Concentration around the mean for maxima of empirical processes

T. Klein and E. Rio, *UMR 8100 CNRS, Université de Versailles Saint Quentin en Yvelines, 45 Avenue des Etats Unis, 78035 VERSAILLES.*

Abstract :   In this paper we give optimal constants in Talagrand's concentration inequalities for maxima of empirical processes associated to independent and eventually non identically distributed random variables. Our approach is based on the entropy method introduced by Ledoux.

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of independent random variables with values in some Polish space $\mathcal{X}$ and let $\mathcal{S}$ be a countable class of measurable functions from $\mathcal{X}$ into $[-1, 1]^n$. For $s = (s^1, \ldots s^n)$ in $\mathcal{S}$, we set

$$(1.1) \qquad S_n(s) = s^1(X_1) + \cdots + s^n(X_n).$$

In this paper we are interested in concentration inequalities for $Z = \sup\{S_n(s) : s \in \mathcal{S}\}$.

Now let us recall the main results in this direction. Starting from concentration inequalities for product measures, Talagrand (1996) obtains Bennett type upper bounds on the Laplace transform of Z via concentration inequalities for product measures. More precisely he proves

$$(1.2) \qquad \log \mathbb{E} \exp(tZ) \le t\mathbb{E}(Z) + V ab^{-2}(e^{bt} - bt - 1)$$

for any positive t. Here

$$V = \mathbb{E}\Big(\sup_{s \in \mathcal{S}} \sum_{k=1}^{n} (s^k(X_k))^2\Big).$$

In order to analyze the variance factor V, set

$$(1.3) \qquad\qquad V_n = \sup_{s \in \mathcal{S}} \text{Var}\, S_n(s).$$

Then, one can derive from the comparison inequalities in Ledoux and Talagrand (1991) that $V_n \leq V \leq V_n + 16\mathbb{E}(Z)$ (see Massart (2000), p. 882). Consequently V is often close to the maximal variance V_n. The conjecture concerning the constants is then $a = b = 1$. The constant a plays a fundamental role: in particular, for Donsker classes, $a = 1$ gives the exact rate function in the moderate deviations bandwidth. Nevertheless is seems difficult to reach $a = 1$ via Talagrand's method (see Panchenko (2001) for more about the constants in Talagrand's concentration inequalities for product measures). In order to obtain concentration inequalities more directly, Ledoux (1996) uses a Log-Sobolev type method together with a powerful argument of tensorization of the entropy. When applied to $\exp(tZ)$, this method yields a differential inequality (this is the so-called Herbst argument) on the Laplace transform of Z and gives (1.2) again. Applying Ledoux's method, Massart (2000) obtained $a = 8$ in (1.2) with Talagrand's variance factor and $a = 4$ in (1.2) with the variance factor $V_n + 16\mathbb{E}(Z)$. Later on, Rio (2002) proved (1.2) for independent and identically distributed (iid) random variables (in the iid case $s^1 = \cdots = s^n$) with $a = 1$, $b = 3/2$ and a variance factor $v = V_n + 2\mathbb{E}(Z)$. Next Bousquet (2002) found a nice trick to improve Rio's inequality. He proved (1.2) with $a = b = 1$ and the variance factor v in the iid case. For negative values of t, Klein (2002) obtained (1.2) with $a = 1$, $b = 4$ and v in the iid case.

Here we are interested in optimal constants in Talagrand's inequalities for non identically distributed random variables. Our approach to obtain the best constants is to apply the lemma of tensorization of the entropy proposed by Ledoux (1996). However the differential inequality on the Laplace transform of Z is more involved than in the iid case. Therefore the results are suboptimal in the large deviations bandwidth. We start by right-hand side deviations.

Theorem 1.1. *Let $\mathcal{S}$ be a countable class of measurable functions with values in $[-1, 1]^n$. Suppose that $\mathbb{E}(s^k(X_k)) = 0$ for any $s = (s^1, \ldots s^n)$ in $\mathcal{S}$ and any integer k in $[1, n]$. Let L denote the logarithm of the Laplace transform of Z. Then, for any positive t,*

$$(a) \qquad\qquad L(t) \leq t\mathbb{E}(Z) + \frac{t}{2}(2\mathbb{E}(Z) + V_n)(\exp((e^{2t} - 1)/2) - 1).$$

Consequently, setting $v = 2\mathbb{E}(Z) + V_n$, for any positive x,

$$(b) \qquad\qquad \mathbb{P}(Z \geq \mathbb{E}(Z) + x) \leq \exp\left(-\frac{x}{4}\log\left(1 + 2\log(1 + x/v)\right)\right)$$

2

and

$$(c) \qquad \mathbb{P}(Z \geq \mathbb{E}(Z) + x) \leq \exp\left(-\frac{x^2}{v + \sqrt{v^2 + 3vx} + (3x/2)}\right) \leq \exp\left(-\frac{x^2}{2v + 3x}\right).$$

Remark 1.1. In the spirit of Massart's paper (2000), Theorem 1.1(b) can be improved for large values of x to get a Bennett's type inequality with $a = 1$.

