The Annals of Probability
1991, Vol. 19, No. 3, 1338–1348

MINIMAX GRID MATCHING AND EMPIRICAL MEASURES

By P. W. Shor and J. E. Yukich

AT & T Bell Laboratories and Lehigh University

In this article we solve the minimax grid matching problem in dimensions greater than two. As a by-product, we settle a long-open problem involving the Glivenko–Cantelli convergence of empirical measures.

1. Introduction. Given two sets of points $X := \{x_1, \dots, x_n\}$ and $Y := \{y_1, \dots, y_n\}$, where x_i and $y_i \in \mathbb{R}^d$ for all $i \geq 1$, let $L(X, Y)$ denote the minimum length such that there exists a perfect matching of the points in X to the points in Y for which the distance between every pair of matched points is at most $L(X, Y)$. In other words, $L(X, Y)$ is the minimum over all perfect matchings of the maximum matching length for points in X to points in Y; $L(X, Y)$ is thus called the *minimax matching length* for the couple (X, Y). L is a metric on the collection of unordered sets of n points.

Let $S := [0, 1]^d$, $d \geq 2$, denote the unit cube and G a regularly spaced $n^{-1/d} x \cdots x n^{-1/d}$ array of n grid points on S ($n = k^d, k \in \mathbb{N}^+$). Partition S into n congruent cubes of volume n^{-1}; call these the grid cubes. Each grid cube is thus centered around a grid point. Let $X := X(\omega) := \{X_1(\omega), \dots, X_n(\omega)\}$ denote a collection of n random points in S. Assume that the X_i, $1 \leq i \leq n$, are i.i.d. random variables with the uniform distribution λ on S.

The problem of finding the expected value of $L(X, G)$ is called the minimax grid matching problem. Leighton and Shor [7] have shown that the minimax grid matching problem is fundamental in the average case analysis of algorithms. When d equals 2, they have shown that with very high probability [i.e., with probability exceeding $1 - n^{-\alpha}$, $\alpha = c_1(\log n)^{1/2}$ for some constant c_1] there are constants c and C such that

$$c \log^{3/4} n \leq n^{1/2} L(X, G) \leq C \log^{3/4} n. \tag{1.1}$$

Here and throughout, c and C denote *finite positive* constants with values possibly changing from line to line. They use this estimate to solve the maximum upright matching problem (cf. [6]) and as a result obtain tight upper bounds on the average case behavior of the best algorithms known for two-dimensional bin packing and one-dimensional on-line packing. When d equals 1, it is easily seen that there are constants c and C such that

$$c \leq n^{1/2} \mathbb{E} L(X, G) \leq C.$$

Received February 1989; revised March 1990.
AMS 1980 *subject classifications*. Primary 60B10, 05C70; secondary 60D05, 68C05.
Key words and phrases. Minimax matching length, empirical measures, Prokhorov metric.

In this article we derive the analog of this result for dimensions greater than two. In fact the main theorem is given by:

THEOREM 1.1. *Let $d \geq 3$. There exist constants c and C depending only upon d such that*

$$c \leq \liminf\left(\frac{n}{\log n}\right)^{1/d} L(X,G) \leq \limsup\left(\frac{n}{\log n}\right)^{1/d} L(X,G) \leq C \quad a.s.$$

A long-open problem in empirical measures involves finding the a.s. exact rate of convergence of the nth empirical measure $\lambda_n := \lambda_n(\omega) := n^{-1}(\delta_{X_1(\omega)} + \cdots + \delta_{X_n(\omega)})$ to Lebesgue measure λ, as measured by the Prokhorov metric ρ (defined in Section 3). This problem was studied originally by Dudley [3] and subsequently by Gaenssler [5], Zuker [13] and more recently, Massart [8], who found the exact order rate up to a power of $\log n$. In the two-dimensional case, Yukich [12] deduced from (1.1) the same rate of convergence for $\rho(\lambda_n, \lambda)$. In dimensions greater than two, Theorem 1.1 will be used to determine exact order rates for $\rho(\lambda_n, \lambda)$; see Section 3. This result, along with that of [12], completely settles Dudley's question. Note that the one-dimensional case is trivial since the inequality ([3]) $\|F_n - F\|_\infty \leq 2\rho(\lambda_n, \lambda) \leq 4\|F_n - F\|_\infty$ and the a.s. bounded law of the iterated logarithm for the centered empirical distribution function $F_n - F$ together imply the existence of constants c and C such that

$$c \leq \limsup\left(\frac{n}{\log\log n}\right)^{1/2} \rho(\lambda_n, \lambda) \leq C \quad \text{a.s.}$$

Concerning other connections between matching problems and empirical measures, it has been noted that tight bounds for the upright matching problem (see Leighton and Shor [7], Rhee and Talagrand [9], Coffman and Shor [2]) are equivalent to discrepancies of empirical measures over lower layers.

