MATCHING THEOREMS AND EMPIRICAL DISCREPANCY COMPUTATIONS USING MAJORIZING MEASURES

M. TALAGRAND

1. INTRODUCTION

Throughout most of this paper, we will denote by $(U_i)_{i \leq n}$ an independent sequence of points uniformly distributed over the unit square $[0, 1]^2$. We will be concerned with the study of the “fluctuations” of the set $\{U_1, \dots, U_n\}$ and by how far these points are to be spread evenly in $[0, 1]^2$. Consider fixed points $V_1, \dots, V_n$ “spread evenly” in $[0, 1]^2$. For example, if m is the largest integer with $m^2 \leq n$, one can take for $V_1, \dots, V_n$ the points $(k/m, \ell/m)$, $0 \leq k < m$, $0 \leq \ell < m$, together with any subset of $n - m^2$ of the points $(k/m, \ell/m)$, $0 < k, \ell \leq m$, $k = m$ or $\ell = m$. A *matching* of $(U_i)_{i \leq n}$ and $(V_i)_{i \leq n}$ is a permutation π of $\{1, \dots, n\}$ (that matches the point U_i with the point $V_{\pi(i)}$). We will study how “closely” the random points $U_1, \dots, U_n$ can be matched to the fixed points $V_1, \dots, V_n$. (The investigation of different measures of “closeness” will be central to this work.) It was Ajtai, Komlos, and Tusnady who discovered the depth of the subject.

Theorem 1.1 (The Ajtai-Komlós-Tusnady matching theorem [AKT]). *With high probability, there exists a matching π such that $\sum_{i \leq n} \|U_i - V_{\pi(i)}\| \leq K\sqrt{n \log n}$.*

In this statement, $\|u - v\|$ is the euclidean distance of u and v and K denotes a universal constant. (Throughout the paper, unless mentioned otherwise, K will denote a universal constant, not necessarily the same at each occurrence.) With high probability means with probability close to one. (Soon we will indicate a general principle that allows one to compute precisely how close to one it is possible to get.) Theorem 1.1 is only half of the result of Ajtai, Komlós, and Tusnady. They prove that with high probability, for each matching π we have $\sum_{i \leq n} \|U_i - V_{\pi(i)}\| \geq K^{-1}\sqrt{n \log n}$. The problem of lower bounds will not be addressed in the present paper. One reason is that much of the paper is devoted to extensions of Theorem 1.1, where the same lower bounds

Received by the editors March 11, 1991 and, in revised form, June 21, 1993.

1991 *Mathematics Subject Classification*. Primary 60G17, 60G55, 60G57, 60B10; Secondary 46C10.

Work partially supported by a National Science Foundation grant.

©1994 American Mathematical Society
0894-0347/94 \$1.00 + \$.25 per page

hold as in Theorem 1.1. It must be mentioned, however, that lower bounds problems are made considerably easier by the upper bounds methods presented in this paper, since lower bounds can be obtained by the analysis of the “near equality” case in the inequalities that yield to upper bounds.

For a function f on $[0, 1]^2$, we denote throughout the paper by $\iint f d\lambda$ the integral $\iint_{[0, 1]^2} f(x, y) dx dy$ and by $\|f\|_p$ the quantity $(\iint |f|^p d\lambda)^{1/p}$. Consider a collection $\mathcal{E}$ of functions on $[0, 1]^2$. We define the empirical discrepancy $D_n(\mathcal{E})$ by

$$D_n(\mathcal{E}) = \sup_{f \in \mathcal{E}} \left| \sum_{i \leq n} \left(f(U_i) - \iint f d\lambda \right) \right|.$$

This is a random variable, at least when the class $\mathcal{E}$ is smooth enough and, in particular, when $\mathcal{E}$ consists of continuous functions, as is the case in most of the present work. The name “empirical discrepancy” reflects the fact that $n^{-1}D_n(\mathcal{E})$ measures how much the integral of a function f of $\mathcal{E}$ for the “empirical measure” $n^{-1} \sum_{i \leq n} \delta_{U_i}$ can differ from its integral for Lebesgue measure. It can be thought of as a way to measure how evenly the points U_i are spread in the square.

Consider the class

$$\mathcal{E}_{\text{ATK}} = \{f : [0, 1]^2 \rightarrow \mathbb{R}; \forall u, v \in [0, 1]^2, |f(u) - f(v)| \leq \|u - v\|\}.$$

It turns out that Theorem 1.1 is equivalent to the fact that

$$(1.1) \quad ED_n(\mathcal{E}_{\text{ATK}}) \leq K\sqrt{n \log n}.$$

However, the original proof of Theorem 1.1 does not establish (1.1) but rather uses a completely different “transportation” method (of considerable interest of its own).

Theorem 1.1 produces a matching such that the average distance of a point to the matched point is at most $K\sqrt{n^{-1} \log n}$. What would happen if one considers the maximum distance instead of the average distance? This is the object of the following.

Theorem 1.2 (The Leighton-Shor grid matching theorem [LS]). *With high probability, there exists a matching π such that*

$$\max_{i \leq n} \|U_i - V_{\pi(i)}\| \leq Kn^{-1/2}(\log n)^{3/4}.$$

It turns out that Theorem 1.2 is (essentially) equivalent to the fact that

$$(1.2) \quad ED_n(\mathcal{E}) \leq Kn^{1/2}(\log n)^{3/4},$$

where $\mathcal{E}$ is the class of the indicator functions of the interiors of the simple curves in $[0, 1]^2$ with a boundary of length ≤ 1 .

Consider a function f on $[0, 1]^2$ and the random variable

$$X_f = \sum_{i \leq n} \left(f(U_i) - \iint f d\lambda \right).$$

Thus, for a class $\mathcal{E}$ of functions, we have

$$(1.3) \quad D_n(\mathcal{E}) = \sup_{f \in \mathcal{E}} |X_f|.$$

The estimation of such a quantity is a problem of importance in probability, which has received considerable attention. (For a modern presentation of the basic tools, one can consult [LT, Chapters 11 and 12]; more sophisticated results appear in [T3].) The tool that is appropriate for us is called the “majorizing measure bound” and was developed in the early seventies by C. Preston and X. Fernique. This bound relies on “majorizing measures”, which are a kind of system of weights. These majorizing measures are a most effective tool in the general theory of stochastic processes—in particular, of gaussian processes [T1]. On the other hand, the explicit construction of majorizing measures is difficult, and this, up to now, restricted their practical use. Independently of the results of Leighton and Shor (but somewhat later), Rhee and Talagrand [RT] proved an important consequence of Theorem 1.2 (the so-called upright-matching theorem) using majorizing measures. While the construction of the majorizing measure they use is simple, it requires complicated computations to check that it has the desired properties. It must also be said that Leighton and Shor prove (1.2) by a method of spirit close to the majorizing measure method. (This fact should not be surprising; it simply indicates that such is the nature of this question at a rather deep level.) Actually Leighton and Shor reinvented, in a complex situation, some very subtle procedures that took a long time to emerge; this is a truly remarkable achievement. It turns out, however, that the construction of majorizing measures on a class of functions is considerably simplified if one tries to rely on general principles rather than on specific features. It is, in particular, of considerable help to have a suitable parametrization of the class of functions under study. It turns out that in the case of Theorems 1.1 and 1.2 such a parametrization is provided by Fourier transforms. This crucial idea of using Fourier transform (that makes this work possible) was introduced by Coffman and Shor, who used it to give a simple proof of the upright matching theorem [CS]. The beauty of Fourier transform is that it represents certain crucial classes of functions as subsets of certain ellipsoids in Hilbert spaces; and since ellipsoids are simple structures, one is then lead to simple constructions. (The content of §§2–5 of this paper was thus discovered within hours of seeing [CS].)

In §2, we recall what are majorizing measures. These involve a certain integral condition. We are led to consider a variation of the idea involving different integral conditions, and we formulate a general theorem (called the ellipsoid theorem) on the existence of these majorizing measures on certain ellipsoids in Hilbert space. The idea of considering a new integral condition is that this allows, for the ellipsoids of interest, to get statements in the infinite-dimensional Hilbert space. The usual integral condition is recovered from the new condition by using Hölder’s inequality, and this use of Hölder’s inequality explains the origin of the mysterious terms $(\log n)^{1/2}$ and $(\log n)^{3/4}$ in Theorems 1.1 and 1.2. The ellipsoid theorem is proved in §3. In §§4 and 5, we demonstrate the power of our approach by showing that Theorems 1.1 and 1.2 follow simply from the ellipsoid theorem.

After the work of §§2–5 was done, P. Shor informed the author that he had an approach to prove the following results.

Theorem 1.3 (Shor’s discrepancy theorem). *Consider the class*

$$\mathcal{E}_S = \left\{ f : [0, 1] \rightarrow \mathbb{R}; \iint \left| \frac{\partial f}{\partial x} \right| d\lambda \leq 1, \sup \left| \frac{\partial f}{\partial y} \right| \leq 1 \right\}.$$

Then

$$(1.4) \quad ED_n(\mathcal{E}_S) \leq K\sqrt{n \log n}.$$

We denote by $U_{1,i}$, $U_{2,i}$ the components of U_i .

Theorem 1.4 (Shor’s matching theorem). *With high probability, there is a matching π such that*

$$\begin{aligned} \max_{i \leq n} |U_{1,i} - V_{1,\pi(i)}| &\leq K\sqrt{\frac{\log n}{n}}, \\ \frac{1}{n} \sum_{i \leq n} |U_{2,i} - V_{2,\pi(i)}| &\leq K\sqrt{\frac{\log n}{n}}. \end{aligned}$$

(Actually Shor’s original approach created an extra $\sqrt{\log \log n}$ term, which he eventually removed.)

Thus, (1.4) improves on (1.1), and Theorem 1.4 improves on Theorem 1.1. (What is not apparent at this stage but should become apparent later is that Shor’s theorems lie considerably deeper than the results they improve.) Shor’s method appears to be unrelated to the method of the present paper. An important feature of his proof is that the matching is “on line” (something that our method cannot achieve) in the sense that the points U_i (resp. V_i) are given in increasing order of their first coordinates and are matched as soon as they are given. This feature is important for some applications of Shor’s matching theorem to bin packing [S3].

The class $\mathcal{E}_S$ (seen as a subset of L^2) does not seem to be contained in a small enough ellipsoid. Thus the ellipsoid theorem by itself does not suffice to prove Shor’s discrepancy theorem. In §5 we develop more tools for constructing majorizing measures, which build upon the ellipsoid theorem. While the basic ideas of these tools are related to those of the ellipsoid theorem, they require a significant further effort, and the knowledge that Shor has essentially proved Theorem 1.3 was of great help in mustering the required energy.

Roughly speaking, for a class $\mathcal{E}$ of functions on $[0, 1]^2$ having the property that $ED_n(\mathcal{E}) \leq K\sqrt{n \log n}$ results of the conjunction of two facts.

- (I) The class $\mathcal{E}$, considered as a subset of L^2 , is small.
- (II) For most of the choices of $U_1, \dots, U_n$ and for all $f, g \in \mathcal{E}$, the sum $\frac{1}{n} \sum |f(U_i) - g(U_i)|^2$ is small whenever $\|f - g\|_2^2$ is small.

Property (I) is quantified by the existence of certain majorizing measures on $\mathcal{E}$; checking that property is certainly the difficult part. On the other hand, while property (II) is easier to deal with, in the cases we have investigated, it

puts considerably more restrictions on $\mathcal{E}$ than property (I). For example, the class

$$(1.5) \quad \mathcal{E}(\alpha) = \left\{ f : [0, 1]^2 \rightarrow \mathbb{R}; \iint f d\lambda = 0, \left\| \frac{\partial f}{\partial x} \right\|_1 \leq 1, \left\| \frac{\partial f}{\partial y} \right\|_\alpha \leq 1 \right\}$$

satisfies property (I) whenever $\alpha > 1$ but satisfies property (II) only when $\alpha = \infty$. (The reason of the condition $\iint f d\lambda = 0$ is to prevent $\mathcal{E}(\alpha)$ from containing the constants, in which case it would not be small; observe that $D_n(\mathcal{E})$ does not change when each function f of $\mathcal{E}$ is replaced by $f - \iint f d\lambda$.)

In §6, we prove Theorem 1.3 and investigate two natural classes of functions $\mathcal{E}$, each of which satisfies $ED_n(\mathcal{E}) \leq K(n \log n)^{1/2}$. One of these classes is the class

$$\mathcal{E}_U = \left\{ f : [0, 1]^2 \rightarrow \mathbb{R} : \left\| \frac{\partial f}{\partial x} \right\|_1 \leq 1, \int_0^1 \sup_{0 \leq x \leq 1} \left| \frac{\partial f}{\partial y}(x, y) \right| dy \leq 1 \right\}.$$

The importance of this class was discovered by P. Shor, who showed that it has application to the bin-packing problems considered in [KLM-S] and kindly suggested its investigation to us.

Theorems 1.3 and 1.4 are connected, but nonetheless there is a significant difference between them. While Theorem 1.3 relies on the fact that the class $\mathcal{E}_S$ has the properties (I) and (II), Theorem 1.4 is essentially equivalent to the fact that this class has property (I). On the other hand, the classes $\mathcal{E}(\alpha)$ of (1.5) still have property (I); thus it is not surprising that Theorem 1.4 can be improved. This is the purpose of §8. In that section, we prove the following.

Theorem 1.5. *For $\alpha < 1/4$, there exists a constant $K(\alpha)$, depending on α only, with the following property. Consider $M \geq K(\alpha)$. Then, with probability $\geq 1 - n^{-M^2}$, one can find a matching π such that*

$$(1.6) \quad \max_{i \leq n} |U_{1,i} - V_{1,\pi(i)}| \leq KM \sqrt{\frac{\log n}{n}},$$

$$(1.7) \quad \frac{1}{n} \sum_{i \leq n} \exp \left(\frac{|U_{2,i} - V_{2,\pi(i)}|}{KM \sqrt{\frac{\log n}{n}}} \right)^\alpha \leq K.$$

The improvement over Theorem 1.4 is that the “moment condition” on the differences $U_{2,i} - V_{2,\pi(i)}$ is replaced by an exponential-type moment condition.

The reader has certainly observed the unnatural condition $\alpha < 1/4$. We conjecture that Theorem 1.5 holds for $\alpha = 2$. It can be shown that this is best possible. As will be shown in §8, this is actually a problem of constructing majorizing measures. While the methods of the present paper might, with further effort, permit incremental gain on the restriction $\alpha < 1/4$ (which came out of a compromise we make between the needs for a simple proof and for a reasonably competent result), a new idea seems to be needed to reach the case $\alpha = 2$.

The conjecture that Theorem 1.5 holds for $\alpha = 2$ is actually a special case of a more general conjecture. We feel that this conjecture is one of the most important problems remaining in this area, so we would like to state it now.

The ultimate matching conjecture. Consider α_1, α_2 with $1/\alpha_1 + 1/\alpha_2 = 1/2$. Then with high probability we can find a matching π such that, for $j = 1, 2$,

$$(1.8) \quad \frac{1}{n} \sum_{i \leq n} \exp \left(\left| \frac{U_{j,i} - V_{j,\pi(i)}}{K \sqrt{\log n/n}} \right|^{\alpha_j} \right) \leq K.$$

This conjecture is so strong that it might be too daring. The case $\alpha_1 = \infty, \alpha_2 = 2$, where (1.8) has to be interpreted as (1.6), would be the improvement of Theorem 1.5 discussed above. The case $\alpha_1 = \alpha_2 = 4$ would contain both Theorems 1.1 and 1.2.

In §9 we consider a three-dimensional problem, when the points U_i are now distributed uniformly in $[0, 1]^3$. Consider the class $\mathcal{C}$ of indicator functions of convex sets. We show

Theorem 1.6. $ED_n(\mathcal{C}) \leq K n^{1/2} (\log n)^{3/4}$.

Convex sets are closely related to the class of functions f on $[0, 1]^2$ that satisfy

$$\iint \left(\left| \frac{\partial^2 f}{\partial x^2} \right| + \left| \frac{\partial^2 f}{\partial y^2} \right| + \left| \frac{\partial^2 f}{\partial x \partial y} \right| \right) d\lambda \leq 1.$$

As one considers classes of functions in dimension ≥ 2 , there is a tradeoff between the number of dimensions and the stringency of the smoothness conditions required to make the class well behaved. (This tradeoff will be made apparent by the ellipsoid theorem.) The methods of the present paper would allow the study of many such classes for which the terms $(\log n)^{1/2}$ and $(\log n)^{3/4}$ appear; but the purpose of this paper is to establish these methods, not to seek encyclopedial knowledge.

2. MAJORIZING MEASURES AND THE MAJORIZING MEASURE BOUND

Consider a function on $[0, 1]^2$ and the r.v. $X_f = \sum_{i \leq n} (f(U_i) - \iint f d\lambda)$. For our purpose, the most important property of this variable is that it has tails that resemble the tails of a gaussian r.v. Namely, it follows from Bernstein's inequality (see, e.g., [H]) that for $t \leq n\sigma^2/\|f\|_\infty$, we have

$$(2.1) \quad P(|X_f| \geq t) \leq \exp - \frac{t^2}{Kn\sigma^2}$$

where $\sigma^2 = \iint (f - \iint f d\lambda)^2 d\lambda$. This unfortunately fails when t is too large. This failure is a considerable nuisance and could be a source of complication. To avoid this problem, we will make use of a (well-known) device called symmetrization. Throughout the paper, we denote by $(\epsilon_i)_{i \leq n}$ an independent Bernoulli sequence (i.e., $P(\epsilon_i = -1) = P(\epsilon_i = 1) = 1/2$) that is independent of the sequence U_i . For a class of functions $\mathcal{C}$ on $[0, 1]^2$, we set

$$D'_n(\mathcal{C}) = \sup_{f \in \mathcal{C}} \left| \sum_{i \leq n} \epsilon_i f(U_i) \right|.$$

The motivation to introduce this quantity is the following well-known "subgaussian inequality".

Proposition 2.1. Consider real numbers $(a_i)_{i \leq n}$. Then

$$P\left(\sum_{i \leq n} \epsilon_i a_i \geq t\right) \leq \exp -\frac{t^2}{2 \sum_{i \leq n} a_i^2}.$$

Proof. Using series expansion, we have

$$E \exp \lambda \epsilon_i = \frac{1}{2}(e^\lambda + e^{-\lambda}) = \sum_{n \geq 0} \frac{\lambda^{2n}}{(2n)!} \leq \exp \frac{\lambda^2}{2}$$

since $(2n)! \geq 2^n n!$. Thus

$$E \exp \lambda \sum_{i \leq n} \epsilon_i a_i \leq \exp \frac{\lambda^2}{2} \left(\sum_{i \leq n} a_i^2 \right)$$

and

$$P\left(\sum_{i \leq n} \epsilon_i a_i \geq t\right) \leq \inf_{\lambda} (\exp -\lambda) E \exp \lambda \sum_{i \leq n} \epsilon_i a_i. \quad \square$$

We denote throughout the paper by E_ϵ (resp. P_ϵ) the conditional expectation (resp. probability) at $U_1, \dots, U_n$ fixed. By Proposition 2.1, the r.v. $X_f = \sum_{i \leq n} \epsilon_i f(U_i)$ now satisfies

$$(2.2) \quad \forall f \geq 0, \quad P_\epsilon(|X_f| \geq t) \leq 2 \exp \left(-\frac{t^2}{2 \|f\|_U^2} \right)$$

where

$$\|f\|_U = \left(\sum_{i \leq n} f^2(U_i) \right)^{1/2}.$$

The gain over (2.1) is that (2.2) holds for all t ; but the gain comes at a cost, since the denominator of right-hand side of (2.2) now depends on the sample U . However, one can expect $\|f\|_U$ to be close to $n^{1/2} \|f\|_2$; checking that this is indeed the case uniformly on sufficiently many functions will be a significant part of our task.

The following well-known fact relates D_n and D'_n . We reproduce the simple proof for the convenience of the reader.

Proposition 2.2. $ED_n(\mathcal{E}) \leq 2ED'_n(\mathcal{E})$.

Proof. Consider another independent sequence $(U'_i)_{i \leq n}$ of r.v. uniformly distributed in $[0, 1]^2$ that is independent of $(U_i)_{i \leq n}$ and $(\epsilon_i)_{i \leq n}$. Denoting by E' the expectation at U_i, ϵ_i fixed, we have, since $\int f d\lambda = E'(f(U'_i))$,

$$\sup_{f \in \mathcal{E}} \left| \sum_{i \leq n} (f(U_i) - E'(f(U'_i))) \right| \leq E' \sup_{f \in \mathcal{E}} \left| \sum_{i \leq n} (f(U_i) - f(U'_i)) \right|,$$

so that

$$ED_n(\mathcal{E}) \leq E \sup_{f \in \mathcal{E}} \left| \sum_{i \leq n} (f(U_i) - f(U'_i)) \right|.$$

Given fixed signs $(\eta_i)_{i \leq n}$, all the variables $\sup_{f \in \mathcal{E}} |\sum_{i \leq n} \eta_i (f(U_i) - f(U'_i))|$ have the same distribution. Thus,

$$ED_n(\mathcal{E}) \leq E \sup_{f \in \mathcal{E}} \left| \sum_{i \leq n} \epsilon_i (f(U_i) - f(U'_i)) \right| \leq 2ED'_n(\mathcal{E})$$

by the triangle inequality. $\square$

The bounds for $ED'_n(\mathcal{E})$ that we will obtain will always be obtained conditionally on U . In view of (2.2), the main problem is now as follows. Consider a metric space (T, d) and a family of random variables $(X_t)_{t \in T}$ that satisfies

$$(2.3) \quad \forall u, v \in T, \quad P(|X_u - X_v| \geq t) \leq 2 \exp - \frac{t^2}{d(u, v)^2}.$$

What are good bounds for $E \sup_{u \in T} |X_u|$? Surely since (2.3) controls only the differences $X_u - X_v$, it cannot control $E \sup_{u \in T} |X_u|$, so we will rather look for bounds of $E \sup_{u, v \in T} |X_u - X_v|$, keeping in mind that for any $u_0 \in T$ we have

$$E \sup_{u \in T} |X_u| \leq E |X_{u_0}| + E \sup_{u, v \in T} |X_u - X_v|.$$

The bounds one needs depend on the geometry of the space (T, d) . If $N(T, d, \epsilon)$ denotes the smallest number of balls of radius ϵ , for the metric d , that can cover T , a very useful bound (due essentially to Dudley [D]) states that (for some universal constant K)

$$E \sup_{u, v \in T} |X_u - X_v| \leq K \int_0^\infty (\log N(T, d, \epsilon))^{1/2} d\epsilon.$$

(Observe that since $\log 1 = 0$, the integral is actually for values of ϵ less than the diameter of T .)

One main feature of the numbers $N(T, d, \epsilon)$ is that they give the same "weight" to each piece of T . This does not present any inconvenience if T is in some sense homogeneous; in that case the bound above is best possible. But in the situations studied in the present work, this bound fails to give the right order of magnitude (i.e., gives a term in $\log n$). We will have to use the sharper bound given by majorizing measures. This paper will consider extensions of the usual notion of majorizing measures. In order to avoid repetition, we start to give our definitions and will come back later to the classical case. Throughout the paper, we will for simplicity denote by $L(t)$ the function $\log(1/t)$. Consider $\alpha, \beta > 0$, a metric space (T, d) , a probability measure μ on T , and a point $x \in T$. We set

$$\gamma_{\alpha, \beta}(T, d, \mu, x) = \left(\int_0^\infty \epsilon^\beta L(\mu(B(x, \epsilon)))^{\alpha\beta} \frac{d\epsilon}{\epsilon} \right)^{1/\beta}.$$

There, $B(x, \epsilon)$ denotes the (closed) ball centered at x of radius ϵ . We observe that $L(\mu(B(x, \epsilon))) = 0$ if $B(x, \epsilon) = T$, so the integral is actually over the values of ϵ less than the diameter $D(T) = \sup\{d(u, v) : u, v \in T\}$ of T .

We set

$$\gamma_{\alpha, \beta}(T, d, \mu) = \sup_{x \in T} \gamma_{\alpha, \beta}(T, d, \mu, x);$$

$$\gamma_{\alpha, \beta}(T, d) = \inf \gamma_{\alpha, \beta}(T, d, \mu),$$

where the infimum is taken over all the possible choices of μ . The “classical case” is the case $\alpha = 1/2$, $\beta = 1$. The two other cases central to this paper are $\alpha = 1/2$, $\beta = 2$; $\alpha = 1$, $\beta = 2$.

Proposition 2.3 (The majorizing measure bound). *Under condition (2.3), for some universal constant K , we have $E \sup_{u,v \in T} |X_u - X_v| \leq K \gamma_{1/2,1}(T, d)$.*

We refer to [RT, Lemma 3] for a simple proof of this statement. There is actually a whole body of results along the lines of Proposition 2.3. We refer the reader, for example, to [LT, Chapter 11].

When there is no ambiguity on which distance we consider, we will omit the mention of the distance. Unless otherwise specified, classes of functions defined on a subset A of $\mathbb{R}^2$ will always be equipped with the distance

$$d(f, g) = \left(\iint_A (f(x, y) - g(x, y))^2 dx dy \right)^{1/2}$$

induced by $L^2(A)$.

While Proposition 2.3 is certainly sharp, the difficulty in using it is that it requires one to construct majorizing measures. A main step in their construction in the present work is the following result.

Theorem 2.4 (The ellipsoid theorem). *Consider a nonincreasing sequence $(a_n)_{n \geq 1}$ of positive numbers and the ellipsoid $\mathcal{E} = \{(x_n)_{n \geq 1}; \sum_{n \geq 1} |x_n/a_n|^2 \leq 1\}$ of $\ell^2 = \ell^2(\mathbb{N})$. (The scalars are either real or complex.) Consider $\alpha \geq 0$, $0 < \beta \leq 2$. Then, if $\beta < 2$, we have*

$$\gamma_{\alpha, \beta}(\mathcal{E}) \leq K(\alpha, \beta) \left(\sum_{n \geq 1} a_n^\delta n^{\alpha\delta-1} \right)^{1/\delta}$$

where $\delta = \frac{2}{2-\beta}$. If $\beta = 2$, we have

$$\gamma_{\alpha, 2}(\mathcal{E}) \leq K(\alpha) \sup_{n \geq 1} a_n n^\alpha.$$

There the constants $K(\alpha, \beta)$ (resp. $K(\alpha)$) depend only on α, β (resp. α).

We leave to the reader the proof that the inequality of Theorem 2.4 can be reversed (for a different constant!). This theorem will be proved in §3. It will be used in the case $\beta = 2$. We have given the proof in the more general case, since this proof is identical and allows the recovery of the interesting inequality $\gamma_{1/2,1}(\mathcal{E}) \leq K(\sum_{n \geq 1} a_n^2)^{1/2}$. (This inequality follows from the general theory of [T1]. However, its discovery, using an earlier version of the method that we will present here, was the first step of the author in the direction of the results of [T1].)

Another worthy application of Theorem 2.4 is the study of Lévy's Brownian motion on Hilbert space [B1, B2]. This is a Gaussian process $(X_t)_{t \in \ell^2}$ such that $E|X_s - X_t|^2 = \|s - t\|$ for all $s, t \in H$. It follows from the general theory of [T1] that the restriction of a subset T of ℓ^2 is sample bounded if and only if $\gamma_{1/2,1}(T, \sqrt{d}) < \infty$, where d is the canonical distance on H . With

a change of variable, one sees that $\gamma_{1/2,1}(T, \sqrt{d}) = \frac{1}{2}\gamma_{1,1/2}(T, d)^{1/2}$. Thus, when T is an ellipsoid $\mathcal{E}$, Lévy's Brownian motion is sample bounded on $\mathcal{E}$ if and only if $\sum_{n \geq 1} a_n^{4/3} n^{1/3} < \infty$. This could be compared with [B2, Theorem 6.1] (see the "added in proof" mention) where it is shown that the condition $\sum_n n^{-3} a_n^{-2} < \infty$ implies that Lévy's Brownian motion is unbounded on $\mathcal{E}$.

We finish this section by collecting simple facts about majorizing measures. These facts will be used numerous times.

Lemma 2.5. *Consider a metric space (T, d) and $\alpha > 0$, $\beta \geq 1$. Then for some constant $K(\alpha, \beta)$, depending on α, β only, we have*

$$D(T) \leq K(\alpha, \beta) \gamma_{\alpha, \beta}(T, d).$$

Proof. Consider $u, v \in T$, with $d(u, v) > 2D(T)/3$. The balls $B(u, D(T)/3)$, $B(v, D(T)/3)$ are disjoint. Consider a probability μ on T . Then we have $\mu(B(w, D(T)/3)) \leq 1/2$ for either $w = u$ or $w = v$. Thus $L(\mu(B(w, \epsilon))) \geq \log 2$ for $\epsilon \leq D(T)/3$, and hence

$$\begin{aligned} \gamma_{\alpha, \beta}(T, d, \mu, w) &\geq \left(\int_0^{D(T)/3} \epsilon^\beta (\log 2)^{\alpha\beta} \frac{d\epsilon}{\epsilon} \right)^{1/\beta} \\ &= \left(\frac{1}{\beta} \right)^{1/\beta} (\log 2)^\alpha D(T)/3. \quad \square \end{aligned}$$

Lemma 2.6. *Consider two metric spaces (T_1, d_1) , (T_2, d_2) , a probability measure μ on T_1 , a (Borel-measurable) map φ from T_1 to T_2 , and $x \in T_1$. Assume that*

$$(2.4) \quad \forall y \in T_1, \quad d_2(\varphi(y), \varphi(x)) \leq M d_1(x, y).$$

Then

$$(2.5) \quad \gamma_{\alpha, \beta}(T_2, d_2, \varphi(\mu), \varphi(x)) \leq M \gamma_{\alpha, \beta}(T_1, d_1, \mu, x).$$

There, the image measure $\varphi(\mu)$ of μ by φ is given by $\varphi(\mu)(A) = \mu(\varphi^{-1}(A))$ for each Borel set $A \subset T_2$.

Proof. We denote by B_i the ball for d_i ($i = 1, 2$). By (2.4), we have

$$\varphi(B_1(x, \epsilon/M)) \subset B_2(\varphi(x), \epsilon)$$

so that

$$B_1(x, \epsilon/M) \subset \varphi^{-1}(B_2(\varphi(x), \epsilon))$$

and hence

$$\varphi(\mu)(B_2(\varphi(x), \epsilon)) \geq \mu(B_1(x, \epsilon/M)).$$

Thus

$$\begin{aligned} \gamma_{\alpha, \beta}(T_2, d_2, \varphi(\mu), \varphi(x)) &\leq \left(\int_0^\infty \epsilon^\beta L\left(\mu\left(B_1\left(x, \frac{\epsilon}{M}\right)\right)\right)^{\alpha\beta} \frac{d\epsilon}{\epsilon} \right)^{1/\beta} \\ &= M \gamma_{\alpha, \beta}(T_1, d_1, \mu, x) \end{aligned}$$

by making the change of variables $\epsilon = M\eta$. $\square$

Corollary 2.7. Consider two metric spaces (T_1, d_1) , (T_2, d_2) and a map φ from T_1 into T_2 that satisfies

$$(2.6) \quad \forall x, y \in T_1, \quad d_2(\varphi(x), \varphi(y)) \leq M d_1(x, y).$$

Then

$$\gamma_{\alpha, \beta}(T_2, d_2) \leq M \gamma_{\alpha, \beta}(T_1, d_1).$$

Proof. Take the supremum over x and then the infimum over μ in (2.5). $\square$

An interesting case will be when $d_2(\varphi(x), \varphi(y)) = M d_1(x, y)$. Then $\gamma_{\alpha, \beta}(T_2, d_2) = M \gamma_{\alpha, \beta}(T_1, d_1)$.

We will say that a map that satisfies (2.6) is M -Lipschitz and that it is Lipschitz if it is M -Lipschitz for some (unspecified) constant M .

Corollary 2.8. Consider a metric space (T, d) , a subset S of T , and a probability μ on T . Then

$$\gamma_{\alpha, \beta}(S, d) \leq 2 \sup\{\gamma_{\alpha, \beta}(T, d, \mu, x); x \in S\}.$$

In particular, $\gamma_{\alpha, \beta}(S, d) \leq 2 \gamma_{\alpha, \beta}(T, d)$.

Proof. Consider $\delta > 1$. For each $x \in T$, we select a point $\varphi(x)$ in S such that

$$d(x, \varphi(x)) < \delta \inf\{d(x, y); y \in S\}.$$

It is well known that this can be done in a measurable way. For $x \in S$, $y \in T$, we then have

$$d(\varphi(y), x) \leq (1 + \delta)d(y, x)$$

so that

$$\gamma_{\alpha, \beta}(S, d, \varphi(\mu), x) \leq (1 + \delta) \gamma_{\alpha, \beta}(T, d, \mu, x)$$

by Lemma 2.6. Thus, for all $\delta > 1$,

$$\gamma_{\alpha, \beta}(S, d) \leq \gamma_{\alpha, \beta}(S, d, \varphi(\mu)) \leq (1 + \delta) \sup\{\gamma_{\alpha, \beta}(T, d, \mu, x); x \in S\}. \quad \square$$

Proposition 2.9. Consider a sequence (T_n, d_n) of metric spaces. Provide the product $T = \prod_{n \geq 1} T_n$ with the distances d, δ given for $x = (x_n), y = (y_n)$ by

$$(2.7) \quad d(x, y) = \left(\sum_{n \geq 1} d_n^2(x_n, y_n) \right)^{1/2},$$

$$(2.8) \quad \delta(x, y) = \sum_{n \geq 1} d_n(x_n, y_n).$$

Consider for $n \geq 1$ a probability measure μ_n on T_n , and let $a_n = \gamma_{\alpha, \beta}(T_n, \mu_n)$. Consider the product measure $\mu = \bigotimes_n \mu_n$ on T .

