Packing Rectangles in a Strip

E. G. Coffman, Jr.
Electrical Engineering Department
Columbia University
New York, NY 10027

Peter J. Downey
Department of Computer Science
University of Arizona
Tucson, AZ 85721

Peter Winkler
Bell Labs, Lucent Technologies
Murray Hill, NJ 07974

Abstract Rectangles with dimensions independently chosen from a uniform distribution on [0,1] are packed on-line into a unit-width strip under a constraint like that of the Tetris^{TM} game: rectangles arrive from the top and must be moved inside the strip to reach their place; once placed, they cannot be moved again. Cargo loading applications impose similar constraints. This paper assumes that rectangles must be moved without rotation. For n rectangles, the resulting packing height is shown to have an asymptotic expected value of at least $(0.31382733\ldots)n$ under any on-line packing algorithm. An on-line algorithm is presented that achieves an asymptotic expected height of $(0.36976421\ldots)n$. This algorithm improves the bound achieved in *Next Fit Level* (NFL) packing, by compressing the items packed on two successive levels of an NFL packing via on-line movement admissible under the Tetris constraint.

1 Introduction

A variant of the classical bin-packing problem is two-dimensional strip packing, in which rectangles of width and height bounded by 1 are packed into a semi-infinite strip of width 1, imagined to form a vertical bin. Packings must be such that no rectangle overlaps another's area and the sides of the rectangles are parallel to the strip sides. The objective is to minimize the height of the packing in the strip, for a given sequence of n rectangles[1] The optimum packing is elusive and technically intractable, so most attention focuses on heuristic packing algorithms. Deterministic or probabilistic analysis can assess how closely such

[1] The term two-dimensional *bin packing* is often reserved for the variant of this problem in which horizontal boundaries are placed at integer strip heights, and rectangles are forbidden to overlap these "bin" boundaries. This paper studies the less restrictive strip packing problem. Many authors refer to the strip as a bin.

algorithms approach the optimum in either an absolute or ratio sense.

In this paper, we study the strip-packing problem under three constraints on the allowable packing algorithm

- Rotation of rectangles (also referred to as *items*) during placement is not allowed.
- The packing algorithm must be *on-line*: the algorithm must inspect rectangles one at a time, make a placement decision for each rectangle at the time it is inspected, and must not later renege on the placement even in light of subsequent information.
- The packing algorithm must obey a *Tetris*-like constraint [1] (Tetris is a registered trademark of The Tetris Company[2]): rectangles descend from the top of the strip; they must be moved within the strip horizontally and vertically to reach their final placement; and they may not overlap the area of any other rectangle during this movement.

In the original formulation of the strip-packing problem [4], an on-line packing algorithm was allowed to consider rectangles as being initially outside of the two-dimensional strip; rectangles could then be placed wherever they fit. The Tetris constraint removes many such space utilization opportunities, when algorithms must operate on-line. See Figure 1 for an example.

Constraints similar to the Tetris packing restriction arise in warehousing and cargo container loading applications, where the objective is to minimize wasted storage area. The on-line Tetris constraint corresponds to requiring that each item (e.g., box) be stored before a new item arrives, and that physical access to each item's storage location exists at the time of its arrival. A requirement to unload in reverse order of on-line arrival also imposes the Tetris constraint.

In This Paper: The objective of the strip packing problem is to pack a list L_n of n rectangles in such a way as to minimize the height of the packing. After all rectangles in L_n have been placed, the *height* of the packing is the maximum distance from the strip bottom to the top of any packed rectangle.

The packing problem can be studied under deterministic or probabilistic assumptions about the rectangle widths and heights. Here we examine the probabilistic model in which the n given rectangle widths and heights (W_i, H_i), $i = 1, 2, \dots n$, are $2n$ independent draws from a uniform distribution on $[0, 1]$, which we denote in the usual way by $U[0,1]$. We are interested in the *expected* height of the packing, asymptotically for large n. For this model it is shown that

[2]Tetris has been one of the world's most popular video games; it was created by Alexey Pajitnov in 1985. The allusion to the Tetris game in strip-packing comes from the requirement that rectangles drop from the top of the strip and avoid already packed rectangles. However, we should note that the objective of strip-packing – height minimization – is different from the Tetris game's objective.

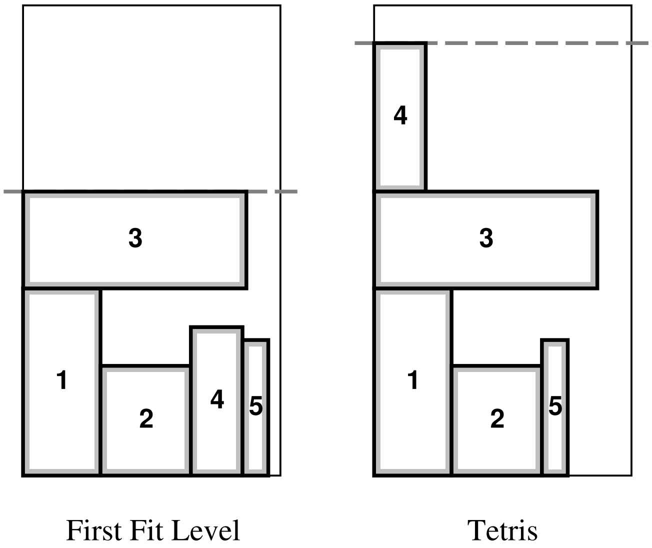


Figure 1: If the items depicted arrive on-line in the order $1, 2, \ldots$, they can be packed on-line using a *First-Fit Level* heuristic [14, 8] as shown at left. Under the Tetris constraint, a best possible packing is shown at right.

- No on-line algorithm[3] operating under the Tetris constraint can achieve a lower expected height, asymptotically, than $(2\ln 2 - \pi^2/12 - 1/4)n = (0.313827\ldots)n$.
- An algorithm exists (a variation of the *Next Fit Level* (NFL) algorithm [14]) that achieves the expected height of $(0.369764\ldots)n$. Since NFL is an on-line algorithm that obeys the Tetris constraint, this result improves the NFL bound of $(0.381338\ldots)n$.

Problem History: The problem of two-dimensional strip packing has had a long history of development since [4]; for a comprehensive overview of the subject up to the late '80s, see [9]. The taxonomy of strip packing divides naturally into the study of off-line algorithms (regarding which a large literature exists) and on-line algorithms (about which a good deal less is known). This paper considers on-line algorithms, and specifically those related to level algorithms, defined below. Within this class, we are interested in the effect of the Tetris constraint on packing density, and on average-case analysis of the problem. We will confine this brief review to the on-line strip packing literature, beginning

[3] A distinction can be made between packing algorithms that enforce "gravity" and those that do not. If gravity is in effect, rectangles must be placed with their bottoms abutting the top of a previously packed rectangle, or the bottom of the strip. Our lower bound applies whether or not gravity is in effect, and our compression algorithm (which does not enforce gravity) provides an upper bound in both cases.

with deterministic results and concluding with probabilistic results.

In the deterministic literature regarding approximation algorithms, the quality of algorithms is most commonly measured by the *asymptotic worst case ratio* R_A^∞ of the algorithm, defined as follows. Let $A(L)$ be the height used by algorithm A in packing the rectangles of list L, and let $OPT(L)$ denote the minimal height used by any algorithm of the class under study (e.g., on-line, on-line with lists sorted by height, etc.). Then the asymptotic performance ratio is:

$$R_A^\infty = \overline{lim}_{n\to\infty} \sup_{|L|=n} \frac{A(L)}{OPT(L)}$$

The asymptotic ratio is an appropriate measure of performance when the number of rectangles is large; when n is small, the *absolute worst-case ratio* $R_A = \sup_L \frac{A(L)}{OPT(L)}$ is used.

