4) holds.

Now substitute (4.3) into (4.2) and apply (4.4) to obtain

$$E \left| \begin{smallmatrix} s \\ b \end{smallmatrix} \right| + 3E \left| \begin{smallmatrix} s \\ s \end{smallmatrix} \right| + E \left| \begin{smallmatrix} | \\ s \end{smallmatrix} \right| \geq \beta_k \frac{n}{6}.$$

Substitution of (4.1) then gives

$$3E \left| \begin{smallmatrix} s \\ s \end{smallmatrix} \right| + E \left| \begin{smallmatrix} | \\ s \end{smallmatrix} \right| = \Omega(\sqrt{nk}). \quad (4.5)$$

Finally, return to the set of all items and all bins. The bound (4.5) shows that, on the average, there are at least $\Omega(\sqrt{nk})$ bins containing only items smaller than $k/2$. Since there is an average of $n/2$ bins with items larger than $k/2$, we conclude that $E[W^{FF}(L)] = \Omega(\sqrt{nk})$, as desired. ■

To prove that the condition in (4.1) indeed holds, we first convert it to an assertion about up-right matchings. Let L^* denote the sublist of L containing just those items in $(k/3, k/2)$ and $(k/2, 2k/3)$; the ordering of these items in L is retained in L^* . Note that the number of items in L^* is Poisson distributed with mean $\beta_k n/3$. The second reduction is given in

Lemma 4.2. *If*

$$E[W^{MFF}(L^*)] = \Omega(\sqrt{nk}) , \quad (4.6)$$

then (4.1) and hence $E[W^{FF}(L)] = \Omega(\sqrt{nk})$ holds.

Proof. In the two-dimensional representation of L^* , match a $+$ with a $-$ if and only if the corresponding items define a type- $\begin{bmatrix} s \\ b \end{bmatrix}$ bin in the FF packing of L . Let M^* be the resulting matching. The expected numbers of $+$'s and $-$'s in L^* are each $\beta_k n/6$, so (4.1) will be proved if we can show that the expected number of items in L^* left unmatched by M^* is $\Omega(\sqrt{nk})$.

It is easy to verify that M^* is an up-right matching with the weak FF property, even though it is not necessarily an MFF matching. (For example, M^* may fail to be an MFF matching because of $+$'s that would be matched to $-$'s by MFF, but are not so matched in M^* because the items corresponding to these $+$'s were packed together with items no larger than $k/3$ in the original FF packing.) Lemma 3.3 then shows that the expected number of points left unmatched by M^* is at least $E[U^{MFF}(L^*)] = 2E[W^{MFF}(L^*)] = \Omega(\sqrt{nk})$, by (4.6). ■

To prove that (4.6) actually holds, we introduce another matching algorithm that is easier to analyze and has the same asymptotic lower bound. Relaxing the “up-right” of MFF to just “right,” we define the *rightward matching* (RM) algorithm as follows: RM scans the $-$'s top-down, matching each to the highest unmatched $+$, if any, to the right of (or in the same column as) the $-$. Our third reduction shows that, if $E[U^{RM}(L^*)] = \Omega(\sqrt{nk})$, then (4.6) holds and hence the theorem is proved.

Lemma 4.3. *For all lists L , $U^{RM}(L) \leq U^{MFF}(L)$.*

Proof. Consider an MFF matching as shown in Fig. 11(a). Shift the instance of $-$'s down to the time interval $(n_*, 2n_*)$ so that all $-$'s are below all $+$'s, and extend all edges so that the same pairs of points are matched. Perform the same operation on the corresponding RM matching shown in Fig. 11(b). The new matching produced from RM is clearly an MFF matching. The new matching produced from MFF is not necessarily an MFF matching (see Fig. 11), but it is easy to see that the weak FF property is preserved. Then Lemma 3.3 proves that $U^{RM}(L) \leq U^{MFF}(L)$. ■

It remains to show that $E[U^{RM}(L^*)] = \Omega(\sqrt{nk})$. It is helpful to simplify notation by recasting the problem in simpler terms. Recall that item sizes in L^* are drawn from a set of $k/3 + O(1)$ sizes symmetric about $k/2$ with no items of size $k/2$. For notational convenience, let us say that the number of item sizes is $k' - 1$, where note that k' is odd. The expected number of items in L^* is $n' = \beta_k n/3$, so consider a new problem with lists L' of item sizes drawn from $1, \dots, k' - 1$, with bin size k' , and with a mean number n' of items. Because of the matching algorithms being analyzed, it is easy to see that $U^{RM}(L^*)$ and $U^{RM}(L')$ are equal in distribution. It remains to observe that the problem with lists L' differs by constant factors (in both the time and size coordinates) from our original problem with parameters k and n , where k is restricted to odd values. Since a return to our original problem formulation with k odd can only affect hidden multiplicative constants, it suffices to show that

$$E[U^{RM}(L)] = \Omega(\sqrt{nk}) \quad (4.7)$$

for k odd. As before $k_* = (k - 1)/2$ and $n_* = n/(k - 1)$ with n_* an integer. We prove (4.7) by a detailed analysis of the RM process.

