Probab. Theory Relat. Fields 120, 585–599 (2001)
Digital Object Identifier (DOI) 10.1007/s004400100136

E.G. Coffman, Jr. · George S. Lueker · Joel Spencer · Peter M. Winkler

Packing random rectangles

Received: 1 September 1999 / Revised version: 3 November 2000 /
Published online: 14 June 2001 –

Abstract. A random rectangle is the product of two independent random intervals, each being the interval between two random points drawn independently and uniformly from [0, 1]. We prove that the number C_n of items in a maximum cardinality disjoint subset of n random rectangles satisfies

$$n^{1/2}/K \leq \mathsf{E}C_n \leq Kn^{1/2},$$

where K is an absolute constant. Although tight bounds for the problem generalized to $d > 2$ dimensions remain an open problem, we are able to show that, for some absolute constant K,

$$n^{1/2}/K \leq \mathsf{E}C_n \leq K(n \log^{d-1} n)^{1/2}.$$

Finally, for a certain distribution of random cubes we show that for some absolute constant K, the number Q_n of items in a maximum cardinality disjoint subset of the cubes satisfies

$$n^{d/(d+1)}/K \leq \mathsf{E}Q_n \leq Kn^{d/(d+1)}.$$

1. Introduction

We estimate the expected number of items in a maximum cardinality disjoint subset of n rectangles chosen at random in the unit square. We say that such a subset is a *packing* of the n rectangles, and stress that a rectangle is specified by its position as well as its sides; it can not be freely moved to any position as in strip packing or two-dimensional bin packing (see [2] and the references therein for the probabilistic analysis of algorithms for these problems). A random rectangle is the product of two independent random intervals on the coordinate axes; each random interval

E.G. Coffman: Electrical Engineering Department, Columbia University, New York, NY 10027, USA. The research of this author was done while at Bell Labs, Lucent Technologies

G.S. Lueker: Information and Computer Science Department, University of California, Irvine, CA 92697-3425, USA. e-mail: lueker@ics.uci.edu. The research of this author was partially supported by NSF grant CCR 91-19999 while he was visiting DIMACS on a sabbatical in Fall 1998

J. Spencer: Mathematics Department, New York University, New York, NY 10003, USA

P.M. Winkler: Bell Labs, Lucent Technologies, 700 Mountain Avenue, Murray Hill, NJ 07974, USA

Mathematics Subject Classification (2000): Primary 52C17; Secondary 05C69, 52C15, 60D05

Key words or phrases: n-dimensional packing – 2-dimensional packing – Intersection graphs – Independent sets – Probabilistic analysis of optimization problems

in turn is the interval between two independent uniform random draws from the interval $[0, 1]$.

This problem is an immediate generalization of the one-dimensional problem of packing random intervals [4]. It generalizes in an obvious way to packing random rectangles (boxes) in $d > 2$ dimensions into the d-dimensional unit cube, where each such box is determined by $2d$ independent random draws from $[0, 1]$, two for every dimension. A later section also studies the case of random cubes in $d \geq 2$ dimensions. For this case, it is convenient to wrap around the dimensions of the unit cube to form a toroid. A random cube is generated by drawing $d + 1$ random variables $v_1, v_2, \dots, v_d$, and w, independently and uniformly from $[0, 1]$, to produce the cube

$$[v_1, v_1 + w) \times [v_2, v_2 + w) \times \cdots \times [v_d, v_d + w).$$

We note that for the rectangle packing problem (but not for the cube packing problem) we can replace the uniform distribution over $[0, 1]$ by any continuous distribution over $[0, 1]$ without any change in the distribution of the maximum cardinality of a packing, because the relevant intersection properties depend only on the relative ordering of the points that determine the intervals in each dimension.

Potential applications of our model appear in jointly scheduling resources, where customers require specific "intervals" of a resource for specific intervals of time. Suppose that in a linear network, we have a set S of call requests, each specifying a pair of endpoints (calling parties) that define an interval of the network. If we suppose also that each request gives a future time interval to be reserved for the call, then a call request is a rectangle in the two dimensions of space and time. In an unnormalized and perhaps discretized form, we can pose our problem as finding the expected value of the number of requests in S that can be accommodated.

We note that the combinatorial version of our problem is equivalent to finding maximum independent sets in intersection graphs. For the case of arbitrary rectangles this is NP-complete [6]; hence it is also NP-complete for rectangles of arbitrary dimension $d \geq 2$. It remains NP-complete even if we only allow the packing of equal size squares, by an approach like that in [1, 6] (which was applied in [1] to equal size circles; the approach is equally applicable to equal size squares); again, this generalizes to any dimension $d \geq 2$. For the case of $d = 1$, our problem is equivalent to finding maximum independent sets in interval graphs; this can be solved in linear time [3, 7] even for a larger class of graphs known as chordal graphs. For the case of interval graphs, the fact that a simple greedy algorithm gives the optimum was used in the precise analysis of [4].

For convenience, we use the notation of [5] for describing asymptotic bounds. Let $f(n)$ and $g(n)$ be real-valued functions, with $g(n)$ positive for large n. We say $f(n) = O\big(g(n)\big)$ if there exist constants c and n_0 such that

$$n \geq n_0 \;\Rightarrow\; |f(n)| \leq cg(n).$$

We say $f(n) = \Omega\big(g(n)\big)$ if there exist constants $c > 0$ and n_0 such that

$$n \geq n_0 \;\Rightarrow\; f(n) \geq cg(n).$$

Finally, we say $f(n) = \Theta\big(g(n)\big)$ if there exist positive constants c_1, c_2, and n_0 such that

$$n \geq n_0 \;\Rightarrow\; c_1 g(n) \leq f(n) \leq c_2 g(n).$$

Note that if f and g are defined on the integers, and are positive for positive integers, then for each of these notations we can constrain the choice of n_0 to be 1 without changing the meaning of the definition; thus the bounds are in fact hard bounds rather than merely asymptotic bounds. This explains why we can claim the hard bounds given in the abstract.