A similar matching problem, the *transportation problem*, is considered in the paper of Ajtai, Komlós and Tusnády [1]. The essential difference between the problem analyzed in their paper and the one in ours is that they minimize the *transportation cost* $T(X, Y)$ instead of the minimax edge length. [Recall that for the sets of points $X := \{x_1, \dots, x_n\}$ and $Y := \{y_1, \dots, y_n\}$, $T(X, Y)$ is the minimum sum of the distances between matched pairs of points, where the minimum is taken over all perfect matchings of points in X to points in Y.] They also consider $T(X, Y)$ instead of $T(X, G)$, where X and Y are collections of n random points, but this changes only constant factors.

In two dimensions, Ajtai, Komlós and Tusnády [1] show that if X and Y are distributed uniformly on the unit square, then with probability $1 - o(1)$,

$$c(n \log n)^{1/2} \leq T(X, Y) \leq C(n \log n)^{1/2}.$$

Thus, one can make the *average* length of an edge in a matching

$n^{-1/2}\log^{1/2} n$, which is less than the minimax edge length by a $\log^{1/4} n$ factor.

In one dimension, it is easily shown that

$$cn^{1/2} \le \mathbb{E}T(X,Y) \le Cn^{1/2},$$

giving the same results as in the minimax grid matching problem.

As Ajtai, Komlós and Tusnády [1] remark, an easy generalization of their proof to d dimensions shows that with probability $1 - o(1)$,

$$cn^{1-1/d} \le T(X,Y) \le Cn^{1-1/d}.$$

In this case, the average edge length in the transportation problem is less than the minimax edge length by a $\log^{1/d} n$ factor.

Notice that for both the minimax matching problem and the transportation problem, the case of two dimensions is different than the case of three or more dimensions. Intuitively, this is because in three or more dimensions, the matching length is determined by the behavior of random points on small scales, that is, locally. In one dimension, the matching length is determined by the behavior of random points on large scales. In two dimensions, however, the matching length is determined by the behavior of the points on all scales, and the interaction of these different scales adds an extra $\log^{1/4} n$ factor for minimax matching and an extra $\log^{1/2} n$ factor for the transportation problem.

Both our proof and the similar proof of Ajtai, Komlós and Tusnády give an upper bound which is the sum of $\log n$ terms, each term representing the behavior of the random points on a different scale. For dimensions three or more, the sum is convergent and the terms corresponding to small scale behavior dominate. Thus, one obtains the correct bound by considering only these terms. In two dimensions, however, all the terms are of size $O(n^{1/2})$, leading to an $O(n^{1/2}\log n)$ bound, which is an overestimate. To obtain the correct bound for two dimensions, one must use a more complicated proof that takes into account the interaction between the terms for different scales.

Although Theorem 1.1 is considered only in the context of Dudley's question, there are also possible applications to bin packing problems and dynamic allocation (see [7]). These applications will not be pursued here. Also, no attempt is made to estimate the constants in Theorem 1.1.

NOTATION. We write $f(x) = O(g(x))$ iff there is a constant C such that $\lim_{x\to\infty} f(x)/g(x) \le C$. Likewise, $f(x) = \Omega(g(x))$ iff there is a constant c such that $\lim_{x\to\infty} f(x)/g(x) \ge c$. If $f(x) = O(g(x))$ and $f(x) = \Omega(g(x))$, then write $f(x) = \Theta(g(x))$.

$K = \text{Bin}(n,p)$ means that K has a binomial distribution with parameters n and p.

2. Proof of the main result. The upper estimate of Theorem 1.1 is the more difficult one and is proved first. To this end, we recursively subdivide S in a manner used first by Ajtai, Komlós and Tusnády [1] and subsequently, by

Shor [10]. For simplicity and ease of exposition assume that the dimension d is equal to 3.