Remark 1.2. Theorem 1.1 applies to set-indexed empirical processes associated to non identically distributed random variables. In that case $s^i(X_i) = 1\!\!1_{X_i \in S} - \mathbb{P}(X_i \in S)$ and consequently the centering constant depends on i. Some different concentration inequalities for set-indexed empirical processes are given in Rio (2001), Theorem 4.2 and Remark 4.1. However, due to the concavity of the polynomial function $u(1 - u)$, the variance factor in Rio (2001) is suboptimal for non identically distributed random variables. Here, as a byproduct of Theorem 1.1(a), we get the upper bound below for the variance of Z.

Corollary 1.1. *Under the assumptions of Theorem 1.1(a),* $\operatorname{Var} Z \leq V_n + 2\mathbb{E}(Z)$.

For left-hand side deviations, the concentration bounds are similar. However the proof is more intricate. We emphasize that the proof of Theorem 1.1 is not relevant for left-hand side deviations. This is the reason why we need to compensate the empirical process for left-hand side deviations.

Theorem 1.2. *Under the assumptions of Theorem 1.1, for any positive t,*

$$(a) \qquad\qquad L(-t) \leq -t\mathbb{E}(Z) + \frac{v}{9}(e^{3t} - 3t - 1).$$

Consequently, for any positive x,

$$(b) \qquad\qquad \mathbb{P}(Z \leq \mathbb{E}(Z) - x) \leq \exp\left(-\frac{v}{9}h\left(\frac{3x}{v}\right)\right),$$

where $h(x) = (1 + x)\log(1 + x) - x$, and

$$(c) \qquad \mathbb{P}(Z \leq \mathbb{E}(Z) - x) \leq \exp\left(-\frac{x^2}{v + \sqrt{v^2 + 2vx} + x}\right) \leq \exp\left(-\frac{x^2}{2v + 2x}\right).$$

Remark 1.3. Theorem 1.2(b) improves on Theorem 1.1, inequality (2) in Klein (2002). However Klein gives additional results for functions with values in $]-\infty, 1]$ and subexponential tails on the left (cf. inequality (3), Theorem 1.1).

Let us now apply Theorems 1.1 and 1.2 to randomized processes, as defined in Ledoux and Talagrand (1991), Section 4.3. Let $X_1, X_2, \ldots X_n$ be a sequence of independent and centered random variables with values in $[-1, 1]$. Let T be some countable set and $\zeta_1, \zeta_2, \ldots, \zeta_n$ be numerical functions on T. Let

$$Z = \sup\left\{ X_1\zeta_1(t) + X_2\zeta_2(t) + \cdots + X_n\zeta_n(t) : t \in T \right\}.$$

The random variable Z corresponds to the class of functions $\mathcal{S} = \{s_t : t \in T\}$, where the components s_t^i of s_t are defined by $s_t^i(x) = x\zeta_i(t)$. Assuming that

$$V_n = \sup_{t \in T} \sum_{k=1}^{n} \zeta_k^2(t) \mathbb{E}(X_k^2) < \infty \text{ and } M = \sup_{k \in [1,n]} \sup_{t \in T} |\zeta_k(t)| < \infty,$$

Corollary 1.1 gives $\operatorname{Var} Z \le V_n + 2\mathbb{E}(Z)$. Let us compare this variance bound with the known results. Theorem 3 in Bobkov (1996) applied to Z yields $\operatorname{Var} Z \le 2V$, where

$$V = \mathbb{E}\left(\sup_{t \in T} \sum_{k=1}^{n} (\zeta_k(t) X_k)^2\right)$$

is Talagrand's variance factor. If the random variables $X_1, X_2, \ldots, X_n$ are symmetric signs, then Z is the maximum of a Rademacher process and $V = V_n$. In that case Corollary 1.1 improves the known bounds on $\operatorname{Var} Z$ as soon as $2\mathbb{E}(Z) < V_n$. For Rademacher processes, the concentration inequality (4.10) in Ledoux and Talagrand (1991) yields

(1.4) $$\mathbb{P}(Z \ge m_Z + x) \le \exp(-x^2/(8V_n)),$$

where m_Z denotes a median of Z. Theorems 1.1 and 1.2 provide exponential bounds with a factor 2 instead of 8. However our variance factor is greater than V_n and our bounds are not subgaussian. Finally, we refer the reader to Bousquet (2002) or Panchenko (2003) for concentration inequalities (with suboptimal variance factor) for randomized or empirical processes in the unbounded case.

2. Tensorization of entropy and related inequalities.

In this section we apply the method of tensorization of the entropy to get an upper bound on the entropy of positive functionnals f of independent random variables $X_1, X_2, \ldots, X_n$.

Notation 2.1. Let $\mathcal{F}_n$ be the σ-field generated by $(X_1, \ldots, X_n)$ and $\mathcal{F}_n^k$ be the σ-field generated by $(X_1, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n)$. Let E_n^k denote the conditional expectation operator associated to $\mathcal{F}_n^k$.

In this paper, the main tool for proving concentration inequalities is the following consequence of the tensorization inequality in Ledoux (1996).