Let S_n be a given set of random boxes, and let C_n be the maximum cardinality of any set of mutually disjoint boxes taken from S_n. After preliminaries in the next section, Section 3 proves that, in the case of cubes in $d \geq 2$ dimensions, $\mathsf{E}[C_n] = \Theta(n^{d/(d+1)})$, and Section 4 proves that, in the case of boxes in d dimensions, $\mathsf{E}[C_n] = \Omega\big(n^{1/2}\big)$ and $\mathsf{E}[C_n] = O\big((n \log^{d-1} n)^{1/2}\big)$. In Section 5, the final section, we first prove an $O(n^{1/2})$ bound for a reduced, discretized version of the two dimensional problem, and then prove that the reduced version has the same upper bound as the original version, thus showing that for the case of boxes in $d = 2$ dimensions, $\mathsf{E}[C_n] = \Theta\big(n^{1/2}\big)$; this is the most difficult result in the paper.

2. Preliminaries

We can Poissonize the problem without affecting our results. In this version, the number of rectangles is a Poisson distributed random variable T_λ with mean λ. Equivalently, we could generate the rectangles from a Poisson process with uniform intensity on

$$\{(x, x', y, y') \mid 0 \leq x \leq x' \leq 1 \text{ and } 0 \leq y \leq y' \leq 1\}$$

by letting each point (x, x', y, y') yield the rectangle $[x, x'] \times [y, y']$. Let $C(\lambda)$ denote the number of rectangles packed in a maximum cardinality disjoint subset.

The fact that upper bounds on the $\mathsf{E}C(n)$ are also upper bounds on $\mathsf{E}C_n$ (modulo constant factors) is easily argued as follows. Since $\mathsf{E}C_n$ is nondecreasing in n, we can write $\mathsf{E}C(n) \geq \Pr[T_n > n/2]\mathsf{E}C_{n/2}$. Thus, an upper bound for $\mathsf{E}C(n)$ yields an upper bound for $\mathsf{E}C_{n/2}$ and hence for $\mathsf{E}C_n$, up to a constant multiplicative factor. (In fact standard arguments easily show that

$$\mathsf{E}C_n \sim \mathsf{E}C(n); \tag{1}$$

see Appendix A.) We will continue to parenthesize arguments in the notation of the Poissonized model to distinguish quantities like C_n in the model where the number of rectangles to pack is fixed at n.

Let $X_1, \ldots, X_n$ be i.i.d. with a distribution F concentrated on $[0, 1]$, and let $(s_n \in (0, 1]; n \geq 1)$ be a given sequence. Let $N_n(F, s_n)$, which we shorten to N_n when the arguments are clear from context, be the random variable giving the maximum number of the X_i that can be chosen such that their sum is at most ns_n; equivalently, N_n is such that the sum of the smallest N_n of the X_i is at most ns_n, but the sum of the smallest $N_n + 1$ of the X_i exceeds ns_n. (If $\sum_{i=1}^{n} X_i \leq ns_n$, let

N_n be n.) The remainder of this section derives an estimate of the expected value of $N_n(F, s_n)$.

The following notion will be useful.

Definition 1. *Say a function F is admissible if, for some $\xi > 0$ and $K < 1$, F is continuous on $[0, \xi]$ and for all $x \in [0, \xi]$ we have $F(x/2) \le K F(x)$.*

Let X_i, $i = 1, 2, \ldots, n$, be independent random variables with mean 0 and variance σ^2, and suppose that $|X_i| \le M$. Then for $x \ge 0$ Bernstein's inequality tells us that

$$\Pr\left\{\sum_{i=1}^{n} X_i \ge x\sqrt{n}\right\} \le \exp\left(-\frac{x^2/2}{\sigma^2 + \frac{Mx}{3\sqrt{n}}}\right).$$

The following easily proved corollary will be useful.

Lemma 1. *Let Y be a random variable concentrated on $[0, b]$, and let n be a positive integer. Define Y_n as the sum of n independent samples distributed as Y. Then for any $y \ge 0$, we have*

$$\Pr\{|Y_n - \mathsf{E}Y_n| \ge y\} \le 2\exp\left(-\min\left(\frac{y^2}{4b\,\mathsf{E}Y_n}, \frac{3y}{4b}\right)\right).$$

While the following technical lemma is easily proved by standard techniques, for completeness we record it here in a form convenient to us.

Lemma 2. *Let F be an admissible distribution function on $[0, 1]$. For each element of a given positive sequence $(s_n, n \ge 1)$, let x_n be a solution to*

$$s_n = \int_0^{x_n} x\, dF(x), \tag{2}$$

and assume that $\lim_{n\to\infty} s_n = \lim_{n\to\infty} x_n = 0$. Then if $nF(x_n) = \Omega(\log^2 s_n^{-1})$, we have

$$\mathsf{E}[N_n(F, s_n)] \sim nF(x_n)\,. \tag{3}$$

Proof. By (2), we have $s_n \le x_n c_n$, where $c_n := F(x_n)$ is introduced to simplify notation. Also, letting K be as given in Definition 1, we have

$$s_n \ge \int_{x_n/2}^{x_n} x\, dF(x) \ge \frac{x_n}{2}\int_{x_n/2}^{x_n} dF(x) \ge \frac{x_n}{2}(1-K)F(x_n) = \frac{(1-K)x_n c_n}{2},$$

where the penultimate step follows from $\lim_{n\to\infty} x_n = 0$ and the fact that F is admissible. (Statements in this proof should all be interpreted as holding for sufficiently large n.) Thus, $s_n = \Theta(x_n c_n)$.

