COMBINATORICA
Akadémiai Kiadó — Springer-Verlag

COMBINATORICA 9 (2) 1989 161—187

TIGHT BOUNDS FOR MINIMAX GRID MATCHING WITH APPLICATIONS TO THE AVERAGE CASE ANALYSIS OF ALGORITHMS

T. LEIGHTON and P. SHOR

Received 12 June, 1986

The minimax grid matching problem is a fundamental combinatorial problem associated with the average case analysis of algorithms. The problem has arisen in a number of interesting and seemingly unrelated areas, including wafer-scale integration of systolic arrays, two-dimensional discrepancy problems, and testing pseudorandom number generators. However, the minimax grid matching problem is best known for its application to the maximum up-right matching problem. The maximum up-right matching problem was originally defined by Karp, Luby and Marchetti—Spaccamela in association with algorithms for 2-dimensional bin packing. More recently, the up-right matching problem has arisen in the average case analysis of on-line algorithms for 1-dimensional bin packing and dynamic allocation.

In this paper, we solve both the minimax grid matching problem and the maximum up-right matching problem. As a direct result, we obtain tight upper bounds on the average case behavior of the best algorithms known for 2-dimensional bin packing, 1-dimensional on-line bin packing and on-line dynamic allocation. The results also solve a long-open question in mathematical statistics.

1. Introduction

Consider a square with area N in the plane that contains N *grid points* arranged in a regularly spaced $\sqrt{N}\times\sqrt{N}$ array and N *random points* located independently and randomly according to the uniform distribution on the square. For example, see Figure 1.

For any particular set of N random points $\mathcal{P}$, let $L(\mathcal{P})$ denote the minimum length such that there exists a perfect matching of the (random) points in $\mathcal{P}$ to the grid points in the square for which the distance between every pair of matched points is at most $L(\mathcal{P})$. In other words, $L(\mathcal{P})$ is the minimum over all perfect matchings of the maximum distance between any pair of matched points, or more simply the *minimax matching length* for $\mathcal{P}$. As an example, Figure 2 illustrates two matchings, one that achieves the minimax matching length and one that does not.

Algorithmically speaking, the problem of computing the minimax matching length for any set $\mathcal{P}$ is relatively straightforward, and is not of concern in this paper. Nor are we interested in the worst case value of $L(\mathcal{P})$ over all $\mathcal{P}$, which is trivially $\Theta(\sqrt{N})$. Rather, we are interested in the *expected* value of $L(\mathcal{P})$ for random $\mathcal{P}$.

This research was supported by Air Force Contracts AFOSR—82—0326 and AFOSR—86—0078, NSF Grant 8120790, and DARPA contract N00014—80—C—0326. In addition, Tom Leighton was supported by an NSF Presidential Young Investigator Award with matching funds from Xerox and IBM.

AMS subject classification (1980): 05C70, 60C05, 60K30, 68K25, 68R05

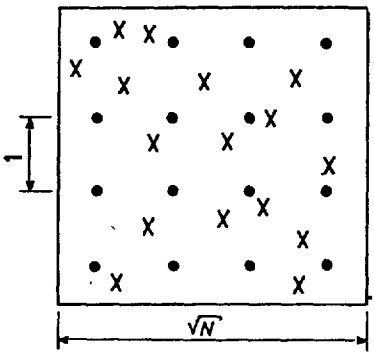

Fig. 1. A square with area N containing a regularly spaced $\sqrt{N}\times\sqrt{N}$ grid (denoted with black dots) and N uniformly distributed random points (denoted with x's)

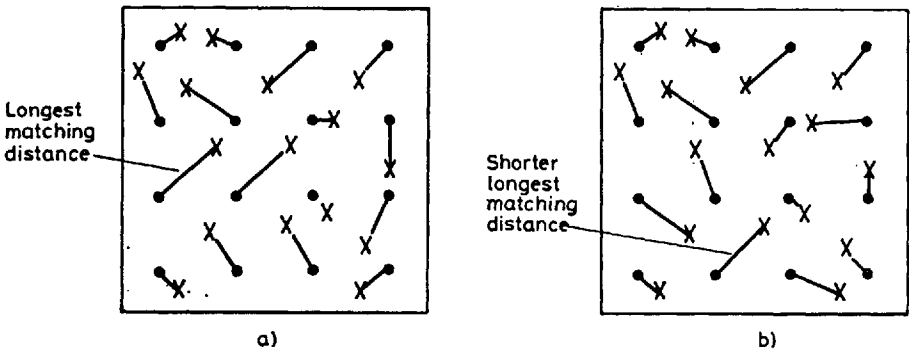

Fig. 2. Two possible matchings for a set of random points, one which achieves the minimax matching length (b) and one which does not (a)

Initially, one might hope that the expected value of $L(\mathscr{P})$ is a constant independent of N since every random point is within distance $\sqrt{2}/2$ of a grid point. With high probability, however, the same is not true for every grid point. In particular, it is not difficult to show that with probability exceeding $1-1/N$ there is a circular region with $\Theta(\log N)$ area in the square that does not contain any random points at all, and hence $L(\mathscr{P})\geq\Omega(\sqrt{\log N})$ with high probability. For example, see Figure 3.

Although the *minimax grid matching problem* (that of determining the expected value of $L(\mathscr{P})$ for random $\mathscr{P}$), has not to our knowledge been directly considered before, Leighton and Leiserson [13] and Ajtai, Komlós and Tusnády [1] considered similar problems, and developed probabilistic divide-and-conquer techniques that show $L(\mathscr{P})\leq O(\log N)$ with high probability. The resulting $\Theta(\sqrt{\log N})$ gap in the bounds for the expected value of $L(\mathscr{P})$ remained for some time until

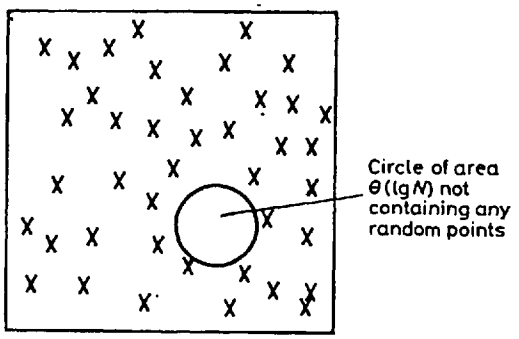

Fig. 3. Collection of N random points $\mathcal{P}$ for which there is a circle of area $\Theta(\log N)$ not containing any random points. The grid point (not shown here) nearest to the center of this circle is $\Omega(\sqrt{\log N})$ away from every random point and hence, $L(\mathcal{P}) \geq \Omega(\sqrt{\log N})$ for this example

Shor [21, 22] proved that $L(\mathcal{P}) \geq \Omega(\log^{3/4} N)$ with *very high probability** (i.e., with probability exceeding $1-1/N^{\alpha}$ where $\alpha=\Omega(\sqrt{\log N})$). Shor's result was substantially more difficult than the rather elementary $\Omega(\sqrt{\log N})$ bound, and suggested that the minimax grid matching problem possessed a deeper structure than one might initially realize.

In this paper, we complete the asymptotic analysis of the minimax grid matching problem by proving that $L(\mathcal{P}) \leq O(\log^{3/4} N)$ with very high probability. Hence $L(\mathcal{P})=\Theta(\log^{3/4} N)$ with very high probability, and the expected value of $L(\mathcal{P})$ is $\Theta(\log^{3/4} N)$.

Aside from being an interesting combinatorial problem, the improved upper bound for typical values of $L(\mathcal{P})$ has important implications for the average case behavior of a wide variety of algorithms. For example, as a direct consequence of this result, we now know that the 1-dimensional Best Fit bin packing algorithm wastes $\Theta(\sqrt{N}\log^{3/4} N)$ space with very high probability when packing N items with sizes determined by identical and independent distributions that are symmetric about 1/2 in [0, 1]. Similar consequences hold for the Karp—Luby—Marchetti—Spaccamela 2-dimensional bin packing algorithm and the Coffman—Leighton Best Fit Aligned dynamic allocation algorithm. Our result also has implications for the maximum up-right matching problem, a long open discrepancy problem, wafer-scale integration of systolic arrays, and potentially for testing pseudorandom number generators.

Establishing the connection between the minimax grid matching problem and the wide variety of seemingly unrelated problems just mentioned has, of course, required a great deal of work by many researchers. Hence, it is not possible to provide detailed explanations for all of the connections in this paper. Instead, we include only the relevant definitions and statements of results for each connection, referring the interested reader to other sources for details.

* Henceforth, we will reserve the phrase "with high probability" to mean "with probability exceeding $1-1/N^{\alpha}$ for any constant $\alpha \geq 1$", and the phrase "with *very* high probability" to mean "with probability exceeding $1-1/N^{\alpha}$ where $\alpha=\Omega(\sqrt{\log N})$.

The remainder of the paper is divided into six sections. In Section 2, we describe some related combinatorial problems. In particular, we discuss the maximum up-right matching problem and several 2-dimensional discrepancy problems. In Section 3, we discuss the application of the minimax grid matching problem to bin packing, dynamic allocation, wafer-scale integration and testing pseudorandom number generators. The proof of the $O(\log^{3/4} N)$ upper bound for minimax grid matching is contained in Section 4. We conclude with some remarks, acknowledgements and references in Sections 5—7.

2. Related Combinatorial Problems

In this section, we discuss several combinatorial problems that are closely related to the minimax grid matching problem. These problems are interesting in their own right, but are included here mostly for their role in the applications described in Section 3 and the proof contained in Section 4. We commence with the maximum up-right matching problem in Section 2.1 and finish with a discussion of two-dimensional discrepancy problems in Section 2.2.

2.1. The Maximum Up-Right Matching Problem

Consider a square with area N in the plane that contains N random *pluses* and N random *minuses* located independently and randomly according to the uniform distribution on the square. Given a set of N pluses and N minuses $\mathscr{P}^{\pm}$, an *up-right matching* on $\mathscr{P}^{\pm}$ is a one-to-one matching of pluses to minuses such that every plus is either unmatched or is matched to a single minus that lies above and to the right of the plus. A *maximum up-right matching* for $\mathscr{P}^{\pm}$ is an up-right matching which minimizes the number of unmatched points. For example, Figure 4 illustrates two up-right matchings. One is maximum, the other is not.