The analysis starts with a sample of +’s and –’s extended over the entire time axis. Define the process $\mathbf{y}(t) = (y_1(t), \dots, y_{k_*}(t))$, $t \geq 0$, specifying the time coordinates of just those +’s available for matching to the next –. Thus, $y_i(t)$, $1 \leq i \leq k_*$, is defined as the time coordinate of the + that would be matched to the next –, if this – were to appear in column i . Figure 12 gives an illustration of $\mathbf{y}(t)$. Note that, in the initial state, $y_{k_*}(0)$ is the position of the highest + in column k_* , and $y_i(0)$, $i = k_* - 1, \dots, 1$, is the minimum of $y_{i+1}(0)$ and the position of the highest + in column i .

Note also that, in general, $\mathbf{y}(t)$ consists of a sequence of runs or *clusters*, each being a maximum-length sequence of one or more contiguous y_i ’s all having the same value, i.e., all specifying the position of the same +, which appears in the rightmost column of the cluster. Thus, if the next – after time t appears in any one of the columns of the cluster $y_{j_1}(t), \dots, y_{j_2}(t)$, then it will be matched to the + in column j_2 . Bearing this in mind, it will also be convenient to say that the – is matched to the cluster $y_{j_1}(t), \dots, y_{j_2}(t)$. Because of the rule “highest and to the right,” an easy induction shows that the y_i ’s are nondecreasing, i.e., $y_1(t) \leq \dots \leq y_{k_*}(t)$.

Transitions in $\mathbf{y}(t)$ occur when –’s are encountered and then matched, thus removing one of the available +’s. Define $y_{k_*+1}(t) = \infty$ for all t . Suppose that a – is encountered at time t in column j and is matched to the cluster $y_{j_1}(t^-), \dots, y_{j_2}(t^-)$, i.e., $j_1 \leq j \leq j_2$. Then after