Lower bounds for on-line strip packing algorithms were first explored by Brown, Baker, and Katseff [3], who proved a number of lower bounds on the absolute worst-case ratio and showed, using a bin-packing result of Liang [16], that any on-line algorithm A for strip packing must obey $R_A^\infty \geq 1.536\ldots$. Van Vliet [17] later improved this to $R_A^\infty \geq 1.540\ldots$. These lower bounds hold even if algorithms are allowed to pre-sort rectangles by width or height. In [5], Baker and Schwarz introduced on-line *shelf* algorithms, which allocate levels of different heights into which items can be packed. The possible strip (shelf) heights are fixed in advance, in contrast to the *level* algorithms described later in this section. Baker and Schwarz give an algorithm *First Fit Shelf* having an asymptotic performance ratio arbitrarily close to 1.7. Csirik and Woeginger [6] showed that any on-line shelf algorithm A must obey $R_A^\infty \geq 1.691\ldots$, and that the *Harmonic shelf* algorithm has an asymptotic performance ratio arbitrarily close to the lower bound.

However, the above shelf algorithms do not adhere to a Tetris constraint, which imposes further restrictions upon on-line placement of rectangles. A deterministic analysis of packings satisfying the Tetris constraint was carried out by Azar and Epstein [1], who found that, for widths W_i chosen from $[\epsilon, 1]$ or from $[0, 1-\epsilon]$, and arbitrary rectangle heights H_i (not limited to [0,1]), there is a lower bound of $\Omega(\log^{1/2}(1/\epsilon))$ on the asymptotic performance ratio for any deterministic or randomized algorithm. This rules out any competitive algorithm for the deterministic model. They present a level algorithm that achieves an asymptotic performance ratio of $O(\log(1/\epsilon)$, for widths chosen as above. If rectangles have both their widths and heights bounded by 1, and can be rotated so that their smaller dimension becomes the width, then there is a shelf algorithm with performance ratio 4 relative to the optimum [1].

Moving now to stochastic models of two-dimensional strip packing, we are given a list L_n of n items (or rectangles) (W_i, H_i), with widths and heights sampled from populations with a given distribution function. The objective criterion is now $\mathbf{E}A(L_n)$, the expected height of the packing after all n items have been packed by algorithm A. For stochastic models, performance of an algorithm is commonly assessed by giving the absolute value $\mathbf{E}A(L_n)$ as a function of n (exactly or asymptotically), rather than in terms of performance ratios. Progress in probabilistic strip packing has been made under the assumption that the W_i and H_i are $2n$ independent $U[0,1]$ random variables.

Under this uniform model, a lower bound for any packing – on or off-line – can easily be derived. The height of the packing must be at least as large as the total area of the rectangles, so that $\mathbf{E}OPT(L_n) \geq n/4$. It must also be as large as the total height of rectangles with widths exceeding $1/2$; use of this latter fact improves the lower bound [8] to $\mathbf{E}OPT(L_n) = n/4 + \Omega(n^{1/2})$. It is interesting to compare this absolute lower bound with that for algorithms having a Tetris constraint, derived in Theorem 1 below.

If the packing algorithm is allowed to know the number of items to be packed and the distribution of item heights, then a shelf algorithm can be used to pack the items on-line. The algorithm *Best Fit Shelf* [10] operates on-line and achieves the result $\mathbf{E}BFS(L_n) \sim n/4$. It does not adhere to the Tetris constraint.

The *level* algorithms pack rectangles into horizontal strips, or levels, where level 1 is the bottom of the strip and level i is a horizontal line drawn through the top of the tallest rectangle on level $i-1$. Level algorithms are especially appropriate when there is no information available *a priori* about the rectangles to be packed, and the packing must be on-line. The *Next Fit Level* (NFL) algorithm was analyzed by in [14], and is most relevant to our results because (a) it is on-line, (b) it obeys the Tetris constraint, and (c) it forms the starting point for our improved bound. We will therefore take some time to describe the NFL algorithm and its analysis.

In packing L_n, NFL starts out by placing items left justified along level 1 (the bottom of the strip). If an item I_i is encountered that is too wide to fit in the remaining space on a level j, then I_i is placed left justified on a horizontal baseline drawn through the top of the tallest rectangle on level j, thus opening a new level $j+1$. If all items are packed by level m, then the height of the packing is at the height reached by the tallest item of that level.

Analysis of NFL uses results on one-dimensional bin packing originally proved for the *Next Fit* (NF) algorithm [11, 15]. The process by which NFL fills successive *levels* is identical to the process by which NF fills successive *bins*. Consider the sequence $W_1, W_2, \ldots$ of item widths packed level by level into the strip of unit width. If Y_i is the width used by all items at the ith level, then the Y_i form a continuous-state Markov process [11] with transition kernel

$$K(x,y) = \mathbf{P}[Y_{i+1} \leq y | Y_i = x] = 1 - \frac{1-y}{x} e^{x-(1-y)}, \quad 1-x < y <\leq 1,$$

and zero elsewhere. This kernel can be readily calculated from the uniform assumptions made on the W_i. The process $\{Y_i\}$ is ergodic [11] and the distributions of the Y_i converge at a geometric rate to an equilibrium distribution function (d.f.) $F_Y(y)$. If we imagine the NFL algorithm proceeding to equilibrium (large n), it is established in [11] that the process tends toward stable behavior characterized by three equilibrium random variables involving the widths of items (see Figure 2).

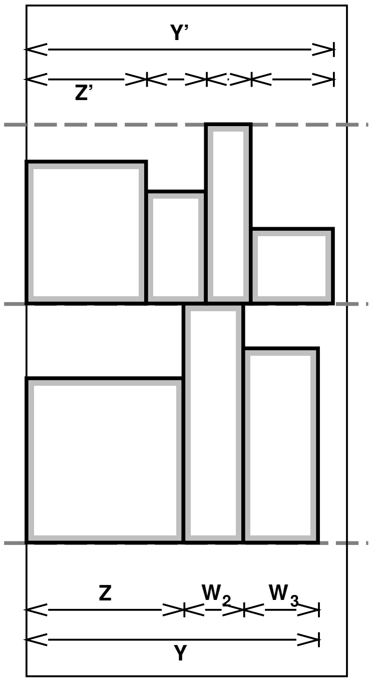


Figure 2: The *Next Fit Level* (NFL) packing process at equilibrium, showing two adjacent levels. The leading item width Z is distributed as in equation (3). The level width Y is distributed as in equation (1). The remaining widths W_i are $U[0, 1]$.

