Probab. Th. Rel. Fields 79, 431–450 (1988)

Probability Theory and Related Fields

About the Prohorov Distance Between the Uniform Distribution Over the Unit Cube in R^d and its Empirical Measure

P. Massart

Université de Paris-Sud, U.A. CNRS 743 "Statistique appliquée" Mathématiques, Bât. 425, F-91405 Orsay Cedex France

Summary. We compute the almost sure order of convergence of the Prokhorov distance between the uniform distribution P over $[0, 1]^d$ and the empirical measure associated with n independent observations with (common) distribution P. We show that this order of convergence is $n^{-1/d}$ up to a power of $\log(n)$. This result extends to the case where the observations are weakly dependent.

1. Introduction

Let $\mathscr{P}(\mathbb{R}^d)$ be the set of all Borel probability laws on $\mathbb{R}^d$.

Let $x_1, x_2, \ldots$ be $\mathbb{R}^d$-valued and bounded random variables with common distribution P.

Let P_n be the empirical measure associated with $x_1, \ldots, x_n$ $P_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}$.

The weak star convergence in $\mathscr{P}(\mathbb{R}^d)$ is metrizable by various metrics. Dudley (1969) considered two such metrics: that of Prokhorov which we call ρ and that of Mourier-Fortet β (see Dudley (1968) for a definition of β and for general relations between β and ρ). In the case where the observations are independent, he proved that the speed of convergence of $E(\rho(P_n, P))$, resp. $E(\beta(P_n, P))$, is $n^{-1/(d+2)}$, resp. $n^{-1/d}$ if $d > 2$. (Note that these results extend to cases where d is defined in terms of metric entropy and need not be an integer (see Dudley (1969)) and to cases where the observations are weakly dependent as in Gäenssler (1970)).

Moreover he showed that these rates of convergence are sharp in the sense that they cannot be improved for some choice of P. But this choice of P in the case of ρ-convergence is very special so Dudley raised the problem to find the exact order of convergence of $E(\rho(P_n, P))$ when choosing a "regular" P such as the Lebesgue measure on the unit cube in $\mathbb{R}^d$. What we intend to do in this paper is to solve this problem up to a power of $\log(n)$.

So, from now on P is the Lebesgue measure on $[0, 1]^d$ and we write $\rho_n^{(d)}$ for $\rho(P_n, P)$ and $\beta_n^{(d)}$ for $\beta(P_n, P)$.

Let us recall what is already known about the subject.

In the Case Where $d=1$

$E(\rho_n^{(1)})$ and $E(\beta_n^{(1)})$ are both of order $n^{-1/2}$ and this is related to the central limit theorem (see Dudley (1969) again).

In the Case Where $d\geqq 2$

$\underline{\lim}\ n^{1/d}\ \rho_n^{(d)}\geqq \frac{1}{2}\ \underline{\lim}\ n^{1/d}\ \beta_n^{(d)}\geqq C>0$ surely, from Bakhvalov (1959). Moreover the following upper bound is available:

$$\rho_n^{(d)}=O(\sqrt{\log(n)}\,n^{-1/(d+1)}) \quad \text{a.s. from Zuker (1974).}$$

From now on we call L the function $x\to\log(\max(x,e))$. Before stating our results, here is a definition.

For any (α,β) with $0<\alpha\leqq 2$ and $\beta\geqq 0$, we call $H(\alpha,\beta)$ the following assumption: there exist some constants, c, c' and c'' such that, for any Borel set A, the inequality:

$$\Pr(\sqrt{n}|(P_n-P)(A)|>t\sigma(L\sigma^{-1})^{\beta})\leqq c'\exp(-ct^{\alpha})$$

holds for any (t,σ) in $\mathbb{R}_+^2$ fulfilling $P(A)(1-P(A))\leqq\sigma^2$ and $t\leqq c''\sqrt{n}\,\sigma/Ln$.

Statement of the results

Theorem 1. *(Speed of the ρ-convergence in the independent case). If the observations are independent, for any integer $d\geqq 2$ there exist two positive constants $C_1(d)$ and $C_2(d)$ such that:*

(a) $\underline{\lim}_n\ (n/Ln)^{1/d}\ \rho_n^{(d)}\geqq C_1(d)$ a.s.

(b) $\overline{\lim}_n (n/(Ln)^2)^{1/d}\ \rho_n^{(d)}\leqq C_2(d)$ a.s.

Theorem 2. *(Generalization to the weakly dependent case). Assume that $H(\alpha,\beta)$ holds for some (α,β) with $0<\alpha\leqq 2$ and $\beta\geqq 0$. Then, for any integer $d\geqq 2$, there exists a positive constant $C_3(d)$ (depending of course on the constants appearing in $H(\alpha,\beta)$) such that:*

$$\overline{\lim_n}(n/(Ln)^{2+(2/\alpha)+2\beta)\,1/d})\,\rho_n^{(d)}\leqq C_3(d) \quad \text{a.s.}$$

Comments

Assumption $H(\alpha,\beta)$ holds with $(\alpha,\beta)=(1/2,\ 1)$ (resp. with $(\alpha,\beta)=(2,0)$) when the strong mixing (resp. φ-mixing) coefficient of the sequence of observations decreases geometrically to zero, see Doukhan and Portal (1987) (resp. Collomb (1984)).

The methods used to prove Theorem 1 and Theorem 2 are quite different. This is the reason why the speed of convergence in Theorem 2 when $(\alpha, \beta)=(2,0)$ is not the same as the rate given in Theorem 1 as we should expect.

As we shall see in the next sections, Theorem 1 and 2 derive from exponential bounds, so the almost sure orders of convergence above also hold in mean.

From the results of Dudley (1969) and Theorem 1 we get that the speeds of mean ρ-convergence or β-convergence are the same up to a necessary power of Ln for $d \geqq 3$. In our opinion the interesting case is $d=2$ because it is critical on the one hand (in the sense of metric entropy exponents as well as for the Donsker property) and because on the other hand, it is the key case for ρ-convergence as we shall see later. When $d=2$ our upper bound for $\rho^{(2)}$ is consistent with that of Dudley for $\beta_n^{(2)}$ both are of order $\frac{Ln}{\sqrt{n}}$; but the problem remains open to find the exact speeds of convergence for $\beta_n^{(2)}$ as well as for $\rho_n^{(2)}$ (unfortunately the method used by Bretagnolle and Massart (1986) to study the critical Hölderian classes of functions in $\mathbb{R}^d$ does not work when d is even).

We write $u \approx v$ when there exist two positive constants C and C' such that $Cu \leqq v \leqq C'u$. Changing the norm in $\mathbb{R}^d$ affects $\rho_n^{(d)}$ only through this equivalence. From now on we choose to work with the supremum norm in $\mathbb{R}^d\, |y| = \max_{1 \leqq i \leqq d} |y_i|$.

The following sections are devoted to the proofs of Theorem 1 and 2.

2. Approximation with Finite Algebras

First of all, let us recall that $\rho_n^{(d)}$ is defined by:

$$\rho_n^{(d)} = \operatorname{Inf}\{\varepsilon>0 : P_n(A) \leqq P(A^\varepsilon)+\varepsilon \quad \text{for any Borel set } A\}$$

where A^ε is the ε-neighbourhood of A, that is

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

We show that it is enough to study finite σ-algebras instead of all Borel sets when computing $\rho_n^{(d)}$.

More precisely, for any integer μ, let $\mathscr{B}_\mu^{(d)}$ be the σ-algebra generated by the cubes $C_i = \{y \in \mathbb{R}^d : i_j 2^{-\mu} < y_j \leqq (i_j+1) 2^{-\mu} \text{ for all } j\}$, $i \in \mathbb{N}^d \cap [0, 2^\mu - 1]^d$. In what follows, a Borel set B in $\mathscr{B}_\mu^{(d)}$ will not be distinguished from its closure $\bar{B}$ which is a union of closed atoms $\{y \in \mathbb{R}^d : i_j 2^{-\mu} \leqq y_j \leqq (i_j+1) 2^{-\mu} \text{ for all } j\}$. This confusion does not affect the boundary of B, nor does it change anything for B with respect to P or P_n (up to a fixed null probability set as far as the latter is concerned).

