Probabilistic analysis of $k$-dimensional packing algorithms

Dawei Hong, Joseph Y-T. Leung *

Department of Computer Science and Engineering, University of Nebraska – Lincoln, Lincoln, NE 68588-0115, USA

Communicated by F.B. Schneider; received 3 June 1993; revised 1 February 1995

Abstract

In the $k$-dimensional packing problem, we are given a set $I = \{b_1, b_2, \ldots, b_n\}$ of $k$-dimensional boxes and a $k$-dimensional box $B$ with unit length in each of the first $k-1$ dimensions and unbounded length in the $k$th dimension. Each box $b_i$ is represented by a $k$-tuple $b_i = (x_{i(1)}, \ldots, x_{i(k-1)}, x_{i(k)}) \in (0, 1]^{k-1} \times (0, \infty)$, where $x_{i(j)}$ denotes its length in the $j$th dimension, $1 \leq j \leq k$. We are asked to find a packing of $I$ into $B$ such that each box is packed orthogonally and oriented in all $k$ dimensions and such that the height in the $k$th dimension of the packing is minimized. The $k$-dimensional packing problem is known to be NP-hard for each $k > 1$. In this note, we study the average-case behavior of a class of algorithms, which includes any optimal algorithm and an on-line algorithm. Let $A$ denote an algorithm in this class. Assume that $b_1, b_2, \ldots, b_n$ are independent, identically distributed according to a distribution $F(x_{1(1)}, \ldots, x_{(k-1)}, x_{(k)})$ over $(0, 1]^{k-1} \times (0, \infty)$, and the marginal distribution $F_k$ of $x_{(k)}$ satisfies the property that there is a positive number $\alpha$ at which the moment generating function $M_k(t)$ has a finite value $C_\alpha > 0$. It is shown that for each given $\epsilon > 0$, there is an $N_{\epsilon, F} > 0$ such that for all $n \geq N_{\epsilon, F}$, $\Pr(|A(b_1, \ldots, b_n)/n - \Gamma| > \epsilon) < (2 + C_\alpha)\exp(- (s\alpha/3)^{2/3}n^{1/3})$, where $\Gamma = \lim_{n \to \infty} \mathbb{E}[A(b_1, \ldots, b_n)/n]$ and $A(b_1, \ldots, b_n)$ denotes the height in the $k$th dimension of the packing of $(b_1, \ldots, b_n)$ produced by $A$.

Keywords: Probabilistic analysis of algorithms; NP-hard; $k$-dimensional packing; On-line algorithm

1. Introduction

In the $k$-dimensional packing problem, we are given a set $I = \{b_1, b_2, \ldots, b_n\}$ of $k$-dimensional boxes and a $k$-dimensional box $B$ with unit length in each of the first $k-1$ dimensions and unbounded length in the $k$th dimension. Each box $b_i$ is represented by the $k$-tuple $b_i = (x_{i(1)}, \ldots, x_{i(k-1)}, x_{i(k)}) \in (0, 1]^{k-1} \times (0, \infty)$, where $x_{i(j)}$ denotes its length in the $j$th dimension, $1 \leq i \leq k$. Our goal is to pack $I$ into $B$ such that each box is packed orthogonally and oriented in all $k$ dimensions and such that the height in the $k$th dimension of the packing is minimized. The $k$-dimensional packing problem is known to be NP-hard for each $k > 1$. In this note, we study the average-case behavior of a class of algorithms, which includes

---

* Research supported in part by the ONR grant N00014-91-J-1383 and in part by the CCIS at UNL.
* Corresponding author. Email: jyl@cse.unl.edu.
any optimal algorithm and an on-line algorithm. Throughout this note, we let $A$ denote an algorithm in this class and $A(b_1, \ldots, b_n)$ denote the height in the $k$th dimension of the packing of $(b_1, b_2, \ldots, b_n)$ produced by $A$. Furthermore, we let OPT denote an optimal algorithm and $OPT(b_1, \ldots, b_n)$ denote the height in the $k$th dimension of an optimal packing of $(b_1, b_2, \ldots, b_n)$.

