ANNALES DE L'I. H. P., SECTION B

PASCAL MASSART

Rates of convergence in the central limit theorem for empirical processes

Annales de l'I. H. P., section B, tome 22, nº 4 (1986), p. 381-423

<http://www.numdam.org/item?id=AIHPB_1986__22_4_381_0>

Ann. Inst. Henri Poincaré,
Vol. 22, n° 4, 1986, p. 381-423. *Probabilités et Statistiques*

Rates of convergence in the central limit theorem for empirical processes

by

Pascal MASSART
Université Paris-Sud
U. A. CNRS 743 « Statistique Appliquée » Mathématiques, Bât. 425,
91405 Orsay (France)

SUMMARY. — In this paper we study the uniform behavior of the empirical brownian bridge over families of functions $\mathscr{F}$ bounded by a function F (the observations are independent with common distribution P). Under some suitable entropy conditions which were already used by Kolčinskii and Pollard, we prove exponential inequalities in the uniformly bounded case where F is a constant (the classical Kiefer's inequality (1961) is improved), as well as weak and strong invariance principles with rates of convergence in the case where F belongs to $\mathscr{L}^{2+\delta}(\mathrm{P})$ with $\delta \in]0, 1]$ (our results improve on Dudley and Philipp's results (1983) whenever $\mathscr{F}$ is a Vapnik-Červonenkis class in the uniformly bounded case and are new in the unbounded case).

Key-words and phrases: Invariance principles, empirical processes, gaussian processes, exponential bounds.

RÉSUMÉ. — Dans cet article nous étudions le comportement uniforme du pont brownien empirique sur des familles de fonctions $\mathscr{F}$ bornées par une fonction F (les observations sont indépendantes, de distribution commune P). Sous des conditions d'entropie convenables qui sont déjà utilisées par Kolčinskii et Pollard, nous prouvons les inégalités exponentielles dans le cas uniformément borné où F est une constante (l'inégalité classique de Kiefer (1961) est améliorée) aussi bien que les principes d'in-

variance faibles et forts avec des vitesses de convergence dans le cas où F appartient à $\mathcal{L}^{2+\delta}(P)$ avec $\delta \in]0,1]$ (nos résultats améliorent ceux de Dudley et Philipp (1983) lorsque $\mathcal{F}$ est une classe de Vapnik-Červonenkis dans le cas uniformément borné et sont nouveaux dans le cas non-borné.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. GENERALITIES. — Let $(X, \mathcal{X}, P)$ be a probability space and $(x_n)_{n\geq 1}$ be some sequence of independent and identically distributed random variables with law P, defined on a rich enough probability space $(\Omega, \mathcal{A}, \Pr)$.

P_n stands for the empirical measure $\frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}$ and we choose to call empirical brownian bridge relating to P the centered and normalized process $v_n = \sqrt{n}(P_n - P)$. Our purpose is to study the behavior of the empirical brownian bridge uniformly over $\mathcal{F}$, where $\mathcal{F}$ is some subset of $\mathcal{L}^2(P)$.

More precisely, we hope to generalize and sometimes to improve some classical results about the empirical distribution functions on $\mathbb{R}^d$ (here $\mathcal{F}$ is the collection of quadrants on $\mathbb{R}^d$), in the way opened by Vapnik, Červonenkis and Dudley.

In particular, the problem is to get bounds for:

$$\Pr(\| v_n \|_{\mathcal{F}} > t), \quad \text{for any positive } t, \tag{1.1.1}$$

where $\|\cdot\|_{\mathscr{F}}$ stands for the uniform norm over $\mathscr{F}$ and to build strong uniform approximations of ν_n by some regular gaussian process indexed by $\mathscr{F}$ with some speed of convergence, say (b_n).

First let us recall the main known results about the subject in the classical case described above.

1.2. THE CLASSICAL BIBLIOGRAPHY. — We only submit here a succinct bibliography in order to allow an easy comparison with our results (for a more complete bibliography see [*26*]). Concerning the real case ($d = 1$), the results mentioned below do not depend on P and are optimal:

1.2.1. (1.1.1) is bounded by $C\exp(-2t^2)$, where C is a universal constant according to Dvoretzky, Kiefer and Wolfowitz [*24*] ($C \leq 4\sqrt{2}$ according to [*17*]).

1.2.2. The strong invariance principle holds with $b_n = \dfrac{\mathrm{Log}(n)}{\sqrt{n}}$, according to Komlos, Major and Tusnady [*37*].

In the multidimensional case ($d \geq 2$):

1.2.3. (1.1.1) is bounded by $C(\varepsilon)\exp(-(2-\varepsilon)t^2)$, for any $\varepsilon > 0$, according to Kiefer [*34*]. In this expression ε cannot be removed (see [*35*] but also [*28*]).

1.2.4. The strong invariance principle holds with $b_n = n^{-\frac{1}{2(2d-1)}}\mathrm{Log}(n)$, according to Borisov [*8*].

This result is not known to be optimal, besides it can be improved when P is uniformly distributed on $[0, 1]^d$. In this case we have:

1.2.5. If $d = 2$, the strong invariance principle holds with $b_n = \dfrac{(\mathrm{Log}(n))^2}{\sqrt{n}}$, according to Tusnady [*50*].

1.2.6. If $d \geq 3$, the strong invariance principle holds with

$$b_n = n^{-\frac{1}{2(d+1)}}(\mathrm{Log}(n))^{\frac{3}{2}},$$

according to Csörgö and Révész [*14*].

1.2.5 and 1.2.6 are not known to the optimal.

Let us note that even the asymptotic distribution of $\|\nu_n\|_{\mathscr{F}}$ is not well known (the case where $d = 2$ and P is the uniform distribution on $[0, 1]^2$ is studied in [*12*]).

Now we describe the way which has already been used to extend the above results.

1.3. THE WORKS OF VAPNIK, ČERVONENKIS, DUDLEY AND POLLARD. Vapnik and Červonenkis introduce in [*51*] some classes of sets—which are generally called V. Č.-classes—for which they prove a strong Glivenko-Cantelli law of large numbers and an exponential bound for (1.1.1).

P. Assouad studies these classes in detail and gives many examples in [*3*] (see also [*40*] for a table of examples).

The functional P-Donsker classes (that is to say those uniformly over which some central limit theorem holds) were introduced and characterized for the first time by Dudley in [*20*] and were studied by Dudley himself in [*21*] and later by Pollard in [*44*].

Some sufficient (and sometimes necessary, see [*27*] in case $\mathcal{F}$ is uniformly bounded) conditions for $\mathcal{F}$ to be a P-Donsker class used in these works are some kinds of *entropy* conditions, as follows.

Conditions where functions are approximated from above and below (bracketing, see [*20*]) are used in case $\mathcal{F}$ is a P-Donsker class whenever P belongs to some restricted set of laws on X (P often has a bounded density with respect to the Lebesgue measure in the applications) whereas Kolčinskii and Pollard's conditions are used in case $\mathcal{F}$ is a P-Donsker class whenever P belongs to some set of laws including any finite support law (the V. Č.-classes are—under some measurability assumptions—the classes of sets of this kind, see [*21*]).

In our study we are interested in the latter kind of the above classes. Let us recall the already existing results in this particular direction. Whenever $\mathcal{F}$ is some V. Č.-class and under some measurability conditions, we have:

1.3.1. (1.1.1) is bounded by $C(\mathcal{F}, \varepsilon)\exp(-(2-\varepsilon)t^2)$ for any ε in $]0, 1]$, according to Alexander in [*1*] and more precisely by:

$$C(\mathcal{F})(1+t^2)^{2048(D+1)}\exp(-2t^2), \qquad \text{in [2] (}^1\text{)},$$

where D stands for the integer density of $\mathcal{F}$ (from Assouad's terminology in [*3*]).

1.3.2. (1.1.1) is bounded by $4e^8\left(\sum_{j=0}^{D}\binom{n^2}{j}\right)\exp(-2t^2)$, according to Devroye in [*16*].

(1) Our result of the same kind (inequality 3.3.1°) *a*) in the present work) seems to have been announced earlier (in [*41*]) than K. Alexander's.

1.3.3. The strong invariance principle holds with $b_n = n^{-\frac{1}{2700(D+1)}}$, according to Dudley and Philipp in [*23*].

Now let us describe the scope of our work more precisely.

2. ENTROPY AND MEASURABILITY

From now on we assume the existence of a non-negative measurable function F such that $|f| \leq F$, for any f in $\mathscr{F}$. We use in this work Kolčinskii's entropy notion following Pollard [*44*] and the same measurability condition as Dudley in [*21*]. Let us define Kolčinskii's entropy notion.

Let p be in $[1, +\infty[$. $\mathscr{A}(X)$ stands for the set of laws with finite support and $\mathscr{P}_F^{(p)}(X)$ for the set of the laws making F^p integrable.

2.1. DEFINITIONS. — Let ε be in $]0, 1[$ and Q be in $\mathscr{P}_F^{(p)}(X)$. $N_F^{(p)}(\varepsilon, \mathscr{F}, Q)$ stands for the maximal cardinality of a subset $\mathscr{G}$ of $\mathscr{F}$ for which:

$$Q(|f - g|^p) > \varepsilon^p Q(F^p)$$

holds for any f, g in $\mathscr{G}$ with $f \neq g$ (such a maximal cardinality family is called an ε-net of $(\mathscr{F}, F)$ relating to Q). We set $N_F^{(p)}(., \mathscr{F}) = \sup_{Q \in \mathscr{A}(X)} N_F^{(p)}(., \mathscr{F}, Q)$. $\operatorname{Log}(N_F^{(p)}(., \mathscr{F}))$ is called the (p)-entropy function of $(\mathscr{F}, F)$. The finite or infinite quantities:

$$d_F^{(p)}(\mathscr{F}) = \inf \{ s > 0 : \limsup_{\varepsilon \to 0} \varepsilon^s N_F^{(p)}(\varepsilon, \mathscr{F}) < \infty \}$$

$$e_F^{(p)}(\mathscr{F}) = \inf \{ s > 0 : \limsup_{\varepsilon \to 0} \varepsilon^s \operatorname{Log}(N_F^{(p)}(\varepsilon, \mathscr{F})) < \infty \}$$

are respectively called the (p)-entropy dimension and (p)-entropy exponent of $(\mathscr{F}, F)$.

Entropy computations. — We can compute the entropy of $\mathscr{F}$ from that of a uniformly bounded family as follows. Let $\mathscr{I} = \left\{ \frac{f}{F} 1_{(F>0)}, f \in \mathscr{F} \right\}$, then:

$$N_F^{(p)}(., \mathscr{F}) \leq N_1^{(p)}(., \mathscr{I}).$$

For, given Q in $\mathscr{A}(X)$, either $Q(F) = 0$ and so $N_F^{(p)}(., \mathscr{F}, Q) = 1$, or $Q(F) > 0$, so $\frac{F^p}{Q(F^p)} Q \in \mathscr{A}(X)$ and then:

$$N_F^{(p)}(., \mathscr{F}, Q) = N_1^{(p)}\left(., \mathscr{I}, \frac{F^p}{Q(F^p)} Q\right) \leq N_1^{(p)}(., \mathscr{I}).$$

Some other properties of the (p)-entropy are collected in [40]. The main examples of uniformly bounded classes with finite (p)-entropy dimension or exponent are described below.

2.2. COMPUTING A DIMENSION: THE V. Č.-CLASSES. — According to Dudley [20] on the one hand and to Assouad [3] on the other we have $d_1^{(p)}(\mathcal{S}) = pd$ whenever $\mathcal{S}$ is some V. Č.-class with real density d (this notion can be found in [3]). Concerning V. Č.-classes of functions, an analogous computation and its applications are given in [45]. See also [21] for a converse.

2.3. COMPUTING AN EXPONENT: THE HOLDERIAN FUNCTIONS. — Let d be an integer and α be some positive real number. We write β for the greatest integer strictly less than α. Whenever x belongs to $\mathbb{R}^d$ and k to $\mathbb{N}^d$, $|k|$ stands for $k_1 + \ldots + k_d$ and D^k for the differential operator $\partial^{|k|}/\partial x_1^{k_1} \ldots \partial x_d^{k_d}$. Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$. Let $\Lambda_{\alpha,d}$ be the family of the restrictions to the unit cube of $\mathbb{R}^d$ of the β-differentiable functions f such that:

$$\max_{|k| \leq \beta} \sup_{x \in \mathbb{R}^d} |D^k f(x)| + \max_{|k| = \beta} \sup_{x \neq y} \frac{|D^k f(x) - D^k f(y)|}{\|x - y\|^{\alpha - \beta}} \leq 1 .$$

Then, according to [36] on the one hand and using Dudley's arguments in [19] on the other, it is easy to see that: $e_1^{(p)}(\Lambda_{\alpha,d}) = \frac{d}{\alpha}$.

Measurability considerations. — Durst and Dudley give in [21] an example of a V. Č.-class $\mathcal{S}$ such that $\| P_n - P \|_{\mathcal{S}} \equiv 1$. So some measurability condition is needed to get any of the results we have in view. So from now on we assume the following measurability condition (which is due to Dudley [21]) to be fulfilled:

($\mathcal{M}$) . $(X, \mathcal{X})$ is a Suslin space.
. There exists some auxiliary Suslin space $(Y, \mathcal{Y})$ and some mapping T from Y onto $\mathcal{F}$ such that:

$$(x, y) \rightarrow T(y)(x) \quad \text{is measurable on } (X \times Y, \mathcal{X} \otimes \mathcal{Y})$$

and we say that $\mathcal{F}$ is image admissible Suslin via (Y, T).

This assumption is essentially used through one measurable selection theorem which is due to Sion [47] (more about Suslin spaces is given in [13]).

2.4. THEOREM. — Let H be some measurable subset of $X \times Y$. We write A for its projection on X. Then A is universally measurable and there exists a universally measurable mapping from A to Y whose graph is included in H.

A trajectory space for brownian bridges. — We set:

$$l_T^\infty(\mathscr{F}) = \{ h : \mathscr{F} \to \mathbb{R} \,;\, h_0 T \text{ is bounded and measurable on } (Y, \mathscr{Y}) \}.$$

We consider $l_T^\infty(\mathscr{F})$ as a measurable space equipped with the σ-field generated by the open balls relating to $\|\cdot\|_{\mathscr{F}}$ (which is generally distinct from the Borel σ-field because $l_T^\infty(\mathscr{F})$ is not separable).

This trajectory space does not depend on P any more (as it was the case in [*20*]) but only on the measurable representation (Y, T) of $\mathscr{F}$.

From now on for convenience we set:

$$(\Omega, \mathscr{A}, \Pr) = (X' \times [0,1], \hat{\mathscr{X}}' \otimes \mathscr{B}([0,1]), \widehat{P^\infty} \otimes \lambda)$$

where λ stands for the Lebesgue measure on $[0,1]$, $\mathscr{B}([0,1])$ for the Borel σ-field on $[0,1]$ and $(X^\infty, \hat{\mathscr{X}}^\infty, \widehat{P^\infty})$ for the completed probability space of the countable product $(X^\infty, \mathscr{X}^\infty, P^\infty)$ of copies of $(X, \mathscr{X}, P)$. The following theorem points out how $l_T^\infty(\mathscr{F})$ is convenient as a trajectory space.

2.5. THEOREM. — For any a in $\mathbb{R}^n$, $\sum_{i=1}^{n} a_i \delta_{x_i}$ is measurable from Ω to $l_T^\infty(\mathscr{F})$. Moreover, setting $\mathscr{U}_b(\mathscr{F}) = \{ h : \mathscr{F} \to \mathbb{R},\ h$ is uniformly continuous and bounded on $(\mathscr{F}, \rho_P) \}$, $\mathscr{U}_b(\mathscr{F})$ is included in $l_T^\infty(\mathscr{F})$. Provided that $(\mathscr{F}, \rho_P)$ is totally bounded this inclusion is measurable. Here $\mathscr{L}^2(P)$ is given the distance $\rho_P : (f,g) \to \sigma_P(f-g)$, with $\sigma_P^2 : f \to P(f^2) - (P(f))^2$.