1. The *equilibrium level width* Y is the amount of the available unit width that is utilized in each level, at equilibrium. It has the following d.f. on [0,1]:
$$\mathbf{P}[Y \leq y] = F_Y(y) = y^3 \tag{1}$$

2. The *equilibrium level number* M is the number of items that will fit in each level, at equilibrium. It has the distribution
$$\mathbf{P}[M = m] = p_M(m) = 3\frac{m^2 + 3m + 1}{(m+3)!} \tag{2}$$

3. The first item packed at each level has a width that is conditioned by the fact that it failed to fit in the space $1 - Y$ remaining on the level below. It is no longer a uniform random variable, but is positively width-biased with expected value $5/8$. This *equilibrium leading item width* Z has the following d.f. on [0,1]:
$$\mathbf{P}[Z \leq z] = F_Z(z) = \frac{3}{2}\left(z^2 - \frac{z^3}{3}\right) \tag{3}$$

From (1) we have $\mathbf{E}Y = 3/4$, and we know that the sum of the n rectangle widths has expectation $n/2$; from these facts it follows that [11, 15]
$$\mathbf{E}[\text{number of levels}] = \frac{2n}{3}. \tag{4}$$

Based on the above process involving only the sequence of widths $W_1, \ldots, W_n$, the independent heights H_i can be used to obtain the height of a level at equilibrium. It is clearly the extremum
$$H_{(M)} = \max(H_1, \ldots, H_M)$$
where the H_i are independent $U[0,1]$ variates and M is the (random) number of items on a level. Conditioned on $M = m$, the expectation of such a uniform maximum is $m/(m+1)$, and so using (2) one obtains the *equilibrium level height* in terms of the exponential integral [2] and Euler's constant, making it readily calculable:
$$\begin{aligned} \mathbf{E}H_{(M)} &= \sum_{m \geq 1} p_M(m) \\ &= \frac{27}{4} - 3e + \frac{3}{2}(\mathrm{Ei}(1) - \gamma) = 0.5720077418\ldots \end{aligned} \tag{5}$$

Finally, by (4) and (5), the expected height of the NFL packing is
$$\begin{aligned} \mathbf{E}[NFL(L_n)] &\sim \frac{2n}{3}(0.5720077418\ldots) \\ &\sim (0.3813384945\ldots)n \quad (n \to \infty) \end{aligned} \tag{6}$$

2 The Lower Bound

For each list L_n of rectangles, there is some on-line Tetris packing (sequence of rectangle placements) that produces the minimal height among all on-line packings of L_n obeying the Tetris constraint. Let OPT denote an algorithm that, for each L_n, produces one such minimal height packing. Let $\mathbf{E}OPT(L_n)$ be the expected height produced by OPT in packing the n rectangles of L_n under the uniform model. Then the *packing constant*, i.e., the asymptotic expected height of the optimal packing per additional item, is given by

$$\gamma = \underline{lim}_{n\to\infty} \frac{\mathbf{E}OPT(L_n)}{n}$$

Theorem 1 *The packing constant has the lower bound*

$$\gamma \geq 2\ln 2 - \frac{\pi^2}{12} - \frac{1}{4} = 0.31382733\ldots \tag{7}$$

Proof: Consider the subsequence of rectangles in L_n having widths greater than $1/2$, and let I_1 and I_2 be two successive rectangles in the subsequence, having widths x and y, respectively. Let C be the possibly empty collection of rectangles in the full sequence L_n between I_1 and I_2, and having a width exceeding $1 - \min(x, y)$. No rectangle in C can be moved past either of I_1 and I_2 under the Tetris constraint. Let H be the height of the tallest rectangle in C, with $H = 0$ if C is empty. There is an average of $n/2$ rectangles wider than $1/2$ and the average height of each such rectangle is $1/2$, so for n large enough, $\mathbf{E}OPT(L_n)$ is at least $(1/2 + \mathbf{E}[H])n/2$, i.e.,

$$\gamma \geq 1/4 + \mathbf{E}H/2.$$

Another look at (7) shows that it is now sufficient to prove that $\mathbf{E}H = 4\ln 2 - 1 - \pi^2/6$.

Define $z = \min(x, y)$ so that all rectangles in C have widths exceeding $1 - z$. To find $\mathbf{P}[|C| = k]$, focus on the subsequence of rectangles with widths in $[1-z, 1]$. We identify the event $[|C| = k]$ with the occurrence in this subsequence of exactly k rectangles with widths in $[1 - z, 1/2]$ from one rectangle wider than $1/2$ to the next such rectangle, i.e.,

$$\mathbf{P}[|C| = k] = p^k(1 - p), \quad k = 0, 1, \ldots$$

where $p = [1/2 - (1 - z)]/z = 1 - 1/(2z)$ is the conditional probability that a width is at most $1/2$ given that it exceeds $1 - z$.

The tallest of $|C| = k$ rectangles has expected height $k/(k + 1)$, so the expected height of the tallest rectangle in C is

$$\mathbf{E}H_{xy} \quad = \quad \sum_{k\geq 0} \frac{k}{k+1} p^k(1-p) = 1 - \sum_{k\geq 0} \frac{1}{k+1} p^k(1-p)$$

$$\begin{aligned}
&= 1 - \frac{1-p}{p} \sum_{k \geq 0} \frac{p^{k+1}}{k+1} = 1 - \frac{1-p}{p} \ln \frac{1}{1-p} \\
&= 1 - \frac{1}{2z-1} \ln 2z
\end{aligned}$$

Without loss of generality, assume that $x \leq y$ ($z = x$), so that to obtain $\mathbf{E}H$ we need to average $\mathbf{E}H_{xy}$ over the (conditional) uniform distribution in the triangular region $1/2 \leq x \leq y \leq 1$. The density in this triangle (with area $1/8$ of that of the unit square) is 8, since we are conditioning on an outcome occurring there. So

$$\begin{aligned}
\mathbf{E}H &= 8 \int_{1/2}^{1} \int_{x}^{1} \left[1 - \frac{\ln 2x}{2x-1}\right] dy dx = 8 \int_{1/2}^{1} (1-x) \left[1 - \frac{\ln 2x}{2x-1}\right] dx \\
&= 2 \int_{0}^{1} (1-v) \left[1 - \frac{\ln(v+1)}{v}\right] dv \quad (v = 2x-1) \\
&= 2 \int_{0}^{1} \left[1 - \frac{\ln(v+1)}{v} - v + \ln(v+1)\right] dv \\
&= 4 \ln 2 - 1 - \frac{\pi^2}{6} \quad \text{(see e.g., [13], p. 555.)}
\end{aligned}$$

3 The Compression Algorithm

This section describes an improvement to the NFL algorithm which we call the *Compression* algorithm and denote by CA. The algorithm is motivated by the observation that a large proportion of levels (over 70%) in the NFL packings at equilibrium contain only one or two items. In any two adjacent levels, it is likely that some of the items from the upper level can be dropped down and repacked into holes in the lower level. Of course, confining attention to levels with at most two items also makes the analysis of the algorithm tractable. As a convenience, we shall often refer to the compression algorithm as "repacking" rectangles in other packings. Bear in mind, however, that no off-line repacking is actually being done: The algorithm achieves the repacking on-line by simple modifications to the steps taken in the NFL packing, as we will see shortly.

The repacking and compression of adjacent levels of the NFL packing described above must first be shown to be possible by an on-line algorithm. In describing how CA operates on-line, it will be easier to describe first a variant of NFL, which we call *Bi-level* NFL, that has identical probabilistic properties at equilibrium.

Bi-level NFL (BNFL), like NFL, packs items into levels, opening new packing levels at the high point of the old level as soon as a rectangle is encountered that will not fit. As its name implies, it packs two levels at a time. We will call the first of the two levels the lower level, and the second the upper level. After packing two levels, BNFL repeats the pattern below.