Proposition 2.1. *Let f be some positive $\mathcal{F}_n$-measurable random variable such that $\mathbb{E}(f \log f) < \infty$ and $g_1, g_2, \ldots, g_n$ be any sequence of positive and integrable random variables such that $\mathbb{E}(g_i \log g_i) < \infty$. Then*

$$\mathbb{E}(f \log f) - \mathbb{E}(f) \log \mathbb{E}(f) \le \sum_{k=1}^{n} \mathbb{E}(g_k \log(g_k/E_n^k g_k)) + \sum_{k=1}^{n} \mathbb{E}((f - g_k) \log(f/E_n^k f)).$$

4

Proof. Set $f_k = E_n^k f$. By the tensorization inequality in Ledoux (1996)

(2.1) $$\mathbb{E}(f \log f) - \mathbb{E}(f) \log \mathbb{E}(f) \le \sum_{k=1}^{n} \mathbb{E}(f \log(f/f_k)).$$

Now

(2.2) $$\mathbb{E}(f \log(f/f_k)) = \mathbb{E}(g_k \log(f/f_k)) + \mathbb{E}((f - g_k) \log(f/f_k)).$$

Since $E_n^k(f/f_k) = 1$, we have:

$$\mathbb{E}(g_k \log(f/f_k)) \le \sup\{\mathbb{E}(g_k h) : \ h \ \mathcal{F}_n\text{-measurable}, \ E_n^k(e^h) = 1\}.$$

Hence, from the duality formula for the relative entropy in Ledoux (1996),

$$\mathbb{E}(g_k \log(f/f_k)) \le \mathbb{E}(g_k \log(g_k/E_n^k g_k)).$$

Together with (2.2), it implies that

(2.3) $$\mathbb{E}(f \log(f/f_k)) \le \mathbb{E}(g_k \log(g_k/E_n^k g_k)) + \mathbb{E}((f - g_k) \log(f/f_k)),$$

and Proposition 2.1 follows.

3. Right-hand side deviations.

To prove Theorems 1.1 and 1.2, we start by proving the results for a finite class of functions. The results in the countable case are derived from the finite case using the Beppo Levi lemma. Consequently, throughout the sequel we may assume that $\mathcal{S} = \{s_1, \ldots, s_m\}$.

As mentioned in the introduction, the deviation of Z on the right is easier to handle than the deviation on the left. In fact, for positive t, the functionnal $\exp(tZ)$ is an increasing and convex function with respect to the variables $s_i^k(X_k)$. This is no more the case for negative values of t. Consequently upper bounds for the Laplace transform of Z via the Herbst-Ledoux method are more difficult to handle for negative values of t. In section 4, we will introduce compensated processes in order to handle the deviation on the left.

Definition 3.1. Let τ be the first integer such that $Z = S_n(s_\tau)$. Set $f = \exp(tZ)$ and $f_k = E_n^k(f)$. Let P_n^k denote the conditional probability measure conditionally to $\mathcal{F}_n^k$.

Set

(3.1) $$g_k = \sum_i P_n^k(\tau = i) \exp(tS_n(s_i)).$$

Let F denote the Laplace transform of Z. From Proposition 2.1,

$$(3.2) \qquad tF'(t) - F(t)\log F(t) \le \sum_{k=1}^{n} \mathbb{E}(g_k \log(g_k/E_n^k g_k)) + \sum_{k=1}^{n} \mathbb{E}((f - g_k)\log(f/f_k)).$$

Since $f - g_k \ge 0$, the upper bound on the second term in (3.3) will be derived from Lemma 3.1 below.

Lemma 3.1. *With the notations of Definition 3.1, $\exp(ts_\tau^k(X_k)) \ge (f/f_k) \ge \exp(-2t)$ a.s.*

Proof. Let $S_n^k(s) = S_n(s) - s^k(X_k)$. Let τ_k be the first integer in $[1, m]$ such that

$$(3.3) \qquad S_n^k(s_{\tau_k}) = \sup\{S_n^k(s) : s \in \mathcal{S}\} =: Z_k.$$

Clearly Z_k is a $\mathcal{F}_n^k$-measurable random variable and

$$(3.4) \qquad \exp(tZ_k + t) \ge f \ge \exp(tZ_k)\exp(ts_{\tau_k}^k(X_k)).$$

Since the stopping time τ_k is $\mathcal{F}_n^k$-measurable, $E_n^k(s_{\tau_k}^k(X_k)) = 0$ by the centering assumption on the elements of $\mathcal{S}$. It follows that

$$(3.5) \qquad E_n^k f \ge \exp(tZ_k)E_n^k(\exp(ts_{\tau_k}^k(X_k)) \ge \exp(tZ_k) \ge \exp(tS_n^k(s_\tau)).$$

Hence $f_k \ge f\exp(-ts_\tau^k(X_k))$, which implies the left inequality in Lemma 3.1.

We now prove the second inequality in Lemma 3.1. From the left inequality in (3.4), $\exp(tZ_k + t) \ge E_n^k(f)$. Next, from the right inequality in (3.4), $\exp(tZ_k) \le \exp(tZ + t)$. Hence $f_k \le f\exp(2t)$, which implies the second part of Lemma 3.1. ∎

From Lemma 3.1 and the facts that $f - g_k \ge 0$ and $ts_\tau^k(X_k) \le t$ we get that

$$(3.6) \qquad \mathbb{E}((f - g_k)\log(f/f_k)) \le t\mathbb{E}(f - g_k).$$

We now bound up the first term in (3.2). Set

$$(3.7) \qquad h_k = \sum_{i=1}^{m} P_n^k(\tau = i)\exp(tS_n^k(s_i)).$$

The random variable h_k is positive and $\mathcal{F}_n^k$-measurable. Hence, from the variational definition of the relative entropy,

$$E_n^k(g_k \log(g_k/E_n^k g_k)) \le E_n^k(g_k \log(g_k/h_k) - g_k + h_k).$$

6

Putting this inequality in (3.2) and using (3.6), we get

$$(3.8) \qquad tF' - F \log F \le \sum_{k=1}^{n} \mathbb{E}(g_k \log(g_k/h_k) + (1+t)(h_k - g_k)) + t \sum_{k=1}^{n} \mathbb{E}(f - h_k).$$

In order to bound up the second term on right hand, we will use Lemma 3.2 below.