For a proof of 2.5, see [*21*] (sec. 9) and [*40*] where it is also shown that many reasonable families (in particular $\Lambda_{\alpha,d}$ and the « geometrical » V. Č.-classes) fulfill $(\mathscr{M})$.

2.6. REMARK. — Since $\mathscr{F}$ fulfills $(\mathscr{M})$ it follows from [*21*] (sec. 12) that $\| P_n - P \|_{\mathscr{F}} \to 0$ a. s. whenever $N_F^{(1)}(.,\mathscr{F}) < \infty$ and therefore:

$$\sup_{Q \in \mathscr{P}_F^{(2)}(X)} N_F^{(2)}(\varepsilon, \mathscr{F}, Q) \leq N_F^{(2)}\left(\frac{\varepsilon}{2}, \mathscr{F}\right) \qquad \text{for any } \varepsilon \text{ in }]0,1[\,.$$

This implies that the local behavior of the entropy function is unchanged when taking the sup in 2.1 over the set of all reasonable laws.

3. EXPONENTIAL BOUNDS FOR THE EMPIRICAL BROWNIAN BRIDGE

We assume in this section that for some constants u and v, $u \leq f \leq v$ for any f in $\mathscr{F}$; we set $U = v - u$ and $\mathscr{F} - u = \{ f - u, f \in \mathscr{F} \}$.

The following entropy conditions are considered:

$$a) \qquad d_U^{(2)}(\mathscr{F} - u) < \infty$$

$$b) \qquad e_U^{(2)}(\mathscr{F} - u) < 2 .$$

Using a single method we find upper bounds for (1.1.1) that are effective in the following two situations:

1°) Observe that $\| \sigma_P^2 \|_{\mathscr{F}} \leq \frac{U^2}{4}$; nothing more is known about the variance over $\mathscr{F}$. In this case we prove some inequalities which are analogous to Hoeffding's inequality [*30*].

2°) We assume that $\| \sigma_P^2 \|_{\mathscr{F}} \leq \sigma^2$. This time our inequalities are analogous to Bernstein's inequality (see Bennett [*5*]).

3.1. Description of the method. — We randomize from a sample whose size is equal to $N = mn$. In Pollard's [*44*], Dudley's [*20*] or Vapnik and Červonenkis' [*51*] symmetrization technique, $m = 2$ but here, following an idea from Devroye [*16*], we choose a large m.

Effecting the change of central law: $P \to P_N$ with the help of a Paul Lévy's type inequality, we may study $\tilde{P}_n - P_N$ instead of $P_n - P$ where $\tilde{P}_n$ stands for the randomized empirical measure.

Choosing some sequence of—measurably selected—nets relating to P_N whose mesh decreases to zero and controlling the errors committed by passing from a net to another via some one dimensional exponential bounds, we can evaluate, conditionally on P_N, the quantity $\| \tilde{P}_n - P_N \|_{\mathscr{F}}$.

Randomization. — Setting $N = nm$ (m is an integer), let w be some random one-to-one mapping from $[1, n]$ into $[1, N]$ whose distribution is uniform (the « sample w is drawn without replacement »).

The inequalities in the next two lemmas are fundamental for what follows:

3.1. LEMMA. — For any ξ in $\mathbb{R}^N$, we set

$$S_N = \sum_{i=1}^{N} \xi_i, \quad \tilde{S}_n = \sum_{i=1}^{n} \xi_{w(i)},$$

$$\sigma_N^2 = \left(\frac{1}{N}\sum_{i=1}^{N} \xi_i^2\right) - \left(\frac{1}{N}\sum_{i=1}^{N} \xi_i\right)^2$$

and $U_N = (\max_{1 \le i \le N}(\xi_i)) - (\min_{1 \le i \le N}(\xi_i))$; the following three quantities are, for any positive ε, lower bounds for $-\operatorname{Log}\left(\frac{1}{2}\operatorname{Pr}\left(\left|\frac{\tilde{S}_n}{n} - \frac{S_N}{N}\right| > \varepsilon\right)\right)$:

1°) $$\frac{2n\varepsilon^2}{U_N^2}$$

2°) $$\frac{n\varepsilon^2}{2\sigma_N^2 + \varepsilon U_N}$$

3°) $$\frac{n\varepsilon^2}{m\sigma_N^2}.$$

These bounds only depend on ξ through numerical parameters (U_N, σ_N). Bound 3°) is new; concerning 1°) (due to Hoeffding [*30*]), Serfling's bound is better (see [*46*]) but brings no more efficiency when m is large.

The proof of lemma 3.1 is given in the appendix.

From now on we write $\tilde{P}_n$ for the randomized empirical process $\frac{1}{n}\sum_{i=1}^{n} \delta_{x_{w(i)}}$. The inequality allowing us to study the randomized process rather than the initial one is the following:

3.2. LEMMA. — The random elements $\| \tilde{P}_n - P_N \|_{\mathscr{F}}$ and $\| P_n - P \|_{\mathscr{F}}$ are measurable.

Besides, whenever $\| \sigma_P^2 \|_{\mathscr{F}} \le \rho^2$, the following holds:

$$\left(1 - \frac{\rho^2}{a^2\varepsilon^2 n'}\right)\operatorname{Pr}(\| P_n - P \|_{\mathscr{F}} > \varepsilon) \le \operatorname{Pr}\left(\| \tilde{P}_n - P_N \|_{\mathscr{F}} > (1-a)\frac{n'}{N}\varepsilon\right)$$

for any positive ε and any a in $]0,1[$, where $n' = N - n$.

For a proof of this lemma see [*16*] using Dudley's measurability arguments in [*21*] (sec. 12).

Statements of the results.

3.3. THEOREM. — The following quantities are, for any positive t and η, upper bounds for $\Pr(\| v_n \|_{\mathscr{F}} > t)$:

1°) *a*) if $d_U^{(2)}(\mathscr{F} - u) = 2d$, $0_{\eta,\mathscr{F}}(1)\left(1 + \frac{t^2}{U^2}\right)^{3(d+\eta)} \exp\left(-2\frac{t^2}{U^2}\right)$;

b) if $e_U^{(2)}(\mathscr{F} - u) = \zeta < 2$, $0_{\eta,\mathscr{F}}(1)\exp\left(0_{\eta,\mathscr{F}}(1)\left(\frac{t}{U}\right)^{k+\eta}\right)\exp\left(-2\frac{t^2}{U^2}\right)$, where $k = \zeta\left(\frac{6-\zeta}{2+\zeta}\right)$ (when ζ increases from 0 to 2 so does k).

2°) Suppose that $\| \sigma_P^2 \|_{\mathscr{F}} \leq \sigma^2$, with $\sigma \leq U$, then

a) if $d_U^{(2)}(\mathscr{F} - u) = 2d$,

$$0_{\eta,\mathscr{F}}(1)\left(\frac{\sigma}{U}\right)^{-4(d+\eta)}\left(1 + \frac{t^2}{\sigma^2}\right)^{3(d+\eta)} \exp\left(-\frac{t^2}{2\left(\sigma^2 + \frac{U}{\sqrt{n}}(3U + t)\right)}\right);$$

b) if $e_U^{(2)}(\mathscr{F} - u) = \zeta < 2$,

$$0_{\eta,\mathscr{F}}(1)\exp\left(0_{\eta,\mathscr{F}}(1)\left(\frac{\sigma}{U}\right)^{-\zeta-\eta}\left(\frac{t}{\sigma}\right)^{2p-\zeta+\eta} + 5\left(\frac{t}{\sigma}\right)^{2p+\eta}\right)\exp\left(-\frac{t^2}{2\left(\sigma^2 + \frac{U}{\sqrt{n}}\left(3U\left(\frac{t}{\sigma}\right)^{p+\eta} + t\right)\right)}\right)$$

where $p = \frac{2\zeta(4-\zeta)}{4 + \zeta(4-\zeta)}$ (when ζ increases from 0 to 2 so does $2p$).

The constants appearing in these bounds depend on $\mathscr{F}$ only through $N_U^{(2)}(., \mathscr{F} - u)$ and of course on η.

Comments.

— Yukich also used in [*54*] Kolčinskii-Pollard entropy notion to prove an analogue of theorem 3.3, but our bounds are sharper because of the use of a randomization from a large sample as described in 3.1.

— From section 2.2, the assumption $d_1^{(2)}(\mathscr{F}) < \infty$ is typically fulfilled whenever $\mathscr{F}$ is some V. Č.-class with real density d.

Thus bound 1°) *a*) is sharper than those of 1.3.1; in an other connection the factor $0(\mathscr{F}, \eta)t^{6(d+\eta)}$ in 1°) *a*) is specified in the appendix.

— In the classical case (i. e. $\mathscr{F}$ is the collection of quadrants in $\mathbb{R}^d$), bound 1°) *a*) improves on 1.2.3 but is less sharp than 1.2.1 in the real case; moreover the optimality of 1°) *a)* is discussed in the appendix where it is shown that

$$\lim_{n\to\infty} \Pr(\| v_n \|_{\mathscr{F}} > t) \geq 2 \sum_{i=0}^{d-1} \frac{(2t^2)^i}{i!} \exp(-2t^2).$$

So, there is a gap for the degree of the polynomial factor in bound 3.3.2°) *a)* between $2(d-1)$ and $6(d+\eta)$.

— Suppose that $\mathscr{F} = \Lambda_{\alpha,d}$ then, from section 2.3 we have $e_1^{(2)}(\mathscr{F}) = \frac{d}{\alpha}$. In other respects, Bakhvalov proves in [*4*] that if P stands for the uniform distribution on $[0, 1]^d$ then:

$$\| v_n \|_{\mathscr{F}} \geq C n^{\frac{1}{2} - \frac{\alpha}{d}} \quad \text{surely}.$$

Thus we cannot get any inequality of the 1°) or 2°) type in the situation where $e_1^{(2)}(\mathscr{F}) > 2$.

The border line case. — For any modulus of continuity ϕ, we can introduce a family of functions $\Lambda_{\phi,d}$ in the same way as $\Lambda_{\alpha,d}$ by changing $u \to u^\alpha$ into ϕ and defining β as the greatest integer for which $\phi(u)u^{-\beta} \to 0$ as $u \to 0$.

It is an easy exercise, using Bakhvalov's method, to show that:

$$\| v_n \|_{\Lambda_{\alpha,d}} \geq C(\text{Log}(n))^\gamma$$

provided that $\phi(u) = u^{d/2}(\log(u^{-1}))^\gamma$ and P is uniformly distributed on $[0, 1]^d$. Of course $e_1^{(2)}(\Lambda_{\phi,d}) = 2$ and we cannot get bounds such as in theorem 3.3.

But the above result is rather rough and we want to go further in the analysis of the families $\Lambda_{\phi,1}$ around the border line.

Then the (2)-entropy plays the same role for $\Lambda_{\phi,1}$ concerning the Donsker property as the metric entropy in a Hilbert space for the Hilbert ellipsoids concerning the pregaussian property, that is to say that the following holds:

i) $\Lambda_{\phi,1}$ is a functional P-Donsker class whenever

$$\int_0^1 (\text{Log}(N_1^{(2)}(\varepsilon, \Lambda_{\phi,1})))^{1/2} d\varepsilon < \infty.$$

ii) $\Lambda_{\phi,1}$ is not a functional λ-Donsker class whenever

$$\phi(u) \simeq \left(\frac{u}{|\operatorname{Log}(u)|}\right)^{1/2} \qquad (^2)$$

and in this case we have $\operatorname{Log}(N_1^{(2)}(\varepsilon, \Lambda_{\phi,1})) \simeq (\varepsilon \operatorname{Log}(\varepsilon))^{-2}$.

i) follows from Pollard's central limit theorem in [*44*].

ii) follows from a result of Kahane's in [*32*] about Rademacher trigonometric series. In fact, if we set $\phi(u) = \sqrt{\frac{u}{|\operatorname{Log}(u)|}}$, we have from [*32*] p. 66 that: $t \to \sum_{n \geq 1} \frac{\varepsilon_n e_n(t)}{Kn \operatorname{Log}(n)}$ belongs to $\Lambda_{\phi,1}$ with some probability $p_K \to 1$ as $K \to \infty$, where (ε_n) is a Rademacher sequence and $e_n(t) = \sqrt{2}\cos(2\pi n t)$.

Let us consider a standard Wiener process on $L^2([0, 1])$, we may write $(W(e_n))$ as $(\varepsilon_n | W(e_n)|)$ with (ε_n) being independent of $(|W(e_n)|)$.

Then, with probability more than p_K, the following holds:

$$\| W \|_{\Lambda_{\phi,1}} \geq \frac{1}{K} \cdot \sum_{n \geq 1} | W(e_n) | \frac{1}{n \operatorname{Log}(n)}.$$

By the three series theorem the series $\Sigma \, | W(e_n) | \frac{1}{n \operatorname{Log}(n)}$ diverges to infinity almost surely and therefore W is almost surely unbounded on $\Lambda_{\phi,1}$.

The same property holds for any brownian bridge G, for $f \to G(f) + \int f W(1)$ is some Wiener process provided that W(1) is some N(0, 1) random variable independent of G. So $\Lambda_{\phi,1}$ is not pregaussian and *ii)* is proved.

An upper bound in situation 2°) is also an oscillation control. — If we set $\mathscr{G}_\sigma = \{ f - g ; \sigma_P(f - g) < \sigma, f, g \in \mathscr{F} \}$, it is not difficult to see that:

3.4 $$N_{2U}^{(2)}(., \mathscr{G}_\sigma + U) \leq (N_U^{(2)}(., \mathscr{F} - u))^2 .$$

Thus changing U into 2U and d into $2d$ if necessary the upper bounds in situation 2°) hold with $\mathscr{G}_\sigma$ instead of $\mathscr{F}$, the constants being independent of σ because of 3.4.

In particular if $\mathscr{F}$ is a V. Č.-class with real density d, we set:

$$A(\sigma, n, t) = \Pr(\| \nu_n \|_{\mathscr{G}_\sigma} > t).$$

(^{2}) We write $f \simeq g$, when $0 < \underline{\lim}(fg^{-1}) \leq \overline{\lim}(fg^{-1}) < \infty$.

As it is summarized in [23], Dudley shows in [20] that $A(\sigma, n, t) \leq t$ whenever t is small enough, $\sigma = 0\left(\frac{t}{|\operatorname{Log}(t)|}\right)$ and $n \geq 0(t^{-r})$ with $r > 8$.

Applying 3.3.2°) *a*) improves on this evaluation for then:

$A(\sigma, n, t) \leq t$ whenever t is small enough, $\sigma = 0\left(\frac{t}{|\operatorname{Log}(t)|}\right)$ and

$$n \geq 0\left(\left(\frac{t}{|\operatorname{Log}(t)|}\right)^{-4}\right).$$

In order to specify in what way the constant in bound 2°) *a*) depends on $\mathscr{F}$, we indicate the following variant of 3.3.2°) *a*).

3.5. PROPOSITION. — If we assume that $N_U^{(2)}(\varepsilon, \mathscr{F} - u) \leq C(\varepsilon_0 \varepsilon)^{-2d}$ for any ε in $]0, 1[$ and some ε_0 in $]0, 1[$ and that $\| \sigma_P^2 \|_{\mathscr{F}} \leq \sigma^2$ with σ not exceeding U, then there exists some ε_1 in $]0, 1[$ depending only on ε_0 and a constant K depending only on C such that:

$$\Pr(\| v_n \|_{\mathscr{F}} > t) \leq K \varepsilon_1^{-d} \left(\frac{\sigma}{U}\right)^{-4d} \left(1 + \frac{t^2}{\sigma^2}\right)^{14d} \exp\left(- \frac{t^2}{2\left(\sigma^2 + \frac{U(3U + t)}{\sqrt{n}}\right)}\right).$$

From now on L stands for the function $x \to \max(1, \operatorname{Log}(x))$.