(a) Suppose $\alpha \beta \geq 1$. Then

$$(2.9) \quad \gamma_{\alpha, \beta}(T, d, \mu) \leq \left(\sum_{n \geq 1} a_n^{2/(1+2\alpha)} \right)^{1/2+\alpha}$$

$$(2.10) \quad \gamma_{\alpha, \beta}(T, \delta, \mu) \leq \left(\sum_{n \geq 1} a_n^{1/(1+\alpha)} \right)^{1+\alpha}.$$

(b) Suppose that $\alpha\beta \leq 1$. Then

$$(2.11) \quad \gamma_{\alpha,\beta}(T, d, \mu) \leq \left(\sum_{n \geq 1} a_n^{2\beta/(2+\beta)} \right)^{1/2+1/\beta},$$

$$(2.12) \quad \gamma_{\alpha,\beta}(T, \delta, \mu) \leq \left(\sum_{n \geq 1} a_n^{\beta/(\beta+1)} \right)^{1+1/\beta}.$$

Proof. We first consider the case of d . Consider $x = (x_n) \in T$, where $x_n \in T_n$. Consider a sequence $(\eta_n)_{n \geq 1}$, such that $\sum_{n \geq 1} \eta_n^2 = 1$, that will be determined later on. Given $\epsilon > 0$, we observe that

$$\forall n, \quad d_n(y_n, x_n) \leq \eta_n \epsilon \Rightarrow d((y_n), (x_n)) \leq \epsilon.$$

Thus

$$\mu(B_d(x, \epsilon)) \geq \prod_{n \geq 1} \mu_n(B(x_n, \eta_n \epsilon)),$$

and thus

$$(2.13) \quad L(\mu(B_d(x, \epsilon))) \leq \sum_{n \geq 1} L(\mu_n(B(x_n, \eta_n \epsilon))).$$

Suppose first $\alpha\beta \geq 1$. For a function h defined on $[0, \infty)$, define

$$Q(h) = \left(\int_0^\infty \epsilon^{\beta-1} |h(\epsilon)|^{\alpha\beta} d\epsilon \right)^{1/\alpha\beta}.$$

Since $\alpha\beta \geq 1$, this defines a norm (an L_p space with a certain weight function). Using (2.13), and the triangle inequality for Q , we get

$$Q(L(\mu(B_d(x, \cdot)))) \leq \sum_{n \geq 1} Q(L(\mu(B(x, \eta_n \cdot)))).$$

After a change of variable, this means that

$$\begin{aligned} \gamma_{\alpha,\beta}(T, d, \mu, x)^{1/\alpha} &\leq \sum_{n \geq 1} \eta_n^{-1/\alpha} \gamma_{\alpha,\beta}(T_n, d_n, \mu_n, x_n)^{1/\alpha} \\ &\leq \sum_{n \geq 1} \eta_n^{-1/\alpha} a_n^{1/\alpha}. \end{aligned}$$

Optimization over the sequence η_n yields

$$\gamma_{\alpha,\beta}(T, d, \mu, x)^{1/\alpha} \leq \left(\sum_{n \geq 1} a_n^{2/(2\alpha+1)} \right)^{(2\alpha+1)/2\alpha}$$

and hence the result.

Suppose now that $\alpha\beta \leq 1$. We use the inequality $(\sum s_i)^{\alpha\beta} \leq \sum s_i^{\alpha\beta}$ to get that

$$\gamma_{\alpha,\beta}(T, d, \mu, x)^\beta \leq \sum_{n \geq 1} \eta_n^{-\beta} \gamma_{\alpha,\beta}(T_n, d_n, \mu_n, x_n)^\beta,$$

and we finish the proof in the same way.

In the case of δ we observe that if $\sum_{n \geq 1} \eta_n \leq 1$, then

$$\forall n \geq 1, \quad d_n(x_n, y_n) \leq \eta_n \epsilon \Rightarrow d((x_n), (y_n)) \leq \epsilon.$$

The proof is then identical to the proof in the case of d , with the only difference that we now optimize under the condition $\sum_{n \geq 1} \eta_n \leq 1$. $\square$

Corollary 2.10. *Consider subsets $(T_n)_{n \geq 1}$ of a Hilbert space H and their sum*

$$(2.14) \quad \Sigma = \sum_{n \geq 1} T_n = \left\{ \sum_{n \geq 1} x_n; x_n \in T_n \right\}$$

(convergence of the series is part of the hypothesis).

Then, if $\delta = \max(\frac{1}{1+\alpha}, \frac{\beta}{\beta+1})$, we have

$$\gamma_{\alpha, \beta}(\Sigma) \leq \left(\sum_{n \geq 1} \gamma_{\alpha, \beta}(T_n)^\delta \right)^{1/\delta}.$$

Comment. We can certainly assume

$$(2.15) \quad \sum_{n \geq 1} \gamma_{\alpha, \beta}(T_n)^\delta < \infty.$$

Since $\delta \leq 1$, (2.15) and Lemma 2.5 imply $\sum_{n \geq 1} D(T_n) < \infty$. In that case, the series $\sum_{n \geq 1} x_n$, $x_n \in T_n$, converges for all choices of $(x_n)_{n \geq 1}$ if and only if there exists one choice of $(x_n)_{n \geq 1}$ for which it converges. In the cases of application, we will always have $0 \in T_n$, so that convergence of the series $\sum x_n$ will not have to be checked.

Proof. The observation is that the summation $(x_n) \rightarrow \sum x_n$ is a 1-Lipschitz map from the product space $T = \prod_{n \geq 1} T_n$, provided with the distance δ of (2.8), into H . Thus the result follows from Proposition 2.9 and Corollary 2.7. $\square$

3. PROOF OF THEOREM 2.4

The proof is based on a general method of construction of majorizing measures, which is of independent interest (Theorem 3.2). Our first result is not needed for the rest of the paper. However, it provides motivation by showing that the method of Theorem 3.2 is generic.

Proposition 3.1. *Consider a metric space (T, d) of diameter $D(T)$. Consider the largest integer k_0 such that $2^{-k_0} \geq D(T)$. For $k \geq k_0$, $s \geq 0$, we can find subsets $A_k(s)$ of T with the following properties:*

$$(3.1) \quad s \leq s' \Rightarrow A_k(s) \subset A_k(s').$$

$$(3.2) \quad 2^{-\beta k} (\log N(A_k(s), d, 2^{-k}))^{\alpha\beta} \leq s.$$

$$(3.3) \quad \text{Given } t \in T, \text{ we can find numbers } s_k(t) \text{ such that } t \in A_k(s_k(t)) \text{ and}$$

$$\sum_{k \geq k_0} s_k(t) \leq K \gamma_{\alpha, \beta}^\beta(T, d).$$

There, as well as in the rest of this section, K denotes a constant that depends only on α, β and that may vary at each occurrence.

Comment. Consider a sequence $\sigma = (s_k)_{k \geq k_0}$ of positive numbers, and consider the set $A(\sigma) = \bigcap_{k \geq k_0} A_k(s_k)$. It follows from (3.2) that

$$I(A(\sigma)) \leq K \left(\sum_{k \geq k_0} s_k \right)^{1/\beta},$$

where, for a subset A of T , we have set

$$I(A) = \left(\int_0^\infty \epsilon^\beta (\log N(A, d, \epsilon))^{\alpha\beta} \frac{d\epsilon}{\epsilon} \right)^{1/\beta}.$$

It then follows from (3.3) that given $t \in T$, we can find a certain sequence $\sigma(t)$ such that $t \in A(\sigma(t))$ while $I(A(\sigma(t))) \leq K\gamma_{\alpha, \beta}(T, d)$. Thus the sets $A(\sigma)$ for which $I(A(\sigma)) \leq K\gamma_{\alpha, \beta}(T, d)$ cover T . Observe that the integral condition that defines $I(A)$ is similar to the integral condition that defines $\gamma_{\alpha, \beta}(T, d)$ but where majorizing measures are replaced by the more manageable control of entropy numbers. In other words, the majorizing measure condition on T implies the corresponding entropy condition of some pieces of T . The point of this observation is that, as will be shown in Theorem 3.2, there are sufficiently few of these pieces to allow reconstruction of T from them.

Proof. Consider a probability measure μ on T such that $\gamma_{\alpha, \beta}(T, d, \mu) \leq 2\gamma_{\alpha, \beta}(T, d)$. For $k \geq k_0, s > 0$, we set

$$A_k(s) = \left\{ t \in T; 2^{-k\beta} \left(\log \frac{1}{\mu(B(t, 2^{-k-1}))} \right)^{\alpha\beta} \leq s \right\}.$$

Then (3.1) is obvious. To prove (3.2), we consider a subset F of $A_k(s)$ that is maximal with respect to the property that $u, v \in F \Rightarrow d(u, v) > 2^{-k}$. Thus, the balls centered on F of radius 2^{-k} cover $A_k(s)$, so that $N(A_k(s), d, 2^{-k}) \leq \text{card } F$. On the other hand, the balls centered on F of radius 2^{-k-1} are disjoint, so, since μ is a probability, one of these balls has measure $\leq (\text{card } F)^{-1}$, and, by definition of $A_k(s)$, we get $2^{-k\beta} (\log \text{card } F)^{\alpha\beta} \leq s$. This proves (3.2).

To prove (3.3) we simply set

$$s_k(t) = 2^{-\beta k} \left(\log \frac{1}{\mu(B(t, 2^{-k-1}))} \right)^{\alpha\beta}$$

so that $t \in A_k(s_k(t))$. We then have

$$\begin{aligned} \sum_{k \geq k_0} s_k(t) &= \sum_{k \geq k_0} 2^{-\beta k} \left(\log \frac{1}{\mu(B(t, 2^{-k-1}))} \right)^{\alpha\beta} \\ &\leq K \int_0^{D(T)} \epsilon^\beta \left(\log \frac{1}{\mu(B(t, \epsilon))} \right)^{\alpha\beta} \frac{d\epsilon}{\epsilon} \leq K\gamma_{\alpha, \beta}^\beta(T, d). \end{aligned}$$

This completes the proof. $\square$

Theorem 3.2. Consider a metric space (T, d) and a subset S of T , of diameter $D(S)$. Denote by k_0 an integer such that $2^{-k_0} \geq D(S)$, and consider a number $B > 0$. For $k \geq k_0$, $0 \leq q \leq k - k_0$, consider subsets $U_{k,q}$ of T . Assume that, given any $t \in S$, one can find a sequence $(q(k))_{k \geq k_0}$, $0 \leq q(k) \leq k - k_0$, such that $t \in U_{k,q(k)}$ for all $k \geq k_0$ and that

$$(3.4) \quad \sum_{k \geq k_0} 2^{-\beta k} (\log N(U_{k,q(k)}, d, 2^{-k}))^{\alpha\beta} \leq B.$$

Then we have

$$\gamma_{\alpha,\beta}(S, d) \leq K(B^{1/\beta} + 2^{-k_0}).$$

Comment. This statement is actually more general than the converse of Proposition 3.1, as can be seen by taking $U_{k,q} = A_k(C2^{-q})$, for a suitable C .

Proof. For $k \geq k_0$ and $0 \leq q \leq k - k_0$, we consider a subset $F_{k,q}$ of $U_{k,q}$ such that

$$\text{card } F_{k,q} = N(U_{k,q}, d, 2^{-k})$$

and that each point of $U_{k,q}$ is within distance 2^{-k} of a point of $F_{k,q}$. Consider the measure ν on T that, for each $k \geq k_0$ and each $0 \leq \ell \leq k - k_0$, gives mass

$$\frac{2^{-k+k_0-1}}{k - k_0 + 1} \frac{1}{\text{card } F_{k,\ell}}$$

to each point of $F_{k,\ell}$ that is within distance 2^{-k_0} of S . Clearly, the total mass of ν is ≤ 1 .

Consider now $t \in S$. By hypothesis, one can find a sequence $q(k)$, $q(k) \leq k - k_0$, such that $t \in U_{k,q(k)}$ and that (3.4) holds. By construction of $F_{k,q(k)}$, $F_{k,q(k)}$ meets the ball of $B(t, 2^{-k})$. By construction of ν , we thus have

$$\nu(B(t, 2^{-k})) \geq \frac{2^{-k+k_0-1}}{k - k_0 + 1} \frac{1}{\text{card } F_{k,q(k)}}$$

so that

$$2^{-k\beta} \left(\log \frac{1}{\nu(B(t, 2^{-k}))} \right)^{\alpha\beta} \leq K 2^{-k\beta} [(k - k_0 + 1)^{\alpha\beta} + (\log \text{card } F_{k,q(k)})^{\alpha\beta}].$$

Since $\text{card } F_{k,q} = N(U_{k,q}, d, 2^{-k})$, it follows from (3.4) that

$$(3.5) \quad \sum_{k \geq k_0} 2^{-k\beta} \left(\log \frac{1}{\nu(B(t, 2^{-k}))} \right)^{\alpha\beta} \leq K(2^{-k_0\beta} + B).$$

Consider now a probability measure μ on T such that $\mu \geq \nu$. Clearly, one can arrange that $\mu(B(t, 2^{-k_0+1})) = 1$ for all $t \in S$. Thus, from (3.5) we get that

$$\forall t \in S, \quad \gamma_{\alpha,\beta}(T, d, \mu, t) \leq K(2^{-k_0} + B^{1/\beta}),$$

and the conclusion follows from Corollary 2.8. $\square$

Remark. Only trivial changes in the proof are required to show that the result still holds if the range of q is extended to $0 \leq q \leq L(k - k_0 + 1)$ for some $L > 0$ (the constant K depending now on L). This will be used repeatedly in the sequel without further comment.

Proof of Theorem 2.4. We assume first that the scalars are real. The approach is to use Theorem 3.2 for $S = \mathcal{E}$, T the Hilbert space, and d its natural distance.

Step 1. Preparation. We let the reader check, by a homogeneity argument based on Corollary 2.7, that there is no loss of generality to assume that $a_1 = 1/2$. Thus $D(\mathcal{E}) = 2a_1 = 1$, and, with the notation of Theorem 3.2, $k_0 = 0$. (This will slightly simplify notation.)

For $\ell \geq 1$, we set

$$I_\ell = \{n \geq 1 : 2^{-\ell-1} < |a_n| \leq 2^{-\ell}\}.$$

Thus the sets $(I_\ell)_{\ell \geq 1}$ form a partition of $\mathbb{N}$. We set $m_\ell = \text{card } I_\ell$ for $\ell \geq 1$, and we set $m_0 = 0$. For $t \in \ell^2$, we set

$$S_\ell(t) = \sum_{n \in I_\ell} t_n^2.$$

Thus, for $t \in \mathcal{E}$, we have $\sum_{\ell \geq 1} 4^\ell S_\ell(t) \leq 1$.

Step 2. Entropy estimates. For $k \geq 0$, $q \geq 1$ we consider the following subset of ℓ^2 :

$$U_{k,q} = \left\{ t; \sum_{\ell \geq 1} \min(1, 2^{4(\ell-q)}) S_\ell(t) \leq 2^{-2k-2} \right\}.$$

In this step, we prove that

$$\log N(U_{k,q}, 2^{-k}) \leq K \sum_{0 \leq \ell < q} (q - \ell) m_\ell.$$

This computation will be made (in a standard way) through volume estimates. The pleasant fact is that we need only consider ellipsoids and that it is trivial to compare the volume of these to volume of balls.

It is obvious that $\log N(U_{k,1}, 2^{-k}) = 0$, so it suffices to consider the case $q \geq 2$.

First we consider the set $V_{k,q}$ of the sequences $t = (t_n)$ that satisfy the following two properties:

$$\begin{aligned} \sum_{1 \leq \ell < q} 2^{4(\ell-q)} S_\ell(t) &\leq 2^{-2k-2}, \\ n \in \bigcup_{k \geq q} I_k &\Rightarrow t_n = 0. \end{aligned}$$

Then each point t of $U_{k,q}$ is within distance 2^{-k-1} of a point t' of $V_{k,q}$. (To see this, consider the point t' obtained from t by replacing the component t_n by zero whenever $n \in I_k$, $k \geq q$, and observe that $\sum_{\ell \geq q} S_\ell(t) \leq 2^{-2k-2}$.)

Thus

$$N(U_{k,q}, 2^{-k}) \leq N(V_{k,q}, 2^{-k-1}).$$

We set $r = \sum_{1 \leq \ell < q} m_\ell$. We can and will view $V_{k,q}$ as a subset of $\mathbb{R}^r$. We denote by B the unit ball of $\mathbb{R}^r$. We now note that

$$N(V_{k,q}, 2^{-k-1}) \leq \frac{\text{Vol}(V_{k,q} + 2^{-k-2}B)}{\text{Vol}(2^{-k-2}B)}$$

where Vol denotes the r -dimensional volume. Indeed, consider a subset F of $V_{k,q}$, which is maximal with respect to the condition that any two points of F are within distance at least 2^{-k-1} . Then the balls of radius 2^{-k-2} centered on F are disjoint and are contained in $V_{k,q} + 2^{-k-2}B$. Thus the cardinality of F is at most $\text{Vol}(2^{-k-2}B)^{-1} \text{Vol}(V_{k,q} + 2^{-k-2}B)$. On the other hand, the definition of F shows that the balls of radius 2^{-k-1} centered on F cover $V_{k,q}$.

It is obvious that $2^{-k-2}B \subset V_{k,q}$, so that $V_{k,q} + 2^{-k-2}B \subset 2V_{k,q}$. Thus we have

$$N(V_{k,q}, 2^{-k-1}) \leq \frac{\text{Vol}(2V_{k,q})}{\text{Vol}(2^{-k-2}B)} = 2^{r(k+3)} \frac{\text{Vol}(V_{k,q})}{\text{Vol}(B)}.$$

Since $V_{k,q}$ is an ellipsoid, it is easy to compute its volume. Indeed

$$\frac{\text{Vol}(V_{k,q})}{\text{Vol}(B)} = \prod_{1 \leq \ell < q} (2^{-k-1} 2^{2(q-\ell)})^{m_\ell} = 2^{-(k+1)r + 2 \sum_{\ell < q} (q-\ell)m_\ell}.$$

Thus

$$N(V_{k,q}, 2^{-k-1}) \leq 2^{2r+2 \sum_{\ell < q} (q-\ell)m_\ell} \leq 2^{4 \sum_{\ell < q} (q-\ell)m_\ell}.$$

The result follows.

Step 3. Construction of $q(k)$. In this step, we fix $t \in \mathcal{E}$, so that $\sum_{\ell \geq 1} 4^\ell S_\ell(t) \leq 1$. We set

$$u_p = \sum_{\ell \geq 1} 2^{-|\ell-p|} 4^\ell S_\ell(t).$$

Thus $\sum_{p \geq 1} u_p \leq 3$, and, clearly, $1/2 \leq u_{p+1}/u_p \leq 2$. Since $u_p \leq \sum_{\ell \geq 1} 4^\ell S_\ell(t) \leq 1$, we have $4^{-p} u_p \leq 2^{-2k-4}$ for $p \geq k+2$. Thus there exists a smallest number $q(k) \geq 1$ such that $4^{-q(k)} u_{q(k)} \leq 2^{-2k-4}$, and we have $q(k) \leq k+2$.

We now prove that $t \in U_{k,q(k)}$.

Indeed, for $\ell \geq 1$, we have

$$S_\ell(t) \leq 4^{-\ell} 2^{|\ell-q(k)|} u_{q(k)} \leq 4^{q(k)-\ell} 2^{|\ell-q(k)|} 2^{-2k-4}.$$

Thus

$$\begin{aligned} \sum_{\ell \geq q(k)} S_\ell(t) &\leq \sum_{\ell \geq q(k)} 2^{q(k)-\ell} 2^{-2k-4} \leq 2^{-2k-3}, \\ \sum_{\ell < q(k)} 2^{4(\ell-q(k))} S_\ell(t) &\leq \sum_{\ell < q(k)} 2^{-(q(k)-\ell)} 2^{-2k-4} \leq 2^{-2k-3}. \end{aligned}$$

This finishes the proof.

Step 4. Proof of (3.4). For convenience we set $b(q) = \sum_{0 \leq \ell < q} (q - \ell) m_\ell$. It follows from Step 2 that

$$(3.6) \quad \sum_{k \geq 0} 2^{-\beta k} (\log N(U_{k, q(k)}, 2^{-k}))^{\alpha\beta} \leq K \sum_{k \geq 0} 2^{-\beta k} b(q(k))^{\alpha\beta}.$$

We observe that, by definition of $q(k)$, if $q(k) > 1$, we have

$$(3.7) \quad 4^{-q(k)} u_{q(k)} \leq 2^{-2k-4} < 4^{-q(k)+1} u_{q(k)-1} \leq 8 \cdot 4^{-q(k)} u_{q(k)}$$

since $u_{q(k)-1} \leq 2u_{q(k)}$. Since $\log N(U_{k, 1}, 2^{-k}) = 0$, it follows from (3.6) that

$$\sum_{k \geq 0} 2^{-\beta k} (\log N(U_{k, q(k)}, 2^{-k}))^{\alpha\beta} \leq K \sum_{k \geq 0} 2^{-\beta q(k)} u_{q(k)}^{\beta/2} b(q(k))^{\alpha\beta}.$$

We now observe from (3.7) that, unless $q(k) = 1$, there are at most three distinct values of k that can produce a given value of $q(k)$. Thus, since $b(1) = 0$, we have

$$\sum_{k \geq 0} 2^{-\beta q(k)} u_{q(k)}^{\beta/2} b(q(k))^{\alpha\beta} \leq 3 \sum_{\ell \geq 2} 2^{-\beta \ell} u_\ell^{\beta/2} b(\ell)^{\alpha\beta}.$$

Case $\beta = 2$. Since $\sum_{\ell \geq 1} u_\ell \leq 3$, we get that

$$\sum_{k \geq 0} 2^{-2k} (\log N(U_{k, q(k)}, 2^{-k}))^{2\alpha} \leq K \sup_{\ell \geq 2} 2^{-2\ell} b(\ell)^{2\alpha},$$

and Theorem 3.2 implies that

$$(3.8) \quad \gamma_{\alpha, 2}(\mathcal{E}) \leq K \left(1 + \sup_{\ell \geq 2} 2^{-\ell} b(\ell)^\alpha \right).$$

Case $\beta < 2$. We use again $\sum_{\ell \leq 1} u_\ell \leq 3$ and Holder's inequality with exponents $2/\beta, 2/(2-\beta)$ to obtain as above

$$(3.9) \quad \gamma_{\alpha, \beta}(\mathcal{E}) \leq K \left(1 + \left(\sum_{\ell \geq 2} 2^{-\ell\delta} b(\ell)^{\alpha\delta} \right)^{1/\delta} \right).$$

Step 5. We are left with the easy task of showing that the bounds (3.8) and (3.9) are dominated by the bounds of Theorem 2.1. If $\beta = 2$ and if $n^\alpha a_n \leq u$ for $n \geq 1$, then $u \geq 1/2$ since $a_1 = 1/2$. For $a_n \geq 2^{-k-1}$, we have $n^\alpha \leq 2^{k+1}u$, so that $m_k^\alpha \leq 2^{k+1}u$, $m_k \leq 2^{(k+1)/\alpha} u^{1/\alpha}$, so that $b(k) \leq K 2^{(k+1)/\alpha} u^{1/\alpha}$, and $\sup_{\ell \geq 2} 2^{-\ell} b(\ell)^\alpha \leq Ku$, so that, by (3.8) we have $\gamma_{\alpha, 2}(\mathcal{E}) \leq K(1+u) \leq Ku$.

If $\beta > 2$, let us set $S = (\sum_{n \geq 1} a_n^\delta n^{\alpha\delta-1})^{1/\delta}$. (Thus $S \geq 1/2$.) Then

$$S^\delta \geq \sum_{k \geq 1} 2^{-(k+1)\delta} \sum_{n \in I_k} n^{\alpha\delta-1} \geq \frac{1}{K} \sum_{k \geq 1} 2^{-(k+1)\delta} m_k^{\alpha\delta}$$

so that

$$(3.10) \quad \sum_{k \geq 1} 2^{-k\delta} m_k^{\alpha\delta} \leq KS^\delta.$$

We observe now that

$$b(\ell)^{\alpha\delta} = \left(\sum_{k < \ell} (\ell - k) m_k \right)^{\alpha\delta} \leq K \sum_{k < \ell} 2^{\delta(\ell-k)} (\ell - k)^{\alpha\delta} m_k^{\alpha\delta}.$$

Regrouping terms and using (3.10) yields

$$\sum_{\ell \geq 2} 2^{-\ell\delta} b(\ell)^{\alpha\delta} \leq KS^\delta,$$

and this finishes the proof. $\square$

4. THE AJTAI-KOMLÓS-TUSNÁDY THEOREM

We first prove the following, essentially equivalent to Theorem 1.1.

Theorem 4.1. *Consider the class*

$$\mathcal{E}_{\text{AKT}} = \{f : [0, 1]^2 \rightarrow \mathbb{R}; \forall u, v \in [0, 1]^2, |f(u) - f(v)| \leq \|u - v\|\}.$$

Then $ED_n(\mathcal{E}_{\text{AKT}}) \leq K(n \log n)^{1/2}$.

The most urgent task is certainly to explain the appearance of the term $(\log n)^{1/2}$. This is the purpose of the next lemma. From now on, K denotes again a universal constant, not necessarily the same at each occurrence.

Lemma 4.2. *Consider a metric space (T, d) and a subset S of T . Set*

$$a = \inf\{d(s, t); s, t \in S, s \neq t\}.$$

Then

$$\gamma_{1/2,1}(S, d) \leq K \left(1 + \log \frac{D(T)}{a}\right)^{1/2} \gamma_{1/2,2}(S, d).$$

Proof. First, using Corollary 2.8, we can assume that $S = T$. Consider a probability μ on S such that

$$\gamma_{1/2,2}(S, d, \mu) \leq 2\gamma_{1/2,2}(S, d).$$

Given $s \in S$, we have

$$I =: \int_0^{D(T)} L(\mu(B(s, \epsilon)))^{1/2} d\epsilon = \int_0^a + \int_a^{D(T)} =: I_1 + I_2.$$

Now

$$\begin{aligned} I_1 &= aL(\mu(\{s\}))^{1/2} = (a^2L(\mu(\{s\})))^{1/2} \\ &\leq \left(2 \int_0^a \epsilon L(\mu(B(s, \epsilon))) d\epsilon\right)^{1/2} \leq 2\gamma_{1/2,2}(S, d, \mu). \end{aligned}$$

Using Cauchy-Schwarz, we have

$$\begin{aligned} I_2 &= \int_a^{D(T)} \frac{1}{\sqrt{\epsilon}} (\epsilon L(\mu(B(s, \epsilon))))^{1/2} d\epsilon \\ &\leq \left(\int_a^{D(T)} \frac{d\epsilon}{\epsilon} \right)^{1/2} \left(\int_a^{D(T)} \epsilon L(\mu(B(s, \epsilon))) d\epsilon \right)^{1/2} \\ &\leq \left(\log \frac{D(T)}{a} \right)^{1/2} \gamma_{1/2,2}(S, d, \mu). \quad \square \end{aligned}$$

To use that lemma in the present case, we certainly need to know that $\gamma_{1/2,2}(\mathcal{E}) < \infty$. This follows from the next lemma, which is the crucial ingredient of our approach to the Ajtai-Komlos-Tusnaday Theorem. The more general formulation is needed for later purposes.

Lemma 4.3. *Consider $M \geq 0$, and*

$$\mathcal{E}_M = \left\{ f : [0, 1]^2 \rightarrow \mathbb{R}, |f| \leq M, \left\| \frac{\partial f}{\partial x} \right\|_2 \leq M, \left\| \frac{\partial f}{\partial y} \right\|_2 \leq M \right\}.$$

Then $\gamma_{1/2,2}(\mathcal{E}_M) \leq KM$.

Proof. By homogeneity, we can assume $M = 1$. The idea is to use Fourier transform to represent $\mathcal{E} = \mathcal{E}_1$ as a subset of a certain ellipsoid and to use Theorem 2.4. Fourier transform is more elegantly studied using complex numbers. To each function f in $L_2([0, 1]^2)$ it associates the sequence $\mathcal{F}(f) = (a_{n,m}(f))_{n,m \in \mathbb{Z}}$ of complex numbers given by

$$a_{n,m}(f) = \iint_{[0,1]^2} f(x, y) \exp 2i\pi(nx + my) dx dy.$$

The fundamental fact for our purposes is that

$$\|f\|_2 = \left(\sum_{n,m \in \mathbb{Z}} |a_{n,m}(f)|^2 \right)^{1/2},$$

or, equivalently, $\|f\|_2 = \|\mathcal{F}(f)\|_2$ where the norm on the right is the norm on the complex Hilbert space of sequences $(a_{n,m})_{n,m \in \mathbb{Z}}$, i.e., $\|(a_{n,m})\| = (\sum |a_{n,m}|^2)^{1/2}$. By Corollary 2.7, it suffices to prove that $\gamma_{1/2,2}(\mathcal{F}(\mathcal{E})) < \infty$.

We observe that

$$\mathcal{F}(\mathcal{E}) \subset T_1 + T_2 + T_3$$

where T_1 is the set of all sequences $(a_{n,m})_{n,m \in \mathbb{Z}}$ such that for some $f \in \mathcal{E}$, $a_{n,m} = a_{n,m}(f)$ if $|n| \geq |m|$, $|n| \geq 1$, and $a_{n,m} = 0$ otherwise; where T_2 is the set of all sequences $a_{n,m}$ such that for some $f \in \mathcal{E}$, we have $a_{n,m} = a_{n,m}(f)$ if $|m| > |n|$, and $a_{n,m} = 0$ otherwise; and when T_3 is the set of all sequences $a_{n,m}$ for which $a_{n,m} = 0$ unless $n = m = 0$, and $|a_{0,0}| \leq 1$. In view of Corollary 2.10, it suffices to show that $\gamma_{1/2,2}(T_j) < \infty$ for $j = 1, 2, 3$. This is obvious for $j = 3$. The arguments for $j = 1, j = 2$ are identical, so we treat only the case $j = 1$.

Set $A = \{(n, m) \in \mathbb{Z}^2; |n| \geq |m|, |n| \geq 1\}$. Consider now $(n, m) \in A$ and $f \in \mathcal{E}$. Then integration by parts in x yields

$$\begin{aligned} a_{n,m}(f) &= \frac{1}{2i\pi n} \int_0^1 (f(1, y) - f(0, y)) e^{2i\pi my} dy \\ &\quad - \frac{1}{2i\pi n} \iint_{[0,1]^2} \frac{\partial f}{\partial x}(x, y) e^{2i\pi(nx+my)} dx dy \\ (4.1) \quad &= \frac{1}{2i\pi n} b_m(g) - \frac{1}{2i\pi n} a_{n,m} \left(\frac{\partial f}{\partial x} \right) \end{aligned}$$

where we have set $g(y) = f(1, y) - f(0, y)$ and

$$b_m(g) = \int_0^1 g(y) e^{2i\pi m y} dy.$$

This shows that T_1 is (isometric to) a subset of $\frac{1}{2i\pi}(T_4 + T_5)$, where

$$T_4 = \{(\frac{1}{n}b_m(g))_{n,m \in A}; |g| \leq 2\},$$

$$T_5 = \{(\frac{1}{n}a_{n,m}(h))_{n,m \in A}; \|h\|_2 \leq 1\}.$$

Thus, by Corollaries 2.7 and 2.8, it suffices to show that $\gamma_{1/2,2}(T_4) < \infty$, $\gamma_{1/2,2}(T_5) < \infty$.

Consider the ellipsoid

$$\mathcal{E}_1 = \left\{ (b_m)_{m \in \mathbb{Z}, m \neq 0}; \sum_{m \neq 0} m b_m^2 \leq 4 \right\}.$$

Then by Theorem 2.4 we have $\gamma_{1/2,2}(\mathcal{E}_1) < \infty$. Consider now the map $U : \ell^2(\mathbb{Z} \setminus \{0\}) \rightarrow \ell^2(A)$ given by $U((b_m)) = (\frac{\sqrt{m}}{n} b_m)_{n,m \in A}$. Since $\sum_{n \geq m} m/n^2 \leq K$, we see that

$$\|U((b_m)) - U((b'_m))\|_2 \leq K \| (b_m) - (b'_m) \|_2.$$

Thus by Corollary 2.7, we have $\gamma_{1/2,2}(U(\mathcal{E}_1)) < \infty$. Now, since $\sum |b_m(g)|^2 \leq \|g\|_2^2$, we see that for $|g| \leq 2$ we have $(\frac{1}{\sqrt{m}} b_m(g)) \in \mathcal{E}_1$; thus $T_4 \subset U(\mathcal{E}_1)$, and thus $\gamma_{1/2,2}(T_4) < \infty$. We now turn to T_5 . Since $\sum_{n,m} |a_{n,m}(h)|^2 \leq \|h\|_2^2$, we see that

$$T_5 \subset \mathcal{E}_2 = \left\{ (a_{n,m})_{n,m \in A}; \sum n^2 |a_{n,m}|^2 \leq 4 \right\}.$$

We recall that $|m| \leq |n|$ for $(n, m) \in A$. Thus the number of terms $a_{n,m}$ in the definition of $\mathcal{E}_2$ that have a coefficient $\leq 2^{2k}$ is $\leq 2^{2k+2}$; thus, Theorem 2.4 shows that $\gamma_{1/2,2}(\mathcal{E}_2) < \infty$. The proof is complete. $\square$

The following lemma is somewhat too powerful for the present needs. The more general form is provided for later purposes, in order to avoid repetition. The random variable r introduced in that lemma will be used at various places in the sequel. All the rectangles considered in the paper have sides parallel to the axes.

Lemma 4.4. *There exists a random variable r (depending on n) such that $Er \leq K$, with the following property. Consider a rectangle $A \subset [0, 1]^2$. Then*

$$(4.2) \quad \text{card}\{i \leq n; U_i \in A\} \leq rn(|A| + (\log n)/n).$$

In particular, if $|A| \geq \log n/n$, we have

$$(4.3) \quad \text{card}\{i \leq n; U_i \in A\} \leq 2rn|A|.$$

Proof. Consider the largest integer ℓ_0 such that $2^{-\ell_0} \geq (\frac{\log n}{n})^{1/2}$. For $-\ell_0 \leq \ell \leq \ell_0$, we consider the paving $\mathcal{P}_\ell$ of $[0, 1]^2$ by rectangles of sides $2^{-\ell-\ell_0}$, $2^{\ell-\ell_0}$.