The $k$-dimensional packing problem models many optimization problems in computer science and operations research. For example, the 2-dimensional packing problem has been used to model job scheduling in a multiprogrammed computer system [4]. In this application, each job is represented by a box, where the first and second dimensions of the box represent the memory requirement and time taken by the job, respectively. Worst-case analysis of fast approximation algorithms for 2-dimensional packing is conducted in [2,4], while in [6] its average-case behavior is considered under the assumption that $x_1^{(1)}, \ldots, x_n^{(1)}, x_1^{(2)}, \ldots, x_n^{(2)}$ are independent, uniform random samples from $[0,1]$. Another example is the 3-dimensional packing problem, which was recently introduced as a model of job scheduling in partitionable mesh connected systems (PMCS) [11]. In this problem, the bottom of the box $B$ represents the PMCS and the unbounded height represents the time dimension. Each box $b_i = (x_i^{(1)}, x_i^{(2)}, x_i^{(3)})$ represents a job, where $x_i^{(1)}$ represents the size of a submesh required by the job and $x_i^{(3)}$ represents its execution time. A fast approximation algorithm for the 3-dimensional packing problem was proposed in [11] and its worst-case performance analyzed. The earliest work in probabilistic analysis of $k$-dimensional packing is due to Karp et al. [9]; other works can be found in [5].

For the 2-dimensional packing problem, Karp et al. [9] have shown that

\[
\lim_{n \to \infty} \frac{E[OPT(b_1, \ldots, b_n)]}{n} = \frac{1}{4}.
\]  

(1.1)

Recently, there is a renewed interest [3,6] in the Next Fit Shelf (NFS) algorithm due to Baker and Schwarz [2]. The generalized version of NFS, as described in [3,6], works as follows: First, design a set $R = \{r_j | j \in \mathbb{Z} \text{ and } r_j < r_{j+1}\}$ of shelf heights. Then, the boxes $b_1, \ldots, b_n$ are packed successively into the box $B$, one after the other. In the course of packing the boxes, $B$ can be viewed as a sequence of shelves with heights drawn from $R$; initially, $B$ consists of no shelf. Suppose we are packing the box $b_i = (x_i^{(1)}, x_i^{(2)})$. Let $j$ be the smallest index such that $x_i^{(2)} \leq r_j$. If $b_i$ can be packed into the topmost shelf with height $r_j$, then pack as far left into this shelf as possible; otherwise, pack into the leftmost position of a newly created shelf with height $r_j$. It is clear that NFS is an on-line algorithm. In [3,6], $R$ is chosen to be the set $\{r_j = j/l | j = 1, \ldots, l\}$ for some positive integer $l$. For this choice of $R$, it is clear that the difference between any two consecutive shelf heights is bounded above by the constant $\lambda = 1/l$. With this choice of $R$, it was shown [3,6] that if $b_1, \ldots, b_n$ is a random sample from the uniform distribution over $(0, 1)^2$, then

\[
\lim_{n \to \infty} \frac{E[NFS(b_1, \ldots, b_n)]}{n} = \frac{1}{3}.\frac{1}{3l}.
\]  

(1.2)

Combining (1.1) and (1.2), we have

\[
\lim_{n \to \infty} \frac{E[OPT(b_1, \ldots, b_n)]}{E[NFS(b_1, \ldots, b_n)]} = \frac{3}{4 + 4l^{-1}}.
\]  

(1.3)

In the probability model considered here, we assume that $b_1, b_2, \ldots, b_n$ are independent, identically distributed according to a distribution $F(x^{(1)}, \ldots, x^{(k)})$ over $(0, 1)^{k-1} \times (0, \infty)$. Furthermore, we assume
that the marginal distribution $F_k$ of $x^{(k)}$ satisfies the property that there is a positive number $\alpha$ at which the moment generating function $M_{F_k}(t)$ has a finite value $C_\alpha > 0$, i.e.,

$$\int_{(0,\infty)} e^{\alpha x^{(k)}} \, dF_k = C_\alpha.$$  \hspace{1cm} (1.4)

No further regularity restriction is made on $F$, unless stated otherwise. It is clear that most of the commonly used distributions, such as uniform and exponential distributions, do meet (1.4).