First, divide S in half horizontally and linearly transform each half so that it has volume equal to the fraction of sample points X_i in it. When each half of S is transformed, the sample points in that half are similarly transformed. Next, subdivide each of these halves vertically and transform each to have volume equal to the ratio of the number of transformed points contained within to the *total number of points* n. This process generates four rectangular solids which partition S. Then divide each of these solids in half laterally and transform. Apply the procedure recursively, alternating horizontal, vertical and lateral divisions. Repeating this refining procedure until each solid contains at most one point yields n nondegenerate rectangular solids. For any random sample $X := \{X_1, \dots, X_n\}$, each of the n transformed sample points $\tilde{X}_i$ is contained in a nondegenerate solid with volume n^{-1}. Notice that any solid generated in this process is a linear transformation of a rectangular solid taken from a $2^i \times 2^j \times 2^k$ subdivision of the cube, with $0 \le i \le j \le k \le i + 1$. Notice also that the points in any solid are uniformly distributed.

Let $\tilde{X}$ denote the collection of points X after the recursive transformations described above. It will be shown with high probability that

$$L(X, G) = O\big((\log n/n)^{1/3}\big);$$

the expression with high probability will be made precise shortly. This estimate is established in two steps. First, it is shown with high probability that

$$L(X, \tilde{X}) = O\big((\log n/n)^{1/3}\big). \tag{2.1}$$

Clearly, (2.1) will follow if it can be shown that the successive transformations shift an arbitrary point by $O((\log n/n)^{1/3})$. Second, it is shown with high probability that

$$L(\tilde{X}, G) = O\big((\log n/n)^{1/3}\big). \tag{2.2}$$

The triangle inequality and the Borel–Cantelli lemma then give the desired result. The estimate (2.2) is proved by showing that the transformed solids have a bounded aspect ratio and then applying Hall's marriage theorem to match each transformed solid to an *overlapping* grid cube. This provides a perfect matching of the transformed points $\tilde{X}_i$ to the grid points.

Before proving (2.1) and (2.2) we establish some lemmas, the first of which is well known and describes the tail behavior of binomial random variables (see, e.g., [4], 2.2.7 and 2.2.8).

LEMMA 2.1. *Let* $K = \mathrm{Bin}(n, p)$.

(i) *If* $p \le \frac{1}{2}$ *and* $k \le np$, *then*

$$\Pr\{K \le k\} \le \exp\big\{-(np - k)^2/2np(1 - p)\big\}.$$

(ii) *If* $k \geq np$, *then*

$$\Pr\{K \geq k\} \leq (np/k)^k \exp\{k - np\}.$$

LEMMA 2.2. *On the first step, let* K *be the number of sample points* $X_1, \dots, X_n$ *falling on a given side of the horizontal bisecting plane. Then for all* $\alpha > 0$,

$$\Pr\{|K - (n - K)| \geq (\alpha n \log n)^{1/2}\} \leq 2n^{-\alpha/2}.$$

PROOF. Notice that $K = \mathrm{Bin}(n, \frac{1}{2})$. By Lemma 2.1, it follows that

$$\Pr\{|K - (n - K)| \geq (\alpha n \log n)^{1/2}\} = 2\Pr\left\{0 \leq K \leq \left(n - (\alpha n \log n)^{1/2}\right)/2\right\}$$
$$\leq 2n^{-\alpha/2}. \qquad \square$$

Fix $\alpha > 2$; its exact value will be specified later. From now on, the expression with high probability will mean with probability *at least* $1 - p_n$, where $p_n = O(n^{1-(\alpha/2)})$.

In the following lemma and in the sequel, S_{2^k} denotes one of the 2^k transformed solids at the kth step of the recursion, where $1 \leq k \leq \log(n/4\alpha \log n)$. Note that the total number of solids S_{2^k}, $1 \leq k \leq \log(n/4\alpha \log n)$ is $O(n)$. Also, $\mathrm{vol}(A)$ denotes the volume of the set A.

LEMMA 2.3. *There exist positive, finite constants* c *and* C *such that with high probability, the number of points in any* S_{2^k} *is between* $c(n/2^k)$ *and* $C(n/2^k)$. *Moreover, with high probability* $c/2^k \leq \mathrm{vol}\, S_{2^k} \leq C/2^k$.