Now let $(\epsilon_n, n \ge 1)$ be a positive sequence to be chosen later with the property that $\epsilon_n \to 0$. Let x_n' be a solution to $F(x_n') = c_n(1-\epsilon_n/2)$. For large enough n, we will have $1-\epsilon_n/2 > K$, so from the fact that F is admissible, we have $x_n' > x_n/2$. Then for some $\beta_n \in [1/2, 1]$,

$$\int_{x_n'}^{x_n} x\, dF(x) = \beta_n x_n \int_{x_n'}^{x_n} dF(x) = \beta_n x_n \epsilon_n c_n/2. \tag{4}$$

By the fact that $s_n = \Theta(x_n c_n)$, we see that this is $o(s_n)$, and hence

$$\int_0^{x_n'} x\, dF(x) \sim s_n\ . \tag{5}$$

Now if $N_n \leq nc_n(1-\epsilon_n)$, we must have

$$|\{i : X_i \leq x_n'\}| \leq nc_n(1-\epsilon_n) \tag{6}$$

or

$$\sum_{i:\ X_i \leq x_n'} X_i \geq ns_n\ . \tag{7}$$

To apply Lemma 1 to (6), let Y be distributed as the indicator $1_{X_i \leq x_n'}$ so that

$$\mathsf{E}[Y_n] = nF(x_n') = nc_n(1-\epsilon_n/2) \sim nc_n\ ,$$

by the definition of x_n'. Next, set

$$\begin{aligned} b &= 1, \\ y &= \mathsf{E}[Y_n] - nc_n(1-\epsilon_n) = n\epsilon_n c_n/2\ , \end{aligned}$$

so by Lemma 1 the probability of the event in (6) is bounded by

$$\exp\Big(-\Omega\big(n\epsilon_n c_n \min(1,\epsilon_n)\big)\Big) = \exp\big(-\Omega\big(n\epsilon_n^2 c_n\big)\big)\ . \tag{8}$$

Similarly, we apply Lemma 1 to (7) with Y distributed as $1_{X_i \leq x_n'} X_i$ and the remaining parameters

$$\begin{aligned} &\mathsf{E}[Y_n] = n\int_0^{x_n'} x\, dF(x) \sim ns_n, \text{ by (5),} \\ &b = x_n' = \Theta(x_n) \text{ since } x_n' > x_n/2, \text{ and} \\ &y = ns_n - \mathsf{E}[Y_n] = \Theta(nx_n\epsilon_n c_n) \text{ by (4)} \end{aligned}$$

to bound the probability of the event in (7) by

$$\exp\left(-\Omega\left(n\epsilon_n c_n \min\left(1, \frac{x_n\epsilon_n c_n}{s_n}\right)\right)\right) = \exp\left(-\Omega\big(n\epsilon_n^2 c_n\big)\right) \tag{9}$$

where the last step follows from $s_n = \Theta(x_n c_n)$. Thus (9) bounds the probability that (6) or (7) holds, and thus bounds the probability that $N_n \leq nc_n(1-\epsilon_n)$. By letting $\epsilon_n = (nc_n)^{-1/5}$, we can make this probability be $o(s_n)$ by the assumption $nF(x_n) = \Omega(\log^2 s_n^{-1})$ of the lemma. Hence

$$\mathsf{E}N_n \geq nc_n(1-\epsilon_n)\big(1-o(s_n)\big) \sim nc_n.$$

A similar lower bound can be shown; combining these two bounds proves (3). $\square$

3. Random cubes

The optimum packing of random cubes is readily analyzed. We work with a d-dimensional unit cube, and allow (toroidal) wrapping in all axes. The n cubes are generated independently as follows: First a vertex $(v_1, v_2, \ldots, v_d)$ is generated by drawing each v_i independently from the uniform distribution on $[0, 1]$. Then one more value w is drawn independently, again uniformly from $[0, 1]$. The cube generated is

$$[v_1, v_1 + w) \times [v_2, v_2 + w) \times \cdots \times [v_d, v_d + w),$$

where each coordinate is taken modulo 1. In this set-up, we have the following result.

Theorem 1. *The expected cardinality of a maximum packing of n random cubes is $\Theta(n^{d/(d+1)})$.*

Proof. We begin by considering the following simple heuristic. Subdivide the cube into c^{-d} cells with sides

$$c = \alpha n^{-1/(d+1)},$$

where α is a parameter that may be chosen to optimize performance. For each cell $\mathscr{C}$, if there are any generated cubes contained in $\mathscr{C}$, include one of these in the packing. Clearly, all of the cubes packed are nonoverlapping.

To analyze this heuristic, we first fix a cell $\mathscr{C}$ with side c and estimate the probability that a generated cube will lie completely within $\mathscr{C}$. Using the cube generation process described above, the probability that the vertex $(v_1, v_2, \ldots, v_d)$ lies in $\mathscr{C}$ is c^d. For the generated cube to lie in $\mathscr{C}$ we must of course also have that the generated size w is in $[0, c)$, which happens with probability c. It is not hard to see that if both of these conditions are met, then the probability that the cube will fit is just the probability that the last of $d + 1$ uniform draws from $[0, c]$ is the smallest, i.e., $1/(d + 1)$. Hence the probability that a generated cube fits into a particular cell $\mathscr{C}$ is $c^{d+1}/(d + 1)$, and the probability that $\mathscr{C}$ remains empty after generating all n cubes is

$$\left(1 - \frac{c^{d+1}}{d+1}\right)^n = \left(1 - \frac{\alpha^{d+1}}{n(d+1)}\right)^n \sim \exp\left(-\frac{\alpha^{d+1}}{d+1}\right).$$

Since the number of cells is

$$\frac{1}{c^d} = \frac{n^{d/(d+1)}}{\alpha^d}$$

the expectation of the total number of cubes packed is asymptotic to

$$\alpha^{-d}\left(1 - \exp\left(-\frac{\alpha^{d+1}}{d+1}\right)\right) n^{d/(d+1)},$$

which gives the desired lower bound.