In the next section, we will define a class of algorithms, which includes both OPT and NFS. We then show that for each given $s > 0$, there is an $N_{s,F} > 0$ such that for all $n \geq N_{s,F}$,

$$\Pr\left( \left| A\left(b_1, \ldots, b_n\right) / n - \Gamma \right| > s \right) < \left(2 + C_\alpha\right) \exp\left(-\left(\frac{s\alpha}{3}\right)^{2/3} n^{1/3}\right),$$

where $\Gamma = \lim_{n \to \infty} E[A(b_1, \ldots, b_n)]/n$ and $A(b_1, \ldots, b_n)$ denotes the height in the $k$th dimension of the packing obtained for $(b_1, \ldots, b_n)$ by an algorithm in this class. As a corollary of our result and by (1.1) and (1.2), we can show that for each $0.3 > s > 0$ and large enough $n$,

$$\Pr\left( \frac{3 - 12s}{4 + 4l^{-1} + 12s} \leq \frac{\text{OPT}(b_1, \ldots, b_n)}{\text{NFS}(b_1, \ldots, b_n)} \leq \frac{3 + 12s}{4 + 4l^{-1} - 12s} \right) \geq 1 - 4 \exp\left(-(s^2n)/3\right).$$ \hspace{1cm} (1.5)

The above probabilistic bound provides more information about the average-case behavior of NFS than (1.3) does, since (1.3) merely says that $3/(4 + 4l^{-1})$ is the asymptotic ratio of $E[\text{OPT}(b_1, \ldots, b_n)]$ over $E[\text{NFS}(b_1, \ldots, b_n)]$, while (1.5) indicates that this ratio is nearly guaranteed with exponentially high probability for every instance $(b_1, \ldots, b_n)$.

2. Large deviation bound

For any algorithm $A$, we may view $A(b_1, \ldots, b_n)$ as an $n$-ary function from $D^n$ into $\mathbb{R}^+$, where $D = (0, 1]^{k-1} \times (0, \infty)$ and $\mathbb{R}^+$ is the set of positive real numbers. Consider the class of algorithms that satisfy the following three properties:

(P1) For any $n,m \geq 1$ and $b_1, \ldots, b_n, b_{n+1}, \ldots, b_{n+m},$

$$A(b_1, \ldots, b_n) + A(b_{n+1}, \ldots, b_{n+m}) \geq A(b_1, \ldots, b_n, b_{n+1}, \ldots, b_{n+m}).$$

(P2) There is a constant $\lambda > 0$ such that for any $1 \leq i \leq n$, $b_i = (x^{(1)}_i, \ldots, x^{(k-1)}_i, x^{(k)}_i)$ and $\tilde{b}_i = (\tilde{x}^{(1)}_i, \ldots, \tilde{x}^{(k-1)}_i, \tilde{x}^{(k)}_i)$,

$$\left| A\left(b_1, \ldots, b_{i-1}, b_i, b_{i+1}, \ldots, b_n\right) - A\left(b_1, \ldots, b_{i-1}, \tilde{b}_i, b_{i+1}, \ldots, b_n\right) \right| \leq \max\{x^{(k)}_i + \lambda, \tilde{x}^{(k)}_i + \lambda\}.$$

(P3) For any $(b_1, \ldots, b_n) \in D^n$, $A(b_1, \ldots, b_n) \leq \sum_{i=1}^{n} x^{(k)}_i + n\lambda$, where $\lambda$ is the constant as defined in (P2).

It is clear that OPT satisfies the above three properties with $\lambda = 0$. It is also easy to see that NFS satisfies (P1). To see that NFS satisfies (P2), let $b_i = (x^{(1)}_i, x^{(2)}_i)$ and $\tilde{b}_i = (\tilde{x}^{(1)}_i, \tilde{x}^{(2)}_i)$ be packed into shelves with heights $r_j$ and $\tilde{r}_j$, respectively. By the nature of NFS,

$$\left| \text{NFS}(b_1, \ldots, b_{i-1}, b_i, b_{i+1}, \ldots, b_n) - \text{NFS}(b_1, \ldots, b_{i-1}, \tilde{b}_i, b_{i+1}, \ldots, b_n) \right| \leq \max\{r_j, \tilde{r}_j\}.$$
Since the difference between two consecutive shelf heights is bounded above by a constant, say \( \lambda \), we have \( r_j - x^{(2)}_j < \lambda \) and \( \bar{r}_j - \bar{x}^{(2)}_j \leq \lambda \). Thus, NFS satisfies (P2). Finally, it is easy to see that NFS satisfies (P3).