For $0 \leq x \leq 1$, we observe that $e^x \leq 1 + (e-1)x \leq 1 + 2x$. Thus, for a r.v. Z that satisfies $0 \leq Z \leq 1$, we have

$$E \exp Z \leq 1 + 2EZ \leq \exp 2EZ.$$

Consider now $A \subset [0, 1]^2$, and set $Z_i = 1_A(U_i)$. Thus

$$E \exp \sum_{i \leq n} Z_i = \prod_{i \leq n} E \exp Z_i \leq \exp 2n|A|$$

so that

$$P\left(\sum_{i \leq n} Z_i \geq t\right) \leq \exp(2n|A| - t),$$

and hence

$$(4.4) \quad P(\text{card}\{i \leq n; U_i \in A\} \geq 2n|A| + t) \leq \exp -t.$$

Each paving $\mathcal{P}_\ell$ consists of $2^{2\ell_0}$ rectangles, so a total of $(2\ell_0 + 1)2^{2\ell_0} \leq Kn$ rectangles. Define

$$r' = \max\{\text{card}\{i \leq n; U_i \in A\}; A \in \mathcal{P}_\ell, -\ell_0 \leq \ell \leq \ell_0\}.$$

Thus, by (4.4) we have

$$P(r' \geq n2^{-2\ell_0} + t) \leq Kn \exp -t.$$

Thus, since $n2^{-2\ell_0} \leq K \log n$, by a standard computation we have $Er' \leq K \log n$. Consider now a rectangle $A \subset [0, 1]^2$, of sides of length a, b . Consider the smallest $\ell \geq -\ell_0$ such that $a \leq 2^{-\ell-\ell_0}$. Thus $2^{-\ell-\ell_0} \leq 2a$ if $\ell > -\ell_0$, so that, in any case $2^{-\ell-\ell_0} \leq 2a + 2^{-2\ell_0}$. It is simple to see that A can be covered by at most $2b2^{\ell-\ell_0} + 4$ rectangles of $\mathcal{P}_\ell$. Thus A contains at most $r'(2b2^{-\ell+\ell_0} + 4)$ points U_i . Now, since $b \leq 1$,

$$\begin{aligned} r'(2b2^{-\ell+\ell_0} + 4) &= r'2^{2\ell_0}(2b2^{-\ell-\ell_0} + 4 \cdot 2^{-2\ell_0}) \\ &\leq r'2^{2\ell_0}(4ab + 6 \cdot 2^{-2\ell_0}) \leq K \frac{r'}{\log n} n \left(|A| + \frac{\log n}{n}\right). \end{aligned}$$

This proves the result, with $r = Kr' / (\log n)$. $\square$

We will now relate, for two functions f, g of $\mathcal{C}$, their L_2 distance and their random distance

$$(4.5) \quad d_U(f, g) = \|f - g\|_U = \left(\sum_{i \leq n} (f - g)(U_i)^2\right)^{1/2}.$$

We introduce some notation which will remain in force throughout the paper. For a function h on $\mathbb{R}^2$ and a measurable set $A \subset \mathbb{R}^2$, we set

$$\|h\|_{\infty, A} = \sup\{|h(x)|; x \in A\}, \quad m_A(h) = \frac{1}{|A|} \iint_A h d\lambda.$$

Lemma 4.5. Consider $f, g \in \mathcal{E}_{\text{AKT}}$ and r defined in Lemma 4.3. Then

$$(4.6) \quad d_U(f, g) \leq \sqrt{rn} \left(\|f - g\|_2 + K \sqrt{\frac{\log n}{n}} \right).$$

Proof. Set $\mathcal{P} = \mathcal{P}_0$. Let $h = f - g$. Obviously, since each set $A \in \mathcal{P}$ contains at most $2nr2^{-2\ell_0}$ points U_i , we have

$$\sum_{i \leq n} h(U_i)^2 \leq 2^{-2\ell_0+1} rn \sum_{A \in \mathcal{P}} \|h\|_{\infty, A}^2.$$

Consider the function $E_{\mathcal{P}}(h)$ defined by

$$E_{\mathcal{P}}(h)(u) = m_A(h) = \frac{1}{|A|} \iint_A h d\lambda$$

for $u \in A \in \mathcal{P}$. Since h is 2-Lipschitz and since A is of diameter $\leq 2^{-\ell_0}$, we have

$$\|h\|_{\infty, A} \leq |m_A(h)| + K2^{-\ell_0}.$$

Thus, by the triangle inequality, we have

$$\begin{aligned} \|h\|_U &\leq \sqrt{rn} \left(\left(\sum_{A \in \mathcal{P}} 2^{-2\ell_0+1} m_A^2(h) \right)^{1/2} + K2^{-\ell_0} \right) \\ &= \sqrt{rn} (\|E_{\mathcal{P}}(h)\|_2 + K2^{-\ell_0}). \end{aligned}$$

This is the result, since $\|E_{\mathcal{P}}(h)\|_2 \leq \|h\|_2$. $\square$

We can now prove Theorem 4.1. Obviously to prove that theorem, one can replace the class $\mathcal{E}_{\text{AKT}}$ by the class $\mathcal{E}' = \{f \in \mathcal{E}_{\text{AKT}}, f(1/2, 1/2) = 0\}$. Since $\mathcal{E}' \subset \mathcal{E}_1$, we have $\gamma_{1/2, 2}(\mathcal{E}') < \infty$ by Lemma 4.3.

Consider now a subset $\mathcal{E}''$ of $\mathcal{E}'$ that is maximal with respect to the property that $\|f - g\|_2 \geq (\log n/n)^{1/2}$ for $f, g \in \mathcal{E}''$. It follows from Lemma 4.2 that $\gamma_{1/2, 1}(\mathcal{E}'') \leq K\sqrt{\log n}$. It follows from Lemma 4.5 that

$$f, g \in \mathcal{E}'' \Rightarrow d_U(f, g) \leq K\sqrt{rn}\|f - g\|_2.$$

Thus, it follows from Corollary 2.7 that

$$\gamma_{1/2, 1}(\mathcal{E}'', d_U) \leq K\sqrt{rn \log n}.$$

The majorizing measure bound (Proposition 2.3) shows now that

$$(4.7) \quad E_{\epsilon} \sup_{f \in \mathcal{E}''} \left| \sum_{i \leq n} \epsilon_i f(U_i) \right| \leq K\sqrt{rn \log n}.$$

Consider $g \in \mathcal{E}'$. We can find $f \in \mathcal{E}''$ such that $\|h\|_2 \leq 2(\log n/n)^{1/2}$, where $h = f - g$. By Lemma 4.5, we have

$$\left(\sum_{i \leq n} h(U_i)^2 \right)^{1/2} \leq K\sqrt{r \log n}.$$

Thus by Cauchy-Schwarz, we have

$$\sum_{i \leq n} |h(U_i)| \leq K \sqrt{rn \log n}$$

so that in particular

$$\left| \sum_{i \leq n} \epsilon_i (f(U_i) - g(U_i)) \right| \leq K \sqrt{rn \log n}.$$

Combined with (4.7), this gives

$$E_\epsilon \sup_{g \in \mathcal{G}'} \left| \sum_{i \leq n} \epsilon_i g(U_i) \right| \leq K \sqrt{rn \log n}.$$

Since, by Lemma 4.4, we have $E\sqrt{r} \leq \sqrt{Er} \leq K$, we have

$$E \sup_{g \in \mathcal{G}'} \left| \sum_{i \leq n} \epsilon_i g(U_i) \right| \leq K \sqrt{n \log n}.$$

The proof is complete by Proposition 2.2. $\square$

We now show how to deduce Theorem 1.1 from Theorem 4.1. The argument is certainly not new (see, e.g., [S1]), but we present it since more elaborate versions will be used later.

The basic tool to construct matchings is the following “duality” method (see, e.g., [PS]).

Proposition 4.5. *Consider a matrix $H = (h_{ij})_{i,j \leq n}$, where $h_{ij} \geq 0$. Set*

$$M(H) = \inf \sum_{i \leq n} h_{i\pi(i)}$$

where the infimum is computed over all permutations π of $\{1, \dots, n\}$. Then

$$(4.8) \quad M(H) = \sup \left(\sum_{i \leq n} w_i - \sum_{i \leq n} w'_i \right)$$

where the supremum is over all families $(w_i)_{i \leq n}, (w'_i)_{i \leq n}$ that satisfy

$$(4.9) \quad \forall i, j \leq n, \quad w_i \leq w'_j + h_{ij}.$$

In our case, we want to compute $M(H)$ when $h_{ij} = \|U_i - V_j\|$. Consider numbers $(w_i)_{i \leq n}, (w'_i)_{i \leq n}$ that satisfy (4.8). For $u \in [0, 1]^2$, we set

$$f(u) = \min_{i \leq n} (w'_i + \|u - V_i\|).$$

Thus $f(V_i) \leq w'_i$, and, by (4.9), $w_i \leq f(U_i)$. It follows that

$$\begin{aligned} \sum_{i \leq n} w_i - \sum_{i \leq n} w'_i &\leq \sum_{i \leq n} f(U_i) - \sum_{i \leq n} f(V_i) \\ &\leq \left| \sum_{i \leq n} \left(f(U_i) - \iint f d\lambda \right) \right| + \left| \sum_{i \leq n} \left(f(V_i) - \iint f d\lambda \right) \right|. \end{aligned}$$

For each i , the function $w'_i + \|u - V_i\|$ of u is 1-Lipschitz by the triangle inequality. Thus f is also 1-Lipschitz, i.e., $f \in \mathcal{C}_{\text{AKT}}$. Thus we have

$$M(H) \leq \sup_{f \in \mathcal{C}_{\text{AKT}}} \left| \sum_{i \leq n} \left(f(U_i) - \iint f d\lambda \right) \right| + \sup_{f \in \mathcal{C}_{\text{AKT}}} \left| \sum_{i \leq n} \left(f(V_i) - \iint f d\lambda \right) \right|.$$

Since the points V_i are evenly distributed, we get

$$M(H) \leq D_n(\mathcal{C}_{\text{AKT}}) + K\sqrt{n}.$$

It follows from Theorem 4.1 that $ED_n(\mathcal{C}_{\text{AKT}}) \leq K\sqrt{n \log n}$. We will have proved Theorem 1.1 if we show that with high probability we have $D_n(\mathcal{C}_{\text{AKT}}) \leq L\sqrt{n \log n}$, for some constant L .

This actually holds with probability $\geq 1 - n^{-L^2/K}$ for L large enough, as is shown by the following principle. (Observe that

$$D_n(\mathcal{C}_{\text{AKT}}) = D_n(\{f \in \mathcal{C}_{\text{AKT}}; f(1/2, 1/2) = 0\}).$$

Proposition 4.6. *Consider a class $\mathcal{C}$ of functions and $M = \sup\{\|f\|_\infty; f \in \mathcal{C}\}$. Then*

$$P(|D_n(\mathcal{C}) - ED_n(\mathcal{C})| \geq t) \leq 2 \exp - \frac{t^2}{4nM}.$$

Proof. This is a straightforward application of the method of martingale difference sequences and of Azuma's inequality as for example in [LT, §1.3]. $\square$

Later, we will need a considerable sharpening of Proposition 4.6 (Proposition 5.4).

5. THE LEIGHTON-SHOR GRID MATCHING THEOREM

We fix n and a parameter $L \geq 1$. We denote by G_n the grid on $[0, 1]^2$ made of squares of side $2^{-\ell_1}$, where ℓ_1 is the largest integer such that $2^{-\ell_1} \geq Ln^{-1/2}(\log n)^{3/4}$. We denote by $\mathcal{C}$ the class of simple curves traced on G_n . For a simple curve C , we denote by $\overset{\circ}{C}$ the bounded region it defines and by $\ell(C)$ its length. We will prove the following two facts.

Fact 1. *With probability $\geq 1 - \exp(-L^2(\log n)^{3/2}/K)$, we have*

$$(5.1) \quad \forall C \in \mathcal{C}, \quad \left| \sum_{i \leq n} 1_{\overset{\circ}{C}}(U_i) - n|\overset{\circ}{C}| \right| \leq K_1 Ln^{1/2}(\log n)^{3/4} \ell(C)$$

where K_1 is universal.

Fact 2. *When (5.1) occurs, we have a matching π such that*

$$\max_{i \leq n} \|U_i - V_{\pi(i)}\| \leq KLn^{-1/2}(\log n)^{3/4}.$$

For the convenience of the reader, we first sketch the proof of Fact 2 [LS, S1].

Using the Marriage Lemma, it suffices to show that, when (5.1) occurs, for some constant K_2 and all subsets I of $\{1, \dots, n\}$ we have

$$\text{card}\{j \leq n; \exists i \in I, \|U_i - V_j\| \leq K_2 L n^{-1/2} (\log n)^{3/4}\} \geq \text{card } I.$$

Consider the grid G'_n made of squares of side $2^{-\ell_2}$, where $\ell_2 \leq \ell_1$ will be chosen later. Given a subset I of $\{1, \dots, n\}$, consider the region R that consists of all the squares of G'_n that contain at least one point U_i , $i \in I$. Consider the region R' consisting of all the squares of G'_n that are adjacent to R . Assuming that $2^{-\ell_2} \leq K_2 L n^{-1/2} (\log n)^{3/4}$, it suffices to show that

$$\text{card}\{j \leq n; V_j \in R'\} \geq \text{card } I.$$

Since the points V_j are spread uniformly, it suffices to show that

$$\text{card } I \leq n(|R| + \frac{1}{2}|R' \setminus R|).$$

A square of G'_n in $R' \setminus R$ is adjacent to at most eight connex components of R . Thus it suffices to show that when R is connex, we have

$$\text{card } I \leq n(|R| + \frac{1}{16}|R' \setminus R|).$$

Since we assume that R is connex, it is the interior of a simple curve minus a number of "holes", that are themselves the interior of simple curves. By difference one sees that it suffices to show that for the interior $\overset{\circ}{C}$ of a simple curve C , one has

$$n(|\overset{\circ}{C}| - \frac{1}{16}|\overset{\circ}{C} \setminus R''|) \leq \text{card}\{i \leq n; U_i \in \overset{\circ}{C}\} \leq n(|\overset{\circ}{C}| + \frac{1}{16}|R' \setminus \overset{\circ}{C}|)$$

where R' (resp. R'') is the union of the squares of G'_n that touch $\overset{\circ}{C}$ (resp. touch $[0, 1]^2 \setminus \overset{\circ}{C}$). We have $|\overset{\circ}{C} \setminus R''| \geq 2^{-\ell_2-2}\ell(C)$, and we have $|R' \setminus \overset{\circ}{C}| \geq 2^{-\ell_2-2}\ell'(C)$, where $\ell'(C)$ is the length of the part of the boundary of C that is not traced on the boundary of $[0, 1]^2$; by replacing if necessary C by $[0, 1]^2 \setminus C$, one sees that it suffices to consider the case where $\ell'(C) \geq \ell(C)/4$. Finally, one has reduced to showing that

$$\left| \text{card}\{i \leq n; U_i \in \overset{\circ}{C}\} - n|\overset{\circ}{C}| \right| \leq 2^{-\ell_2-8}\ell(C).$$

This follows from (5.1) provided $2^{-\ell_2-8} \geq K_1 L n^{-1/2} (\log n)^{3/4}$, where K_1 is the constant of (5.1). (Note that a curve traced on G'_n is a curve traced on G_n .)

We now turn to the principle that will create the unusual term $(\log n)^{3/4}$.

Lemma 5.1. *Consider a metric space (T, d) , and a finite subset S of T . Then if $D(T)$ denotes the diameter of T , we have*

(a) $\forall a > 0$,

$$\gamma_{1/2,1}(S, \sqrt{d}) \leq K \left(1 + \log \frac{D(T)}{a} \right)^{3/4} \gamma_{1,2}(T, d)^{1/2} + (a \log(2 \text{ card } S))^{1/2}.$$

(b) If $a = \inf\{d(s, t); s, t \in S, s \neq t\}$, we have

$$\gamma_{1/2,1}(S, \sqrt{d}) \leq K \left(1 + \log \frac{D(T, d)}{a}\right)^{3/4} \gamma_{1,2}(T, d)^{1/2}.$$

Comment. Only (b), which is somewhat simpler, will be needed here. The proof of (a), which will be needed later, is given at this point in order to avoid repetition.

Proof. (a) Using Corollary 2.8 we reduce to the case $T = S$. Consider a probability μ on S such that $\gamma_{1,2}(S, d, \mu) \leq 2\gamma_{1,2}(S, d)$. Let $N = \text{card } S$. Consider the probability ν on S given by $\nu = \frac{1}{2}(\mu + \mu')$, where μ' gives mass $\frac{1}{N}$ to each point of S . Consider $s \in S$. We observe that, with obvious notation, $B_{\sqrt{d}}(s, \epsilon) = B_d(s, \epsilon^2)$. Thus setting $D = D(T, d)$, we have

$$\begin{aligned} \int_0^{D(T, \sqrt{d})} L(\nu(B_{\sqrt{d}}(s, \epsilon)))^{1/2} d\epsilon &= \int_0^{D^{1/2}} L(\nu(B_d(s, \epsilon^2)))^{1/2} d\epsilon \\ &= \frac{1}{2} \int_0^D L(\nu(B_d(s, \epsilon)))^{1/2} \frac{d\epsilon}{\sqrt{\epsilon}} \end{aligned}$$

by change of variable. We have

$$\frac{1}{2} \int_0^a L(\nu(B_d(s, \epsilon)))^{1/2} \frac{d\epsilon}{\sqrt{\epsilon}} \leq (\log 2N)^{1/2} \int_0^a \frac{d\epsilon}{2\sqrt{\epsilon}} = (a \log 2N)^{1/2}.$$

We have

$$\frac{1}{2} \int_a^D L(\nu(B_d(s, \epsilon)))^{1/2} \frac{d\epsilon}{\sqrt{\epsilon}} \leq \sqrt{D}(\log 2)^{1/2} + \frac{1}{2} \int_a^D L(\mu(B_d(s, \epsilon)))^{1/2} \frac{d\epsilon}{\sqrt{\epsilon}}.$$

Using Holder's inequality, we have

$$\begin{aligned} \int_a^D L(\mu(B_d(s, \epsilon)))^{1/2} \frac{d\epsilon}{\sqrt{\epsilon}} &= \int_a^D \frac{1}{\epsilon^{3/4}} \epsilon^{1/4} (L(\mu(B_d(s, \epsilon))))^{1/2} d\epsilon \\ &\leq \left(\int_a^D \frac{d\epsilon}{\epsilon} \right)^{3/4} \left(\int_a^D \epsilon L(\mu(B_d(s, \epsilon)))^2 d\epsilon \right)^{1/4} \\ &\leq \left(\log \frac{D}{a} \right)^{3/4} \gamma_{1,2}(S, \mu, d)^{1/2}. \end{aligned}$$

The result then follows from the fact that $D \leq K\gamma_{1,2}(S, \mu, d)$, as was noted in Lemma 2.5.

(b) For $s \in S$, we have

$$2a^2 L(\mu(\{s\}))^2 \leq \int_0^a \epsilon L(\mu(B_d(s, \epsilon)))^2 d\epsilon \leq \gamma_{1,2}(S, \mu)^2.$$

Taking s such that $\mu(\{s\}) \leq 1/N$ yields

$$(a \log N)^{1/2} \leq \gamma_{1,2}(S, \mu)^{1/2},$$

so (b) follows from (a). $\square$

To use that result, we need to control the functional $\gamma_{1,2}$ of certain sets. The following is the basic step.

Lemma 5.2. *Consider the class $\mathcal{S}$ of functions on $[0, 1]$ such that $|f| \leq 1$, $\int_0^1 f'^2(x) dx \leq 1$. Then $\gamma_{1,2}(\mathcal{S}) < \infty$.*

Comments. (1) For the purposes of the present section, we could moreover require that $f(0) = f(1)$.

(2) Using methods that will be presented later in this article, one can show that the condition $\int_0^1 f'^2(x) dx \leq 1$ can be weakened into $\int_0^1 |f'(x)| dx \leq 1$. This is, however, not easy, and since we do not have applications of this result, we do not present the proof.

Proof. We use Fourier transform again. Integration by part shows that the Fourier coefficients $b_n(f)$ satisfy $b_n(f) = \frac{1}{n}(f(1) - f(0)) + b'_n(f)$, where $b'_n(f)$ satisfy $\sum_{n \in \mathbb{Z}} n^2 |b'_n(f)|^2 \leq 1$; and since $|f| \leq 1$, we have $|b_0(f)| \leq 1$. Thus the result follows from Theorem 2.4. (and Corollary 2.10). $\square$

We now turn to curves. We denote by $\mathcal{R}$ the class of piecewise smooth simple closed curves of length ≤ 1 that are contained in $[-1, 1]^2$. We provide $\mathcal{R}$ with the distance given by $\delta(C_1, C_2) = |\dot{C}_1 \Delta \dot{C}_2|$.

Proposition 5.3. $\gamma_{1,2}(\mathcal{R}, \delta) < \infty$.

Proof. Given $f, g \in \mathcal{S}$, we consider the curve $C(f, g)$ that is the set of points $(f(t), g(t))$, $0 \leq t \leq 1$. We denote by $\mathcal{T}$ the set of couples (f, g) such that $C(f, g) \in \mathcal{R}$. We observe that each $C \in \mathcal{R}$ is of the type $C(f, g)$ (as is clear if one parametrizes C using the arc length). Also, it follows from Lemma 4.2 and Proposition 2.9 that $\gamma_{1,2}(\mathcal{S} \times \mathcal{S}) < \infty$, so that $\gamma_{1,2}(\mathcal{T}) < \infty$ by Corollary 2.8. We will now show that

$$(5.2) \quad \forall f_1, f_2, g_1, g_2 \in \mathcal{S}, \\ \delta(C(f_1, g_1), C(f_2, g_2)) \leq (\|f_1 - f_2\|_2^2 + \|g_1 - g_2\|_2^2)^{1/2}.$$

Appealing to Corollary 2.7 then concludes the proof.

Consider the function $u(s, t)$ from $[0, 1]^2$ to $[-1, 1]$ given by

$$u(s, t) = (sf_1(t) + (1-s)f_2(t), sg_1(t) + (1-s)g_2(t)).$$

The Jacobian $J(s, t)$ is

$$(5.3) \quad \begin{aligned} J(s, t) &= (sg'_1(t) + (1-s)g'_2(t))(f_1(t) - f_2(t)) \\ &\quad \cdot (sf'_1(t) + (1-s)f'_2(t))(g_1(t) - g_2(t)) \\ &\leq (sf_1'^2(t) + (1-s)f_2'^2(t) + sg_1'^2(t) + (1-s)g_2'^2(t))^{1/2} \\ &\quad \cdot ((f_1(t) - f_2(t))^2 + (g_1(t) - g_2(t))^2)^{1/2}. \end{aligned}$$

using Cauchy-Schwarz and the convexity of the function $x \rightarrow x^2$. The area A of the image domain by u satisfies

$$A \leq \iint_{[0,1]^2} |J(s, t)| ds dt.$$

Computation using (5.3) and Cauchy Schwarz yields

$$A \leq (\|f_1 - f_2\|_2^2 + \|g_1 - g_2\|_2^2)^{1/2}.$$

It remains to show that the domain of u contains $\overset{\circ}{C}_1 \Delta \overset{\circ}{C}_2$, where $C_i = C(f_i, g_i)$, $i = 1, 2$. Consider the curve B_s of parametrization $t \rightarrow u(s, t)$. If a point v does not belong to the domain of u , the number of times that B_s goes around it is independent of s . Since $B_0 = C_1$, $B_1 = C_2$, we have either $v \in \overset{\circ}{C}_1 \cap \overset{\circ}{C}_2$ or $v \notin \overset{\circ}{C}_1$, $v \notin \overset{\circ}{C}_2$. The proof is complete. $\square$

Proof of Fact 1. Step 1. For $k \in \mathbb{Z}$, we consider the set H_k of points of $[0, 1]^2$, coordinates of which are multiples of 2^k (so that H_k is reduced to the origin for $k > 0$). For $u \in H_k$, we denote by $\mathcal{R}_{k,u}$ the set of curves traced on G_n that are of length $\leq 2^k$ and that are included in the intersection of $[0, 1]^2$ and of $u + [-2^k, 2^k]^2$. We consider the random variables

$$Y_{k,u} = \frac{1}{2^k} \sup_{C \in \mathcal{R}_{k,u}} \left| \sum_{i \leq n} (1_{\overset{\circ}{C}}(U_i) - |\overset{\circ}{C}|) \right|,$$

$$X_{k,u} = \frac{1}{2^k} \sup_{C \in \mathcal{R}_{k,u}} \left| \sum_{i \leq n} \epsilon_i 1_{\overset{\circ}{C}}(U_i) \right|.$$

Consider the set S of couples (k, u) , where $-\ell_1 \leq k \leq 2 \log 2n$, $u \in H_n$. We observe by crude estimates that $\text{card } S \leq n$. We set

$$Y = \sup_{(k,u) \in S} Y_{k,u}.$$

We show that

$$\forall C \in \mathcal{R}, \quad \left| \sum_{i \leq n} (1_{\overset{\circ}{C}}(U_i) - |\overset{\circ}{C}|) \right| \leq 2Y\ell(C).$$

Indeed, consider $C \in \mathcal{R}$ and k with $2^{k-1} \leq \ell(C) \leq 2^k$. Since C is a simple curve traced on G_n , its length must satisfy $\ell(C) \leq 2^{-\ell_1} (2^{\ell_1} + 1)^2 \leq n$, so that $2^k \leq 2n$, and hence $k \leq 2 \log 2n$. It should be clear that $C \in \mathcal{R}_{k,u}$ for a certain $u \in H_n$. Since $\ell(C) \geq 2^{k-1}$, we have

$$\frac{1}{\ell(C)} \left| \sum_{i \leq n} 1_{\overset{\circ}{C}}(U_i) - |\overset{\circ}{C}| \right| \leq 2Y_{k,u} \leq 2Y.$$

Step 2. We have reduced the proof to showing the following: with probability $\geq 1 - \exp(-L^2(\log n)^{3/2}/K)$, we have $Y \leq KL\sqrt{n}(\log n)^{3/4}$.

Since $\text{card } S \leq n$, it suffices that each variable $Y_{k,u}$ has this property. In this step, we investigate $X_{k,u}$. We observe that a dilatation of factor t multiplies areas by t^2 . So, if we appeal to Corollaries 2.7 and 2.8 and Proposition 5.3, we see that $\gamma_{1,2}(\mathcal{R}_{k,u}, \delta) \leq K2^{2k}$ for all $u \in H_k$. Two different curves traced

on G_n are at a distance $\geq 2^{-2\ell_1} \geq 1/n$. The diameter of $\mathcal{R}_{k,u}$ is ≤ 1 . Thus, Lemma 5.1(b) shows that

$$\gamma_{1/2,1}(\mathcal{R}_{k,u}, \sqrt{\delta}) \leq K(\log n)^{3/4} 2^k.$$

On $\mathcal{R}_{k,u}$, consider the random distance

$$\begin{aligned} d_U(C, C') &= \left(\sum_{i \leq n} (1_{\hat{C}'}(U_i) - 1_{\hat{C}}(U_i))^2 \right)^{1/2} \\ &= (\text{card}\{i \leq n : U_i \in \hat{C} \Delta \hat{C}'\})^{1/2}. \end{aligned}$$

Since $\hat{C} \Delta \hat{C}'$ is union of squares of G_n , we see that

$$d_U(C, C') \leq \sqrt{r'n} \sqrt{\delta(C, C')},$$

where

$$r' = \max\left\{\frac{1}{n|A|} \text{card}\{i \leq n : U_i \in A\}; A \text{ square of } G_n\right\}.$$

Thus, by Corollary 2.7 we have

$$\gamma_{1/2,1}(\mathcal{R}_{k,u}, d_U) \leq \sqrt{r'n} \gamma_{1/2,1}(\mathcal{R}_{k,u}, \sqrt{\delta}) \leq K 2^k \sqrt{r'n} (\log n)^{3/4}.$$

It then follows from the majorizing measure bound (Proposition 2.3) that

$$(5.4) \quad E_\epsilon X_{k,u} \leq K \sqrt{r'n} (\log n)^{3/4}.$$

Step 3. We want to show that the large values of $X_{k,u}$ are exceptional. A key observation is that for $C \in \mathcal{R}_{k,u}$, $\hat{C}$ is of area $\leq \pi 2^{-2k}$ (by the isoperimetric inequality); thus, $\hat{C}$ contains at most $K r' n 2^{-2k}$ points U_i . The use of Azuma's inequality as in Proposition 2.6 does not take advantage of this fact and thus would not yield the correct result. Instead we will use the following principle, proved in [T2], that we reformulate in the present setting.

Proposition 5.4. *Consider a class of functions $\mathcal{G}$ on $[0, 1]^2$ and $U_1, \dots, U_n \in [0, 1]^2$. Then, for $t > 0$*

$$P_\epsilon \left(\sup_{f \in \mathcal{G}} \left| \sum_{i \leq n} \epsilon_i f(U_i) \right| \geq 2E_\epsilon \sup_{f \in \mathcal{G}} \left| \sum_{i \leq n} \epsilon_i f(U_i) \right| + t\sigma \right) \leq 2 \exp \left(-\frac{t^2}{8} \right)$$

where $\sigma^2 = \sup_{f \in \mathcal{G}} \sum_{i \leq n} f^2(U_i)$.

In the present case, we have observed that $\sigma \leq K \sqrt{r'n} 2^{-k}$ if $\mathcal{G} \in \mathcal{R}_{k,u}$. Thus we get, using (5.4),

$$P_\epsilon(X_{k,u} \geq K(1+L)\sqrt{r'n}(\log n)^{3/4}) \leq 2 \exp(-L^2(\log n)^{3/4}).$$

If we look at the proof of Lemma 4.4, we see that $P(r' \geq 5) \leq \exp -L^2(\log n)^{3/4}$. So, we finally get

$$P(X_{k,u} \geq KL\sqrt{n}(\log n)^{3/4}) \leq 3 \exp(-L^2(\log n)^{3/4}).$$

What we desire is a similar inequality for $Y_{k,u}$ instead of $X_{k,u}$. To obtain it, we will appeal again to a general principle.

Lemma 5.5 [GZ, Lemma 2.7(b)]. *Consider a class $\mathcal{C}$ of functions on $[0, 1]^2$. Then, for $t \geq 2 \sup_{f \in \mathcal{C}} \|f\|_2$, we have*

$$P\left(\sup_{f \in \mathcal{C}} \left| \sum_{i \leq n} f(U_i) - \iint f d\lambda \right| \geq t\right) \leq 4P\left(\sup_{f \in \mathcal{C}} \left| \sum_{i \leq n} \epsilon_i f(U_i) \right| \geq t/4\right).$$

This completes the proof. $\square$

In our formulation of the grid matching theorem the two coordinates played the same role. In the spirit of the “ultimate matching conjecture” it is, however, of interest to give a more general statement. We leave the reader to check that only minor modifications in our proof are needed to obtain that result.

Theorem 5.6. *Consider $n \geq 1$ and numbers a, b such that $a, b \leq 1$ and $ab \geq K(\log n)^{3/2}/n$. Then with large probability ($\geq 1 - \exp -(\log n)^{3/2}$) we can find a matching π such that*

$$\begin{aligned} \max_{i \leq n} \|U_{1,i} - V_{1,\pi(i)}\| &\leq a, \\ \max_{i \leq n} \|U_{2,i} - V_{2,\pi(i)}\| &\leq b. \end{aligned}$$

Observe that a, b can depend on n . A case of interest is when a is of order $n^{-1/2}(\log n)^{1/2}$ and b of order $n^{-1/2} \log n$. This would also follow from the conjecture that Theorem 1.5 holds for $\alpha = 2$.

6. MORE MAJORIZING MEASURES

It is an unfortunate fact that many classes of functions of interest do not resemble ellipsoids; for those classes the existence of majorizing measures cannot be deduced from Theorem 2.4 alone. A different principle of difficulty comparable to that of Theorem 2.4 will be needed.

Rather than giving an abstract formulation, we will first present a simple occurrence of the basic idea. The same idea will then occur again in the proof of Proposition 6.3, which is the main step towards Theorems 1.3 and 1.6.

Proposition 6.1. *Consider for $n \geq 1$, $1 \leq k \leq n$ a subset $T_{n,k}$ of a Hilbert space H . Consider $\alpha, \beta, \delta' > 0$. Consider a set $\mathcal{S}$ of sequences $\sigma = (\sigma(n))_{n \geq 1}$, such that $1 \leq \sigma(n) \leq n$, and assume that*

$$(6.1) \quad \sup_{\sigma \in \mathcal{S}} \left(\sum_{n \geq 1} \gamma_{\alpha, \beta}(T_{n, \sigma(n)})^\delta \right)^\delta < \infty,$$

where $\delta = \max(\frac{1}{1+\alpha}, \frac{\beta}{\beta+1})$.

Assume, moreover, that $\delta' > \alpha$, that $0 \in T_{n,k}$ for all n, k , and that for all $\sigma \in \mathcal{S}$ we have

$$(6.2) \quad \sum_{n \geq 1} n^{\delta'} \text{diam}(T_{n, \sigma(n)}) \leq 1.$$

Then

$$\gamma_{\alpha, \beta} \left(\bigcup_{\sigma \in \mathcal{S}} \sum_{n \geq 1} T_{n, \sigma(n)} \right) < \infty.$$

Proof. Step 1. Consider the set $\mathcal{X} = \prod_{n \geq 1} \prod_{k \leq n} T_{n,k}$. We will denote by x the generic element of $\mathcal{X}$. Thus x is a family $(x_{n,k})$, $n \geq 1$, $1 \leq k \leq n$. Consider the map φ , defined on $\mathcal{S} \times \mathcal{X}$, valued in H , that is given by

$$\varphi(\sigma, x) = \sum_{n \geq 1} x_{n, \sigma(n)}.$$

For each $n \geq 1$, $1 \leq k \leq n$, we consider a probability measure $\mu_{n,k}$ on $T_{n,k}$ such that

$$\gamma_{\alpha, \beta}(T_{n,k}, \mu_{n,k}) \leq 2\gamma_{\alpha, \beta}(T_{n,k}),$$

and we denote by μ the product measure on $\mathcal{X}$.

It is a rather trivial fact that on $\mathcal{S}$ there exists a probability measure ν such that, for each $\sigma \in \mathcal{S}$, for each $m \geq 1$

$$(6.3) \quad \nu(\{\rho \in \mathcal{S}; \forall n \leq m, \rho(n) = \sigma(n)\}) \geq \frac{2^{-m-1}}{m^m}.$$

We denote by T the image of φ . Consider the measure η on T , which is the image of $\nu \otimes \mu$ under φ . We aim to prove that

$$(6.4) \quad \gamma_{\alpha, \beta}(T, \eta) = \sup_{y \in T} \left(\int_0^\infty \epsilon^\beta L(\eta(B(y, \epsilon)))^{\alpha\beta} \frac{d\epsilon}{\epsilon} \right)^{1/\beta} < \infty.$$

This will finish the proof, since $T = \bigcup_{\sigma \in \mathcal{S}} \sum_{n \geq 1} T_{n, \sigma(n)}$.

Step 2. Consider $y \in T$. Thus y is the image of a couple (σ, x) of $\mathcal{S} \times \mathcal{X}$.

We consider the distance d_σ on $\mathcal{X}$, given for two elements $x = (x_{n,k})$, $x' = (x'_{n,k})$ of $\mathcal{X}$ by

$$(6.5) \quad d_\sigma(x, x') = \sum_{n \geq 1} \|x_{n, \sigma(n)} - x'_{n, \sigma(n)}\|_2.$$

In that step we show that, for $\rho \in \mathcal{S}$, $x' \in \mathcal{X}$, we have

$$(6.6) \quad \forall n \leq m, \quad \rho(n) = \sigma(n) \Rightarrow \|\varphi(\sigma, x) - \varphi(\rho, x')\| \leq d_\sigma(x, x') + 2m^{-\delta'}.$$

To see this, we simply write

$$(6.7) \quad \begin{aligned} \|\varphi(\sigma, x) - \varphi(\rho, x')\| &\leq \sum_{n \leq m} \|x_{n, \sigma(n)} - x'_{n, \rho(n)}\| \\ &\quad + \sum_{n > m} (\|x_{n, \sigma(n)}\| + \|x'_{n, \rho(n)}\|). \end{aligned}$$

The first term is less than $d_\sigma(x, x')$ since $\rho(n) = \sigma(n)$ for $n \leq m$. To bound, the second term, we simply use that

$$\sum_{n \geq m} \|x_{n, \sigma(n)}\| \leq m^{-\delta'} \sum_{n \geq 1} n^{\delta'} \text{diam } T_{n, \sigma(n)} \leq m^{-\delta'}.$$

The central idea of (6.6) is that only the distance d_σ occurs, rather than the larger distance $\sum_{n,k} \|x_{n,k} - x'_{n,k}\|_2$.