$$k_* = 3$$

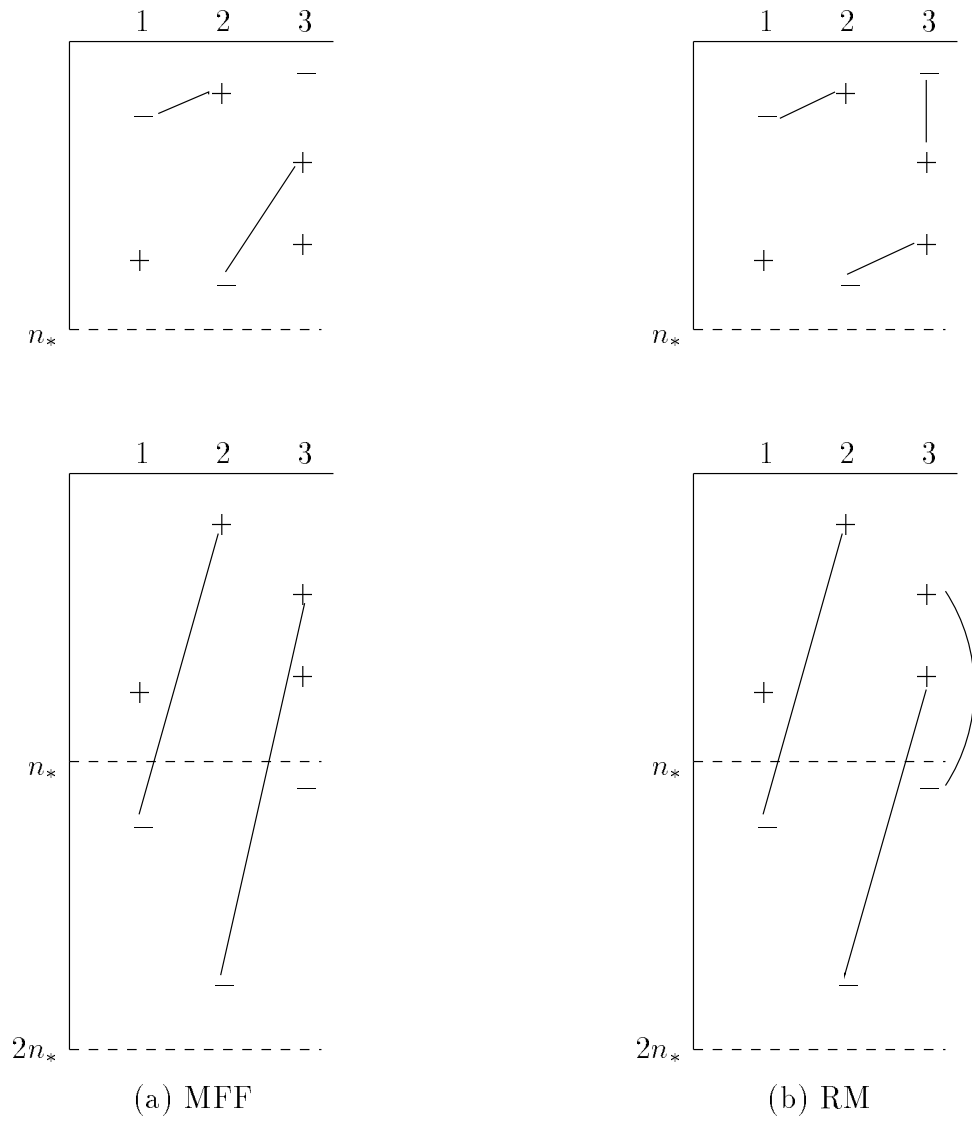


Figure 11: Comparison of RM and MFF

the $-$ is matched to the $+$ in column j_2 , $y_{j_2}(t)$ becomes the minimum of $y_{j_2+1}(t^-)$ and the time coordinate of the highest $+$ in column j_2 after time t . Then, in the decreasing order $i = j_2 - 1, \dots, j_1$, $y_i(t)$ becomes the minimum of $y_{i+1}(t)$ and the time coordinate of the highest $+$ in column i after time t . Observe that a transition affects only the components $y_i(t^-)$ of the cluster matched to the $-$ encountered at time t .

It is also convenient to think of $\mathbf{y}(t)$ as a Markov interacting-particle process. Each column has a particle whose initial position is $y_i(0)$. Thereafter, the particle in column i jumps to a new position whenever the next $-$ appears in a column in the same cluster as column i . Interactions of the particles take the form of the “no-passing” rules defined by the jumps of $\mathbf{y}(t)$; whenever particle i attempts to jump to a $+$ lower than the particle at $y_{i+1}(t)$, it stops at position $y_{i+1}(t)$. We say that particle i *collides* with particle $i + 1$ in these circumstances.

The proof of (4.7) balances two different lower bounds expressed in terms of a parameter that measures the dispersion in the particle process $\mathbf{y}(t)$. Let $d_i(t) = y_{i+1}(t) - y_i(t)$, $t \geq 0$, $1 \leq i < k_*$, denote the difference in the positions of adjacent particles, and define the dispersion of the middle third of the particles by $d(t) = \sum_{i=\frac{k_*}{3}+1}^{2k_*/3} d_i(t) = y_{2k_*/3+1}(t) - y_{k_*/3+1}(t)$. Instead of attempting to estimate $E[d(t)]$, we prove the desired bound, first assuming that $E[d(t)]$ is large for some t , $0 \leq t \leq n_*$, and then assuming that it is small for all t , $0 \leq t \leq n_*$.

Specifically, consider any constant $\alpha > 0$, and let $\{n_l\}$ denote the subset of the positive integers such that for every $n \in \{n_l\}$,

$$E[d(t)] \geq \alpha \sqrt{\frac{n}{k}} \quad \text{for some } t, 0 < t \leq n_* . \quad (4.8)$$

Let $\{n'_l\}$ denote the positive integers not in $\{n_l\}$, so that for every $n \in \{n'_l\}$

$$E[d(t)] \leq \alpha \sqrt{\frac{n}{k}} \quad \text{for all } t, 0 < t \leq n_* . \quad (4.9)$$

Case 1 $n \in \{n_l\}$. We assume that $\{n_l\}$ is infinite; otherwise, (4.7) needs to be proved only for Case 2 with $n \in \{n'_l\}$. To continue with compact asymptotic notation, we now assume that $\{n_l\}$ is the entire set, i.e., (4.8) holds for all $n \geq 1$. Adapting the arguments below to restricted n is only a notational matter.

Let U_t^{RM} denote the number of unmatched points in $[0, t]$. These include unmatched $+$'s above t , and $-$'s above t that are matched to $+$'s below t . We first show that (4.8) implies

$$E[U_t^{RM}] = \Omega(\sqrt{nk}) . \quad (4.10)$$

The remainder of the proof then shows that $E[U_t^{RM}]$ increases with t and hence that $E[U^{RM}] \equiv E[U_{n_*}^{RM}]$ has the same lower bound.