By the theory of subadditive processes (see p. 186 of [10]), (P1) implies that there is a constant to which \( A(b_1, \ldots, b_n)/n \) converges almost surely, as \( n \to \infty \). By (P3) and the assumption that the moment generating function \( M_{f_i}(t) \) has a finite value at \( \alpha > 0 \), \( E[A(b_1, \ldots, b_n)/n] \) converges to this constant too. (The proof is an application of Lebesgue’s dominated convergence theorem, see [7]). Let this constant be denoted by \( \Gamma \).

Our proof of the probabilistic concentration property of \( A \) is an extension of the basic ideas by Rhee and Talagrand, who first introduced martingale inequalities to the analysis of the classical one-dimensional bin packing problem, the traveling salesman problem, as well as other problems [12, 13]. In this note we define a martingale as in [8]. Let \( \{X_i | i = 0, 1, \ldots, n\} \) and \( \{Y_i | i = 0, 1, \ldots, n\} \) be two sequences of random variables defined on a probability space. We say that \( \{X_i | i = 0, 1, \ldots, n\} \) is a martingale with respect to \( \{Y_i | i = 0, 1, \ldots, n\} \) if for each \( i = 0, 1, \ldots, n \), \( E[X_i | Y_0, \ldots, Y_i] = X_i \).

We need a martingale inequality due to Azuma [11].

**Azuma Lemma.** Suppose the sequence \( X_0, X_1, \ldots, X_n \) is a martingale. Let \( c_i = \sup |X_i - X_{i-1}|, i = 1, \ldots, n, \) and assume that \( \Sigma_{i=1}^n c_i^2 > 0 \). Then, for all \( t > 0 \),

\[
Pr(|X_n - X_0| > t) \leq 2e^{t^2/2(\Sigma_{i=1}^n c_i^2)}.
\]

Note that \( A(b_1, \ldots, b_n) \) is Turing-computable and hence, as an \( n \)-ary function defined on \( D^n \), it is Borel-measurable. Now we use \( A(b_1, \ldots, b_n) \) to define a martingale. The idea behind the definition is quite simple. In the proof of the theorem below, we seek an estimation of \( Pr(|A(b_1, \ldots, b_n) - E[A(b_1, \ldots, b_n)]| > t) \). Let \( X_n = A(b_1, \ldots, b_n) \) and \( X_0 = E[A(b_1, \ldots, b_n)] \). Then, we create \( X_1, \ldots, X_{n-1} \) along which \( E[A(b_1, \ldots, b_n)] \) gradually becomes \( A(b_1, \ldots, b_n) \). The most natural way to do this is to let \( X_i \) be as follows: For each \( i, 1 \leq i \leq n \), \( X_i \) is obtained by taking the conditional expected value of \( A(b_1, \ldots, b_n) \) for each of those subsets of \( D^n \), where two elements in \( D^n \) are in the same subset if and only if they agree on \( b_i \). This suggests that \( Y_i \) be chosen in the following way: Let \( Y_0 \) be a constant function from \( D^n \) into \( \mathbb{R}^+ \), and for each \( i, 1 \leq i \leq n, Y_i \) is a Borel-measurable function from \( D^n \) into \( \mathbb{R}^+ \) that maps any two elements in \( D^n \) to the same number in \( \mathbb{R}^+ \) if and only if these two elements agree on \( b_i \). With this definition of \( Y_i \)'s, we define for each \( i, 0 \leq i \leq n, \)

\[
X_i = E[A(b_1, \ldots, b_n) | Y_0, \ldots, Y_i].
\]