Lemma 3.2. Let $(h_k)_{k \le n}$ be the finite sequence of random variables defined in (3.7). Then

$$\sum_{k=1}^{n} \mathbb{E}(f - h_k) \le e^{2t} F(t) \log F(t).$$

Proof. Since the random variables $S_n^k(s)$ are $\mathcal{F}_n^k$ measurable,

$$h_k = E_n^k \Big(\sum_{i=1}^{m} \mathbb{1}_{\tau=i} \exp(tS_n^k(s_i)) \Big) = E_n^k \Big(\exp(tS_n^k(s_\tau)) \Big).$$

It follows that

$$(3.9) \qquad \sum_{k=1}^{n} \mathbb{E}(f - h_k) = \sum_{k=1}^{n} \mathbb{E}\Big(f(1 - \exp(-ts_\tau^k(X_k)) - e^{2t}ts_\tau^k(X_k)) \Big) + te^{2t}F'(t).$$

Now, from Lemma 3.1, $ts_\tau^k(X_k) \ge \log(f/f_k) \ge -2t$. Since $1 - \exp(-x) - e^{2t}x$ is a nonincreasing function of x on the interval $[-2t, +\infty[$, it follows that

$$\mathbb{E}(f(1 - \exp(-ts_\tau^k(X_k)) - te^{2t}s_\tau^k(X_k))) \le \mathbb{E}(f - f_k - e^{2t}f \log(f/f_k)).$$

From the equality $\mathbb{E}(f_k) = \mathbb{E}(f)$, we get that

$$\mathbb{E}(f - f_k - e^{2t}f \log(f/f_k)) = -e^{2t}\mathbb{E}(f \log(f/f_k)).$$

Hence, summing on k and applying (2.1),

$$\sum_{k=1}^{n} \mathbb{E}\Big(f(1 - \exp(-ts_\tau^k(X_k)) - te^{2t}s_\tau^k(X_k)) \Big) \le e^{2t}(F \log F - tF'),$$

which, together with (3.9), implies Lemma 3.2. ∎

Next, we bound up the first term on right hand in (3.8).

Definition 3.2. Let $r(t, x) = x \log x + (1+t)(1-x)$.

With the above definition

$$g_k \log(g_k/h_k) + (1+t)(h_k - g_k) = h_k r(t, g_k/h_k).$$

7

From the convexity of r with respect to x,

$$h_k r(t, g_k/h_k) \leq \sum_i P_n^k(\tau = i) \exp(tS_n^k(s_i)) r(t, \exp(ts_i^k(X_k))).$$

which ensures that

(3.10) $\qquad E_n^k(h_k r(t, g_k/h_k)) \leq \sum_i P_n^k(\tau = i) \exp(tS_n^k(s_i)) \mathbb{E}\left(r(t, \exp(ts_i^k(X_k)))\right)$

Here we need the bound below.

Lemma 3.3. *Let r be the function defined in definition 3.2. For any function s in $\mathcal{S}$ and any positive t,*

$$\mathbb{E}\, r(t, \exp(ts^k(X_k))) \leq \frac{t^2}{2} \mathbb{E}(s^k(X_k))^2.$$

Proof. Let $\eta(x) = r(t, e^{tx}) = txe^{tx} + (t+1)(1 - e^{tx})$. We will prove that, for any $x \leq 1$,

(3.11) $\qquad\qquad\qquad\qquad \eta(x) \leq x\eta'(0) + (tx)^2/2.$

Set $\delta(x) = \eta(x) - x\eta'(0) - (tx)^2/2$. Then $\delta(0) = 0$ and $\delta'(x) = t^2(x-1)(e^{tx}-1)$. Consequently $\delta'(x)$ has the same sign as $x(x-1)$, which leads to (3.11). Since the random variables $s^k(X_k)$ are centered, taking $x = s^k(x_k)$ and integrating with respect to the marginal law of X_k, we get Lemma 3.3. ∎

From Lemma 3.3 and (3.10) we have:

(3.12) $\qquad E_n^k(h_k r(t, g_k/h_k)) \leq \frac{t^2}{2} E_n^k\left(\sum_i \mathbb{I}_{\tau=i} \exp(tS_n^k(s_i)) \mathbb{E}(s_i^k(X_k))^2\right).$

Now $\exp(tS_n^k(s_i)) \leq \exp(2t + tS_n(s_i))$, and therefrom

$$\sum_{k=1}^n \mathbb{E}(h_k r(t, g_k/h_k)) \leq \frac{t^2 e^{2t}}{2} \mathbb{E}\left(\sum_i \mathbb{I}_{\tau=i} \exp(tS_n(s_i)) \sum_{k=1}^n \mathbb{E}(s_i^k(X_k))^2\right).$$