BNFL Algorithm:

- *Lower Level:* When a new bi-level is opened, BNFL places the first rectangle I_1 left-justified in a newly opened level (the lower level). If the next rectangle I_2 does not fit into the lower level, BNFL proceeds to the *Upper Level* stage described below.

  If I_2 fits into the lower level, BNFL *right*-justifies it there. As long as subsequent items will fit on the lower level, BNFL packs them right justified against the last item packed. As soon as an item is encountered whose width will not fit into the lower level, the packing proceeds to the Upper Level stage below.

- *Upper Level:* Let I_j, $j \geq 2$, be the first item that will not fit into the lower level. If $j = 2$, so there is only one item in the lower level, then BNFL packs I_2 left justified above I_1.

  If $j \geq 3$, BNFL packs I_j *above the shorter of items* I_1 *and* I_2 and justified against the nearer strip wall.

  Each of the items $I_{j+1}, \ldots$ is packed on the upper level, justified against the last item packed there, as long as it fits.

  As soon as an item I_k is encountered that will not fit into the upper level, a new bi-level is opened, and BNFL repeats its packing cycle, beginning with I_k.

Observe that BNFL is on-line, has the same items in the same levels as NFL, and both utilize the same amount of width (Y) at each level at equilibrium. The first item packed at each level is identical in BNFL and NFL (although in BNFL that first item may be right justified if it occurs in an upper level of a bi-level). Therefore, the equilibrium processes described for NFL also apply to BNFL. Also note that the expected height of a bi-level in BNFL, at equilibrium, will be twice the expected height of an NFL level, as given in (5).

We are now ready to describe the CA algorithm. CA proceeds as BNFL, but takes advantage of certain patterns of items in a bi-level, repacking these patterns to reduce the height of the bi-level. In order to make the analysis tractable, only certain bi-level patterns containing exactly 3 or 4 items will be repacked by CA[4].

CA Algorithm:

- *Lower Level:* The algorithm proceeds to pack the lower level exactly as for BNFL. If only one item fits at the lower level, it is left justified. If there are at least two items that fit at the lower level, CA leaves the lower level with I_1 left justified and I_2 right justified.

[4]Bi-levels with exactly 2 items cannot be improved by CA

- *Upper Level:* Let I_j, $j \geq 2$, be the first item that will not fit into the lower level.

  If $j = 2$, so there is only one item packed in the lower level, CA left justifies I_2. If I_3 fits in the upper level, CA right justifies it, and if I_3 can be slid down to the bottom of the lower level by a legal Tetris move, CA does so. Otherwise, CA continues exactly as for BNFL. Items $I_4, \ldots$ that fit in the upper level are packed exactly as for BNFL.

  If $j > 3$, then CA operates exactly as BNFL in the upper level.

  As soon as an item I_k is encountered that will not fit into the upper level, a new bi-level is opened, and CA repeats its packing cycle, beginning with I_k.

As for BNFL, the distribution of widths of items packed by CA is exactly the same as for NFL. The only change is that with positive probability, I_3 and possibly I_4 may be slid downward in the CA packing. This will result in a certain overall expected compression in the height of a bi-level that will be analyzed in section 4 below.

There are three patterns of items in a bi-level that will lead to a positive expectation of height compression under CA. (There may be more such patterns, but we have limited ourselves to these three for reasons of tractability and because they are the patterns with highest probability.) These three patterns are shown in Figure 3, and consist of numbers of items in the lower and upper levels corresponding to 2–2, 2–1, and 1–2. The figure also suggests how CA can realize compression, provided the widths of individual items are amenable.

4 Upper Bound Analysis

In this section, we take in turn each of the patterns of Figure 3, calculate probabilities of occurrence, and compute the expected compression in the height of a bi-level that ensues. At the end of the section, we accumulate these expected rewards and compute the asymptotic expected height of the packing under CA.

4.1 Two Items in the Lower and Two Items in the Upper Level

The four events whose probability and expected compression we wish to calculate are depicted in Figure 4. In each case, one must imagine another item of width W_5, marking the beginning of the next bi-level, and placed left justified on a line through the highest point of the bi-levels shown.

Although case 1a and 2a are symmetric (and likewise cases 1b and 2b), their probabilities differ. This is because at equilibrium the distribution of the width Z_1 of the first item packed is not uniform; it is given in (3). Variates with the d.f. (3) will be denoted generically by Z in the sequel, and variates with the $U[0,1]$ distribution will be denoted by W.

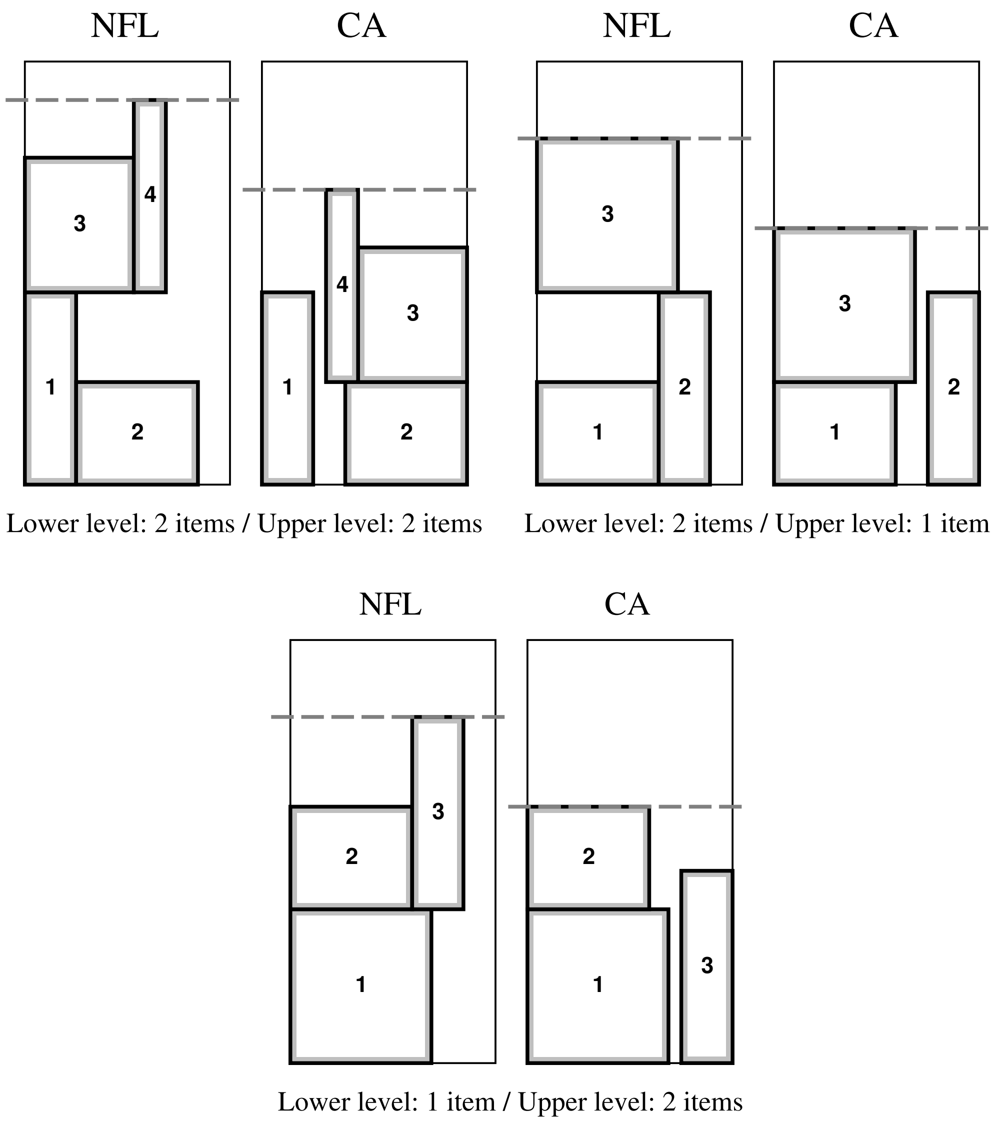


Figure 3: Bi-level patterns in NFL packings that will lead to a positive expectation of compression under the CA algorithm. These bi-level events are disjoint at equilibrium.