For any Borel set B in $\mathscr{B}_\mu^{(d)}$ we call *perimeter* of B the hypersurface area of its boundary. Let $s_{d-1}(B)$ denote this quantity. The following relations between P and s_{d-1} are available (we shall give a short proof of the first one in the appendix):

(2.1) $$P(B) \leqq \left(\frac{s_{d-1}(B)}{2d}\right)^{d/(d-1)} \quad (\textit{isoperimetric inequality})$$

(2.2) $$s_{d-1}(B) \leqq (2d)\, 2^\mu P(B).$$

(2.3) *Notation.* Whenever f is a function from a finite set $\mathscr{I}$ into $\mathbb{R}$, $\bigvee_{\mathscr{I}} f$ denotes the supremum of f over $\mathscr{I}$.

The next lemma means that $\rho_n^{(d)}$ and $\bigvee_{\mathscr{B}_\mu^{(d)}} \dfrac{P_n - P}{1 + s_{d-1}}$ are simultaneously of order $2^{-\mu}$.

Lemma 1. *For any integer* μ, *the two following inclusions hold:*

(a) $\{\rho_n^{(d)} > \theta\, 2^{-\mu}\} \supseteq \left\{ \bigvee_{\mathscr{B}_\mu^{(d)}} \dfrac{P_n - P}{1 + s_{d-1}} > \theta(1 + 2\theta)^{d-1}\, 2^{-\mu} \right\}$ for any positive θ,

(b) $\{\rho_n^{(d)} > \frac{3}{2}\, 2^{-\mu}\} \subseteq \left\{ \bigvee_{\mathscr{B}_\mu^{(d)}} \dfrac{P_n - P}{1 + s_{d-1}} > 2^{-\mu}/(2d) \right\}$.

Proof of lemma 1. Let $\varepsilon = \theta\, 2^{-\mu}$. It will be shown that the following inequalities hold for any B in $\mathscr{B}_\mu^{(d)}$:

(2.4) $$P(B^\varepsilon \setminus B) \leqq ((1 + 2\theta)^d - 1)/(2d)\, 2^{-\mu}\, s_{d-1}(B) \quad \text{for any positive } \theta.$$

(2.5) $$P(B^\varepsilon \setminus B) \geqq (1 - (1 - 2\theta)^d)/(2d)\, 2^{-\mu}(s_{d-1}(B) - 2d) \quad \text{for } 0 \leqq \theta \leqq \tfrac{1}{2}.$$

In fact $B^\varepsilon \setminus B$ may be described as a union of disjoint prismoids, one of the two bases of any of these prismoids being an elementary face (of a cube) composing the boundary of B. The maximal (resp. minimal) content of any of these prismoids is equal to $2^{-\mu d}((1 + 2\theta)^d - 1)/(2d)$ (resp. $2^{-\mu d}((1 - (1 - 2\theta)^d)/(2d))$, so this quantity represents the maximal (resp. minimal) contribution to $P(B^\varepsilon \setminus B)$ of one of the elementary faces composing the boundary of B (resp. of one of the elementary faces composing the boundary of B which is not included in the boundary of $[0, 1]^d$) and the number of such faces is equal to $2^{\mu(d-1)}\, s_{d-1}(B)$ (resp. greater than or equal to $2^{\mu(d-1)}(s_{d-1}(B) - 2d)$).

Given $\theta = \frac{1}{2}$, set $\alpha = \frac{3}{2} 2^{-\mu}$, then $\rho_n^{(d)} > \alpha$ means that there exists a Borel set B' such that:

$$P_n(B') > P(B'^\alpha) + \alpha.$$

This inequality is preserved when replacing B' with B'' which is the intersection of B' with the support of P_n.

We call then B the Borel set in $\mathscr{B}_\mu^{(d)}$ which is composed by the atoms that intersect B''. As $B''^\alpha \supset B^{\alpha - 2^{-\mu}}$, we have:

$$P_n(B) \geqq P_n(B') > P(B^\varepsilon) + \alpha.$$

Using 2.5 we get:

$$P_n(B) - P(B) > (2^{-\mu}/(2d))(s_{d-1}(B) + 3d - 2d) \geqq (2^{-\mu}/(2d))(1 + s_{d-1}(B))$$

giving (b). The proof of (a) is straightforward using 2.4.

A. The Independent Case

Throughout part A, the observations will be assumed to be independent. For technical reasons of construction of random variables with given distribution it will be convenient to assume that the probability space $(\Omega, \mathscr{A}, \Pr)$ on which the observations are defined is "richenough" in the sense that there exists a random variable defined on $(\Omega, \mathscr{A}, \Pr)$ which is independent of the observations and whose distribution is the Lebesgue measure on $[0, 1]$.

3. Approximation by a Brownian Bridge via the "Hungarian Theorem"

The process $\sqrt{n}(P_n - P)$ is called the empirical Brownian bridge and is denoted by Z_n.

Let D and C denote respectively the space of functions on $[0, 1]$ that are right-continuous and have left-hand limits on the one hand and the space of functions on $[0, 1]$ that are continuous on the other hand. We give C the uniform topology and D the Skorohod topology. Then both spaces are Polish.

A generic point in $D \times C$ is written (φ, ψ); the uniform norm over $[0, 1]$ is denoted by $\|\cdot\|_\infty$. Let us recall the statement of the "Hungarian theorem" (Komlós, Major, Tusnády (1975)).

Theorem 3. *Let $\hat{F}_n$ be the law on D of the distribution function of Z_n (here $d=1$). Let $\hat{B}_0$ be the law on C of a continuous version of a Brownian bridge on $[0, 1]$. Then, there exists a probability law Q_n defined on $D \times C$ (depending on the sample size n) with given marginals $\hat{F}_n$ and $\hat{B}_0$ and some positive constants C, λ and Λ such that, given $H_n(s) = Q_n(\|\varphi - \psi\|_\infty > s)$, we have:*

$$H_n(s) \leqq \Lambda \exp(-\lambda(\sqrt{n}\, s - C\, Ln)),$$

for all positive s.

(According to Bretagnolle and Massart (1987), *we may take $C=12$, $\lambda = 1/6$ and $\Lambda = 2$).*

Before deriving a Gaussian approximation lemma from theorem 3, we state the following definition.

(3.1) *Definition.* Let $\mathscr{J}$ be a family of Borel sets (assumed here to be finite). We say that Z is a *Brownian bridge indexed by* $\mathscr{J}$ if Z is a Gaussian process indexed by $\mathscr{J}$ sucht that $E(Z(B)) = 0$ and $E(Z(B)\, Z(B')) = P(B \cap B') - P(B)\, P(B')$ for any B, B' in $\mathscr{J}$.

From now on H_n is the function that is defined in Theorem 3.

Lemma 2. *There exists a Brownian bridge Z (depending on n) indexed by $\mathscr{B}_\mu$, such that for any positive U:*

$$\Pr\left(\bigvee_{\mathscr{B}_\mu^{(d)}} \frac{|Z_n - Z|}{1 + s_{d-1}} > U\right) \leqq H_n(U\, 2^{-\mu(d-1)}).$$

Proof of Lemma 2. For each cube $c_i = \prod_{j=1}^{d}]i_j 2^{-\mu}, (i_j+1)\, 2^{-\mu}]$, let $T(C_i)$ be the interval $T(G_i) =](i_1 2^{\mu(d-1)} + \ldots + i_d)\ 2^{-\mu d},\ (i_1 2^{\mu(d-1)} + \ldots + i_d + 1)\, 2^{-\mu d}]$. The mapping T extends to $\mathscr{B}_\mu^{(d)}$ using additivity (the atoms of $\mathscr{B}_\mu^{(d)}$ and $\mathscr{B}_\mu^{(1)}$ are ordered with respect to the lexicographical ordering, T preserves this ordering). Given ξ in D, define $\tilde{\xi}(]s,t])=\xi(t)-\xi(s)$ and then $\tilde{\xi}$ on $\mathscr{B}_{\mu d}^{(1)}$ using additivity again. Then, given $T^*: D\times C\to \mathbb{R}^{\mathscr{B}_\mu^{(d)}}\times \mathbb{R}^{\mathscr{B}_\mu^{(d)}}$

$$T^*:\ (\varphi,\psi)\to(\tilde{\varphi}\circ T,\tilde{\psi}\circ T)$$

let QT^{*-1} be the distribution of T^* under Q, where we denote by Q for short, the probability law Q_n which is defined in Theorem 3. The processes that are mentioned below are all indexed by $\mathscr{B}_\mu^{(d)}$. The first marginal of QT^{*-1} is exactly the distribution of Z_n (just because the underlying multinomial distribution is the right one). So, using a lemma from Skorohod (1976), there exists a process Z such that the joint distribution of (Z_n, Z) is QT^{*-1}.