Since $\sum_k \mathbb{E}(s_i^k(X_k))^2 \leq V_n$, we infer that

(3.13) $\qquad\qquad\qquad \sum_{k=1}^n \mathbb{E}(h_k r(t, g_k/h_k)) \leq \frac{1}{2} t^2 e^{2t} V_n F(t).$

Together with Lemma 3.2 and (3.8), (3.13) leads to the differential inequality :

(3.14) $\qquad\qquad\qquad tL' - (te^{2t} + 1)L \leq t^2 e^{2t}(V_n/2).$

8

Let $\gamma(t) = t^{-2} \exp((1 - e^{2t})/2)$. Multiplying (3.14) by γ, we get

(3.15)
$$(t\gamma L)' \le (V_n/2)e^{2t} \exp((1 - e^{2t})/2).$$

Since $t\gamma(t) \sim (1/t)$ as t tends to 0, integrating (3.15) gives

$$t\gamma(t)L(t) \le \mathbb{E}(Z) + (V_n/2)(1 - \exp((1 - e^{2t})/2)),$$

which implies Theorem 1.1(a). ∎

To prove Theorem 1.1(b) we apply both Markov's inequality to the random variable $\exp(tZ)$ and Theorem 1.1(a) with $t = \frac{1}{2}\log(1 + 2\log(1 + x/v))$.

To prove Theorem 1.1(c), we bound up the log-Laplace transform of $Z - \mathbb{E}(Z)$ via Lemma 3.4 below and next we apply Markov's exponential inequality.

Lemma 3.4. *Under the assumptions of Theorem 1.1, for any t in $]0, 2/3[$,*

$$L(t) \le t\mathbb{E}(Z) + (2\mathbb{E}(Z) + V_n)\frac{t^2}{2 - 3t}.$$

Proof. From Theorem 1.1(a), it is enough to prove that

$$\exp((e^{2t} - 1)/2) \le 1 + 2t/(2 - 3t).$$

This inequality holds if and only if

$$\lambda(t) := \log(2 - t) - \log(2 - 3t) - (e^{2t} - 1)/2 \ge 0.$$

Expanding λ in power series yields $\lambda(t) = \sum_{j \ge 2} b_j t^j / j!$, where

$$b_j = (j - 1)!\left((3/2)^j - (1/2)^j\right) - 2^{j-1} \ge 2(j - 1)! - 2^{j-1} \ge 0.$$

Hence $\lambda(t) \ge 0$, which implies Lemma 3.4. ∎

Theorem 1.1(c) follows from Lemma 3.4 by noting that the Legendre transform of the function $t \to t^2/(2 - 3t)$ (here $t < 2/3$) is equal to $\frac{4}{9}(1 + (3x/2) - \sqrt{1 + 3x})$.

4. Compensated empirical processes.

In this section we prove Theorem 1.2. We start by proving Theorem 1.2(a). Throughout the section, t is any positive real. For i in $\{1, \ldots, m\}$, let

$$L_i(t) = \log \mathbb{E}\left(\exp(-tS_n(s_i))\right).$$

9

Let us define the exponentially compensated empirical process $T(s_i, t)$ by

(4.1) $$T(s_i, t) = S_n(s_i) + t^{-1} L_i(t)$$

We set

(4.2) $$Z_t = \sup_{1 \leq i \leq m} T(s_i, t) \text{ and } f_t = \exp(-t Z_t).$$

Let

(4.3) $$F(t) = \mathbb{E}(f_t) = \mathbb{E}(\exp(-t Z_t)) \text{ and } \Lambda(t) = \log F(t).$$

Our purpose is to obtain a differential inequality for Λ via the log-Sobolev method.

Before that, we link the log-Laplace L_{-Z} of $-Z$ with Λ.

Lemma 4.1. *For any positive t,*

$$L_{-Z}(t) - \sup_i L_i(t) \leq \Lambda(t) \leq \min(L_{-Z}(t), 0).$$

Proof. By definition of Z_t,

$$\exp(-t Z_t) = \exp\left(\inf_i (-t S_n(s_i) - L_i(t))\right) \geq \exp(-tZ - \sup_i L_i),$$

Consequently, for any positive t,

$$\exp(\Lambda(t)) \geq \exp(-\sup_i L_i(t)) \mathbb{E} \exp(-tZ),$$

which gives the first inequality. Next, by definition of Z_t,

$$\exp(\Lambda(t)) = \mathbb{E}\left(\inf_i \exp(-t S_n(s_i) - L_i(t))\right) \leq \mathbb{E}\left(\exp(-t S_n(s_1) - L_1(t))\right) = 1,$$

which ensures that $\Lambda(t) \leq 0$. Moreover $L_i(t) \geq 0$ by the centering assumption on the random variables $S_n(s)$. Hence

$$\exp(\Lambda(t)) \leq \mathbb{E}\left(\inf_i \exp(-t S_n(s_i))\right) = \mathbb{E}(\exp(-tZ)),$$

which completes the proof of Lemma 4.1. ∎

Definition 4.1. Let τ_t denote the first integer i such that $Z_t = T(s_i, t)$, where Z_t is defined in (4.2).