The upper bound is based on the simple observation that the sum of the volumes of the packed cubes is at most 1. First we consider the probability distribution of the volume of a single generated cube. The side of this cube is a uniform random variable U over [0, 1]. Thus the probability that its volume is bounded by x is

$$F(x) = \Pr\left\{U^d \le x\right\} = \Pr\left\{U \le x^{1/d}\right\} = x^{1/d} .$$

Then applying Lemma 2 with

$$s_n = \frac{1}{n}, \quad x_n = \left(\frac{d+1}{n}\right)^{d/(d+1)} , \quad c_n = \left(\frac{d+1}{n}\right)^{1/(d+1)} ,$$

we conclude that the expected number of cubes selected before their total volume exceeds 1 is asymptotic to

$$(d+1)^{1/(d+1)} n^{d/(d+1)} ,$$

which gives the desired matching upper bound. □

4. Bounds for $d \ge 2$ dimensional boxes

In the remainder of the paper, let $\mathscr{H}_d$ denote the unit hypercube in $d \ge 1$ dimensions. This section is confined to the following result.

Theorem 2. *Fix d and draw n boxes uniformly at random from $\mathscr{H}_d$. The maximum number that can be packed without overlap is asymptotically bounded from below by $\Omega(\sqrt{n})$ and from above by $O(\sqrt{n \ln^{d-1} n})$.*

Proof. We start by calculating the distribution function F_d for the volume of a d-dimensional box, recalling that each dimension is distributed as the absolute difference between two independent random draws, and thus has the density $2(1-z)$. F_d is the distribution of the product of d independent such variables. For $d = 1$ we readily obtain

$$F_1(x) = 2x - x^2 .$$

For higher dimensions we may use the recurrence

$$\begin{aligned} F_{d+1}(x) &= \int_0^x 2(1-z)\,dz + \int_x^1 2(1-z)\,F_d\left(\frac{x}{z}\right) dz \\ &= 2x - x^2 + 2\int_x^1 (1-z)\,F_d\left(\frac{x}{z}\right) dz . \end{aligned} \tag{10}$$

In particular, one computes

$$F_2(x) = 4x \ln x^{-1} + 2x^2 \ln x^{-1} - 4x + 5x^2 . \tag{11}$$

The exact form of F_d becomes complicated as d increases, so we will settle for an asymptotic estimate. Some notation will be useful. Let $L_k(z)$ denote the set of all functions which can be formed by taking linear combinations (with absolute-constant coefficients) of terms of the form $\ln^m z$, where $m \in \{0, 1, \ldots, k\}$. Thus, in particular, $L_0(z)$ just represents the set of all constants. An expression containing one or more instances of this notation is to be interpreted as the set of all possible functions that can be obtained by replacing each instance by any element of the set it represents.

The following is proved in Appendix B.

Lemma 3. *For $d \geq 2$, we have*

$$F_d(x) \in \frac{2^d}{(d-1)!}\, x \ln^{d-1} x^{-1} + x L_{d-2}(x^{-1}) + x^2 L_{d-1}(x^{-1})\,.$$

To apply Lemma 2 we solve

$$\frac{1}{n} = s_n = \int_0^{x_n} x\, dF_d(x) \sim \int_0^{x_n} \frac{2^d}{(d-1)!}\, x \ln^{d-1} x^{-1} dx \sim \frac{2^{d-1}}{(d-1)!}\, x_n^2 \ln^{d-1} x_n^{-1}$$

to obtain

$$x_n \sim \sqrt{\frac{(d-1)!}{n \ln^{d-1} n}}\,,$$

and then

$$F_d(x_n) \sim 2\sqrt{\frac{\ln^{d-1} n}{(d-1)!\, n}}\,.$$

Together with Lemma 2, this yields the desired upper bound; more precisely, this gives the asymptotic upper bound

$$2\sqrt{\frac{n \ln^{d-1} n}{(d-1)!}}\,.$$

The lower bound argument is the same as that for cubes, except that $\mathscr{H}_d$ is partitioned into cells with sides on the order of $n^{-1/(2d)}$. It is easy to verify that, on average, there is a constant fraction of the $\Theta(n^{1/2})$ cells in which each cell wholly contains at least one of the given rectangles. The details are left to the reader. □

5. Tight bound for $d=2$

Closing the gaps left by the bounds on $\mathsf{E}[C_n]$ for $d \geq 3$ remains an interesting open problem. However, we can show that the lower bound for $d = 2$ is tight, i.e., $\mathsf{E}[C_n] = \Theta(n^{1/2})$. To prove the $O(n^{1/2})$ bound, we analyze the following reduced, discretized version. A *canonical* interval is an interval that, for some $i \geq 0$, has length 2^{-i} and has a left endpoint at some multiple of 2^{-i}. A canonical rectangle is the product of two canonical intervals. We consider a Poissonized model of canonical rectangles in which the number of instances of each possible canonical rectangle of area a is independently Poisson distributed with mean λa^2. We refer to this version as the *reduced problem with parameter* λ, and let $C^*(\lambda)$ denote the cardinality of a maximum packing.