Step 3. Consider the distance d on $\mathcal{S}$ given by

$$d(\sigma, \rho) = (\inf\{m; \forall n \leq m, \sigma(n) = \rho(n)\})^{-\delta'}.$$

Consider the distance δ_σ on $\mathcal{S} \times \mathcal{X}$ given by

$$\delta_\sigma((\sigma', x'), (\sigma'', x'')) = d(\sigma', \sigma'') + d_\sigma(x', x'').$$

Using (6.6) and Lemma 2.6, we have

$$(6.8) \quad \begin{aligned} \gamma_{\alpha, \beta}(T, \eta, y) &\leq K\gamma_{\alpha, \beta}(\mathcal{S} \times \mathcal{X}, \delta_\sigma, \nu \otimes \mu, (\sigma, x)) \\ &\leq K\gamma_{\alpha, \beta}(\mathcal{S}, d, \nu, \sigma) + K\gamma_{\alpha, \beta}(\mathcal{X}, d_\sigma, \mu, x), \end{aligned}$$

where the second inequality follows from Proposition 2.9. Consider now $\mathcal{X}_\sigma = \prod_{n \geq 1} T_{n, \sigma(n)}$, and on $\mathcal{X}_\sigma$ consider the distance d'_σ given by the right-hand side of (6.5). Consider the measure μ_σ on $\mathcal{X}_\sigma$ that is the product of the measures $\mu_{n, \sigma(n)}$, $n \geq 1$. Consider the point x_σ of $\mathcal{X}_\sigma$ given by $x_\sigma = (x_{n, \sigma(n)})_{n \geq 1}$. It should be obvious that for all $\epsilon > 0$, we have

$$\mu(B_{d_\sigma}(x, \epsilon)) = \mu_\sigma(B_{d'_\sigma}(x_\sigma, \epsilon)),$$

where the ball on the left is in $\mathcal{X}$ and the ball on the right in $\mathcal{X}_\sigma$. Thus

$$\gamma_{\alpha, \beta}(\mathcal{X}, d_\sigma, \mu, x) = \gamma_{\alpha, \beta}(\mathcal{X}_\sigma, d'_\sigma, \mu_\sigma, x_\sigma) \leq \gamma_{\alpha, \beta}(\mathcal{X}_\sigma, d'_\sigma, \mu_\sigma).$$

By Proposition 2.9 and (6.1) we have

$$\sup_{\sigma \in \mathcal{S}} \gamma_{\alpha, \beta}(\mathcal{X}_\sigma, d'_\sigma, \mu_\sigma) < \infty.$$

To conclude, it suffices to show that

$$\gamma_{\alpha, \beta}(\mathcal{S}, d, \nu) < \infty,$$

which is elementary using (6.2) and the fact that $\delta' > \alpha$.

This concludes the proof. $\square$

In Proposition 6.3 the set that will play the same role as the set $\mathcal{S}$ of Proposition 6.1 will have a much less trivial structure. We turn to its study.

Proposition 6.2. *For $k \geq 1$, we consider a partition $\mathcal{B}_k$ of $[0, 1]^2$ in 2^{2k} pieces of measure 2^{-2k} ; we assume that the sequence $(\mathcal{B}_k)$ of partitions is increasing. Consider $\alpha, \beta > 0$ such that $(\frac{1}{2} + \alpha)\beta > 1$. Consider the set $\mathcal{T}$ of sequences $(B_k)_{k \geq 1}$ such that B_k is $\mathcal{B}_k$ measurable and that $\sum_{k \geq 1} |B_k| \leq b$. We provide $\mathcal{T}$ with the distance*

$$(6.9) \quad d((B_k), (B'_k)) = \left(\sum_{k \geq 1} 2^{-4k\alpha} |B_k \Delta B'_k| \right)^{1/2}.$$

Then

$$\gamma_{\alpha, \beta}(\mathcal{T}, d) \leq Kb^{1/2+\alpha} (\log \frac{K}{b})^\alpha.$$

Here, and in the proof below, K denotes a constant depending on α, β only.

Proof. We will use Theorem 3.2. We denote by k_1 the smallest integer k such that $2^{-2k} \leq b$. For $(B_k) \in \mathcal{T}$, we must have $B_k = \emptyset$ for $k < k_1$. Consider the

largest integer k_0 such that $2^{-4k_1\alpha+1}b \leq 2^{-2k_0}$. Then the diameter of $(\mathcal{T}, d)$ is $\leq 2^{-k_0}$. Note also that $2^{-k_0} \leq Kb^{\alpha+1/2}$.

We set $\alpha' = \min(2\alpha, 1)$ and $A = \sum_{i \geq 0} 2^{-i\alpha'}$. For $k \geq k_0$, $q \geq 0$, we set

$$U_{k,q} = \{(B_\ell) \in \mathcal{T} : \forall \ell \geq 1, |B_\ell| \leq A^{-1} 2^{-2k+\alpha'|\ell-q|+4\alpha q}\}.$$

We now estimate $N(U_{k,q}, d, 2^{-k})$. Consider $(B_\ell) \in U_{k,q}$. Then, since $\alpha' \leq 2\alpha$, we have

$$\begin{aligned} \sum_{\ell \geq q} 2^{-4\ell\alpha} |B_\ell| &\leq A^{-1} \sum_{\ell \geq q} 2^{-4\ell\alpha-2k+\alpha'(\ell-q)+4\alpha q} \\ &\leq A^{-1} 2^{-2k} \sum_{\ell \geq q} 2^{-2\alpha(\ell-q)} \leq A^{-1} 2^{-2k} \sum_{i \geq 0} 2^{-\alpha' i} \leq 2^{-2k}. \end{aligned}$$

It follows that $N(U_{k,q}, d, 2^{-k})$ is bounded by the number of possible sequences $(B_\ell)_{1 \leq \ell < q}$.

For $\ell < q$, we set

$$\theta_\ell = \min(1, A^{-1} 2^{-2k+\alpha'(q-\ell)+4\alpha q})$$

and $n_\ell = \lfloor 2^{2\ell} \theta_\ell \rfloor$. Thus

$$N(U_{k,q}, d, 2^{-k}) \leq \prod_{k_1 \leq \ell < q} \sum_{0 \leq p \leq n_\ell} \binom{2^{2\ell}}{p} \leq \prod_{k_1 \leq \ell < q} \left(\frac{e 2^{2\ell}}{n_\ell} \right)^{n_\ell} \leq \prod_{k_1 \leq \ell < q} \left(\frac{e}{\theta_\ell} \right)^{2^{2\ell} \theta_\ell}$$

where we have used that the function $(e/t)^t$ increases for $t \leq 1$. Thus

$$\log N(U_{k,q}, d, 2^{-k}) \leq \sum_{k_1 \leq \ell < q} 2^{2\ell} \theta_\ell \log \left(\frac{e}{\theta_\ell} \right).$$

Now, we observe that the sequence (θ_ℓ) decreases and that, since $\alpha' \leq 1$, the sequence $2^\ell \theta_\ell$ increases, so

$$\begin{aligned} \log N(U_{k,q}, d, 2^{-k}) &\leq \sum_{k_1 \leq \ell < q} 2^{2\ell} \theta_\ell \log \left(\frac{e}{\theta_q} \right) \\ &\leq \log \left(\frac{e}{\theta_q} \right) \sum_{\ell < q} 2^\ell (2^q \theta_q) \leq 2^{2q} \theta_q \log \left(\frac{e}{\theta_q} \right) \end{aligned}$$

where $\theta_q = \min(1, A^{-1} 2^{-2k+4\alpha q})$.

Consider now any sequence $(B_\ell)_{\ell \geq 1}$ in $\mathcal{T}$. We set $u_p = \sum_{\ell \geq 1} 2^{-\alpha'|\ell-p|} |B_\ell|$. Thus $\sum_p u_p \leq 2Ab$, since $\sum_{\ell \geq 1} |B_\ell| \leq b$. Since $u_p \leq 2Ab \leq A2^{4k_1\alpha-2k_0}$, we have

$$(6.10) \quad 2^{-4q\alpha} u_q \leq A^{-1} 2^{-2k}$$

as soon as $2^{-4q\alpha} \leq A^{-2} 2^{-4k_1\alpha - 2(k-k_0)}$. Thus, if we denote by $q(k)$ the smallest integer $\geq k_1$ such that (6.10) occurs, we have $q(k) \leq k_1 + \frac{1}{2\alpha}(k - k_0) + L$, where L depends only on α, β .

Consider now k such that $q(k) > k_1$. Then, by the definition of $q(k)$, we have

$$2^{-2k} \leq A 2^{-4(q(k)-1)\alpha} u_{q(k)-1} \leq A 2^{4\alpha+\alpha'} 2^{-4q(k)\alpha} u_{q(k)}$$

since $u_{p-1} \leq 2^{\alpha'} u_p$ by the definition of u_p . Thus we have

$$\theta_{q(k)} \leq \min(1, K u_{q(k)})$$

and

$$\begin{aligned} 2^{-k\beta} (\log N(U_{k,q(k)}, d, 2^{-k}))^{\alpha\beta} &\leq 2^{-\beta k} 2^{2\alpha\beta q(k)} \theta_{q(k)}^{\alpha\beta} \left(\log \frac{e}{\theta_{q(k)}} \right)^{\alpha\beta} \\ &\leq K u_{q(k)}^{\beta/2} \theta_{q(k)}^{\alpha\beta} \left(\log \frac{e}{\theta_{q(k)}} \right)^{\alpha\beta}. \end{aligned}$$

Since $\theta_{q(k)} \leq K u_{q(k)} \leq K b \leq K$ and since the function $t \rightarrow t \log(Ke/t)$ increases for $t \leq K$, we get

$$2^{-k\beta} (\log N(U_{k,q(k)}, d, 2^{-k}))^{\alpha\beta} \leq K u_{q(k)}^{\beta(\alpha+1/2)} \left(\log \frac{eK}{u_{q(k)}} \right)^{\alpha\beta}.$$

We observe (as in the proof of Theorem 3.2) that each integer $q > k_1$ is of the type $q(k)$ for at most a bounded number of values of q . We also observe that

$$\begin{aligned} \sum_{q(k)=k_1} 2^{-k\beta} (\log N(U_{k,k_1}, d, 2^{-k}))^{\alpha\beta} &\leq \sum_{k \geq k_0} 2^{-k\beta} \left(2^{2k_1} \theta_{k_1} \log \frac{e}{\theta_{k_1}} \right)^{\alpha\beta} \\ &\leq K 2^{-k_0\beta+2k_1\alpha\beta} \left(b \log \frac{e}{b} \right)^{\alpha\beta} \leq K b^{\beta(\alpha+1/2)} \left(\log \frac{e}{b} \right)^{\alpha\beta} \end{aligned}$$

since 2^{-k_1} is of order $b^{1/2}$ and 2^{-k_0} is of order $b^{\alpha+1/2}$.

Finally, we have shown that, setting $S = b^{\beta(\alpha+1/2)} (\log \frac{eK}{b})^{\alpha\beta}$ we have

$$\sum_{k \geq k_0} 2^{-k\beta} (\log N(U_{k,q(k)}, d, 2^{-k}))^{\alpha\beta} \leq K \left(S + \sum_{q \geq k_1} u_q^{\beta(\alpha+1/2)} \left(\log \frac{eK}{u_q} \right)^{\alpha\beta} \right).$$

We recall now that $\sum_{q \geq k_1} u_q \leq K b$. Since we assume $\delta' = \beta(\alpha + 1/2) - 1 > 0$ and since, as easily seen, the function $t^{\delta'} (\log \frac{eK}{t})^{\alpha\beta}$ has for $t \leq b$ a maximum $\leq K b^{\delta'} (\log \frac{eK}{b})^{\alpha\beta}$, we conclude that

$$\sum_{k \geq k_0} 2^{-k\beta} (\log N(U_{k,q(k)}, d, 2^{-k}))^{\alpha\beta} \leq K S.$$

It follows by Theorem 3.2 that

$$\gamma_{\alpha,\beta}(\mathcal{T}, d) \leq K(S^{1/\beta} + 2^{-k_0}) \leq K S^{1/\beta}. \quad \square$$

Comment. The condition $(\frac{1}{2} + \alpha)\beta > 1$ is essential; the result actually fails for $\alpha = \frac{1}{2}$, $\beta = 1$.

The following result is at the heart of Theorems 1.3 and 1.6.

Proposition 6.3. Consider a class $\mathcal{F}$ of continuous functions on $[0, 1]^2$, an integer $k_0 \geq 0$, and a number $b \leq 2^{-2k_0-1}$. Assume that $|\{f \neq 0\}| \leq b$ whenever $f \in \mathcal{F}$. For a rectangle $A \subset [0, 1]^2$, set

$$\mathcal{F}_A = \{f1_A, f \in \mathcal{F}, |\{f = 0\} \cap A| \geq |A|/2\}.$$

Consider α, β such that $(\frac{1}{2} + \alpha)\beta > 1$. Suppose that for each $a > 0$ and each rectangle $A \subset [0, 1]^2$ of sides $[2^{-k_0}a, 2^{k_0}a]$ we have

$$(6.11) \quad \gamma_{\alpha, \beta}(\mathcal{F}_A) \leq Ma^{1+2\alpha}.$$

Then the diameter of $\mathcal{F}$ is $\leq K Mb^{1/2+\alpha}$ and we have

$$\gamma_{\alpha, \beta}(\mathcal{F}) \leq K Mb^{1/2+\alpha} (\log \frac{K}{b})^\alpha$$

where K depends only on α, β .

Proof. By homogeneity, we can assume $M = 1$.

Step 1. For $k \geq k_0$, we consider the paving $\mathcal{P}_k$ obtained by partitioning $[0, 1]^2$ into 2^{2k} rectangles of sides $2^{-k_0-k}, 2^{k_0-k}$ and the algebra $\mathcal{A}_k$ it generates. We consider the set $\mathcal{T}_0$ of sequences $(B_k)_{k \geq k_0}$ such that B_k is $\mathcal{A}_k$ -measurable, the sets B_k are disjoint, and $\sum_{k \geq k_0} |B_k| \leq 8b$.

We first prove that given $f \in \mathcal{F}$, one can find a sequence $B = (B_k)_{k \geq k_0}$ in $\mathcal{T}_0$ with the following properties:

$$(6.12) \quad f = 0 \text{ outside } \bigcup_{k \geq k_0} B_k.$$

$$(6.13) \quad \text{If } A \text{ is a rectangle of } \mathcal{P}_k \text{ contained in } B_k, \text{ then } |\{f = 0\} \cap A| \geq |A|/2 \\ \text{(so that } f1_A \in \mathcal{F}_A \text{)}.$$

To construct B , we construct the sets B_k inductively. We define first B_{k_0} as the union of all the rectangles $A \in \mathcal{P}_{k_0}$ for which $|A \cap \{f \neq 0\}| \geq |A|/8$. Observe that

$$|\{f \neq 0\}| \leq 2^{-2k_0-1} = |A|/2.$$

When $B_{k_0}, \dots, B_k$ have been constructed, we define B_{k+1} as the union of the rectangles A of $\mathcal{P}_{k+1}$ that satisfy $|A \cap \{f \neq 0\}| \geq |A|/8$ and that are not included in $\bigcup_{\ell \leq k} B_\ell$. We must have $|A \cap \{f \neq 0\}| \leq |A|/2$, for otherwise the rectangle A' of $\mathcal{P}_k$ that contains A would have been selected in the construction of B_k . We have that $|B_k| \leq 8|B_k \cap \{f \neq 0\}|$, so that $\sum_{k \geq k_0} |B_k| \leq 8|\{f \neq 0\}| \leq 8b$. Consider $u \notin \bigcup_{k \geq k_0} B_k$. For each k the rectangle A of $\mathcal{A}_k$ that contains u must satisfy $|A \cap \{f \neq 0\}| < |A|/8$, for otherwise, for the smallest k for which this inequality would be wrong, we would have $A \subset B_k$ by construction. Since $\{f \neq 0\}$ is open, this shows that $u \notin \{f \neq 0\}$ and finishes the proof.

We have succeeded to write $f = \sum 1_A f$, where $1_A f \in \mathcal{F}_A$, and the rectangles A are disjoint and of total area at most $8b$. Combining Lemma 2.5 and (6.11),

it follows easily that $\|f\|_2^2 \leq KM^2b^{1+2\alpha}$, and the claim about the diameter of $\mathcal{F}$ follows.

Step 2. Consider the set $\mathcal{X} = \prod_{k \geq k_0} \prod_{A \in \mathcal{P}_k} \mathcal{F}_A$.

We will denote by X the generic element of $\mathcal{X}$. Thus X is a family (f_A) , where $f_A \in \mathcal{F}_A$, and $A \subset \bigcup_{k \geq k_0} \mathcal{P}_k$. We will denote by B the generic element of $\mathcal{T}_0$ so that B is a sequence $(B_k)_{k \geq k_0}$. We consider the map φ , defined on $\mathcal{T}_0 \times \mathcal{X}$, valued on the set of functions on $[0, 1]^2$, which is given by

$$\varphi(B, X) = \sum_{k \geq k_0} \sum_{A \subset B_k} f_A.$$

(We make the convention that $f_A = 0$ outside A . Thus, if $u \notin \bigcup_{k \geq k_0} B_k$, we have $\varphi(B, X)(u) = 0$. If $u \in B_k$, consider $A \in \mathcal{P}_k$ for which $u \in A \subset B_k$. Then $\varphi(B, X) = f_A(u)$.)

For each $A \in \mathcal{P}_k$, $k \geq k_0$, we denote by μ_A a probability measure on $\mathcal{F}_A$ such that

$$\gamma_{\alpha, \beta}(\mathcal{F}_A, \mu_A) \leq 2\gamma_{\alpha, \beta}(\mathcal{F}_A) \leq 2 \cdot 2^{-k(1+2\alpha)}.$$

We denote by μ the product measure on $\mathcal{X}$.

On $\mathcal{T}_0$, consider the distance d given by (6.9). From Proposition 6.2, and Corollary 2.8 we can find a probability measure ν on $\mathcal{T}_0$ such that

$$(6.14) \quad \gamma_{\alpha, \beta}(\mathcal{T}_0, d, \nu) \leq Kb^{1/2+\alpha}(\log \frac{K}{b})^\alpha.$$

We denote by $\mathcal{F}'$ the image of φ . Consider the measure η on $\mathcal{F}'$, which is the image of $\nu \otimes \mu$ under φ . We aim to prove that

$$\forall f \in \mathcal{F}', \quad \left(\int_0^\infty \epsilon^\beta L(\eta(B(f, \epsilon)))^{\alpha\beta} \frac{d\epsilon}{\epsilon} \right)^{1/\beta} \leq Kb^{1/2+\alpha} \left(\log \frac{K}{b} \right)^\alpha.$$

This will finish the proof by Corollary 2.8.

Step 3. Consider $f \in \mathcal{F}'$. Thus, by definition, f is the image under φ of a couple (B, X) of $\mathcal{T}_0 \times \mathcal{X}$.

We consider the distance d_B on $\mathcal{X}$, given, for two elements $X = (f_A)$, $X' = (f'_A)$ of $\mathcal{X}$ by

$$(6.15) \quad d_B(X, X') = \left(\sum_{k \geq k_0} \sum_{A \subset B_k} \|f_A - f'_A\|_2^2 \right)^{1/2}.$$

In that step, we show that for $B' \in \mathcal{T}_0$, $X' \in \mathcal{X}$,

$$(6.16) \quad d_B(X, X') \leq \epsilon; \quad d(B, B') \leq \epsilon \Rightarrow \|\varphi(B', X') - f\|_2 \leq K\epsilon.$$

The central idea of the whole proof is that in (6.16) we need only have control of $d_B(X, X')$ and not of the larger distance $(\sum_{k \geq k_0} \sum_{A \in \mathcal{P}_k} \|f_A - f'_A\|_2^2)^{1/2}$.

We start with some observations. By hypothesis, for $A \in \mathcal{P}_k$, we have $\gamma_{\alpha, \beta}(\mathcal{F}_A) \leq 2^{-k(1+2\alpha)}$. Since $0 \in \mathcal{F}_A$, we have, by Lemma 2.5 that

$$(6.17) \quad \iint_A |f_A|^2 d\lambda \leq K2^{-2k(1+2\alpha)} = K|A|2^{-4k\alpha}.$$

Consider now any function $f' = \varphi(B', X')$ for $B' \in \mathcal{T}_0$, $X' \in \mathcal{X}$. Consider $k \geq k_0$, $A \in \mathcal{P}_k$, and assume that $A \cap \bigcup_{\ell < k} B'_\ell = \emptyset$. For $\ell \geq k$, set $A_\ell = B'_\ell \cap A$. We have, by (6.17), and since for $C \in \mathcal{P}_\ell$, $C \subset B'_\ell$ we have $f' = f'_C$ on C ,

$$(6.18) \quad \begin{aligned} \iint_A f'^2 d\lambda &= \sum_{\ell \geq k} \sum_{\substack{C \in \mathcal{P}_\ell \\ C \subset A_\ell}} \iint_C f'^2 d\lambda \\ &\leq K \sum_{\ell \geq k} \sum_{\substack{C \in \mathcal{P}_\ell \\ C \subset A_\ell}} |C| 2^{-4\ell\alpha} \leq K \sum_{\ell \geq k} |A_\ell| 2^{-4\ell\alpha} \leq K |A| 2^{-4\alpha k}. \end{aligned}$$

Consider now $B' \in \mathcal{T}_0$. For $k \geq k_0$, we set

$$H_k = B_k \cap B'_k, \quad L_k = B_k \setminus \bigcup_{\ell_0 \leq \ell \leq k} B'_\ell; \quad L'_k = B'_k \setminus \bigcup_{\ell_0 \leq \ell \leq k} B_\ell.$$

We observe that $L_k \subset B_k \setminus B'_k$; $L'_k \subset B'_k \setminus B_k$. Thus

$$(6.19) \quad |L_k| + |L'_k| \leq |B_k \Delta B'_k|.$$

The function $f - f'$ is zero outside the set $\bigcup_{k \geq k_0} (B_k \cup B'_k)$. This set is contained in $\bigcup_{k \geq k_0} (H_k \cup L_k \cup L'_k)$. Indeed, if k is the smallest index $\geq k_0$ for which $u \in B_k \cup B'_k$, it is clear that $u \in H_k \cup L_k \cup L'_k$. Thus we can write

$$\begin{aligned} \iint (f - f')^2 d\lambda &\leq \sum_{k \geq k_0} \iint_{H_k} (f - f')^2 d\lambda + \sum_{k \geq k_0} \iint_{L_k} (f - f')^2 d\lambda \\ &\quad + \sum_{k \geq k_0} \iint_{L'_k} (f - f')^2 d\lambda \\ &=: \text{I} + \text{II} + \text{III}. \end{aligned}$$

Since $H_k \subset B_k$, we have by (6.15) that $\text{I} \leq d_B(X, X')^2$. By the triangle inequality, we have

$$(6.20) \quad \iint_{L_k} (f - f')^2 d\lambda \leq 4 \sum_{\substack{A \subset L_k \\ A \in \mathcal{P}_k}} \left(\iint_A f^2 d\lambda + \iint_A f'^2 d\lambda \right).$$

By (6.17), for $A \in \mathcal{P}_k$, $A \subset L_k$, since $f = f_A$ on A , we have $\iint_A f^2 d\lambda \leq K |A| 2^{-4\alpha k}$, and by (6.18) we have $\iint_A f'^2 d\lambda \leq K |A| 2^{-4\alpha k}$. Thus, by (6.20), we have

$$\iint_{L_k} (f - f')^2 d\lambda \leq K |L_k| 2^{-4\alpha k}.$$

Similarly

$$\iint_{L'_k} (f - f')^2 d\lambda \leq K |L'_k| 2^{-4\alpha k}.$$

Thus by (6.19) we have

$$\text{II} + \text{III} \leq K \sum_{k \geq k_0} 2^{-4\alpha k} |B_k \Delta B'_k| \leq K d^2(B, B').$$

This finishes the proof.

Step 4. Consider the distance δ_B on $\mathcal{T}_0 \times \mathcal{X}$ given by

$$\delta_B((B', X'), (B'', X''))^2 = d(B', B'')^2 + d_B(X', X'')^2.$$

Using (6.16) and Lemma 2.6, we have

$$(6.21) \quad \begin{aligned} \gamma_{\alpha, \beta}(\mathcal{T}', \eta, f) &\leq K \gamma_{\alpha, \beta}(\mathcal{T}_0 \times \mathcal{X}, \delta_B, \nu \otimes \mu, (B, X)) \\ &\leq K \gamma_{\alpha, \beta}(\mathcal{T}_0, d, \nu, B) + K \gamma_{\alpha, \beta}(\mathcal{X}, d_B, \mu, X) \end{aligned}$$

where the second inequality follows from Proposition 2.9. Consider now $\mathcal{X}_B = \prod_{k \geq k_0} \prod_{A \subset B_k, A \in \mathcal{P}_k} \mathcal{F}_A$. On $\mathcal{X}_B$ consider the distance d'_B given by the right-hand side of (6.15). Consider the measure μ_B on $\mathcal{X}_B$ that is the product of the measures μ_A , $A \in \mathcal{P}_k$, $A \subset B_k$, $k \geq k_0$. Consider the point X_B of $\mathcal{X}_B$ given by $X_B = (f_A)_{A \in \mathcal{P}_k, A \subset B_k, k \geq k_0}$. It should be obvious that for all $\epsilon > 0$ we have

$$\mu(B_{d_B}(X, \epsilon)) = \mu_B(B_{d'_B}(X_B, \epsilon)),$$

where the ball on the left is in $\mathcal{X}$ and the ball on the right is in $\mathcal{X}_B$. Thus

$$\gamma_{\alpha, \beta}(\mathcal{X}, d_B, \mu, X) = \gamma_{\alpha, \beta}(\mathcal{X}_B, d'_B, \mu_B, X_B) \leq \gamma_{\alpha, \beta}(\mathcal{X}_B, d'_B, \mu_B).$$

By Proposition 2.9, we have

$$\gamma_{\alpha, \beta}(\mathcal{X}_B, d'_B, \mu_B) \leq \left(\sum_{k \geq k_0} \sum_{A \subset B_k} \gamma_{\alpha, \beta}(\mathcal{F}_A, \mu_A)^{2/(1+2\alpha)} \right)^{\alpha+1/2}.$$

By (6.11) we have

$$\sum_{A \subset B_k} \gamma_{\alpha, \beta}(\mathcal{F}_A, \mu_A)^{2/(1+2\alpha)} \leq |B_k|$$

and thus $\gamma_{\alpha, \beta}(\mathcal{X}_B, d'_B, \mu_B) \leq b^{\alpha+1/2}$.

Going back to (6.21) and using (6.14), the proof is finished. $\square$

We now apply Proposition 6.3 to an important step of the proof of Theorem 1.3. In §9, we will give another application, in the same spirit, to the proof of Theorem 1.6.

Proposition 6.4. Consider $b, c, d \geq 0$, with $b \leq 1/8$. Consider the class

$$(6.22) \quad \mathcal{G}(b, c, d) = \{f : [0, 1]^2 \rightarrow \mathbb{R}; |\{f = 0\}| \leq b, \|\frac{\partial f}{\partial x}\|_\infty \leq c, \|\frac{\partial f}{\partial y}\|_\infty \leq d\}.$$

Then, the diameter of $\mathcal{G}(b, c, d)$ is $\leq Kb \max(\sqrt{cd}, c\sqrt{b}, d\sqrt{b})$ and we have

$$(6.23) \quad \gamma_{1/2, 2}(\mathcal{G}(b, c, d)) \leq Kb (\log \frac{K}{b})^{1/2} \max(\sqrt{cd}, c\sqrt{b}, d\sqrt{b}).$$

Proof. We can and do assume $c \geq d$. We can also assume that $d \geq 8bc$, by replacing d by $\max(d, 8bc)$. Consider the integer $k_0 \geq 0$ such that $\frac{c}{d} \leq 2^{2k_0} \leq 4\frac{c}{d}$. Thus $b \leq d/8c \leq 2^{-2k_0-1}$.

To deduce (6.23) from Proposition 6.3, it suffices to show that if we have a rectangle $A \subset [0, 1]^2$, of sides $[2^{-k_0}a, 2^{k_0}a]$, then we have $\gamma_{1/2,2}(\mathcal{G}_A) \leq Ka^2\sqrt{cd}$ where

$$\mathcal{G}_A = \{f : A \rightarrow \mathbb{R}, \exists u \in A, f(u) = 0, \|\frac{\partial f}{\partial x}\|_\infty \leq c, \|\frac{\partial f}{\partial y}\|_\infty \leq d\}.$$

For simplicity, let us assume that $A = [0, 2^{-k_0}a] \times [0, 2^{k_0}a]$. To each function g defined on A , we associate the function $\tilde{g}$ defined on $[0, 1]^2$ by $\tilde{g}(x, y) = ag(2^{-k_0}ax, 2^{k_0}ay)$. This transformation preserves the L^2 distance and sends $\mathcal{G}_A$ to the class

$$\begin{aligned} \mathcal{G} = \{f \in [0, 1]^2; \exists u \in [0, 1]^2, f(u) = 0, \\ \|\frac{\partial f}{\partial x}\|_\infty \leq a^2c2^{-k_0}, \|\frac{\partial f}{\partial y}\|_\infty \leq a^2d2^{k_0}\}. \end{aligned}$$

We observe that, since $2^{-k_0} \leq \sqrt{d/c}$, $2^{k_0} \leq 2\sqrt{c/d}$, we have

$$\max(c2^{-k_0}, d2^{k_0}) \leq 2\sqrt{cd}.$$

For $f \in \mathcal{G}$, we certainly have $\|f\|_\infty \leq Ka^2\sqrt{cd}$. Thus by Lemma 4.3, we have $\gamma_{1/2,2}(\mathcal{G}) \leq Ka^2\sqrt{cd}$. The proof is complete. $\square$

For $a > 0$, we set $\Phi_a(x) = |x|(\log(1 + |x|))^a$.

Theorem 6.5. Consider $a, b \geq 0$ such that $a + b > 4$. Consider the class

$$(6.24) \quad \mathcal{E} = \mathcal{E}(a, b) = \left\{ f : [0, 1]^2 \rightarrow \mathbb{R}; |f| \leq 1, \right. \\ \left. \iint \Phi_a\left(\frac{\partial f}{\partial x}\right) d\lambda \leq 1; \iint \Phi_b\left(\frac{\partial f}{\partial y}\right) d\lambda \leq 1 \right\}.$$

Then $\gamma_{1/2,2}(\mathcal{E}) < \infty$.

Comment. The condition $a + b > 4$ certainly looks strange. We conjecture that the condition $a + b \geq 1/2$ suffices to imply that $\gamma_{1/2,2}(\mathcal{E}) < \infty$. (It is a simple fact to show that $\gamma_{1/2,2}(\mathcal{E}(a, b)) = \infty$ if $a + b < 1/2$.) However we believe that a radically new idea would be needed to prove this fact. The research of the best possible condition on a, b is motivated by Theorem 1.5.

Proof. The proof will use Proposition 6.1. The sets $T_{n,k}$ will be (with the notation of (6.22)) of the type $\mathcal{G}(2^{-k}, 2h_1(n+1), 2h_2(n+1))$. The numbers $h_1(n), h_2(n)$ will be chosen later.

Step 1. We describe the basic approximation procedure. Consider $f \in \mathcal{E}$. For $x \in [0, 1]$, $t \geq 0$, we denote the interval $[0, 1] \cap [x - t, x + t]$ by $I(x, t)$.

Consider the sets

$$A_\ell = \left\{ (x, y) \in [0, 1]^2; \forall t \geq 0, \int_{I(x, t)} \left| \frac{\partial f}{\partial x}(z, y) \right| dz \leq h_1(\ell)t \right\},$$

$$B_\ell = \left\{ (x, y) \in [0, 1]^2; \forall t \geq 0, \int_{I(y, t)} \left| \frac{\partial f}{\partial y}(x, z) \right| dz \leq h_2(\ell)t \right\},$$

and set $C_\ell = A_\ell \cup B_\ell$.

We consider the function f_ℓ given by

$$f_\ell(x, y) = \inf\{f(r, s) + h_1(\ell)|x - r| + h_2(\ell)|y - s|; (r, s) \notin C_\ell\}.$$

First, we prove that $f_\ell = f$ outside C_ℓ . Consider $(x, y) \notin C_\ell$. The definition of f_ℓ shows that $f_\ell(x, y) \leq f(x, y)$. Consider now $(r, s) \notin C_\ell$. Since $(r, s) \notin A_\ell$, for $t = |x - r|$, we have

$$\int_{I(r, t)} \left| \frac{\partial f}{\partial x}(z, s) \right| dz \leq h_1(\ell)t.$$

Thus

$$f(x, s) \leq f(r, s) + \int_{I(r, t)} \left| \frac{\partial f}{\partial x}(z, s) \right| dz \leq f(r, s) + h_1(\ell)|x - r|.$$

A similar argument, now using the fact that $(x, y) \notin B_\ell$, shows that

$$f(x, y) \leq f(x, s) + h_2(\ell)|y - s|.$$

Thus

$$f(x, y) \leq f(r, s) + h_1(\ell)|x - r| + h_2(\ell)|y - s|.$$

Since this holds for all $(r, s) \notin C_\ell$, we have $f(x, y) \leq f_\ell(x, y)$, so $f_\ell = f$ outside C_ℓ .

The function

$$g(x, y) = f(r, s) + h_1(\ell)|x - r| + h_2(\ell)|y - s|$$

satisfies

$$|g(x, y) - g(x', y)| \leq h_1(\ell)|x - x'|,$$

$$|g(x, y) - g(x, y')| \leq h_2(\ell)|y - y'|.$$

It is then routine to see that f_ℓ satisfies

$$|f_\ell(x, y) - f_\ell(x', y)| \leq h_1(\ell)|x - x'|,$$

$$|f_\ell(x, y) - f_\ell(x, y')| \leq h_2(\ell)|y - y'|,$$

so $\|\partial f_\ell / \partial x\|_\infty \leq h_1(\ell)$, $\|\partial f_\ell / \partial y\|_\infty \leq h_2(\ell)$.

Step 2. We write

$$(6.25) \quad f = f_1 + \sum_{\ell \geq 1} (f_{\ell+1} - f_\ell).$$

We now assume that $h_1(\ell+1) \geq h_1(\ell)$, $h_2(\ell+1) \geq h_2(\ell)$. Then $C_\ell \subset C_{\ell+1}$. Thus $f_{\ell+1} - f_\ell = 0$ outside C_ℓ . With the notation of (6.14) we have by the result of Step 1 that

$$f_{\ell+1} - f_\ell \in \mathcal{G}(|C_\ell|, 2h_1(\ell+1), 2h_2(\ell+1)).$$

The set C_ℓ depends on f only. Let us denote by $k(\ell, f)$ the largest integer $\leq 2\ell$ such that $|C_\ell| \leq 2^{-k(\ell, f)}$. Thus, if we set

$$T_{\ell, k} = \mathcal{G}(2^{-k}, 2h_1(\ell + 1), 2h_2(\ell + 1)),$$

we have

$$f_{\ell+1} - f_\ell \in T_{\ell, k(\ell, f)}.$$

Let us set $\mathcal{G}_1 = \{f; f \in \mathcal{G}\}$. From (6.25), we have

$$(6.26) \quad C \subset \mathcal{G}_1 + \bigcup_{f \in \mathcal{G}} \sum_{\ell \geq 1} T_{\ell, k(\ell, f)}.$$

We consider the set $\mathcal{S}$ of all sequences of the type $(k(\ell, f))_{\ell \geq 1}$ for $f \in \mathcal{G}$. It remains mostly to prove (6.1), (6.2).