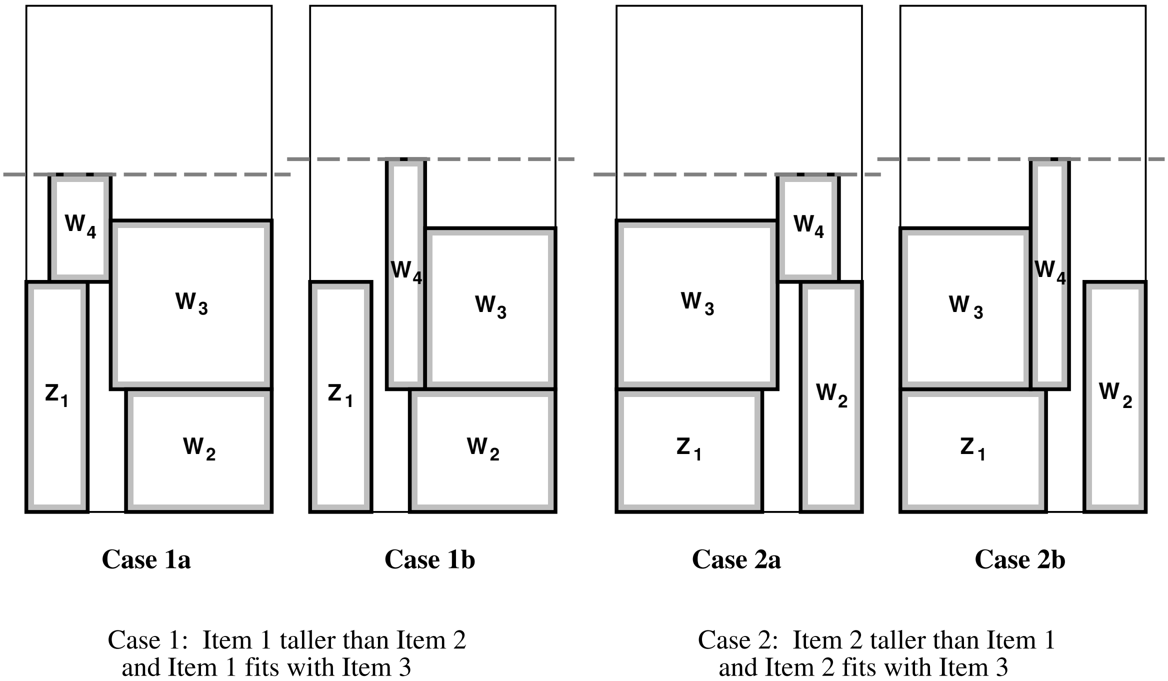


Figure 4: Two items on the lower level and two items on the upper level, shown as compressed by CA. The captions summarize the conditioning on the widths of items, and on the relative heights of items 1 and 2 that allow compression to take place. Not shown is an item of width W_5 that marks the start of the next bi-level above the one depicted.

4.1.1 Event Probabilities

Variates Z have the density function $f_Z(z) = (3/2)(2z - z^2)$ on [0,1], and variates W have the density function $f_W(w) = 1$ on [0,1].

In all four cases, we have the following conditions, which together form an event to be called A:

$$\begin{array}{rcll} Z_1 + W_2 & \leq & 1: & I_1 \text{ and } I_2 \text{ fit on the lower level} \\ Z_1 + W_2 + W_3 & > & 1: & I_3 \text{ does not fit on the lower level} \\ W_3 + W_4 & \leq & 1: & I_3 \text{ and } I_4 \text{ fit on the upper level} \\ W_3 + W_4 + W_5 & > & 1: & I_3 \text{ and } I_4 \text{ are alone on the upper level} \end{array}$$

Case 1a: In addition to event A, we have the conditions that I_1 and I_3 can fit together across the strip, but that I_4 will not fit with them. We also have that the (independent) heights of the first two items make I_1 the taller: [5]

$$\begin{aligned} \mathbf{P}[\text{Case 1a}] &= \mathbf{P}[H_2 < H_1] \cdot \mathbf{P}[A, Z_1 + W_3 \leq 1, Z_1 + W_3 + W_4 > 1] \qquad (8) \\ &= \frac{1}{2} \int_0^1 f_Z(z_1) dz_1 \int_0^{1-z_1} dw_2 \int_{1-z_1-w_2}^{1-z_1} dw_3 \int_{1-z_1-w_3}^{1-w_3} dw_4 \int_{1-w_3-w_4}^{1} dw_5 \\ &= \frac{1}{2} \frac{1}{35} = \frac{1}{70} \end{aligned}$$

Case 1b: In addition to event A, we have the conditions that I_1 and I_3 can fit together across the strip, and that I_4 will also fit with them[6]. We also have that I_1 is taller than I_2:

$$\begin{aligned} \mathbf{P}[\text{Case 1b}] &= \mathbf{P}[H_2 < H_1] \cdot \mathbf{P}[A, Z_1 + W_3 \leq 1, Z_1 + W_3 + W_4 \leq 1] \qquad (9) \\ &= \frac{1}{2} \int_0^1 f_Z(z_1) dz_1 \int_0^{1-z_1} dw_2 \int_{1-z_1-w_2}^{1-z_1} dw_3 \int_{0}^{1-z_1-w_3} dw_4 \int_{1-w_3-w_4}^{1} dw_5 \\ &= \frac{1}{2} \frac{3}{280} = \frac{3}{560} \end{aligned}$$

Case 2a: In addition to event A, we have the condition that I_2 and I_3 can fit together across the strip, but that I_4 will not fit with them. In this case, we have that item I_2 is taller than I_1.

$$\begin{aligned} \mathbf{P}[\text{Case 2a}] &= \mathbf{P}[H_2 \geq H_1] \cdot \mathbf{P}[A, W_2 + W_3 \leq 1, W_2 + W_3 + W_4 > 1] \qquad (10) \\ &= \frac{1}{2} \int_0^1 f_Z(z_1) dz_1 \int_0^{1-z_1} dw_2 \int_{1-z_1-w_2}^{1-z_1} dw_3 \int_{1-w_2-w_3}^{1-w_3} dw_4 \int_{1-w_3-w_4}^{1} dw_5 \\ &= \frac{1}{2} \frac{13}{420} = \frac{13}{840} \end{aligned}$$

[5] All of the multiple integrals of this section have been checked by *Mathematica* [18].

[6] No use is made of the potential gap between items 1 and 2. A similar remark will apply to Case 2b.