3.6. COROLLARY. — Let $(\mathscr{F}_n)$ be some sequence of V. Č.-classes fulfilling $(\mathscr{M})$ with integer densities (D_n). Then (with the above notations) $\Pr(\| v_n \|_{\mathscr{G}_{\sigma(n)}} > t) \to 0$ as $n \to \infty$ for any positive t whenever $\sigma(n)^2 = o(1/(D_n L(D_n)))$ and $\sigma(n)^{-2} = 0(\sqrt{n})$.

$\left(\text{Provided that } D_n = o\left(\frac{\sqrt{n}}{Ln}\right), \text{ such a choice of } \sigma(n) \text{ does exist}\right)$.

Comment. — According to Le Cam [38] (Lemma 2) and applying 3.6 the process $\{ v_n(f), f \in \mathscr{F}_n \}$ admits finite dimensional approximations whenever $D_n = o\left(\frac{\sqrt{n}}{\operatorname{Log}(n)}\right)$ and provided that Le Cam's assumption (A1) is fulfilled.

This result improves on Le Cam's corollary of proposition 3 where $D_n = 0(n^{-\gamma})$ for some $\gamma < \frac{1}{2}$ is needed.

Proof of 3.6. — Let $\mathscr{F}$ be a V. Č.-class with entire density D and real density d. Using Dudley's proof in [20] (more details are given in [40]) it is easy to show that, for any $w > d$ (or $w \geq d$ if d is « achieved »), we have:

$$N_1^{(2)}(\varepsilon, \mathscr{F}) \leq K^{1+(1/2|\mathrm{Log}\,\varepsilon|)} \exp(2w)(1 + 2\,|\,\mathrm{Log}(\varepsilon)\,|)^w \varepsilon^{-2w}$$

for any ε in $]0, 1[$, with in particular when $w = D$, $K = \frac{3}{2D!}(2D)^{2D}$. So from Stirling's formula we get: $N_1^{(2)}(\varepsilon, \mathscr{F}) \leq C_1^3 e^{5D} 2^{3D} \varepsilon^{-4D}$ for any ε in $\left]0, \frac{1}{\sqrt{2}}\right]$ and some universal constant C_1. Hence. for any ε in $]0, 1[$ we have:

$$N_1^{(2)}(\varepsilon, \mathscr{F}) \leq C_1^3 (2e)^{5D} \varepsilon^{-4D}.$$

Thus, applying 3.5 to the class $\mathscr{G}_{\sigma(n)}$ yields 3.6.

We propose below another variant of inequality 3.3.2°) *a*), providing an alternative proof of a classical result about the estimation of densities.

3.7. PROPOSITION. — If we assume that $d_U^{(2)}(\mathscr{F} - u) = 2d < \infty$ and $\| \sigma_P^2 \|_{\mathscr{F}} \leq \sigma^2$ with $\frac{UV}{\sqrt{n}} \leq \sigma \leq U$ for some positive V, then there exists some positive constant C such that an upper bound for $\Pr(\| \nu_n \|_{\mathscr{F}} > t)$ is, for any positive t, given by:

$$0_{V,\mathscr{F},\eta}(1)\left(\eta\left(1 + \frac{t^2}{\sigma^2}\right)\right)^{3(d+\eta)}\left(\frac{\sigma}{U}\right)^{-4(d+\eta)} \exp - \frac{t^2}{2\left(\sigma^2 + \frac{U}{\sqrt{n}}\left(\mathrm{CLL}n\left(\frac{U}{\sqrt{n}} + \sigma\right) + t\right)\right)}.$$

In the situation where U is large this inequality may be more efficient than 3.3.2°) *a*).

Application to the estimation of densities: minimax risk.

Let K_M be the following kernel on $\mathbb{R}^k$: $K_M(y) = \psi(y'My)$ for any y in $\mathbb{R}^k$, where ψ is some continuous function with bounded variation from $\mathbb{R}$ into $\left[-\frac{1}{2}, \frac{1}{2}\right]$ and M is some $k \times k$ matrix. Pollard shows in [*45*] that the class

$$\mathscr{K} = \{ K_M(\cdot - x), \quad M \in \mathbb{R}^{k^2}, \quad x \in \mathbb{R}^k \}$$

is a V. Č.-class of functions and so:

$$N_1^{(2)}(\varepsilon, \mathcal{K}) \leq C\varepsilon^{-w} \quad \text{for any } \varepsilon \text{ in }]0, 1[$$

where C and w depend only on k.

Now if we assume that P is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^k$, the classical kernel estimator of its density f is:

$$\hat{f}_n(x) = h^{-k} P_n\left(K\left(\frac{\cdot - x}{h}\right)\right)$$

where K is a K_M with fixed ψ and M so that $\int K^2(x)dx < \infty$. We set $\bar{f} = E(\hat{f}_n)$. Proposition 3.7 gives a control of the *random expression* $\hat{f}_n - \bar{f}$ by choosing:

$$\mathcal{F} = \left\{ h^{-k/2} K\left(\frac{\cdot - x}{h}\right), x \in \mathbb{R}^k \right\}, \quad \sigma = C \quad \text{and} \quad U = h^{-k/2}$$

where $C^2 > \| f \|_\infty \int K^2(x)dx$.

So, if we assume that $\frac{n}{Ln} \geq h^{-k} \geq C^2$, we get, setting $D_n = \sup_x |\hat{f}_n(x) - \bar{f}(x)|$:

$$\Pr(\sqrt{nh^k} D_n > t) \leq 0(n^\alpha) t^\beta \exp\left(- \frac{t^2}{2(C^2 + 0(LLn/\sqrt{nh^k}) + t/\sqrt{nh^k})} \right)$$

for any t in $[1 + \infty[$ and some positive α and β. Hence, after an integration:

$$E(\sqrt{nh^k} D_n) \leq T + 0(n^\alpha) T^\beta \exp\left(- \frac{T^2}{2(C^2 + 0(LLn/\sqrt{nh^k}) + T/\sqrt{nh^k})} \right)$$

for any T in $[1, +\infty[$, provided that $nh^k \geq 4\beta^2$. We choose $T = 0(\sqrt{Ln})$, thus:

$$E(D_n) = 0(Ln/(nh^k))^{1/2}.$$

Provided that f belongs to some subset of regular functions Θ, the *bias expression* $\bar{f} - f$ can be evaluated so that the minimax risk associated to the uniform distance on $\mathbb{R}^k$ and to Θ can be controlled with the same speed of convergence as in [*29*], via an appropriate choice of h.

3.8. PROOFS OF 3.3, 3.5, 3.7. — Lemma 3.1 is proved in the appendix. Let us prove theorem 3.3. First, we reduce the problem to the case where $u = 0$ and $v = 1$ by studying the class $\mathcal{G} = \left\{ \frac{f - u}{U}, f \in \mathcal{F} \right\}$ instead of $\mathcal{F}$; $\mathcal{G}$ fulfills ($\mathcal{M}$) and $N_U^{(2)}(., \mathcal{F}) = N_1^{(2)}(., \mathcal{G})$. Finally $\| v_n \|_{\mathcal{F}} = U \| v_n \|_{\mathcal{G}}$ and $\| \sigma_P^2 \|_{\mathcal{F}} = U^2 \| \sigma_P^2 \|_{\mathcal{G}}$.

In the course of the proof we will need to introduce some parameters such as: a, μ (in $]0, 1[$); r, m (in $\mathbb{N}$); α (in $]1, +\infty[$); q (in $]0, 2]$) and some positive s, β and γ. Let (τ_j) be a positive sequence decreasing to zero. These parameters will be chosen in due time, sometimes differently in different cases. We set $N = mn$, we write $\Pr^{(N)}(\cdot)$ for the probability distribution conditional on $(x_1, \ldots, x_N)$ and $\| \cdot \|$ instead of $\| \cdot \|_{\mathcal{F}}$ for short. We set $\varepsilon = \frac{t}{\sqrt{n}}$ and $\varepsilon' = \left(1 - \frac{1}{m}\right)(1 - a)\varepsilon$. A bound for $\Pr(\| v_n \| > t)$ will follow, via lemma 3.2, from a bound for $\Pr(\| \tilde{P}_n - P_N \| > \varepsilon')$ which is at first performed conditionally on $(x_1, \ldots, x_N)$.

Conditional approximation by a series of projections.

For each integer j a τ_j-net $\mathcal{F}_j$ can be measurably selected (with the help of 2.4, see [*21*] p. 120). So we can define a projection Π_j from $\mathcal{F}$ onto $\mathcal{F}_j$ such that on the one hand $P_N((\Pi_j(f) - f)^2) \leq \tau_j^2$ holds for any f in $\mathcal{F}$ and on the other hand $(\tilde{P}_n - P_N) \circ \Pi_j$ belongs to $l_T^\infty(\mathcal{F})$. We show that the development:

$$(\tilde{P}_n - P_N) \circ \mathrm{Id} - \Pi_r) = \sum_{j \geq r+1} (\tilde{P}_n - P_N) \circ (\Pi_j - \Pi_{j-1})$$

holds uniformly over $\mathcal{F}$ and over the realizations of $(x_{w(i)})_{1 \leq i \leq n}$

In fact, because each realization of w is one-to-one we have:

$$\tilde{P}_n(g) \leq m P_N(g) \quad \text{for any positive function } g \text{ defined on X.}$$

Hence:

$$\| (\tilde{P}_n - P_N) \circ (\mathrm{Id} - \Pi_J) \| \leq (2 \sup_{f \in \mathcal{F}} (\tilde{P}_n + P_N)(\Pi_J(f) - f)^2)^{1/2} \leq \sqrt{2(1+m)}\tau_J$$

for any J in $\mathbb{N}^*$.

Therefore we get:

$$\| (\tilde{P}_n - P_N) \circ (\mathrm{Id} - \Pi_r) \| \leq \sum_{j \geq r+1} \| (\tilde{P}_n - P_N) \circ (\Pi_j - \Pi_{j-1}) \| .$$

So, provided that (η_j) is a positive series such that $\sum_{j \geq r+1} \eta_j \leq \mu$, we have

$$\Pr^{(N)}(\| \tilde{P}_n - P_N \| > \varepsilon') \leq \Pr^{(N)}(\| (\tilde{P}_n - P_N) \circ \Pi_r \| > (1 - \mu)\varepsilon')$$
$$+ \sum_{j \geq r+1} \Pr^{(N)}(\| (\tilde{P}_n - P_N) \circ (\Pi_j - \Pi_{j-1}) \| > \eta_j \varepsilon') \qquad P^{\otimes N}\text{-a. s.}$$

But the cardinalities of the ranges of Π_j and $\Pi_j - \Pi_{j-1}$ are respectively not greater than N_j and $N_{j-1}N_j$ (where N_j stands for $N_1^{(2)}(\tau_j, \mathscr{F})$) so that:

$$\Pr{}^{(N)}(\| \tilde{P}_n - P_N \| > \varepsilon') \leq A + B \qquad P^{\otimes N}\text{-a. s.}$$

where A and B are the $(x_1, \ldots, x_N)$-measurable variables:

$$A = N_r \| \Pr{}^{(N)}(|(\tilde{P}_n - P_N) \circ \Pi_r| > (1 - \mu)\varepsilon') \| ,$$

$$B = \sum_{j \geq r+1} N_j^2 \| \Pr{}^{(N)}(|(\tilde{P}_n - P_N) \circ (\Pi_j - \Pi_{j-1})| > \eta_j \varepsilon') \| .$$

A is the principal part of the above bound and B is the sum of the errors. We use lemma 3.1 to bound A and B: inequalities 1°) or 2°) are needed to control A according to whether case 1°) or 2°) is investigated. Setting $t' = \sqrt{n}\varepsilon'$, we use bound 3°) to control B, so:

$$B \leq 2 \sum_{j \geq r+1} N_j^2 \exp\left(- \frac{t'^2 \eta_j^2}{4m\tau_{j-1}^2} \right) \tag{3.8.1}$$

for $P_N((\Pi_j(f) - \Pi_{j-1}(f))^2) \leq (\tau_j + \tau_{j-1})^2 \leq 4\tau_{j-1}^2$ holds for any f in $\mathscr{F}$.

We choose $\eta_j = (j-1)^{-\alpha}$ and $r = 2 + [\mu^{-1/(\alpha-1)}]$ $\left(\text{so } \sum_{j \geq r+1} \eta_j \leq \mu\right.$ holds whenever $\alpha \geq 2$).

The control of the tail of the series in 3.8.1 is performed via the following elementary lemma:

3.8.2. LEMMA. — Let $\psi : [r, +\infty[\to \mathbb{R}$. Provided that ψ is an increasing convex function, the following inequality holds:

$$\sum_{j \geq r+1} \exp(-\psi(j)) \leq \frac{1}{\psi'_d(r)} \exp(-\psi(r))$$

where ψ'_d stands for the right-derivative of ψ.

HINT: Note that $\int_r^\infty \exp(-\psi(x))dx \leq (\psi'_d(r))^{-1} \int_r^\infty \psi'_d(x) \exp(-\psi(x))dx.$

In each case it will be enough to prove the inequality for $t \geq t_0 = t_0(\eta, \mathscr{F})$ for some t_0. We choose $\beta = 1$ under assumption *a*) and $\beta = \left(\frac{\zeta}{2 - \zeta}\right)\alpha$ under assumption *b*).

Proof of theorem 3.3 in case 1°). — Applying 3.1.1°) we get:

$$A \leq 2N_r \exp(-2t'^2(1-\mu)^2) \qquad P^{\otimes N}\text{-a. s.}$$

We choose: $a = t^{-2}$, $m = [t^2]$ and $\tau_j = \frac{1}{\sqrt{m}} j^{-(\alpha+\beta)}$.

If we set $\rho = \frac{1}{2}$, the variance factor $1 - \frac{\rho^2}{a^2 \varepsilon^2 n'}$ in lemma 3.2 is stabilized for it is not less than $\frac{1}{2}$ whenever $t \geq 2$.

Besides, whenever $t^2 \geq 3$, $t'^2 \geq t^2 - 5$ hence:

$$A \leq 2N_r e^{10} \exp(-2t^2(1 - 2\mu)) \qquad P^{\otimes N}\text{-a. s.}$$

Under assumption a). — If we prove that for any given positive η' and d' such that $N_1^{(2)}(\nu, \mathscr{F}) \leq C(d')\nu^{-2d}$ holds for any ν in $]0, 1[$, an upper bound for $\Pr(\| \nu_n \| > t)$ is given by: $K(d', \eta')(1 + t^2)^{3(d'+\eta')} \exp(-2t^2)$. Then, setting $d' = d + \frac{\eta}{2}$ and $\eta' = \frac{\eta}{2}$ we get 3.3.1°) *a)*.

So, writing « d, η » instead of « d', η' » for short, we may assume that $N_j \leq Ct^{2d} j^{2(\alpha+1)d}$. We choose $\mu = t^{-2}$ and $\alpha = \text{Max}\left(2, 1 + \frac{4d}{3\eta}\right)$, so:

$$A \leq 0_{\eta, \mathscr{F}}(1)(1 + t^2)^{3(d+\eta)} \exp(-2t^2) \qquad P^{\otimes N}\text{-a. s.}$$

In order to evaluate B we apply lemma 3.8.2, setting:

$$\psi(x) = \frac{t'^2}{4}(x - 1)^2 - 4d(\alpha + 1) \operatorname{Log}(x),$$

then the condition $t^2 \geq 7 + 4d(\alpha + 1)$ (which can be assumed) ensures that $\psi'(r) \geq 1$, hence

$$B \leq 2C^2 t^{4d}(2 + t^{2/(\alpha-1)})^{4d(\alpha+1)} \exp\left(-\frac{t'^2}{4}(r - 1)^2\right) \qquad P^{\otimes N}\text{-a. s.}$$

As $r \to \infty$ whenever $t \to \infty$, we get:

$$B \leq 0_{\eta, \mathscr{F}}(1) \exp(-2t^2) \qquad P^{\otimes N}\text{-a. s.}$$

But the above estimates of A and B are deterministic, so, using lemma 3.2, theorem 3.3 is proved in situation 1°) *a)*.