Of course Z is a Brownian bridge. In other respects, we have straightforwardly:

(3.2) $$|\tilde{\xi}(B)|\leqq F(B)\,\|\xi\|_\infty \quad \text{for any } B \text{ in } \mathscr{B}_{d\mu}^{(1)} \text{ and } \xi \text{ in } D,$$

where $F(B)$ stands for the cardinality of the boundary of B. Given $\bar{i}=(i_1,\dots,i_{d-1})$ and B in $\mathscr{B}_\mu^{(d)}$, define

$$B_{\bar{i}}=B\cap\left(\prod_{j=1}^{d-1}\,]i_j 2^{-\mu},(i_j+1)\,2^{-\mu}]\times[0,1]\right).$$

Then $F(T(B_{\bar{i}}))$ is less than or equal to the number of elementary faces (of cubes) composing the boundary of $B_{\bar{i}}$, that are parallel to the hyperplane $y_d=0$. The number of such faces (for all possible values of $\bar{i}$) is not greater than or equal to $2^{\mu(d-1)}\,s_{d-1}(B)$.

Clearly $F(T(B))\leqq\Sigma_{\bar{i}}\,F(T(B_{\bar{i}}))$, so we get that

$$F(T(B))\leqq 2^{\mu(d-1)}\,s_{d-1}(B).$$

Thus, using 3.2:

$$\bigvee_{\mathscr{B}_\mu^{(d)}}\frac{|\tilde{\xi}\circ T|}{1+s_{d-1}}\leqq 2^{\mu(d-1)}\,\|\xi\|_\infty$$

for any ξ in D, giving Lemma 2 via the identity:

$$\Pr\left(\bigvee_{\mathscr{B}_\mu^{(d)}}\frac{|Z_n-Z|}{1+s_{d-1}}>U\right)=Q\left(\bigvee_{\mathscr{B}_\mu^{(d)}}\frac{|(\tilde{\varphi}-\tilde{\psi})\circ T|}{1+s_{d-1}}>U\right)$$

and theorem 3.

4. A Lower Bound for the Prohorov Distance (Proof of Theorem 1 a))

The following estimate yields Theorem 1 a via the Borel-Cantelli lemma.

Proposition 1. *There exists a positive constant* $C_1(d)$ such that:

$$\Pr\left(\left(\frac{n}{Ln}\right)^{1/d}\rho_n^{(d)}<C_1(d)\right)=O(n^{-2}).$$

Proof of Proposition 1. Let Z be a Brownian bridge indexed by $\mathscr{B}_\mu^{(d)}$ approximating Z_n in the sense of lemma 2. Using a usual trick, Z may be written as $Z = W - \zeta P$, where W is a Wiener process indexed by $\mathscr{B}_\mu^{(d)}$ (that is W is a centered Gaussian process indexed by $\mathscr{B}_\mu^{(d)}$ with covariance function $B, B' \to P(B \cap B')$) and ζ is a random variable independent of Z with standard normal distribution.

Given $\bar{\phi}(s) = \Pr(\zeta > s)$ for all s, we set $t = \bar{\phi}^{-1}(2^{-\mu+1})$. An atom C in $\mathscr{B}_\mu^{(d)}$ is said to be "*fair*" if $W(C) > t\, 2^{-\mu d/2}$. Let B^0 be the random set in $\mathscr{B}_\mu^{(d)}$ composed of the $2^{\mu(d-1)}$ first (with respect to the lexicographical ordering) fair atoms, if the number of such atoms is sufficient, otherwise setting $B^0 = \emptyset$.

We call N the number of fair atoms. Let Θ be the event

$$\left\{N \geqq 2^{\mu(d-1)}, \zeta \leqq \frac{2^{\mu d} Ln}{\sqrt{n}}, \bigvee_{\mathscr{B}_\mu^{(d)}} \frac{|Z_n - Z|}{1 + s_{d-1}} \leqq 2C \frac{2^{\mu(d-1)} Ln}{\sqrt{n}}\right\}$$

where the constant C is that of Theorem 3.

But,

$$\frac{Z_n(B^0)}{1 + s_{d-1}(B^0)} \geqq \frac{W(B^0)}{1 + s_{d-1}(B^0)} - 2^{-\mu} \zeta\, \mathbb{1}_{(\zeta \geqq 0)} - \bigvee_{\mathscr{B}_\mu^{(d)}} \frac{|Z_n - Z|}{1 + s_{d-1}}.$$

So, using 2.2, the following holds on Θ:

$$\frac{Z_n(B^0)}{1 + s_{d-1}(B^0)} \geqq 2^{-\mu} \left(\frac{t\, 2^{\mu d/2}}{1 + 2d} - (2C + 1) \frac{2^{\mu d} Ln}{\sqrt{n}}\right).$$

Given a such that $0 < a < \sqrt{\frac{2}{d}}$, we set $\alpha = \left(\frac{a\, 2^{-d/2}}{2(C+1)(1+2d)}\right)^2$ and then choose μ such that $\alpha\, 2^{-d} \frac{n}{Ln} \leqq 2^{\mu d} < \alpha \frac{n}{Ln}$.

The behaviour of $\bar{\phi}$ is well known, so it is easy to see that $t \geqq a \sqrt{Ln}$ for $n \geqq n_0$.

Thus, we have on Θ and for $n \geqq n_0$:

$$\frac{(P_n - P)(B^0)}{1 + s_{d-1}(B^0)} \geqq 2^{-\mu} \left(\frac{t}{1 + 2d} \left(\frac{\alpha\, 2^{-d}}{Ln}\right)^{1/2} - (2C + 1)\, \alpha\right) \geqq \alpha\, 2^{-\mu}.$$

Now we have to bound $\Pr(\Theta^c)$. From Lemma 2 we get:

$$\Pr(\Theta^c) \leqq \Lambda n^{-\lambda C} + \bar{\phi}\left(\frac{2^{\mu d} Ln}{\sqrt{n}}\right) + \Pr(N < 2^{\mu(d-1)})$$

where, according to Theorem 3, we may take $\Lambda = 2$, $C = 12$ and $\lambda = 1/6$. Moreover N has the binomial distribution $\mathscr{B}(2^{d\mu}, 2^{-\mu+1})$, so using an inequality that is due to Okamoto (1958) (see also Hoeffding (1963)), we get:

$$\Pr(N < 2^{\mu(d-1)}) \leqq \exp\left(-\frac{2^{\mu(d-1)}}{4}\right)$$

thus $\Pr(\Theta^c)=O(n^{-2})$.

The conclusion follows from Lemma 1(a) with $\theta=\alpha 2^{-d+1}$ (a suitable choice of a leads to a constant $C_1(d)=2^{-2d}/(d(13+26d)^2)$).

Let us start proving Theorem 1(b). Lemmas 1 and 2 ensure that, in order to solve the initial problem of bounding $\rho_n^{(d)}$, it is enough to solve the following Gaussian problem: controling the quantity $\frac{|Z|}{s_{d-1}}$ uniformly over $\mathscr{B}_\mu^{(d)}$ whenever Z is a Brownian bridge.

5. Homogeneity of the Gaussian Problem

The next lemma establishes a homogeneity principle for the Gaussian problem. This principle brings out that the solution of the Gaussian problem in dimension 2 is the key for bounding $\rho_n^{(d)}$ in higher dimensions.

Lemma 3. *Given an integer $d \geqq 2$, for any version Z^k of a Wiener process indexed by $\mathscr{B}_\mu^{(k)}$ with $k=d-1$ or d, we have, for any positive t:*

$$\Pr\left(\bigvee_{\mathscr{B}_\mu^{(d)}} \frac{|Z^d|}{s_{d-1}} > t\, 2^{\mu/2}\right) \leqq 2^\mu \Pr\left(\bigvee_{\mathscr{B}_\mu^{(d-1)}} \frac{|Z^{d-1}|}{s_{d-2}} > t\right).$$

Proof of lemma 3. For each integer j such that $0 \leqq j < 2^\mu$, let $\mathscr{B}_\mu^{(d)}(j)$ be the σ-algebra composed by the products: $\tilde{B} \times]j\,2^{-\mu},(j+1)\,2^{-\mu}]$, $\tilde{B} \in \mathscr{B}_\mu^{(d-1)}$. Given B in $\mathscr{B}_\mu^{(d)}(j)$ we call $t_j(B)$ the unique Borel set in $\mathscr{B}_\mu^{(d-1)}$ such that $B=t_j(B) \times]j\,2^{-\mu},(j+1)\,2^{-\mu}]$. Let Θ be the event $|Z^d(B)| \leqq t\,2^{-\mu/2}\, s_{d-2} \circ t_j(B)$ for any in $\mathscr{B}_\mu^{(d)}(j)$ and all integers j such that $0 \leqq j < 2^\mu$.