Let $U(\mathscr{P}^{\pm})$ denote the number of unmatched pluses in a maximum up-right matching for $\mathscr{P}^{\pm}$. For example, $U(\mathscr{P}^{\pm})=1$ for the collection of pluses and minuses shown in Figure 4. As with minimax grid matching, the algorithmic problems of determining $U(\mathscr{P}^{\pm})$ and finding a maximum up-right matching for a particular collection of pluses and minuses are elementary and not of interest here. Rather, we are interested in determining the expected and/or high probability values of $U(\mathscr{P}^{\pm})$ for random $\mathscr{P}^{\pm}$.

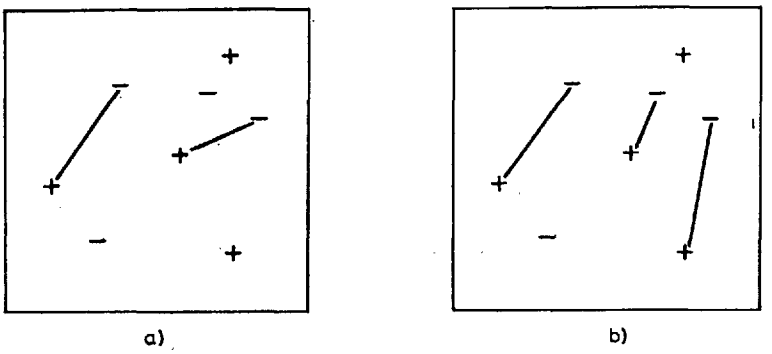

Fig. 4. Two up-right matchings. Only the matching in (b) is maximum

It is not difficult to show that the expected value of $U(\mathscr{P}^{\pm})$ is $\Omega(\sqrt{N})$. Simply consider the pluses and minuses that lie in the upper half of the square. Since pluses in the upper half of the square can only be matched to minuses in the upper half of the square, the number of upper-half pluses less the number of upper half minuses is always a lower bound on the number of unmatched pluses in a maximum up-right matching. The numbers of upper-half pluses and minuses are governed by simple binomial distributions with mean $N/2$ and variance $N/4$. Hence, the number of upper-half pluses will exceed the number of upper-half minuses by at least $\Omega(\sqrt{N})$ with some constant positive probability. Thus, the expected value of $U(\mathscr{P}^{\pm})$ is at least $\Omega(\sqrt{N})$.

Karp, Luby and Marchetti—Spaccamela [10] were the first to explicitly define the *maximum up-right matching problem* (i.e., the problem of determining the expected value of $U(\mathscr{P}^{\pm})$). By applying Hall's Theorem [9] and analyzing the dual problem, they showed that $U(\mathscr{P}^{\pm}) \leqq O(\sqrt{N}\log N)$ with high probability. By applying an elegant argument of Ajtai, Komlós and Tusnády [1], they also showed that $U(\mathscr{P}^{\pm}) \geqq \Omega(\sqrt{N}\sqrt{\log N})$ with high probability.

The resulting $\Theta(\sqrt{\log N})$ gap in the bounds remained until Shor [21, 22] proved that $U(\mathscr{P}^{\pm}) \geqq \Omega(\sqrt{N}\log^{3/4} N)$ with *very* high probability. Not surprisingly, this result combined with the $O(\sqrt{N}\log N)$ upper bound led to the conjecture that $U(\mathscr{P}^{\pm}) = \Theta(\sqrt{N}\log N)$ with high probability. As it turns out, however, this is not the case. Instead, it is the $\Omega(\sqrt{N}\log^{3/4} N)$ lower bound that is correct. In fact, we prove in this paper that $U(\mathscr{P}^{\pm}) = \Theta(\sqrt{N}\log^{3/4} N)$ with *very* high probability.

Judging from the similarity between the claimed bounds for the minimax grid matching problem and the maximum up-right matching problem, it should not be surprising that the two problems are closely related. In fact, any very high probability upper bound on $L(\mathscr{P})$ can be transformed into a very high probability upper bound for $U(\mathscr{P}^{\pm})$ by simply multiplying by $\Theta(\sqrt{N})$. To see this, consider an up-right matching problem with N random pluses $\mathscr{P}^{+}$ and N random minuses $\mathscr{P}^{-}$. (Here $\mathscr{P}^{\pm} = \mathscr{P}^{+} \cup \mathscr{P}^{-}$.) Let $d(N)$ be a very high probability upper bound on $L(\mathscr{P})$. Then, with very high probability, $L(\mathscr{P}^{+}) \leqq d(N)$ and $L(\mathscr{P}^{-}) \leqq d(N)$. In other words, the pluses and minuses can each be matched to N regularly spaced grid points so that each grid point is matched to a plus and a minus that are within distance at most $d(N)$. Next form a matching on $\mathscr{P}^{\pm}$ by matching the plus identified with grid point (i,j) to the minus identified with grid point $(i+2d(N), j+2d(N))$ for $\{(i,j) | 1 \leqq i \leqq \sqrt{N}-2d(N),\ 1 \leqq j \leqq \sqrt{N}-2d(N)\}$. For example, see Figure 5.

From inspection of Figure 5, it is clear that the procedure just described forms an up-right matching. The only pluses not matched by the procedure are those identified with grid points in the topmost $2d(N)$ rows and the rightmost $2d(N)$ columns. As there are less than $4d(N)\sqrt{N}$ such points, we can conclude that $U(\mathscr{P}^{\pm}) \leqq O(\sqrt{N}d(N))$ with very high probability. Hence the claimed very high probability $O(\sqrt{N}\log^{3/4} N)$ bound for $U(\mathscr{P}^{\pm})$ immediately follows from the very high probability $O(\log^{3/4} N)$ bound for $L(\mathscr{P})$ proved in Section 4.

It is conceivable that a symmetric condition is also true: namely, that a very high probability upper bound for $U(\mathscr{P}^{\pm})$ can be easily transformed into a very high probability upper bound for $L(\mathscr{P})$. We doubt that this is the case, however, since the minimax grid matching problem appears (at least to us) to be fundamentally

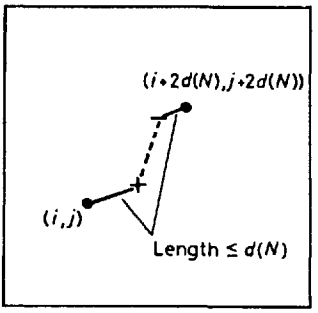

Fig. 5. The matching of a plus to a minus via association with grid points. Since the solid matching lines have distance at most $d(N)$, the minus is above and to the right of the plus

harder and more general than the maximum up-right matching problem*. Nevertheless, the upper bound for $U(\mathscr{P}^{\pm})$ obtained from $L(\mathscr{P})$ is tight up to constant factors.

2.2. Two-Dimensional Discrepancy Problems

Discrepancy problems have a long and rich history in mathematics and statistics [6, 19, 23, 25]. In fact, much of probability theory and mathematical statistics involves estimating the likelihood of certain discrepancies (or deviations from the norm) occurring in random data. As a familiar example, consider an unbiased random walk of length N. It is a well-known consequence of the reflection principal that largest deviation from the origin during the walk is expected to be $\Theta(\sqrt{N})$. Hence, the largest "swing" observed in any subwalk of a walk of length N is also expected to be $\Theta(\sqrt{N})$.

This basic result has many consequences, and can be stated more generally as an example of a 1-dimensional discrepancy problem. For example, consider a fixed interval of length N that contains N random random points selected according to independent and identical uniform distributions. Define the *discrepancy* of any subinterval to be the absolute value of the difference between the expected number of points contained in the interval (i.e., the length of the interval) and the actual number of points contained in the interval. Using the reflection principle, it is not difficult to show that the expected maximum discrepancy over all subintervals is $\Theta(\sqrt{N})$. Moreover, with high probability every subinterval has discrepancy at most $O(\sqrt{L}\sqrt{\log N}+\log N)$, where L is the length of the subinterval.

In this paper, we are interested in two-dimensional generalizations of the preceding one-dimensional results on intervals. Several different generalizations have been studied, depending on how the subintervals are generalized. For example, consider a square with *area* N that contains N random points generated according to independent and identical uniform distributions. For any region R of the square, define the discrepancy of the region $\Delta(R)$ to be the absolute value of the difference

* This is true in higher dimensions. In d-dimensional space, for $d \geq 3$, Karp et al. [10] prove that for up-right matching, the log factor disappears and $U(\mathscr{P}^{\pm})=\Theta(N^{1-1/d})$; whereas there is an easy lower bound of $\Omega(\log^{1/d} N)$ for d-dimensional minimax grid matching. Hence the relationship between $U(\mathscr{P}^{\pm})$ and $L(\mathscr{P})$ breaks down in higher dimensions.

between the expected number of points contained in R (i.e., the area of R) and the actual number of points in R. In a classic paper, Kiefer [11] showed that the expected maximum discrepancy of any oriented rectangle is $\Theta(\sqrt{N})$ for fixed N. Generalizing the law of the iterated logarithm, he also showed that the limsup of the maximum discrepancy of any oriented rectangle over all $N \to \infty$ is $\Theta(\sqrt{N}\sqrt{\log\log N})$ with probability $1-o(1)$. These results were later extended by Philipp [16] who proved the same bounds for arbitrary convex regions, and by Leighton [12] who showed that with high probability, every convex region has discrepancy $O(\sqrt{A}\sqrt{\log N}+\log N)$ where A is the area of the region.

Of greatest interest in this paper, however, is the work on maximum discrepancies of *up-right* regions (also known as *lower layers*) in the square. Up-right regions are defined by monotone decreasing curves. For example, Figure 6 illustrates a rectilinear up-right region with $N=8$, area 2 and discrepancy 1.