Note that there are $i+1$ shapes possible for a rectangle of area 2^{-i}, and that for each of these shapes there are 2^i canonical rectangles. The mean number of each of these is $\lambda/2^{2i}$. Thus, the total number $T(\lambda)$ of rectangles in the reduced problem with parameter λ is Poisson distributed with mean

$$\sum_{i=0}^{\infty}(i+1)2^i\left(\lambda 2^{-2i}\right) = \lambda\sum_{i=0}^{\infty}(i+1)2^{-i} = 4\lambda\,. \tag{12}$$

Theorem 3. *We have the asymptotic upper bound*

$$\mathsf{E}[C^*(\lambda)] = O\left(\lambda^{1/2}\right).$$

Proof. Let Z_1 be the maximum number of rectangles that can be packed if we disallow packings that use rectangles spanning the height of the square. Define Z_2 similarly when packings that use rectangles spanning the width of the square are disallowed. By symmetry, Z_1 and Z_2 have the same distribution, although they may not be independent. To investigate this distribution, we begin with two observations, the second relying on the properties of canonical rectangles: (i) a rectangle spanning the width of $\mathscr{H}_2$ and a rectangle spanning the height of $\mathscr{H}_2$ must intersect and hence can not coexist in a packing; (ii) rectangles spanning the height of $\mathscr{H}_2$ are the only rectangles crossing the horizontal line separating the top and bottom halves of $\mathscr{H}_2$ and rectangles spanning the width of $\mathscr{H}_2$ are the only ones crossing the vertical line separating the left and right halves of $\mathscr{H}_2$. It follows that, if a maximum cardinality packing is not just a single 1×1 square, then it consists of a pair of disjoint maximum cardinality packings, one in the bottom half and one in the top half of $\mathscr{H}_2$, or a similar pair of subpackings, one in the left half and one in the right half of $\mathscr{H}_2$. After rescaling, these subpackings become solutions to our original problem on $\mathscr{H}_2$ with the new parameter λ times the square of half the area of $\mathscr{H}_2$, i.e., $\lambda/4$. We conclude that Z_1 and Z_2 are both distributed as the sum of two independent samples of $C^*(\lambda/4)$, and that

$$C^*(\lambda) \le Z_0 + \max(Z_1, Z_2)\,, \tag{13}$$

where Z_0 is the indicator function of the event that the entire square $\mathscr{H}_2$ is one of the given rectangles. Note that Z_0 is independent of Z_1 and Z_2.

At this point it is convenient to consider the transform

$$S(\lambda) := \mathsf{E}\left[e^{\alpha C^*(\lambda)}\right].$$

Since $C^*(\lambda) \le T(\lambda)$ and $T(\lambda)$ is Poisson distributed with parameter 4λ, we have

$$S(\lambda) \le \mathsf{E}\left[e^{\alpha T(\lambda)}\right] = \exp\left(4\lambda(e^\alpha - 1)\right). \tag{14}$$

Using (13) we have

$$\begin{aligned} S(\lambda) &= \mathsf{E}\left[e^{\alpha C^*(\lambda)}\right] \le \mathsf{E}\left[e^{\alpha Z_0}\right]\mathsf{E}\left[e^{\alpha \max(Z_1,Z_2)}\right] \\ &\le e^\alpha\,\mathsf{E}\left[\max\left(e^{\alpha Z_1}, e^{\alpha Z_2}\right)\right] \le e^\alpha\left(\mathsf{E}[e^{\alpha Z_1}] + \mathsf{E}[e^{\alpha Z_2}]\right). \end{aligned} \tag{15}$$

Recalling that Z_1, Z_2 are equal in distribution to sums of two independent samples of $C^*(\lambda/4)$, we have $\mathsf{E}[e^{\alpha Z_1}] = \mathsf{E}[e^{\alpha Z_2}] = \left(S(\lambda/4)\right)^2$, and so

$$S(\lambda) \le 2e^\alpha\left(S(\lambda/4)\right)^2. \tag{16}$$

Elementary bounds and calculations are all that remain. We readily compute from (16) that

$$S(\lambda) \le (2e^{\alpha})^{2^M-1}\big(S(1)\big)^{2^M},$$

where $M := \lceil \log_4 \lambda \rceil$. By (14), we have $S(1) \le \exp\big(4(e^{\alpha}-1)\big)$, and so

$$S(\lambda) \le (2e^{\alpha})^{2^M-1} \exp\left(2^{M+2}\big(e^{\alpha}-1\big)\right). \tag{17}$$

Then, using the bound $\mathsf{E}[C^*(\lambda)] \le \alpha^{-1} \ln \mathsf{E}\big[e^{\alpha C^*(\lambda)}\big]$, we obtain

$$\mathsf{E}[C^*(\lambda)] \le \alpha^{-1}\left(\big(2^M-1\big)(\alpha+\ln 2) + 2^{M+2}\big(e^{\alpha}-1\big)\right).$$

Since $2^M = \Theta(\lambda^{1/2})$, we can put $\alpha = 1$ and conclude that $\mathsf{E}[C^*(\lambda)] = O(\lambda^{1/2})$, as desired. □

Before continuing we recall some basic facts about the Poisson distribution. Let Π_λ denote the Poisson distribution with mean λ.