Given B in $\mathscr{B}_\mu^{(d)}$, we introduce the partition:

$$B=\sum_{j=0}^{2^\mu-1} B_j,$$

where B_j is the union of the atoms which compose B, belonging to $\mathscr{B}_\mu^{(d)}(j)$.

An elementary face (of cube) composing the boundary of B_j that is parallel to the last coordinate axis, is also a part of the boundary of B. For each B_j, the total area for such faces is equal to $2^{-\mu}\, s_{d-2} \circ t_j(B)$. Thus, we get:

$$s_{d-1}(B) \geqq 2^{-\mu} \sum_{j=0}^{2^\mu-1} s_{d-2} \circ t_j(B). \tag{5.1}$$

So the following holds on Θ:

$$|Z^d(B)| = \left|\sum_{j=0}^{2^\mu-1} Z^d(B_j)\right| \leqq t \cdot 2^{-\mu/2} \left(\sum_{j=0}^{2^\mu-1} s_{d-2} \circ t_j(B)\right) \leqq t \cdot 2^{\mu/2}\, s_{d-1}(B),$$

from 5.1.

In order to control $\Pr(\Theta^c)$, we notice that, for each j, the processes $(2^{-\mu/2} Z^{d-1}(\tilde{B}))_{\tilde{B}\in\mathscr{B}_\mu^{(d-1)}}$ and $(Z^d(B))_{B\in\mathscr{B}_\mu^{(d)}(j)}$ are isomorphic via the mapping t_j. Then, we have:

$$\Pr(\Theta^c)\leq \sum_{j=0}^{2^\mu-1} \Pr\left(\bigvee_{\mathscr{B}_\mu^{(d)}(j)} \frac{|Z^d|}{s_{d-2}\circ t_j} > t\cdot 2^{-\mu/2}\right) \leq 2^\mu \Pr\left(\bigvee_{\mathscr{B}_\mu^{(d-1)}} \frac{|Z^{d-1}|}{s_{d-2}} > t\right)$$

and the proof of Lemma 3 is complete.

We derive straightforwardly from lemma 3 the following corollary (whose proof will be omitted).

Corollary 1. *Given an integer $d\geq 2$, for any version Z^k of a Wiener process indexed by $\mathscr{B}_\mu^{(k)}$ with $k=d$ or 2, we have for any positive t:*

$$\Pr\left(\bigvee_{\mathscr{B}_\mu^{(d)}} \frac{|Z^d|}{s_{d-1}} > t\cdot 2^{\mu(d-2)/2}\right) \leq 2^{\mu(d-2)} \Pr\left(\bigvee_{\mathscr{B}_\mu^{(2)}} \frac{|Z^2|}{s_1} > t\right).$$

It is worth noticing that the homogeneity principle (lemma 3) does not allow us to derive a sharp upper bound for $\rho_n^{(2)}$ from Dudley's one dimensional result. In fact, up to possible powers of $\log(n)$, such an approach yields an upper bound for $\rho_n^{(2)}$ of order only $n^{-1/3}$. As we shall see later, the sharpest upper bound for $\rho_n^{(2)}$ is of order of $Ln^\alpha\, n^{-1/2}$, with $1/2\leq\alpha\leq 1$.

6. Solution of the Two-Dimensional Gaussian Problem

Throughout this section d is equal to 2. We shall write "$\mathscr{B}_\mu$" instead of "$\mathscr{B}_\mu^{(2)}$" and "l" instead of "s_1" which is here a length.

We have in view to prove the following exponential inequality.

Theorem 4. *Let Z be a Wiener process indexed by $\mathscr{B}_\mu$. For any ε in $]0,1[$, we have:*

$$\Pr\left(\bigvee_{\mathscr{B}_\mu} \frac{|Z|}{l} > t\right) \leq 82\cdot 2^{4\mu} \exp(-8(1-\varepsilon)^2 t^2)$$

whenever $t\geq 100\,\varepsilon^{-3/2}\,\mu$.

The basic ideas of the proof of Theorem 4 are the following:

– The supremum of $\frac{|Z|}{l}$ over $\mathscr{B}_\mu$ is the same as over the smaller class of "Jordan μ-domains" (this notion is defined below).

– For each fixed length λ, the metric entropy of the class of "Jordan μ-domains with perimeter λ" (with respect to the $L^2(P)$-metric) can be calculated: it is critical.

– As l is a discrete function, an exponential control for the supremum of $\frac{|Z|}{\lambda}$ for each possible value λ of l means an exponential control for the supremum of $\frac{|Z|}{l}$.

We now make precise the notion of a "Jordan μ-domain".

(6.1) *Definition.* The class $\mathscr{D}_\mu$ of *Jordan μ-domains* is the class of Borel sets B in $\mathscr{B}_\mu$ such that the interiors (in the usual topological sense) of B and $\mathbb{R}^2 \setminus B$ are both connected.

$\mathscr{D}_\mu$ is extremal for $\frac{|Z|}{l}$ in the following sense.

Proposition 2. *(A maximum principle). Given an integer μ, for any version Z of a Wiener process indexed by $\mathscr{B}_\mu$ we have:*

$$\bigvee_{\mathscr{B}_\mu} \frac{|Z|}{l} = \bigvee_{\mathscr{D}_\mu} \frac{|Z|}{l} \quad \text{a.s.}$$

The proof of Proposition 2 is straightforward (via the elementary inequality $\frac{|\sum a_i|}{\sum |b_i|} \leqq \bigvee_{\mathscr{I}} \left|\frac{a}{b}\right|$ which holds for all finite sequences $(a_i, b_i)_{i \in \mathscr{I}}$ of real numbers) using the additive decomposition lemma stated below.

Lemma 4. *For any Borel set B in $\mathscr{B}_\mu$, there exists a finite collection $(P_i)_{i \in \mathscr{I}}$ of Jordan μ-domains and ε in $\{-1, 1\}^{\mathscr{I}}$* such that:

$$l(B) = \sum_{i \in \mathscr{I}} l(P_i) \tag{6.2}$$

$$\mathbb{1}_B = \sum_{i \in \mathscr{I}} \varepsilon_i \mathbb{1}_{P_i} \qquad P\text{–a.s.} \tag{6.3}$$

Proof of lemma 4. Given B in $\mathscr{B}_\mu$, let $\mathscr{C}$ be the (finite) collection of connected components of the topological interior $\mathring{B}$ of B. Then $\mathring{B} = \sum_{C \in \mathscr{C}} C$, $l(B) = \sum_{C \in \mathscr{C}} l(\bar{C})$ and $\mathbb{1}_B = \sum_{C \in \mathscr{C}} \mathbb{1}_{\bar{C}}$ a.s.

Now let C be one of these connected components; as C is bounded, $\mathbb{R}^2 \setminus \bar{C}$ has a unique unbounded connected component. Removing that component from $\mathbb{R}^2 \setminus \bar{C}$ and calling G the result of this operation, we get two cases:

– either $G = \emptyset$ and in that case $\bar{C}$ is a Jordan μ-domain and we have finished.

– either $G \neq \emptyset$. Then, let $\mathscr{S}(C)$ be the collection of connected components of G.

We have $G = \sum_{S \in \mathscr{S}(C)} S$. Moreover each component S is a Jordan μ-domain as well as $\tilde{C} = \bar{C} \cup \bar{G}$ by construction. Now the boundary of C may be partitioned as follows: the boundary of $\tilde{C}$ which represents the "outside" boundary of C on the one hand, the boundaries of the S's whose union represents the "inside" boundary of C on the other hand; so $l(\bar{C})=l(\tilde{C})+\sum_{S\in\mathscr{S}(C)} l(\bar{S})$ and $\mathbb{1}_{\bar{C}}=\mathbb{1}_{\tilde{C}}-\sum_{S\in\mathscr{S}} \mathbb{1}_S$ a.s.

The proof of Lemma 4 is complete when setting $(P_i)_{i\in\mathscr{I}} =(\tilde{C})_{C\in\mathscr{C}}\cup(\bar{S})_{S\in\mathscr{S}(C),\,C\in\mathscr{C}}$.

Metric Entropy of the Class of Jordan μ-domains with Perimeter λ

Given λ among the possible lengths, that is the $k\,2^{-\mu}$'s when k varies between 4 and $2^{2\mu+1}$, let $\mathscr{D}_\mu^\lambda$ be the class of Jordan μ-domains with perimeter λ. Our approach to calculate the entropy of $\mathscr{D}_\mu^\lambda$ is basically that of Dudley (1974). Before going further in that calculation, we need some notations and definitions.