PROOF. Since S_{2^k} is an affine transformation of a grid cube of volume 2^{-k}, the number of sample points in S_{2^k} equals the number of sample points in a grid cube of volume $V := 2^{-k} \geq (4\alpha \log n)/n$. The proof of the lower bound will be complete if we can show that such a grid cube contains at least $nV/2$ sample points with high probability. This, however, follows easily from Lemma 2.1(i) since if K denotes the number of points in a solid with volume V, then $K = \mathrm{Bin}(n, V)$ and

$$\Pr\{K \leq nV/2\} \leq \exp\left\{-(nV - nV/2)^2/2nV(1 - V)\right\}$$
$$\leq \exp\{-nV/8\}$$
$$\leq n^{-\alpha/2}.$$

The upper bound is proved similarly, where we may actually show that $C = 2$. By Lemma 2.1(ii),

$$\Pr\{K \geq 2nV\} \leq (nV/2nV)^{2nV} \exp\{nV\}$$
$$= (e/4)^{nV}$$
$$\leq (e/4)^{4\alpha \log n}$$
$$\leq n^{-\alpha},$$

where the last step follows from $(e/4)^4 < 1/e$. The desired high probability bounds hold uniformly over *all* S_{2^k}, since

$$\sum_{k=1}^{\log(n/4\alpha \log n)} 2^k n^{-\alpha/2} = O(n^{1-(\alpha/2)}).$$

This completes the proof of the first statement of the lemma. The second statement follows immediately, since, by construction, the volume of S_{2^k} is proportional to the number of points in it. □

The next lemma shows that with high probability, the aspect ratios of the rectangular solids change very little. By aspect ratio of a rectangle, we mean the ratio of the longest side to the shortest; by aspect ratio of a rectangular solid we mean the largest aspect ratio of any of its six faces.

LEMMA 2.4. *Suppose that the rectangular solids are constructed as above, by dividing the solid at the previous stage in half and moving the boundaries so that the volume of the rectangular solid is equal to the fraction of points in it. Then with high probability, the aspect ratio at the* $\log(n/4\alpha \log n)$ *stage is bounded uniformly over all solids by some finite constant.*

PROOF. This lemma is proved by bounding the change of the aspect ratio of a rectangular solid at each step and then multiplying these changes together. These factors will form a convergent product dominated by the last term.

Consider the kth stage of the recursion, $1 \le k \le \log(n/4\alpha \log n)$. At this stage, one edge of S_{2^k} is stretched by a factor of $\frac{1}{2}$ and then by a factor of $2K/m$, where $K = \text{Bin}(m, \frac{1}{2})$ and $m := \Theta(n/2^k)$ by Lemma 2.3. Thus the aspect ratio at the kth stage is essentially multiplied by a factor of either $m/2K$ if the shortest side is halved or by $2K/m$ if the longest side is halved, where it is noted that the factors of $\frac{1}{2}$ may be ignored since they cancel on successive stages. Lemma 2.2(i) and the lower bound of Lemma 2.3 imply that $|(2K/m) - 1| = |(K - (m - K))/m|$ is bounded by

$$(\alpha \log n/m)^{1/2} \le (\alpha 2^{k+1} \log n/n)^{1/2}$$

except on a set with probability at most $2n^{-\alpha/2} + O(n^{1-(\alpha/2)})$. Thus the aspect ratio is at worst multiplied by $(1 - (\alpha 2^{k+1} \log n/n)^{1/2})^{-1}$.

At the $\log(n/4\alpha \log n)$ stage, the aspect ratio of the $O(n/\log n)$ transformed solids is bounded by the product of the collective changes, that is, by

$$\prod_{k=1}^{\log(n/4\alpha \log n)} \left(1 - (\alpha 2^{k+1} \log n/n)^{1/2}\right)^{-1}.$$

Note, however, that the above product is $O(1)$ since $(\alpha 2^{k+1} \log n/n)^{1/2} \ge 2^{-1/2}$, $1 - x \ge e^{-2x}$ when $0 \le x \le 2^{-1/2}$ and

$$\sum_{k=1}^{\log(n/4\alpha \log n)} (\alpha 2^{k+1} \log n/n)^{1/2} = O(1).$$

Moreover, noting that the high probability lower bound of Lemma 2.3 holds over all S_{2^k}, the above $O(1)$ bound holds for all $O(n/\log n)$ transformed solids except on a set with probability at most

$$O(n^{1-(\alpha/2)}) + \sum_{k=1}^{\log(n/4\alpha \log n)} 2^{k+1} n^{-\alpha/2} = O(n^{1-(\alpha/2)});$$

that is, the bound holds with high probability. □

The last lemma is used only to prove (2.1) and justifies the presence of the $\log(n/4\alpha \log n)$ bound on the number of recursion steps.