By applying Hall's Matching Theorem in the usual way, it is not difficult to see that the up-right discrepancy problem is closely related to the up-right matching problem. More precisely, if we consider a square with area N that contains N random pluses $\mathscr{P}^+$ and N random minuses $\mathscr{P}^-$, and we define $\Delta^{\pm}(R)$ to be the number of pluses in R less the number of minuses in R, then the maximum value of $\Delta^{\pm}(R)$ over all up-right regions is precisely the number of unmatched pluses in a maximum up-right matching for $\mathscr{P}^{\pm}=\mathscr{P}^+\cup\mathscr{P}^-$. For random distributions, the maximum values of $\Delta^{\pm}(R)$ and $\Delta(R)$ are within constant factors of each other with very high probability, so we can conclude from the bounds on $U(\mathscr{P}^{\pm})$ cited in Section 1 that the maximum value of $\Delta(R)$ over all up-right regions is $\Theta(\sqrt{N}\log^{3/4} N)$ with very high probability.

Asymptotically characterizing the expected maximum value of $\Delta(R)$ over all up-right regions has been an open problem in mathematical statistics for some time. Perhaps Blum [3] was the first to study the problem, proving a $o(N)$ upper bound in 1955. This was later improved to $O(N^{3/4})$ by Steele [23] in 1977 and then to $O(\sqrt{N}\log N(\log\log N)^{\alpha})$ for some constant α by Philipp [17]. The first nontrivial lower bound was proved by Dudley [7] in 1982. He showed that the expected maximum value of $\Delta(R)$ for all up-right regions is at least $\Omega(\sqrt{N}\sqrt{\log N}/(\log\log N)^{1/2+\epsilon})$ for any constant $\varepsilon>0$. Karp, Luby and Marchetti—Spaccamela [10] were the first

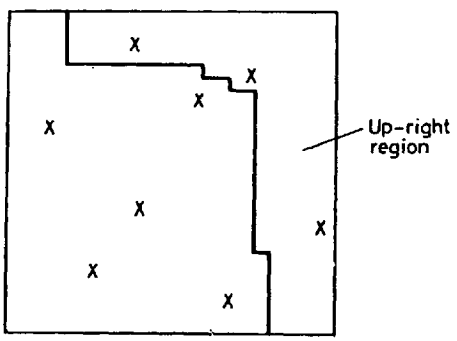

Fig. 6. An up-right region with $N=8$, area 2 and discrepancy 1

to bring the problem into the computer science community, proving an upper bound of $O(\sqrt{N}\log N)$ and (using an elegant argument of Ajtai, Komlós and Tusnády [1]) a lower bound of $\Omega(\sqrt{N}\sqrt{\log N})$. The $O(\sqrt{N}\log N)$ upper bound also follows directly from prior work of Leighton and Leiserson [13] and Ajtai, Komlós and Tusnády [1]. The lower bound was later improved to $\Omega(\sqrt{N}\log^{3/4} N)$ by Shor [21, 22], who independently developed a method similar to that followed by Dudley. The results in this paper provide final resolution to the problem by establishing a very high probability $O(\sqrt{N}\log^{3/4} N)$ upper bound on the maximum discrepancy of any up-right region.

A priori, it is not clear why we should restrict ourselves to either convex or up-right regions of the square when generalizing the 1-dimensional discrepancy problem. For example, why not consider any simply connected region? (A region is said to be *simply connected* if it is connected and contains no holes.) The answer, of course, is that we can always construct a simply connected region that contains all N points, but that has area arbitrarily close to zero. Hence, the worst possible discrepancy is always achieved.

Not all is lost, however, since the discrepancy of a simply connected region can be bounded in terms of its perimeter. In fact, we will prove in Section 4 that with very high probability, the discrepancy of every simply connected region is at most $O(p\log^{3/4} N+\log^{3/2} N)$, where p is the perimeter of the region. Note that this result naturally generalizes the bound on discrepancies of up-right regions since $p=O(\sqrt{N})$ for any up-right region. More importantly, however, the result will be sufficient to prove the $O(\log^{3/4} N)$ upper bound for minimax grid matching. The proof of this fact is not difficult and again uses Hall's Theorem. The details are deferred until Section 4.

Although the topic is not of direct concern in this paper, two-dimensional discrepancy problems have also been studied in a worst-case setting. The most notable result in this area is due to Schmidt [20] who resolved a decades old open question by proving that no matter how N points are arranged in a square of area N, there is always an oriented rectangle in the square which contains $\Omega(\log N)$ fewer or $\Omega(\log N)$ more points than its area. In other words, there is always an oriented rectangle with discrepancy $\Omega(\log N)$. Constructions which achieve a maximum discrepancy of $O(\log N)$ for this problem were known long ago [14, 24] and have recently been rediscovered in the computer science literature [8, 15]. Curiously, it is still not known how to prove either the "$\Omega(\log N)$ fewer" or the "$\Omega(\log N)$ more" result individually. Even more startling is the fact that there is always an oriented right triangle in the square with discrepancy $\Omega(N^{1/4})$ [10]!

3. Applications

In this section, we briefly describe the applications of our work to problems involving bin packing, dynamic allocation, wafer-scale integration, and testing pseudorandom number generators. The applications range widely in difficulty, and some are quite elegant. Due to space limitations, however, we will mostly just refer the reader to relevant papers in the literature.

3.1. Wafer-Scale Integration

The minimax grid matching problem was first studied in the context of wafer-scale integration of two-dimensional systolic arrays [13]. Only later was its association with up-right matching noticed.

When constructing two dimensional arrays on a single wafer, the designer is typically presented with a regular $\sqrt{N}\times\sqrt{N}$ array of chip-size cells, some fraction of which are functional. The task is to connect the functional cells into a smaller square array in a way that minimizes the length of the longest wire needed to connect adjacent cells in the array. For example, Figure 7 illustrates the connection of a 3×3 array of functional cells from a 4×4 array containing 9 functional cells.

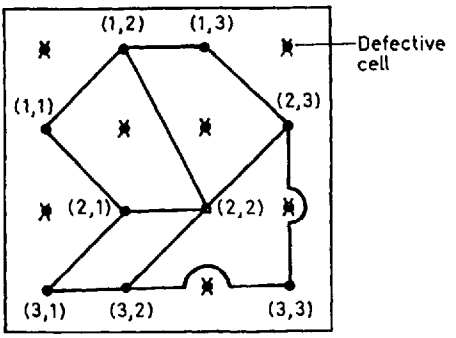

Fig. 7. Connection of a 3×3 array of functional cells

Unfortunately, the designer's task is *NP*-hard in general. However, the impact of this difficulty is mitigated by two practical considerations. First, the defective cells tend to be located randomly. Second, an algorithm need only work well for most (not all) wafers. Hence, an algorithm which works with high probability for randomly located faults is very acceptable in practice.

The minimax grid matching problem suggests exactly such an algorithm. First match the functioning cells one-to-one with an imaginary regularly spaced grid of the same cardinality in a way that minimizes the longest matching length. By the work in this paper, this length will be $O(\log^{3/4} N)$ with very high probability.

The one-to-one association of functioning cells with imaginary grid points assigns a label (i,j) to each functioning cell. It remains only to hook up cells whose labels differ by one in precisely one coordinate. Since the distance between any two adjacent cells (say (i,j) and $(i,j+1)$, for example) is at most the distance between (i,j) and its matching grid point (this is $O(\log^{3/4} N)$ with very high probability) plus the distance between $(i,j+1)$ and its matching grid point (also $O(\log^{3/4} N)$ with very high probability) plus the distance between the two matching grid points (constant), every wire will have length $O(\log^{3/4} N)$ with very high probability.

Of course, this analysis assumes nearly widthless wires, but for some applications this is reasonable. Previously, the best upper bound proved for this problem was $O(\log N)$ by Leighton and Leiserson [13]. The best lower bound known remains the trivial bound of $\Omega(\sqrt{\log N})$.

3.2. Two-Dimensional Bin Packing

The maximum up-right matching problem was first defined in the context of two-dimensional bin packing [10]. In fact, the algorithm proposed by Karp, Luby and Marchetti—Spaccamela for two-dimensional bin packing specifically uses the maximum up-right matching algorithm as a subroutine. In what follows, we briefly describe the problem and the results. We refer the interested reader to [10] for a description of the algorithm and the details of the analysis.

Given a collection of N items from $[0, 1]\times[0, 1]$, the two-dimensional bin packing problem is to pack the N items into the minimum possible number of unit-square bins. Of course, the two-dimensional bin packing problem is NP-complete. However, when the N items are chosen independently and uniformly from $[0, 1]\times$ $\times[0, 1]$, the Karp—Luby—Marchetti—Spaccamela algorithm uses $\frac{N}{4}+\Theta(W)$ bins and wastes $\Theta(W)$ space with very high probability, where $W=\Theta(\sqrt{N}\log^{3/4} N)$ is the anticipated number of unmatched pluses in a random maximum up-right matching problem. This compares favorably with the $\Omega(\sqrt{N})$ lower bound on expected wasted space for any two-dimensional bin packing algorithm. Whether or not there is an algorithm which achieves $o(\sqrt{N}\log^{3/4} N)$ expected wasted space remains unknown.

3.3. One-Dimensional Bin Packing

Somewhat surprisingly, the apparently two-dimensional up-right matching problem also arises in the context of one-dimensional bin packing. In particular, Shor [21, 22] showed that the space wasted by the Best Fit bin packing heuristic when packing N items uniformly selected from $[0, 1]$ is $\Theta(W)$ with high probability, where again $W=\Theta(\sqrt{N}\log^{3/4} N)$ is the anticipated number of unmatched pluses in a random maximum up-right matching problem. Moreover, the same upper bound holds when the items are selected from any distribution on $[0, 1]$ that is symmetric about 1/2.

The Best Fit algorithm is one of the most common on-line algorithms used in practice. The algorithm packs each item as it "arrives" (never looking ahead at subsequent items) in the fullest bin into which it fits. Although better off-line algorithms are known (First Fit Decreasing achieves the optimal bound of $\Theta(\sqrt{N})$ expected wasted space [2]), no better on-line algorithm is known. In fact, Shor [21, 22] used a related matching problem to prove an $\Omega(\sqrt{N}\sqrt{\log N})$ lower bound on the expected wasted space of any on-line algorithm, assuming that the number of items N is not known in advance.