(6.4) *Notation.* Let S^1 denote the unit circle in $\mathbb{R}^2$. Given $\Lambda=\lambda\,2^\mu$, S^1 is divided in Λ parts of equal length, with successive extremities $\theta_0, \theta_1, \ldots, \theta_\Lambda=\theta_0$.

(6.5) **Definition.** A *μ-curve* is a continuous map C: $S^1\to\mathbb{R}^2$, whose range is composed of edges of atoms of $\mathscr{B}_\mu$; moreover C is called a *Jordan μ-curve* whenever C is one-to-one, furthermore C is said to be *uniform with length λ* whenever C follows the edge of vertices $C(\theta_j)$, $C(\theta_{j+1})$ between time θ_j and time θ_{j+1}, with a constant speed. The collection of uniform Jordan μ-curves with length λ is denoted by $\mathscr{J}_\mu^\lambda$.

It turns out that $\mathscr{D}_\mu^\lambda$ is exactly the collection of the "closed interiors" of the C's belonging to $\mathscr{J}_\mu^\lambda$ (the "closed interior" of a C being the union of the unique bounded connected component of $\mathbb{R}^2\backslash\mathrm{range}(C)$ with its boundary range(C) and there exists a unique uniform Jordan μ-curve with length λ whose range is exactly the boundary of a given set in $\mathscr{D}_\mu^\lambda$, up to the choice of an origin). Thus, the approximation of a Jordan μ-domain will be derived from that of a Jordan μ-curve giving its boundary.

Fitting a ν-Curve to a Uniform Jordan μ-Curve

For what follows, ν is a given integer $0\leqq\nu<\mu$.

Lemme 5. *Let C be a uniform Jordan μ-curve with length λ. A ν-curve (denoted by C_ν) exists such that:*

(6.6) $$\|C-C_\nu\|_\infty<2^{-\nu}.$$

(6.7) *range (C_ν) is composed of at most $2+4\lambda 2^\nu$ edges of atoms of $\mathscr{B}_\nu$.*

Moreover, the cardinality of the collection of the C_ν's when C varies in $\mathscr{J}_\mu^\lambda$, is at most $(2^\nu+1)^2\,9^{2\lambda 2^\nu+1}$.

Proof of Lemma 5. We set $t=2^{\mu-\nu}$. We define a new partition on S^1: $(\tilde{\theta}_0, \ldots, \tilde{\theta}_{k-1}, \tilde{\theta}_k=\tilde{\theta}_0)$ where $\tilde{\theta}_j=\theta_{jt/2}$ for any integer j in $[0, k-1]$, with $k=1+[\lambda 2^{\nu+1}]$. For each j, we call $C_\nu(\tilde{\theta}_j)$ the nearest point to $C(\tilde{\theta}_j)$ on the regular grid of $[0,1]^2$

with mesh $2^{-\nu}$ (in case of ambiguities we choose $C_\nu(\tilde{\theta}_j)$ to be the nearest point of (0, 0)), also:

(i) $$\|C_\nu(\tilde{\theta}_j)-C_\nu(\tilde{\theta}_{j+1})\| \leqq 2^{-\nu}.$$

We now define C_ν between time $\tilde{\theta}_j$ and $\tilde{\theta}_{j+1}$ enforcing the following rules:

- Follow as few edges of atoms of $\mathscr{B}_\nu$ as possible (possibly stay on the spot)
- Because of the first rule and because of (i), C_ν has to follow at most two edges. If exactly two edges have to be followed, choose the horizontal one to be followed first.
- Join $C_\nu(\tilde{\theta}_j)$ to $C_\nu(\tilde{\theta}_{j+1})$ with a constant speed.

The ν-curve C_ν defined above fulfills of course 6.6. Moreover, since C_ν follows at most two edges between time $\tilde{\theta}_j$ and time $\tilde{\theta}_{j+1}$, 6.7 holds.

To count the class $\{C_\nu\colon C\in\mathscr{J}_\mu^\lambda\}$, note that (because of our rules), given $C_\nu(\tilde{\theta}_0), \ldots, C_\nu(\tilde{\theta}_{k-1})$, the ν-curve C_ν is entirely known. Now the number of possible origins $C_\nu(\tilde{\theta}_0)$ is equal to $(2^\nu+1)^2$ and, for each step j, the number of possible values for $C_\nu(\tilde{\theta}_{j+1})$ given $C_\nu(\tilde{\theta}_j)$ is equal to 9 because of (i), completing the proof of lemma 5.

Remark. In general, the ν-curves C_ν in Lemma 5 need not to be one-to-one, so we shall use the quite unusual notion of "interior" of a curve that was introduced and studied by Dudley (1974). We recall his definition in the appendix where we also show that this notion (which has no intrinsic geometrical meaning anymore) extends that of "interior" in the sense of Jordan which we already used above.

The following entropy computation for $\mathscr{D}_\mu^\lambda$ is available.

Lemma 6. *There exists a map $\pi^\nu\colon \mathscr{D}_\mu^\lambda \to \mathscr{B}_\nu$ such that, for any Jordan μ-domain B with perimeter λ, we have $\pi^\nu B\subset B$ and $P(B\backslash\pi^\nu B)\leqq 17\,\lambda\, 2^{-\nu}$. Moreover the cardinality of the range of π^ν is at most $(2^\nu+1)^2\, 9^{2\lambda 2^\nu+1}$.*

Proof of Lemma 6. We write $I(C)$, $J(C)$ for the open or closed interior of a curve C (as defined in the appendix).

Let B be a Jordan μ-domain with perimeter λ whose boundary is given by the uniform Jordan μ-curve C with length λ, then $J(C)=\bar{B}$ (see the appendix). Following lemma 5, let C_ν be a curve fitted to C. Recalling that the ε-interior of a set A is defined by $\complement(\complement A)^\varepsilon$, let $\pi^\nu B$ be the $2^{-\nu}$-interior of $I(C_\nu)$. Then we have (see Dudley (1978) p. 917): $\pi^\nu B\subset B$ and $B\backslash\pi^\nu B\subset(\mathrm{range}(C_\nu))^{2^{-\nu}}$. As in 2.4 we get, since C follows at most $4\lambda 2^\nu+2$ edges of atoms of B_ν: $P((\mathrm{range}(C_\nu))^{2^{-\nu}})\leqq 4\cdot 2^{-2\nu}(4\lambda 2^\nu+2)+2^{-2\nu+1}$, thus $P(B\backslash\pi^\nu B)\leqq 17\,\lambda\, 2^{-\nu}$ whenever $2^{-\nu}\leqq\lambda/10$; otherwise by 2.1, $P(B\backslash\pi^\nu B)\leqq P(B)\leqq(\lambda/4)^2\leqq\lambda\, 2^{-\nu}$. □

Before proving Theorem 4, let us mention a corollary of Lemmas 4 and 6 which will be useful in part B.

Corollary 2. *There exists a map $\pi^\nu\colon \mathscr{B}_\mu\to\mathscr{B}_\nu$ such that $P(B\triangle\pi^\nu B)\leqq 17\, l(B)\, 2^{-\nu}$ for any Borel set B in $\mathscr{B}_\mu$.*

Proof of Corollary 2. Let B be a Borel set in $\mathscr{B}_\mu$. Using lemma 4 there exists a collection $(P_i)_{i\in\mathscr{I}}$ of Jordan μ-domains such that 6.2 and 6.3 hold.

Let π^ν: $\mathscr{D}_\mu \to \mathscr{B}_\nu$ have the properties of Lemma 6 for all possible λ. Then define $\pi^\nu B$ by: $\pi^\nu B = (\bigcup_{\varepsilon_i=1} \pi^\nu P_i) \setminus (\bigcup_{\varepsilon_i=-1} \pi^\nu P_i)$. Since

$$((\bigcup_{\varepsilon_i=1} P_i) \setminus (\bigcup_{\varepsilon_i=-1} P_i))\,\Delta\,((\bigcup_{\varepsilon_i=1} \pi^\nu P_i) \setminus (\bigcup_{\varepsilon_i=-1} \pi^\nu P_i)) \subset \bigcup_i (P_i \triangle \pi^\nu P_i)$$

we get from Lemma 6:

$$P(B \triangle \pi^\nu B) \leqq \sum_i P(P_i \triangle \pi^\nu P_i) \leqq 17 (\sum_i l(P_i))\, 2^{-\nu}$$

giving corollary 2 via 6.2.