Consider $\delta = a + b - 3 > 1$. Thus $\sum_{\ell \geq 1} \ell^{-\delta} < \infty$. Writing

$$\gamma^{2/3} = (\gamma \ell^{\delta/2})^{2/3} \ell^{-\delta/3}$$

and using Hölder's inequality with exponents $\frac{3}{2}, 3$, we get that for all $f \in \mathcal{G}$ we have

$$(6.27) \quad \left(\sum_{\ell \geq 1} \gamma_{1/2, 2}(T_{\ell, k(\ell, f)})^{2/3} \right)^{3/2} \leq K \left(\sum_{\ell \geq 1} \ell^{\delta/2} \gamma_{1/2, 2}(T_{\ell, k(\ell, f)}) \right).$$

Now, if we assume that

$$(6.28) \quad h_1(\ell + 1) \geq h_2(\ell + 1), \quad 2^{-k(\ell, f)} \leq \frac{h_2(\ell + 1)}{8h_1(\ell + 1)},$$

we can apply Proposition 6.4 to get

$$\begin{aligned} \gamma_{1/2, 2}(T_{\ell, k(\ell, f)}) &\leq K 2^{-k(\ell, f)} (\log K 2^{k(\ell, f)})^{1/2} (h_1(\ell + 1) h_2(\ell + 1))^{1/2} \\ &\leq K 2^{-k(\ell, f)} \sqrt{\ell} (h_1(\ell + 1) h_2(\ell + 1))^{1/2}. \end{aligned}$$

Thus, since $\frac{1}{2} + \frac{\delta}{2} = \frac{a+b}{2} - 1$, the left-hand side of (6.27) is bounded by

$$(6.29) \quad \sum_{\ell \geq 1} 2^{-k(\ell, f)} \ell^{(a+b)/2-1} (h_1(\ell + 1) h_2(\ell + 1))^{1/2}.$$

To prove (6.1), it suffices to prove (6.28) and that the quantity (6.29) is bounded independently of f . Since (under (6.28)) by Proposition 6.3 the diameter of $T_{\ell, k(\ell, f)}$ is at most $K 2^{-k(\ell, f)} (h_1(\ell + 1) h_2(\ell + 1))^{1/2}$, this will also prove (6.2).

Step 3. Consider a parameter $L \geq 1$ to be determined later. We now take

$$h_1(\ell) = L 2^\ell \ell^{(b-a)/2}, \quad h_2(\ell) = L 2^\ell \ell^{-(b-a)/2}.$$

Thus $h_1(\ell) h_2(\ell) = L^2 2^{2\ell}$.

Lemma 6.5. Consider a function $g \geq 0$ on $[0, 1]$ and $\xi > 0$. Let

$$A = \left\{ x \in [0, 1]; \exists t > 0, \int_{I(x, t)} g(s) ds \geq \xi t \right\}.$$

Then

$$\xi |A| \leq K \int_0^1 g(s) 1_{\{g(s) \geq \xi/4\}} ds.$$

Proof. The Vitali covering theorem implies that we can find disjoint intervals $I_p = I(x_p, t_p)$ such that $|\bigcup I_p| \geq \frac{1}{K} |A|$ and

$$\int_{I_p} g(s) ds \geq \xi t_p$$

for all $p \geq 1$. Since

$$\int_{I_p} g(s) 1_{\{g(s) \leq \xi/4\}} ds \leq \xi t_p / 2,$$

we must have

$$\int_{I_p} g(s) 1_{\{g(s) \geq \xi/4\}} ds \geq \xi t_p / 2.$$

Since the intervals I_p are disjoint, we get

$$\frac{\xi}{2} \left| \bigcup I_p \right| \leq \sum_p \frac{\xi t_p}{2} \leq \int g(s) 1_{\{g(s) \geq \xi/4\}} ds.$$

This implies the result. $\square$

We use the lemma on the function $g(s) = \left| \frac{\partial f}{\partial x}(s, y) \right|$ at y given for $\xi = h_1(\ell)$, and we integrate the corresponding inequality in y . Using the Fubini theorem, we get

$$h_1(\ell) |A_\ell| \leq K \iint \left| \frac{\partial f}{\partial x} \right| 1_{\{|\partial f / \partial x| \geq h_1(\ell)/4\}} d\lambda.$$

Thus

$$(6.30) \quad \sum_{\ell \geq 1} h_1(\ell) \ell^{a-1} |A_\ell| \leq K \iint \left| \frac{\partial f}{\partial x} \right| \sum_{\ell \geq 1} \ell^{a-1} 1_{\{|\partial f / \partial x| \geq h_1(\ell)/4\}} d\lambda.$$

We now observe that (since $L \geq 1$, $b \geq a$)

$$\sum \{ \ell^{a-1}; h_1(\ell)/4 \leq B \} \leq K (\log(1+B))^a.$$

It then follows from (6.30), since $h_1(\ell) = L 2^\ell \ell^{(b-a)/2}$ that

$$\sum_{\ell \geq 1} 2^\ell \ell^{(a+b)/2-1} |A_\ell| \leq \frac{K}{L} \iint \left| \frac{\partial f}{\partial x} \right| \left(\log \left(1 + \left| \frac{\partial f}{\partial x} \right| \right) \right)^a d\lambda \leq \frac{K}{L}.$$

A similar argument gives (we observe that there is no loss of generality to assume $b \leq 5$, so that $a - b \geq -5$)

$$\sum_{\ell \geq 1} 2^\ell \ell^{(a+b)/2-1} |B_\ell| \leq \frac{K}{L}$$

so that

$$(6.31) \quad \sum_{\ell \geq 1} 2^\ell \ell^{(a+b)/2-1} |C_\ell| \leq \frac{K}{L}.$$

Now, we have

$$\frac{h_2(\ell+1)}{h_1(\ell+1)} = (\ell+1)^{-(b-a)}.$$

Thus, we can take L large enough to ensure that (6.28) holds for each $\ell \geq 1$. (L can be taken independently of a, b , since, as already observed, there is no loss of generality to assume $b-a \leq 5$.) This in turn ensures that the left-hand side of (6.26) is bounded by (6.28). And (6.30) and the definition of $k(\ell, f)$, $h_1(\ell+1)$, $h_2(\ell+1)$ ensure that (6.28) is bounded independently of f . Using (6.26), Proposition 6.1, and Corollary 2.10, it suffices to show that $\gamma_{1/2,1}(\mathcal{G}_1) < \infty$.

Now, from (6.31) we see that we have $|C_1| \leq 1/2$ provided L has been taken large enough. Thus $|\{f_1 \neq f\}| < 1/2$, so that $|f_1(u)| \leq 1$ for some $u \in [0, 1]^2$, and thus, since $\|\partial f_1/\partial x\|_\infty, \|\partial f_1/\partial y\|_\infty \leq K$, we have

$$\mathcal{G}_1 \subset \{f: [0, 1]^2 \rightarrow \mathbb{R}^2, \|f\|_\infty \leq K, \|\frac{\partial f}{\partial x}\|_\infty, \|\frac{\partial f}{\partial y}\|_\infty \leq K\}$$

and thus $\gamma_{1/2,2}(\mathcal{G}_1) < \infty$ by Lemma 4.3. $\square$

7. SHOR'S DISCREPANCY THEOREM

In this section, we prove Theorem 1.3 and two related results.

Theorem 7.1. *Consider the class*

$$\mathcal{E}_U = \left\{ f: [0, 1]^2 \rightarrow \mathbb{R}; \right. \\ \left. \forall y \in [0, 1], \int_0^1 \left| \frac{\partial f}{\partial x}(s, y) \right| ds \leq 1; \int_0^1 \sup_x \left| \frac{\partial f}{\partial y}(x, s) \right| ds \leq 1 \right\}.$$

Then $ED_n(\mathcal{E}_U) \leq K(n \log n)^{1/2}$.

In our next example, we need the notion of a pair of conjugated Young functions Φ, Ψ . Use of these functions is not motivated by a vain search of generality but by the fact that these are the exact tools needed to write the best possible integral conditions. We refer to [Z] for definitions. The reader needs only to know that Φ, Ψ are functions from $\mathbb{R}$ to $\mathbb{R}$, are convex, symmetric, and increasing on $\mathbb{R}^+$, and satisfy Young's inequality

$$(7.1) \quad \forall s, t \in \mathbb{R}^+, \quad st \leq \Phi(s) + \Psi(t).$$

The most important examples of conjugated Young functions are the pairs $(t^p/p, t^q/q)$, $\frac{1}{p} + \frac{1}{q} = 1$.

We define the class $\mathcal{E}_Y$ as the class of functions f from $[0, 1]^2$ to $\mathbb{R}$ that satisfy

$$\iint \left(\Psi_1 \left(\frac{\partial f}{\partial x} \right) + \Psi_2 \left(\frac{\partial f}{\partial y} \right) \right) d\lambda \leq 1,$$

where the functions Ψ_1, Ψ_2 are Young functions (depending on f) such that the conjugated Young functions Φ_1, Φ_2 satisfy for some $0 < \alpha \leq 1$ (depending on f)

$$(7.2) \quad \int_0^1 t^\alpha \Phi_1(t^{-\alpha}) dt \leq 1,$$

$$(7.3) \quad \int_0^1 t^\alpha \Phi_2(\alpha t^{-1}) dt \leq 1,$$

$$(7.4) \quad \forall t > 0, \quad \Phi_2(t) \leq \exp(t^{1/5}).$$

The role of α is to establish the relative “weights” of the conditions on $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$. (The role of condition (7.4) is rather technical.) For $\alpha = 1$, these weights are equal. In that case, one can take, for example, $\Phi_1(t) = \Phi_2(t) = (2-p)t^p$ where $p < 2$ (in which case $\Psi_1(t) = t^q C(p)$ for $1/p + 1/q = 1$, $C(p) = q^{-1}(p(2-p)^{-1/(p-1)})$). When α decreases, the integrability of $t^\alpha \Phi_1(t^{-\alpha})$ becomes a weaker condition (as can be seen by a change of variable) while the integrability of $t^\alpha \Phi_2(\alpha t^{-1})$ becomes a stronger condition. Thus one can allow Φ_1 to increase faster (and thus Ψ_1 to increase slower) while one needs Ψ_2 to increase faster. A case of particular interest is when $1/p + 1/q < 1$, $q > 2$, where one can take $\Psi_1(t) = K(p, q)t^p$, $\Psi_2(t) = K(p, q)t^q$, as is seen with a choice of α with $q - 1 < \alpha < (p - 1)^{-1}$. The “limit case” $\alpha = 0$ is the case of Shor’s class, where the condition on $\frac{\partial f}{\partial x}$ is the weakest ($\iint |\frac{\partial f}{\partial x}| d\lambda \leq 1$) and the condition on $\frac{\partial f}{\partial y}$ the strongest ($\|\frac{\partial f}{\partial y}\|_\infty \leq 1$). The class $\mathcal{E}_Y$ contains all these cases *together*.

Theorem 7.2. $ED_n(\mathcal{E}_Y) \leq K(n \log n)^{1/2}$.

To prove Theorems 1.3, 7.1, and 7.2, one can replace the class $\mathcal{E}$ under study by the subclass $\mathcal{E}_0$ of functions that have zero integral. One then has to prove the following two statements.

Proposition 7.3. *If $\mathcal{E}$ is one of the classes of Theorems 1.3, 7.1, and 7.2, then $\gamma_{1/2, 2}(\mathcal{E}) < \infty$.*

Proposition 7.4. *If $\mathcal{E}$ is one of the classes of Theorems 1.3, 7.1, and 7.2, for $f_1, f_2 \in \mathcal{E}$ we have*

$$d_U(f_1, f_2) \leq K\sqrt{rn}(\|f_1 - f_2\|_2 + (\frac{\log n}{n})^{1/2}),$$

where r is given by (4.3).

Once this is known, the proof of these theorems is identical to that of Theorem 4.1.

We first prove Proposition 7.3. In the case of the class of Theorem 1.3, this follows directly from Theorem 6.5. (We leave to the reader the proof of the easy fact that $\|f - \iint f d\lambda\|_\infty \leq K$ for $f \in \mathcal{E}$.) We consider next the case of Theorem 7.2. Since $\Phi_1(t^{-\alpha}) \geq \Phi_1(1)$ for $t \leq 1$, it follows from (7.2) that $\Phi_1(1) \leq \alpha + 1 \leq 2$. By Young’s inequality we have $x \leq \Psi_1(x) + 2$, so that

$\Psi_1(x) \geq x - 2$. Thus, for $f \in \mathcal{E}_Y$, we have

$$(7.5) \quad \iint \left| \frac{\partial f}{\partial x} \right| d\lambda \leq \iint \left(\Psi_1 \left(\frac{\partial f}{\partial x} \right) + 2 \right) d\lambda \leq 3.$$

By (7.4) and Young's inequality we have

$$xy \leq \Psi_2(x) + \exp y^{1/5}.$$

Taking $y = (\log x)^5$ (when $x \geq 1$), we get

$$x(\log x)^{1/5} \leq \Psi_2(x) + x$$

so that, for some constant K , we have $x(\log(x+1))^5 \leq K(\Psi_2(x) + 1)$, so that, for $f \in \mathcal{E}_Y$, we have

$$(7.6) \quad \iint \left| \frac{\partial f}{\partial y} \right| \left(\log \left(\left| \frac{\partial f}{\partial y} \right| + 1 \right) \right)^5 d\lambda \leq K.$$

We should point out that (7.3) implies, by change of variable,

$$\int_0^{1/\alpha} \alpha^{1-\alpha} t^\alpha \Phi_2(t^{-1}) dt \leq 1.$$

Thus for $y \geq 1$, we get

$$\Phi_2(y) \int_0^{1/y} \alpha^{1-\alpha} t^\alpha dt \leq 1$$

so that

$$(7.7) \quad \Phi_2(y) \leq (\alpha + 1) \alpha^{\alpha-1} y^{\alpha+1} \leq 8\alpha^{-1} y^2.$$

When α stays bounded away from zero, this is much stronger than (7.4) for the large values of y that matter. However, when α goes to zero, (7.7) does not control $\Phi_2(y)$. This creates the need for condition (7.4).

In view of (7.5), (7.6), and Theorem 6.5, to prove that $\gamma_{1/2,2}(\mathcal{E}_{Y,0}) < \infty$ it suffices to know that $\|f - \iint f d\lambda\|_\infty \leq K$ for $f \in \mathcal{E}_Y$; this will follow from Lemma 7.9.

We now turn to the case of the class $\mathcal{E}_U$ of Theorem 7.1. The proof that $\gamma_{1/2,2}(\mathcal{E}_{U,0}) < \infty$ will not rely on Theorem 6.5 but rather will use related ideas and will rely on Proposition 6.4.

Consider a function f on $[0, 1]^2$, and assume that

$$(7.8) \quad \iint f d\lambda = 0,$$

$$(7.9) \quad \int_0^1 h(t) dt \leq 1, \quad \text{where } h(t) = \sup_x \left| \frac{\partial f}{\partial y}(x, t) \right|,$$

$$(7.10) \quad \forall y \in [0, 1], \quad \int \left| \frac{\partial f}{\partial x}(s, y) \right| ds \leq 1.$$

Consider $\ell \geq 1$, and for $y \in [0, 1]$ consider the set

$$A_\ell(y) = \left\{ x \in [0, 1], \exists t > 0, \int_{I(x,t)} \left| \frac{\partial f}{\partial x}(s, y) \right| ds \geq Lt2^\ell \right\}$$

where $I(x, t) = [0, 1] \cap [x - t, x + t]$ and where the parameter L will be adjusted presently. Lemma 6.5 and (7.10) imply that $|A_\ell(y)| \leq K2^{-\ell}/L$, so one can fix L such that $|A_\ell(y)| \leq 2^{-\ell}$.

We set $H(t) = \int_0^t h(s) ds$. We observe that, by (7.9), we have

$$|f(x, y) - f(x, y')| \leq |H(y) - H(y')|.$$

Consider now the function f_ℓ given by

$$f_\ell(x, y) = \inf\{f(r, s) + |H(s) - H(y)| + L2^\ell|x - r|; (r, s) \in [0, 1]^2\}.$$

The argument of the first step of the proof of Theorem 6.5 shows that $f = f_\ell$ outside $A_\ell = \bigcup_y A_\ell(y) \times \{y\}$ and that

$$(7.11) \quad \forall x, x', y \in [0, 1], \quad |f_\ell(x, y) - f_\ell(x', y)| \leq L2^\ell|x - x'|,$$

$$(7.12) \quad \forall x, y, y' \in [0, 1], \quad |f_\ell(x, y) - f_\ell(x, y')| \leq |H(y) - H(y')|.$$

We write

$$(7.13) \quad f = f_1 + \sum_{\ell \geq 1} (f_{\ell+1} - f_\ell).$$

Since $A_{\ell+1} \subset A_\ell$, we have $f_1 \in \mathcal{G}_0$, $f_{\ell+1} - f_\ell \in \mathcal{G}_\ell$, where $\mathcal{G}_\ell$ is the class of functions g that satisfy

$$(7.14) \quad \forall y \in [0, 1], |\{x, g(x, y) \neq 0\}| \leq 2^{-\ell};$$

$$(7.15) \quad \forall x, x', y \in [0, 1], |g(x, y) - g(x', y)| \leq L2^{\ell+1}|x - x'|;$$

and

$$(7.16) \quad \begin{aligned} &\exists h \geq 0, \int_0^1 h(t) dt \leq 1, \forall x, y, y' \in [0, 1], \\ &y \leq y' \Rightarrow |g(x, y) - g(x, y')| \leq 2 \int_y^{y'} h(t) dt. \end{aligned}$$

For $\ell \geq 1$, it follows from (7.14), (7.15) that $\|g\|_\infty \leq 2L$ whenever $g \in \mathcal{G}_\ell$, so that, by (7.14) again, $\|g\|_1 \leq L2^{-\ell+1}$. Since $\iint f d\lambda = 0$, it follows from (7.13) that we can, moreover, require that $|\iint f_1 d\lambda| \leq 4L$ for $f_1 \in \mathcal{G}_0$.

It follows from Corollary 2.10 and the inclusion $\mathcal{G}_0 \subset \sum_{\ell \geq 0} \mathcal{G}_\ell$ that

$$(7.17) \quad \gamma_{1/2, 2}(\mathcal{G}_0) \leq \left(\sum_{\ell \geq 0} \gamma_{1/2, 2}(\mathcal{G}_\ell)^{2/3} \right)^{1/3}.$$

To estimate $\gamma_{1/2, 2}(\mathcal{G}_\ell)$, we apply again the same approximation method. Consider $g \in \mathcal{G}_\ell$, $k \geq 1$, and h as in (7.16). To avoid irrelevant technicalities, we will assume that $\frac{\partial g}{\partial y}$ exists (an assumption that can easily be removed) so that (7.16) implies $|\frac{\partial g}{\partial y}(x, y)| \leq 2h(y)$ a.e. Consider the set $B = \{g \neq 0\}$. By (7.14) and Fubini theorem, we have $|B| \leq 2^{-\ell}$. Observe that if $\frac{\partial g}{\partial y}(x, y) \neq 0$, then $(x, t) \in B$ for $t \neq y$, $t - y$ small enough. The set B' of couples (x, y)

with this property that $(x, t) \in B$ for $t \neq y$, $t - y$ small enough is easily seen to satisfy $|B'| = |B|$. And we have

$$\left| \frac{\partial g}{\partial y}(x, y) \right| \leq 2h(y)1_{B'}(x, y) \leq 2h(y)1_B(x, y) \quad \text{a.e.}$$

Thus, if we integrate in x using (7.14) we get

$$\int_0^1 \left| \frac{\partial g}{\partial y}(x, y) \right| dx \leq 2^{-\ell+1} h(y) \quad \text{a.e.}$$

so that, by Fubini theorem $\left\| \frac{\partial g}{\partial y} \right\|_1 \leq 2^{-\ell+1}$.

Comment. This inequality is the key to our success. To obtain it, we have made essential use of the fact that for all $y \in [0, 1]$ we have $\int_0^1 \left| \frac{\partial f}{\partial x}(s, y) \right| ds \leq 1$. It is an interesting question to know if we could instead relax this condition to $\left\| \frac{\partial f}{\partial x} \right\|_1 \leq 1$. (This would not, however, provide a common generalization to Theorems 1.3 and 7.1, since the corresponding class fails Proposition 7.4.) To reformulate this question, consider $p, q \geq 1$, and the class $\mathcal{H}_{p,q}$ of functions f on $[0, 1]^2$ that satisfy $\iint f d\lambda = 0$ and

$$(7.18) \quad \text{the function } h(y) = \sup_x \left| \frac{\partial f}{\partial y}(x, y) \right| \text{ satisfies } \|h\|_p \leq 1,$$

$$(7.19) \quad \text{the function } g(y) = \int \left| \frac{\partial f}{\partial x}(s, y) \right| ds \text{ satisfies } \|g\|_q \leq 1.$$

Thus $\mathcal{H}_{\infty,1}$ is the class of Theorem 1.3, while $\mathcal{H}_{1,\infty}$ is the class of Theorem 7.1. The question is to decide whether $\gamma_{1/2,2}(\mathcal{H}_{1,1}) < \infty$. We leave the reader to check that the obvious adaptation of the arguments presented in this paper show that $\gamma_{1/2,2}(\mathcal{H}_{p,q}) < \infty$ whenever $pq > 1$.

We continue the proof. Consider $k \geq 1$ and the set

$$B_k = \left\{ (x, y); \exists t > 0, \int_{I(y,t)} \left| \frac{\partial g}{\partial y}(x, s) \right| ds \geq L2^k t \right\}$$

where L is a parameter to be adjusted presently. As shown during the proof of Theorem 6.5, we have

$$|B_k| \leq \frac{K}{L2^k} \left\| \frac{\partial g}{\partial y} \right\|_1 \leq \frac{K}{L} 2^{-\ell-k+1} \leq 2^{-\ell-k-2}$$

by a suitable choice of L . Consider now the function g_k given by

$$g_k(x, y) = \inf \{ g(r, s) + L2^k |y - s| + 2^\ell |x - r|; (r, s) \notin B_k \}.$$

Then (as in the proof of Theorem 6.5) we have $g_k = g$ outside B_k while $\|\partial g_k / \partial x\|_\infty \leq 2^\ell$, $\|\partial g_k / \partial y\|_\infty \leq L2^k$. Thus writing $g = g_1 + \sum_{k \geq 1} g_{k+1} - g_k$, we see that we have $\mathcal{G}_\ell \subset \sum_{k \geq 0} \mathcal{G}_{\ell,k}$, where for $k \geq 1$, $\mathcal{G}_{\ell,k}$ is the class of functions h that satisfy

$$|\{h \neq 0\}| \leq 2^{-\ell-k-2}, \quad \left\| \frac{\partial h}{\partial x} \right\|_\infty \leq 2^{\ell+1} \leq L2^{\ell+1}, \quad \left\| \frac{\partial h}{\partial y} \right\|_\infty \leq L2^{k+1}$$

and where $\mathcal{G}_{\ell,0}$ is the class of functions h that satisfy

$$|\{h \neq 0\}| \leq 2^{-\ell} + 2^{-\ell-3}, \quad \|\frac{\partial h}{\partial x}\|_{\infty} \leq 2^{\ell+1}, \quad \|\frac{\partial h}{\partial y}\|_{\infty} \leq 2L.$$

If $k \geq 1$, or if $k = 0, \ell \geq 4$, we can use Proposition 6.4 to get

$$(7.20) \quad \gamma_{1/2,2}(\mathcal{G}_{\ell,k}) \leq K\sqrt{\ell+k}2^{-(\ell+k)/2}, \quad \gamma_{1/2,2}(\mathcal{G}_{\ell,0}) \leq K\sqrt{\ell}2^{-\ell/2}.$$

Using Corollary 2.10, we see that $\gamma_{1/2,2}(\mathcal{G}_{\ell}) \leq K\sqrt{\ell}2^{-\ell/2}$ for $\ell \geq 4$. To prove the finiteness of the right-hand side of (7.17), it suffices to show that $\gamma_{1/2,2}(\mathcal{G}_{\ell}) < \infty$ for $\ell = 0, 1, 2, 3$. In view of (7.20) and Corollary 2.10, it suffices to show that $\gamma_{1/2,2}(\mathcal{G}_{\ell,0}) < \infty$ for $\ell = 0, 1, 2, 3$. For $\ell = 1, 2, 3$, this follows from Lemma 4.3 since a function in $\mathcal{G}_{\ell,0}$ must take the value 0. For $\ell = 0$, it suffices, in view of the same lemma, to prove that $\mathcal{G}_{0,0}$ is uniformly bounded. But we remember that $|\iint f_1 d\lambda| \leq K$ for $f_1 \in \mathcal{G}_0$. And the same argument as was used to prove that fact implies that $|\iint g d\lambda| \leq K$ for $g \in \mathcal{G}_{0,0}$. $\square$

We now turn to the proof of Proposition 7.4. The arguments in the cases we consider have a lot in common, but extra complications occur in the case of Theorem 7.1. We will first cover the cases of Theorems 1.3 and 7.2. In order to avoid repetition of the complete argument in the case of Theorem 7.1, we formulate some abstract principles.

Lemma 7.5. *Consider a function f on a probability space (Ω, μ) and a sequence $(\mathcal{P}_k)_{k \in \mathbb{Z}}$ of finite partitions of Ω (note that the sequence $(\mathcal{P}_k)$ is not assumed to be increasing). Set*

$$B_k = \bigcup \{A \in \mathcal{P}_k; m_A(f) \geq 2^k\}.$$

Then $\sum_{k \in \mathbb{Z}} 2^{2k} \mu(B_k) \leq 8 \int f^2 d\mu$.

Proof. Consider $A \in \mathcal{P}_k$, $A \subset B_k$. Then $\int_A f d\mu \geq 2^k \mu(A)$. Since

$$\int_A f 1_{\{f \leq 2^{k-1}\}} d\mu \leq 2^{k-1} \mu(A),$$

we must have

$$\int_A f 1_{\{f \geq 2^{k-1}\}} d\mu \geq 2^{k-1} \mu(A).$$

Summing this inequality over all the sets $A \in \mathcal{P}_k$, $A \subset B_k$ we get

$$2^{k-1} \mu(B_k) \leq \int_{B_k} f 1_{\{f \geq 2^{k-1}\}} d\mu$$

so that

$$\sum_k 2^{2k} \mu(B_k) \leq 2 \sum_k \int_{B_k} 2^k f 1_{\{f \geq 2^{k-1}\}} d\mu.$$

For $\omega \in \Omega$, $f(\omega) > 0$, consider the largest $\ell \in \mathbb{Z}$ such that $2^{\ell-1} \leq f(\omega)$. Then $\sum_{k \leq \ell} 2^{2k} \leq 2^{\ell+1} \leq 4f(\omega)$. Thus $\sum_k 2^{2k} \mu(B_k) \leq 8 \int f^2 d\mu$. $\square$

Lemma 7.6. Consider a function g on a probability space (Ω, μ) and a sequence $(\mathcal{P}_k)_{k \leq k_0}$ of finite partitions of Ω . (Observe that the sequence $(\mathcal{P}_k)$ is not assumed to be increasing.) Assume that the following holds:

$$(7.21) \quad \text{Given } A \in \mathcal{P}_k, \text{ the union } A' \text{ of the sets } B \text{ of } \mathcal{P}_{k-1} \text{ that satisfy } \mu(A \cap B) > 0 \text{ satisfies } \mu(A') \leq 2\mu(A).$$

For $k \leq k_0$, consider the set

$$V_k = \bigcup \{A \in \mathcal{P}_k; m_A(g) \geq 2^{2k}\}.$$

Then $\sum_{k \leq k_0} 2^{2k} \mu(V_k) \leq 2 \int |g| d\mu$.

Proof. Step 1. Denote by $\mathcal{A}_k$ the algebra generated by $\mathcal{P}_k$. For $A \in \mathcal{A}_k$, we set

$$T_{k-1}(A) = \bigcup \{B \in \mathcal{P}_{k-1}; |\mu(A \cap B)| > 0\}$$

and define $S_p(A)$ by induction for $p \leq k$ as $S_k(A) = A$ and $S_p(A) = T_p(S_{p+1}(A))$ for $p < k$. By induction over $k - p$, one sees from (7.21) that

$$(7.22) \quad \mu(S_p(A)) \leq 2^{k-p} \mu(A).$$

Observe also that $A \subset S_p(A)$ and that for $A \in \mathcal{A}_k$

$$S_p(A) = \bigcup \{S_p(B); B \subset A, B \in \mathcal{P}_k\}.$$

Step 2. We set $H_p = V_p \setminus \bigcup_{p < \ell \leq k_0} S_p(V_\ell)$. Thus the sets (H_p) are disjoint. Since H_p is $\mathcal{A}_p$ -measurable, the definition of V_p shows that

$$\int_{H_p} g d\mu \geq 2^{2p} \mu(H_p).$$

Thus, since the sets H_p are disjoint, we have

$$(7.23) \quad \sum_{p \leq k_0} 2^{2p} \mu(H_p) \leq \int |g| d\mu.$$

Step 3. We show that

$$V_k \subset H_k \cup \bigcup_{k < p \leq k_0} S_k(H_p).$$

Indeed, consider $A \in \mathcal{A}_k$ such that $A \subset V_k$, $A \not\subset H_k$. Thus $A \subset S_k(V_p)$ for some $p > k$. Take $p \leq k_0$, the largest possible. Consider $B \subset V_p$, $B \in \mathcal{P}_p$ such that $A \subset S_k(B)$. We cannot have $B \subset S_p(V_\ell)$ for any $p < \ell \leq k$, for otherwise we would have $A \subset S_k(V_\ell)$, which would contradict the maximality of p . Thus $B \subset H_p$, which finishes the proof.

Step 4. We have, by (7.22)

$$\mu(V_k) \leq \mu(H_k) + \sum_{k < p \leq k_0} \mu(S_k(H_p)) \leq \sum_{k \leq p \leq k_0} 2^{p-k} \mu(H_p).$$

Thus,

$$\sum_{k \leq k_0} 2^{2k} \mu(V_k) \leq \sum_{k \leq p \leq k_0} 2^{p+k} \mu(H_p) \leq \sum_{p \leq k_0} 2^{2p+1} \mu(H_p) \leq 2 \int |g| d\mu$$

where the last inequality follows from (7.23). $\square$

Proposition 7.7. *Consider a function f on $[0, 1]^2$ and $L \geq 1$. Assume $\|f\|_\infty \leq L$. Consider a function g on $[0, 1]^2$, such that $g \geq 0$, $\iint g d\lambda \leq L$. Consider the largest integer ℓ_0 such that $2^{-2\ell_0} \geq \frac{\log n}{n}$. Suppose that for $0 \leq \ell \leq \ell_0$ there exists a partition $\mathcal{P}_\ell$ of $[0, 1]^2$ in $2^{2\ell_0}$ rectangles of area $2^{-2\ell_0}$ with the property*

$$(7.24) \quad A \in \mathcal{P}_\ell, \quad m_A(g) \leq 2^{2\ell} \Rightarrow \|f - m_A(f)\|_{\infty, A} \leq L2^{\ell-\ell_0}.$$

Suppose, moreover, that the following holds.

$$(7.25) \quad \text{Given } A \in \mathcal{P}_\ell, \ell > 0, \text{ the union of the elements } B \text{ of } \mathcal{P}_{\ell-1} \text{ for which } |A \cap B| > 0 \text{ has measure } \leq 2|A|.$$

Then

$$\left(\frac{1}{n} \sum_{i \leq n} f(U_i)^2 \right)^{1/2} \leq K(L) r \left(\|f\|_2 + \left(\frac{\log n}{n} \right)^{1/2} \right),$$

where r is the random variable introduced in Lemma 4.3, and $K(L)$ depends on L only.

Proof. For $\ell \leq \ell_0$, we consider the set B_ℓ (resp. V_ℓ) that is the union of the sets $A \in \mathcal{P}_\ell$ for which $m_A(f) \geq 2^{\ell-\ell_0}$ (resp. $m_A(g) \geq 2^{2\ell}$).