It is easy to see that \( \{X_i | i = 0, 1, \ldots, n\} \) forms a martingale with respect to \( \{Y_i | i = 0, 1, \ldots, n\} \), where \( X_0 = E[A(b_1, \ldots, b_n)] \) and \( X_n = A(b_1, \ldots, b_n) \). This martingale enables us to estimate \( |X_i - X_{i-1}| \) easily. Indeed, since the \( b_i \)'s are independent, the difference between \( X_i \) and \( X_{i-1} \) can only be due to \( b_i \). By (P2), we have

\[
|X_i - X_{i-1}| \leq \sup \{x^{(k)}_i + \lambda | b_i = (x^{(1)}_i, \ldots, x^{(k)}_i, x^{(k)}_i) \}.
\]

Thus, if we could control \( x^{(k)}_i \) in \( b_i \) for each \( i, 1 \leq i \leq n, \) the desired result would follow from the Azuma Lemma. An interesting case is when \( x^{(k)} \) is bounded.

**Proposition 1.** Suppose that the marginal distribution \( F_k \) has a bounded support within \( (0, U] \); i.e., \( x^{(k)} \in (0, U] \). Then, for each \( s > 0 \), there is an \( N_{s,F} > 0 \) such that for all \( n \geq N_{s,F}, \)

\[
Pr\left( \left| \frac{A(b_1, \ldots, b_n)}{n} - \Gamma \right| > s \right) < 2 \exp\left(-\frac{(s^2n)}{3U^2}\right).
\]
Proof. Define a martingale as in (2.1). Then we have \(|X_i - X_{i-1}| \leq U\). The result can be derived from the Azuma Lemma. □

Now we consider the case where \(x^{(k)}\) is unbounded. The following is the main result of this note.

**Theorem 2.** Suppose that the marginal distribution \(F_k\) of \(x^{(k)}\) satisfies the property (1.4); i.e., there is a positive number \(\alpha\) at which the moment generating function \(M_{F_k}(t)\) has a finite value \(C_\alpha > 0\). Then, for each given \(\varepsilon > 0\), there is an \(N_{\varepsilon,F} > 0\) such that for all \(n > N_{\varepsilon,F}\),

\[
\Pr\left(\left|\frac{A(b_1, \ldots, b_n)}{n} - \Gamma\right| > \varepsilon\right) < (2 + C_\alpha) \exp\left(-\left(\frac{s\alpha}{3}\right)^{2/3} n^{1/3}\right).
\tag{2.4}
\]

Proof. To bound \(x_i^{(k)}\) so that the Azuma Lemma can be applied, we define \(f(n)\), for each \(n\), as the largest real root of the equation

\[
\xi(\xi + \lambda) - \frac{\log n}{\alpha}(\xi + \lambda)^2 - \frac{s^2n}{3\alpha} = 0.
\]

Then we transform each \(b_i = (x_i^{(1)}, \ldots, x_i^{(k-1)}, x_i^{(k)})\) into \(\bar{b}_i = (x_i^{(1)}, \ldots, x_i^{(k-1)}, \bar{x}_i^{(k)})\) as follows. For each \(b_i = (x_i^{(1)}, \ldots, x_i^{(k-1)}, x_i^{(k)})\), we define \(\bar{b}_i\) as

\[
\bar{b}_i = \begin{cases} 
 b_i, & \text{if } x_i^{(k)} \leq f(n), \\
 (x_i^{(1)}, \ldots, x_i^{(k-1)}, f(n)), & \text{if } x_i^{(k)} > f(n). 
\end{cases}
\]

Note that \(f(n)\) is chosen so that \(s^2n/(3f(n) + \lambda)^2 = \alpha f(n) - \log n\), which will enable us to obtain a good bound in (2.4). Moreover, it is easy to see that

\[
\lim_{n \to \infty} \frac{f(n)}{n^{1/3}} = \frac{s^{2/3}}{(3\alpha)^{1/3}}. \tag{2.5}
\]

The difference between \(b_i\) and \(\bar{b}_i\) depends on whether \(x_i^{(k)} > f(n)\) or not. By the assumption in (1.4), we have

\[
\Pr(b_i \neq \bar{b}_i) = \int_{(f(n),\infty)} dF_k \leq \frac{C_\alpha}{e^{\alpha f(n)}}. \tag{2.6}
\]

Furthermore, (2.6) and the independence of the \(b_i\)'s imply that

\[
\Pr(A(b_1, \ldots, b_n) = A(\bar{b}_1, \ldots, \bar{b}_n)) \geq \left(1 - \frac{C_\alpha}{e^{\alpha f(n)}}\right)^n.
\]