Case 2b: In addition to event A, we have the condition that I_2 and I_3 can fit together across the strip, and that I_4 will also fit with them. Again, we have that item I_2 is taller than I_1.

$$\begin{aligned}
\mathbf{P}[\text{Case 2b}] &= \mathbf{P}[H_2 \geq H_1] \cdot \mathbf{P}[A, W_2 + W_3 \leq 1, W_2 + W_3 + W_4 \leq 1] \qquad (11) \\
&= \frac{1}{2} \int_0^1 f_Z(z_1) dz_1 \int_0^{1-z_1} dw_2 \int_{1-z_1-w_2}^{1-z_1} dw_3 \int_0^{1-w_2-w_3} dw_4 \int_{1-w_3-w_4}^{1} dw_5 \\
&= \frac{1}{2} \frac{13}{420} = \frac{13}{840}
\end{aligned}$$

4.1.2 Expected Compressions

The compression achieved in the four cases depends entirely upon the independent $U[0,1]$ heights $H_i,\ i = 1,2,3,4$. Compression is better in the "b" cases, where items I_3 and I_4 can be treated as a unit. See Figure 5.

Lemma 1 *In cases 1a and 2a, provided $H_3 > H_4$, the compression is $\Delta_a = \min(R_{12}, R_{34})$, where the R_{ij} are independent 2-sample ranges:*

$$R_{12} = \max(H_1, H_2) - \min(H_1, H_2), \qquad R_{34} = \max(H_3, H_4) - \min(H_3, H_4)$$

Thus,

$$\mathbf{E}\Delta_a = \frac{1}{10} \qquad (12)$$

Proof: The depth of the "hole" into which I_3 can be slid is R_{12}. However, since I_4 is not slid down in these cases, the compression is zero if $H_3 \leq H_4$ and is at most R_{34} if $H_3 > H_4$. See Figure 5.

The d.f. of R_{ij} is $1 - (1-t)^2$ [12], so $\mathbf{P}[\min(R_{12}, R_{34}) > t] = (1-t)^4$, with an integral of $1/5$. Since $\mathbf{P}[H_3 > H_4] = 1/2$, the expected compression is $1/10$.

Lemma 2 *In cases 1b and 2b, the compression is $\Delta_b = \min(R_{12}, H_{(2:2)})$, where R_{12} is the range of a 2-sample, and $H_{(2:2)}$ is the maximum of an independent 2-sample:*

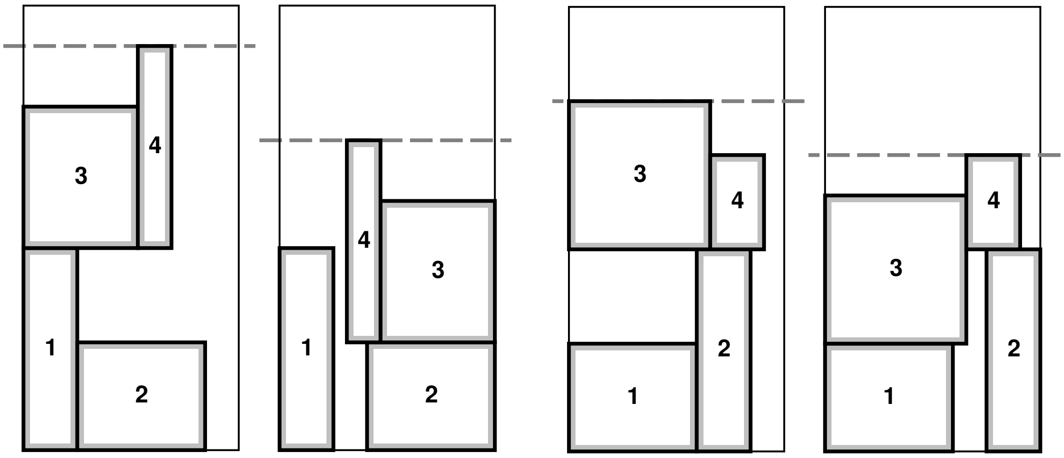
$$R_{12} = \max(H_1, H_2) - \min(H_1, H_2), \qquad H_{(2:2)} = \max(H_3, H_4).$$

Thus,

$$\mathbf{E}\Delta_b = \frac{3}{10}. \qquad (13)$$

Proof: The depth of the "hole" into which I_3 and I_4 can be slid is R_{12}. However, at most $H_{(2:2)}$ gain can be realized, since the top of I_1 is fixed. See Figure 5.

The d.f. of $H_{(2:2)}$ is t^2, so $\mathbf{P}[\min(R_{12}, H_{(2:2)}) > t] = (1-t)^2 \cdot (1-t^2)$, with an integral of $3/10$.

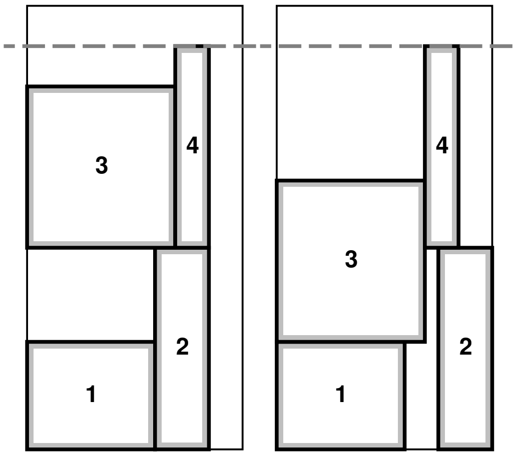


Case 1b: NFL **Case 1b: Compressed** **Case 2a: NFL** **Case 2a: Compressed**

Case 2b is similar

Illustrates $H_3 > H_4$

Case 1a is similar



Case 2a: NFL **Case 2a: Compressed**

Illustrates $H_3 \leq H_4$

Case 1a is similar

Figure 5: Illustrating the compression gained in various cases, where NFL packs two items in the lower level and two in the upper.

4.1.3 Overall Expected Compression: 2 Lower and 2 Upper

For the above patterns, a weighting of the individual expected gains from (12) and (13) with their probabilities from (8)– (11) yields

$$\begin{aligned} \mathbf{E}\Delta_{2,2} &= \frac{1}{70}\frac{1}{10} + \frac{3}{560}\frac{3}{10} + \frac{13}{840}\frac{1}{10} + \frac{17}{1120}\frac{3}{10} \\ &= \frac{307}{33600} = 0.0091369047\ldots \end{aligned} \tag{14}$$

4.2 Two Items in the Lower and One Item in the Upper Level

The two events whose probability and expected compression we wish to calculate are shown in Figure 6. In each case, there is another item of width W_4, not shown, marking the beginning of the next bi-level.