LEMMA 2.5. *To show* (2.1) *it suffices to repeat the subdivision of the cube until the* $\log(n/4\alpha \log n)$ *stage.*

PROOF. Recall that with high probability the aspect ratios of the rectangular solids are uniformly bounded and their volumes roughly equivalent. We may stop the recursion when the rectangular solids have diameter $\Theta((\log n/n)^{1/3})$ since successive recursions would shift points at most $O((\log n/n)^{1/3})$.

However, since the aspect ratios are uniformly bounded, a solid with diameter $\Theta((\log n/n)^{1/3})$ has volume $\Theta(\log n/n)$. Thus with high probability, we may stop when there are $O(n/\log n)$ rectangular solids; that is, we may stop at the $\log(n/4\alpha \log n)$ stage. □

FIRST STEP: PROOF OF (2.1). On the first step of the recursion, Lemma 2.2 implies that with high probability the difference between the fractions representing the proportion of the points in the two respective halves of the cube is at most $n^{-1/2}(\alpha \log n)^{1/2}$. Thus, with high probability we need to shift the horizontal bisecting plane by at most $n^{-1/2}(\alpha \log n)^{1/2}$ to insure that the resulting volumes equal the fraction of points within.

With high probability on the kth step of the recursion, $1 \le k \le \log(n/4\alpha \log n)$, the number m of points in S_{2^k} satisfies

$$c(n/2^k) \le m \le C(n/2^k).$$

The number K of points falling in one-half of S_{2^k} satisfies $K = \mathrm{Bin}(m, \frac{1}{2})$. By Lemma 2.1 and the above upper bound on m, the difference between the fraction of points in the two respective halves of S_{2^k} (i.e., between K/n and $m - K/n$) is bounded above by

$$(\alpha m \log n)^{1/2}/n \le (C\alpha \log n/2^k n)^{1/2}, \tag{2.3}$$

except perhaps on a set with probability $2n^{-\alpha/2} + O(n^{1-(\alpha/2)})$. Note that (2.3) represents an upper bound on the required volume change.

Next, let s_k denote the distance through which the bisecting plane of S_{2^k} must be moved in order that the resulting volumes equal the fraction of points within. By Lemma 2.4, the aspect ratios of all of the transformed solids are uniformly bounded [except on a set with probability $O(n^{1-(\alpha/2)})$] and since $\mathrm{vol}(S_{2^k}) = \Theta(2^{-k})$, it follows that the edge length of S_{2^k} is $\Theta(2^{-k/3})$. The required volume change is thus $\Theta(s_k(2^{-2k/3}))$. Bounding this by (2.3) yields, modulo a constant factor independent of k, the estimate

$$s_k \le \left(C\alpha 2^{k/3} \log n/n\right)^{1/2}. \tag{2.4}$$

Clearly, $s_k = O((2^{k/3} \log n/n)^{1/2})$ except on a set with probability $2n^{-\alpha/2} + O(n^{1-(\alpha/2)})$.

Thus, each point is clearly shifted through a net distance of at most

$$\sum_{k=1}^{\log(n/4\alpha \log n)} s_k = O\left(\sum_{k=1}^{\log(n/4\alpha \log n)} \left(2^{k/3} \log n/n\right)^{1/2}\right) = O\left(\left(\log n/n\right)^{1/3}\right),$$

as claimed. As in the proof of Lemma 2.4, these bounds hold everywhere except perhaps on a set with probability $O(n^{1-(\alpha/2)})$. This concludes the first step.

SECOND STEP: PROOF OF (2.2). In this step the recursion is carried out until each solid contains one point. It is easy to show that with high probability, $O(\log n)$ steps suffice.

Notice that after recursing $\log(n/4\alpha \log n)$ times, with high probability each transformed point $\tilde{X}_i$ belongs to one of the $O(n/\log n)$ rectangular solids of diameter $\Theta((\log n/n)^{1/3})$ (since the aspect ratios are uniformly bounded with high probability and since the volumes are roughly equivalent). Moreover, after recursing $O(\log n)$ times, with high probability each point $\tilde{X}_i$ belongs to one of n transformed solids, each with diameter $O((\log n/n)^{1/3})$ and volume n^{-1}. It only remains to show that there is a perfect matching between the n transformed rectangular solids and the original grid cubes of diameter $O(n^{-1/3})$, such that each transformed solid is matched to an *overlapping* grid cube. This is accomplished with the following lemma, which is actually a simple corollary of the marriage theorem.