3.4. Dynamic Allocation

The up-right matching problem also arises in the context of dynamic algorithms for file storage. In particular, the Coffman—Leighton [4] Best-Fit-Aligned (BFA) algorithm for dynamic allocation wastes $\Theta(W)$ space with very high probability where W is the expected number of unmatched pluses in an N-point random up-right matching problem, when files arrive and depart according to a Poisson process where N is the expected number of files in storage at any point in time. Since the amount of used (i.e., occupied) space in storage at any time is easily bounded, this means that the capacity of the storage device can be predetermined so as to almost never waste more than $O(\sqrt{N}\log^{3/4} N)$ space and so as to almost

never need compaction. Moreover, the results hold for any distribution of file sizes in [0, 1]. These results are much better than those for the worst-case algorithms used in practice, which commonly waste $\Theta(N)$ space and spend a constant fraction of their time in amortized compaction.

Although, we do not have room to describe the BFA algorithm in detail here, we can mention a simpler balls and boxes problem with related performance. Consider an infinite collection of boxes in a line 1, 2, 3, ... Assume initially that precisely the first N boxes are filled. At each step, remove one of the balls at random, and then try to insert a new ball into one of the first N boxes (chosen at random). Boxes can contain at most one ball each, so if we are unable to insert the new ball into its desired box, then insert the ball into the leftmost empty box to the right of the originally desired box. Continue this process indefinitely.

It is clear that the system just described always contains exactly N balls, but the position of the rightmost ball will vary with time. The question of interest in dynamic allocation involves the expected location of the rightmost ball. This, is, of course, identical to N plus the number of empty boxes before the rightmost ball (i.e., the wasted space). By a rather complicated reduction to maximum up-right matching, Coffman and Leighton [4] showed that the rightmost position at any time $T \cong N \log N$ is $N + \Theta(\sqrt{N} \log^{3/4} N)$ with very high probability. Without up-right matching to rely upon, this result might well have been impossible to obtain.

3.5. Testing Pseudorandom Number Generators

All of the applications discussed thus far are directly tied to one of the matching problems. The corresponding discrepancy problems also have applications, but most are to problems in mathematical statistics (e.g., see [23, 26]) and are of limited interest to computer scientists. There is one potential application that could be of interest, however: using the minimax grid matching problem to test pseudorandom number generators. As an example, consider the plot of "random numbers" from the IBM PC random number generator shown in Figure 8.

The points in Figure 8 (which is copied from [18]) were obtained by taking consecutive pairs of "random numbers" and using them as x and y coordinates. In

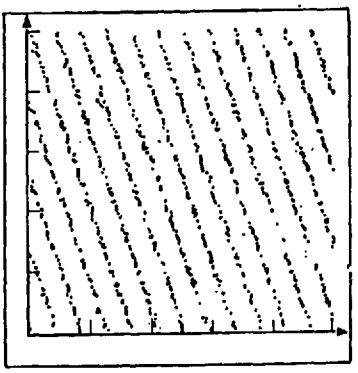

Fig. 8. Plot of 2000 random pairs $(x_i, x_{i+1}), (x_{i+4}, x_{i+5}), \ldots$ *generated from the IBM PC Random Number Generator with IBM PC Basic. (Taken from* [18].)

this case, it is readily apparent that the data are not random. One way of detecting this formally is to observe that there are simply connected regions with area A and perimeter p that contain far more than $A+\Theta(p\log^{3/4}N)$ points. (Just draw the boundary around one of the diagonal clusters.) Hence, by the arguments of Sections 2.2 and 4, the minimax grid matching length for this set of points would be far in excess of $\Theta(\log^{3/4}N)$. Hence, we would have formal grounds for deciding that the IBM PC random number generator is faulty.

There are several good reasons to use the minimax grid matching length as a barometer for randomness. First, it is a polynomial time test which in some sense is doing an exponential amount of work. Because of Hall's Theorem, we know that if the minimax matching length is d, then every simply connected region in the square (of which there are more than an exponential number) is verified to contain about the right number of random points (i.e., between $A-\Theta(pd)$ and $A+\Theta(pd)$). Hence large discrepancies in odd shaped regions will be readily detected, even if they are not easily characterizable or readily apparent to the naked eye. Second, the test will also identify generators that are "too regular". If the numbers are too regular, they will form a nearly perfect grid resulting in a minimax matching length that is too small. Third, the distribution of the minimax matching length is very sharply peaked at $\Theta(\log^{3/4}N)$. The probability of deviating from the peak by even a constant factor is vanishingly (more than polynomially) small. Hence small deviations from the peak would provide high confidence that the pseudorandom number generator is not random.

The major drawback to using the minimax matching length as a statistical test is that we have not been able to completely characterize its distribution. This would have to be done in order for the test to be usable for fixed values of N. In the mean time, the test can only be used as a well-motivated heuristic.

The idea of using a matching problem as a test for a pseudorandom number generator is not completely new. The same idea (but for a different problem) motivated the work of Ajtai, Komlós and Tusnády [1]. Unfortunately, they also were unable to completely characterize the distribution around the peak. More generally, related methods can be found in the literature on spectral and Fourier tests [5].

4. Proofs

In this section, we prove that the minimax grid matching length $L(\mathcal{P})$ for N random points $\mathcal{P}$ is $O(\log^{3/4}N)$ with very high probability. The section is divided into four subsections. In Section 4.1, we formally convert the minimax grid matching problem into a dual discrepancy problem with a straightforward application of Hall's Theorem. We prove the necessary bounds for the discrepancy problem in Section 4.3. Section 4.2 provides some intuition and motivation for the rather complicated proof in Section 4.3, and gives some insight into why the eventual answer is $\Theta(\log^{3/4}N)$ instead of $\Theta(\sqrt{\log N})$ or $\Theta(\log N)$. In Section 4.4, we extend the result of Section 4.3 to arbitrary regions in the plane. The result is not needed for the rest of the paper, but is mathematically more interesting and natural.

4.1. Conversion to the Dual Discrepancy Problem

In the next few subsections, we will restrict our attention to simply connected regions with a special rectilinear form. In particular, we define a partition Γ of the square with area N into $\frac{N}{\log^{3/2} N}$ subsquares, each with side length $\log^{3/4} N$. In Section 4.3, we will prove that there is a constant $c \geqq 0$ such that with very high probability, the discrepancy of every simply connected region whose boundary lies along the edges of Γ is at most $cp \log^{3/4} N$ where p is the perimeter of the region. As a consequence, we can conclude that with very high probability, the discrepancy of every region (not necessarily connected or simple) whose boundary lies along the edges of Γ is at most $cp \log^{3/4} N$ where p is the perimeter of the region. (This corollary is easily proved by decomposing an arbitrary region into the sum and difference of simply connected regions.) In this subsection, we show how to use this result to prove that the minimax grid matching length $L(\mathscr{P})$ for N random points $\mathscr{P}$ is at most $d = O(\log^{3/4} N)$ with very high probability.

By Hall's Theorem [9], it is sufficient to show that for every set $\mathscr{A} \subset \mathscr{P}$ with x random points, there are at least x grid points within distance d of $\mathscr{A}$. To do this, we define a coarser partition Γ' of the square into $\frac{4N}{d^2}$ subsquares with side length $\frac{d}{2}$ where (for now) $d = 8c \log^{3/4} N$. For any subset $\mathscr{A} \subset \mathscr{P}$, let the region R consist of all the subsquares in Γ' containing a point of $\mathscr{A}$. Let R' be the slightly larger region formed from R by adding an isosceles right triangle with hypothenuse $\frac{d}{2}$ to every $\frac{d}{2}$-length segment along the perimeter of R. For example, see Figure 9. (For now, we overlook the special case when the boundary of R coincides with the boundary of the overall square.)

By construction, every grid point of R' is within distance d of some point in $\mathscr{A}$. Hence it remains only to show that the number of grid points in R' is at least as large as the number of random points in $\mathscr{A}$. This can be accomplished by observing that the number of grid points in R' is at least the area of R' which is

$$Area(R) + \frac{Per(R)}{d/2} \cdot \frac{1}{4}\left(\frac{d}{2}\right)^2 = Area(R) + Per(R)\, c \log^{3/4} N$$

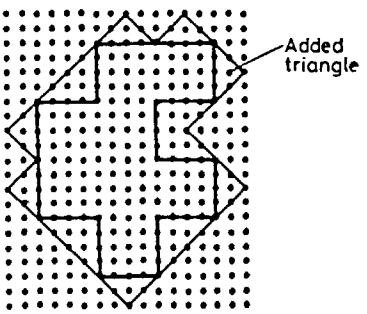

Fig. 9. The extension of R to R'

whereas

$$|\mathscr{A}| \leqq Area(R)+\Delta(R)$$
$$\leqq Area(R)+Per(R)c\log^{3/4}N.$$

This completes the proof. (Accounting for the case when the boundary of R coincides with the boundary of the square only multiplies d by a constant factor since at least one of R and $\bar{R}$ has at most a constant fraction of its perimeter on the boundary of the square.)

4.2. Decomposition Into Triangles — The Intuition

In Section 4.3, we will prove that with very high probability, every simply connected region whose boundary lies along the edges of Γ has discrepancy at most $O(p\log^{3/4}N)$ where p is the perimeter of the region. In what follows we provide some intuition for the result by proving a much simpler result with the same $\log^{3/4}n$ bound. This simpler result deals with the decomposition of a polygonal region into triangles and is interesting in its own right. The result does not imply the desired discrepancy bound, but it does give some of the basic ideas behind it. In Section 4.3, these ideas are obscured by technical details.

Theorem 1. *Any polygonal region R with n vertices and perimeter p can be decomposed into a sum and difference of triangles T_j such that* $\sum_j \sqrt{Area(T_j)}=O(p\log^{3/4}n)$.

Proof. The algorithm we use to decompose the region into a sum and difference of triangles is simple. At each step, we reduce the number of vertices of the polygon by half by cutting off $n/2$ consecutive triangles. (See Figure 10.) In particular, we cut off $n/2$ triangles formed by pairs of adjacent sides. We concentrate on what happens when we remove the jth triangle, which is formed, say, by vertices v_{i-1}, v_i and v_{i+1}. Let the distance between v_{i-1} and v_{i+1} be d_j, and let the two edges $v_{i-1}v_i$ and v_iv_{i+1} have lengths e_i and e_{i+1}. Let $c_j=e_i+e_{i+1}$. In removing the triangle $v_{i-1}v_iv_{i+1}$, we remove two edges totalling length c_j, and add an edge of length d_j. Thus, we reduce the perimeter by $\delta_j=c_j-d_j$.