By (7.24), for $A \in \mathcal{P}_\ell$, $A \not\subset B_\ell \cup V_\ell$, we have for $u \in A$

$$f(u) \leq \|f - m_A(f)\|_{\infty, A} + m_A(f) \leq (L+1)2^{\ell-\ell_0} \leq L2^{\ell-\ell_0+1}.$$

Thus

$$\text{card}\{i \leq n; f(U_i) \geq L2^{\ell-\ell_0+1}\} \leq \text{card}\{i \leq n; U_i \in B_\ell \cup V_\ell\}.$$

We now recall from Lemma 4.3 that if a set Z is union of disjoint rectangles of area $2^{-2\ell_0}$ (in particular, if Z is $\mathcal{A}_\ell$ measurable) we have

$$\text{card}\{i \leq n; U_i \in Z\} \leq 2rn|Z|.$$

Thus we have

$$\text{card}\{i \leq n; f(U_i) \geq L2^{\ell-\ell_0+1}\} \leq 2rn(|B_\ell| + |V_\ell|)$$

so that

$$\begin{aligned} \sum_{0 \leq \ell \leq \ell_0} 2^{2(\ell-\ell_0)} \text{card}\{i \leq n; f(U_i) \geq L2^{\ell-\ell_0+1}\} \\ \leq 2rn \left[2^{-2\ell_0} \sum_{0 \leq \ell \leq \ell_0} 2^{2\ell} |V_\ell| + \sum_{0 \leq \ell \leq \ell_0} 2^{2(\ell-\ell_0)} |B_\ell| \right]. \end{aligned}$$

We now appeal to Lemma 7.5 (resp. 7.6) to get that

$$\sum_{0 \leq \ell \leq \ell_0} 2^{2(\ell - \ell_0)} |B_\ell| \leq 8 \|f\|_2^2, \quad \sum_{0 \leq \ell \leq \ell_0} 2^{2\ell} |V_\ell| \leq 2 \|g\|_1.$$

So we get

$$\sum_{-\ell_0 \leq \ell \leq 0} 2^{2\ell} \text{card}\{i \leq n; f(U_i) \geq L2^{\ell+1}\} \leq K r n (\|f\|_2^2 + L2^{-2\ell_0}).$$

Since $\|f\|_\infty \leq L$, we have

$$\begin{aligned} & \sum_{i \leq n} \{f(U_i)^2; f(U_i) \geq 0\} \\ & \leq 4L^2 \sum_{-\ell_0 \leq \ell \leq 0} 2^{2\ell} \text{card}\{i \leq n; f(U_i) \geq L2^{\ell+1}\} + 4L^2 n 2^{-2\ell_0}. \end{aligned}$$

To finish the proof, we apply again the same argument to $-f$ instead of f , and we remember that $2^{-2\ell_0} \leq 4 \frac{\log n}{n}$. $\square$

Lemma 7.8. Consider a function f on $I \times J$. Assume that

$$(7.26) \quad \forall y, y' \in J, \forall x \in I, \quad |f(x, y) - f(x, y')| \leq a,$$

$$(7.27) \quad \exists z \in J, \quad \frac{1}{|J|} \int_J \left| \frac{\partial f}{\partial x}(s, z) \right| ds \leq b.$$

Then we have

$$(7.28) \quad \|f - m_{I \times J}(f)\|_{\infty, I \times J} \leq 2a + |J|b.$$

Proof. Consider $x, x' \in I, y, y' \in J$. By (7.27) we have

$$|f(x, z) - f(x', z)| \leq |J|b,$$

while $|f(x, y) - f(x, z)| \leq a, |f(x', z) - f(x', y')| \leq a$. Thus $|f(u) - f(v)| \leq 2a + |J|b$ for $u, v \in I \times J$, whence the result. $\square$

We now prove that if a function f satisfies $\|\frac{\partial f}{\partial x}\|_1 \leq 1, \|\frac{\partial f}{\partial y}\|_\infty \leq 1, \iint f d\lambda = 0$, then it satisfies the condition of Proposition 7.7. We set $g = |\frac{\partial f}{\partial x}|$. We observe that if $A = I \times J$, then by the Fubini theorem condition (7.27) holds for $b = m_A(g)$. In particular, taking $A = [0, 1]^2$, we see from Lemma 7.8 that $\|f\|_\infty \leq 3$. For $0 \leq \ell \leq \ell_0$, we consider the partition $\mathcal{P}_\ell$ of $[0, 1]$ in $2^{2\ell_0}$ rectangles of sides $2^{-\ell_0-\ell}, 2^{-\ell_0+\ell}$, so (7.25) holds. We see that if $A = I \times J \in \mathcal{P}_\ell$, then (7.26) holds with $a = 2^{-\ell_0+\ell}$ while, as observed, (7.27) holds for $b = m_A(g)$. Thus, since $|J| = 2^{-\ell_0-\ell}$, we have $\|f - m_A(f)\|_{\infty, A} \leq 3 \cdot 2^{\ell-\ell_0}$ whenever $m_A(g) \leq 2^{2\ell}$. This proves (7.24), and this finishes the proof of Theorem 1.3.

We now turn to the case of Theorem 7.2. The key is a Sobolev-type inequality. (A similar inequality occurs in [T3].)

Lemma 7.9. Consider conjugated Young functions (Φ_1, Ψ_1) , (Φ_2, Ψ_2) , and $0 < \alpha \leq 1$ and a function f on $[0, 1]^2$. Assume that (7.2), (7.3) hold, and consider a rectangle A of sides a, b contained in $[0, 1]^2$. Then for all numbers $\eta, \delta \leq 1$ we have

$$(7.29) \quad \|f - m_A(f)\|_{\infty, A} \leq Ka \left(\eta^{-1/\alpha} + \eta m_A \left(\Psi_1 \left(\frac{\partial f}{\partial x} \right) \right) \right) + Kb \left(\delta^{-\alpha} + \delta m_A \left(\Psi_2 \left(\frac{\partial f}{\partial y} \right) \right) \right).$$

Proof. To simplify the notation, we will assume that $A = [0, a] \times [0, b]$, and we consider $(x, y) \in A$. For definiteness, we assume $x \leq a/2$. For $x \leq s \leq a$, we consider the two functions

$$\varphi(s) = y + \left(\frac{s-x}{a-x} \right)^\alpha (-y), \quad \psi(s) = y + \left(\frac{s-x}{a-x} \right)^\alpha (b-y).$$

Since $x \leq a/2$, for $0 \leq t \leq 1$, we have

$$(7.30) \quad |t\psi'(s) + (1-t)\varphi'(s)| = \alpha \left| \frac{tb-y}{a-x} \right| \cdot \left(\frac{s-x}{a-x} \right)^{\alpha-1} \leq 2\alpha \frac{b}{a} \left(\frac{s-x}{a-x} \right)^{\alpha-1}.$$

For $x \leq s \leq a$, we set

$$u(s, t) = (s, \varphi(s) + t(\psi(s) - \varphi(s)))$$

and

$$(7.31) \quad h(s) = \int_0^1 f(u(s, t)) dt.$$

Thus $h(x) = f(x, y)$, while $h(a) = \int_0^1 f(a, tb) dt$.

We now have, by differentiation of (7.31) under the integral that

$$|h(a) - h(x)| \leq \int_x^a |h'(s)| ds \leq I_1 + I_2$$

where

$$I_1 = \int_x^a ds \int_0^1 \left| \frac{\partial f}{\partial x}(u(s, t)) \right| dt$$

and, by (7.30),

$$\begin{aligned} I_2 &= \int_x^a ds \int_0^1 |t\psi'(s) + (1-t)\varphi'(s)| \cdot \left| \frac{\partial f}{\partial y}(u(s, t)) \right| dt \\ &\leq 2\alpha \frac{b}{a} \int_x^a ds \int_0^1 \left(\frac{s-x}{a-x} \right)^{\alpha-1} \left| \frac{\partial f}{\partial y}(u(s, t)) \right| dt. \end{aligned}$$

To relate these integrals to integrals with respect to Lebesgue measure, we have to make the Jacobian $\psi(s) - \varphi(s)$ appear. To that purpose, we observe that by Young's inequality (7.1) we have

$$\frac{z}{\eta w^\alpha} \leq \Phi_1 \left(\frac{1}{\eta w^\alpha} \right) + \Psi_1(z)$$

so that

$$z \leq \eta w^\alpha \Phi_1 \left(\frac{1}{\eta w^\alpha} \right) + \eta w^\alpha \Psi_1(z).$$

We use this for $z = |\frac{\partial f}{\partial x}(u(s, t))|$, $w = \frac{s-x}{a-x}$, and we obtain $I_1 \leq I_3 + I_4$, where

$$I_3 = \int_x^a \eta \left(\frac{s-x}{a-x} \right)^\alpha \Phi_1 \left(\frac{1}{\eta(s-x)/(a-x)^\alpha} \right) ds,$$

$$I_4 = \eta \int_x^a ds \int_0^1 \left(\frac{s-x}{a-x} \right)^\alpha \Psi_1 \left(\frac{\partial f}{\partial x}(u(s, t)) \right) dt.$$

To study I_3 , we set $s = \eta^{-1/\alpha} v(a-x) + x$ to get

$$I_1 = (a-x) \eta^{-1/\alpha} \int_0^{\eta^{1/\alpha}} v^\alpha \Phi_1(v^{-\alpha}) dv \leq a \eta^{-1/\alpha}$$

by (7.2) and since $\eta^{1/\alpha} \leq 1$.

To study I_4 , we set $z = \varphi(s) + t(\psi(s) - \varphi(s))$, so that, since $\psi(s) - \varphi(s) = b(\frac{s-x}{a-x})^\alpha$, we have

$$I_4 = \frac{\eta}{b} \int_x^a ds \int_{\varphi(s)}^{\psi(s)} \Psi_1 \left(\frac{\partial f}{\partial x}(s, z) \right) dz = \frac{\eta}{b} \iint_W \Psi_1 \left(\frac{\partial f}{\partial x} \right) d\lambda$$

where W is the region $\{(s, z); x \leq s \leq a, \varphi(s) \leq z \leq \psi(s)\} \subset A$. Thus $I_4 \leq \eta a m_A(\Psi_1(\frac{\partial f}{\partial x}))$.

We now turn to the study of I_2 . Using the same method, we write now

$$\frac{\alpha z}{\delta w} \leq \Phi_2 \left(\frac{\alpha}{\delta w} \right) + \Psi_2(z)$$

so that

$$\alpha z w^{\alpha-1} \leq w^\alpha \delta \Phi_2 \left(\frac{\alpha}{\delta w} \right) + \delta w^\alpha \Psi_2(z).$$

We use this for $z = |\frac{\partial f}{\partial y}(u(s, t))|$, $w = \frac{s-x}{a-x}$, and we obtain $I_2 \leq I_5 + I_6$, where

$$I_5 = 2 \frac{b}{a} \delta \int_x^a \left(\frac{s-x}{a-x} \right)^\alpha \Phi_2 \left(\frac{\alpha}{\delta(s-x)/(a-x)} \right) ds,$$

$$I_6 = 2 \frac{b}{a} \delta \int_x^a ds \int_0^1 \left(\frac{s-x}{a-x} \right)^\alpha \Psi_2 \left(\frac{\partial f}{\partial y}(u(s, t)) \right) dt.$$

To study I_5 , we set $s = x + \delta^{-1} v(a-x)$ to get, by (7.3) and since $\delta \leq 1$,

$$I_5 \leq 2 \frac{b}{a} (a-x) \delta^{-\alpha} \int_0^\delta v^\alpha \Phi_2(\alpha v^{-1}) dt \leq 2b \delta^{-\alpha}.$$

To study I_6 , we set $z = \varphi(s) + t(\psi(s) - \varphi(s))$ to get

$$I_6 = 2 \frac{\delta}{a} \int_x^a ds \int_{\varphi(s)}^{\psi(s)} \Psi_2 \left(\frac{\partial f}{\partial y}(s, z) \right) dz.$$

As in the case of I_4 , this yields

$$I_6 \leq 2b \delta m_A(\Psi_2(\frac{\partial f}{\partial y})).$$

We have now shown that $|f(x, y) - \frac{1}{b} \int_0^b f(a, y) dy| \leq E$ is bounded by the right-hand side of (7.29). We now give an argument in the spirit of the previous

one, but much simpler, to compare $m_A(f)$ and $\frac{1}{b} \int_0^b f(a, y) dy$. For $x_1 \in [0, a]$, we have

$$|f(x_1, y) - f(a, y)| \leq \int_0^a \left| \frac{\partial f}{\partial x}(x, y) \right| dx$$

so that, setting $\theta(x) = \frac{1}{b} \int_0^b f(x, y) dy$, we have

$$|\theta(x_1) - \theta(a)| \leq \frac{1}{b} \iint_A \left| \frac{\partial f}{\partial x}(x, y) \right| dx dy \leq a m_A \left(\left| \frac{\partial f}{\partial x} \right| \right).$$

Averaging over $x_1 \in [0, a]$ gives

$$|\theta(a) - m_A(f)| \leq a m_A \left(\left| \frac{\partial f}{\partial x} \right| \right).$$

Now, by Young's inequality, we have

$$\left| \frac{\partial f}{\partial x} \right| \leq \eta \Phi_1 \left(\frac{1}{\eta} \right) + \eta \Psi_1 \left(\frac{\partial f}{\partial x} \right)$$

so that

$$m_A \left(\left| \frac{\partial f}{\partial x} \right| \right) \leq \eta \Phi_1 \left(\frac{1}{\eta} \right) + \eta m_A \left(\Psi_1 \left(\frac{\partial f}{\partial x} \right) \right).$$

Since $\Phi_1(1/z^\alpha) \geq \Phi_1(\frac{1}{\eta})$ for $z \leq \eta^{1/\alpha}$, we have

$$\begin{aligned} \eta^{1+1/\alpha} \Phi_1 \left(\frac{1}{\eta} \right) &= (\alpha + 1) \int_0^{\eta^{1/\alpha}} z^\alpha \Phi_1 \left(\frac{1}{\eta} \right) dz \\ &\leq (\alpha + 1) \int_0^{\eta^{1/\alpha}} z^\alpha \Phi_1 \left(\frac{1}{z^\alpha} \right) dz \leq 2 \end{aligned}$$

so that $\eta \Phi_1(\frac{1}{\eta}) \leq K \eta^{-1/\alpha}$. The proof is complete. $\square$

We now prove Proposition 7.4 in the case of $\mathcal{E}_Y$. This is done by checking the hypothesis of Proposition 7.7. Consider f on $[0, 1]^2$, such that $\iint f d\lambda = 0$ and $\iint g d\lambda \leq 1$, where $g = \Psi_1(\frac{\partial f}{\partial x}) + \Psi_2(\frac{\partial f}{\partial y})$. If we use (7.29) with $A = [0, 1]^2$, $\eta = \delta = 1$, we see that $\|f\|_\infty \leq K$. We set now $c = 2^{(1-\alpha)/(1+\alpha)} \leq 2$. For $0 \leq \ell \leq \ell_0$, denote by $h(\ell)$ the smallest integer such that $h(\ell) \geq \frac{1-\alpha}{1+\alpha} \ell$. Thus $c^{2\ell} \leq 2^{2h(\ell)} \leq 4c^{2\ell}$. Consider the partition $\mathcal{P}_\ell$ of $[0, 1]^2$ in rectangles of sides $a = 2^{-h(\ell)-\ell_0}$, $b = 2^{h(\ell)-\ell_0}$. Then (7.25) holds. Assume that for a rectangle A of $\mathcal{P}_\ell$ we have $m_A(g) \leq 2^{2\ell}$. Thus by (7.29), for all $\eta, \delta \leq 1$ we have

$$\|f - m_A(f)\|_{\infty, A} \leq K a (\eta^{-1/\alpha} + \eta 2^{2\ell}) + K b (\delta^{-\alpha} + \delta 2^{2\ell}).$$

Taking $\eta = (c/2)^\ell$, $\delta = (1/2c)^\ell$ yields

$$\|f - m_A(f)\|_{\infty, A} \leq K (a c^\ell 2^\ell + b 2^\ell c^{-\ell}).$$

Now $a = \sqrt{ab} \sqrt{a/b} \leq 2^{-\ell_0} c^{-\ell}$ and $b = \sqrt{ab} \sqrt{b/a} \leq 2^{-\ell_0+1} c^\ell$. Thus

$$\|f - m_A(f)\|_{\infty, A} \leq K 2^{\ell-\ell_0}.$$

The proof of Theorem 7.2 is complete.

We now turn to the proof of Proposition 7.4 in the case of $\mathcal{E}_U$. Consider a function f on $[0, 1]$ such that $\int f d\lambda = 0$ and that

$$(7.32) \quad \forall y \in [0, 1], \quad \int \left| \frac{\partial f}{\partial x}(s, y) \right| ds \leq 1,$$

$$(7.33) \quad \forall x, y, y' \in [0, 1], \quad |f(x, y) - f(x, y')| \leq |H(y) - H(y')|$$

where $H : [0, 1] \rightarrow [0, 1]$ is nondecreasing. For $0 \leq \ell \leq \ell_0$, $1 \leq k \leq 2^{\ell_0 - \ell}$ we define

$$I_{\ell, k} = H^{-1}((k-1)2^{\ell - \ell_0}, k2^{\ell - \ell_0}).$$

Thus, for all x and $y, y' \in I_{\ell, k}$, we have $|f(x, y) - f(x, y')| \leq 2^{\ell - \ell_0}$.

Consider the partition $\mathcal{P}_\ell$ of $[0, 1]^2$ created by the rectangles obtained when $[0, 1]^2$ is partitioned in $2^{\ell + \ell_0}$ equal vertical strips of width $2^{-\ell - \ell_0}$ and in the horizontal strips defined by the intervals $I_{\ell, k}$. As in the case of $\mathcal{E}_S$ (Lemma 7.8), we have for $A \in \mathcal{P}_\ell$

$$(7.34) \quad m_A(|\frac{\partial f}{\partial x}|) \leq 2^{2\ell} \Rightarrow \|f - m_A(f)\|_{\infty, A} \leq 3 \cdot 2^{\ell - \ell_0}.$$

The only reason for which we cannot use Proposition 7.7 is that the rectangles of $\mathcal{P}_\ell$ are not all of area $2^{-2\ell_0}$. Some are of smaller area and possibly contain abnormally many points U_i . We observe, however, that if a subset A of $[0, 1]$ is the union of q disjoint rectangles, then by (4.3) we have

$$(7.35) \quad \text{card}\{i \leq n; U_i \in A\} \leq rn(|A| + q \frac{\log n}{n}).$$

The proof will follow the pattern of the proof of Proposition 7.7. What we have to show is that the contribution in (7.35) of the term $q \frac{\log n}{n}$ does not change the result. This is the purpose of the next two lemmas. For $y \in [0, 1]$, we define g_y by $g_y(x) = |\frac{\partial f}{\partial x}(x, y)|$.

We define W_ℓ as the union of the sets A in $\mathcal{P}_\ell$ with the following property. A is a rectangle $B \times I_{\ell, k}$, and for all y in $I_{\ell, k}$, we have

$$\frac{1}{|B|} \int_B g_y(s) ds \geq 2^\ell.$$

Lemma 7.10. *We have*

$$\sum_{0 \leq \ell \leq \ell_0} 2^{2\ell} \text{card}\{i \leq n; U_i \in W_\ell\} \leq Krn.$$

We define B_ℓ as the union of the sets A in $\mathcal{P}_\ell$ for which $m_A(f) \geq 2^{\ell - \ell_0 + 3}$.

Lemma 7.11. *We have*

$$\sum_{0 \leq \ell \leq \ell_0} 2^{2(\ell - \ell_0)} \text{card}\{i \leq n; U_i \in B_\ell\} \leq Krn \left[\|f\|_2^2 + \frac{\log n}{n} \right].$$

Once these two lemmas are proved, the conclusion follows using (7.34) and repeating (the easy) part of the proof of Proposition 7.7.

We first prove Lemma 7.10. Consider the algebra $\mathcal{B}_\ell$ on $[0, 1]$ obtained by partitioning $[0, 1]$ in $2^{\ell + \ell_0}$ equal intervals. Consider $y \in [0, 1]$, and define

$V_\ell(y)$ as the union of the sets $B \in \mathcal{B}_\ell$ for which $m_B(g_y) \geq 2^{2\ell}$. It follows from Lemma 7.6 and the fact that $\int_0^1 g_y(s) ds \leq 1$ that

$$(7.36) \quad \sum_{0 \leq \ell \leq \ell_0} 2^{2\ell} |V_\ell(y)| \leq 2.$$

For $1 \leq k \leq 2^{\ell_0}$, consider a fixed point $y_k \in I_{0,k}$. Thus we have

$$\sum_{0 \leq \ell \leq \ell_0, k \leq 2^{\ell_0}} 2^{2\ell} |V_\ell(y_k)| \leq 2^{\ell_0+1}$$

so that

$$(7.37) \quad \sum_{0 \leq \ell \leq \ell_0, k \leq 2^{\ell_0}} 2^\ell \text{card}\{B; B \in \mathcal{B}_\ell, B \subset V_\ell(y_k)\} \leq 2^{2\ell_0+1}.$$

Consider $A \in \mathcal{P}_\ell$ such that $A \subset W_\ell$. By definition of W_ℓ , $A = B \times I$ for $B \in \mathcal{B}_\ell$, I of the type $I_{\ell,p}$, and we have $B \subset V_\ell(y_k)$ whenever $y_k \in I$, which occurs for 2^ℓ values of k . Thus we get, from (7.37)

$$(7.38) \quad \sum_{0 \leq \ell \leq \ell_0} 2^{2\ell} \text{card}\{A \in \mathcal{P}_\ell; A \subset W_\ell\} \leq 2^{2\ell_0+1}.$$

On the other hand, from (7.36), and since $\sum_{k \leq 2^{\ell_0}} |I_{0,k}| = 1$, we have

$$(7.39) \quad \sum_{0 \leq \ell \leq \ell_0, k \leq 2^{\ell_0}} 2^{2\ell} |I_{0,k}| |V_\ell(y_k)| \leq 2.$$

For $A = B \times I \in \mathcal{P}_\ell$, we have

$$|I| = \sum \{|I_{0,k}|; y_k \in I\}.$$

Since $B \subset V_\ell(y_k)$ whenever $y_k \in I$, we get from (7.39) that

$$(7.40) \quad \sum_{0 \leq \ell \leq \ell_0} 2^{2\ell} |W_\ell| \leq 2.$$

The conclusion of Lemma 7.10 now follows from (7.35), (7.38), (7.40), since $2^{2\ell_0} \frac{\log n}{n} \leq K$. $\square$

We now prepare for the proof of Lemma 7.11. Consider a continuous function h on $[0, 1]$. We denote by $J(h, a)$ the number of down-crossings of h from a to $a/2$, i.e.,

$$J(h, a) = \sup\{q; \exists x_1 < y_1 < x_2 < \cdots < x_q < y_q; \\ \forall i \leq q, h(x_i) \geq a, h(y_i) \leq a/2\}.$$

Lemma 7.12. *Consider a differentiable function h on $[0, 1]$. Then*

$$\sum_{k \in \mathbb{Z}} 2^k J(h, 2^k) \leq 2 \int_0^1 |h'(x)| dx.$$

Proof. If $J(h, 2^k) = q_k$, one can find disjoint intervals $J_1^k, \dots, J_{q_k}^k$ such that $2^{k-1} < h < 2^k$ in the interior of J_i^k ($i \leq q_k$) and that, moreover, $h(x) = 2^k$ (resp. 2^{k-1}) when x is the left (resp. right) end point of J_i^k . Thus the intervals $(J_i^k)_{k \in \mathbb{Z}, i \leq q_k}$ have disjoint interiors, and $\int_{J_i^k} |h'(x)| dx \geq 2^{k-1}$, so that

$$\sum_{k \in \mathbb{Z}} 2^{k-1} q_k \leq \int_0^1 |h'(x)| dx. \quad \square$$

We denote by $\mathcal{P}_{\ell,k}$ the set of rectangles of $\mathcal{P}_\ell$ included in $[0, 1] \times I_{\ell,k}$. (Thus, these rectangles are the type $C \times I_{\ell,k}$, where C has length $2^{-\ell-\ell_0}$.) For $y \in [0, 1]$, we denote by f_y the function given by $f_y(x) = f(x, y)$.

Lemma 7.13. *We fix ℓ and k . Consider the set $B_{\ell,k}$, that is the union of the rectangles $A \in \mathcal{P}_{\ell,k}$ for which $m_A(f) \geq 2^{\ell-\ell_0+3}$. Then we can find a set $B'_{\ell,k} \supset B_{\ell,k}$ with the following properties.*

(7.41) $B'_{\ell,k}$ is a union of rectangles of $\mathcal{P}_{\ell,k}$.

(7.42) If $A \in \mathcal{P}_{\ell,k}$, $A \subset B'_{\ell,k}$, then $m_A(f) \geq 2^{\ell-\ell_0}$.

(7.43) $B'_{\ell,k}$ is the union of $q_{\ell,k}$ disjoint rectangles (each of which being the union of consecutive rectangles of $\mathcal{P}_{\ell,k}$), where for all $y \in I_{\ell,k}$ we have

$$q_{\ell,k} \leq 1 + J(f_y, 2^{\ell-\ell_0+2}).$$

Proof. Consider the set W that is the union of the sets $A \in \mathcal{P}_{\ell,k}$ for which $m_A(f) \geq 2^{\ell-\ell_0}$. We define $B'_{\ell,k}$ as the union of the connex components of W that meet $B_{\ell,k}$. Denote by $q_{\ell,k}$ the number of these components. All we have to show is that

$$q_{\ell,k} \leq 1 + J(f_y, 2^{\ell-\ell_0+2})$$

whenever $y \in I_{k,\ell}$. We recall that $|f_y(x) - f(x, y')| \leq 2^{\ell-\ell_0}$ whenever $y' \in I_{k,\ell}$. Thus if $m_A(f) \geq 2^{\ell-\ell_0+3}$, we must have

$$f_y(x) \geq 2^{\ell-\ell_0+3} - 2^{\ell-\ell_0} \geq 2^{\ell-\ell_0+2}$$

for some x with $(x, y) \in A$, while if $m_A(f) \leq 2^{\ell-\ell_0}$, a similar argument shows that we must have $f_y(x) \leq 2^{\ell-\ell_0+1}$ for some x with $(x, y) \in A$. The result follows easily. $\square$

We now observe that $B_\ell = \bigcup_k B_{\ell,k}$. We set $B'_\ell = \bigcup_k B'_{\ell,k}$, and we prove Lemma 7.11.

Using (7.35) we have

$$\begin{aligned}
 \text{card}\{i \leq n; U_i \in B_\ell\} &\leq \text{card}\{i \leq n; U_i \in B'_\ell\} \\
 &= \sum_{k \leq 2^{\ell_0 - \ell}} \text{card}\{i \leq n; U_i \in B'_{\ell, k}\} \\
 &\leq \sum_k rn \left(|B'_{\ell, k}| + q_{\ell, k} \frac{\log n}{n} \right) \\
 &\leq rn \left(|B'_\ell| + \left(\sum_{k \leq 2^{\ell_0 - \ell}} q_{\ell, k} \right) \frac{\log n}{n} \right).
 \end{aligned}$$

By Lemma 7.5 and (7.42), we have

$$\sum_{\ell \leq \ell_0} 2^{2(\ell - \ell_0)} |B'_\ell| \leq K \|f\|_2^2.$$

Thus it remains to show that

$$\sum_{\ell \leq \ell_0} 2^{2(\ell - \ell_0)} \sum_{k \leq 2^{\ell_0 - \ell}} q_{\ell, k} \leq K.$$

For $p \leq 2^{\ell_0}$, consider $y_p \in I_{0, p}$. By Lemma 7.12 and since $\int_0^1 |f'_{y_p}(s)| ds \leq 1$, we have

$$\sum_{0 \leq \ell \leq \ell_0} 2^{\ell - \ell_0} J(f_{y_p}, 2^{\ell - \ell_0}) \leq 2$$

so that, by summation over $p \leq 2^{\ell_0}$, we have

$$(7.44) \quad \sum_{p \leq 2^{\ell_0}} \sum_{0 \leq \ell \leq \ell_0} 2^{\ell - 2\ell_0} J(f_{y_p}, 2^{\ell - \ell_0}) \leq 2.$$

For $\ell \leq \ell_0$, $k \leq 2^{\ell_0 - \ell}$, we have by (7.43) that

$$q_{\ell, k} \leq 1 + J(f_{y_p}, 2^{\ell - \ell_0})$$

whenever $y_p \in I_{\ell, k}$, which occurs for 2^ℓ values of p .

Thus, from (7.44) we get

$$\sum_{0 \leq \ell \leq \ell_0} 2^{2\ell - 2\ell_0} \sum_{k \leq 2^{\ell_0 - \ell}} (q_{\ell, k} - 1) \leq 2$$

so that

$$\sum_{0 \leq \ell \leq \ell_0} 2^{2\ell - 2\ell_0} \sum_{k \leq 2^{\ell_0 - \ell}} q_{\ell, k} \leq K. \quad \square$$

8. SHOR'S MATCHING THEOREM

We develop the tools that will enable us to prove that theorem. It will be convenient to “discretise” the problem, so, for a while, we will study a discrete

problem. Consider $p \geq 0$ and the "grid" $\{1, \dots, p\}^2$. For $1 \leq k, \ell \leq p$ we consider an integer $n(k, \ell)$. We assume that for some integer m_0 we have

$$(8.1) \quad m_0 \leq n(k, \ell) \leq 2m_0.$$

We set $n = \sum_{k, \ell \leq p} n(k, \ell)$.

For $i \leq n$, we consider points $Y_i = (k(i), \ell(i))$, $Z_i = (k'(i), \ell'(i))$ on the grid. We assume that exactly $n(k, \ell)$ points Z_i are located at (k, ℓ) . We want to match the points Y_i with the points Z_i ; i.e., we try to find a one-to-one map π from $\{1, \dots, n\}$ to $\{1, \dots, n\}$ such that

$$(8.2) \quad \forall i \leq n, \quad |k(i) - k'(\pi(i))| \leq 1$$

and

$$W(\pi) = \sum_{i \leq n} \varphi(\ell(i) - \ell'(\pi(i)))$$

is minimum, where φ is a certain weight function to be specified later. We assume $\varphi(0) = 0$, φ symmetric, convex, and strictly increasing on $\mathbb{R}^+$. It follows from Proposition 4.5 that

$$(8.3) \quad M_1 = \inf W(\pi) = \sup \sum_{i \leq n} (w'_i - w_i)$$

where the supremum is taken over all the n -tuples $(w_i)_{i \leq n}$, $(w'_i)_{i \leq n}$ that satisfy

$$(8.4) \quad \forall i, j \leq n, \quad |k(i) - k'(j)| \leq 1 \Rightarrow w'_i \leq w_j + \varphi(\ell(i) - \ell'(j)).$$

We fix a family $(w_i)_{i \leq n}$, $(w'_i)_{i \leq n}$ for which the supremum is obtained in (8.3). Consider the function h' given by

$$h'(k, \ell) = \inf \{w_i + \varphi(\ell - \ell'(i)); |k - k'(i)| \leq 1, i \leq n\}.$$

By (8.4), we have

$$M_1 \leq \sum_{i \leq n} h'(Y_i) - \sum_{i \leq n} w_i.$$

Suppose now that for each point Z_i we replace its weight w_i by the average of the weights of the points having the same location. We do not change the value of $\sum_{i \leq n} w_i$ while h' increases. Thus we can write

$$M_1 \leq \sum_{i \leq n} h(Y_i) - \sum_{k, \ell \leq p} n(k, \ell) u(k, \ell)$$

where

$$(8.5) \quad h(k, \ell) = \inf \{u(k', \ell') + \varphi(\ell - \ell'); |k - k'| \leq 1, 1 \leq \ell' \leq p\}.$$

We set

$$\theta(x) = x \max(\varphi^{-1}(x) - 2, 0).$$

An essential step is the following.

Proposition 8.1. *We have*

$$\begin{aligned} m_0 \left(\sum_{k < p, \ell \leq p} |h(k+1, \ell) - h(k, \ell)| + \sum_{k \leq p, \ell < p} \theta(|h(k, \ell+1) - h(k, \ell)|) \right) \\ \leq 18 \left(\sum_{k, \ell \leq p} n(k, \ell) u(k, \ell) - \sum_{k, \ell \leq p} n(k, \ell) h(k, \ell) \right). \end{aligned}$$

We set $v(k, \ell) = \min(u(k', \ell); |k - k'| \leq 1)$. Thus from (8.5), we have

$$(8.6) \quad h(k, \ell) = \inf\{v(k, \ell') + \varphi(\ell - \ell'); 1 \leq \ell' \leq p\}.$$

Lemma 8.2.

$$\begin{aligned} m_0 \left(\sum_{k < p, \ell \leq p} |v(k+1, \ell) - v(k, \ell)| \right) \\ (8.7) \quad \leq 6 \left(\sum_{k, \ell \leq p} n(k, \ell) u(k, \ell) - \sum_{k, \ell \leq p} n(k, \ell) v(k, \ell) \right). \end{aligned}$$

Proof. We observe the elementary inequality

$$2 \min(a, b) \leq a + b - |b - a|$$

so that

$$\begin{aligned} v(k, \ell) + v(k+1, \ell) &\leq 2 \min(u(k, \ell), u(k+1, \ell)) \\ &\leq u(k, \ell) + u(k+1, \ell) - |u(k+1, \ell) - u(k, \ell)|. \end{aligned}$$

We sum these inequalities for $k < p, \ell \leq p$ and get

$$\sum_{k, \ell \leq p} a(k) v(k, \ell) \leq \sum_{k, \ell \leq p} a(k) u(k, \ell) - \sum_{k < p, \ell \leq p} |u(k+1, \ell) - u(k, \ell)|$$

where $a(1) = a(p) = 1, a(k) = 2$ for $1 < k < p$. Since $v(1, \ell) \leq u(1, \ell), v(k, \ell) \leq u(k, \ell)$, we get

$$\frac{1}{2} \sum_{k < p, \ell \leq p} |u(k+1, \ell) - u(k, \ell)| \leq \sum_{k, \ell \leq p} u(k, \ell) - \sum_{k, \ell \leq p} v(k, \ell).$$

Thus

$$\begin{aligned} \frac{m_0}{2} \sum_{k < p, \ell \leq p} |u(k+1, \ell) - u(k, \ell)| \\ \leq m_0 \sum_{k, \ell \leq p} u(k, \ell) - m_0 \sum_{k, \ell \leq p} v(k, \ell) \\ \leq \sum_{k, \ell \leq p} n(k, \ell) u(k, \ell) - \sum_{k, \ell \leq p} n(k, \ell) v(k, \ell). \end{aligned}$$

Now we use the elementary inequality

$$|\min(a, b, c) - \min(a', b', c')| \leq |a - a'| + |b - b'| + |c - c'|$$

for $a = u(k-1, \ell), a' = b = u(k, \ell), b' = c = u(k+1, \ell), c' = u(k+2, \ell)$ to get by summation that

$$\sum_{k < p, \ell \leq p} |v(k+1, \ell) - v(k, \ell)| \leq 3 \sum_{k < p, \ell \leq p} |u(k+1, \ell) - u(k, \ell)|. \quad \square$$

Consider now two sequences $(v_\ell)_{\ell \leq p}$, $(v'_\ell)_{\ell \leq p}$ and define, for $\ell \leq p$

$$g(\ell) = \min(v_{\ell'} + \varphi(\ell - \ell'); 1 \leq \ell' \leq p),$$

$$g'(\ell) = \min(v'_{\ell'} + \varphi(\ell - \ell'); 1 \leq \ell' \leq p).$$

Lemma 8.3. *We have*

$$\sum_{\ell \leq p} |g(\ell) - g'(\ell)| \leq \sum_{\ell \leq p} (v_\ell + v'_\ell - g(\ell) - g'(\ell) + |v_\ell - v'_\ell|).$$

Proof. Observe that $v_\ell \geq g(\ell)$, $v'_\ell \geq g'(\ell)$. If $g'(\ell) \geq g(\ell)$, we have

$$g'(\ell) - g(\ell) \leq v'_\ell - g(\ell) = v'_\ell - v_\ell + v_\ell - g(\ell) \leq |v'_\ell - v_\ell| + v_\ell - g(\ell).$$

A similar argument in the case where $g(\ell) \geq g'(\ell)$ shows that

$$|g(\ell) - g'(\ell)| \leq |v'_\ell - v_\ell| + v'_\ell - g'(\ell).$$

The result follows by summation over $\ell \leq p$. $\square$

Lemma 8.4. *We have*

$$\sum_{\ell < p} \theta(|g(\ell+1) - g(\ell)|) \leq 16 \left(\sum_{\ell \leq p} v_\ell - g(\ell) \right).$$

Proof. Consider the set $A \subset \{1, \dots, p\}$ of the integers ℓ for which $g(\ell+1) \geq g(\ell)$. For $\ell \in A$, we define $x_\ell = \ell + 1 - \varphi^{-1}(g(\ell+1) - g(\ell))$.