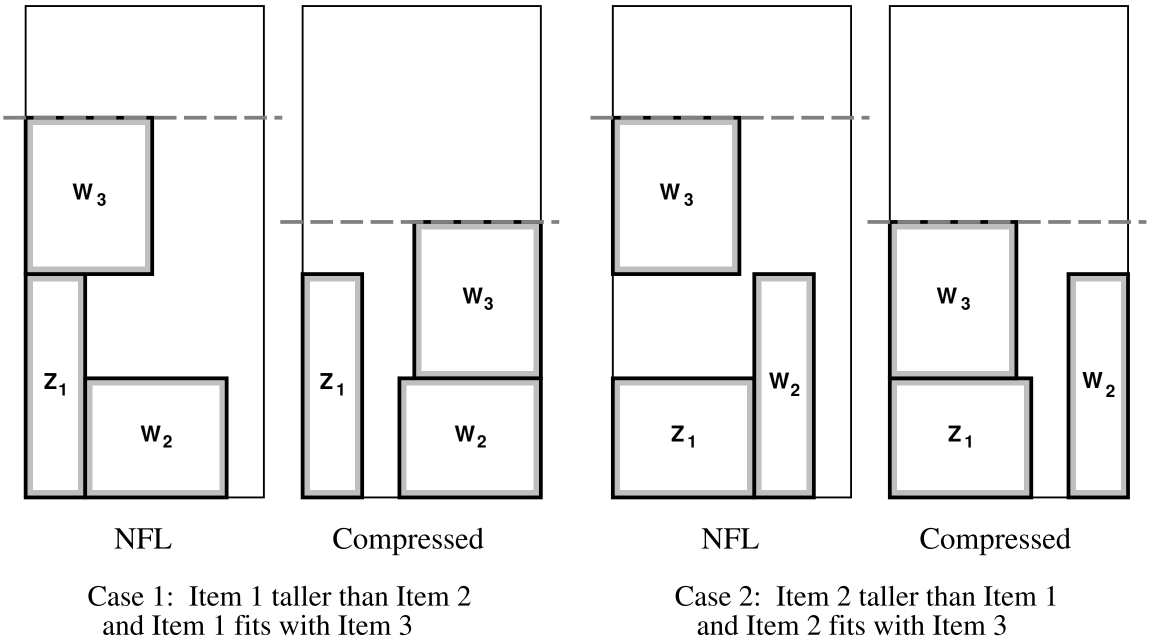


Figure 6: Two items on the lower level and one item on the upper level, shown before compression (as NFL packs them), and shown as compressed by CA. The captions summarize the conditioning on the heights and widths of items 1 and 2 that allow compression to take place. Not shown is an item of width W_4 that makes the start of the next bi-level above the one depicted.

Cases 1 and 2 are symmetric, but their probabilities differ as before because of the width bias in Z_1.

4.2.1 Event Probabilities

In both cases, we have the following conditions, which together form an event to be called B:

$$\begin{array}{rcll} Z_1 + W_2 & \leq & 1: & I_1 \text{ and } I_2 \text{ fit on the lower level} \\ Z_1 + W_2 + W_3 & > & 1: & I_3 \text{ does not fit on the lower level} \\ W_3 + W_4 & > & 1: & I_3 \text{ is alone on the upper level} \end{array}$$

Case 1: We add to event B the conditions that I_1 and I_3 can fit together across the strip. We also have that the (independent) heights of the first two items make I_1 the taller:

$$\begin{aligned} \mathbf{P}[\text{Case 1}] &= \mathbf{P}[H_1 < H_2] \cdot \mathbf{P}[B, Z_1 + W_3 \leq 1] \\ &= \frac{1}{2} \int_0^1 f_Z(z_1) dz_1 \int_0^{1-z_1} dw_2 \int_{1-z_1-w_2}^{1-z_1} dw_3 \int_{1-w_3}^1 dw_4 \\ &= \frac{1}{2} \frac{1}{24} = \frac{1}{48}. \end{aligned} \tag{15}$$

Case 2: In addition to event B, we have the conditions that I_2 and I_3 can fit together across the strip. In this case, we have that item I_2 is taller than I_1.

$$\begin{aligned} \mathbf{P}[\text{Case 2}] &= \mathbf{P}[H_1 \leq H_2] \cdot \mathbf{P}[B, W_2 + W_3 \leq 1] \\ &= \frac{1}{2} \int_0^1 f_Z(z_1) dz_1 \int_0^{1-z_1} dw_2 \int_{1-z_1-w_2}^{1-w_2} dw_3 \int_{1-w_3}^1 dw_4 \\ &= \frac{1}{2} \frac{7}{80} = \frac{7}{160}. \end{aligned} \tag{16}$$

4.2.2 Expected Compressions

The compression achieved in the two cases depends entirely upon the independent $U[0,1]$ heights $H_i,\ i = 1, 2, 3$. As the reasoning here is similar to that in Lemmas 1 and 2 of section 4.1, no figure is provided.

Lemma 3 *In cases 1 and 2, the compression is $\Delta = \min(R_{12}, H_3)$, where R_{12} is the range of a two-sample, and H_3 is an independent uniform random variable:*

$$R_{12} = \max(H_1, H_2) - \min(H_1, H_2).$$

Thus,

$$\mathbf{E}\Delta = \frac{1}{4} \tag{17}$$

Proof: The depth of the "hole" into which I_3 can be slid is R_{12}. However, at most H_3 gain can be realized, since the taller of I_1 and I_2 has its top at a fixed height.

The d.f. of H_3 is t, so $\mathbf{P}[\min(R_{12}, H_3) > t] = (1-t)^2(1-t)$, with an integral of $1/4$.

4.2.3 Overall Expected Compression: 2 Lower and 1 Upper

For the above patterns, a weighting of the individual expected gains from (17) together with their probabilities form (15) and (16) yields

$$\begin{aligned} \mathbf{E}\Delta_{2,1} &= \frac{1}{48}\frac{1}{4} + \frac{7}{160}\frac{1}{4} \\ &= \frac{31}{1920} = 0.0161458333\ldots \end{aligned} \tag{18}$$

4.3 One Item in the Lower and Two Items in the Upper Level

The event whose probability and expected compression we wish to calculate is shown in Figure 7. There is another item of width W_4, not shown, marking the beginning of the next bi-level.

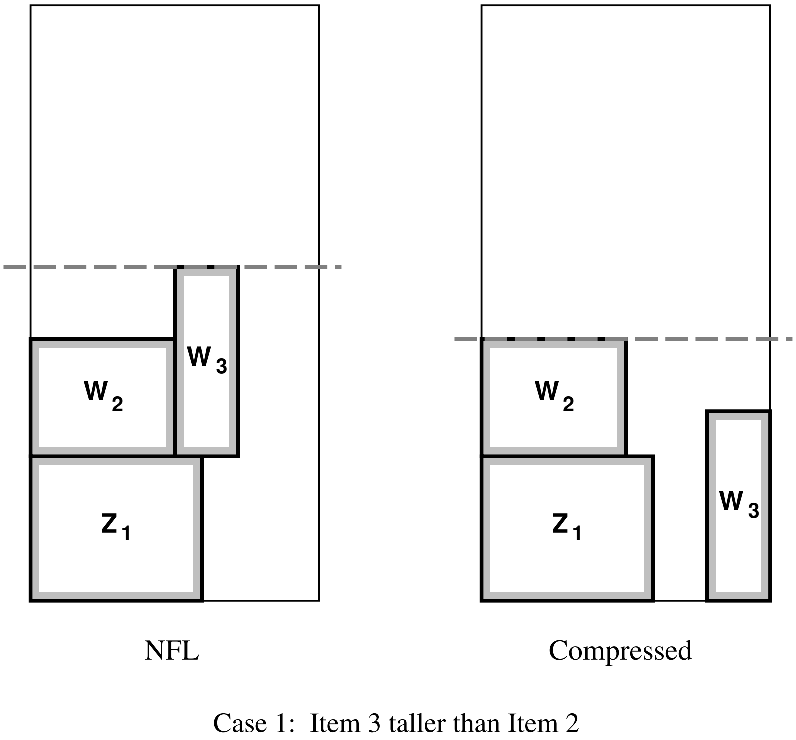


Figure 7: One item on the lower level and two items on the upper level, shown before compression (as NFL packs them), and shown as compressed by CA. Only when $H_3 > H_2$ is any compression realized. Not shown is an item of width W_4 that marks the start of the next bi-level above the one depicted.