Assume for the moment that $y_1(t) \leq t \leq y_{k_*}(t)$ with $d(t)$ satisfying (4.8), as illustrated in Fig. 13. Let A_1 and A_2 denote the two regions between t and $\mathbf{y}(t)$, the first being above t and the second below t (see Fig. 13). A_1 and A_2 will also denote the areas of these regions. U_t^{RM} is the sum of the number of +’s in A_1 , none of which are matched at time t , and the number of +’s in A_2 , all of which are matched to –’s above t , but are not themselves in the problem instance above t . The +’s occur at rate 1 in each column, so $E[U_t^{RM}] \geq E[A_1 + A_2]$. But by (4.8), $E[A_1 + A_2] \geq \frac{k_*}{3}\alpha\sqrt{\frac{n}{k}}$ (see Fig. 13). This gives $E[U_t^{RM}] = \Omega(\sqrt{nk})$, as desired. Similar arguments apply to the simpler cases $t < y_1(t)$ and $t > y_{k_*}(t)$ (A_1 or A_2 is empty), and yield the same result.

It remains to show that (4.10) implies a similar lower bound at time n_* . Consider the extension of the matching from time t to time n_* , and assume first that $y_1(t) \leq t \leq y_{k_*}(t) \leq n_*$, as illustrated in Fig. 13. Let R_1 and R_2 be the regions between t and n_* to the left of column j and to the right of column $j - 1$, respectively, with j being the column where $\mathbf{y}(t)$ crosses t . At time t , the –’s in A_1 are already matched, but the +’s in A_1 are still available. The intensity of +’s is the same as that of –’s, so the expected number of available +’s in $R_1 \cup A_1$ exceeds by $E[A_1]$ the expected number of –’s in R_1 that need to be matched. Then the expected number of unmatched +’s in columns 1 through $j - 1$ at time n_* is at least the expected number of unmatched +’s in these columns at time t . Indeed, the number at time n_* will tend to be even larger because some of the –’s in R_1 may be matched to +’s in R_2 .

A complementary argument applies to columns j through k_* . The +’s but not the –’s are matched in A_2 , so the expected number of –’s in R_2 that need matching exceeds by $E[A_2]$ the expected number of +’s available in R_2 . Then the expected number of unmatched –’s at time n_* in columns j through k_* is at least the expected number of unmatched –’s in these columns at time t . Again, the expected number of unmatched –’s at time n_* will tend to be even greater because some +’s in R_2 are matched to –’s in R_1 . We conclude that $E[U_{n_*}^{RM}] \geq E[U_t^{RM}]$.

Similar area arguments apply to the various cases when $n_* < y_{k_*}(t)$, $t < y_1(t)$, or $t > y_{k_*}(t)$. These are left to the reader. ■

Case 2 $n \in \{n'_i\}$. As before, we assume that $\{n'_i\}$ is infinite, since otherwise we are done. And again, to simplify notation we assume that (4.9) holds for all $n \geq 1$. We prove that, if (4.9)