LEMMA 2.6. ([10]). *Suppose that there are two partitions of* $[0,1]^3$ *into* n *solids of equal volumes* $Q_1, \dots, Q_n$ *and* $R_1, \dots, R_n$. *Then there is a matching* σ *between the* R_i*'s and the* Q_i*'s such that* $Q_i \cap R_{\sigma(i)} \neq \varnothing$.

PROOF. By Hall's marriage theorem, it suffices to show that any j of the Q_i solids can always be matched to j of the R_i solids. This will be true unless we can find $j-1$ R_i's and j Q_i's such that

$$\bigcup_{i=1}^{j} Q_i \subset \bigcup_{i=1}^{j-1} R_i.$$

However, since all of the solids have volume n^{-1}, this implies that a solid of volume j/n is contained in a solid of volume $(j-1)/n$, a contradiction. □

This completes the second step. Having shown (2.1) and (2.2), the claimed upper bound

$$\limsup\left(\frac{n}{\log n}\right)^{1/3} L(X_n, G) \le C \quad \text{a.s.}$$

follows immediately from the Borel–Cantelli lemma, provided that α is chosen large enough so that

$$\sum_{n=1}^{\infty} O(n^{1-(\alpha/2)}) < \infty.$$

Clearly it suffices to choose $\alpha = 5$ to ensure convergence. This concludes the proof of the upper estimate; a simple proof of the lower estimate is at the end of the next section.

3. Application to Glivenko–Cantelli convergence of empirical measures. On the space $\mathbb{P}([0,1]^d)$ of probability measures on $[0,1]^d$ define the Prokhorov metric:

$$\rho(\mu, \nu) := \inf\{\varepsilon > 0 : \mu(A) \le \nu(A^\varepsilon) + \varepsilon \text{ for all Borel sets } A\},$$

where $\mu, \nu \in \mathbb{P}([0,1]^d)$ and A^ε denotes the ε-neighborhood of A; that is,

$$A^\varepsilon := \{y \in \mathbb{R}^d : \|y - z\| < \varepsilon \text{ for some } z \in A\}.$$

As in Section 1, $\lambda_n(\omega)$ denotes the nth empirical measure for Lebesgue measure λ on $[0,1]^d$. It is well known that $\lambda_n(\omega)$ converges weakly to λ a.s. [11] and, as noted, Dudley's question concerns finding the exact order of convergence for $\rho(\lambda_n, \lambda)$. Massart [8] has recently obtained the estimates

$$c \le \liminf\left(\frac{n}{\log n}\right)^{1/d} \rho(\lambda_n, \lambda) \quad \text{a.s.} \quad \text{and}$$

$$\limsup\left(\frac{n}{\log^2 n}\right)^{1/d} \rho(\lambda_n, \lambda) \le C \quad \text{a.s.},$$

where c and C are constants depending only on d.

Using the upper estimate of the main theorem, it is now a simple matter to determine the exact order of convergence of $\rho(\lambda_n, \lambda)$ for $d \ge 3$. Let G_n be the discrete uniform measure based on the n grid points g_i, $1 \le i \le n$, that is $G_n := n^{-1}\sum_{i=1}^{n}\delta_{g_i}$. First, from the definition of ρ it is clear that if $n = k^d$, $k \in \mathbb{N}^+$, then

$$\rho(\lambda_n, G_n) \le L(X, G) \quad \text{and} \quad \rho(G_n, \lambda) = O(n^{-1/d}).$$

Thus, with high probability $\rho(\lambda_n, \lambda) \le \rho(\lambda_n, G_n) + \rho(G_n, \lambda) = O((\log n/n)^{1/d})$.