Only the case shown, where I_3 is taller than I_2, is of any interest. This is because, in any on-line algorithm, we will not be able to repack I_2, should it prove to be taller than I_3. In the event that happens, repacking I_3 does not reduce the height of the bi-level.

The event of interest is C, the conjunction of the following conditions:

$$\begin{array}{rcll} Z_1 + W_2 & > & 1: & I_1 \text{ is alone on the lower level} \\ W_2 + W_3 & \leq & 1: & I_2 \text{ and } I_3 \text{ fit on the upper level} \\ W_2 + W_3 + W_4 & > & 1: & I_2 \text{ and } I_3 \text{ are alone on the upper level} \end{array}$$

Added to condition C, we have the condition that I_1 and I_3 can fit together across the strip. We also have that the (independent) heights of items I_2 and I_3 make I_3, the one that was encountered later, the taller of the two. The two constraints on W_3 together imply that $W_3 \leq \min(1 - W_2, 1 - Z_1)$, which is reflected in the integral below:

$$\begin{aligned} \mathbf{P}[H_2 \leq H_3] \cdot \mathbf{P}[C, Z_1 + W_3 \leq 1] &= \frac{1}{2} \int_0^1 f_Z(z_1) dz_1 \int_{1-z_1}^1 dw_2 \cdot \\ &\qquad \int_0^{\min(1-w_2, 1-z_1)} dw_3 \int_{1-w_2-w_3}^1 dw_4 \\ &= \frac{1}{2} \frac{29}{384} = \frac{29}{768}. \end{aligned} \tag{19}$$

The compression achieved is calculated in analogy with the proof of Lemma 3 in section 4.2.2, where now we have $\Delta = \min(H_1, R_{23})$. The result is again $\mathbf{E}\Delta = 1/4$.

Weighting the expected gain by the probability from (19) yields

$$\begin{aligned} \mathbf{E}\Delta_{1,2} &= \frac{29}{768} \frac{1}{4} \qquad (20) \\ &= \frac{29}{3072} = 0.0094401041 \ldots \end{aligned}$$

for the overall expected compression in the case: 1 lower and 2 upper.

4.4 Expected Height of the Compression Algorithm

Now that we have the expected bi-level height compressions from (14), (18), and (20), we use them to obtain the asymptotic expected height of a packing by the CA algorithm.

Theorem 2

$$\mathbf{E}CA(L_n) \sim (0.3697642137 \ldots \qquad (n \to \infty) \tag{21}$$

Proof: Since the expected number of levels in NFL is $2n/3$, the expected number of bi-levels in BNFL, and hence in CA, is $n/3$.

Since the expected height of a level in NFL is $0.5720077418\ldots$, the expected height of a bi-level in BNFL is $1.1440154836\ldots$. From the sum of (14), (18), and (20), the expected savings in bi-level height achieved by CA is

$$\frac{307}{33600} + \frac{31}{1920} + \frac{29}{3072} = 0.0347228422\ldots$$

The expected height of a compressed bi-level is thus

$$\begin{aligned}\mathbf{E}[\text{compressed bi-level height}] &= (1.1440154836\ldots) - \frac{18667}{537600} \\ &= 1.1092926413\ldots\end{aligned}$$

whereupon multiplication by $n/3$ yields the desired result.

Comparison of Theorem 2 with (6) shows that CA is an on-line algorithm that improves over NFL by an amount $(0.01157\ldots)n$. The gap between this upper bound and the lower bound of Theorem 1 remains $(0.05593\ldots)n$.

5 Conclusion and Further Work

Other directions can be explored, with a view to improving the bound of $0.36976\ldots$ of Theorem 2. For example, Theorem 2 improves the equilibrium NFL process by operating two levels at a time. One can attempt to exploit compressible patterns at each individual NFL level at equilibrium. The difficulty here is assuring disjointness of the events that form compressible patterns. When the single pattern of a level with 2 items followed by a level with 1 item is applied at every matching level at equilibrium, the result is an expected height of $(0.37058\ldots)n$. This is a simple way to improve on NFL, but it is not as good as the technique leading to Theorem 2.

Another tack is to take items k at a time, pack them in an optimal or at least very dense way, then begin again with k more items at the height of this packing. These *group packing* algorithms are in fact off-line, and only the case $k = 3$ has been worked out in detail. While the analysis is simple, the resulting expected height of $(37/96)n = (0.3854166666\ldots)n$, is worse than NFL. The group packing algorithm with $k = 4$ is more complex to analyze, and its expected height is not known.

References

[1] Azar, Y. and L. Epstein, "On two dimensional bin packing," *J. Algorithms* **25**(1997), 290–310.

[2] Abramowitz, M. and I. A. Stegun, eds. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables,* National Bureau of Standards, Washington D.C., 1970.

[3] Baker, B. S., D. Brown, and H. P. Katseff, "Lower bounds for on-line two-dimensional packing algorithms," *Acta Informatica*, **18** (1982), 207–225.

[4] Baker, B. S., E. G. Coffman, Jr., and R. L. Rivest, "Orthogonal packings in two dimensions," *SIAM J. Comput.,* **9** (1980), 846–855.

[5] Baker, B. S. and J. S. Schwarz, "Shelf algorithms for two-dimensional packing problems," *SIAM J. Comput.,* **12** (1983), 508–525.

[6] Csirik, J. and G. Woeginger, "Shelf algorithms for on-line strip packing," *Infor. Proc. Lett.*, **63**(1997), 171–175.

[7] Coffman, E. G., Jr., M. R. Garey, D. S. Johnson, and R. E. Tarjan, "Performance bounds for level-oriented two-dimensional' packing algorithms," *SIAM J. Comput.,* **9** (1980), 808–826.

[8] Coffman, E. G., Jr., and G. S. Lueker, *Probabilistic analysis of packing and partitioning problems,* Wiley, New York, 1991.

[9] Coffman, E. G., Jr., and P. W. Shor, "Average-case analysis of cutting and packing in two dimensions," *European J. Oper. Res.,* **44** (1990), 134–144.

[10] Coffman, E. G., Jr. and P. W. Shor, "Packing in two dimensions: Asymptotic average-case analysis of algorithms," *Algorithmica*, **9** (1993), 253–277.

[11] Coffman, E. G., Jr., K. So, M. Hofri, and A. C. Yao, "A stochastic model of bin packing," *Information and Control,* **44** (1980), 104–115.

[12] David, H. A., *Order Statistics, 2nd Ed.*, Wiley, New York, 1981.

[13] Gradshteyn, I. S. and I. M. Ryzhik, *Tables of integrals, series and products, corrected and enlarged edition,* Academic Press, New York, 1980.

[14] Hofri, M., "Two dimensional packing: Expected performance of simple level algorithms," *Information and Control*, **45** (1980), 1–17.

[15] Hofri, M., "A probabilistic analysis of the next fit bin packing algorithm," *J. Algorithms,* **5** (1984), 547–556.

[16] Liang, F. M., "A lower bound for on-line bin packing," *Infor. Proc. Lett.,* **10**(1980), 76–79.

[17] Van Vliet, A., "An improved lower bound for on-line bin packing algorithms," *Infor. Proc. Lett.,* **43**(1992), 277–284.

[18] Wolfram, S., *Mathematica*, Version 3, 3rd Edition, Cambridge University Press, New York, 1996.