Fact 1. *Let X and Y be independent random variables with distributions Π_λ and $\Pi_{\lambda'}$, respectively. Then $X+Y$ has distribution $\Pi_{\lambda+\lambda'}$.*

Fact 2. *Let X be a positive integer-valued random variable, and interpret each value as a color; let $p_c = \Pr\{X = c\}$. Also let N have distribution Π_λ. Suppose we take N balls and color each one independently distributed as X. Then the number of balls of color c is independent of the numbers of balls of the other colors and has distribution $\Pi_{\lambda p_c}$.*

(Note that Fact 2 follows from the readily verified equalities

$$e^{-\lambda}\frac{\lambda^k}{k!}\binom{k}{x_1, x_2, \ldots}\prod_{c=1}^{\infty} p_c^{x_c} = e^{-\lambda}\lambda^k \prod_{c=1}^{\infty}\frac{p_c^{x_c}}{x_c!} = \prod_{c=1}^{\infty} e^{-p_c\lambda}\,\frac{(p_c\lambda)^{x_c}}{x_c!},$$

where $k = \sum_{c=1}^{\infty} x_c$.)

Theorem 3, together with the following Lemma, implies our main result. In the proof we say that a random variable X *is stochastically dominated by* a random variable Y if for all real z we have

$$\Pr\{X \le z\} \ge \Pr\{Y \le z\}.$$

Lemma 4. *The following bound holds for $d = 2$ dimensions.*

$$\mathsf{E}C(\lambda) \le \mathsf{E}C^*(9\lambda/4).$$

Proof. We define a randomized mapping $\mathscr{M}$ which maps random instances of the original problem to instances of the reduced problem in such a way that a) the solution to the reduced problem is an upper bound on the solution to the original problem, and b) the distribution of the solution to the reduced problem is simply $C^*(9\lambda/4)$.

Any interval in $\mathscr{H}_1$ contains either one or two canonical intervals of maximal length. For example, $[0.2, 0.7)$ contains only $\left[\frac{1}{4}, \frac{2}{4}\right)$ while $[0.3, 0.7)$ contains both $\left[\frac{3}{8}, \frac{4}{8}\right)$ and $\left[\frac{4}{8}, \frac{5}{8}\right)$. (Note that the latter pair of canonical intervals combine into an interval with length $1/4$ but this interval is not canonical since its left endpoint is not at a multiple of $1/4$. Note also that no interval can contain three distinct maximal canonical intervals, since two of them could be joined to form a larger canonical interval.) Let the *canonical subinterval* I' of an interval I be the maximal canonical interval in I, if only one exists, and one such interval chosen uniformly and randomly if two exist. If $R = I_1 \times I_2$ is any rectangle, we define the *canonical subrectangle* of R to be $I_1' \times I_2'$, where I_j' is the canonical subinterval of I_j, $j = 1, 2$. The mapping $\mathscr{M}$ simply replaces each original rectangle by its canonical subrectangle. Note that shrinking the rectangles can never produce new intersections between rectangles, so the solution to the new problem is at least as large as the solution to the original problem.

Next, we investigate the distribution of the solution to the new problem. To begin, we determine the distribution of the canonical subinterval of an interval I between two points chosen independently and uniformly at random from $[0, 1]$. Consider the canonical interval $I' = [k2^{-i}, (k+1)2^{-i})$, and assume that

$$k \text{ is odd and } 0 \le (k-1)2^{-i} \le (k+3)2^{-i} \le 1 \tag{18}$$

or

$$k \text{ is even and } 0 \le (k-2)2^{-i} \le (k+2)2^{-i} \le 1. \tag{19}$$

Assume first that k is odd. (See Figure 1(a), where the thick black interval represents I'.) For I' to be the canonical subinterval of I, we must clearly have $I' \subseteq I$. Moreover, the left endpoint of I cannot be at or to the left of $(k-1)2^{-i}$, since then the larger interval $[(k-1)2^{-i}, (k+1)2^{-i}]$ would be canonical, contradicting the maximality of I'. Hence the left endpoint must be in $\left((k-1)2^{-i}, k2^{-i}\right]$. (This is shown as a shaded interval in the figure.) Given that the left endpoint lies in this interval, if the right endpoint lies in $\left[(k+1)2^{-i}, (k+2)2^{-i}\right)$, I' will be the canonical interval. (Again, this interval is shaded in the figure.) If the right endpoint lies in $\left[(k+2)2^{-i}, (k+3)2^{-i}\right)$, then there is a 50% chance that I' is the canonical subinterval. (This is indicated by a thinner shaded region in the figure.) Finally, if the right endpoint of I is at or to the right of $(k+3)2^{-i}$, then the interval

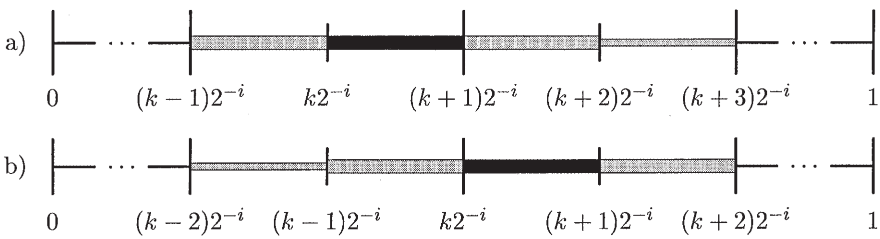

Fig. 1. The vertical lines represent endpoints of canonical intervals of length 2^{-i}, and the tall vertical lines represent endpoints of canonical intervals of length 2^{-i+1}.

$\left[(k+1)2^{-i}, (k+3)2^{-i}\right]$ is canonical, contradicting the maximality of I'. Thus the probability that I' is the canonical subinterval of I is $2^{-i} \cdot \frac{3}{2}2^{-i} = \frac{3}{2}2^{-2i}$. If condition (18) is not satisfied, some of the shaded regions shown in the figure will lie outside the interval [0, 1], but $\frac{3}{2}2^{-2i}$ is still an upper bound on the probability that I' is the canonical interval. A similar argument holds for even k, and is illustrated in Figure 1(b); we omit the details.