The above entropy computation in Lemma 6 allows us to prove Theorem 4 using the so called "chain argument".

Proof of theorem 4. By hypothesis we have:

$$t \geqq 100\, \varepsilon^{-3/2} \mu. \tag{6.8}$$

Given a fixed length λ, let ν_0 be the smallest integer such that either $\nu_0 = \mu$ or $2^{\nu_0} \geqq \dfrac{\varepsilon t^2}{K^2 \lambda}$ holds, where $K = 2\sqrt{L3}$. For each integer ν such that $0 \leqq \nu \leqq \mu$, let π^ν be the map defined in Lemma 6 (π^μ may be taken as the identity map). Let $|\pi^\nu|$ denote the cardinality of the range of π^ν.

The identity $\mathbb{1}_{Id} = \mathbb{1}_{\pi^{\nu_0}} + \sum_{\nu=\nu_0}^{\mu-1} (\mathbb{1}_{\pi^{\nu+1}} - \mathbb{1}_{\pi^\nu})$ leads to the inequality:

$$\Pr(\bigvee_{\mathscr{D}_\mu^\lambda} |Z| > \lambda t) \leqq \mathbb{P}_1 + \mathbb{P}_2 \quad \text{where} \quad \mathbb{P}_1 = |\pi^{\nu_0}| \left(\bigvee_{\mathscr{D}_\mu^\lambda} \Pr\left(|Z \circ \pi^{\nu_0}| > \left(1 - \frac{\varepsilon}{2}\right) \lambda t\right)\right):$$

$\mathbb{P}_1 = 0$ whenever $\nu_0 = 0$ and whenever $(\mu - \nu_0)\, \eta \leqq \frac{\varepsilon}{2} \lambda t$:

$$\mathbb{P}_2 = \sum_{\nu=\nu_0}^{\mu-1} |\pi^\nu|\,|\pi^{\nu+1}| (\bigvee_{\mathscr{D}_\mu^\lambda} \Pr(|Z \circ \pi^{\nu+1} - Z \circ \pi^\nu| > \eta));$$

$\mathbb{P}_2 = 0$ whenever $\nu_0 = \mu$.

We choose $\eta = 4 K \lambda (34/\varepsilon)^{1/2}$, then the latter condition on η is fulfilled because of 6.8.

Control of $\mathbb{P}_1$ ($\nu_0 > 0$)

Given B in $\mathscr{D}_\mu^\lambda$, we get from the isoperimetric inequality 2.1:

$$P(\pi^{\nu_0} B) \leqq P(B) \leqq \lambda^2/16.$$

Besides, whenever ζ is a random variable with standard normal distribution, the following classical inequality is available:

$$\Pr(|\zeta|>s)\leqq 2\exp(-s^2/2) \quad \text{for all } s. \tag{6.9}$$

Then, using lemma 6 and 6.9 we get:

$$\mathbb{P}_1\leqq 18(2^\mu+1)^2\exp(K^2\lambda 2^{\nu_0})\exp\left(-8\left(1-\frac{\varepsilon}{2}\right)^2 t^2\right)$$

thus, since $2^{\nu_0}<2\varepsilon t^2/(K^2\lambda)$,

$$\mathbb{P}_1\leqq 18(2^\mu+1)^2\exp(-8(1-\varepsilon)^2 t^2).$$

Control of $\mathbb{P}_2(\nu_0<\mu)$

From Lemma 6 we get $P(\pi^{\nu+1}B\triangle\pi^\nu B)\leqq 34\lambda\, 2^{-\nu}$ for any B in $\mathscr{D}^\lambda_\mu$, then, using 6.9 again, we have:

$$\mathbb{P}_2\leqq 2\sum_{\nu=\nu_0}^{\mu-1}|\pi^{\nu+1}|^2\exp\left(-\frac{\eta^2}{68\lambda\, 2^{-\nu}}\right),$$

thus, since $|\pi^\nu|\leqq 9(2^\mu+1)^2\exp(K^2\lambda\, 2^\nu)$,

$$\mathbb{P}_2\leqq 81\cdot 2^{4\mu+1}\sum_{j\geqq 0}\exp\left(\frac{2^{\nu_0}K^2\lambda}{\varepsilon}(-8+4\varepsilon)\,2^j\right).$$

As $2^{\nu_0}K^2\lambda/\varepsilon\geqq t^2$ and $2^j\geqq j+1$, we get:

$$\mathbb{P}_2\leqq 162\cdot(2^\mu+1)^2\exp(-(8-4\varepsilon)\,t^2)\leqq\exp(-8(1-\varepsilon)^2 t^2)$$

because of 6.8.

Collecting the above estimates of $\mathbb{P}_1$ and $\mathbb{P}_2$ gives:

$$\mathbb{P}_1+\mathbb{P}_2\leqq 41\cdot 2^{2\mu}\exp(-8(1-\varepsilon)^2 t^2).$$

Noticing that the number of possible values of the function l is at most $2^{2\mu+1}$, we get:

$$\Pr\left(\bigvee_{\mathscr{D}_\mu}\frac{|Z|}{l}>t\right)\leqq 82\cdot 2^{4\mu}\exp(-8(1-\varepsilon)^2 t^2)$$

whenever condition 6.8 is fulfilled; we derive theorem 4 from the above inequality via the maximum principle (proposition 2).

7. An Upper Bound for the Prokhorov Distance (End of the Proof of Theorem 1b)

Using the Borel Cantelli lemma, Theorem 1b derives from the following exponential bound:

Theorem 5.

$$\Pr(\rho_n^{(d)} > 3(36\, d^2\, U(Ln)^2/n)^{1/d}) \leqq 2(1+n^2)\exp(-U(Ln)^2)$$

for any $d \geqq 2$ *and* U *such that*

$$U \geqq (5/3)^3\, 10^4 \tag{7.1}$$

Proof of Theorem 5. Let $\alpha_n = 3(36\, d^2\, U(Ln)^2/n)^{1/d}$. We set $A_n = \Pr(\rho_n^{(d)} > \alpha_n)$. Since $\rho_n^{(d)} \leqq 1$, we may assume that $\alpha_n \leqq 1$. Then, let μ be the integer such that $\frac{3}{2}\, 2^{-\mu} < \alpha_n \leqq 3 \cdot 2^{-\mu}$. Using Lemmas 1 and 2, there exists a Brownian bridge Z indexed by $\mathscr{B}_\mu^{(d)}$ such that:

$$A_n \leqq H_n(\sqrt{n}\, 2^{-\mu d}/(12 d)) + \Pr\left(\bigvee_{\mathscr{B}_\mu^{(d)}} \frac{Z}{s_{d-1}} > 5\sqrt{n}\, 2^{-\mu}/(12 d)\right)$$

then, applying Corollary 1 and Theorem 4 with $\varepsilon = 3/5$ and $t = (2^{-\mu d}\, n)^{1/2}/(6 d)$ we get:

$$A_n \leqq H_n(4 d t^2 n^{-1/2}) + 82 \cdot 2^{\mu(d+2)} \exp(-32\, t^2/25) + \bar{\phi}(3t/2)$$

whenever condition 6.8 is fulfilled. Since $t^2 \geqq U(Ln)^2$ and $\mu \leqq Ln$, condition 7.1 implies 6.8. Moreover, using Theorem 3, we get:

$$A_n \leqq (2n^2 + 82 \cdot 2^{2\mu d} \exp(-7t^2/25) + 1)\exp(-t^2)$$

which leads to theorem 5 by noting that $t^2 \geqq U(Ln)^2$ and using the fact that 7.1 implies:

$$82 \cdot 2^{2\mu d} \exp(-7t^2/25) \leqq 1.$$

B. The Weakly Dependent Case

Throughout part B assumption $H(\alpha, \beta)$ is supposed to be fulfilled. We shall give up here the numbering of the approximating sets in $\mathscr{B}_\nu^{(2)}$ which was an essential point in part A (when $\alpha < 2$, inequality $H(\alpha, \beta)$ does not allow the "chaining" in critical entropy cases). So, we shall take directly advantage of the finite atomic structure of the approximating σ-algebras we are dealing with and of the additivity property of Z_n.

8. Regularization by Changing of Scale

The regularization lemma below derives from Corollary 2 via the following idea: see a Borel set B in $\mathscr{B}_\mu^{(d)}$ as a "stack" of products of a Borel set in $\mathscr{B}_\mu^{(2)}$ with an atom of $\mathscr{B}_\mu^{(d-2)}$, the point is that the perimeter of B is greater than the sum of the "lateral" perimeters of the slices composing the stack; then each slice is regularized via Corollary 2.