Since the random functions $T(s_i, t)$ are analytic functions of t, the random function f_t defined in (4.2) is continuous and piecewise analytic, with derivative with respect to t, almost everywhere (a.e.)

$$(4.4) \qquad f_t' = -Z_t f_t - (L_{\tau_t}'(t) - t^{-1} L_{\tau_t}(t)) f_t = -(Z_t + t Z_t') f_t,$$

where $t Z_t' = L_{\tau_t}'(t) - t^{-1} L_{\tau_t}(t)$ by convention. Consequently the Fubini theorem applies and

$$(4.5) \qquad F(t) = 1 - \int_0^t \mathbb{E}((Z_u + u Z_u') f_u) du.$$

Therefrom the function F is absolutely continuous with respect to the Lebesgue measure, with a.e. derivative in the sense of Lebesgue

$$(4.6) \qquad F'(t) = -\mathbb{E}((Z_t + t Z_t') f_t).$$

Moreover, from the elementary lower bound $f_t \geq \exp(-2nt)$, the function $\Lambda = \log F$ is absolutely continuous with respect to the Lebesgue measure, with a.e. derivative F'/F if F' is the above defined function.

Definition 4.2. Let $f^k = E_n^k f_t$.

We now apply Proposition 2.1 to the random function f_t. Clearly

$$(4.7) \qquad \mathbb{E}(f_t \log f_t) - \mathbb{E}(f_t) \log \mathbb{E}(f_t) = \mathbb{E}(t^2 Z_t' f_t) + t F'(t) - F(t) \log F(t) \quad \text{a.e.}$$

Hence, applying Proposition 2.1 with $f = f_t$

$$(4.8) \quad t F' - F \log F \leq -\mathbb{E}(t^2 Z_t' f) + \sum_{k=1}^n \mathbb{E}(g_k \log(g_k/E_n^k g_k)) + \sum_{k=1}^n \mathbb{E}((g_k - f)) \log(f^k/f)).$$

Now choose

$$(4.9) \qquad g_k = \sum_i P_n^k(\tau_t = i) \exp(-t S_n(s_i) - L_i(t)).$$

By definition of Z_t,

$$\exp(-t S_n(s_i) - L_i(t)) \geq \exp(-t Z_t),$$

which implies that $g_k \geq f$. Therefore the upper bound on the second term in (4.8) will be derived from Lemma 4.2 below.

Notation 4.1. For sake of brievity, throughout the sequel we note $\tau = \tau_t$.

11

Lemma 4.2. *Let $\psi(t) = (\exp(2t) + 1)/2$. Set $l_{ki}(t) = \log \mathbb{E}(\exp(-ts_i^k(X_k)))$. Then a.s.*

$$(f^k/f) \leq \exp(ts_\tau^k(X_k) + l_{k\tau}(t)) \leq \psi(t).$$

Proof. Let $S_n^k(s) = S_n(s) - s^k(X_k)$. Set

$$Z^k = \sup\{S_n^k(s) + t^{-1} \log \mathbb{E}(\exp(-tS_n^k(s))) : s \in \mathcal{S}\}.$$

Let τ_k be the first integer in $[1, m]$ such that

$$S_n^k(s_{\tau_k}) + t^{-1} \log \mathbb{E}(\exp(-tS_n^k(s_{\tau_k}))) = Z^k$$

Clearly

$$f_t \leq \exp(-tZ^k) \exp(-ts_{\tau_k}^k(X_k)) - l_{k\tau_k}(t)).$$

Since the stopping time τ_k is $\mathcal{F}_n^k$-measurable, it follows that

(4.10) $$E_n^k f_t \leq \exp(-tZ^k).$$

Now, by definition of Z^k,

(4.11) $$\exp(-tZ^k) \leq \exp(-tZ + ts_\tau^k(X_k) + l_{k\tau}(t)),$$

which ensures that $(f^k/f) \leq \exp(ts_\tau^k(X_k) + l_{k\tau}(t))$. To conclude the proof of Lemma 4.2, recall that $\mathbb{E}(\exp(tX)) \leq \cosh(t)$, for any centered random variable X with values in $[-1, 1]$, which implies the second part of Lemma 4.2. ∎

The next step to bound up the second term on right hand is Lemma 4.3 below. However, due to technical difficulties, we are able to bound up this term only on some finite interval.

Lemma 4.3. *Let (g_k) be the finite sequence of random variables defined in (4.9). Set $\varphi = \psi \log \psi$. Let t_0 be the positive solution of the equation $\varphi(t) = 1$. Then, for any t in $[0, t_0[$,*

$$\sum_{k=1}^n \mathbb{E}\left((g_k - f) \log(f^k/f)\right) \leq \frac{\varphi(t)}{1 - \varphi(t)} \left(\sum_{k=1}^n \mathbb{E}(g_k \log(g_k/E_n^k g_k)) - \mathbb{E}(f \log f)\right).$$

Proof. Since the random variables $S_n^k(s)$ are $\mathcal{F}_n^k$ measurable,

(4.12) $$E_n^k(g_k) = \sum_{i=1}^m P_n^k(\tau = i) \exp(-tS_n^k(s_i) - L_i(t) + l_{ki}(t)).$$