Consider $1 \leq \ell' \leq p$ such that $g(\ell) = v_{\ell'} + \varphi(\ell - \ell')$. Since

$$g(\ell) \leq g(\ell+1) \leq v_{\ell'} + \varphi(\ell+1 - \ell'),$$

we see that $\varphi(\ell - \ell') \leq \varphi(\ell+1 - \ell')$ so that $\ell' \leq \ell$. Also,

$$\varphi(\ell+1 - x_\ell) = g(\ell+1) - g(\ell) \leq \varphi(\ell+1 - \ell') - \varphi(\ell - \ell') \leq \varphi(\ell+1 - \ell')$$

so that $x_\ell \geq \ell' \geq 1$. Also, for $\ell' \leq m \leq \ell$ we have

$$g(m) \leq v_{\ell'} + \varphi(m - \ell') \leq v_{\ell'} + \varphi(\ell - \ell') = g(\ell).$$

Thus

$$(8.8) \quad x_\ell \leq m \leq \ell \Rightarrow g(m) \leq g(\ell).$$

Consider, for $\ell \in A$, the domain

$$D_\ell = \{(m, y); x_\ell \leq m \leq \ell; g(\ell) \leq y < g(\ell+1) - \varphi(\ell+1 - m)\}.$$

First we observe that for $\ell, q \in A$, $\ell \neq q$, we have $D_\ell \cap D_q = \emptyset$. Indeed, suppose for definiteness that $q < \ell$. If $q < x_\ell$, we certainly have $D_\ell \cap D_q = \emptyset$. If $q \geq x_\ell$, by (8.8) and since $q+1 \leq \ell$, we have $g(q+1) \leq g(\ell)$. Now, if $(m, y) \in D_\ell$, we have $y \geq g(\ell)$, while if $(m, y) \in D_q$, we have

$$y < g(q+1) - \varphi(q+1 - m) \leq g(q+1) \leq g(\ell).$$

Consider now the measure μ on $\mathbb{R}^2$, which is the sum of the linear measures on each of the vertical lines through the points $(0, \ell)$, $1 \leq \ell \leq p$. Thus

$$\mu(D_\ell) = \sum_{x_\ell \leq m \leq \ell} g(\ell+1) - g(\ell) - \varphi(\ell+1 - m).$$

We observe that the function $\varphi(z)/z$ increases for $z > 0$. Thus, for $m \geq x_\ell$,

$$\frac{\varphi(\ell + 1 - m)}{\ell + 1 - m} \leq \frac{\varphi(\ell + 1 - x_\ell)}{\ell + 1 - x_\ell}$$

so that

$$\varphi(\ell + 1 - m) \leq \frac{\ell + 1 - m}{\ell + 1 - x_\ell} (g(\ell + 1) - g(\ell))$$

and

$$g(\ell + 1) - g(\ell) - \varphi(\ell + 1 - m) \geq (g(\ell + 1) - g(\ell)) \left(\frac{m - x_\ell}{\ell + 1 - x_\ell} \right).$$

We observe now that

$$m \geq \frac{1}{2}(x_\ell + \ell + 1) \Rightarrow \frac{m - x_\ell}{\ell + 1 - x_\ell} \geq \frac{1}{2}.$$

Thus, provided $x_\ell \leq \ell - 1$, there are at least $(\ell + 1 - x_\ell)/4$ values of $m \leq \ell$ for which $(m - x_\ell)/(\ell + 1 - x_\ell) \geq 1/2$. Thus, in that case

$$\begin{aligned} \mu(D_\ell) &\geq \frac{1}{8}(\ell + 1 - x_\ell)(g(\ell + 1) - g(\ell)) \\ &= \frac{1}{8}\varphi^{-1}(g(\ell + 1) - g(\ell))(g(\ell + 1) - g(\ell)) \\ &\geq \frac{1}{8}\theta(g(\ell + 1) - g(\ell)). \end{aligned}$$

In the case where $x_\ell \geq \ell - 1$, we have $\varphi^{-1}(g(\ell + 1) - g(\ell)) \leq 2$, so

$$\mu(D_\ell) \geq 0 = \frac{1}{8}\theta(g(\ell + 1) - g(\ell)).$$

Then, in any case, we have

$$(8.9) \quad \mu(D_\ell) \geq \frac{1}{8}\theta(g(\ell + 1) - g(\ell)).$$

Consider now the region

$$(8.10) \quad R = \{(\ell, y); 1 \leq \ell \leq p, g(\ell) \leq y \leq v_\ell\}.$$

For $(m, y) \in D_\ell$, it follows from (8.7) that $g(\ell) \leq y$. Also, we have, by definition of g

$$v_m + \varphi(\ell + 1 - m) \geq g(\ell + 1)$$

so that $y \leq g(\ell + 1) - \varphi(\ell + 1 - m) \leq v_m$. This shows that $D_\ell \subset R$. Since the sets $(D_\ell)_{\ell \in A}$ are disjoint, we get $\sum_{\ell \in A} \mu(D_\ell) \leq \mu(R)$, so that, by (8.9)

$$\sum_{\ell \in A} \frac{1}{8}\theta(g(\ell + 1) - g(\ell)) \leq \mu(R) = \sum_{\ell \leq p} v_\ell - g(\ell).$$

Using the same argument for the set $B = \{\ell \leq p; g(\ell + 1) < g(\ell)\}$ completes the proof. $\square$

We now prove Proposition 8.1. Given $1 \leq k < p$, we use Lemma 8.3 for $v_\ell = v(k, \ell)$, $v'_\ell = v(k + 1, \ell)$. Thus $g(\ell) = h(k, \ell)$, $g'(\ell) = h(k + 1, \ell)$.

Summing the corresponding inequalities for $k < p$ yields

$$(8.11) \quad \sum_{k < p, \ell \leq p} |h(k+1, \ell) - h(k, \ell)| \\ \leq 2 \sum_{k, \ell \leq p} (v(k, \ell) - h(k, \ell)) + \sum_{k < p, \ell \leq p} |v(k+1, \ell) - v(k, \ell)|.$$

From Lemma 8.4, we get, by summation over $k \leq p$

$$(8.12) \quad \sum_{k \leq p, \ell < p} \theta(|h(k, \ell+1) - h(k, \ell)|) \leq 16 \left(\sum_{k, \ell \leq p} v(k, \ell) - h(k, \ell) \right).$$

We denote by S the sum of the left-hand sides of (8.11), (8.12). Thus, using (8.7)

$$\begin{aligned} m_0 S &\leq 18m_0 \left(\sum_{k, \ell \leq p} v(k, \ell) - h(k, \ell) \right) \\ &\quad + m_0 \left(\sum_{k < p, \ell \leq p} |v(k+1, \ell) - v(k, \ell)| \right) \\ &\leq 18 \sum_{k, \ell \leq p} n(k, \ell)(v(k, \ell) - h(k, \ell)) \\ &\quad + 18 \sum_{k, \ell \leq p} n(k, \ell)(u(k, \ell) - v(k, \ell)) \\ &= 18 \sum_{k, \ell \leq p} n(k, \ell)(u(k, \ell) - h(k, \ell)). \quad \square \end{aligned}$$

We set $\xi(x) = x(\varphi^{-1}(x) + 1)$ and

$$a(\varphi) = \sup_{x \geq 0} \xi(x) - 2\theta(x) = \sup_{x \geq 0} (x(\varphi^{-1}(x) + 1) - 2\varphi^{-1}(x)(x - 2)^+)$$

so that $a(\varphi) < \infty$ and

$$(8.13) \quad \xi(x) \leq 2\theta(x) + a(\varphi).$$

For a function g on $\{1, \dots, p\}^2$, we set

$$E(g) = \frac{1}{n} \sum_{k, \ell \leq p} n(k, \ell)g(k, \ell).$$

Consider the class $\mathcal{X}_\varphi$ of functions g from $\{1, \dots, p\}^2$ to $\mathbb{R}$ that satisfy

$$(8.14) \quad E(g) = 0,$$

$$(8.15) \quad \sum_{k < p, \ell \leq p} |g(k+1, \ell) - g(k, \ell)| \\ + \sum_{k \leq p, \ell < p} \xi(|g(k, \ell+1) - g(k, \ell)|) \leq 2p^2 a(\varphi).$$

We can now state and prove another important step.

Proposition 8.5. *Suppose that*

$$(8.16) \quad \sup_{g \in \mathcal{H}_\varphi} \left| \sum_{i \leq n} g(Y_i) \right| \leq \frac{na(\varphi)}{72}.$$

Then $M_1 \leq \frac{na(\varphi)}{72}$.

Proof. Consider h given by (8.5). Then

$$(8.17) \quad \begin{aligned} M_1 &\leq \sum_{i \leq n} h(Y_i) - \sum_{k, \ell \leq p} n(k, \ell) u(k, \ell) \\ &\leq \left| \sum_{i \leq n} (h(Y_i) - Eh) \right| + \sum_{k, \ell \leq p} n(k, \ell) (h(k, \ell) - u(k, \ell)). \end{aligned}$$

Note that the second term is negative. Observe that the function $g = h - Eh$ satisfies (8.14). If $g \in \mathcal{H}_\varphi$, by (8.16) we have $M_1 \leq \frac{na(\varphi)}{72}$, and there is nothing to prove. If $g \notin \mathcal{H}_\varphi$, there is a smallest number $\lambda > 1$ such that $g' = g/\lambda \in \mathcal{H}_\varphi$. We then have

$$\sum_{k < p, \ell \leq p} |g'(k+1, \ell) - g'(k, \ell)| + \sum_{k \leq p, \ell < p} \xi(|g'(k, \ell+1) - g'(k, \ell)|) = 2p^2 a(\varphi).$$

By (8.13) we get

$$p^2 a(\varphi) \leq \sum_{k < p, \ell \leq p} |g'(k+1, \ell) - g'(k, \ell)| + 2 \sum_{k \leq p, \ell < p} \theta(|g'(k, \ell+1) - g'(k, \ell)|).$$

We observe that $\theta(\frac{x}{\lambda}) \leq \frac{\theta(x)}{\lambda}$ for $\lambda \geq 1$, by definition of θ . Thus, remembering that $g = h - Eh$, we have

$$\lambda p^2 a(\varphi) \leq \sum_{k < p, \ell \leq p} |h(k+1, \ell) - h(k, \ell)| + 2 \sum_{k \leq p, \ell < p} \theta(|h(k, \ell+1) - h(k, \ell)|).$$

We now use Proposition 8.1 to obtain, since $n = \sum n(k, \ell) \leq 2m_0 p^2$,

$$\frac{\lambda na(\varphi)}{2} \leq m_0 \lambda p^2 a(\varphi) \leq 36 \sum_{k, \ell \leq p} n(k, \ell) (u(k, \ell) - h(k, \ell)).$$

Now by (8.16) we have, since $g/\lambda \in \mathcal{H}_\varphi$,

$$\left| \sum_{i \leq n} \frac{g(Y_i)}{\lambda} \right| \leq \frac{na(\varphi)}{72}$$

so that

$$\left| \sum_{i \leq n} (h(Y_i) - E(h)) \right| \leq \frac{\lambda na(\varphi)}{72}.$$

Thus by (8.17), we have $M_1 \leq 0 \leq \frac{na(\varphi)}{72}$. $\square$

Consider the probability P on $\{1, \dots, p\}^2$ given by $P(\{(k, \ell)\}) = \frac{n(k, \ell)}{n}$. We assume now that the points $(Y_i)_{i \leq n}$ are random distributed independently according to P .

We consider the class

$$\mathcal{E}_\varphi = \left\{ f : [0, 1]^2 \rightarrow \mathbb{R}; \left| \iint f d\lambda \right| \leq 1, \iint \xi \left(\frac{\partial f}{\partial y} \right) d\lambda \leq 1 \right\}.$$

Proposition 8.6. *There exists a universal constant K with the following property. Suppose that $m_0 \geq K\gamma_{1/2,2}(\mathcal{E}_\varphi)^2 \log n$. Then*

$$P \left(\sup_{g \in \mathcal{H}_\varphi} \left| \sum_{i \leq n} g(Y_i) \right| \geq \frac{na(\varphi)}{72} \right) \leq K \exp \left(-\frac{m_0}{K} \right).$$

Before we start the proof, we need more technical steps.

Lemma 8.7. *Consider $g \in \mathcal{H}_\varphi$. Then*

$$(8.18) \quad \sum_{k, \ell \leq p} |g(k, \ell)| \leq 6p^3 a(\varphi).$$

Proof. Consider $q \leq p$. Set

$$a_q = \sum_{k < p} |g(k+1, q) - g(k, q)|.$$

Thus for $k, m \leq p$, we have

$$|g(k, q) - g(m, q)| \leq a_q.$$

Thus

$$\begin{aligned} & \left| g(k, q) \left(\sum_{m \leq p} n(m, q) \right) - \left(\sum_{m \leq p} n(m, q) g(m, q) \right) \right| \\ & \leq a_q \left(\sum_{m \leq p} n(m, q) \right) \leq 2pm_0 a_q \leq \frac{2n}{p} a_q. \end{aligned}$$

Summing over q and using the fact that $E(g) = 0$, we now have

$$(8.19) \quad \left| \sum_{q, m \leq p} g(k, q) n(m, q) \right| \leq 2 \frac{n}{p} \sum_{q \leq p} a_q.$$

We set now

$$b_k = \sum_{q < p} |g(k, q+1) - g(k, q)|$$

so that

$$|g(k, q) - g(k, \ell)| \leq b_k$$

for all $q, \ell \leq p$. Thus from (8.19) we get

$$(8.20) \quad n|g(k, \ell)| \leq nb_k + \frac{2n}{p} \sum_{q \leq p} a_q.$$

Since $g \in \mathcal{H}_\varphi$ and since $\xi(x) \geq x$, we have

$$\sum_{k \leq p} b_k \leq 2p^2 a(\varphi), \quad \sum_{q \leq p} a_q \leq 2p^2 a(\varphi).$$

Thus the result from (8.20), by summation over $k, \ell \leq p$. $\square$

Lemma 8.8. *There exists a linear map H that associates to each function h on $\{1, \dots, p\}^2$ a function $H = H(h)$ on $[0, 1]^2$, such that the following hold.*

$$(8.21) \quad \iint |H| d\lambda \leq \frac{4}{p^3} \sum_{k, \ell \leq p} |h(k, \ell)|.$$

$$(8.22) \quad \iint \left| \frac{\partial H}{\partial x} \right| d\lambda \leq \frac{4}{p^2} \sum_{k < p, \ell \leq p} |h(k+1, \ell) - h(k, \ell)|.$$

(8.23) *For all convex functions η*

$$\iint \eta \left(\left| \frac{\partial H}{\partial y} \right| \right) d\lambda \leq \frac{4}{p^2} \sum_{k \leq p, \ell < p} \eta(|h(k, \ell+1) - h(k, \ell)|).$$

$$(8.24) \quad \|H\|_2 \geq \frac{1}{8p} \|h\|_{L^2(P)}.$$

Proof. *Step 1.* We divide $[0, 1]$ in $2p-1$ consecutive intervals $(I_\ell)_{1 \leq \ell \leq 2p-1}$ of length $1/(2p-1)$. To a function g on $\{1, \dots, p\}$, we associate the function $G(g)$ that is affine on each interval I_ℓ and has constant value $g(\ell)$ on each interval $I_{2\ell-1}$. Observe that if η is any convex function, we thus have

$$(8.25) \quad \eta(G(g)(x)) \leq G(\eta \circ g)(x).$$

It is simple to see, by crude estimates, that

$$(8.26) \quad \int_0^1 |G(g)(x)| dx \leq \frac{2}{p} \sum_{\ell \leq p} g(\ell).$$

$$(8.27) \quad \int_0^1 (G(g)(x))^2 dx \geq \frac{1}{2p} \sum_{\ell \leq p} g(\ell)^2.$$

(8.28) *For any convex increasing function η on $\mathbb{R}^+$*

$$\int_0^1 \eta \left(\frac{|G(g)'(x)|}{2p} \right) dx \leq \frac{1}{p} \sum_{\ell < p} \eta(|g(\ell+1) - g(\ell)|).$$

Step 2. Consider now a function h on $\{1, \dots, p\}^2$. Given $x \in [0, 1]$, consider the function f_x on $\{1, \dots, p\}$ given by $f_x(\ell) = G(g_\ell)(x)$, where $g_\ell(k) = h(k, \ell)$.

We define $\bar{H} = \bar{H}(h)$ by

$$\bar{H}(x, y) = G(f_x)(y),$$

and we set $H = \frac{1}{2p} \bar{H}$.

By (8.26), we have

$$\int_0^1 |H(x, y)| dy \leq \frac{2}{p} \sum_{\ell \leq p} |f_x(\ell)| = \frac{2}{p} \sum_{\ell \leq p} |G(g_\ell)(x)|.$$

Integrating in x and using (8.26) again gives

$$\iint |H| d\lambda \leq \frac{4}{p^2} \sum_{k, \ell \leq p} |g_\ell(k)| = \frac{4}{p^2} \sum_{k, \ell \leq p} |h(k, \ell)|.$$

This proves (8.21). The same iteration argument now using (8.27) yields

$$\begin{aligned} \|H\|_2 &\geq \frac{1}{2} \left(\frac{1}{p^2} \sum_{k, \ell \leq p} h(k, \ell)^2 \right)^{1/2} \\ &\geq \frac{1}{2} \left(\frac{1}{2n} \sum_{k, \ell \leq p} n(k, \ell) h(k, \ell)^2 \right)^{1/2} \geq \frac{1}{4} \|h\|_{L^2(P)} \end{aligned}$$

from which (8.24) follows. We observe, since the map G is linear, that $\frac{d}{dx} G(f_x) = G(\partial f_x / \partial x)$ so that, by (8.25) we have

$$\left| \frac{\partial \bar{H}}{\partial x}(x, y) \right| \leq G \left(\left| \frac{\partial f_x}{\partial x} \right| \right)(y).$$

Using (8.26), we have

$$\int_0^1 \left| \frac{\partial \bar{H}}{\partial x}(x, y) \right| dy \leq \frac{2}{p} \sum_{\ell \leq p} \left| \frac{\partial f_x}{\partial x}(\ell) \right| = \frac{2}{p} \sum_{\ell \leq p} |G(g_\ell)'(x)|.$$

Integrating with respect to x and using (8.28) with $\eta(x) = x$ proves (8.22), since $H = \frac{1}{2p} \bar{H}$.

From (8.28) we have

$$\begin{aligned} \int_0^1 \eta \left(\left| \frac{\partial H}{\partial y} \right|(x, y) \right) dy &= \int_0^2 \eta(|G(f_x)'(y)|) dy \leq \frac{1}{p} \sum_{\ell < p} \eta(|f_x(\ell+1) - f_x(\ell)|) \\ &= \frac{1}{p} \sum_{\ell < p} \eta(|G(g_{\ell+1} - g_\ell)(x)|) \leq \frac{1}{p} \sum_{\ell < p} G(\eta(|g_{\ell+1} - g_\ell|))(x) \end{aligned}$$

where we have used (8.25) (for the convex function $\eta(|\cdot|)$) in the last inequality. Integrating in x and using (8.26) proves (8.23). $\square$

Lemma 8.9. Consider $h \in \mathcal{H}_\varphi$. Then $\|h\|_{L^2(P)} \leq Kpa(\varphi)$ where K is universal.

Proof. If we use Lemma 8.8 (with $\eta(x) = x$), Lemma 8.7, and the definition of $\mathcal{H}_\varphi$ and remember that $\xi(x) \geq x$, we see that the function $f = H(h)/a(\varphi)$ satisfies

$$\begin{aligned} \left| \iint f d\lambda \right| &\leq \iint |f| d\lambda \leq 24, \\ \iint \left| \frac{\partial f}{\partial x} \right| d\lambda &\leq 8, \quad \iint \left| \frac{\partial f}{\partial y} \right| d\lambda \leq 8. \end{aligned}$$

We appeal to Sobolev's inequality (e.g., as in [O]) to conclude that $\|f - \iint f d\lambda\|_2 \leq K$, so that $\|f\|_2 \leq K$. The result then follows from (8.24).

We now start to prove Proposition 8.6.

Step 1. We show that (when $\mathcal{H}_\varphi$ is provided with the metric induced by $L^2(P)$) we have $\gamma_{1/2,2}(\mathcal{H}_\varphi) \leq Kp\gamma_{1/2,2}(\mathcal{E}_\varphi)a(\varphi)$. Indeed, let us use Lemma 8.8 with $\eta = \xi$. If we recall that $\eta(x/a) \leq \eta(x)/a$ for $a \geq 1$ and use Lemma 8.7, we see that for $h \in \mathcal{H}_\varphi$ we have $T(h) := T(h)/Ka(\varphi) \in \mathcal{E}_\varphi$. Consider the map T^{-1} for $T(\mathcal{H}_\varphi)$ onto $\mathcal{H}_\varphi$. By (8.24) it satisfies

$$\|T^{-1}(H) - T^{-1}(H')\|_{L^2(P)} \leq Kpa(\varphi)\|H - H'\|_2.$$

Thus the result follows from Corollary 2.8.

Step 2. We set

$$r = \max\{\text{card}\{i \leq n; Y_i = (k, \ell)\}; k, \ell \leq p\}.$$

It follows from the proof of Lemma 4.4 that

$$(8.29) \quad P(r \geq 6m_0) \leq \exp(-m_0).$$

We assume $r \leq 6m_0$ and consider a Bernoulli sequence $(\epsilon_i)_{i \leq n}$. We estimate

$$P_1 =: P_\epsilon \left(\sup_{g \in \mathcal{H}_\varphi} \left| \sum_{i \leq n} \epsilon_i g(Y_i) \right| \geq \frac{na(\varphi)}{300} \right).$$

For any function g on $\{1, \dots, p\}^2$, we consider

$$\|g\|_Y = \left(\sum_{i \leq n} g(Y_i)^2 \right)^{1/2}.$$

Thus

$$\|g\|_Y^2 \leq 6m_0 \sum_{k, \ell \leq p} g(k, \ell)^2 \leq 6n \sum_{k, \ell \leq p} \frac{n(k, \ell)}{n} g(k, \ell)^2$$

and hence

$$\|g\|_Y \leq 3\sqrt{n}\|g\|_{L^2(P)}.$$

By Lemma 8.9, we thus have

$$(8.30) \quad g \in \mathcal{H}_\varphi \Rightarrow \|g\|_Y \leq Kpa(\varphi)\sqrt{n}.$$

We consider the distance d_Y on $\mathcal{H}_\varphi$ given by $d_Y(g, g') = \|g - g'\|_Y$.

Consider a subset S of $\mathcal{H}_\varphi$ that is maximal with respect to the condition $g, g' \in S \Rightarrow \|g - g'\|_Y \geq pa(\varphi)/600\sqrt{n}$. It follows from Lemma 4.2 and (8.30) that

$$(8.31) \quad \gamma_{1/2,1}(S, d_Y) \leq Kpa(\varphi)\gamma_{1/2,2}(\mathcal{E}_\varphi)\sqrt{n(1 + \log n)}.$$

For $g \in \mathcal{H}_\varphi$, by definition of S , we can find $g' \in S$ with $\|g - g'\|_Y \leq pa(\varphi)/600\sqrt{n}$. By Cauchy-Schwarz, we have

$$\left| \sum_{i \leq n} \epsilon_i (g - g')(Y_i) \right| \leq \sqrt{n}\|g - g'\|_Y \leq pa(\varphi)/600 \leq na(\varphi)/600.$$

Thus we have

$$P_1 \leq P_2 =: P_\epsilon \left(\sup_{g \in S} \left| \sum_{i \leq n} \epsilon_i g(Y_i) \right| \geq \frac{na(\varphi)}{600} \right).$$

By (8.31) and the majorizing measure bound (Proposition 2.3) we have

$$E =: E_\epsilon \sup_{g \in S} \left| \sum_{i \leq n} \epsilon_i g(Y_i) \right| \leq Kpa(\varphi)\gamma_{1/2,2}(\mathcal{E}_\varphi)\sqrt{n(1+\log n)}.$$

We now appeal to Proposition 5.4 to obtain, using (8.29),

$$Kp\gamma_{1/2,2}(\mathcal{E}_\varphi)\sqrt{n(1+\log n)} \leq n \Rightarrow P_2 \leq 2 \exp - \frac{n}{Kp^2} \leq 2 \exp - \frac{m_0}{K}.$$

Since $p^2 \leq n/m_0$, recalling (8.30), we have shown that

$$m_0 \geq K\gamma_{1/2,2}(\mathcal{E}_\varphi)^2 \log n \Rightarrow P \left(\sup_{g \in \mathcal{H}_\varphi} \left| \sum_{i \leq n} \epsilon_i g(Y_i) \right| \geq \frac{na(\varphi)}{300} \right) \leq 3 \exp - \frac{m_0}{K}.$$

The conclusion now follows from Lemma 5.5. $\square$

We now finish the proof of Theorem 1.5. Consider φ such that $\gamma_{1/2,2}(\mathcal{E}_\varphi) < \infty$ and consider a number L . Set $p = \lfloor \frac{1}{L} \sqrt{\frac{n}{\log n}} \rfloor$ so that

$$\frac{n}{p^2} \geq L^2 \log n.$$

For $1 \leq \ell \leq p$, we set

$$n(\ell) = \lfloor \ell n/p \rfloor - \lfloor (\ell-1)n/p \rfloor.$$

For $1 \leq k \leq p$, we set

$$n(k, \ell) = \lfloor kn(\ell)/p \rfloor - \lfloor (k-1)n(\ell)/p \rfloor.$$

Thus $\sum_{k, \ell \leq p} n(k, \ell) = n$. Also, we have

$$n/p \leq n(\ell) \leq n/p + 1$$

so that

$$n/p^2 \leq n(k, \ell) \leq n/p^2 + 2.$$

Thus, for n large enough, we have $m_0 \leq n(k, \ell) \leq 2m_0$, where $m_0 \geq \frac{1}{2}L^2 \log n$. We divide $[0, 1]^2$ into p horizontal consecutive strips, such that the ℓ th strip has width $n(\ell)/n$. This ℓ th strip is in turn divided into p rectangles $R(k, \ell)$ of lengths $n(k, \ell)/n(\ell)$.

We define the points Y_i as follows: $Y_i = \{k, \ell\}$ if and only if $U_i \in R(k, \ell)$. Thus the points $(Y_i)_{i \leq n}$ are independent, distributed according to P . Consider points $(Z_i)_{i \leq n}$, such that $n(k, \ell)$ of these points are located at (k, ℓ) . From Propositions 8.5 and 8.6, it follows that with probability

$\geq 1 - K \exp -m_0/K$, we can find a permutation π of $\{1, \dots, n\}$ such that, if we set $Y_i = (k(i), \ell(i))$; $Z_i = (k''(i), \ell''(i))$, we have

$$|k(i) - k''(\pi(i))| \leq 1, \\ \sum_{i \leq n} \varphi(|\ell(i) - \ell''(\pi(i))|) \leq \frac{na(\varphi)}{72}.$$

Consider now the points $Y'_i = (k'(i), \ell'(i))$ defined from the points V_i as Y_i is defined from U_i . Since the points V_i are evenly spread, it is clear that we can find a matching π' of the points Y'_i and of the points Z_i such that

$$|k'(i) - k''(\pi'(i))| \leq 1, \quad |\ell'(i) - \ell''(\pi'(i))| \leq 1.$$

Thus we can find a matching (called π again) of the points Y_i and of the points Y'_i such that

$$|k(i) - k'(\pi(i))| \leq 2, \quad \sum_{i \leq n} \varphi(|\ell(i) - \ell'(\pi(i))| - 1) \leq \frac{na(\varphi)}{72}.$$

Now, it is simple to see that, for all i, j ,

$$p|U_{1,i} - V_{1,j}| \leq 2(|k(i) - k'(j)| + 5).$$

A similar inequality holds for the second component. Thus we have

$$|U_{1,i} - V_{1,\pi(i)}| \leq KL \sqrt{\frac{\log n}{n}}, \\ \sum_{i \leq n} \varphi\left(\frac{p}{2}|U_{2,i} - V_{2,\pi(i)}| - 5\right) \leq \frac{na(\varphi)}{72}.$$

To finish the proof, for $\alpha < \frac{1}{4}$, we consider the function

$$\varphi_\alpha(x) = \exp(|x| + 1)^\alpha - e.$$

Calculus shows that this function is convex and that $\xi_\alpha(x) = x(\log(e+x))^{1/\alpha}$ is also convex. It follows from Theorem 6.5, taking $a = 0$ and $b = 1/\alpha$, that $\gamma_{1/2,2}(\mathcal{E}_{\varphi_\alpha}) < \infty$.

Finally, one checks that $a(\varphi_\alpha)$ is bounded independently of $\alpha \leq 2$. This completes the proof, taking $L = KM$ for K large enough. $\square$

The reader has noticed that we have written the proof of Theorem 1.5 so that the condition $\alpha < 1/4$ could be improved just by improving Theorem 6.4.

9. CONVEX BODIES

In this section we denote by $\mathcal{E}$ the class

$$\mathcal{E} = \left\{ f : [0, 1]^2 \rightarrow [-1, 1]; \iint \left(\left| \frac{\partial^2 f}{\partial x^2} \right| + \left| \frac{\partial^2 f}{\partial y^2} \right| + \left| \frac{\partial^2 f}{\partial x \partial y} \right| \right) d\lambda \leq 1 \right\}.$$

The main step in the proof of Theorem 1.6 will be to show that $\gamma_{1,2}(\mathcal{E}) < \infty$. But first, we relate $\mathcal{E}$ with the class of convex functions.

Proposition 9.1. Consider a convex function $f : [0, 1]^2 \rightarrow [-1, 1]$, such that $\|\frac{\partial f}{\partial x}\|_\infty \leq M$, $\|\frac{\partial f}{\partial y}\|_\infty \leq M$. Then $f \in 6M\mathcal{C}$.

Proof. Since f is convex, $\partial^2 f / \partial x^2$ is nonnegative, so that

$$\int_0^1 \left| \frac{\partial^2 f}{\partial x^2}(t, y) \right| dt = \int_0^1 \frac{\partial^2 f}{\partial x^2}(t, y) dt = \frac{\partial f}{\partial x}(1, y) - \frac{\partial f}{\partial x}(0, y) \leq 2M.$$

Thus $\iint |\partial^2 f / \partial x^2| d\lambda \leq 2M$. In a similar way, $\iint |\partial^2 f / \partial y^2| d\lambda \leq 2M$. Now, since f is convex, we have $|\partial^2 f / \partial x \partial y|^2 \leq \partial^2 f / \partial x^2 \cdot \partial^2 f / \partial y^2$, so that, taking the square root and using Cauchy-Schwarz, we have

$$\iint \left| \frac{\partial^2 f}{\partial x \partial y} \right| d\lambda \leq \left(\iint \frac{\partial^2 f}{\partial x^2} d\lambda \iint \frac{\partial^2 f}{\partial y^2} d\lambda \right)^{1/2} \leq 2M. \quad \square$$

If, in the definition of $\mathcal{C}$, the integrability conditions on the second derivatives were replaced by integrability conditions on their squares, we could appeal to Theorem 2.4. But this is not the case, and the method of proof will follow that of Theorem 6.4. We first develop an approximation procedure. The starting point of this procedure is a Sobolev-type inequality. We will again denote by K a universal constant, not necessarily the same at each occurrence.

Proposition 9.2. Consider a rectangle R of sides a, b , such that $\frac{1}{2} \leq \frac{b}{a} \leq 2$. Consider a function $f : R \rightarrow \mathbb{R}$, and set $g = |\partial^2 f / \partial x^2| + |\partial^2 f / \partial y^2| + |\partial^2 f / \partial x \partial y|$. Then we can find an affine function ξ on R such that

$$\|f - \xi\|_{\infty, R} \leq K \iint_R g d\lambda.$$

Proof. For simplicity, we take $R = [0, a] \times [0, b]$.

Step 1. We will show the existence of an affine function ξ that approximates f on the set $\{(x, y) \in R; |x| \leq \frac{2a}{3}\}$. The same argument works to produce an affine function ξ' that approximates f on the set $\{(x, y) \in R; |x| \geq \frac{a}{3}\}$. Observe that ξ' must approximate ξ on all R , so ξ must approximate f on all R .

Step 2. We fix $(x, y) \in R$, $x \leq \frac{2a}{3}$. We set $a' = -\frac{y}{a-x}$, $b' = \frac{b-y}{a-x}$. We note that, since $x \leq \frac{2a}{3}$, we have $|a'|, |b'| \leq 6$. For $x \leq s \leq a$, we set

$$\varphi(s) = a'(s-x) + y, \quad \psi(s) = b'(s-x) + y.$$

We note that $\psi(s) - \varphi(s) = \frac{b(s-x)}{(a-x)}$. For $x \leq s \leq a$, we set

$$h(s) = \int_0^1 f(u(t, s)) ds,$$

where, for simplicity, we set $u(t, s) = (s, \varphi(s) + t(\psi(s) - \varphi(s)))$. Thus $h(x) = f(x, y)$, and $h(a)$ does not depend on x, y . Now, by integration by parts,

$$h(a) - h(x) = \int_x^a h'(s) ds = (a-x)h'(a) - \int_x^a (s-x)h''(s) ds.$$

Differentiation under the integral gives

$$\begin{aligned} h'(s) &= \int_0^1 \frac{\partial f}{\partial x}(u(t, s)) dt + (a' + t(b' - a')) \frac{\partial f}{\partial y}(u(t, s)) dt, \\ h''(s) &= \int_0^1 \frac{\partial^2 f}{\partial x^2}(u(t, s)) dt + (a' + t(b' - a')) \frac{\partial^2 f}{\partial x \partial y}(u(t, s)) dt \\ &\quad + \int_0^1 (a' + t(b' - a'))^2 \frac{\partial^2 f}{\partial y^2}(u(t, s)) dt. \end{aligned}$$

Since $a', b' \leq 3$, we have $|h''(s)| \leq 18 \int_0^1 g(u(t, s)) dt$. Thus we have, since $s - x = (a - x)(\psi(s) - \varphi(s))/b \leq 2(\psi(s) - \varphi(s))$,

$$\begin{aligned} \left| \int_x^a (s - x) h''(s) ds \right| &\leq 36 \int_x^a ds (\psi(s) - \varphi(s)) \int_0^1 g(u(t, s)) dt \\ &\leq 36 \int_x^a ds \int_{\varphi(s)}^{\psi(s)} g(s, y) dy \leq 36 \iint_R g d\lambda \end{aligned}$$

by change of variable.

Now we have

$$h'(a) = \int_0^1 \frac{\partial f}{\partial x}(a, bt) dt + (a' + t(b' - a')) \frac{\partial f}{\partial y}(a, bt) dt.$$

By definition of a' and b' , $(x - a)h'(a)$ is a linear function of x, y . Thus, we have shown that

$$|f(x, y) - h(a) - (x - a)h'(a)| \leq K \iint_C g d\lambda$$

and that $h(a) + (x - a)h'(a)$ is an affine function of x, y . $\square$

We denote by ∇f the gradient of f .