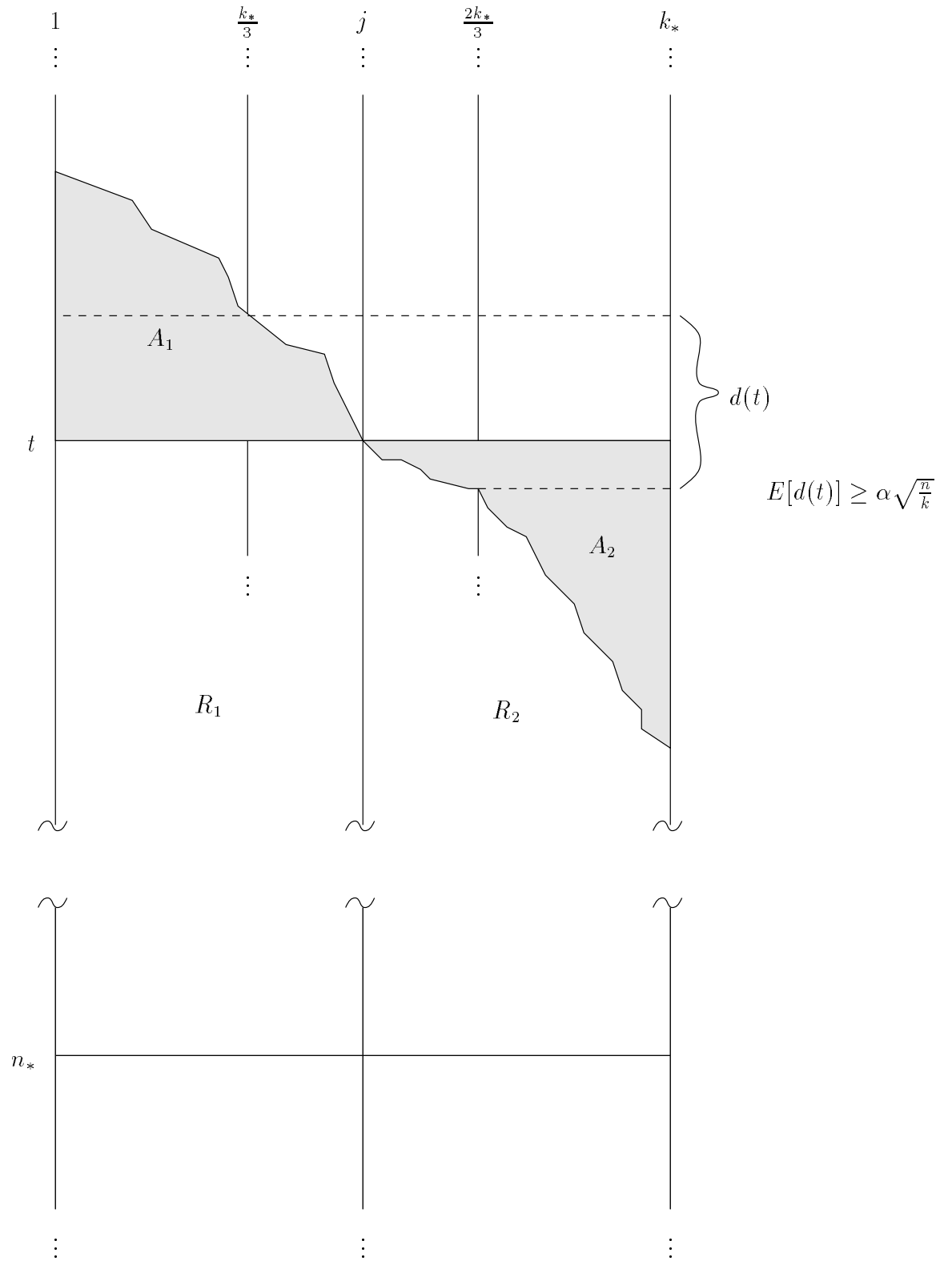


Figure 13: Example for case 1.

holds, then $\Omega(\sqrt{nk^3})$ is a lower bound on the expected sum of horizontal components H_i in an RM matching. Lemma 2.2 shows that, after dividing by k , we get $E[U^{RM}] = \Omega(\sqrt{nk})$ as desired.

Consider a cluster C of size h just after the j^{th} jump of $\mathbf{y}(t)$, and suppose the $(j+1)^{\text{st}}$ jump is caused by a $-$ in a column of C . Cluster C was created by the $h-1$ most recent particle collisions in the first $h-1$ columns of C . The $-$ causing the $(j+1)^{\text{st}}$ jump will be matched to a $+$ in the last column of C ; the horizontal component of this matching will be an integer in $\{0, \dots, h-1\}$, with each choice equally likely. Thus, the expected horizontal component is $(h-1)/2$. This shows that, except for the at most $k_* - 1$ collisions creating the clusters in state $y(n_*)$, each collision contributes an amount $\Omega(1)$ to the expected sum of horizontal components in the final matching. From these observations and Lemma 2.2, we see that

$$E[U^{RM}(L)] = \frac{1}{k} \Omega(E[c_n]) , \quad (4.11)$$

where c_n is the total number of collisions in the interval $[0, n_*]$. The remainder of the proof shows that, under the condition (4.9), $E[c_n] = \Omega(\sqrt{nk^3})$. Then $E[U^{RM}(L)] = \Omega(\sqrt{nk})$ will follow from (4.11).

First, we prove a lower bound on the number of collisions of a single particle, say particle i , $i < k_*$, with its neighbor, particle $i+1$. Let $f_r^{(i)}(m)$ denote the number of collisions of particle i in a time interval of duration m , given that particles i and $i+1$ are at a distance $r \geq 0$ apart at the start of the interval.