Lemma 7. *For any integer v such that $0 \leqq v \leqq \mu$, there exists a map π^v from $\mathscr{B}_\mu^{(d)}$ into $\mathscr{B}_v^{(2)} \otimes \mathscr{B}_\mu^{(d-2)}$ such that the inequality:*

$$P(\pi^v B \triangle B) \leqq 17\, s_{d-1}(B)\, 2^{-v}, \tag{8.1}$$

holds for any Borel set B in $\mathscr{B}_\mu^{(d)}$.

Proof of Lemma 7. Given an atom C of $\mathscr{B}_\mu^{(d-2)}$, we consider the collection $\mathscr{B}_\mu^{(d)}(C)$ of those Borel sets of $\mathscr{B}_\mu^{(d)}$ which are of the form $\tilde{B} \times C$ with $\tilde{B}$ in $\mathscr{B}_\mu^{(2)}$. For each B in $\mathscr{B}_\mu^{(d)}$, let $B(C)$ be the union of the atoms composing B that belong to $\mathscr{B}_\mu^{(d)}(C)$, clearly: $B = \sum_C B(C)$. Besides, each elementary boundary face of $B(C)$ that is orthogonal to the plane $(y_3=0, \ldots, y_d=0)$ is also a part of the boundary of B. The total area of such faces is equal to $2^{-\mu(d-2)}\, l(\tilde{B}(C))$ for each $B(C)$, where $\tilde{B}(C)$ is such that $B(C) = \tilde{B}(C) \times C$. Thus, we get:

$$s_{d-1}(B) \geqq 2^{-\mu(d-2)} \sum_C l(\tilde{B}(C)). \tag{8.2}$$

From Corollary 2, there exists a map: $\tilde{\pi}^v$: $\mathscr{B}_\mu^{(2)} \to \mathscr{B}_v^{(2)}$ with the nice properties stated in the corollary.

Given B in $\mathscr{B}_\mu^{(d)}$, we set $\pi^v B = \sum_C (\tilde{\pi}^v \tilde{B}(C) \times C)$, then $\pi^v B$ belongs to $\mathscr{B}_v^{(2)} \otimes \mathscr{B}_\mu^{(d-2)}$ and we have from Corollary 2:

$$\begin{aligned} P(B \triangle \pi^v B) &\leqq \sum_C P((\tilde{\pi}^v \tilde{B}(C) \times C) \triangle (\tilde{B}(C) \times C)) \\ &\leqq 2^{-\mu(d-2)}\, 17 \cdot 2^{-v} \sum_C l(\tilde{B}(C)). \end{aligned}$$

Thus, lemma 7 follows using 8.2.

9. Proof of Theorem 2

As in Sect. 7, the proof of Theorem 2 is based upon an exponential inequality.

Theorem 6. *There exist some positive constants C and C' such that, for any $d \geqq 2$, any positive U and any integer n, we have:*

$$\Pr\left(\rho_n^{(d)} > \left(\frac{U (Ln)^{2\beta+2+(2/\alpha)}}{n}\right)^{1/d}\right) \leqq C'\, n^{-CU^{\alpha/2}+1}.$$

Let us specify that if c, c' and c'' are the given constants of assumption $H(\alpha, \beta)$, the calculations yield $C' = c' \vee 1$ and $C = (c \wedge 1).\,((c'' \wedge (1/(102.d))) \cdot 3^{-d/2})^\alpha$.

Proof of Theorem 6. Let $\beta_n=\left(\frac{(Ln)^{2\beta+2+(2/\alpha)}\,U}{n}\right)^{1/d}$. The structure of the inequality we intend to prove allows us to assume that $\beta_n\leqq 1$ and $U^{\alpha/2}\geqq C^{-1}$. Then, let μ be the integer such that $\frac{3}{2}\,2^{-\mu}<\beta_n\leqq 3\cdot 2^{-\mu}$. We set $t=\left(\frac{\sqrt{n}}{Ln}\,2^{-\mu d/2}\right).\,(c''\wedge(1/(102.d)))$. From Lemma 1 we get (since $\mu\leqq Ln$):

$$\Pr(\rho_n^{(d)}>\beta_n)\leqq\Pr\left(\bigvee_{\mathscr{B}_\mu^{(d)}}\frac{|Z_n|}{s_{d-1}}>51\,\mu\,2^{\mu(d-2)/2}\,t\right)=A_\mu.$$

In order to bound A_μ, we consider the collection $\mathscr{A}_{\nu,\mu}$ of the atoms of $\mathscr{B}_\nu^{(2)}\otimes\mathscr{B}_\mu^{(d-2)}$ where ν is an integer such that $0\leqq\nu\leqq\mu$. Let Θ be the event defined by:

$$\Theta=\{\bigvee_{\mathscr{A}_{\nu,\mu}}|Z_n|\leqq t\cdot 2^{-\nu}\,2^{-\mu(d-2)/2}\quad\text{for all } \nu,\ 0\leqq\nu\leqq\mu\}$$

($2^{-\nu}\,2^{-\mu(d-2)/2}$ represents the square roof of the Lebesgue measure of each atom of $\mathscr{B}_\nu^{(2)}\otimes\mathscr{B}_\mu^{(d-2)}$).

Note that $t\leqq c''\sqrt{n}\,2^{-\mu d/2}/Ln$, so, using assumption $H(\alpha,\beta)$ (and the inequality: $2^{\mu d}\leqq n$), we get:

$$\Pr(\Theta^c)\leqq c'\,2^{d\mu+1}\exp(-c(t(Ln)^{-\beta})^\alpha)$$

because each $\mathscr{A}_{\nu,\mu}$ has a cardinality equal to $2^{2\nu+\mu(d-2)}$. The point is that, for each B in $\mathscr{B}_2^{(2)}\otimes\mathscr{B}_\mu^{(d-2)}$, the following inequality holds on Θ:

$$(9.1)\qquad |Z_n(B)|\leqq(2^{2\nu}\,2^{\mu(d-2)}\,P(B))\times(t\cdot 2^{-\nu}\,2^{-\mu(d-2)/2})\leqq t\cdot 2^\nu\,2^{\mu(d-2)/2}\,P(B).$$

Now, let π^ν be as in lemma 7, taking π^μ to be the identity map and $\pi^0\colon B\to\emptyset$ (this choice of π^0 satisfies 8.1 because of 2.1). Then, taking advantage of the decomposition $Z_n=\sum_{\nu=0}^{\mu-1}(Z_n\circ\pi^{\nu+1}-Z_n\circ\pi^\nu)$, we get:

$$|Z_n|\leqq\sum_{\nu=0}^{\mu-1}|Z_n\circ(\pi^{\nu+1}\backslash\pi^\nu)|+|Z_n\circ(\pi^\nu\backslash\pi^{\nu+1})|.$$

Now, the ranges of the maps $\pi^{\nu+1}\backslash\pi^\nu$ and $\pi^\nu\backslash\pi^{\nu+1}$ are included in the σ-algebra $\mathscr{B}_{\nu+1}^{(2)}\otimes\mathscr{B}_\mu^{(d-2)}$ for each given ν, thus, from 9.1, we have on Θ:

$$|Z_n|\leqq t\,2^{\mu(d-2)/2}\sum_{\nu=0}^{\mu-1}2^{\nu+1}\,P\circ(\pi^{\nu+1}\triangle\pi^\nu)$$

giving, using 8.2:

$$|Z_n|\leqq 51\,t\mu\cdot 2^{\mu(d-2)/2}\,s_{d-1}\quad\text{on}\quad\Theta.$$

Thus

$$A_\mu \leqq \Pr(\theta^c) \leqq 2c' \, 2^{\mu d} \exp(-c(t(Ln)^{-\beta})^\alpha)$$

so, since $c(t(Ln)^{-\beta})^\alpha \geqq C \, Ln \, U^{\alpha/2}$ and $2^{\mu d} \geqq (3^d n)/U$ we get: $A_\mu \leqq c' \left(\frac{2 \cdot 3^d}{U}\right) n^{-CU^{\alpha/2}}$ leading to theorem 6 via the inequality $U^{\alpha/2} \geqq c^{-1}$. □

Comment. It is worth mentioning that the conclusions of theorem 1b still hold when assuming $H(2,0)$ instead of independence (using the same approach as in Bretagnolle and Massart (1986)), that is to say that the use of the "Hungarian" theorem is not absolutely necessary. Nevertheless the approach that we developed here seems to us to be much more illuminating (especially because of the homogeneity principle (Lemma 3 above)).