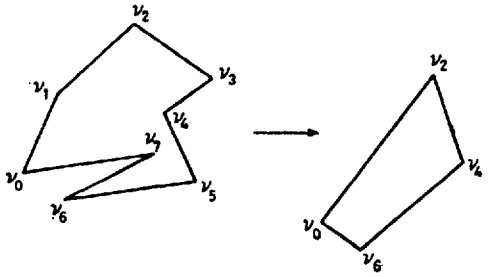

Fig. 10. Removing four triangles

We must show that this algorithm produces a decomposition into triangles such that $\sum_j \sqrt{Area(T_j)}\leqq O(p\log^{3/4}n)$. Removing the jth triangle reduces the perimeter by $\delta_j=c_j-d_j$. The area of this triangle is at most $\frac{1}{4}d_j\sqrt{c_j^2-d_j^2}$, since this area is maximized when $e_i=e_{i+1}$.

We therefore get

$$\sqrt{Area(T_j)} \leqq \frac{1}{2} d_j^{1/2}(c_j^2 - d_j^2)^{1/4}$$

$$= \frac{1}{2} d_j^{1/2}(c_j + d_j)^{1/4}(c_j - d_j)^{1/4}$$

$$\leqq c_j^{3/4}\delta_j^{1/4}.$$

We repeat this step of cutting off half the vertices until we reduce the number of sides to 2, and thus have no area left. At each step the total sum of the c_j's is at most the perimeter p, so since there are $\log n$ steps, we have $\sum_{j=1}^{n-2} c_j \leqq p \log n$. Since δ_j was the change in the perimeter produced by removing the jth triangle, $\sum_{j=1}^{n-2} \delta_j \leqq p$. Thus, we have

$$\sum_{j=1}^{n-2} \sqrt{Area(T_j)} \leqq \sum_{j=1}^{n-2} c_j^{3/4}\delta_j^{1/4}$$

$$\leqq \left(\sum_{j=1}^{n-2} c_j\right)^{3/4} \left(\sum_{j=1}^{n-2} \delta_j\right)^{1/4}$$

$$\leqq p \log^{3/4} n.$$

The second step is a special case of Hölder's inequality. ∎

The expected discrepancy of a triangle T is $\Theta(\sqrt{Area(T)})$. If we could assume that the average discrepancy of a triangle in the decomposition of the polygonal region R was this expected discrepancy, then by Theorem 1 the total discrepancy of the region R would be $O(p \log^{3/4} n)$. Proving this for a general class of regions R is the most difficult part of the paper and is done in Section 4.3.

The decomposition used in the proof of Theorem 1 is basically the one that we will use for the general result. We will bound, with very high probability, the total discrepancy of all the $n/2^k$ triangles removed at the kth stage. We do this by counting the number of possible sets of triangles that could be removed at the kth stage.

Although we have not investigated the matter formally, the bound in Theorem 1 should be tight in the worst case.

4.3. Formally Bounding the Discrepancy — The Proof

In this section, we prove the following central result.

Theorem 2. *Consider a set of N points uniformly and independently distributed in the $\sqrt{N}\times\sqrt{N}$ square, and define Γ to be a grid of squares with edge length $\log^{3/4} N$. Then there is a constant c such that with probability at least $1-N^{-\log^{1/2} N}$, every simply connected region R whose boundary follows Γ has discrepancy $\Delta(R) \leqq cPer(R)\log^{3/4} N$.*

The proof is divided into two sections, one deterministic and one probabilistic In the deterministic section, we show that the discrepancy of any R satisfying the hypothesis can be bounded by the sum of the discrepancies of $2 \log p$

disjoint regions, most of which are drawn from "small enough" classes of "small enough" regions. In the probabilistic part of the proof, we establish very high probability bounds on the discrepancies of the regions in the classes, thus obtaining an upper bound on the discrepancy of any R.

4.3.1. The Deterministic Part of the Proof

The deterministic part of the proof of Theorem 2 consists of showing that there are collections of regions $\mathcal{D}_{i,s,p}$ of the $\sqrt{N}\times\sqrt{N}$ square and a constant c_0 independent of N such that:

1) the area of every region in $\mathcal{D}_{i,s,p}$ is at most $2^{-i}p^2\sqrt{s}$,
2) the number of regions in $\mathcal{D}_{i,s,p}$ is at most $N^2(s\log N)^{2^{i+1}}$, and
3) for any p and any simply connected region R with boundary on Γ and perimeter p in the $\sqrt{N}\times\sqrt{N}$ square, there exist real numbers $s_1, \ldots, s_m$, and regions $D_1, \ldots, D_m$ and $D_1^*, \ldots, D_m^*$ such that
   3a) $D_i, D_i^*\in\mathcal{D}_{i,s_i,p}$ for $1\leqq i\leqq m$,
   3b) $\Delta(R)\leqq\sum_{i=1}^{m}(\Delta(D_i)+\Delta(D_i^*))$, and
   3c) $s_1+\ldots+s_m\leqq c_0$.

The $\mathcal{D}_{i,s,p}$ are defined for values of i that are integers in the range $[1, m]$, where $m=\log p$, values of s that are integer multiples of $1/\log N$ in the range $[1/\log N, c_0]$, and values of p that are integer multiples of $\log^{3/4}N$ in the range $[\log^{3/4}N, N]$. In fact, we will assume without loss of generality that p is a power-of-two multiple of $\log^{3/4}N$.

The regions in a particular $\mathcal{D}_{i,s,p}$ need not be connected or simple, but they will consist of the sum and difference of easily described polygons. Very roughly speaking, the regions in $\mathcal{D}_{i,s,p}$ will correspond to the union of triangles removed in the $(\log n-i)$th halving operation described in Section 4.2, where s is related to the total area of the removed triangles and p is the perimeter of the original region being decomposed. The precise characterization of the $\mathcal{D}_{i,s,p}$ is fairly difficult, and will be described as we go along.

The $\{D_i\}$ and $\{D_i^*\}$ for a region R are derived from a sequence of approximations $R_1, \ldots, R_{m+1}$ to R where $R_1=\emptyset$ and $R_{m+1}=R$. To construct the sequence, we use two different sequences of piecewise linear closed curves. The first sequence has vertices lying on the boundary of R. We call the ith curve of this kind A_i. The A_i are then approximated further by closed curves B_i with vertices on a grid G_i. If B_i is a *simple* closed curve, then R_i is its interior. Although we begin with a simple closed curve, neither the A_i nor the B_i approximations are necessarily simple closed curves, which causes further problems, forcing us to define a region "enclosed" by the curve B_i. This "enclosed" region is R_i, which is not necessarily connected, even though R is connected.

To obtain the approximation A_i, we mark 2^i points at equal distances along the boundary of R. We call these points $a_{i0}, a_{i1}, \ldots, a_{i2^i-1}$. The starting point $a_{i0}=a_0$ is the same for all A_i and is chosen to be a vertex of Γ. We then join these points in order by edges. Half the vertices of A_{i+1} are also vertices of A_i, specifically, $a_{ij}=a_{i+1,2j}$. We let the length of the edge between $a_{i,j-1}$ and a_{ij} be e_{ij}.

The curve B_i is obtained by approximating the curve A_i with points on a grid G_i. We let the jth vertex b_{ij} of B_i be the nearest grid point to the jth vertex

a_{ij} of A_i. If several grid points are equidistant to some vertex, we break the tie arbitrarily.

The grid G_i has points spaced evenly at distance $g_i=p/(2^i\sqrt{\log n})$. The grid G_{i+1} is a refinement of G_i, so a fourth of the points of G_{i+1} are also points of G_i. We denote the *edge length* of G_i by $g_i=g_1/2^{i-1}$.

We have now produced the curve B_i. If it is a simple closed curve, then the area inside it is our ith approximation R_i to the region R. Otherwise, we must do some more work. If an area is enclosed by B_i twice or more (i.e., has winding number ≥ 2), we still wish to count each point inside it at most once when calculating the discrepancy. We do this in the following manner. If the winding number of a point is positive with respect to B_i, we include it in our region. If the winding number is zero or negative, we do not include it. This gives the region R_i which we use as our approximation to the region R.

By the assumption that p be a power-of-two multiple of $\log^{3/4} N$, we know that A_{m+1} coincides with the boundary of R. Since the vertices of Γ are subset of the grid points in G_{m+1}, this means that B_{m+1} is identical to the boundary of R and thus $R_{m+1}=R$. Since B_1 contains only 2 vertices, $R_1=\emptyset$. Hence, we have constructed the desired sequence of approximations to R. The $\{D_i\}$ and $\{D_i^*\}$ are now easily defined for $1\leq i\leq m$ as follows:

$$D_i = R_{i+1}/R_i$$

and

$$D_i^* = R_i/R_{i+1},$$

where the notation S_1/S_2 is used to denote the set of points contained in S_1 but not in S_2.

It is readily observed that

$$\Delta(S_2) \leq \Delta(S_1)+\Delta(S_1/S_2)+\Delta(S_2/S_1)$$

for any regions S_1 and S_2. Thus for any i,

$$\Delta(R_{i+1})-\Delta(R_i) \leq \Delta(D_i)+\Delta(D_i^*).$$

Summing over $1\leq i\leq m$ and using $R_{m+1}=R$ and $R_1=\emptyset$, we find that

$$\Delta(R) \leq \sum_{i=1}^{m} [\Delta(D_i)+\Delta(D_i^*)],$$

as desired.

Each D_i and D_i^* is contained in the union of "triangles" removed from B_{i+1} to produce B_i. That is because any point which has a positive winding number with respect to B_{i+1} but not with respect to B_i must be contained in one of the triangle-like objects removed from B_{i+1} to form B_i. (The objects removed from A_{i+1} to form A_i *are* triangles, albeit possibly overlapping, but slight deformations of this basic structure are possible when passing to B_{i+1} and B_i because of the coarseness of the approximation.) We will eventually use this fact to show that the areas of the D_i's and D_i^*'s are small. Before doing this, however, we must classify D_i and D_i^* as elements of $\mathcal{D}_{i,s,p}$ for some s and p. The choice of p is easy; it is simply the perimeter of the original region R. Since the boundary of R follows Γ, p must be an integer multiple of $\log^{3/4} N$ in the range $[\log^{3/4} N, N]$, as needed for the

definition of $\mathcal{D}_{i,s,p}$. In fact, we can increase p to a power-of-two multiple of $\log^{3/4} N$ by appending to R a zero-area region with sufficiently large perimeter.