Lemma 9.3. Consider $f \in \mathcal{C}$ and $M > 0$. Then we can find a set $B \subset [0, 1]^2$ with the following properties.

$$(9.1) \quad |B| \geq 1 - K/M.$$

$$(9.2) \quad u, v \in B \Rightarrow \left| f(u) - f(v) - \left\langle u - v, \frac{\nabla f(u) + \nabla f(v)}{2} \right\rangle \right| \leq KM \|u - v\|^2.$$

$$(9.3) \quad u, v \in B \Rightarrow \|\nabla f(u) - \nabla f(v)\| \leq KM \|u - v\|.$$

Proof. For $u \in [0, 1]$, $s \geq 0$, we denote by $R(u, s)$ the intersection of $[0, 1]^2$ and of the square of center u and side $2s$. It follows from the two-dimensional version of Lemma 6.5 that if one sets

$$B_1 = \{u \in [0, 1]^2, \forall s \geq 0, m_{R(u, s)}(g) \leq M\},$$

then $|B_1| \geq 1 - K/M$.

Given a function h on $\mathbb{R}$ such that $\int_{[0,1]} h d\lambda \leq 1$, Lemma 6.5 shows that

$$\left| \left\{ x; \forall t \geq 0, \int_{I(x,t)} h d\lambda \leq M|I(x,t)| \right\} \right| \geq 1 - \frac{K}{M} \int h d\lambda,$$

where $I(x,t) = [0,1] \cap [x-t, x+t]$. Thus, by the Fubini theorem, we have $|B_2| \geq 1 - K/M$, where

$$B_2 = \left\{ (x,y) \in [0,1]^2, \forall t \geq 0, \int_{I(x,t)} \left| \frac{\partial^2 f}{\partial x^2}(s,y) \right| ds \leq M|I(x,t)| \right\}$$

and similarly, $|B_3| \geq 1 - K/M$, where

$$B_3 = \left\{ (x,y) \in [0,1]^2, \forall t \geq 0, \int_{I(y,t)} \left| \frac{\partial^2 f}{\partial y^2}(x,s) \right| ds \leq M|I(y,t)| \right\}.$$

We prove that $B = B_1 \cap B_2 \cap B_3$ satisfies (9.2), (9.3). Consider $u, v \in B$, and the smallest s for which $v \in R(u,s)$ (so that $s \leq \|v - u\|$). Since $u \in B_1$, we have $\iint_{R(u,s)} g d\lambda \leq 4Ms^2$. Clearly the sides a, b of $R(u,s)$ satisfy $\frac{1}{2} \leq \frac{b}{a} \leq 2$. Thus we can use Proposition 9.2 to find an affine function h on $R(u,s)$ such that

$$(9.4) \quad \|f - h\|_{\infty, R(u,s)} \leq KMs^2.$$

Since $u = (x,y) \in B_2$, we have for $x' \in I(x,s)$ that

$$\left| \frac{\partial f}{\partial x}(u) - \frac{\partial f}{\partial x}(x', y) \right| \leq \int_{I(x,s)} \left| \frac{\partial^2 f}{\partial x^2}(t, y) \right| dt \leq KMs.$$

Then averaging over $x' \in I(x,s)$ gives

$$\left| \frac{\partial f}{\partial x}(u) - \frac{1}{|I(x,s)|} \int_{I(x,s)} \frac{\partial f}{\partial x}(x', y) dx' \right| \leq KMs.$$

Writing $I(x,s) = [x_1, x_2]$, we get

$$\left| \frac{\partial f}{\partial x}(u) - \frac{1}{x_2 - x_1} (f(x_2, y) - f(x_1, y)) \right| \leq KMs.$$

Now, since h is affine, we have

$$\left| \frac{\partial h}{\partial x}(u) - \frac{1}{x_2 - x_1} (h(x_2, y) - h(x_1, y)) \right| = 0.$$

Using (9.4), and since $x_2 - x_1 \geq s$, we get

$$\left| \frac{\partial f}{\partial x}(u) - \frac{\partial h}{\partial x}(u) \right| \leq KMs.$$

Since $u \notin B_3$, we get $|\frac{\partial f}{\partial y}(u) - \frac{\partial h}{\partial y}(u)| \leq KMs$, by the same argument, so that

$$(9.5) \quad \|\nabla f(u) - \nabla h(u)\| \leq KMs.$$

Observe that the same relation holds at v . Since h is affine, $\nabla h(u) = \nabla h(v)$; so (9.3) is proved. We observe that

$$h(u) - h(v) = \langle u - v, \frac{\nabla h(u) + \nabla h(v)}{2} \rangle$$

since h is affine. Combined with (9.4) and (9.5), we get (9.2). $\square$

Lemma 9.4. Consider $u_1, u_2 \in \mathbb{R}^2$, $a_1, a_2 \in \mathbb{R}$. Then we have

$$\forall w \in \mathbb{R}^2, M\|w - u_1\|^2 + a_1 \geq -M\|w - u_2\|^2 + a_2$$

if and only if $\frac{M}{2}\|u_1 - u_2\|^2 + a_1 - a_2 \geq 0$.

Proof. $\|w - u_1\|^2 + \|w - u_2\|^2 = 2\|w - \frac{(u_1 + u_2)}{2}\|^2 + \frac{1}{2}\|u_1 - u_2\|^2$. $\square$

Lemma 9.5. For $i = 0, 1$, consider $u_i, v_i, b_i \in \mathbb{R}^2$. Set

$$f_i(w) = (-1)^i M\|w - u_i\|^2 + \langle w - u_i, v_i \rangle + b_i.$$

Then we have

$$\forall w \in \mathbb{R}^2, f_0(w) \geq f_1(w)$$

if and only if

$$(9.6) \quad \frac{M}{2}\|u_0 - u_1\|^2 - \frac{1}{4M}\|v_0 - v_1\|^2 - \left\langle u_0 - u_1, \frac{v_0 + v_1}{2} \right\rangle + b_0 - b_1 \geq 0.$$

Proof. We write

$$f_i(w) = (-1)^i M\|w - u_i + (-1)^i v_i/2M\|^2 + (-1)^i \|v_i\|^2/4M + b_i$$

and use Lemma 9.4 and algebra. $\square$

Consider $f \in \mathcal{C}$. For $u \in [0, 1]^2$, $i = 0, 1$, and $w \in \mathbb{R}$, we set

$$f_{i,u}(w) = (-1)^i M\|w - u\| + \langle w - u, \nabla f(u) \rangle + f(u).$$

Lemma 9.6. Consider $f \in \mathcal{C}$. Then, for $M > 0$, we can find a set $B \subset [0, 1]^2$ such that $|B| \geq 1 - K/M$ and $\forall u, v \in B$, $\forall w \in \mathbb{R}^2$, $f_{0,u}(w) \geq f_{1,v}(w)$.

Proof. Combine Lemmas 9.5 and Lemma 9.3 used for M/K rather than M . (With the notation of Lemma 9.3, in Lemma 9.5 we take $u_1 = u$, $u_2 = 0$, $v = \nabla(u)$, $v_1 = \nabla f(u)$, $v_2 = \nabla f(v)$, $b_1 = f(u)$, $b_2 = f(v)$.) $\square$

For a function h defined on $\mathbb{R}^2$, we define the function $\hat{h}$ by

$$\hat{h}(u) = \inf\{t; \exists v \in \mathbb{R}^2, \forall w \in \mathbb{R}^2, t + \langle w - u, v \rangle + M\|w - u\|^2 \geq h(w)\}.$$

We define $\check{h}$ by $\check{h} = -(-h)^\wedge$.

Lemma 9.7. For $u, u' \in \mathbb{R}^2$, we have

$$\hat{h}(u - u') + \hat{h}(u + u') \leq 2\hat{h}(u) + 2M\|u'\|^2.$$

Proof. This inequality is satisfied by the function $g(u) = M\|u - v\|^2 + a$ and, hence, by any infimum of such functions; $\hat{h}$ is such an infimum. $\square$

Corollary 9.8. Consider a function h on $\mathbb{R}^2$. If $h = \hat{h} = \check{h}$, we have

$$\forall u, u' \in \mathbb{R}^2, |h(u + u') + h(u - u') - 2h(u)| \leq 2M\|u'\|^2.$$

We now state the main approximation procedure.

Proposition 9.9. Consider $f \in \mathcal{C}$ and $M > 0$; then we can find a function θ defined on all $\mathbb{R}^2$ such that

$$(9.7) \quad |\{u \in [0, 1]^2; f(u) \neq \theta(u)\}| \leq K/M,$$

$$(9.8) \quad \forall u, u' \in \mathbb{R}^2, \quad |\theta(u + u') + \theta(u - u') - 2\theta(u)| \leq 2M\|u'\|^2.$$

Proof. Consider first the set B of Lemma 9.6. We set for $w \in \mathbb{R}^2$

$$h(w) = \sup\{f_{1,v}(w); v \in B\},$$

and we set $\theta = \hat{h}$. We first prove (9.7). Given $u, v \in B$, by Lemma 9.6, we have $f_{0,u}(w) \geq f_{1,v}(w)$ for all $w \in \mathbb{R}^2$. Thus, by definition of $\hat{h}$, we have $\hat{h}(u) \leq f_{0,u}(u) = f(u)$. On the other hand, for $u \in B$, we have $\hat{h}(u) \geq h(u) \geq f_{1,u}(u) = f(u)$. Thus $\hat{h}(u) = f(u)$ for $u \in B$, which proves (9.7). To obtain (9.8), we will apply Corollary 9.8 to θ . It is obvious by definition that $\hat{\hat{h}} = \hat{h}$, so that $\hat{\theta} = \theta$. It remains to prove that $\hat{\theta} = \theta$. This unfortunately will require tedious work.

Lemma 9.10. (a) Consider a continuous function h on $\mathbb{R}^2$. Then given $u \in \mathbb{R}^2$ we can find $v \in \mathbb{R}^2$ such that

$$\forall w \in \mathbb{R}^2, \quad \xi(w) \geq h(w),$$

where

$$(9.9) \quad \xi(w) = \hat{h}(u) + \langle w - u, v \rangle + M\|w - u\|^2.$$

(b) Suppose that $h(w) \leq M\|w\|^2/2$ for $\|w\|$ large enough. Then $u \in \text{conv}\{h = \xi\}$.

Proof. (a) By definition of $\hat{h}$ for each n we can find $v_n \in \mathbb{R}^2$ such that

$$\forall w \in \mathbb{R}^2, \quad \xi_n(w) \geq h(w),$$

where

$$\xi_n(w) = \hat{h}(u) + 2^{-n} + \langle w - u, v_n \rangle + M\|w - u\|^2.$$

Taking $w_n = u - v_n/M\|v_n\|$, we get that

$$h(w_n) \leq \hat{h}(u) + 2^{-n} - \|v_n\|/M + 1/M.$$

Since $\|w_n - u\| = 1/M$, $-h(w_n)$ stays bounded, so $\|v_n\|$ is bounded. Thus we can find a subsequence (v_n) that converges to some $v \in \mathbb{R}^2$, and the function ξ given by (9.9) satisfies $\xi(w) \geq h(w)$ for all w .

(b) Since we suppose now that $h(u) \leq M\|u\|^2/2$ for $\|u\|$ large enough, the set $H = \{\xi = h\}$ is bounded closed; so it compact, and its convex hull is closed. Suppose, for contradiction, that u does not belong to this convex hull. Then, by the Hahn-Banach Theorem, we can find $v' \in \mathbb{R}^2$ and $\delta > 0$ such that

$$w \in H \Rightarrow \langle v', w - u \rangle > \delta > 0.$$

For $\epsilon > 0$, consider the function

$$\xi_\epsilon(w) = \hat{h}(u) + \langle w - u, v + \epsilon v' \rangle + M\|w - u\|^2.$$

If $\langle w - u, v' \rangle \geq \delta/2$, we have

$$\xi_\epsilon(w) \geq \xi(w) + \epsilon\delta/2.$$

Since we assume $|h(w)| \leq M\|w\|^2/2$ for $\|w\|$ large enough, there exists a such that

$$\|w\| > a \Rightarrow \xi(w) - h(w) \geq M\|w\|^2/3.$$

Since the set $\{\xi = h\}$ does not meet the set $\{\langle w - u, v' \rangle \leq \delta/2\}$, we have

$$\gamma = \inf\{\xi(w) - h(w); \langle w - u, v' \rangle \leq \delta/2; \|w\| \leq a\} > 0.$$

If $\langle w - u, v' \rangle \leq \delta/2$, $\|w\| \leq a$, we then have

$$\xi_\epsilon(w) = \xi(w) + \epsilon\langle w - u, v' \rangle \geq h(w) + \gamma - \epsilon\|v'\|(\|u\| + a).$$

If $\|w\| \geq a$, we have

$$\begin{aligned} \xi_\epsilon(w) &= \xi(w) + \epsilon\langle w - u, v' \rangle \\ &\geq h(w) + \frac{M}{3}\|w\|^2 - \epsilon\|v'\|(\|w\| + \|u\|). \end{aligned}$$

Thus, if ϵ is chosen small enough, there is $\eta > 0$ such that

$$\forall w \in \mathbb{R}^2, \quad \xi_\epsilon(w) \geq h(w) + \eta$$

so that $\hat{h}(u) \leq \xi_\epsilon(u) - \eta = \hat{h}(u) - \eta$, a contradiction. $\square$

We now finish the proof that $\theta = \check{\theta}$. There is no loss of generality to assume that ∇f stays bounded on B . Then it is clear that h satisfies $\|h(w)\| \leq M\|w\|^2/2$ for $\|w\|$ large enough (actually $h(w) \leq 0$ for $\|w\|$ large enough). Consider $u_0 \in \mathbb{R}^2$. Since $\theta = \hat{h}$, by Lemma 9.10(a) we can find v such that $h(w) \leq \xi(w)$ for all $w \in \mathbb{R}^2$, where

$$\xi(w) = \theta(u_0) + \langle w - u_0, v \rangle + M\|w - u_0\|^2.$$

We set $H = \{\xi = h\}$. Thus by Lemma 9.10(b) we have $u \in \text{conv } H$.

By the definition of h and the argument of Lemma 9.10(a), given $u \in H$, we can find a function h_u of the type

$$h_u(w) = h(u) + \langle w - u, v_u \rangle - M\|w - u\|^2$$

such that $h_u(w) \leq h(w)$ for all $w \in \mathbb{R}^2$. Since $\xi(w) \geq h(w)$ for all $w \in \mathbb{R}^2$, and since $\xi(u) = h(u) = h_u(u)$, we see that ξ and h_u must have the same gradient at u , so that

$$h_u(w) = \xi(u) + \langle w - u, v + 2M(u - u_0) \rangle - M\|w - u\|^2.$$

For $v' \in \mathbb{R}^2$, $a \in \mathbb{R}$, consider the function

$$\eta(w) = \langle w, v' \rangle + M\|w\|^2 + a$$

and assume that $\eta(w) \geq h(w)$ for $w \in \mathbb{R}^2$. Thus, for all $u \in H$, we have $\eta(w) \geq h_u(w)$ for all $w \in \mathbb{R}^2$. If we use Lemma 9.5 and algebra we see that this is equivalent to saying that

$$(9.10) \quad G(u, v', a) \geq 0$$

where $G(u, v', a)$ is affine in u . Thus (9.10) continues to hold for all $u \in \text{conv } H$ and, in particular, for $u = u_0$. This implies in turn that, for all $w \in \mathbb{R}^2$, we have $\eta(w) \geq h_{u_0}(w)$, where

$$h_{u_0}(w) = \xi(u_0) + \langle w - u_0, v \rangle - M\|w - u_0\|^2.$$

Since $\eta \geq h_{u_0}$ whatever the choice of v', a such that $\eta \geq h$, this shows that $h_{u_0} \leq \hat{h} = \theta$, so that $h_{u_0}(u_0) = \xi(u_0) \geq \check{\theta}(u_0)$ by definition of $\check{\theta}$. Proposition 9.9 is proved. $\square$

Consider the class $\mathcal{E}_\ell$ of functions on $[0, 1]^2$ that have the following properties.

$$(9.11) \quad |\{f \neq 0\}| \leq K2^{-\ell}.$$

$$(9.12) \quad f \text{ is the restriction to } [0, 1]^2 \text{ of a function } \theta \text{ on } \mathbb{R}^2 \text{ such that}$$

$$(9.13) \quad \forall u, u' \in \mathbb{R}^2, \quad |\theta(u + u') + \theta(u - u') - 2\theta(u)| \leq K2^\ell \|u'\|^2.$$

It follows from Proposition 9.9 that each $f \in \mathcal{E}$ can be written $f = \sum_{\ell \geq 1} f_\ell$ where $f_\ell \in \mathcal{E}_\ell$ (provided the constant K of (9.13) has been chosen suitably). To prove that $\gamma_{1,2}(\mathcal{E}) < \infty$, it thereby suffices to prove that $\gamma_{1,2}(\mathcal{E}_\ell) \leq K\ell 2^{-\ell/2}$. To that effect we will use Proposition 6.2 with $k_0 = 0$, $b = K2^{-\ell}$, $\alpha = 1$, $\beta = 2$, $M = K2^\ell$. In order to check (6.3), we need the following lemma.

Lemma 9.11. *Consider a square A of side a and the class $\mathcal{F}_A$ of functions f on A that satisfy $|\{f = 0\}| \geq |A|/2$ and that are restrictions to A of a function θ on $\mathbb{R}^2$ that satisfies (9.13). Then $\gamma_{1,2}(\mathcal{F}_A) \leq K2^\ell a^3$.*

Proof. *Step 1.* We can assume $A = [0, a]^2$. For a function f on $[0, a]^2$ (resp. $\mathbb{R}^2$) we define $\bar{f}$ on $[0, 1]^2$ (resp. $\mathbb{R}^2$) by $\bar{f}(x, y) = af(ax, ay)$; this is an isometry for the L^2 norm. If θ satisfies (9.13), then $\bar{\theta}$ satisfies (9.13) with $a^3 2^\ell$ instead of 2^ℓ . Thus by homogeneity we can assume $a = 1$, and we can assume $\ell = 0$. We write $\mathcal{G}$ for $\mathcal{F}_{[0,1]^2}$.

Step 2. We show that if $f \in \mathcal{G}$ then $|f| \leq K$. Consider $u \in [0, 1]^2$. Then we can find a unit vector v such that $|\{t \in [0, \sqrt{2}]; f(u + tv) = 0\}| \geq 1/K_3$. This reduces to the following statement: If a function h on $[0, a]$ ($a \leq \sqrt{2}$) satisfies $|\{h = 0\}| \geq 1/K_3$ and $|h(t+s) + h(t-s) - 2h(t)| \leq s^2$ for $s-t, s+t \in [0, a]$, then $\|h\|_\infty \leq K$. The proof of this elementary fact is left to the reader.

Step 3. Consider the class $\mathcal{H}$ of functions f on $[0, 1]^2$ that satisfy $\|f\|_\infty \leq 1$, $\|\partial^2 f / \partial x^2\|_\infty \leq 1$, and $\|\partial^2 f / \partial y^2\|_\infty \leq 1$. We show that it suffices to prove that $\gamma_{1,2}(\mathcal{H}) < \infty$. It is sufficient to prove that $\mathcal{G}$ is contained in the L^2 closure of $K\mathcal{H}$ for K large enough. This follows from the following observations. If θ on $\mathbb{R}^2$ satisfies (9.13), then so does $\theta * \varphi$, where $\varphi \geq 0$, $\int \varphi d\lambda = 1$; when the support of φ goes to $\{0\}$, $\theta * \varphi$ converges uniformly to θ on each compact set; If φ is infinitely differentiable, so is $\theta * \varphi$.

Finally if θ is 2 times differentiable and satisfies (9.13) (with $\ell = 0$) then $\|\partial^2 \theta / \partial x^2\|_\infty \leq K$, $\|\partial^2 \theta / \partial y^2\|_\infty \leq K$.

Step 4. We prove using Fourier transform that $\gamma_{1,2}(\mathcal{H}) < \infty$. The proof resembles that of Lemma 4.3. We use the notation of that lemma. It suffices to show that $\gamma_{1,2}(T) < \infty$, where T is the sets of sequences $a_{n,m}(f)$, $f \in \mathcal{H}$, $(n, m) \in A = \{(n, m); |n| \geq |m|, |n| \geq 1\}$. Integrating by parts (4.1) one more time shows that it suffices to prove that $\gamma_{1,2}(T_i) < \infty$ for $i = 1, 2, 3$, where

$$\begin{aligned} T_1 &= \left\{ \left(\frac{1}{n} b_m(g) \right)_{n,m \in A}, |g|, |g'| \leq 2 \right\}, \\ T_2 &= \left\{ \left(\frac{1}{n^2} b_m(g) \right)_{n,m \in A}, |g|, |g'| \leq 2 \right\}, \\ T_3 &= \left\{ \left(\frac{1}{n^2} a_{n,m}(h) \right)_{n,m \in A}, \|h\|_\infty \leq 1 \right\}. \end{aligned}$$

The proof of Lemma 5.2 consists of showing that $\gamma_{1,2}(U) < \infty$, where

$$U = \{(b_m(g))_{m \in \mathbb{Z}}; |g|, |g'| \leq 1\}.$$

Thus, since T_1, T_2 are images of U by a Lipschitz map, we have $\gamma_{1,2}(T_1), \gamma_{1,2}(T_2) < \infty$. Thus it suffices to prove that $\gamma_{1,2}(T_3) < \infty$. But this follows from Theorem 2.4. $\square$

We have now proved that $\gamma_{1,2}(\mathcal{E}) < \infty$.

The proof of Theorem 1.6 will also use a bracketing argument. Recall that if $f \leq f'$, the bracket $[f, f']$ is the set of functions h such that $f \leq h \leq f'$.

Proposition 9.12. *Given $n \geq 1$ the class $\mathcal{E}$ can be covered by at most $\exp Kn$ brackets $[f, f']$ such that $\|f' - f\|_1 \leq K/n$.*

Proof. *Step 1.* Consider the largest integer k_0 such that $2^{-k_0} \geq 1/\sqrt{n}$. For $k \leq k_0$ consider the paving $\mathcal{P}_k$ of $[0, 1]^2$ by squares of side 2^{-k} . Thus $\mathcal{P}_{k_0}$ has at most n elements. Consider a family $e = (e_A)_{A \in \mathcal{P}_{k_0}}$ of positive numbers such that $\sum_{A \in \mathcal{P}_{k_0}} e_A \leq 3$. Consider the family

$$\mathcal{E}(e) = \left\{ f \in \mathcal{E}, \forall A \in \mathcal{P}_{k_0}, \iint_A \left(\left| \frac{\partial^2 f}{\partial x^2} \right| + \left| \frac{\partial^2 f}{\partial y^2} \right| + \left| \frac{\partial^2 f}{\partial x \partial y} \right| \right) d\lambda \leq e_A \right\}.$$

In that step, we show that it suffices to prove the result for $\mathcal{E}(e)$ rather than $\mathcal{E}$. It suffices to show that $\mathcal{E}$ can be covered by at most $\exp Kn$ classes $\mathcal{E}(e)$. We recall the following easy fact.

Lemma 9.13 (see, e.g., [Sc]). *The unit ball of ℓ_m^1 can be covered by $\leq \exp Km$ translates of B_∞/m , where B_∞ is the unit ball of ℓ_m^∞ .*

This lemma implies that one can find a subset Y of ℓ_m^1 , with $\text{card } Y \leq \exp Km$, such that given $x \in \ell_m^1$, $\|x\|_1 \leq 1$ there is $y \in Y$ for which $|x_i - y_i| \leq$

$1/m$ for each $i \leq n$. Thus $x_i \leq y_i + 1/m \leq x_i + 2/m$ and $\sum_{i \leq m} (y_i + 1/m) \leq 3$. This implies that $\mathcal{E}$ can be covered by $\leq \exp Kn$ classes $\mathcal{E}(e)$.

Step 2. Consider, for $k \geq 0$, the set H_k of points of $[0, 1]^2$ that are vertices of squares of $\mathcal{P}_k$. Set

$$G_k = \{f; \forall u \in H_k, |f(u)| \leq 2^{-2k}\}.$$

We show that for $k \leq k_0$ we have

$$N(\mathcal{E}(e), G_k) \leq \exp K 2^{2k}.$$

This is proved by induction over k . To perform the induction, it suffices to show that if $f \in \mathcal{E}(e)$, we have

$$N(\mathcal{E}(e) \cap (f + 2G_k), G_{k+1}) \leq \exp K 2^{2k}.$$

If for a square B of $\mathcal{P}_k$, we set $e_B = \sum_{A \subset B, A \in \mathcal{P}_{k_0}} e_A$, we have by Proposition 9.2 that whenever $f, h \in \mathcal{E}(e)$, $f - h \in 2G_k$

$$\|f - h\|_{\infty, B} \leq K(2^{-2k} + e_B).$$

It follows that

$$\sum_{u \in H_{k+1}} (f - h)(u) \leq K.$$

The conclusion then follows from Lemma 9.13.

Step 3. To conclude the proof, we consider a set $\mathcal{E}(e) \cap (f + 2G_{k_0})$, where $f \in \mathcal{E}(e)$, and we show that this set itself is contained in a suitable bracket. For $h \in \mathcal{E}(e) \cap (f + 2G_{k_0})$ and $A \in \mathcal{P}_{k_0}$, we have by Proposition 9.2 that

$$\|f - h\|_{\infty, A} \leq K(2^{-2k_0} + e_A).$$

Thus $\mathcal{E}(e) \cap (f + 2G_{k_0}) \subset [f_0, f_1]$ where

$$f_i = f + (-1)^i K(2^{-2k_0} + e_A)$$

on A . Thus $\|f_1 - f_0\|_1 \leq K/n$. $\square$

Consider now the class $\mathcal{H}$ of subsets of $[0, 1]^3$ that are the intersection of $[0, 1]^3$ and of the subgraph of a function $f \in 20\mathcal{E}$. Consider six rotations $R_1, \dots, R_6$ that leave the cube $[0, 1]^3$ invariant. Choose these rotations in order that each face of the cube is obtained as the image of the top face under one of $R_1, \dots, R_6$. Consider the class $\mathcal{G}$ of subsets of $[0, 1]^3$ that are of the type $\bigcup_{i \leq 6} R_i(H_i)$, where $H_i \in \mathcal{H}$.

Theorem 9.13. $ED_n(\mathcal{G}) \leq Kn^{1/2}(\log n)^{3/4}$.

To motivate this statement, we show that $\mathcal{G}$ contains the complement of each convex set $C \subset [0, 1]^3$, thereby proving Theorem 1.6. Consider the set W of unit vectors of $\mathbb{R}^3$ that have a third component $\geq 1/\sqrt{3}$. For $w \in W$, we

define the subset S_w of $\mathbb{R}^3$ given by $S_w = \{u \in \mathbb{R}^3; \langle u, w \rangle \leq \inf\{\langle v, w \rangle; v \in C\}\}$. Consider the function f on $[0, 1]^2$, whose subgraph is the union H of all the half spaces S_w for $w \in W$. Thus f is convex, and the definition of W easily implies that $|\frac{\partial f}{\partial x}|, |\frac{\partial f}{\partial y}| \leq \sqrt{6}$. Thus $f \in 20\mathcal{E}$ by Proposition 9.1. Consider the sets H'_i obtained in this manner with $R_i(C)$ instead of C and the sets $H_i = R_i^{-1}(H'_i) \cap [0, 1]^3$. We note that each point not in C is contained in a half space that does not meet C and that, for a unit vector of $\mathbb{R}^3$, one of the components has absolute value $\geq 1/\sqrt{3}$. These observations imply that $\bigcup_{i \leq 6} H_i$ is the complement of C .

Proof of Theorem 9.13. If we use Proposition 9.12 with $\sqrt{n}$ instead of n , we see that we can find a family $\mathcal{X}$ of subsets of $[0, 1]^3$ and $\mathcal{G}' \subset \mathcal{G}$, with the following properties.

$$(9.14) \quad \text{card } \mathcal{X} \leq \exp K\sqrt{n}, \quad \text{card } \mathcal{G}' \leq \exp K\sqrt{n},$$

$$(9.15) \quad \forall X \in \mathcal{X}, \quad |X| \leq K/\sqrt{n},$$

$$(9.16) \quad \forall G \in \mathcal{G}, \exists G' \in \mathcal{G}', \exists X \in \mathcal{X}, \quad G \Delta G' \subset X.$$

It is then clear that

$$(9.17) \quad D_n(\mathcal{G}) \leq D_n(\mathcal{G}') + \sup_{X \in \mathcal{X}} \text{card}\{i \leq n; U_i \in X\} + K\sqrt{n}.$$

By (4.4) we have

$$P(\{\forall X \in \mathcal{X}, \text{card}\{i \leq n; U_i \in X\} \geq 2n|X| + t\}) \leq \exp(K\sqrt{n} - t)$$

so that

$$(9.18) \quad E \sup_{X \in \mathcal{X}} \text{card}\{i \leq n; U_i \in X\} \leq K\sqrt{n}.$$

We now observe that if d denotes the distance on the subsets of $[0, 1]^3$ given by $d(A, B) = |A \Delta B|$ then $(\mathcal{G}, d)$ is a Lipschitz image of the product of six copies of $\mathcal{E}$. Thus, since $\gamma_{1,2}(\mathcal{E}) < \infty$, we have $\gamma_{1,2}(\mathcal{G}, d) < \infty$. It follows from Lemma 5.1(a) taking a there equal to say, $1/n$, that

$$(9.19) \quad \gamma_{1/2,1}(\mathcal{G}', \sqrt{d}) \leq K(\log n)^{3/4}.$$

Consider now the random distance d_U of $\mathcal{G}'$ given by

$$d_U(A, B) = (\text{card}\{i \leq n; U_i \in A \Delta B\})^{1/2}.$$

It follows from (4.4) and (9.14) that

$$P(\forall A, B \in \mathcal{G}', d_U^2(A, B) \geq t + 2n d(A, B)) \leq \exp(K\sqrt{n} - t).$$

Thus, we can find a r.v. s such that $Es < \infty$ and

$$(9.20) \quad \forall A, B \in \mathcal{G}', \quad d_U(A, B) \leq n^{1/4}s + \sqrt{nd(A, B)}.$$

Consider now a subset $\mathcal{G}''$ of $\mathcal{G}'$, maximal with respect to the property that $\forall A, B \in \mathcal{G}'', d(A, B) \geq n^{-1/2}$. From (9.20) follows that for $A, B \in \mathcal{G}''$

$$(9.21) \quad d_U(A, B) \leq (1+s)\sqrt{nd(A, B)}.$$

By Corollaries 2.7 and 2.8 and (9.19) we have

$$\gamma_{1/2,1}(\mathcal{G}'', d_U) \leq K(1+s)\sqrt{n}(\log n)^{3/4}.$$

Consider a probability μ on $\mathcal{G}''$ such that

$$\gamma_{1/2,1}(\mathcal{G}'', d_U, \mu) \leq 2(s+1)\sqrt{n}(\log n)^{3/4}.$$

Consider the probability ν on $\mathcal{G}'$ given by $\nu = \frac{1}{2}(\mu + \eta)$, where η gives mass $\frac{1}{2}(\text{card } \mathcal{G}')^{-1}$ to each point of $\mathcal{G}'$. It is easily seen that

$$\gamma_{1/2,1}(\mathcal{G}', d_U, \nu) \leq K(s+1)\sqrt{n}(\log n)^{3/4}.$$

Finally we have shown that

$$\gamma_{1/2,1}(\mathcal{G}', d_U) \leq K(s+1)\sqrt{n}(\log n)^{3/4}.$$

This in turn implies as usual that $ED_n(g') \leq K\sqrt{n}(\log n)^{3/4}$. The conclusion then follows by combining this result with (9.17), (9.18).

REFERENCES

- [AKT] M. Ajtai, J. Komlós, and G. Tusnády, *On optimal matchings*, *Combinatorica* **4** (1984), 259–264.
- [B1] S. M. Berman, *Some continuity properties of Brownian motion with the time parameter in Hilbert space*, *Trans. Amer. Math. Soc.* **131** (1968), 182–198.
- [B2] ———, *Harmonic analysis of local time on sample functions of Gaussian processes*, *Trans. Amer. Math. Soc.* **143** (1969), 269–281.
- [CS] E. Coffman and P. W. Shor, *A simple proof of the $O(n^{1/2} \log^{3/4} n)$ up-right matching bound*, *SIAM J. Discrete Math.* **4** (1991), 68–57.
- [D] R. M. Dudley, *The size of compact subsets of Hilbert space and continuity of Gaussian processes*, *J. Funct. Anal.* **1** (1967), 290–330.
- [GZ] E. Giné and J. Zinn, *Limit theorems for empirical processes*, *Ann. Probab.* **12** (1984), 929–989.
- [H] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, *J. Amer. Statist. Assoc.* **58** (1963), 13–30.
- [KLM-S] R. M. Karp, M. Luby, and A. Marchetti-Spaccamela, *Probabilistic analysis of multidimensional bin packing problems*, *Proceedings of the 16th ACM Sympos. on Theory of Computing*, A.C.M., New York, 1984, pp. 289–298.
- [LT] M. Ledoux and M. Talagrand, *Probability in a Banach space*, Springer-Verlag, New York, 1991.
- [LS] T. Leighton and P. W. Shor, *Tight bounds for minimax grid matching with applications to the average case analysis of algorithms*, *Combinatorica* **9** (1989), 161–187.
- [O] R. Osserman, *The isoperimetric inequality*, *Bull. Amer. Math. Soc.* **84** (1978), 1182–1238.
- [PS] C. H. Papadimitriou and K. Steiglitz, *Combinatorial optimization, algorithms and complexity*, Prentice-Hall, Englewood Cliffs, NJ, 1982.

- [RT] W. T. Rhee and M. Talagrand, *Exact bounds for the stochastic upward matching problem*, Trans. Amer. Math. Soc. **307** (1988), 109–125.
- [Sc] C. Schütt, *Entropy numbers of diagonal operators between Banach spaces*, J. Approx. Theory **40** (1984), 121–128.
- [S1] P. W. Shor, *Random planar matching and bin packing*, Ph.D. thesis, M.I.T., 1985.
- [S2] ———, private communications.
- [S3] ———, *How to pack better than Best fit: Tight bounds for average-case on-line bin packing*, Proc. 32nd Sympos. on Foundations of Computer Science, IEEE Computer Soc. Press, New York, 1991, pp. 752–759.
- [SY] P. W. Shor and J. Yukich, *Minimax grid matching and Empirical measures*, Ann. Probab. (to appear).
- [T1] M. Talagrand, *Regularity of Gaussian processes*, Acta. Math **159** (1987), 99–147.
- [T2] ———, *An isoperimetric theorem on the cube and the Kinchine-Kahane inequalities*, Proc. Amer. Math. Soc. **104** (1989), 905–909.
- [T3] ———, *Sample boundedness of stochastic processes under increment conditions*, Ann. Probab. **18** (1990), 1–49.
- [Za] A. C. Zaneen, *Linear analysis*, North-Holland, Amsterdam, 1953.

ABSTRACT. We give explicit constructions of certain majorizing measures. These constructions allow us to give a unified proof of deep matching theorems of Ajtai, Komlós, and Tusnády, of Leighton and Shor, and of Shor, as well as of more precise and more general results in a similar spirit.

EQUIPE D'ANALYSE, TOUR 46, U.A. NO754 AU C.N.R.S., UNIVERSITÉ PARIS VI, 75230 PARIS CEDEX 05, FRANCE

E-mail address: mit@frunip62.bitnet

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, 231 WEST 18TH AVENUE, COLUMBUS, OHIO 43210