Appendix

First let us recall some definitions of algebraic topology.

(A.1) Definition. Let f and g be two maps from a topological space X into another topological space Y.

A *homotopy* of f and g is a continuous map F from $[0,1] \times X$ into Y such that $F(0,\cdot) \equiv f$ and $F(1,\cdot) \equiv g$. f and g are called *homotopic* if there exists an homotopy of f and g. f is *inessential* if it is homotopic to a constant map.

Let the unit Euclidean sphere of $\mathbb{R}^d$ be denoted by S^{d-1}. Following Dudley (1974), the "interior" of a continuous map from S^{d-1} into $\mathbb{R}^d$ can be defined as follows:

(A.2) Definition. If f is a continuous map from S^{d-1} into $\mathbb{R}^d$, the (*open*) *interior* of f is defined by:
$I(f) = \{x \in \mathbb{R}^d \backslash \text{range}(f)$: for any homotopy F of f and a constant map we have $x \in \text{range}(F)\}$.

We call *closed interior* of f the set $J(f) = I(f) \cup \text{range}(f)$. Besides, the definition of the notion of interior in the sense of Jordan is given below.

(A.3) Definition. A continuous one-to-one map from S^{d-1} into $\mathbb{R}^d$ is called a *Jordan hypersurface.*

The *open interior in the sense of Jordan* of a Jordan hypersurface f is the unique bounded connected component of $\mathbb{R}^d \backslash \text{range}(f)$ (the Jordan separation theorem, see Hocking and Young (1961) p. 363, gives a sense to that definition) which we denote by $I'(f)$. Moreover we call *closed interior of f in the sense of Jordan,* the set $J'(f) = I'(f) \cup \text{range}(f)$.

We intend to show that the two notions of interior defined above, are the same for Jordan hypersurfaces. That property comes from the following separation theorem (whose proof can be found in Hocking and Young p. 275).

(A.4) Notations. Let $|\cdot|$ denote the canonical Euclidean norm on $\mathbb{R}^d$. Given $x \in \mathbb{R}^d$, let β_x be the map from $\mathbb{R}^d \backslash \{x\}$ into S^{d-1} such that $\beta_x(y) = \dfrac{y-x}{\|y-x\|}$.

(A.5) Theorem (Borsuk's separation theorem). *Let K be a compact subset of $\mathbb{R}^d$, and let x be a point of $\mathbb{R}^d \setminus K$. For x to lie in the unbounded connected component of $\mathbb{R}^d \setminus K$, it is necessary and sufficient that the mapping $\beta_x|_K$ be inessential as a function into S^{d-1}.*

The first aim of the appendix is the following property which can be derived from the above separation theorem.

(A.6) Corollary. *Let f be a Jordan hypersurface, then $I(f)=I'(f)$ and $J(f)=J'(f)$.*

Proof of Corollary A.6. We set $K=\text{range}(f)$. Clearly, it is enough to prove $I(f)=I'(f)$. Proof of $I'(f)\subset I(f)$.

Let $x\in I'(f)$, if $x\notin I(f)$ it would mean the existence of a homotopy F: $[0,1]\times S^{d-1}\to\mathbb{R}^d\setminus\{x\}$ such that $F(0,\cdot)\equiv f$ and $F(1,\cdot)\equiv x_0$ for some x_0. We consider the map G: $[0,1]\times K\to S^{d-1}$, G: $(t,a)\to\beta_x(F(t,f^{-1}(a)))$. G is a homotopy of $\beta_x|_K$ and $\beta_x(x_0)$. So $\beta_x|_K$ is inessential, a contradiction (with A.5).

– Proof of $I(f)\subset I'(f)$ (or equivalently $\mathbb{R}^d\setminus I'(f)\subset\mathbb{R}^d\setminus I(f)$).

As K is included in both $\mathbb{R}^d\setminus I'(f)$ and $\mathbb{R}^d\setminus I(f)$, it is enough to prove that the unbounded component of $\mathbb{R}^d\setminus K$ is included in $\mathbb{R}^d\setminus I(f)$.

So let x lying in that component. According to theorem A.5, $\beta_x|_K$ is inessential. So there exists a continuous mapping G: $[0,1]\times K\to S^{d-1}$ such that $G(0,\cdot)\equiv\beta_x|_K$ and $G(1,\cdot)\equiv y_0$. We define a continuous mapping F: $[0,1]\times S^{d-1}\to\mathbb{R}^d$ by:

$$F:\ (t,u)\to G(t,u)(t+(1-t)|f(u)-x|)+x,$$

then F is a homotopy of f and y_0+x, moreover $x\notin\text{range}(F)$ because $G(t,u)(t+(1-t)|f(u)-x|)$ cannot be equal to zero.

Thus $x\in\mathbb{R}^d\setminus I(f)$ and the proof is complete.

Proof of the isoperimetric inequality 2.1. We follow Federer (1969) p. 278. Let λ_d denote the Lebesgue measure in $\mathbb{R}^d$. Given $B\in B_\mu^{(d)}$ and $\varepsilon>0$, we have: $B^\varepsilon=B+]-\varepsilon,+\varepsilon[^d$. The Brunn-Minkowski inequality (see Federer (1969) p. 277) then gives on the one hand: $\lambda_d(B^\varepsilon)\geqq(\lambda_d(B)^{1/d}+2\varepsilon)^d\geqq\lambda_d(B)+2d\varepsilon(\lambda_d(B))^{(d-1)/d}$. On the other hand, inequality 2.4 obviously still holds with λ_d instead of P, so: $2d(\lambda_d(B))^{(d-1)/d}\leqq(1+2\varepsilon\,2^\mu)^{d-1}\,s_{d-1}(B)$. Letting ε tend to zero, we get 2.1.

References

1. Bakhvalov, N.S.: On approximate calculation of multiple integrals (in Russian) Vestnik Mosk. Ser. Mat. Mekh. Astron. Fiz. Khim, **4**, 3–18 (1959)
2. Bretagnolle, J., Massart, P.: Classes de fonctions d'entropie critique. C.R. Acad. Sci., Paris, Ser. 1 **302**, 363–366 (1986)
3. Bretagnolle, J., Massart, P.: Hungarian constructions from the non-asymptotic view point. Ann. Probab. (in press)
4. Collomb, G.: Uniform complete convergence of the kernel predictor. Z. Wahrscheinlichkeitstheor. Verw. Geb. **66**, 441–460 (1984)
5. Doukhan, P., Portal, F.: Principe d'invariance faible pour la fonction de répartition empirique, dans un cadre multidimensionnel et mélangeant. Probab. Math. Statist. **8**, 117–132 (1987)

6. Dudley, R.M.: Distances of probability measures and random variables. Ann. Math. Stat. **38**, 1563–1572 (1968)
7. Dudley, R.M.: The speed of mean Glivenko-Cantelli convergence. Ann. Math. Stat. **40**, 40–50 (1969)
8. Dudley, R.M.: Metric entropy of some classes of sets with differentiable boundaries. J. Approximation Theory, **10**, 227–236 (1974)
9. Dudley, R.M.: Central limit theorems for empirical measures. Ann. Probab. **6**, 899–929 (1978): correction **7**, 909–911 (1979)
10. Federer, H.: Geometric measure theory. Berlin Heidelberg: Springer 1969
11. Gäenssler, P.: A note on a result of Dudley on the speed of Glivenko-Cantelli convergence. Ann. Math. Stat. **41**, 1339–1343 (1970)
12. Hocking, J.G., Young, G.S.: Topology. Reading, Mass.: Addison-Wesley (1961)
13. Hoeffding, W.: Probability inequalities for sums of bounded random variables. J. Am. Stat. Assoc. **58**, 13–30 (1963)
14. Komlós, J., Major, P., Tusnády, G.: An approximation of partial sums of independent RV's and the sample D.F.I. Z. Wahrscheinlichkeitstheor. Verw. Geb. **32**, 111–131 (1975)
15. Okamoto, M.: Some inequalities relating to the partial sum of binomial probabilities. Ann. Inst. Stat. Math. **10**, 29–35 (1958)
16. Skorohod, A.V.: On a representation of random variables. Theor. Probab. Appl. **21**, 628–632 (1976)
17. Zuker, M.: Speeds of convergence of random probability measures. Ph. D. dissertation, M.I.T. 1974

Received June 26, 1986; in revised form April 7, 1988