The selection of $s=s_i$ is much more complicated and depends on the triangle decomposition of R. In particular, we will need to examine the triangles T_{ij} that are removed from A_{i+1} to form A_i. Specifically, we define T_{ij} to be the triangle with vertices $a_{i,j-1}$, $a_{i,j}$ and $a_{i+1,2j-1}$. Recall that e_{ij} is defined to be the distance between $a_{i,j-1}$ and $a_{i,j}$ and define

$$q_i = \frac{2^i}{p^2} \sum_{j=1}^{2^i} e_{ij}^2.$$

The value of q_i is a normalized, dimensionless parameter that is closely related to the perimeter of A_i as well as the area of R_{i+1}/R_i. It is hard to provide better motivation for this definition except to say that it has a sufficiently strong form so that the inequalities will all work out right later. Unfortunately, the same cannot be said of more natural definitions. The following fact, in particular, will prove useful later.

Fact 1. $q_i \leqq q_{i+1} \leqq 1$ *for* $0 \leqq i \leqq m-1$.

Proof. To show that $q_m \leqq 1$, we observe that the 2^m edges of A_m each have length at most $p/2^m$. Thus

$$q_m \leqq \frac{2^m}{p^2} 2^m (p/2^m)^2 = 1.$$

To show that $q_i \leqq q_{i+1}$, we apply the triangle inequality to T_{ij} to show that $e_{ij}^2 \leqq \leqq 2(e_{i+1,2j-1}^2 + e_{i+1,2j}^2)$. Summing over $1 \leqq j \leqq 2^i$ yields that

$$\sum_{j=1}^{2^i} e_{ij}^2 \leqq 2 \sum_{j=1}^{2^{i+1}} e_{i+1,j}^2$$

and thus that

$$\frac{2^i}{p^2} \sum_{j=1}^{2^i} e_{ij}^2 \leqq \frac{2^{i+1}}{p^2} \sum_{j=1}^{2^{i+1}} e_{i+1,j}^2$$

which means that $q_i \leqq q_{i+1}$. ∎

It is also worthwhile defining

$$k_{ij} = 2(e_{i+1,2j-1}^2 + e_{i+1,2j}^2) - e_{ij}^2$$

which is proportional to the contribution of triangle T_{ij} to q_i, and

$$r_i = q_{i+1} - q_i$$

which is the difference in q for A_{i+1} and A_i. By Fact 1, $r_i \geqq 0$ for $0 \leqq i \leqq m-1$ and

$$\sum_{i=1}^{m-1} r_i = q_m \leqq 1.$$

Also note that

$$\sum_{j=1}^{2^i} k_{ij} = p^2 r_i / 2^i.$$

Lastly, we define

$$r_i' = r_i + \frac{r_{i-1}}{2} + \frac{r_{i-2}}{4} + \ldots + \frac{r_1}{2^{i-1}}$$

and set s_i to be the smallest integer multiple of $1/\log N$ that is greater than $\alpha r_i' + \beta/\log N$ where α and β are sufficiently large constants to be determined later. (It will turn out that $\alpha = 9e\pi$ and $\beta = 9e(1+16\pi)$ are sufficient where $e = 2.718\ldots$.) The class $\mathscr{D}_{i,s,p}$ is then defined to be the set of all D_i and D_i^* resulting from a region R with perimeter p for which $s = s_i$. It now remains to prove the hypotheses stated at the beginning of this section. In particular, we still must show that:

1) $Area(D_i)$ and $Area(D_i^*)$ are at most $2^{-i}p^2\sqrt{s_i}$,

2) the number of D_i or D_i^* for which $\frac{t-1}{\log N} \leqq s_i \leqq \frac{t}{\log N}$ for some integer t is at most $N^2 t^{2i+1}$, and

3) $s_i + \ldots + s_m \leqq c_0$ where c_0 is a constant independent of N, p and R.

We start by showing that the areas of D_i and D_i^* are at most $2^{-i}p^2\sqrt{s_i}$. From before, we know that D_i and D_i^* are contained in the union of the triangle-like objects removed from B_{i+1} to form B_i. In particular, define $E_i = \bigcup_{j=1}^{2^i} T_{ij}$ and define F_i to be the set of all points within distance $\sqrt{2}g_{i+1}$ of a point in E_i. Then $D_i, D_i^* \subseteq F_i$ and it is sufficient to show $Area(F_i) \leqq 2^{-i}p^2\sqrt{s_i}$. Since E_i can be bounded with two closed curves each with length at most p, it is easily seen that

$$Area(F_i) \leqq Area(E_i) + 2\sqrt{2}g_{i+1}p + 2\pi g_{i+1}^2.$$

The area of E_i is bounded by the sum of the areas of the triangles T_{ij} it contains. The area of T_{ij}, in turn, is at most $\frac{1}{2}e_{ij}$ times the altitude of the triangle. A simple geometric argument shows that the altitude is at most $\frac{1}{2}\sqrt{k_{ij}}$ where, as before, $k_{ij} = 2(e_{i+1,2j-1}^2 + e_{i+1,2j}^2) - e_{ij}^2$. Since $e_{ij} \leqq p/2^i$, this means that $Area(T_{ij}) \leqq p\sqrt{k_{ij}}/2^{i+2}$ and thus that

$$\begin{aligned} Area(E_i) &\leqq \frac{p}{2^{i+2}} \sum_{j=1}^{2^i} \sqrt{k_{ij}} \\ &\leqq \frac{p}{4} \sqrt{\sum_{j=1}^{2^i} k_{ij}/2^i} \\ &= \frac{p^2\sqrt{r_i}}{2^{i+2}}. \end{aligned}$$

We can now conclude that

$$\begin{aligned} Area(F_i) &\leqq \frac{p^2\sqrt{r_i}}{2^{i+2}} + 2\sqrt{2}g_{i+1} + 2\pi g_{i+1}^2 \\ &\leqq 2^{-i}p^2\left(\frac{\sqrt{r_i}}{4} + \frac{\sqrt{2}}{p\sqrt{\log N}} + \frac{\pi}{2^{i+1}\log N}\right) \\ &\leqq 2^{-i}p^2\sqrt{s_i} \end{aligned}$$

provided only that $\alpha > \frac{1}{16}$ and $\beta > \frac{\pi}{8}$. (Recall that $g_i = p/2^i\sqrt{\log N}$ and $p \geq \log^{3/4} N$ since the boundary of R coincides with Γ.)

We next show that the number of D_i and D_i^* for which $\frac{t-1}{\log N} \leq s_i \leq \frac{t}{\log N}$ for some integer t is at most $N^2 t^{2^{i+1}}$. This will be reasoned with a Kolmogorov complexity type of argument. In particular, we will associate every set of sequential approximations $\{B_1, \ldots, B_{i+1}\}$ for which $s_i \leq \frac{t}{\log N}$ with a unique triple of lists. The first list will consist of two integers from $[1, N]$. The second list will consist of 2^{i+1} integers drawn from $[1, 9]$. The third list will consist of 2^{i+1} integers that sum to at most $\frac{\pi}{\alpha}\, 2^{i+1}(t-\beta+16\alpha)$, where α and β are the constants in the definition of $s_i = \alpha r_i' + \beta/\log N$. This will be sufficient to prove the desired result since the number of such triples of lists is at most

$$N^2(9)^{2^{i+1}} \binom{\frac{\pi}{\alpha} 2^{i+1}(t-\beta+16\alpha)+2^{i+1}}{2^{i+1}}$$

$$= N^2(9)^{2^{i+1}} \binom{\frac{\pi}{\alpha} 2^{i+1}\left(t-\beta+16\alpha+\frac{\alpha}{\pi}\right)}{2^{i+1}}$$

$$\leq N^2(9)^{2^{i+1}} \left[\frac{\pi e}{\alpha}\left(t-\beta+16\alpha+\frac{\alpha}{\pi}\right)\right]^{2^{i+1}}$$

$$= N^2 \left[\frac{9\pi e}{\alpha}\left(t-\beta+16\alpha+\frac{\alpha}{\pi}\right)\right]^{2^{i+1}}$$

$$\leq N^2 t^{2^{i+1}}$$

provided that $\alpha > 9e\pi$ and $\beta \geq 16\alpha + \frac{\pi}{\alpha} \geq 9e(1+16\pi)$. Notice that we used the well-known inequality $\binom{x}{y} \leq \left(\frac{xe}{y}\right)^y$ in the third step of the calculation.

We start describing the association of $\{B_1, \ldots, B_{i+1}\}$ with lists by considering the ways that B_i can be extended to form B_{i+1}. First we note that for every vertex b_{ij} of B_i, the corresponding vertex $b_{i+1,2j}$ of B_{i+1} is either the same point or an adjacent point of the grid G_{i+1}. Hence there are 9 possible ways to refine the approximation of each vertex of B_i in B_{i+1}.

The remaining vertices of B_{i+1} fall "between" consecutive vertices of B_i. To specify the location of vertex $b_{i+1,2j-1}$, we label all the grid points of G_{i+1}, starting with the midpoint of the edge $b_{i,j-1}b_{ij}$ of B_i. The labelling is done in increasing order of distance from the midpoint. For example, see Figure 11. Since the edges of A_{i+1} have length at most $p/2^{i+1}$, the next vertex of B_{i+1} conceivably could be as far as $p/2^{i+1}$ away from this midpoint (up to the $O(g_{i+1})$ error induced by the approximation). This is essentially $\sqrt{\log N}$ grid points of G_{i+1} away. How-

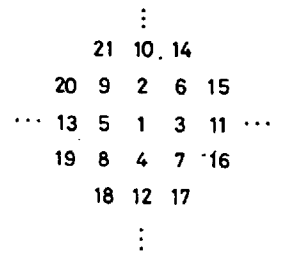

Fig. 11. Numbering the grid points

ever, most points will be considerably closer, resulting in a much narrower range of choices for most $b_{i+1,2j-1}$. We will prove this using the following elementary fact from geometry.