Lemma 4.4. *For all i , $1 \leq i < k_*$, if $r = O(\sqrt{m})$, then*

$$E[f_r^{(i)}(m)] = \Omega(\sqrt{m}) .$$

We briefly postpone the proof of the lemma until after we have shown how it yields the estimate $E[c_n] = \Omega(\sqrt{nk^3})$. In the interval $[0, n_*]$ focus on columns $i, i+1$ for some i , $\frac{k_*}{3} < i < \frac{2k_*}{3}$. Partition the time interval $[0, n_*]$ into k^2 subintervals of length n_*/k^2 , and consider the expected number of collisions of particle i in any one of these subintervals, say one starting at time t . If $d_i(t) = y_{i+1}(t) - y_i(t) \leq \beta \sqrt{\frac{n}{k^3}}$ for some $\beta > 0$, then by Lemma 4.4 with $m = n_*/k^2$ ($n_* = n/(k-1)$), the expected number of collisions of particle i in the subinterval is $\Omega(\sqrt{\frac{n}{k^3}})$. But by the condition in (4.9), there is a positive probability that $d(t) \leq \alpha' \sqrt{\frac{n}{k}}$, for $\alpha' > \alpha$. It follows that, with positive probability, a constant fraction of the

$\frac{k}{6} + O(1)$ differences $d_i(t)$, $k_*/3 < i < 2k_*/3$, satisfy $d_i(t) \leq 6\alpha' \sqrt{\frac{n}{k^3}}$. Then the expected total number of collisions in the subinterval is $\Omega(k_* \sqrt{\frac{n}{k^3}})$, or $\Omega(\sqrt{\frac{n}{k}})$. There are k^2 subintervals, so the expected total number of collisions in $[0, n_*]$ is $E[c_n] = k^2 \Omega(\sqrt{\frac{n}{k}}) = \Omega(\sqrt{nk^3})$, as desired.

Proof of Lemma 4.4. We first prove the lemma for $k_* = 2$ ($k = 5$), and then use this result in a bounding argument for general $k_* > 2$ ($k > 5$ and odd). Thus, to start, we want a lower bound on the number of collisions of particle 1 during the interval $[t_0, t_0 + m]$ of a two-particle process, assuming that $y_2(t_0) - y_1(t_0) = r \geq 0$. Let $d(j) \geq 0$ denote the distance between the two particles just after the j^{th} jump following t_0 , and define $d(0) = r$ as the separation at time t_0 . Since the $+$ and $-$ processes are Poisson, the intervals between successive $+$'s are independent, exponentially distributed random variables with mean 1, as are the intervals between successive $-$'s. By the memoryless property of the exponential law, these properties also hold starting at time t_0 .

If $d(j) > 0$, then the $(j + 1)^{\text{st}}$ jump will be positive or negative with equal probability; it will be positive if particle 2 makes the jump and negative if particle 1 makes the jump. We have $d(j + 1) = d(j) + \xi$ or $d(j + 1) = (d(j) - \xi)^+$ according as the jump is positive or negative, respectively, where ξ is a time between consecutive $+$'s. Thus, ξ is exponentially distributed with mean 1 and independent of all previous jumps.

If $d(j) = 0$, then both particles jump; if particle 1 tries to jump farther than particle 2, then particle 1 collides with particle 2 and $d(j + 1) = 0$; otherwise, $d(j + 1) = d(j) + \delta$, where δ is distributed as the (positive) difference between two independent exponentially distributed jumps with mean 1. By the memoryless property of the exponential distribution, we see that, if $d(j) = 0$, then the events $d(j + 1) = 0$ and $d(j + 1) > 0$ are equally likely, and in the latter case, δ is exponentially distributed with mean 1 and independent of earlier jumps. Based on these observations, we can write

$$\begin{aligned} d(j) &= (d(j-1) + \xi_j)^+, \quad j = 1, 2, \dots \\ d(0) &= r, \end{aligned} \tag{4.12}$$

where the ξ_j 's are i.i.d. random variables with $E[\xi_j] = 0$ and with $|\xi_j|$ exponentially distributed with mean 1. Now the number of collisions of particle 1 is distributed as the number of transitions $d(j) \rightarrow d(j + 1) = 0$ in $[t_0, t_0 + m]$. The total number of transitions in $[t_0, t_0 + m]$ is bounded from below by the number of jumps of particle 2 in $[t_0, t_0 + m]$, which is Poisson distributed with mean m . Then the lemma (with $k_* = 2$) is an easy consequence of

Corollary 2.1. ■

Now consider a general $k_* > 2$ (k odd) and the motion of particles i and $i + 1$, $1 \leq i < k_*$, during $[t_0, t_0 + m]$, given that $y_{i+1}(t_0) - y_i(t_0) = r \geq 0$. To bound the number $f_r^{(i)}(m)$ of collisions of particle i , we will study a simple two-particle process. Imagine that the particles of an isolated two-particle process $\mathbf{z}(t) = (z_1(t), z_2(t))$ are activated at time t_0 ; particles 1 and 2 of \mathbf{z} move in parallel with particles i and $i + 1$ of \mathbf{y} according to the given sequences of $+$'s and $-$'s in columns i and $i + 1$. Particle 1 of \mathbf{z} starts out in the same position as particle i of \mathbf{y} . This is also true of particles 2 and $i + 1$ of \mathbf{z} and \mathbf{y} unless the latter is in a collision state with particle $i + 2$, in which case particle 2 starts out in the position of the next $+$ in column $i + 1$ below particle $i + 2$. The key observation is that, if t_1 is the time of the first collision of particle i after t_0 , then

$$z_1(t) \leq y_i(t), \quad z_2(t) \geq y_{i+1}(t), \quad t_0 \leq t \leq t_1. \quad (4.13)$$