Extend the interval calculation to rectangles and note that the probability of obtaining a given canonical rectangle of width w, height h, and area $a = wh$ from a random rectangle is bounded by

$$\frac{3}{2}w^2 \cdot \frac{3}{2}h^2 = \frac{9}{4}a^2. \tag{20}$$

Sublemma 1. *For each possible canonical subrectangle, the number of copies produced by this mapping is Poisson-distributed, and independent of the numbers of the other canonical subrectangles produced.*

Proof. To view the process more formally, let G be a function which maps an arbitrary rectangle into the list of maximum-size canonical rectangles it contains, and let $\mathscr{R}_G$ be the range of G. Thus each element $l \in \mathscr{R}_G$ is a list of 1, 2, or 4 canonical rectangles. When $|l|$ is 2 or 4, coin flips are used to select one of the elements of l. Let $\mathscr{F}$ be the set $\{1, 2, 3, 4\}$, representing the outcome of two coin flips. Let H be the mapping whose domain is $\mathscr{R}_G \times \mathscr{F}$, with the image defined as

$$H(l, c) = \begin{cases} \text{the element of } l & \text{if } |l| = 1 \\ \text{the first element of } l & \text{if } |l| = 2 \text{ and } c \in \{1, 2\} \\ \text{the second element of } l & \text{if } |l| = 2 \text{ and } c \in \{3, 4\} \\ \text{the } c\text{-th element of } l & \text{if } |l| = 4. \end{cases}$$

Informally, H is the function which selects which of the canonical rectangles to use when there is more than one possibility. Then the process of mapping a rectangle R to its canonical subrectangle can be viewed as follows: select a random element $c \in \mathscr{F}$ and then compute $H\big(G(R), c\big)$.

We now use Fact 2 to establish that, after mapping all of the initial rectangles by G, and flipping all of the coins, the number of copies of each element (l, c) of the domain $\mathscr{R}_G \times \mathscr{F}$ of H is independent and has a Poisson distribution. To do so, we recall that the total number of rectangles in the input has a Poisson distribution. (Assume that we first determine the number of rectangles without looking at which rectangles are actually generated.) These rectangles correspond to the "balls." The "color" of a rectangle R is the specification of both the list $l = G(R)$ and the value c in $\mathscr{F}$ giving the coin flips. So the probability of color (l, c) is the probability that the rectangle contains precisely the list l of canonical rectangles, times 1/4 since the two coin flips are fair. Thus Fact 2 tells us that the total number of instances of (l, c), for each $l \in \mathscr{R}_G$ and $c \in \mathscr{F}$, is Poisson distributed and independent.

These "colors" do not correspond 1-1 to the canonical subrectangles, but for any two distinct canonical rectangles R and R', the sets $H^{-1}(R)$ and $H^{-1}(R')$ are disjoint, so when we have completed the construction of all the canonical subrectangles the number of copies of the various possibilities for canonical subrectangles are sums of disjoint sets of independent Poisson-distributed variables, and hence using Fact 1 are themselves Poisson-distributed and independent. □

Completion of Proof of Lemma 4. From (20) and the sublemma we now know that the number of copies of each canonical rectangle R is independent and of the form $\Pi_{\lambda'}$, where $\lambda' \leq (9/4)a^2\lambda$ and a is the area of R. Now add a number of additional copies of R with distribution $\Pi_{9a^2\lambda/4-\lambda'}$ (independently for each R). Of course this can only increase the number of rectangles in the packing, since any packing which was feasible before these additions will remain feasible after these additions. By Fact 1 the total number of instances of R will now have the exact distribution $\Pi_{(9/4)\lambda a^2}$. Thus the distribution of rectangles is now identical to that of $C^*(9\lambda/4)$.

We conclude that $C(\lambda)$ is stochastically dominated by $C^*(9\lambda/4)$, from which the lemma is an immediate consequence. □

Theorem 4. *For $d = 2$, $\mathsf{E}C_n = \Theta(n^{1/2})$.*

Proof. The lower bound $\mathsf{E}C_n = \Omega(n^{1/2})$ was proved in the previous section. Now combining (1), Lemma 4, and Theorem 3, we have

$$\mathsf{E}C_n \sim \mathsf{E}C(n) \leq \mathsf{E}C^*(9n/4) = O(n^{1/2}). \qquad \square$$

Acknowledgements. In the early stages of this research, we had useful discussions with Richard Weber, which we gratefully acknowledge. We also very much appreciate the referee's prompt report and helpful remarks.

Appendix A: Proof of (1)

Let $p \in (0, 1)$. Suppose that after solving an instance of rectangle packing with n rectangles, we randomly select $\lceil pn \rceil$ of the original rectangles to form a new problem. Then the selected rectangles that were used in the original packing must be a valid packing for the new problem. It follows that

$$\mathsf{E}C_n \geq \mathsf{E}C_{\lceil np \rceil} \geq p\,\mathsf{E}C_n,$$

and hence

$$\mathsf{E}C_{n(1+o(1))} = \big(1 + o(1)\big)\mathsf{E}C_n.$$

Similarly,

$$\mathsf{E}C\big(n(1 + o(1))\big) = \big(1 + o(1)\big)\mathsf{E}C(n).$$

Now let Π_n denote a random variable with a Poisson distribution and mean n. It is well-known that, for any $k \geq 1$,

$$\Pr\left\{|\Pi_n - n| \geq \sqrt{n}\log n\right\} = o(n^{-k})$$

and

$$\Pr\left\{\Pi_n \geq n + \sqrt{n}\log n\right\} \mathsf{E}\big[\Pi_n \,\big|\, \Pi_n \geq n + \sqrt{n}\log n\big] = o(n^{-k}).$$