Fact 2. *Let UV denote a line segment and let $|UV|$ denote its length. The locus of points W determined by $2(|UW|^2+|VW|^2)-|UV|^2=k$ is a circle of radius $\frac{1}{2}\sqrt{k}$ about the midpoint of UV.* ∎

By Fact 2 the point $b_{i+1,2j-1}$ of B_{i+1} will be in a circle of radius $\varrho=\frac{1}{2}(\sqrt{k_{ij}}+3\sqrt{2}g_{i+1})$ centered on the midpoint of the edge $b_{i,j-1}b_{ij}$. (Recall that $k_{ij}=2(e^2_{i+1,2j-1}+e^2_{i+1,2j})-e^2_{ij}$ and that B_{i+1} is an approximation to A_{i+1}.) Since we labelled the points in order of increasing distance from the midpoint, the label l_{ij} of the point $b_{i+1,2j-1}$ will be less than the number of grid points in that circle. The number of grid points of G_{i+1} in a circle of radius ϱ is at most $\pi(\varrho+g_{i+1}/\sqrt{2})^2/g^2_{i+1}$. It follows that the label l_{ij} of the point $b_{i+1,2j-1}$ satisfies the equation

$$l_{ij} \leqq \frac{1}{2}\pi k_{ij}/g^2_{i+1}+16\pi.$$

Summing this inequality for l_{ij} above, we get that

$$\begin{aligned}\sum_{j=1}^{2^i} l_{ij} &\leqq \left(\frac{1}{2}\pi/g^2_{i+1}\right)\sum_{j=1}^{2^i} k_{ij}+16\pi 2^i\\ &\leqq \frac{\pi}{2g^2_{i+1}}\left(\frac{p^2 r_i}{2^i}\right)+16\pi 2^i\\ &= \pi 2^{i+1} r_i \log N+16\pi 2^i.\end{aligned}$$

The preceding analysis indicates that the number of options for B_{i+1} given B_i is limited in a fundamental way by r_i. In particular B_{i+1} can be completely specified by B_i, r_i, a list of 2^i numbers from [1, 9], and a list of 2^i positive integers that sum to at most $\pi 2^{i+1} r_i \log N+16\pi 2^i$. Applying this process recursively, we find that $\{B_2, \ldots, B_{i+1}\}$ can be completely specified by B_1, r_i, a list of $2^i+2^{i-1}+\ldots+2 \leqq \leqq 2^{i+1}$ numbers from [1, 9], and a list of 2^{i+1} positive integers that sum to at most

$$\sum_{i'=1}^{i} (\pi 2^{i'+1} r_{i'} \log N+16\pi 2^{i'}) \leqq \pi r'_i 2^{i+1} \log N+16\pi 2^{i+1}.$$

If we restrict our attention to those $\{B_1, \ldots, B_{i+1}\}$ for which $s_i \leqq t/\log N$, then the preceding sum is at most

$$\pi 2^{i+1}\left(\frac{s_i \log N - \beta}{\alpha}\right) + 16\pi 2^{i+1} \leqq \frac{\pi 2^{i+1}}{\alpha}(t - \beta + 16\pi).$$

Since B_1 contains only two points from G_1, it can be completely specified by naming the points. Since G_1 has at most N points, this is accomplished with a list of two numbers from $[1, N]$. Hence, every $\{B_1, \ldots, B_{i+1}\}$ for which $s_i \leqq \frac{t}{\log N}$ can be uniquely specified by a triple of lists: one with two numbers from $[1, N]$, one with 2^{i+1} numbers from $[1, 9]$, and one with 2^{i+1} integers that sum to at most $\frac{\pi}{\alpha} 2^{i+1}(t - \beta + 16\pi)$. As argued previously, this means that the number of D and D_i^* for which $s_i \leqq \frac{t}{\log N}$ is at most $N^{2i+1} t^{i+1}$.

The last step in the deterministic part of the proof is to show that the maximum of $s_1 + \ldots + s_m$ over all regions R is bounded by a constant independent of N and p. Fortunately, this is easy to do since

$$\begin{aligned}
\sum_{i=1}^{m} s_i &\leqq \frac{m}{\log N} + \sum_{i=1}^{m}\left(\alpha r_i' + \frac{\beta}{\log N}\right) \\
&= \frac{(\beta+1)m}{\log N} + \alpha \sum_{i=1}^{m} \sum_{j=1}^{i} \frac{r_j}{2^{i-j}} \\
&= \frac{(\beta+1)\log p}{\log N} + \alpha \sum_{j=1}^{m} \sum_{i=j}^{m} \frac{r_j}{2^{i-j}} \\
&\leqq \beta + 1 + \alpha \sum_{j=1}^{m} 2 r_j \\
&\leqq \beta + 1 + 2\alpha
\end{aligned}$$

which is a constant independent of N and p.

4.3.2. The Probabilistic Part of the Proof

The probabilistic part of the proof of Theorem 2 consists of bounding the discrepancies of the regions in the $\mathcal{D}_{i,s,p}$. To do this, we will make use of the following simple fact, which can be proved by an elementary counting argument.

Fact 3. *Given a region D with area A in the $\sqrt{N} \times \sqrt{N}$ square, the probability that $\Delta(D) \geqq q$ is at most*

$$O(e^{-c_1 q^2/A}) \quad \textit{for} \quad q \leqq A$$

and

$$O(e^{-c_1 q}) \quad \textit{for} \quad q \geqq A,$$

where c_1 is a constant independent of D and N.

In order to bound the maximum discrepancy of a region in $\mathscr{D}_{i,s,p}$ with very high probability, it is sufficient to find a q such that

$$N^2 2^{2^{i+1}\log(s\log N)} e^{-c_1 q^2/A} \leqq O(2^{-\log^{3/2} N})$$

and

$$N^2 2^{2^{i+1}\log(s\log N)} e^{-c_1 q} \leqq O(2^{-\log^{3/2} N}),$$

where $A \leqq 2^{-i} p^2 \sqrt{s}$. This because the probability that a fixed region in $\mathscr{D}_{i,s,p}$ has discrepancy q or greater is $O(e^{-c_1 q^2/A})$ or $O(e^{-c_1 q})$ depending on whether or not $q \leqq A$, and because $\mathscr{D}_{i,s,p}$ only contains at most $N^2 2^{2^{i+1}\log(s\log N)}$ regions.

Solving for $q = q_1$ in the first inequality we find that it is sufficient for

$$\begin{aligned} q_1 &= O\left(\sqrt{A[\log^{3/2} N + 2^i \log(s\log N)]}\right) \\ &\leqq O\left(2^{-i/2} p s^{1/4} \log^{3/4} N + p s^{1/4} \sqrt{\log(s\log N)}\right) \\ &\leqq O\left(2^{-i/2} p \log^{3/4} N + p s^{1/4} \log(s\log N)\right). \end{aligned}$$

Solving for $q = q_2$ in the second inequality, we find that it is sufficient for

$$\begin{aligned} q_2 &= O\left(\log^{3/2} N + 2^i \log(s\log N)\right) \\ &= O(\log^{3/2} N + 2^i \log\log N). \end{aligned}$$

Hence with very high probability, the maximum discrepancy of a region in $\mathscr{D}_{i,s,p}$ ist at most

$$\begin{aligned} q &= q_1 + q_2 \\ &= O\left(2^{-i/2} p \log^{3/4} N + p s^{1/4} \log(s\log N) + \log^{3/2} N + 2^i \log\log N\right). \end{aligned}$$

Since there are only a polynomial number of combinations of i, s and p, we can conclude that the preceding bound holds with very high probability for all $\mathscr{D}_{i,s,p}$. From Section 4.3.1, we know that for every region R,

$$\Delta(R) \leqq \sum_{i=1}^{m} \left(\Delta(D_i) + \Delta(D_i^*)\right)$$

where $D_i, D_i^* \in \mathscr{D}_{i,s,p}$ for $1 \leqq i \leqq m$ and $s_1 + \ldots + s_m \leqq c_0$ for some constant c_0. Thus, with very high probability, the discrepancy of any region R whose boundary follows Γ is at most

$$O\Big(\sum_{i=1}^{\log p} \left[2^{-i/2} p \log^{3/4} N + p s_i^{1/4} \log(s_i \log N) + \log^{3/2} N + 2^i \log\log N\right]\Big)$$

where $\sum_{i=0}^{\log p} s_i \leqq c_0$.

We will consider the impact of each of the four terms in the sum individually. The first term in the sum forms a geometric series that sums to $O(p \log^{3/4} N)$. The second term in the sum is the most important and lies at the heart of the proof. Because $x^{1/4} \log(x \log N)$ is convex in x, we know by Jensen's inequality that the value of $\sum_{i=1}^{m} p s_i^{1/4} \log(s_i \log N)$ is maximized when the s_i are all equal. Hence

these terms sum to

$$O(\log Np(c_0/\log N)^{1/4} \log c_0) = O(p \log^{3/4} N).$$

The third term in the sum trivially adds up to $O(\log p \log^{3/2} N)$ which is $O(p \log^{3/4} N)$ provided that $p \geq \log^{3/4} N \log\log N$. The fourth term in the sum forms another geometric series that converges to $O(p \log\log N)$.

The preceding argument concludes the proof of Theorem 2 except in the special case when $p \leq \log^{3/4} N \log\log N$. For this case, we can forget Section 4.3.1 and apply Fact 3 directly.

Since such regions can contain at most $(\log\log N)^2$ interlinked square cells of Γ, there are at most $O(N^2)$ such regions. Hence, with very high probability, the maximum discrepancy of such a region is at most $q = q_1 + q_2$ where

$$N^2 e^{-c_1 q_1^2/A} \leq O(2^{-\log^{3/2} N})$$

and

$$N^2 e^{-c_1 q_2} \leq O(2^{-\log^{3/2} N}).$$

Solving the equations yields

$$\begin{aligned} q_1 &= O(\sqrt{A} \log^{3/4} N) \\ &\leq O(p \log^{3/4} N) \end{aligned}$$

and

$$\begin{aligned} q_2 &= O(\log^{3/2} N) \\ &\leq O(p \log^{3/4} N) \end{aligned}$$

since $p \geq \log^{3/4} N$ and $A \leq O(p^2)$. Thus, with very high probability, the discrepancy of any region with $p \leq \log^{3/4} \log\log N$ is $O(p \log^{3/4} N)$ and we are done with the proof of Theorem 2.