To see the first inequality, note that $y_i(t) < y_{i+1}(t)$, $t_0 < t < t_1$, so $-$'s encountered in column i during $[t_0, t_1]$ are matched to $+$'s in column i . Each such $-$ causes both particle 1 of \mathbf{z} and particle i of \mathbf{y} to jump down to the next $+$. These are the only jumps of particle 1, so $z_1(t) \leq y_i(t)$, $t_0 \leq t \leq t_1$, with strict inequality when $+$'s in column i are matched to $-$'s to the left of column i according to the process \mathbf{y} .

To see the second inequality in (4.13), note that while $y_i(t) < y_{i+1}(t) < y_{i+2}(t) \leq \dots \leq y_{k_*}(t)$ holds, particle $i + 1$ jumps downwards from a $+$ only when that $+$ is matched to a $-$ encountered in column $i + 1$. Such $-$'s cause particle 2 to jump all the way down to the next $+$; particle $i = 1$ can also jump this far, but it may fall short if it collides with particle $i + 2$. We conclude that, since particle 2 starts out at a $+$ located at or below particle $i + 1$, it can not be passed by particle $i + 1$ in $[t_0, t_1]$.

By a standard coupling argument (see e.g. Ross [9], p. 155), (4.13) shows that, if the initial particle separation of \mathbf{z} is set to $r + \zeta$, where the random variable ζ is exponentially distributed with mean 1, then the time to the next collision of particle 1 in \mathbf{z} is stochastically at least as large as that of particle i in \mathbf{y} . The inter-collision intervals after t_1 begin with the separation state $r = 0$. Thus, extending the bounding process \mathbf{z} to any interval $[t_0, t_0 + m]$, we can define the following difference sequence for \mathbf{z} :

$$\begin{aligned} \tilde{d}(0) &= r + \zeta_0 \\ \tilde{d}(j) &= \zeta_j \text{ if } \tilde{d}(j-1) + \xi_j < 0 \quad j \geq 1 \end{aligned} \quad (4.14)$$

$$= \tilde{d}(j-1) + \xi_j, \text{ otherwise ,}$$

where the ξ_j are as in (4.12) and the ζ_j , $j \geq 0$, are independent and exponentially distributed with mean 1. Here, a ζ_j corresponds to an initial move of particle 2 to the next + so as to guarantee (4.13). Note that $\{d(j)\}$ and $\{\tilde{d}(j)\}$ differ only in their behavior near the origin; $\{\tilde{d}(j)\}$ has a reflecting barrier there, whereas $\{d(j)\}$ has the elastic barrier of the Lindley process. By our earlier observations, the number $f_r^{(i)}(m)$ of collisions of particle i in $[t_0, t_0 + m]$ with $d_i(t_0) = r$ is stochastically at least as large as the number $\tilde{f}_r(m)$ of times $\{\tilde{d}(j)\}$ reflects at the origin in $[t_0, t_0 + m]$, with $\tilde{d}(0) = r + \zeta_0$. Thus,

$$E[f_r^{(i)}(m)] \geq E[\tilde{f}_r(m)] , \quad (4.15)$$

and it remains to bound $E[\tilde{f}_r(m)]$.

It is easily verified that, if the sample paths of $\{\tilde{d}(j)\}$ and $\{d(j)\}$ are constructed from the same samples ξ_j , $j \geq 1$, then for each j , $\tilde{d}(j) - d(j)$ is nonnegative and bounded by the distance that $\{\tilde{d}(j)\}$ reflected from the origin on its most recent reflection there. It follows that $\tilde{d}(j)$ is stochastically no larger than $d(j) + \zeta$, where ζ is exponentially distributed with mean 1. As an easy consequence of this fact, the expected number of reflections of $\{\tilde{d}(j)\}$ in $[t_0, t_0 + m]$ is a constant fraction of the expected time spent by $\{d(j)\}$ at the origin during $[t_0, t_0 + m]$. Then $E[\tilde{f}_r(m)] = \Omega(\sqrt{m})$ by the result for $k_* = 2$, and the lemma is proved by (4.15). ■

This completes the proof of part (i) of Theorem 4.1. We now prove part (ii), where $k = \Omega(n^{1/3})$ is assumed; this case is much simpler. The proof for the special case $k = \Theta(n^{1/3})$ is already in hand, so assume that k grows strictly faster than $n^{1/3}$. We adapt the technique of Theorem 3.2 and Corollary 3.1 as follows.