With the idea of proving proposition 3.5 we remark that, setting $\alpha = 2$, the above method gives, under the hypothesis in 3.5, that $\Pr(\| \nu_n \| > t)$ is bounded by:

$$K_1(\varepsilon_0^{-1} t)^{4d}(2 + t^2)^{12d} \exp(-2t^2) \qquad 3.8.3$$

with K_1 depending only on C, whenever $t^2 \geq 7 + 12d$.

Under assumption b). — Using the same arguments as above we may suppose that: $N_j \leq \exp(Ct^\zeta j^{\zeta(\alpha+\beta)})$.

We set $\mu = t^{-2\gamma}$, then the following inequalities hold $P^{\otimes N}$-a. s.:

$$A \leq 2\exp(Ct^\zeta(2 + t^{2\gamma/(\alpha-1)})^{\zeta(\alpha+\beta)} + 4t^{2(1-\gamma)} + 10)\exp(-2t^2),$$

$$B \leq 2\sum_{j\geq r+1} \exp\left(2Ct^\zeta j^{\zeta(\alpha+\beta)} - \frac{t'^2}{4}(j-1)^2\right).$$

In order to balance the above bound of A we choose γ so that $\zeta\left(1 + 2\gamma\left(\frac{\alpha+\beta}{\alpha-1}\right)\right) = 2(1-\gamma)$ and α large enough for $\gamma \geq \gamma(\zeta) - \frac{\eta}{2}$ and $\beta > 1$ to hold, where $\gamma(\zeta)$ is the solution of the above equation when $\alpha = +\infty$ (Note that $2(1-\gamma(\zeta)) = k$, where k is defined in the statement of 3.3). So

$$A \leq 0_{\eta,\mathscr{F}}(1)\exp(0_{\eta,\mathscr{F}}(1)t^{k+\eta})\exp(-2t^2) \qquad P^{\otimes N}\text{-a. s.}$$

The evaluation of B is performed via lemma 3.8.2 with the choice:

$$\psi(x) = \frac{t'^2}{4}(x-1)^{2\beta} - 2Ct^\zeta x^{2\beta}.$$

Then, whenever $t^2 \geq \frac{2}{\beta} + 5 + Ct^\zeta 2^{2\beta+4}$, $\psi'(r) \geq 1$, hence (since $r \to \infty$ whenever $t \to \infty$)

$$B \leq 0_{\eta,\mathscr{F}}(1)\exp(-2t^2) \qquad P^{\otimes N}\text{-a. s.}$$

We finish the proof as under assumption *a*), so theorem 3.3 is proved in case 1°).

Proof of theorem 3.3 in case 2°). — We set $\varphi = \frac{t}{\sigma}$ and choose $m = [\varphi^q]$, $a = 2\varphi^{-q}$, $\mu = \varphi^{-q}$ and $\tau_j = \frac{\sigma}{\sqrt{m}} j^{-(\alpha+\beta)}$. If we set $\rho = \sigma$, the variance factor $1 - \frac{\sigma^2}{a^2\varepsilon^2 n'}$ in lemma 3.2 is stabilized for it is not less than $\frac{1}{2}$ whenever $\varphi^q \geqslant 4$. The variable A is this time controlled with the help of 3.1.2°), so now *the probleme is to replace* σ_N^2 *by* σ_P^2. In fact, let Ξ_N be the $(x_1, \ldots, x_N)$-measurable event:

$$\Xi_N = \{\ \|\sigma_N^2 - \sigma_P^2\| > s\}, \quad \text{where} \quad \sigma_N^2(f) = P_N(f^2) - (P_N(f))^2 \quad \text{for any } f \text{ in } \mathscr{F}.$$

Each term of the following estimate is studied in the sequel:

$$\Pr(\|\tilde{P}_n - P_N\| > \varepsilon') \leq \Pr(\Xi_N) + A' + B' \qquad 3.8.4$$

where $A' = E(A\mathbb{1}_{\Xi_N^c})$ and $B' = E(B\mathbb{1}_{\Xi_N^c})$.

The control of $\Pr(\Xi_N)$ *reduces to a problem of type 1°).* — For, setting $\mathcal{F}^2 = \{ f^2, f \in \mathcal{F} \}$, we have:

$$\| \sigma_N^2 - \sigma_P^2 \| \leq \| P_N - P \|_{\mathcal{F}^2} + 2 \| P_N - P \| .$$

Since $N_1^{(2)}(., \mathcal{F}^2) \leq N_1^{(2)}\left(\frac{\cdot}{2}, \mathcal{F}\right)$ and $\mathcal{F}^2$ fulfills $(\mathcal{M})$, we may use the bounds in 3.3.1°), so

$$\Pr(\Xi_N) \leq \Psi\left(\frac{s}{3}\sqrt{N}\right) \exp\left(-\frac{2Ns^2}{9}\right)$$

where Ψ is a polynomial or exponential function (according to whether case *a*) or *b*) is studied). Anyway

$$\Pr(\Xi_N) \leq C_0 \exp\left(-\frac{Ns^2}{9}\right).$$

We choose $s = \dfrac{3\varphi}{\sqrt{2nm}}$, so $\Pr(\Xi_N) = 0\left(\exp\left(-\dfrac{\varphi^2}{2}\right)\right)$.

The evaluation of A′ *and* B′. — The inequality $\| \sigma_N^2 \| \leq \sigma^2 + s$ holds on Ξ_N^c, thus applying 3.1.2°):

$$A \mathbb{1}_{\Xi_N^c} \leq 2N_r \exp\left(-\frac{t'^2(1-\mu)^2}{2(\sigma^2 + \varepsilon + s)}\right) \qquad P^{\otimes N}\text{-a. s.}$$

so, since $(1-\mu)^2 \dfrac{t'^2}{\sigma^2} \geq \varphi^2 - 10\varphi^{2-q}$ whenever $\varphi^q \geq 2$, we get:

$$A' \leq 2N_r \exp(5\varphi^{2-q}) \exp\left(-\frac{t^2}{2(\sigma^2 + (t + 3\varphi^{1-q/2})/\sqrt{n})}\right),$$

$$B' \leq 2 \sum_{j \geq r+1} N_j^2 \exp\left(\left(-\frac{t'^2}{4\sigma^2}(j-1)^{2\beta}\right)\right).$$

Under assumption a). — As in the proof of 3.3.1°) we may assume that $N_j \leq C\sigma^{-2d}\varphi^{2d}j^{2d(\alpha+1)}$ and we choose $q = 2$ and $\alpha = \operatorname{Max}\left(2, 1 + \dfrac{4d}{3}\right)$ (recall that $\beta = 1$). Then:

$$A' \leq 0_{\eta,\mathcal{F}}(1)\sigma^{-2d}(\varphi^2)^{3(d+\eta)} \exp\left(-\frac{t^2}{2(\sigma^2 + (3+t)/\sqrt{n})}\right).$$

In another connection the control of B′ follows from the control of the same tail of series as in 1°) (via a modification of parameters), so:

$$B' \leq 0_{\eta,\mathcal{F}}(1)\sigma^{-4d} \exp\left(-\frac{\varphi^2}{2}\right)$$

whenever $\varphi^2 \geq 8 + 4d(\alpha + 1)$. Collecting the estimates of $\Pr(\Xi_N)$, A′ and B′, bound 3.3.2°) *a)* is established via lemma 3.2 and inequality 3.8.4.

Proof of 3.5. — Under the hypothesis in 3.5 we choose this time $\alpha = 2$, so:

$$N_j \leq C\varepsilon_0^{-2d}\sigma^{-2d}\varphi^{2d}j^{4d},$$

hence

$$A' \leq 2C\varepsilon_0^{-2d}e^5\sigma^{-2d}\varphi^{2d}(2+\varphi^2)^{6d}\exp\left(-\frac{t^2}{2(\sigma^2+(3+t)/\sqrt{n})}\right),$$

$$B' \leq 2C^2\varepsilon_0^{-4d}\sigma^{-4d}\varphi^{4d}(2+\varphi^2)^{12d}\exp\left(-\frac{(\varphi^2-8)^2}{4}\right)$$

whenever $\varphi^2 \geq 8 + 12d$. The evaluation of $\Pr(\Xi_N)$ is performed by using bound 3.8.3 $\left(\text{changing } \varepsilon_0 \text{ into } \frac{\varepsilon_0}{2} \text{ when studying } \mathscr{F}^2\right)$ at the point $\frac{\varphi}{2}$. Hence

$$\Pr(\Xi_N) \leq 2K_1\left(\frac{\varepsilon_0}{2}\right)^{-4d}\left(\frac{\varphi}{2}\right)^{4d}\left(2+\frac{\varphi^2}{2}\right)^{12d}\exp\left(-\frac{\varphi^2}{2}\right)$$

whenever $\frac{\varphi^2}{4} \geq 7 + 12d$. Applying inequality 3.8.4 and lemma 3.2 gives proposition 3.5.

Proof of 3.3.2°). — *Under assumption* b). Let us choose q so that:

$$\zeta + \zeta q\left(\frac{1}{2} + \left(\frac{\alpha+\beta}{\alpha-1}\right)\right) = 2 - q.$$

We choose α large enough for $\beta > 1$ and $1 - \frac{q}{2} \leq p + \eta$ (where p is defined in the statement of theorem 3.3) to hold. We may assume that:

$$N_j \leq \exp(C\sigma^{-\zeta}(\varphi)^{\zeta q/2}j^{\zeta(\alpha+\beta)}).$$

Then:

$$A' \leq 2\exp(0_{\eta,\mathscr{F}}(1)\sigma^{-\zeta}\varphi^{2p-\zeta+\eta} + 5\varphi^{2p+\eta})\exp\left(-\frac{t^2}{2\left(\sigma^2+\frac{1}{\sqrt{n}}(t+3\varphi^{p+\eta})\right)}\right).$$

The control of B′ is of the same kind as that of B in situation 1°), so:

$$B' \leq 2\exp\left(2C\sigma^{-\zeta}\varphi^{\zeta q/2}r^{2\beta} + 2\varphi^{2-q} - \frac{\varphi^2}{4}(r-1)^{2\beta}\right)$$

whenever $\varphi^2 \geq 8\varphi^{2p+\eta} + \frac{2}{\beta} + 8C\sigma^{-\zeta}\varphi^{(p-1)\zeta}$. Since $r \to \infty$ whenever $\varphi \to \infty$:

$$B' \leq 0_{\eta,\mathscr{F}}(1)\exp(2C\sigma^{-\zeta}\varphi^{2p-\zeta+\eta})\exp\left(-\frac{\varphi^2}{2}\right).$$

Using inequality 3.8.4 and lemma 3.2 gives 3.3.2°) *b*) and finishes the proof of theorem 3.3.

Proof of proposition 3.7. — The above method which allows deduction of bounds of type 2°) from bounds of type 1°) is iterated here. We shall assume that $u = 0$ and $v = 1$. Inequality 3.3.2°) *a*) may be written (in view of its proof):

$$\mathrm{M}_0: \quad \Pr(\| v_n \| > t) \leq \mathrm{K}\sigma^{-\alpha_1}\left(\frac{t^2}{\sigma^2}\right)^{\alpha_2} \exp\left(-\frac{t^2}{2(\sigma^2 + (3+t)/\sqrt{n})}\right)$$

whenever $t^2/\sigma^2 \geq 5$. Let us define by induction the following sequences

$$a_{j+1} = 54^{4d+1} a_j \left(2 + \frac{2b_j}{\sigma^2} + \frac{1}{n\sigma^2}\right),$$

$$b_{j+1} = \frac{2}{\sqrt{n}}\left(\frac{1}{\sqrt{n}} + \sqrt{b_j}\right)$$

with $a_0 = 1$ and $b_0 = \frac{3}{\sqrt{n}}$.

We suppose that the following inequality (M_j) holds whenever $t^2/\sigma^2 \geq 5$:

$$\Pr(\| v_n \| > t) \leq \mathrm{K} a_j \sigma^{-\alpha_1}\left(\frac{t^2}{\sigma^2}\right)^{\alpha_2} \exp\left(-\frac{t^2}{2(\sigma^2 + b_j + t/\sqrt{n})}\right). \quad (\mathrm{M}_j)$$

We want to deduce (M_{j+1}) from (M_j) by the same way as 3.3.2°) *a*) from 3.3.1°) *a*). So, let ξ be positive so that:

$$\frac{m}{2} = \frac{\mathrm{N}\xi^2}{2(\sigma^2 + b_j(\mathrm{N}) + \xi)} \qquad \text{(recall that } mn = \mathrm{N} \text{ and } m = [t^2/\sigma^2]\text{)}.$$

Then

$$\left[\frac{t^2}{\sigma^2}\right] \leq \frac{\mathrm{N}\xi^2}{\sigma^2} \leq \frac{t^2}{\sigma^2}\left(2 + \frac{2b_j(n)}{\sigma^2} + \frac{1}{n\sigma^2}\right).$$

Besides, setting $\tilde{\mathcal{F}} = \{ f - \mathrm{P}(f), f \in \mathcal{F} \}$, we have for any law Q

$$\mathrm{N}_2^{(2)}(.,\tilde{\mathcal{F}} + 1, \mathrm{Q}) \leq \mathrm{N}_1^{(2)}\left(.,\mathcal{F},\frac{\mathrm{P}+\mathrm{Q}}{2}\right) \leq \mathrm{N}_1^{(2)}\left(\frac{.}{2},\mathcal{F}\right)$$

according to remark 2.6, hence:

$$\mathrm{N}_1^{(2)}(.,\tilde{\mathcal{F}}^2) \leq \mathrm{N}_1^{(2)}\left(\frac{.}{4},\mathcal{F}\right).$$

Moreover it is easy to show that

$$\| \sigma_\mathrm{N}^2 - \sigma_\mathrm{P}^2 \|_{\mathcal{F}} \leq \| \mathrm{P}_\mathrm{N} - \mathrm{P} \|_{\tilde{\mathcal{F}}^2} + \| \mathrm{P}_\mathrm{N} - \mathrm{P} \|_{\mathcal{F}}.$$

Therefore, using the notations of the above proof of 3.3.2°) *a*) and applying (M_j) to $\mathcal{F}$ and $\tilde{\mathcal{F}}^2$ at the point $\sqrt{N}\xi$, we get (since $\| P \|_{\tilde{\mathcal{F}}^4} \le \sigma^2$)

$$\Pr(\Xi_N) \le (1 + 4^{4d+1}) K a_j \sigma^{-\alpha_1} \left(\frac{t^2}{\sigma^2}\right)^{\alpha_2} \left(2 + \frac{2b_j}{\sigma^2} + \frac{1}{n\sigma^2}\right)^{\alpha_2} \sqrt{e} \exp\left(-\frac{t^2}{2\sigma^2}\right)$$

and also (here $s = 2\xi$)

$$A' + B' \le (K/2)\sigma^{-\alpha_1}\left(\frac{t^2}{\sigma^2}\right)^{\alpha_2} \exp\left(-\frac{t^2}{2(\sigma^2 + 2\xi + t/\sqrt{n})}\right).$$

So, via lemma 3.2 and inequality 3.8.4 we get (M_{j+1}) whenever $t^2/\sigma^2 \ge 5$. Therefore inequalities (M_j) are proved by induction.