From (2.5), it is not difficult to verify that for large enough \(n\), say for all \(n > N_1\),

\[
\left(1 - \frac{C_\alpha}{e^{\alpha f(n)}}\right)^n > 1 - \frac{C_\alpha n}{e^{\alpha f(n)}}.
\]

Thus, for all \(n > N_1\),

\[
\Pr(A(b_1, \ldots, b_n) = A(\bar{b}_1, \ldots, \bar{b}_n)) > 1 - \frac{C_\alpha n}{e^{\alpha f(n)}}. \tag{2.7}
\]
Using the $\bar{b}_i$'s, we define a martingale as in (2.1). By the definition of $\bar{b}_i$, it is clear from (2.2) that for $i = 1, \ldots, n$, $|X_i - X_{i-1}| \leq f(n) + \lambda$. Letting $t = \sqrt{6} \ln n / 3$ in the Azuma Lemma, we have for all $s > 0$,

$$
\Pr\left(\left| \frac{A(\bar{b}_1, \ldots, \bar{b}_n)}{n} - \mathbb{E}\left[ A(\bar{b}_1, \ldots, \bar{b}_n) \right] \right| > \frac{\sqrt{6} s}{3} \right) \leq 2e^{-s^2 n / 3(f(n) + \lambda)^2}.
$$

Combining (2.7) and (2.8), we have

$$
\Pr\left(\left| \frac{A(b_1, \ldots, b_n)}{n} - \mathbb{E}\left[ A(\bar{b}_1, \ldots, \bar{b}_n) \right] \right| > \frac{\sqrt{6} s}{3} \right) \leq 2e^{-s^2 n / 3(f(n) + \lambda)^2} + C_{\alpha} e^{-\alpha f(n) - \log n}.
$$

As mentioned before, $f(n)$ was chosen so that $s^2 n / 3(f(n) + \lambda)^2 = \alpha f(n) - \log n$. Define the function $g(n)$ as follows: $g(n) = s^2 n / 3(f(n) + \lambda)^2 = \alpha f(n) - \log n$. From (2.5), we have $\lim_{n \to \infty} g(n) / n^{1/3} = (s\alpha)^{2/3} / 3^{1/3}$. Thus, when $n$ is large enough, say for all $n > N_2$, such that $g(n) / n^{1/3} > (s\alpha)^{2/3} / 3^{2/3}$, we have

$$
2e^{-s^2 n / 3(f(n) + \lambda)^2} + C_{\alpha} e^{-\alpha f(n) - \log n} = (2 + C_{\alpha}) e^{-g(n)} < (2 + C_{\alpha}) \exp \left( -\left( \frac{s\alpha}{3} \right)^{2/3} n^{1/3} \right).
$$

From the above inequality and (2.9), we have for all $n > \max\{N_1, N_2\}$,

$$
\Pr\left(\left| \frac{A(b_1, \ldots, b_n)}{n} - \mathbb{E}\left[ A(\bar{b}_1, \ldots, \bar{b}_n) \right] \right| > \frac{\sqrt{6} s}{3} \right) < (2 + C_{\alpha}) \exp \left( -\left( \frac{s\alpha}{3} \right)^{2/3} n^{1/3} \right).
$$

So far, we have shown that $A(b_1, \ldots, b_n) / n$ concentrates on $\mathbb{E}[A(\bar{b}_1, \ldots, \bar{b}_n)] / n$ with high probability. To complete the proof, all we need show is that $\lim_{n \to \infty} \mathbb{E}[A(\bar{b}_1, \ldots, \bar{b}_n)] / n = \Gamma$. To do this, we consider (2.7) again, which gives a relationship between $A(b_1, \ldots, b_n)$ and $A(\bar{b}_1, \ldots, \bar{b}_n)$. From (2.7), we have for all $n > N_1$,

$$
\Pr\left( \bigcup_{j = n}^{\infty} \left\{ A(b_1, \ldots, b_j) \neq A(\bar{b}_1, \ldots, \bar{b}_j) \right\} \right) \leq C_{\alpha} \sum_{j = n}^{\infty} \frac{j}{e^{\alpha f(j)}},
$$