12

It follows that $E_n^k(g_k) = E_n^k(f \exp(ts_\tau^k(X_k) + l_{k\tau}(t)))$. Hence

$$
(4.13) \qquad \sum_{k=1}^n \mathbb{E}(g_k - f) = \sum_{k=1}^n \mathbb{E}\Big(f(\exp(ts_\tau^k(X_k) + l_{k\tau}(t)) - 1)\Big).
$$

Setting $\eta_k = ts_\tau^k(X_k) + l_{k\tau}(t)$, we have

$$
\sum_{k=1}^n \mathbb{E}(g_k - f) = \sum_{k=1}^n \mathbb{E}\Big(f(e^{\eta_k} - 1 - \psi(t)\eta_k)\Big) + \psi(t)\mathbb{E}\Big(f\sum_{k=1}^n \eta_k\Big)
$$

$$
= \sum_{k=1}^n \mathbb{E}\Big(f(e^{\eta_k} - 1 - \psi(t)\eta_k)\Big) - \psi(t)\mathbb{E}(f\log f),
$$

since $\sum_{k=1}^n \eta_k = -\log f$. Now, for x in $]-\infty, \log \psi(t)]$, the function $x \to e^x - 1 - x\psi(t)$ is nonincreasing. Since $\log \psi(t) \ge \eta_k \ge \log(f^k/f)$ by Lemma 4.2, we infer that

$$
\sum_{k=1}^n \mathbb{E}(g_k - f) \le \sum_{k=1}^n \mathbb{E}\Big(f\big((f^k/f) - 1 - \psi(t)\log(f^k/f)\big)\Big) - \psi(t)\mathbb{E}(f\log f)
$$

$$
\le \psi(t)\left(\sum_{k=1}^n \mathbb{E}(f\log(f/f^k)) - \mathbb{E}(f\log f)\right).
$$

Hence, applying (2.3), we obtain
$$
(4.14)
$$
$$
\sum_{k=1}^n \mathbb{E}(g_k - f) \le \psi(t)\left(\sum_{k=1}^n \mathbb{E}\Big(g_k \log(g_k/E_n^k g_k) + (g_k - f)\log(f^k/f)\Big) - \mathbb{E}(f\log f)\right).
$$

Now, from Lemma 4.2 we know that $\log(f^k/f) \le \log \psi(t)$. Since $(g_k - f) \ge 0$, it follows that

$$
\sum_{k=1}^n \mathbb{E}\Big((g_k - f)\log\Big(\frac{f^k}{f}\Big)\Big) \le \log \psi(t) \sum_{k=1}^n \mathbb{E}(g_k - f)
$$

$$
(4.15) \qquad \le \varphi(t)\left(\sum_{k=1}^n \mathbb{E}\Big(g_k \log\Big(\frac{g_k}{E_n^k g_k}\Big) + (g_k - f)\log\Big(\frac{f^k}{f}\Big)\Big) - \mathbb{E}(f\log f)\right).
$$

Since $1 - \varphi(t) > 0$ for any t in $[0, t_0[$, inequality (4.15) then implies Lemma 4.3. ∎

From Lemma 4.3 and the differential inequality (4.8) we then get that

$$
(1 - \varphi)(tF' - F\log F) \le \varphi \mathbb{E}(t^2 Z_t' f - f\log f) - \mathbb{E}(t^2 Z_t' f) + \sum_{k=1}^n \mathbb{E}(g_k \log(g_k/E_n^k g_k)).
$$

13

Now from (4.7) we know that $\mathbb{E}(t^2 Z_t' f - f \log f) = -tF'$, whence

$$(4.16) \qquad tF' - (1 - \varphi)F \log F \leq -\mathbb{E}(t^2 Z_t' f) + \sum_{k=1}^{n} \mathbb{E}(g_k \log(g_k/E_n^k g_k)).$$

Let us now bound up the first term on right hand in (4.16). Set $w_k = (g_k/E_n^k g_k)$. Then

$$E_n^k(g_k \log(g_k/E_n^k g_k)) = E_n^k(g_k)E_n^k(w_k \log w_k).$$

From (4.12), by convexity of the function $x \log x$,

$$E_n^k(g_k)w_k \log w_k \leq \sum_i P_n^k(\tau = i)(-ts_i^k(X_k) - l_{ki}(t)) \exp(-tS_n(s_i) - L_i(t)).$$

Consequently

$$E_n^k(g_k \log(g_k/E_n^k g_k)) \leq \sum_i P_n^k(\tau = i) \exp(-tS_n^k(s_i) - L_i(t) + l_{ki}(t))(tl_{ki}'(t) - l_{ki}(t)).$$

Since

$$\sum_i P_n^k(\tau = i) \exp(-tS_n^k(s_i) - L_i + l_{ki})(tl_{ki}' - l_{ki}) =$$

$$E_n^k\left(\sum_i \mathbb{1}_{\tau = i} \exp(-tS_n^k(s_i) - L_i + l_{ki})(tl_{ki}' - l_{ki})\right),$$

it implies that

$$\mathbb{E}(g_k \log(g_k/E_n^k g_k)) \leq \mathbb{E}\left(\exp(-tZ_t + ts_\tau^k(X_k) + l_{k\tau})(tl_{k\tau}' - l_{k\tau})\right),$$