Let us consider the step $J = 1 + [LLn/L2]$ of the iteration. Then, with a few calculations it can be shown that:

$$b_J \le \frac{4}{n}\left(e + J\left(1 + \frac{\sigma\sqrt{n}}{2}\right)\right)$$

and

$$a_J \le (5\,4^{4d+1})^J n^{\alpha_2}\left(\frac{9}{n\sigma^2} + 2 + \frac{8J}{n\sigma^2}\left(1 + \frac{\sigma\sqrt{n}}{2}\right)\right)^{J\alpha_2}.$$

Hence, using the condition $\sqrt{n\sigma} \ge V$ and the above estimates of a_J and b_J, inequality (M_J) gives proposition 3.7.

4. EXPONENTIAL BOUNDS FOR THE BROWNIAN BRIDGE

We assume that $P(F^2) < \infty$. We want to show that the bounds in the preceding section still hold for the brownian bridge.

4.1. THEOREM. — If $e_F^{(2)}(\mathcal{F}) < 2$, then there exists some version G_P of a brownian bridge relating to P whose trajectories are uniformly continuous and bounded on $(\mathcal{F}, \rho_P)$. Moreover, setting $\zeta = e_F^{(2)}(\mathcal{F})$, if $\| \sigma_P^2 \|_{\mathcal{F}} \le \sigma^2 \le P(F^2)$, an upper bound for $\Pr(\| G_P \|_{\mathcal{F}} > t)$ is, for any positive t and η, given by:

$$0_{\eta,\mathcal{F}}(1)\exp\left(0_{\eta,\mathcal{F}}(1)\sigma^{-\zeta-2\eta}(P(F^2))^{\zeta/2+\eta}\left(\frac{t}{\sigma}\right)^{2p-\zeta+\eta} + \left(\frac{t}{\sigma}\right)^{2p+\eta}\right)\exp(-t^2/2\sigma^2) \quad 4.1.1$$

or, if more precisely $d_F^{(2)}(\mathcal{F}) = 2d < \infty$, by:

$$0_{\eta,\mathcal{F}}(1)(P(F^2))^{2d+\eta}\sigma^{-4d-2\eta}\left(1 + \frac{t^2}{\sigma^2}\right)^{2d+\eta} \exp(-t^2/2\sigma^2) \quad 4.1.2$$

where p is defined in the statement of theorem 3.3. The constants appearing in the above bounds depend only on η and $N_F^{(2)}(.,\mathscr{F})$.

Comments. — In the framework of theorem 4.1 the existence of some regular version of a brownian bridge is an easy consequence of the proof of 4.1.1 and 4.1.2 but is of course a well known result (see [*18*]). Moreover the bounds in 4.1 are in this case sharper than the more general Fernique-Landau-Shepp inequality (see [*25*]) that can be written:

$$\Pr(\| G_P \|_{\mathscr{F}} > t) \leq C(\alpha) \exp(-t^2/2\alpha^2)$$

for any $\alpha > \| \sigma_P \|_{\mathscr{F}}$.

Proof of theorem 4.1. — If $\mathscr{F}$ is countable. Let G_P be some brownian bridge relating to P. The calculations are similar to those in the proof of theorem 3.3. Of course here a randomization would not mean anything and a sequence of nets in $(\mathscr{F}, F)$ relating to P is directly given. Moreover the following single inequality is used instead of the bounds in 3.1:

4.2. LEMMA. — Let V be a real and centered gaussian random variable, then:

$$\Pr(| V | > s) \leq 2 \exp(-s^2/(2v^2)) \qquad \text{for any positive } s$$

where v^2 stands for the variance of V.

So, as in 3.8 we have:

$$\Pr(\| G_P \|_{\mathscr{F}} > t) \leq A + B$$

with this time:

$$A \leq 2N_r \exp\left(-\frac{t^2(1-\mu)^2}{2\sigma^2}\right),$$

$$B \leq 2 \sum_{j \geq r+1} N_j^2 \exp\left(-\frac{t^2\eta_j^2}{8P(F^2)\tau_{j-1}^2}\right).$$

The choices of parameters $\left(\text{except } \tau_j^2 = \frac{\sigma^2}{P(F^2)} j^{-2(\alpha+\beta)}\right)$ and the estimates of A and B are then performed as in the proof of 3.3.2°) giving 4.1.1 and 4.1.2. Since 4.1.1 is also an oscillation control the almost sure regularity of G_P is ensured via the Borel-Cantelli lemma. So theorem 4.1 is proved in the case where $\mathscr{F}$ is countable.

The general case. — Since $(\mathscr{F}, \rho_P)$ is totally bounded any version of a brownian bridge which is almost surely regular on some countable and dense subset of $(\mathscr{F}, \rho_P)$ may be extended to $(\mathscr{F}, \rho_P)$ giving some version of a brownian bridge which is almost surely regular on $(\mathscr{F}, \rho_P)$. Inequalities 4.1.1 and 4.1.2 hold for this version.

Comment. — The optimality of bound 4.1.2 is discussed in the appendix. The degrees of the polynomial factors are different in 3.3.1.2°) *a*) and in 4.1.2; the reason is that bound 3.1.3°) is less efficient than bound 4.2.

5. WEAK INVARIANCE PRINCIPLES WITH SPEEDS OF CONVERGENCE

We assume from now on that $P(F^{2+\delta}) < \infty$ for some δ in $]0, 1]$. Using the results in sections 3 and 4, we can evaluate the oscillations of the empirical brownian bridge and of a regular version of the brownian bridge over $\mathscr{F}$, so we can control the approximations of these processes by some E_k-valued processes (where E_k is a vector space with finite dimension k). The Prokhorov distance between the distributions of these two processes is estimated via an inequality from Dehling [*15*] allowing reasonable variations of k with n.

Oscillations of the empirical brownian bridge over $\mathscr{F}$. — The oscillations of ν_n over $\mathscr{F}$ are controlled with the help of a truncation from 3.3.2°) *a*) (the proof in this case is straightforward) on the one hand and of a slight modification in the proof of 3.3.2°) *b*) (truncating twice) on the other hand. The control of the tails of these various truncations is performed via the following inequalities (the second is elementary).

5.1. LEMMA (Marcinkiewicz-Zygmund's inequality). — Let $\xi_1, \ldots, \xi_n$ be independent and identically distributed real valued random variables such that $E(\xi_1) = 0$ and $E(|\xi_1|^r) < \infty$ for some r in $]1, +\infty[$.

We set $S_n = \sum_{i=1}^{n} \xi_i$, then:

$$E(|S_n/n|^r) \le \Gamma_r E(|\xi_1|^r) n^{-\min(r-1, r/2)}.$$

See [*42*] for a proof of the above lemma.

Besides: let ξ be a positive random variable such that $E(\xi^r) < \infty$ for some r in $]1, +\infty[$. The following inequality holds for any positive M:

$$E(\xi I_{(\xi \ge M)}) \le E(\xi^r)/M^{r-1}. \qquad 5.2$$

(Note that $M^{r-1}\xi I_{(\xi \ge M)} \le \xi^r$ to prove 5.2). The next theorem is an analogue of theorem 3.3 in the case where F is not necessarily bounded.

5.3. THEOREM. — We set $\mu_\delta = \mathrm{P}(\mathrm{F}^{2+\delta})$. If we assume that $\|\sigma_{\mathrm{P}}^2\|_{\mathscr{F}} \leq \sigma^2 \leq \mathrm{P}(\mathrm{F}^2)$ with $\sqrt{n}\sigma \geq 1$, then an upper bound for $\Pr(\|\nu_n\|_{\mathscr{F}} > t)$ is, for any positive t such that $t^2/\sigma^2 \geq 1$, given by:

a) if $d_{\mathrm{F}}^{(2)}(\mathscr{F}) = 2d < \infty$,

$$0_{\mathscr{F}}(1) n^{7d} \left(\frac{t}{\sigma}\right)^{8d} \exp\left(-\frac{t^2}{8\sigma^2\left(2 + 0\left(\frac{\mathrm{LL}n\sigma}{t}\right)\right)}\right) + 128 n^{-\delta/2}\sigma^{-2\delta}t^{-2+\delta}$$

whenever $$n^{\delta/2}\sigma^{2+2\delta}t^{-\delta} \geq 64\mu_\delta. \qquad 5.3.1$$

b) if $e_{\mathrm{F}}^{(2)}(\mathscr{F}) = \zeta < 2$, an upper bound is

$$0_{\eta,\mathscr{F}}(1) \exp\left(0_{\eta,\mathscr{F}}(1)\left(\frac{t}{\sigma}\right)^{2p+\eta}\left(1 + (\mathrm{P}(\mathrm{F}^2))^{\frac{\zeta}{2}+\eta}\sigma^{-\zeta-2\eta}\left(\frac{t}{\sigma}\right)^{-\zeta}\right)\exp\left(-\frac{t^2}{24\sigma^2}\right)\right.$$

$$+ 0_{\eta,\mathscr{F}}(1)\exp\left(-\frac{1}{16}\left(\frac{t}{\sigma}\right)^{2-2p+\eta}\right)$$

$$+ 0_\eta(1)(1 + \mu_\delta)^2\left(n\left(\frac{t}{\sigma}\right)^{2-2p-\eta}\right)^{-\delta/2}(\sigma^{-2-\delta} + \sigma^{-\delta})$$

$$+ 0(1)\mu_\delta n^{-\delta/2}\sigma^{-2\delta}t^{-2+\delta} \quad \text{for any positive } \eta$$

(p is defined in the statement of 3.3) whenever 5.3.1 holds and:

$$n^{\delta/4}\sigma^{2+\delta} \geq 512\mu_\delta. \qquad 5.3.2$$

Proof of theorem 5.3. — Under assumption a). Note that we may assume that $\mathscr{F}$ is centered under P as in the proof of proposition 3.7, for, setting $\tilde{\mathscr{F}} = \{ f - \mathrm{P}(f), f \in \mathscr{F} \}$ and $\tilde{\mathrm{F}} = \mathrm{F} + (\mathrm{P}(\mathrm{F}^2))^{1/2}$, we have:

$$\mathrm{N}_{\tilde{\mathrm{F}}}^{(2)}(\,.\,, \tilde{\mathscr{F}}) \leq \mathrm{N}_{\mathrm{F}}^{(2)}\left(\frac{\cdot}{2\sqrt{2}}, \mathscr{F}\right)$$

with $\mathrm{P}(\tilde{\mathrm{F}}^2) \leq 4\mathrm{P}(\mathrm{F}^2)$ and $\mathrm{P}(\tilde{\mathrm{F}}^{2+\delta}) \leq 8\mathrm{P}(\mathrm{F}^{2+\delta})$.

We make the truncation of $\mathscr{F}$ from 3.7 with the speed $\mathrm{M} = \sigma^2\sqrt{n}/t$. We set:

$f(\mathrm{M}) = f\mathrm{I}_{(\mathrm{F} \leq \mathrm{M})}$ and $\bar{f}(\mathrm{M}) = f - f(\mathrm{M})$ for any f in $\mathscr{F} \cup \{\mathrm{F}\}$ and then $$\mathscr{F}(\mathrm{M}) = \{ f(\mathrm{M}), f \in \mathscr{F} \}.$$

It is easy to see that:

$$\mathrm{N}_{\mathrm{F(M)}}^{(2)}(\,.\,, \mathscr{F}(\mathrm{M})) \leq \mathrm{N}_{\mathrm{F}}^{(2)}(\,.\,, \mathscr{F}), \quad \text{hence} \quad \mathrm{N}_{2\mathrm{M}}^{(2)}(\,.\,, \mathrm{M} + \mathscr{F}(\mathrm{M})) \leq \mathrm{N}_{\mathrm{F}}^{(2)}(\,.\,, \mathscr{F}).$$

Moreover $\mathscr{F}(\mathrm{M})$ fulfills ($\mathscr{M}$). Note that:

$$\|\nu_n\|_{\mathscr{F}} \leq \|\nu_n\|_{\mathscr{F}(\mathrm{M})} + \sqrt{n}(\mathrm{P}_n + \mathrm{P})(\bar{\mathrm{F}}(\mathrm{M}))$$

hence $$\Pr(\|\nu_n\|_{\mathscr{F}} > t) \leq C + D$$

where $C = \Pr\left(\|\nu_n\|_{\mathscr{F}(M)} > \frac{t}{2}\right)$ and $D = \Pr\left(\sqrt{n}(P_n + P)(\overline{F}(M)) > \frac{t}{2}\right)$.

Control of D. — Using Bienaymé-Tchebycheff's inequality and bound 5.2 we get:

$$D \leq P((\overline{F}(M))^2)\frac{16}{t^2} \leq 16t^{-2}\mu_\delta M^{-\delta}$$

whenever $P(\overline{F}(M)) \leq \frac{t}{8\sqrt{n}}\cdot$ Since $P(\overline{F}(M)) \leq \mu_\delta M^{-1-\delta}$ according to 5.2 the above estimate of D holds whenever $\mu_\delta 8\sqrt{n}M^{-1-\delta} \leq t$.

Control of C. — Applying proposition 3.7 to $\mathscr{F}(M)$ we get:

$$C \leq Q_{\mathscr{F}}(1)n^{4d}\left(\frac{t^2}{\sigma^2}\right)^{4d}(\sqrt{n})^{5d}\exp\left(-\frac{t^2}{8\sigma^2\left(2 + 0\left(\frac{\mathrm{LL}n\sigma}{t}\right)\right)}\right).$$

Collecting the above estimates and changing μ_δ into $8\mu_\delta$ give 5.3.*a*).

Under assumption b). — Bound 3.3.2°) *b*) is not efficient enough to allow any direct truncation of $\mathscr{F}$. Moreover an inequality such as in prosition 3.7 is here not available (our proof of 3.7 does not work under assumption *b*)). So our plan to avoid this difficulty is as follows: truncating from 3.3.1°) *b*) to deduce a new bound of type 2°) via the method in 3.8 in the case where F is bounded, finishing the proof in the general case by truncating as under assumption *a*).

5.4. Lemma. — Let $\mathscr{H}$ be $\mathscr{F}$ or $\mathscr{F}^2$. Then an upper bound for $\Pr(\|P_n - P\|_{\mathscr{H}} > \varepsilon)$ is, for any positive ε, given by:

i) If $\mathscr{H} = \mathscr{F}$, $0_{\mathscr{F}}(1)\exp(-\theta) + 64\left(\frac{n}{\theta}\right)^{-\delta/2} n^{-1}\varepsilon^{-2-\delta}\mu_\delta$ for any positive θ fulfilling the condition:

$$\left(\frac{n}{\theta}\right)^{\frac{1+\delta}{2}}\varepsilon^{2+\delta} \geq 2^7\mu_\delta. \qquad 5.4.1$$

ii) If $\mathscr{H} = \mathscr{F}^2$, $0_{\mathscr{F}}(1)\exp(-\theta) + 0(1)\mu_\delta n^{-\delta/2}\varepsilon^{-1-(\delta/2)}$ for any positive θ fulfilling the condition:

$$\left(\frac{n}{\theta}\right)^{\delta/4}\varepsilon^{1+(\delta/2)} \geq 16\mu_\delta. \qquad 5.4.2$$

Proof of lemma 5.4 (first truncation). — Let H be F or F^2. The notations for the truncations are the same as above. It can be shown that $N_{2M}^{(2)}(., M + \mathscr{H}(M)) \leq N_F^{(2)}(., \mathscr{F})$ and that $\mathscr{H}(M)$ fulfills $(\mathscr{M})$, so theorem 3.3 may be applied to $\mathscr{H}(M)$. This time we choose the speed of truncation: $M = \left(\frac{n}{\theta}\right)^{1/2} \frac{\varepsilon}{4}$. Then: $\Pr(\| P_n - P \|_{\mathscr{H}} > \varepsilon) \leq C + D$ with $C = \Pr\left(\| P_n - P \|_{\mathscr{H}(M)} > \frac{\varepsilon}{2}\right)$ and $D = \Pr\left((P_n + P)(\overline{H}(M)) > \frac{\varepsilon}{2}\right)$.