Combining this inequality, (2.5), the Borel–Cantelli Lemma (see [7]), and the fact that $\lim_{n \to \infty} A(b_1, \ldots, b_n) / n = \Gamma$ almost surely, we have

$$
\lim_{n \to \infty} \frac{A(\bar{b}_1, \ldots, \bar{b}_n)}{n} = \Gamma,
$$

almost surely. It is clear that for any $n$ and $\bar{b}_1, \ldots, \bar{b}_n$, $A(\bar{b}_1, \ldots, \bar{b}_n) / n \leq \sum_{k = 1}^{n} x(k) / n + \lambda$. Noting that $\lim_{n \to \infty} \sum_{k = 1}^{n} x(k) / n - \mathbb{E}[x^{(k)}] < \infty$ almost surely, with (2.11) and Lebesgue's dominated convergence theorem (see [7]), we have $\lim_{n \to \infty} \mathbb{E}[A(\bar{b}_1, \ldots, \bar{b}_n)] / n = \Gamma$. Therefore, for $(1 - \sqrt{6}/3)s > 0$, there is an $N_3 > 0$ such that for all $n > N_3$,

$$
\left| \mathbb{E}\left[ \frac{A(\bar{b}_1, \ldots, \bar{b}_n)}{n} \right] - \Gamma \right| < \left( 1 - \frac{\sqrt{6}}{3} \right) s.
$$

Letting $N_{s,F} = \max\{N_1, N_2, N_3\}$, the theorem follows from the above inequality and (2.10).  \(\square\)
Corollary 3. Suppose $A$ is an algorithm that satisfies (P1)–(P3). Let $\Gamma_{OPT} = \lim_{n \to \infty} E[OPT(b_1, \ldots, b_n)]/n$ and $\Gamma_A = \lim_{n \to \infty} E[A(b_1, \ldots, b_n)]/n$. For each $\Gamma_A > s > 0$, there is an $N_{s,F} > 0$ such that for all $n \geq N_{s,F}$,

$$\Pr\left(\frac{\Gamma_{OPT} - s}{\Gamma_A + s} \leq \frac{OPT(b_1, \ldots, b_n)}{A(b_1, \ldots, b_n)} \leq \frac{\Gamma_{OPT} + s}{\Gamma_A - s}\right)$$

$$\geq 1 - 2(2 + C_n) \exp\left(-\left(\frac{s \alpha}{3}\right)^{2/3} n^{1/3}\right).$$

(2.12)

Moreover, if the marginal distribution $F_k$ has a bounded support within $(0, U]$, then

$$\Pr\left(\frac{\Gamma_{OPT} - s}{\Gamma_A + s} \leq \frac{OPT(b_1, \ldots, b_n)}{A(b_1, \ldots, b_n)} \leq \frac{\Gamma_{OPT} + s}{\Gamma_A - s}\right)$$

$$\geq 1 - 4 \exp(- (s^2 n)/(3U^2)).$$

(2.13)

Consequently, almost surely

$$\lim_{n \to \infty} \frac{OPT(b_1, \ldots, b_n)}{A(b_1, \ldots, b_n)} = \frac{\Gamma_{OPT}}{\Gamma_A}.$$  

(2.14)

Proof. From (2.4) and (2.3), we have (2.12) and (2.13) by a straightforward calculation. (2.14) follows from (2.12), (2.13) and the Borel–Cantelli Lemma (see [7]).

Taking NFS as $A$ and substituting $1/4$ and $1/3 + 1/3l$ for $\Gamma_{OPT}$ and $\Gamma_{NFS}$, respectively, into (2.13), we obtain (1.5) immediately.

3. Concluding remarks

In this note, we have shown a probabilistic concentration property of a class of $k$-dimensional packing algorithms. This class includes any optimal algorithm and an on-line algorithm that has been studied in the literature. Our results provide more information about the average-case behavior of these algorithms than those found in the literature. As can be seen from the proofs, our results are also applicable to the packing of $k$-dimensional, irregular shaped objects into a $k$-dimensional box.

Acknowledgment

We gratefully acknowledge the suggestions by the referees, which have led to a major improvement in the presentation over our earlier version.

References