Consider a two-dimensional representation for large k and n . For convenience, assume that k_* is a multiple of $n^{1/3}$. This assumption is not essential; the argument below is easily modified to handle general values of k_* . Partition the columns into $n^{1/3}$ equal-size groups of consecutive columns. Insert $n^{1/3} - 1$ new columns between the groups and place one new column just to the left of the first group and one to the right of the last group. Next, as in the proof of Theorem 3.2, shift +'s to the left and -'s to the right, stopping in each case at the nearest *new* column. Focusing now on all but the first and last of the new columns, we have a random instance for a new number of columns $\hat{k}_* = \frac{k_*}{n^{1/3}} - 1$.

Construct the MFF matching M for this new, reduced instance, then shift the +'s and -'s right and left back to their original position, extending the edges of M so as to keep the same

pairs of points matched. In analogy with Theorem 3.2, we obtain the desired result from the following three observations: (i) By the proof of the theorem for $k = O(n^{1/3})$, the expected number of points left unmatched by M is at least $\Omega(n^{2/3})$, (ii) the first and last new columns which were excluded from the reduced instance have on the order of $n^{2/3}$ points, and (iii) the weak FF property of M is preserved in the shift of $+$'s and $-$'s back to their original positions.

Note that this technique also yields a proof for the $\Omega(n^{2/3})$ lower bound for the continuous case ($k \rightarrow \infty$), originally proved in [10]. ■

5. Final Remarks

The techniques of this paper, in particular the reductions to matching problems, can also be applied to proofs of asymptotic bounds for symmetric distributions on $\{1, \dots, k-1\}$. It is not difficult to show that the symmetry of a given distribution guarantees the \sqrt{n} dependence of both upper and lower bounds, except for the trivial case where all item sizes are $1/2$. These bounds will also depend on shape parameters and how they vary with k or n .

Very few results exist on FF bin packing under more general distributions, discrete or continuous. For example, consider a uniform distribution on $\{1, \dots, j\}$, with a bin size $k \geq j+2$. It is known that if j is sufficiently small relative to k (roughly at most \sqrt{k}), then $E[FF(L)] = O(1)$ (see [1]). Simulations give convincing evidence that for many j , $\sqrt{k} < j < k-2$, the expected wasted space grows linearly in n . However, the proof of this result for any such j remains an intriguing open problem. As shown recently in [6], results of this type *do* exist for best-fit bin packing.

References

- [1] 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 23rd Annual ACM Symposium on Theory of Computing*, pages 230–240, New York, 1991. ACM Press.
- [2] E. G. Coffman, Jr., C. Courcoubetis, M. R. Garey, D. S. Johnson, P. W. Shor, R. R. Weber, and M. Yannakakis. Bin packing with discrete item sizes, Part I: Perfect packing theorems and the average case behavior of optimal packings. (In preparation).

- [3] E. G. Coffman, Jr., D. S. Johnson, P. W. Shor, and R. R. Weber. Bin packing with discrete item sizes, Part III: Tight bounds on Best Fit. (In preparation).
- [4] E. G. Coffman, Jr., D. S. Johnson, P. W. Shor, and R. R. Weber. Bin packing with discrete item sizes, Part V: 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 IV: Average case behavior of FFD and BFD. (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 25th Annual ACM Symposium on Theory of Computing*, pages 412–421, New York, 1993. ACM Press.
- [7] W. Feller. *An Introduction to Probability Theory and Its Applications, Volume II*. Wiley & Sons, New York, New York, 1966.
- [8] C. Kenyon, Y. Rabani, and A. Sinclair. Biased random walks, Lyapunov functions, and stochastic analysis of best fit bin packing. In *Proc. Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 351–358, Philadelphia, 1996. Society for Industrial and Applied Mathematics.
- [9] Sheldon M. Ross. *Introduction to Stochastic Dynamic Programming*. Academic Press, New York, 1983.
- [10] P. W. Shor. The average case analysis of some on-line algorithms for bin packing. *Combinatorica*, 6:179–200, 1986.
- [11] P. W. Shor. How to pack better than Best Fit: Tight bounds for average-case on-line bin packing. In *32nd Annual Symposium on Foundations of Computer Science*, pages 752–759, New York, 1991. IEEE Computer Society Press.