Using 3.3.1°) *b*), an upper bound for C is given by: $0_{\mathscr{F}}(1) \exp(-\theta)$. Besides, whenever $P(\overline{H}(M)) \leq \frac{\varepsilon}{8}$ and using Bienaymé-Tchebycheff's inequality

$$D \leq P((\overline{F}(M))^2) n^{-1} \left(\frac{\varepsilon}{4}\right)^{-2}$$

holds in case *i*) on the one hand; on the other hand we apply lemma 5.1 to get:

$$D \leq 0(1) \mu_\delta n^{-\delta/2} \varepsilon^{-1-(\delta/2)}$$

in case *ii*).

Bounding $P(\overline{H}(M))$ with the help of 5.2 and collecting the various estimates above, we obtain lemma 5.4. Let us return to the proof of 5.3.*b*).

We assume that $F \leq M$ *for some positive* M.

The slight modifications that we introduce in the proof of 3.3.2°) *b*) are as follows (the notations are those of this proof):

— We consider nets of $(\mathscr{F}, F)$ relating to P_N.

— The event Ξ_N is defined by:

$$\Xi_N = \{ \| \sigma_N^2 - \sigma_P^2 \|_{\mathscr{F}} > \sigma^2 \} \cup \{ (P_N - P)(F^2) > P(F^2) \}.$$

— We choose $\tau_j^2 = \frac{\sigma^2}{m P(F^2)} j^{-2(\alpha+\beta)}$.

Then the inequalities: $P_N((\Pi_j(f) - f)^2) \leq \frac{2\sigma^2}{m} j^{-2(\alpha+\beta)}$ for any f in $\mathscr{F}$ and $\| \sigma_N^2 \|_{\mathscr{F}} \leq 2\sigma^2$, hold on Ξ_N^c, so, recalling estimate 3.8.4:

$$\Pr(\| P_n - P_N \|_{\mathscr{F}} > \varepsilon') \leq \Pr(\Xi_N) + A' + B'.$$

The control of A′ and B′ is performed as in 3.8, giving:

$$A' + B' \leq 0_{\eta,\mathscr{F}}(1) \exp\left(0_{\eta,\mathscr{F}}(1) \sigma^{-\zeta-2\eta} \left(\frac{t}{\sigma}\right)^{2p-\zeta+\eta} (P(F^2))^{(\zeta/2)+\eta} + 5\left(\frac{t}{\sigma}\right)^{2p+\eta}\right)$$
$$\cdot \exp\left(-\frac{t^2}{4\left(\sigma^2 + \frac{Mt}{\sqrt{n}}\right)}\right).$$

Control of $\Pr(\Xi_N)$. — We have $\Pr(\Xi_N) \le \mathrm{I} + \mathrm{II}$ where

$$\mathrm{I} = \Pr(\{ \| \sigma_N^2 - \sigma_P^2 \|_{\mathscr{F}} > \sigma^2 \} \cap \{ (P_N - P)(F^2) \le P(F^2) \})$$

and $$\mathrm{II} = \Pr(|(P_N - P)(F^2)| > P(F^2)).$$

The control of I may be performed via lemma 5.4, for:

$$\| \sigma_N^2 - \sigma_P^2 \|_{\mathscr{F}} \le \| P_N - P \|_{\mathscr{F}^2} + \| P_N - P \|_{\mathscr{F}} (2(P_N + P)(F^2))^{1/2}$$

so, on the event $\{ (P_N - P)(F^2) \le P(F^2) \}$, the following inequality holds:

$$\| \sigma_N^2 - \sigma_P^2 \|_{\mathscr{F}} \le \| P_N - P \|_{\mathscr{F}^2} + (6P(F^2))^{1/2} \| P_N - P \|_{\mathscr{F}} .$$

Hence, applying lemma 5.4 with $\theta = m/4$, we get:

$$\mathrm{I} \le 0_{\mathscr{F}}(1) \exp\left(- \frac{m}{4} \right) + 0(1)\left(N^{-1-(\delta/2)} \mu_\delta^2 \sigma^{-4-2\delta} + N^{-\delta/2} \mu_\delta^2 \sigma^{-2-\delta}\right)$$

whenever $\sigma^{2+\delta} N^{\delta/4} \ge 64\mu_\delta$. Besides, from lemma 5.1 we get:

$$\mathrm{II} \le 0(1)(P(F^2))^{-\delta/2} \mu_\delta N^{-\delta/2} .$$

Since $m \ge \left(\frac{t}{\sigma}\right)^{2-2p+\eta} - 1$, collecting the above estimates and using lemma 3.2 we get the following upper bound for $\Pr(\| \nu_n \|_{\mathscr{F}} > t)$:

$$\begin{aligned}
&0_{\eta,\mathscr{F}}(1) \exp\left(0_{\eta,\mathscr{F}}(1) \sigma^{-\zeta-2\eta} (P(F^2))^{(\zeta/2)+\eta} \left(\frac{t}{\sigma}\right)^{2p-\zeta+\eta} \right. \\
&+ 5\left(\left(\frac{t}{\sigma}\right)^{2p+\eta}\right) \exp\left(- \frac{t^2}{4\left(\sigma^2 + \frac{Mt}{\sqrt{n}}\right)} \right) + 0_{\mathscr{F}}(1) \exp\left(- \frac{1}{4}\left(\frac{t}{\sigma}\right)^{2-2p-\eta} \right) \\
&+ 0(1)(1+\mu_\delta)^2 \left(n \left(\frac{t}{\sigma}\right)^{2-2p-\eta} \right)^{-\delta/2} (\sigma^{-2-\delta} + \sigma^{-\delta}) \qquad 5.5
\end{aligned}$$

whenever $\sigma^{2+\delta} n^{\delta/4} \ge 64\mu_\delta$.

The general case (second truncation). — We proceed here as in the proof of 5.3.*a*) using 5.5 instead of proposition 3.7. So theorem 5.3 is proved.

Speed of convergence in the central limit theorem in finite dimension. — We recall below an evaluation of the Prokhorov distance between the distribution of the centered and normalized sum of independent and identically distributed $\mathbb{R}^k$-valued and $(2 + \delta)$-integrable random variables and the corresponding gaussian distribution, that is due to Dehling [*15*] (the first result in the same direction is due to Yurinskii [*53*]).

5.6. THEOREM. — Let $(X_i)_{1 \le i \le n}$ be a sample of centered $\mathbb{R}^k$-valued random variables. We write F_n for the distribution of the normalized sum of these variables and G for the centered gaussian distribution whose covariance is that of X_1.

Let $\|\cdot\|_2$ be an euclidean pseudo-norm on $\mathbb{R}^k$ and π_2 be the Prokhorov distance that is associated to $\|\cdot\|_2$. If $E(\|X_1\|_2^{2+\delta}) = \mu < \infty$, then:

$$\pi_2(F_n, G) \le Kn^{-\delta/8}k^{1/4}\mu^{1/4}(1 + |L(n^{-\delta/2}k^{-1}\mu)|^{1/2}).$$

Comment. — Comparing the above evaluation and the one in [*52*] (assuming that $\delta = 1$), a power of n is lost in the speed of convergence for in [*52*] we have $\pi_2(F_n, G) = 0(n^{-1/2})$ but in 5.6 the estimate is a polynomial function of k; on the contrary the estimate in [*52*] depends exponentially on k.

Weak invariance principles for the empirical brownian bridge. — In order to build regular versions of brownian bridges with given projection on a finite dimensional vector space (or further in section 6 on a countable product of such spaces), we need two lemmas.

5.7. LEMMA (Berkes, Philipp [*6*]). — Let R_1, R_2, R_3 be Polish spaces, Q_1 and Q_2 be some distributions respectively defined on $R_1 \times R_2$ and $R_2 \times R_3$ with common marginal on R_2. Then there exists a distribution Q on $R_1 \times R_2 \times R_3$ whose marginals on $R_1 \times R_2$ and $R_2 \times R_3$ are respectively Q_1 and Q_2.

Remember that $l_T^\infty(\mathscr{F})$ is generally not separable. The following lemma is fundamental to avoid this difficulty (see [*23*]). The space Ω to be mentionned below is defined in Section 2.

5.8. LEMMA (Skorohod [*48*]). — Let R_1, R_2 be Polish spaces and Q be some distribution on $R_1 \times R_2$ with marginal q on R_2. If V is a random variable from X^∞ to R_2 whose distribution is q, then there exists a random variable Y from Ω to R_1 such that the distribution of (Y, V) is Q.

Concerning our problems of construction the point in the sequel is that the distribution on $l_T^\infty(\mathscr{F})$ of a regular version of a brownian bridge is concentrated on a separable space. Now we can state some weak invariance principles for the empirical brownian bridge with speeds of convergence.

5.9. NOTATIONS. — From now on γ and β are positive functions that are respectively defined on $[0,1] \times \mathbb{R}_+$ and $[0,2]$ by:

$$\gamma(x, y) = \frac{x}{8 + 2y(4 + x)} \qquad \text{and} \qquad \beta(z) = \frac{2(1 - p(z))}{z(2 - 2p(z) + z)}$$

where, as in the statement of theorem 3.3, $p(z) = \dfrac{2z(4 - z)}{4 + z(4 - z)}$.

5.10. THEOREM. — Under each of the following assumptions there exists some continuous version on $(\mathscr{F}, \rho_P)$ of a brownian bridge relating to P, $G_P^{(n)}$, such that $\Pr(\| \nu_n - G_P^{(n)} \|_{\mathscr{F}} \geq \alpha_n) \leq \beta_n$ where (α_n) and (β_n) are defined below (we recall that $F \in \mathscr{L}^{2+\delta}(P)$ and that $e_F^{(2)}$ and $d_F^{(2)}$ are defined in Section 2):

a) if $d_F^{(2)}(\mathscr{F}) = 2d < \infty$, $\alpha_n = \beta_n = 0(n^{-\tau})$ for any $\tau < \gamma(\delta, d)$;
a') if $N_F^{(2)}(\varepsilon, \mathscr{F}) \leq C\varepsilon^{-2d}(L\varepsilon^{-1})^d$ for any ε in $]0, 1[$,

$$\alpha_n = \beta_n = 0((Ln)^{(1/2)+d} n^{-\gamma(\delta,d)});$$

b) if $e_F^{(2)}(\mathscr{F}) = \zeta < 2$, $\alpha_n = 0((Ln)^{-\tau})$ and $\beta_n = 0((Ln)^{-s})$ for any $\tau < \beta(\zeta)$ and any positive s.

Proof of theorem 5.10. — Let σ be an oscillation rate (depending on n) and Π_σ be a projection of $\mathscr{F}$ on a σ-net $\mathscr{F}(\sigma)$ relating to P. We approximate ν_n uniformly over $\mathscr{F}$ by $\nu_n \circ \Pi_\sigma$.

Setting $\mathscr{G}_\sigma = \{ f - g, \rho_P(f, g) \leq \sigma \}$, we may apply theorem 5.3 to $\mathscr{G}_\sigma$ (changing d into $2d$ if necessary), hence the quantity $\| \nu_n - \nu_n \circ \Pi_\sigma \|_{\mathscr{F}} \leq \| \nu_n \|_{\mathscr{G}_\sigma}$ can be evaluated. Besides, let $F_{n,\sigma}$ be the distribution of $\nu_{n|\mathscr{F}(\sigma)}$ on the k-dimensional vector space $l^\infty(\mathscr{F}(\sigma))$ and let G_σ be the corresponding gaussian distribution.

Writing π_∞ for the Prokhorov distance associated to $\| \cdot \|_{\mathscr{F}(\sigma)}$ and applying Strassen's theorem [49], there exists a probability space $(\Omega', \mathscr{A}', \Pr')$ and two random variables $\nu_n(\sigma)$ and $G(\sigma)$ with respective distributions $F_{n,\sigma}$ and G_σ on $l^\infty(\mathscr{F}(\sigma))$ such that:

$$\Pr'(\| \nu_n(\sigma) - G(\sigma) \|_{\mathscr{F}(\sigma)} > B) \leq B$$

where $B = \pi_\infty(F_{n,\sigma}, G_\sigma)$.

So, using lemma 5.7, we may ensure the existence of some regular version of a brownian bridge G_P relating to P such that $G(\sigma) = G_{P|\mathscr{F}(\sigma)}$ and then, applying lemma 5.8 with $V : \omega \to \nu_{n|\mathscr{F}(\sigma)}$, we may assume that G_P is constructed on Ω with $\Pr(\| \nu_{n|\mathscr{F}(\sigma)} - G_{P|\mathscr{F}(\sigma)} \geq B) \leq B$.

Hence, noticing that $\| \nu_{n|\mathscr{F}(\sigma)} - G_{P|\mathscr{F}(\sigma)} \|_{\mathscr{F}(\sigma)} = \| (\nu_n - G_P) \circ \Pi_\sigma \|_{\mathscr{F}}$, we get:

$$\Pr(\| \nu_n - G_P \|_{\mathscr{F}} > 2t + B) \leq A + B + C$$

where $A = \Pr(\| \nu_n \|_{\mathscr{G}_\sigma} > t)$ and $C = \Pr(\| G_P \|_{\mathscr{G}_\sigma} > t)$.

Theorem 5.6 is used to control B (with $\| \cdot \|_{\mathscr{F}(\sigma)} \leq \| \cdot \|_2 \leq \sqrt{k} \| \cdot \|_{\mathscr{F}}$) noticing that $k \leq N_F^{(2)}\left(\frac{\sigma}{2}, \mathscr{F}\right)$ according to remark 2.6.

Moreover C is evaluated with the help of theorem 4.1, so the calculations are completed via an appropriate choice of t and σ.

Under assumption a) *or* a'). — it is enough to prove 5.10.*a'*). We choose $t = 0(n^{-\gamma(\delta,d)}(Ln)^{(1/2)+d})$ and $\sigma = 0(n^{-\gamma(\delta,d)})$.

Under assumption b). — We may assume that $k \leq \exp(C\sigma^{-\zeta})$ with $\zeta > 0$. We choose $t = 0((Ln)^{-(a-b)})$ with $b = (2 - 2p(\zeta) + \zeta)^{-1}$, $a < \frac{1}{\zeta}$ and a in the neighbourhood of $\frac{1}{\zeta}$.

Comment. — It is shown in [*40*] that a Prokhorov or Lipschitz distance may be defined on $l_T^\infty(\mathscr{F})$. In the notations of the above theorem $\mathrm{Max}(\alpha_n, \beta_n)$ is an upper bound for these distances between the distributions of ν_n and G_P.

6. STRONG INVARIANCE PRINCIPLES WITH SPEEDS OF CONVERGENCE

The method to deduce strong approximations from the preceding weak invariance principles is the one used in [*43*] to prove theorem 2: the weak estimates are used locally, giving strong approximations with the help of maximal inequalities and via the Borel-Cantelli lemma.

Maximal inequalities. — As was noticed in [*23*], the proofs of the following inequalities may be deduced from those given in [*10*] and in [*32*] (and hold without any measurability assumption).

Notation. — We set $X_j = \delta_{x_j} - P$ for any integer j.

6.1. LEMMA *(Ottaviani's inequality)*. — We set $S_k = \sum_{j=1}^{k} X_j$. Then, for any positive α, the following inequality holds:

$$(1 - c) \Pr^* (\max_{k \leq n} \| S_k \|_{\mathscr{F}} > 2\alpha) \leq \Pr^*(\| S_n \|_{\mathscr{F}} > \alpha)$$

where $c = \max_{k \leq n} \Pr^* (\| S_k \|_{\mathscr{F}} > \alpha)$.