Next, for index values between k^d and $(k+1)^d$, notice that if $j = O(n^{(d-1)/d})$, $n = k^d$, then for all ω, $\rho(\lambda_n, \lambda_{n+j}) = O(n^{-1/d})$. Thus, with high probability,

$$\begin{aligned}\rho(\lambda_{n+j},\lambda) &\le \rho(\lambda_{n+j},\lambda_n)+\rho(\lambda_n,\lambda)\\ &= O(n^{-1/d})+\rho(\lambda_n,\lambda)\\ &= O\big((\log n/n)^{1/d}\big)\\ &= O\big((\log(n+j)/(n+j))^{1/d}\big).\end{aligned}$$

Therefore, for arbitrary n the upper estimate $\rho(\lambda_n,\lambda)=O((\log n/n)^{1/d})$ holds with high probability. Thus $\limsup(n/\log n)^{1/d}\rho(\lambda_n,\lambda)\le C$ a.s. Combining this with Massart's lower estimate settles Dudley's question for $d\ge 3$:

COROLLARY 3.1. *There exist constants c and C depending only upon d such that*

$$c\le \liminf\left(\frac{n}{\log n}\right)^{1/d}\rho(\lambda_n,\lambda)\le \limsup\left(\frac{n}{\log n}\right)^{1/d}\rho(\lambda_n,\lambda)\le C\quad a.s.$$

Next, notice that the lower estimate of Theorem 1.1 follows from the triangle inequality

$$L(X,G)\ge \rho(G_n,\lambda_n)\ge |\rho(G_n,\lambda)-\rho(\lambda,\lambda_n)|,$$

the estimate $\rho(G_n,\lambda)=O(n^{-1/d})$ and Massart's lower bound. A second and more direct proof simply involves showing that with high probability there exists a cube of volume $\alpha\log n/n$ centered at a grid point but not containing any of the sample points.

Finally, it would be worthwhile to extend both Theorem 1.1 and Corollary 3.1 to general probability measures.

Acknowledgments. This paper originates in conversations with Michel Talagrand, who pointed out to the second author the articles [7] and [10] and their connections to empirical measures. In particular, Talagrand was aware that the determination of Glivenko–Cantelli convergence was a simple consequence of tight estimates for minimax grid matching distances. In dimensions three and higher these estimates were discussed by Tom Leighton and the first author in the course of preparing [7] and [10]. The outline of the proof of these estimates was worked out then, although the details were not written down until now.

REFERENCES

[1] AJTAI, M., KOMLÓS, J. and TUSNÁDY, G. (1984). On optimal matchings. *Combinatorica* **4** 259–264.

[2] COFFMAN, E. and SHOR, P. W. (1991). A simple proof of the $O(n^{1/2}\log^{3/4} n)$ up-right matching bound. *SIAM J. Discrete Math.* To appear.

[3] DUDLEY, R. M. (1969). The speed of mean Glivenko–Cantelli convergence. *Ann. Math. Statist.* **40** 40–50.

[4] DUDLEY, R. M. (1984). A course in empirical processes. *Ecole d'Été de Probabilités de*

Saint-Flour XII—1982. Lecture Notes in Math. **1097** 1–142.

[5] GAENNSLER, P. (1970). A note on a result of Dudley on the speed of Glivenko–Cantelli convergence. *Ann. Math. Statist.* **41** 1339–1343.

[6] KARP, R. M., LUBY, M. and MARCHETTI-SPACCAMELA, A. (1984). Probabilistic analysis of multidimensional bin packing problems. In *Proc. Sixteenth ACM Symp. Theory Computing* 289–298. Assoc. Comp. Mach., New York.

[7] LEIGHTON, T. and SHOR, P. W. (1989). Tight bounds for minimax grid matching, with applications to the average case analysis of algorithms. *Combinatorica* **9** 161–187.

[8] MASSART, P. (1988). About the Prokhorov distance between the uniform distribution over the unit cube in $\mathbb{R}^d$ and its empirical measure. *Probab. Theory Related Fields* **79** 431–450.

[9] RHEE, W. S. and TALAGRAND, M. (1988). Exact bounds for the stochastic upward matching problem. *Trans. Amer. Math. Soc.* **307** 109–125.

[10] SHOR, P. W. (1985). Random planar matching and bin packing. Ph.D. dissertation, MIT.

[11] VARADARAJAN, V. S. (1958). On the convergence of sample probability distributions. *Sankhyā Ser. A* **19** 23–26.

[12] YUKICH, J. E. (1989). Optimal matching and empirical measures. *Proc. Amer. Math. Soc.* **107** 1051–1059.

[13] ZUKER, M. (1974). Speeds of convergence of random probability measures. Ph.D. dissertation, MIT.

AT & T BELL LABORATORIES
600 MOUNTAIN AVE.
MURRAY HILL, NEW JERSEY 07974

DEPARTMENT OF MATHEMATICS
LEHIGH UNIVERSITY
CHRISTMAS-SAUCON HALL 14
BETHLEHEM, PENNSYLVANIA 18015