It is worth remarking that the probability bound in Theorem 2 is essentially tight. For example, the discrepancy of the upper half of the $\sqrt{N} \times \sqrt{N}$ square will exceed $c_1 \sqrt{N} \log^{3/4} N$ with probability $2^{-c_2 \log^{3/2} N}$ for some constant c_1 and c_2. In fact, this example illustrates the relationship between the constants in the $\Theta(p \log^{3/4} N)$ bound and the constant in the exponent of the very high probability bound. With further work, it is conceivable that the exact relationship could be specified.

4.4. Extension to Arbitrary Regions

In what follows, we show how to extend the discrepancy bound in Theorem 2 to arbitrary simply connected regions. The result is not essential to the rest of the paper (Theorem 2 is sufficient), but it is more natural than Theorem 2 and may eventually prove to be equally important.

Theorem 3. *Consider a set of N points uniformly and independently distributed in the* $\sqrt{N} \times \sqrt{N}$ *square. There is a constant c such that with probability* $1 - N^{-\log^{1/2} N}$, *every simply connected region R has discrepancy* $\Delta(R) \leq cPer(R) \log^{3/4} N + c \log^{3/2} N$.

Proof. The proof is not difficult given Theorem 2 so we will only sketch the argument here. Divide R into two regions R' and R'' where R' consists of all $\log^{3/4} N \times \log^{3/4} N$ square subregions of Γ that are entirely contained in R and where $R'' =$

$=R/R'$. Although R' need not be connected, we can conclude immediately from Theorem 2 that $\Delta(R') \leqq O(p' \log^{3/4} N)$ with very high probability where p' is the perimeter of R'. By elementary geometry, $p' \leqq 6p$ where p is the perimeter of R and hence $\Delta(R') \leqq O(p \log^{3/4} N)$ with very high probability. Since $\Delta(R) \leqq \Delta(R') + \Delta(R'')$, it remains only to show that $\Delta(R'') \leqq O(p \log^{3/4} N + \log^{3/2} N)$ with very high probability.

With very high probability, every square subregion of Γ contains at most $O(\log^{3/2} N)$ random points. Since any square containing part of R'' must contain part of the simple closed curve bounding R, at most $O(1 + p/\log^{3/4} N)$ squares can contain part of R''. Any such square can contribute at most $O(\log^{3/2} N)$ to the discrepancy of R'' since all squares have $O(\log^{3/2} N)$ area and $O(\log^{3/2} N)$ random points with very high probability. Thus $\Delta(R'') \leqq O(\log^{3/2} N + p \log^{3/4} N)$ with very high probability, as claimed. ∎

It is worth remarking that a slightly stronger result can be proved if the very high probability assumption is relaxed. In particular, with high probability, every region with perimeter p has discrepancy at most $O(\log N + p \log^{3/4} N)$. (This can be proved by noting that Theorem 3 handles the case when $p \geqq \log^{3/4} N$, and then observing that there are at most a polynomial number of regions with $p \leqq \log^{3/4} N$.) The $\log^{3/4} N$ factor cannot be improved since Shor [22] has demonstrated the existence of regions having discrepancy $\Omega(p \log^{3/4} N)$ for $p = \Theta(\sqrt{N})$ with very high probability.

5. Remarks

Upon reflection, it really should not be surprising that the minimax grid matching problem arises in so many diverse and useful applications. After all, its dual discrepancy problem captures a very important measure of expected discrepancies in random data. In fact, it is precisely the measure that is important to the analysis of many algorithms.

It is likely that other applications of this work will be discovered, and even more likely still that other matching and/or discrepancy problems will arise in the analysis of algorithms. One such problem that still remains unsolved is the *rightward matching problem.* In rightward matching, we have N random pluses and N random minuses $\mathscr{P}^{\pm}$ in a unit square, and are asked to match pluses rightward to minuses in a way which minimizes the average vertical distance $V(\mathscr{P}^{\pm})$ of the matching edges. (Unmatched pluses are considered to match to the top or bottom of the square.) As a consequence of the minimax grid matching result in this paper, we can conclude that $V(\mathscr{P}^{\pm}) \leqq O(\log^{3/4} N)$ with very high probability. Improvement of the lower bound of $\Omega(\sqrt{\log N})$ would directly lead to improved lower bounds on the expected wasted space of any on-line bin packing algorithm [21, 22].

Discrepancy problems are still of interest in mathematical statistics. For example, Dudley [7] leaves open the question concerning the maximum expected discrepancy of any convex 3-dimensional region. Curiously, this appears to be analogous to the 2-dimensional up-right discrepancy problem analyzed in this paper. We suspect that the bound is again $\Theta(\sqrt{N} \log^{3/4} N)$ but have not proved it.

Multidimensional matching and discrepancy results will also be likely to have consequences for multidimensional packing and allocation problems. For example, see [10].

Acknowledgements. We are indebted to Noga Alon, Jon Bentley, Ed Coffman, Persi Diaconis, Richard Dudley, Dave Johnson, Richard Karp, János Komlós, Mike Luby, Cathy McGeogh, Larry Shepp, Mike Steele and Endre Szemerédi for helpful comments and references.

Note added in proof: The results of this paper were presented at the 18th Annual Symposium on Theory of Computing in 1986 (pp. 91—103). Our result for the special case of up-right matching was independently discovered by Rhee and Talagrand [W. T. RHEE and M. TALAGRAND, Exact bound for the stochastic upward matching problem, *Trans. AMS* **307** (1988), pp. 109—125]. The more general grid matching result was recently used by Yukich (unpublished) to resolve an open question of Dudley [R. M. DUDLEY, The speed of mean Glivenko—Cantelli convergence, *Ann. Math. Stat.* **40** (1969), pp. 40—50] concerning the Prokhorov distance between the uniform measure and empirical measure of random points in a unit square.

References

[1] M. AJTAI, J. KOMLÓS and G. TUSNÁDY, On optimal matchings, *Combinatorica*, **4**, 259—264, 1983.
[2] J. L. BENTLEY, D. S. JOHNSON, F. T. LEIGHTON, C. C. MCGEOCH and L. MCGEOCH, Some unexpected expected behavior results for bin packing, *Proceedings of the 16th ACM Symp. on the Theory of Computing*, 279—288, 1984.
[3] J. BLUM, On convergence of empirical distribution functions, *Annals of Mathematical Statistics*, **26**, 527—529, 1955.
[4] E. G. COFFMAN, JR. and F. T. LEIGHTON, A provably efficient algorithm for dynamic storage allocation, *Proceedings of the 18th ACM Symp. on Theory of Computing*, May 1986, *to appear*.
[5] P. DIACONIS, personal communication, 1985.
[6] R. M. DUDLEY, A course in empirical processes, *École d'Été de Probabilités de Saint-Flour XII*, 1982, Lecture Notes in Math. No. 1097, 1—142, Springer Verlag, NY, 1984. (The relevant part is Chapter 8.)
[7] R. M. DUDLEY, Empirical and Poisson Processes on classes of sets or functions too large for central limit theorems, *Z. Wahrsch. verw. Gebiete*, **61**, 355—368, 1982.
[8] A. FIAT and A. SHAMIR, Polymorphic arrays: a novel VLSI layout for systolic computers, *Proceedings of the 25th Symp. on Foundations of Computer Science*, 37—45, 1984.
[9] P. HALL, On representatives of subsets, *J. London Math. Soc.*, **10**, 26—30, 1935.
[10] R. M. KARP, M. LUBY and A. MARCHETTI-SPACCAMELA, Probabilistic analysis of multi-dimensional bin packing problems, *Proceedings of the 16th ACM Symp. on Theory of Computing*-289—298, 1984.
[11] J. KIEFER, On the large deviation of the empiric d.f. of vector chance variables and a law of the iterated logarithm, *Pacific J. Math.*, **11**, 649—660, 1961.
[12] F. T. LEIGHTON, 18.419 class notes, 1984.
[13] F. T. LEIGHTON and C. E. LEISERSON, Wafer-scale integration of systolic arrays, *IEEE Trans. on Computers*, **C-34, No. 5**, 448—461, 1985.
[14] M. LERCH, Question 1547, *L'Intermédiare Math.*, **11**, 145—146, 1904.
[15] M. Luby, personal communication, 1985.
[16] W. PHILIPP, Empirical distribution functions and uniform distribution mod 1, *Diophantine Approximation and its Applications*, C.F. Osgood, ed., Academic Press, NY, 1973.
[17] W. PHILIPP, personal communication quoted in [6].
[18] G. SAWITZKI, Another random number generator which should be avoided, *Statistical Software Newsletters*, **11**, No. 2, 81—82, 1985.

[19] W. SCHMIDT, *Lectures on Irregularities of Distribution*, Tata Institute of Fundamental Research, Bombay, India, 1977.
[20] W. SCHMIDT, Irregularities of distribution, VII, *Acta Arithmetica*, **21**, 45—50, 1972.
[21] P. W. SHOR, The average-case analysis of some on-line algorithms for bin packing, *Proceedings of the 25th Symp. on Foundations of Computer Science*, 193—200, 1984.
[22] P. W. SHOR, *Random Planar Matching and Bin Packing*, Ph. D. Thesis, MIT Math. Dept., 1985.
[23] M. STEELE, Limit properties of random variables associated with a partial ordering of R_d, *Annals of Probability*, **5**, No. 3, 395—403, 1977.
[24] J. VAN DER CORPUT, Verteilungsfunctionen I., *Proc. Kon. Ned. Akad. v. Wetensch*, **38**, 813—821, 1935.
[25] V. N. VAPNIK and A. YA. CERVONENKIS, Necessary and sufficient conditions for the uniform convergences of means to their expectations, *Theory of Probability and Applications*, **26**, 532—553, 1981.
[26] F. T. WRIGHT, The empirical discrepancy over lower layers and a related law of large numbers, *Annals of Probability*, **9**, 323—329, 1981.

Tom Leighton

Mathematics Department
and Laboratory for Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139
U.S.A.

Peter Shor

AT&T Bell Labs.
Murray Hill, NY 07974
U.S.A.