More precisely, for symmetrical variables, the following sharper inequality is available:

6.2. LEMMA *(Paul Lévy's inequality)*. — Let $(Y_i)_{1 \leq i \leq n}$ be independent and identically distributed B-valued random variables where $(B, \|\cdot\|)$ is a normed vector space. If we assume that Y_1 is symmetrical then: $\Pr^*(\max_{k \leq n} \| S_k \| > \alpha) \leq 2 \Pr^*(\| S_n \| > \alpha)$ holds for any positive α, where

$$S_k = \sum_{j=1}^{k} Y_j.$$

Strong approximations for the empirical brownian bridge.

6.3. THEOREM. — Let $\gamma(\cdot,\cdot)$ and $\beta(\cdot)$ be as defined in 5.9.

Under each of the following assumptions some sequence $(Y_j)_{j\geq 1}$ of independent versions of brownian bridges relating to P that are uniformly continuous on $(\mathscr{F}, \rho_P)$ may be defined on Ω such that:

a) if $d_F^{(2)}(\mathscr{F}) = 2d < \infty$,

$$\frac{1}{\sqrt{n}} \left\| \sum_{j=1}^{n} (X_j - Y_j) \right\|_{\mathscr{F}} = 0(n^{-\alpha}) \qquad \text{a. s.}$$

for any $\alpha < \dfrac{\gamma(\delta, d)}{2(1 + \gamma(\delta, d))}$;

a') if, more precisely $N_F^{(2)}(\varepsilon, \mathscr{F}) \leq C\varepsilon^{-2d}(1 + L\varepsilon^{-1})^d$ for any ε in $]0, 1[$,

$$\frac{1}{\sqrt{n}} \left\| \sum_{j=1}^{n} (X_j - Y_j) \right\|_{\mathscr{F}} = 0(n^{-\gamma(\delta,d)/(2(1+\gamma(\delta,d)))}((Ln)^{(1/2)+d} + (Ln)^{(5/4)+(d/2)})) \quad \text{a. s.};$$

b) if $e_F^{(2)}(\mathscr{F}) = \zeta < 2$,

$$\frac{1}{\sqrt{n}} \left\| \sum_{j=1}^{n} (X_j - Y_j) \right\|_{\mathscr{F}} = 0(Ln^{-(\beta/2)}) \qquad \text{a. s.}$$

for any $\beta < \beta(\zeta)$.

Proof of theorem 6.3. — Note that in order to prove 6.3.*a*) it is enough to prove 6.3.*a'*). We decompose the time into pieces: $H_1, \ldots, H_i, \ldots$ with $H_i = [t_i, t_{i+1}[$.

We choose $t_i = [i^a(Li)^b]$ under assumption *a'*) and $t_i = [\exp(i^{1-\tau})]$

under assumption $b)$, writing γ instead of $\gamma(\delta, d)$ for short: $a = \frac{1+\gamma}{\gamma}$, $b = a(d + 3/2)$ and $\tau = \frac{\beta(\zeta)}{1+\beta(\zeta)}$.

Moreover we set $n_i = t_{i+1} - t_i$. Let σ_i be the rate of oscillation depending on n_i in the same way as σ depends on n in the proof of 5.10. We write Π_i for the projection of $\mathscr{F}$ on some σ_i-net $\mathscr{F}_i$ and we set $D_i = l^\infty(\mathscr{F}_i)$. Then, from the proof of 5.10, we have (choosing $s = 2/(1-\tau)$ in case $b)$):

$$\pi_\infty(F_{n_i,\sigma_i}, G_{\sigma_i}) \leq \beta_{n_i},$$

$$\Pr\left(\left\| \sum_{j \in H_i} (X_j - X_j \circ \Pi_i) \right\|_{\mathscr{F}} \geq \sqrt{n_i}\alpha_{n_i}\right) \leq \beta_{n_i}. \tag{1}$$

So, using Strassen's theorem, some random variables $(V_{i,j}, W_{i,j})_{j \in H_i}$ can be built on $(\Omega', \mathscr{A}', \Pr')$ so that the distributions of $V_{i,j}$ and $W_{i,j}$ are respectively those of $X_{j|\mathscr{F}_i}$ and G_{σ_i} and such that:

$$\Pr'\left(\left\| \sum_{j \in H_i} (V_{i,j} - W_{i,j}) \right\|_{\mathscr{F}_i} \geq \sqrt{n_i}\beta_{n_i}\right) \leq \beta_{n_i}.$$

Moreover we may assume that these variables are independent for different values of i. Writing $i(j)$ for the index fulfilling $j \in H_{i(j)}$, we set $V_j = V_{i(j),j}$ and $W_j = W_{i(j),j}$. Besides, theorem 4.1 ensures the existence of some distribution g_P, defined on the Polish space $\mathscr{U}_P(\mathscr{F})$ of real valued uniformly continuous functions on $(\mathscr{F}, \rho_P)$, of a brownian bridge relating to P.

Then, we can apply lemma 5.7 with $R_1 = (\mathscr{U}_P(\mathscr{F}))^{\mathbb{N}^*}$ and $R_2 = R_3 = \prod_{j \geq 1} D_{i(j)}$, writing Q_1 for the distribution of $(Y_j, Y_{j|\mathscr{F}_{i(j)}})$ (where $(Y_j)_{j \geq 1}$ is the random variable whose distribution is canonically $g_P^{\otimes \mathbb{N}^*}$) and Q_2 stands for the distribution of $(W_j, V_j)_{j \geq 1}$. So, we may assume that $W_j = Y_{j|\mathscr{F}_{i(j)}}$.

Then, using lemma 5.8, we may assume that $(Y_j)_{j \geq 1}$ is built on Ω and that $V_j = X_{j|\mathscr{F}_{i(j)}}$.

As $s > \beta(\zeta)$ we may suppose that $\beta_n \leq \alpha_n$ for any integer n.

Since bound (1) still holds when replacing X_j by Y_j we get:

$$\Pr\left(\left\| \sum_{j \in H_i} (X_j - Y_j) \right\|_{\mathscr{F}} \geq 3\sqrt{n_i}\alpha_{n_i}\right) \leq 3\beta_{n_i} \quad \text{for any integer } i. \tag{2}$$

So, using the inequality:

$$\Big\| \sum_{j=1}^{n} (X_j - Y_j) \Big\|_{\mathcal{F}} \leq A_{r(n)} + B_{r(n)} + C_{r(n)}$$

where $A_r = \sum_{i=1}^{r} \Big\| \sum_{j\in H_i} (X_j - Y_j) \Big\|_{\mathcal{F}}$, $B_r = \max_{m\in H_r} \Big\| \sum_{j=t_r}^{m} X_j \Big\|_{\mathcal{F}}$ and $C = \max_{m\in H_r} \Big\| \sum_{j=t_r}^{m} Y_j \Big\|_{\mathcal{F}}$ the problem reduces to the almost sure control of A, B and C. We notice that:

$n_i = 0(i^{a-1}(Li)^b)$ and then $\beta_{n_i} = 0(i^{-1}(Li)^{-1-\varepsilon})$ with $\varepsilon = b\gamma - \frac{3}{2} - d$ in case a').
$n_i = 0(i^{-\tau}\exp(i^{1-\tau}))$ and then $\beta_{n_i} = 0(i^{-2})$ in case b).

So, in both cases a') and b), the series $\sum_i \beta_{n_i}$ is convergent.

Control of A. — The Borel-Cantelli lemma ensures, because of (2), that $A_r = 0\Big(\sum_{i=1}^{r} \sqrt{n_i}\alpha_{n_i}\Big)$ a. s.

Under assumption a').

$$\sum_{i=1}^{r} \sqrt{n_i}\alpha_{n_i} = 0(r^{(1/2)}(Lr)^{(1/2)+d+b((1/2)-\gamma)}).$$

So, since $(r(n))^a(Lr(n))^b \leq n + 1$ we get:

$$A_{r(n)} = 0((Ln)^{(1/2)+d} n^{1/(2(1+\gamma))}) \qquad \text{a. s.}$$

Under assumption b)

$$\sum_{i=1}^{r} \sqrt{n_i}\alpha_{n_i} = 0\left(\int_1^r x^{-(\tau/2)-(1-\tau)\beta} \exp\left(\frac{x^{1-\tau}}{2}\right) dx\right),$$

hence, integrating by parts:

$$\sum_{i=1}^{r} \sqrt{n_i}\alpha_{n_i} = 0\left(r^{(\tau/2)-(1-\tau)\beta} \exp\left(\frac{r^{1-\tau}}{2}\right)\right).$$

Then, since $r(n) = 0(\mathrm{L}n^{1/(1-\tau)})$, we get:

$$\mathrm{A}_{r(n)} = 0(\sqrt{n}(\mathrm{L}n)^{(\tau/(2(1-\tau)))-\beta}).$$

Control of B. — Using Ottaviani's inequality 6.1, we get (because of ($\mathscr{M}$) Pr* is here unuseless):

$$\Pr(\mathrm{B}_r > 2\mathrm{T}_r)(1 - c_r) \leq \Pr\left(\left\|\sum_{j\in\mathrm{H}_r}\mathrm{X}_j\right\|_{\mathscr{F}} > \mathrm{T}_r\right)$$

where $c_r = \max_{m\in\mathrm{H}_r}\Pr\left(\left\|\sum_{j=1}^{m-t_r}\mathrm{X}_j\right\|_{\mathscr{F}} > \mathrm{T}_r\right)$. With $\mathrm{T}_r = \mathrm{K}\sqrt{\mathrm{L}n_r}\sqrt{n_r}$ in case a') and $\mathrm{T}_r = \mathrm{K}\sqrt{\mathrm{LL}n_r}\sqrt{n_r}$ in case b).

As a consequence of theorem 5.3 applied with $\sigma^2 = \mathrm{P}(\mathrm{F}^2)$, we have that $c_r \to 0$ for $\mathrm{T}_r n_r^{-(1/2)} \to \infty$ as $r \to \infty$. So:

$$\Pr(\mathrm{B} > 2\mathrm{T}_r) = 0(\mathrm{B}' + \mathrm{B}'')$$

where $\mathrm{B}' = \Pr\left(\left\|\sum_{j\in\mathrm{H}_r}(\mathrm{X}_j - \mathrm{Y}_j)\right\|_{\mathscr{F}} > \frac{\mathrm{T}_r}{2}\right)$ and $\mathrm{B}'' = g_{\mathrm{P}}\left(y\in\mathscr{U}_{\mathrm{P}}(\mathscr{F}), \|y\|_{\mathscr{F}} > \frac{\mathrm{T}_r}{2\sqrt{n_r}}\right)$.

Then, using inequality (2) on the one hand and theorem 4.1 with $\sigma^2 = \mathrm{P}(\mathrm{F}^2)$ and via an appropriate choice of K on the other hand, we get that in both cases a') and b): $\mathrm{B}' = 0(\beta_{n_r})$ and $\mathrm{B}'' = 0(r^{-2})$.

Hence, using the Borel-Cantelli lemma again, we get $\mathrm{B}_r = 0(\mathrm{T}_r)$ a. s. So:

Under assumption a')

$$\mathrm{T}_r = 0(\sqrt{\mathrm{L}r}(r^{a-1}(\mathrm{L}r)^b)^{1/2})$$

hence, since $(r(n))^a(\mathrm{L}r(n))^b \leq n + 1$:

$$\mathrm{B}_{r(n)} = 0(n^{1/(2(1+\gamma))}(\mathrm{L}n)^{(5/4)+(d/2)}) \qquad \text{a. s.}$$

Under assumption b)

$$\mathrm{T}_r = 0\left(\sqrt{\mathrm{L}r}\, r^{-(\tau/2)}\exp\left(\frac{r^{1-\tau}}{2}\right)\right)$$

hence, since $r(n) = 0(\mathrm{L}n^{1/(1-\tau)})$:

$$\mathrm{B}_{r(n)} = 0(\sqrt{n\mathrm{LL}n}(\mathrm{L}n)^{-\tau/(2(1-\tau))}) \qquad \text{a. s.}$$

Control of C. — Paul Lévy's inequality 6.2 gives:

$$\Pr(\mathrm{C}_r > \mathrm{T}_r) \leq 2g_{\mathrm{P}}(y\in\mathscr{U}_{\mathrm{P}}(\mathscr{F})/\|y\|_{\mathscr{F}} > \mathrm{T}_r/\sqrt{n_r})$$

so, the above control of B still holds for C.

Then 6.3 follows from collecting the almost sure estimates of A, B and C.

Comments. — When passing from weak invariance principles to strong ones, the speeds of convergence are transformed as follows within our framework.

i) $n^{-\gamma} \to n^{-\gamma(2(1+\gamma))}$ in case *a)*,

ii) $Ln^{-\beta} \to Ln^{-\beta/2}$ in case *b)*.

Transformation *ii)* appears in theorem 6.1 (under 6.3) from [*23*], but it is not the case for transformation *i)* in the same theorem (under 6.4).

On the contrary transformation *i)* is present in finite dimensional principles and appears to be optimal in that case: more precisely, the rate of weak convergence towards the gaussian distribution for 3-integrable variables is ranging about $n^{-1/2}$ when the rate of strong convergence is ranging about $n^{-1/6}$ (see [*39*] for the upper bound and [*9*] for the lower bound), in the real case.

Application to V. Č.-classes. — Applying theorem 6.3 with $\delta = 1$ in the case where $\mathscr{F}$ is a V. Č.-class with real density d, we get a speed of convergence towards the brownian bridge that is $0(n^{-\alpha})$ for any $\alpha < \dfrac{1}{18 + 20d}$. This improves on 1.3.3 but is less sharp than 1.2.4 in the classical case of quadrants in $\mathbb{R}^d$.

6.4. INVARIANCE PRINCIPLES IN C(S). — Following an idea from Dudley in [*21*] (sec. 11), the study of the general empirical processes theoretically allows one to deduce some results about random walks in general Banach spaces. As an application of this principle let us consider a compact metric space (S, κ) and the space C(S) of real continuous functions on S, equipped with the uniform norm $\|\cdot\|_\infty$. Let X be the space of Lipschitz-functions on S equipped with the Lipschitz-norm:

$$\|\cdot\|_{BL} : x \to \|x\|_\infty + \sup_{t \neq s} \frac{|x(t) - x(s)|}{\kappa(s,t)}.$$

We write $N(\varepsilon, S, \kappa)$ for the maximal cardinality of a subset R of S such that $\kappa(s,t) > \varepsilon$ for any $s \neq t$ in R.

We may apply our results through the following choices:

$$\mathscr{F} = \{ \delta_s, s \in S \} \quad \text{and} \quad F = \|\cdot\|_L .$$

Then $(X, \|\cdot\|_\infty)$ is a Suslin space (but is not Polish in general), so $\mathscr{F}$ fulfills $(\mathscr{M})$. Moreover, for any distribution Q in $\mathscr{P}_F^{(2)}(X)$ we have:

$$Q((\delta_s - \delta_t)^2) \leq \kappa^2(s, t)Q(F^2),$$

so $N_F^{(2)}(., \mathscr{F}) \leq N(., S, \kappa)$. Besides $\|\cdot\|_\infty = \|\cdot\|_{\mathscr{F}}$.

Therefore, considering a sequence $(X_j)_{j \geq 1}$ of independent and identically distributed C(S)-valued random variables such that:

$$|X_1(s) - X_1(t)| \leq M\kappa(s, t), \quad \text{for any } s, t \text{ in } S,$$

with $E(M^{2+\delta}) < \infty$ and $E(X_1^{2+\delta}(t_0)) < \infty$ for one t_0 in S, we can apply theorem 5.10 or 6.3 to get speeds of convergence towards the gaussian distribution, whose structure depends on $N(., S, \kappa)$ (the central limit theorem for such uniformly Lipschitzian processes as above is due to Jain and Marcus in [*31*]).