Hence

$$\mathsf{E}C(n) \geq \Pr\left\{\Pi_n \geq \lfloor n - \sqrt{n}\log n \rfloor\right\} C_{\lfloor n-\sqrt{n}\log n\rfloor} \sim C_n,$$

and

$$\begin{aligned}\mathsf{E}C(n) &\leq \Pr\left\{\Pi_n \leq \lceil n + \sqrt{n}\log n\rceil\right\} C_{\lceil n+\sqrt{n}\log n\rceil} \\ &\qquad + \Pr\left\{\Pi_n > \lceil n + \sqrt{n}\log n\rceil\right\} \mathsf{E}\left[\Pi_n \,\middle|\, \Pi_n > \lceil n + \sqrt{n}\log n\rceil\right] \\ &\sim \mathsf{E}C_n. \end{aligned}$$

□

Appendix B: Proof of Lemma 3

We need the following integrals; assume $k \geq 1$. An integration by parts with $u = \ln^k(z/x)$ and $v = z$ shows that

$$\int_x^1 \ln^k \frac{z}{x}\,dz = z\ln^k \frac{z}{x}\Big|_x^1 - \int_x^1 z\frac{k}{z}\ln^{k-1}\frac{z}{x}\,dz = \ln^k\frac{1}{x} - k\int_x^1 \ln^{k-1}\frac{z}{x}\,dz\,. \tag{21}$$

By substitution,

$$\int_x^1 \frac{1}{z}\ln^k\frac{z}{x}\,dz = \frac{1}{k+1}\ln^{k+1}\frac{z}{x}\Big|_x^1 = \frac{1}{k+1}\ln^{k+1}\frac{1}{x}\,. \tag{22}$$

Finally, by integration by parts with $u = \ln^k(z/x)$ and $v = -1/z$, we have

$$\begin{aligned}\int_x^1 \frac{1}{z^2}\ln^k\frac{z}{x}\,dz &= \frac{-1}{z}\ln^k\frac{z}{x}\Big|_x^1 - \int_x^1 \frac{-1}{z}\frac{k}{z}\ln^{k-1}\frac{z}{x}\,dz \\ &= -\ln^k\frac{1}{x} + k\int_x^1 \frac{1}{z^2}\ln^{k-1}\frac{z}{x}\,dz\,. \end{aligned}\tag{23}$$

Using these three integrals and some simple inductions one can verify that, for $k \geq 0$,

$$\int_x^1 L_k\left(\frac{z}{x}\right)dz \subseteq L_k(x^{-1}) + xL_0(x^{-1})\,, \tag{24}$$

$$\int_x^1 \frac{1}{z}L_k\left(\frac{z}{x}\right)dz \subseteq L_{k+1}(x^{-1})\,, \tag{25}$$

and

$$\int_x^1 \frac{1}{z^2}L_k\left(\frac{z}{x}\right)dz \subseteq L_k(x^{-1}) + x^{-1}L_0(x^{-1})\,. \tag{26}$$

We can now prove Lemma 3 by induction on d. The basis follows from (11). For the induction step, we calculate

$$\begin{aligned}
&\int_x^1 (1-z) F_d\left(\frac{x}{z}\right) dz \\
&\in \int_x^1 (1-z)\left(\frac{2^d}{(d-1)!}\frac{x}{z}\ln^{d-1}\left(\frac{z}{x}\right) + \frac{x}{z}L_{d-2}\left(\frac{z}{x}\right) + \frac{x^2}{z^2}L_{d-1}\left(\frac{z}{x}\right)\right) dz \\
&\subseteq \int_x^1 \left(\frac{2^d}{(d-1)!}\frac{x}{z}\ln^{d-1}\left(\frac{z}{x}\right) + \frac{x}{z}L_{d-2}\left(\frac{z}{x}\right) + \left(x + \frac{x^2}{z} + \frac{x^2}{z^2}\right)L_{d-1}\left(\frac{z}{x}\right)\right) dz \\
&\subseteq \frac{2^d}{d!} x \ln^d x^{-1} + x L_{d-1}(x^{-1}) + x^2 L_d(x^{-1}),
\end{aligned}$$

where we have used (22), (24), (25), and (26). Substituting into (10) completes the induction. □

References

[1] Clark, B.N., Colburn, C.J., Johnson, D.S.: “Unit Disk Graphs,” Discrete Mathematics **86**, 165–177 (1990)

[2] Coffman, E.G., Jr., Lueker, G.S.: An Introduction to the Probabilistic Analysis of Packing and Partitioning Algorithms, Wiley & Sons, New York, (1991)

[3] Gavril, F.: “Algorithms for Minimum Coloring, Maximum Clique, Minimum Covering by Cliques and Maximum Independent Set of a Chordal Graph,” SIAM Journal on Computing **1**, 180–187 (1972)

[4] Justicz, J., Scheinerman, E.R., Winkler, P.M.: “Random Intervals,” Amer. Math. Monthly **97**, 881–889 (1990)

[5] Knuth, D.E.: “Big Omicron and Big Omega and Big Theta,” SIGACT News **8**(2), 18–24 (April–June, 1976)

[6] Rim, C.S., Nakajima, K.: “On Rectangle Intersection and Overlap Graphs,” IEEE Trans on Circuits and Systems—I: Fundamental Theory and Applications **42**(9), 549–553 (September 1995)

[7] Rose, D.J., Tarjan, R.E., Lueker, G.S.: “Algorithmic Aspects of Vertex Elimination on Graphs,” SIAM Journal on Computing **5**(2), 266–283 (1976)