APPENDIX

1. Proof of lemma 3.1.

First let us recall Hoeffding's lemma (see [*29*]).

Hoeffding's lemma. — Let S be a centered and $[u, v]$-valued random variable, then:

$$E(\exp(tS)) \leq \exp\left(\frac{t^2(v-u)^2}{8}\right), \quad \text{for any } t \text{ in } \mathbb{R}.$$

We may assume that w is chosen as follows:

. drawing—with uniform distribution—a partition $\mathcal{J} = (J_i)_{1 \leq i \leq n}$ such that $|J_i| = m$ for each i in $[1, n]$.

. then, drawing an index $w(i)$ independently in each J_i—with uniform distribution—. The following evaluations are conditional on $\mathcal{J}$ but the last bound will not depend on $\mathcal{J}$, giving 3.1.

We set $Z = \frac{\tilde{S}_n}{n} - \frac{S_N}{N}$ and we write Λ for the logarithm of the conditional Laplace transform of Z. Then setting $\bar{\xi}_i = \frac{1}{m} \sum_{j \in J_i} \xi_j$, we have, for any s in $\mathbb{R}$:

$$\Lambda(s) = \sum_{i=1}^{n} \operatorname{Log}\left(\frac{1}{m} \sum_{j \in J_i} \exp\left(\frac{s}{n}(\xi_j - \bar{\xi}_i)\right)\right).$$

Then, since the logarithm is a concave function:

$$\Lambda(s) \leq n \operatorname{Log}\left(\frac{1}{N} \sum_{j=1}^{N} \exp\left(\frac{s}{n}\left(\xi_j - \frac{S_N}{N}\right)\right)\right) = n\Lambda_N\left(\frac{s}{n}\right)$$

where, writing Q_N for the uniform distribution on $\{\xi_1, \ldots, \xi_N\}$, Λ_N stands for the logarithm of the Laplace transform under Q_N of $x \to x - E_{Q_N}(x)$. Therefore the Cramér-Chernoff transform of Z is larger than that of $\frac{S'_n}{n} - E_{Q_N^{\otimes n}}(S'_n/n)$ under $Q_N^{\otimes n}$ where S'_n stands for the sum of n i. i. d. random variables with common distribution Q_N.

Then, Hoeffding [*29*] and Bernstein [*5*] inequalities yield 3.1.1°) and 3.1.2°). In order to prove 3.1.3°) we may assume that $S_N = 0$ (otherwise changing ξ_j into $\xi_j - S_N/N$).

Then, applying Hoeffding's lemma to the conditionally centered random variables $\xi_{w(i)} - \bar{\xi}_i$ and setting $u_i = \min_{j \in J_i} \xi_j$ and $v_i = \max_{j \in J_i} \xi_j$, we get:

$$l_i(t) = \operatorname{Log} E^{\mathcal{J}}(\exp(t(\xi_{w(i)} - \bar{\xi}_i))) \leq \frac{(v_i - u_i)^2}{8} t^2, \quad \text{for any } t \text{ in } \mathbb{R}.$$

Hence

$$\Lambda(s) = \sum_{i=1}^{n} l_i\left(\frac{s}{n}\right) \leq \frac{s^2}{8n^2} \sum_{i=1}^{n} (v_i - u_i)^2 \leq \frac{s^2}{4n^2} \sum_{i=1}^{n} \sum_{j \in J_i} \xi_j^2$$

and therefore

$$\Lambda(s) \leq \frac{s^2}{4n} m\sigma_N^2$$

yielding 3.1.3°) via Markov's inequality.

2. The distribution of the supremum of a *d*-dimensional parameter brownian bridge

Goodman's work in [*28*] and Cabaña's in [*11*] give a lower bound of the probability for the supremum of a brownian bridge to cross a barrier.

Notations. — We set $I = [0, 1]$ and write for any integer d, 1_d for the element $(1, \ldots, 1)$ of $\mathbb{R}^d$. Moreover, for any s in I^d, we set $p(s) = s_1 \ldots s_d$.

A.1. THEOREM. — Let d be an integer and W_d be some standard d-dimensional parameter Wiener process, then, on the one hand:

i)
$$\Pr(\sup_{s \in I^d} W_d(s) \leq t \mid W_d(1_d) = at) \leq h_d(a, t)$$

for almost any real number a (in Lebesgue sense) and any positive t, where

$$h_d(a, t) = \left(1 + \exp(2t^2(a-1)) \sum_{i=0}^{d-1} (-1)^{i+1} \frac{(2t^2(a-1))^i}{i!}\right) I_{]-\infty, 1]}(a)$$

and on the other hand:

ii)
$$\Pr(\sup_{s \in I^d} W_d(s) - p(s)W_d(1_d) > t) \geq \sum_{i=0}^{d-1} \frac{(2t^2)^i}{i!} . \exp(-2t^2).$$

Proof of theorem A.1. — If $d = 2$ the whole proof is contained in [*28*]. For $d \geq 3$ it follows easily from Cabaña [*11*].

Comment. — Theorem A.1 was proved by ourself (see [*40*] and [*41*]) as well as by E. Cabaña in [*11*] (3). In another connection, inequality A.1 *ii)* ensures that some polynomial factor $t^{2h(d)}$ with $h(d) \geq d - 1$ cannot be removed in bounds 3.3.1°) *a)* and 4.1.2.

3. Making an exponential bound explicit.

The calculations yielding 3.3.1°) *a)* are slightly modified here, where the entropy condition *a)* is replaced with a more explicit one.

A.2. THEOREM. — If we assume that $\mathscr{F}$ is $[0, 1]$-valued and that

a') $N_1^{(2)}(\varepsilon, \mathscr{F}) \leq K^{1 + 1/\mathrm{Log}(\varepsilon^{-2})}(1 + \mathrm{Log}(\varepsilon^{-2}))^d \varepsilon^{-2d}$ for any ε in $]0, 1[$ then, an upper bound for $\Pr(\| v_n \|_{\mathscr{F}} > t)$ is, for any t in $[1, +\infty[$, given by:

$$4H(t) \exp(13) \exp(-2t^2) + 4H^2(t) \exp(-(t^2 - 5)(Lt)^2)$$

where
$$H(t) = K^{6/5} \exp(16d)(1 + Lt^2)^{5d} t^{6d}.$$

(3) Thanks to M. Wschebor and J. Leon for communicating this reference to us.

Proof of A.2. — In the proof of 3.3.1°) *a*) we choose $\alpha = 1 + Lt^2/LLt^2$, then

$$A \leq 2H(t) \exp(13) \exp(-2t^2) \qquad P^{\otimes N}\text{-a. s.}$$
$$B \leq 2H^2(t) \exp(-(t^2 - 5)(Lt)^2) \qquad P^{\otimes N}\text{-a. s.}$$

whenever $t^2 \geq 6 + 4d$, yielding A.2 via lemma 3.2.

Comment. — Assumption *a'*) is typically fulfilled whenever $\mathcal{F}$ is a V. Č.-class. In that case *d* may be the real density of $\mathcal{F}$ (if it is « achieved ») or the integer density of $\mathcal{F}$ (see the proof of 3.6).

ACKNOWLEDGEMENT

Professor J. Bretagnolle played an essential part in the conception as well as in the realization of this work. The author wishes to express his gratitude to him especially for his patient and numerous readings of this paper.

REFERENCES

[1] K. Alexander, *Ph. D. Dissertation*, Mass. Inst. Tech., 1982.
[2] K. Alexander, Probability inequalities for empirical processes and a law of iterated logarithm. *Annals of Probability*, t. **12**, 4, 1984, p. 1041-1067.
[3] P. Assouad, Densité et dimension, *Ann. Inst. Fourier*, Grenoble, t. **33**, 3, 1983, p. 233-282.
[4] N. S. Bakhvalov, On approximate calculation of multiple integrals (in Russian). *Vestnik Mosk. Ser. Mat. Mekh. Astron. Fiz. Khim.*, t. **4**, 1959, p. 3-18.
[5] G. Bennett, Probability inequalities for sums of independent random variables. *J. Amer. Statist. Assoc.*, t. **57**, 1962, p. 33-45.
[6] I. Berkes, W. Philipp, Approximation theorems for independent and weakly dependent random vectors. *Ann. Probability*, t. **7**, 1979, p. 29-54.
[7] P. Billingsley, *Convergence of probability measures*. Wiley, New York.
[8] I. S. Borisov, *Abstracts of the Colloquium on non parametric statistical inference*, Bundapest, 1980, p. 77-87.
[9] L. Breiman, On the tail behavior of sums of independent random variables. *Z. Warschein. Verw. Geb.*, t. **9**, 1967, p. 20-25.
[10] L. Breiman, *Probability. Reading Mass.*, Addison-Wesley, 1968.
[11] E. Cabaña, On the transition density of a multidimensional parameter Wiener process with one barrier. *J. Appl. Prob.*, t. **21**, 1984, p. 197-200.
[12] E. Cabaña, M. Wschebor, The two-parameter Brownian bridge. *Annals of Probability*, t. **10**, 2, 1982, p. 289-302.
[13] D. L. Cohn, *Measure theory*. Birkhaüser, Boston, 1980.
[14] M. Csörgo, P. A. Révész, A new method to prove Strassen type laws of invariance principle II. *Z. Warschein. Verw. Geb.*, t. **31**, 1975, p. 261-269.
[15] H. Dehling, Limit theorems for sums of weakly dependent Banach space valued random variables. *Z. Warschein. Verw. Geb.*, 1983, p. 391-432.
[16] L. Devroye, Bounds for the uniform deviations of empirical measures. *J. of Multivar. Anal.*, t. **13**, 1982, p. 72-79.
[17] Hu Inchi, A uniform bound for the tail probability of Kolmogorov-Smirnov statistics. *The Annals of Statistics*, t. **13**, 2, 1985, p. 821-826.

[18] R. M. DUDLEY, The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Functional Analysis*, t. **1**, 1967, p. 290-330.
[19] R. M. DUDLEY, Metric entropy of some classes of sets with differential boundaries. *J. Approximation Theory*, t. **10**, 1974, p. 227-236.
[20] R. M. DUDLEY, Central limit theorems for empirical measures. *Ann. Probability*, t. **6**, 1978, p. 899-929 ; *correction*, t. **7**, 1979, p. 909-911.
[21] R. M. DUDLEY, Saint-Flour, 1982. *Lecture Notes in Mathematics* n° 1097.
[22] R. M. DUDLEY, DURST, Empirical Processes, Vapnik-Červonenkis classes and Poisson processes. *Proba. and Math. Stat. (Wroclaw)*, t. **1**, 1981, p. 109-115.
[23] R. M. DUDLEY, W. PHILIPP, Invariance principles for sums of Banach spaces valued random elements and empirical processes. *Z. Warschein. Verw. Geb.*, t. **82**, 1983, p. 509-552.
[24] A. DVORETZKY, J. C. KIEFER, J. WOLFOWITZ, Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Stat.*, t. **33**, 1956, p. 642-669.
[25] X. FERNIQUE, Régularité de processus gaussiens. *Invent. Math.*, t. **12**, 1971, p. 304-320.
[26] P. GAENSSLER, W. STUTE, Empirical processes: a survey of results for independent and identically distributed random variables. *Ann. Proba.*, t. **7**, p. 193-243.
[27] M. E. GINE, J. ZINN, On the central limit theorem for empirical processes. *Annals of Probability*, t. **12**, 4, 1984, p. 929-989.
[28] V. GOODMAN, Distribution estimations for functionals of the two parameter Wiener process. *Annals of Probability*, t. **4**, 6, 1976, p. 977-982.
[29] W. HOEFFDING, Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, t. **58**, 1963, p. 13-30.
[30] I. A. IBRAGIMOV, R. Z. KHASMINSKII, On the non-parametric density estimates. *Zap. Naucha. Semin. LOMI*, t. **108**, 1981, p. 73-81. In Russian.
[31] N. JAIN, M. B. MARCUS, Central limit theorem for C(S)-valued random variables. *J. Functional Analysis*, t. **19**, 1975, p. 216-231.
[32] J. P. KAHANE, *Some random series of functions*. Lexington, Mass. D. C. Heuth, 1968.
[33] S. KARLIN, H. M. TAYLOR, *A first course in Stochastic Processes*, 1971, Academic Press, New-York.
[34] J. C. KIEFER, On large deviations of the empirical d. f. of vector chance variables and a law of iterated logarithm. *Pacific J. Math.*, t. **11**, 1961, p. 649-660.
[35] J. C. KIEFER, J. WOLFOWITZ, On the deviations of the empiric distribution function of vector variables. *Trans. Amer. Math. Soc.*, t. **87**, 1958, p. 173-186.
[36] A. N. KOLMOGOROV, V. M. TIKHOMIROV, ε-entropy and ε-capacity of sets in functional spaces. *Amer. Math. Soc. Transl.* ser. 2, t. **17**, 1961, p. 277-364.
[37] J. KOMLOS, P. MAJOR, G. TUSNADY, An approximation of partial sums of independent RV' s and the sample DF. 1. *Z. Warschein. Verw. Geb.*, t. **32**, 1975, p. 111-131.
[38] L. LE CAM, *A remark on empirical measures*. In *A Fetscheift for Erich L. Lehmam in Honor of his Sixty-Fifth Birthday*, 1983, p. 305-327 Wadsworth, Belmont, California.
[39] P. MAJOR, The approximation of partial sums of independent RV' s. *Z. Warschein. Verw. Geb.*, t. **35**, 1976, p. 213-220.
[40] P. MASSART, *Vitesse de convergence dans le théorème de la limite centrale pour le processus empirique*. Thèse de 3^e^ cycle n° 3545 de l'Université de Paris-Sud, 1983.
[41] P. MASSART, Vitesses de convergence dans le théorème central limite pour des processus empiriques. *Note aux C. R. A. S.*, t. **296**, 20 juin 1983, Série I, p. 937-940.
[42] P. A. MEYER, Martingales and stochastic integrals I. *Lecture Notes in Mathematics*, t. **284**.
[43] W. PHILIPP, Almost sure invariance principles for sums of B-valued random variables. *Lecture Notes in Mathematics*, t. **709**, p. 171-193.

[44] D. POLLARD, A central limit theorem for empirical processes. *J. Australian Math. Soc. Ser. A*, t. **33**, 1982, p. 235-248.

[45] D. POLLARD, *Rates of strong uniform convergence*, 1982. Preprint.

[46] R. J. SERFLING, Probability inequalities for the sum in sampling without replacement. *Ann. Stat.*, t. **2**, 1, 1974, p. 39-48.

[47] M. SION, On uniformization of sets in topological spaces. *Trans. Amer. Math. Soc.*, t. **96**, 1960, p. 237-245.

[48] A. V. SKOROHOD, *Theory Prob. Appl.*, t. **21**, 1976, p. 628-632.

[49] V. STRASSEN, The existence of probability measures with given marginals. *Ann. Math. Stat.*, t. **36**, 1965, p. 423-439.

[50] G. TUSNADY, A remark on the approximation of the sample DF in the multidimensional case. *Periodica Math. Hung.*, t. **8**, 1977, p. 53-55.

[51] V. N. VAPNIK, A. Y. ČERVONENKIS, On the uniform convergence of relative frequencies of events to their probabilities. *Theor. Prob. Appl.*, t. **16**, 1971, p. 264-328.

[52] V. V. YURINSKII, A smoothing inequality for estimates of the Lévy-Prohorov distance. *Theory Prob. Appl.*, t. **20**, 1975, p. 1-10.

[53] V. V. YURINSKII, On the error of the gaussian approximation for convolutions. *Theor. Prob. Appl.*, t. **22**, 1977, p. 236-247.

[54] J. E. YUKICH, *Uniform exponential bounds for the normalized empirical process*, 1985. Preprint.

(Manuscrit reçu le 16 avril 1985)
(révisé le 7 mai 1986)