From the convexity of the functions l_{ki}, we know that $tl_{k\tau}' - l_{k\tau} \geq 0$. Hence, applying Lemma 4.2, we get

$$\mathbb{E}(g_k \log(g_k/E_n^k g_k))) \leq \psi(t)\mathbb{E}((tl_{k\tau}' - l_{k\tau})f).$$

Since $t^2 Z_t' = tL_\tau' - L_\tau$, it follows that

$$(4.17) \qquad -\mathbb{E}(t^2 Z_t' f) + \sum_{k=1}^{n} \mathbb{E}(g_k \log(g_k/E_n^k g_k)) \leq (\psi(t) - 1)\mathbb{E}((tL_\tau' - L_\tau)f).$$

Both (4.16) and (4.17) yield, for t in $[0, t_0[$,

$$(4.18) \qquad tF' - (1 - \varphi)F \log F \leq (\psi(t) - 1)\mathbb{E}((tL_\tau' - L_\tau)f).$$

14

Since $tL'_\tau - L_\tau \leq \sup_i(tL'_i - L_i)$, dividing by F, we infer that

$$(4.19) \qquad t\Lambda' - (1 - \varphi)\Lambda \leq (\psi(t) - 1) \sup_i(tL'_i - L_i).$$

Next we derive an upper bound on $tL'_i - L_i$ from the lemma below.

Lemma 4.4. *Let Y be a random variable with values in $]-\infty, 1]$, such that $\mathbb{E}(Y^2) < +\infty$. Then, for any positive t,*

$$\mathbb{E}(tYe^{tY}) - \mathbb{E}(e^{tY}) \log \mathbb{E}(e^{tY}) \leq \mathbb{E}(Y^2)(1 + (t - 1)e^t).$$

Proof. From the variational definition of the entropy, we know that, for any positive constant c and any positive random variable T,

$$\mathbb{E}(T \log T) - \mathbb{E}(T) \log \mathbb{E}(T) \leq \mathbb{E}(T \log(T/c) - T + c).$$

Taking $c = 1$ and $T = \exp(tY)$, we then get that

$$\mathbb{E}(tYe^{tY}) - \mathbb{E}(e^{tY}) \log \mathbb{E}(e^{tY}) \leq \mathbb{E}((tY - 1)e^{tY} + 1).$$

Now, from De l'Hospital rule for monotonicity the function $x \to x^{-2}(1 + (x - 1)e^x)$ is nondecreasing on the real line. Hence, for any positive t,

$$(tY - 1)e^{tY} + 1 \leq Y^2(1 + (t - 1)e^t),$$

which implies Lemma 4.4. ∎

Let $Y_k = s^k_i(X_k)$. From the centering assumption, $\mathbb{E} \exp(tY_k) \geq 1$. Hence we have:

$$tl'_{ki}(t) - l_{ki}(t) \leq \mathbb{E}(tY_ke^{tY_k}) - \mathbb{E}(e^{tY_k}) \log \mathbb{E}(e^{tY_k}) \leq (1 + (t - 1)e^t) \operatorname{Var} Y_k$$

by Lemma 4.4. Since $tL_i - l'_i = \sum_k(tl'_{ki} - l_{ki})$, it ensures that

$$(4.20) \qquad tL'_i - L_i \leq V_n(1 + (t - 1)e^t).$$

Both the above bound and (4.19) lead to the differential inequality below.

Proposition 4.1. *For any t in $[0, t_0[$,*

$$t\Lambda' - (1 - \varphi)\Lambda \leq \tfrac{1}{2}V_n(e^{2t} - 1)(1 + (t - 1)e^t).$$

It remains to bound up Λ. Set $\widetilde{\Lambda}(t) = t^{-1}\Lambda(t)$ and

$$(4.22) \qquad I(t) = \int_0^t \frac{\varphi(u)}{u} du.$$

15

Then $\left(\widetilde{\Lambda}e^I\right)' = t^{-2}(t\Lambda' - (1-\varphi)\Lambda)e^I$. Consequently, from Proposition 4.1,

$$(4.23) \qquad \left(\widetilde{\Lambda}e^I\right)' \le \frac{V_n}{2t^2}(e^{2t}-1)(1+(t-1)e^t)e^I.$$

Since $\widetilde{\Lambda}e^I$ is absolutely continuous with respect to the Lebesgue measure, integrating (4.23) yields

$$(4.24) \qquad \widetilde{\Lambda}(t) \le \widetilde{\Lambda}(\varepsilon)e^{I(\varepsilon)-I(t)} + \frac{V_n}{2}\int_\varepsilon^t u^{-2}(e^{2u}-1)(1+(u-1)e^u)e^{I(u)-I(t)}\,du$$

for $0 < \varepsilon < t$. The control of the integral on right hand side will be done via the bounds for φ below, whose proof is carried out in section 5.

Lemma 4.5. *For any t in $[0,t_0]$, $t \le \varphi(t) \le t\exp(2t) - (t^2/2)$.*

By Lemma 4.5, $\lim_0 I(\varepsilon) = 0$. Furthermore $\widetilde{\Lambda}(\varepsilon) \le \varepsilon^{-1}L_{-Z}(\varepsilon)$ by Lemma 4.1. Therefore

$$(4.25) \qquad \limsup_{\varepsilon \to 0} \widetilde{\Lambda}(\varepsilon)e^{I(\varepsilon)-I(t)} \le -\mathbb{E}(Z)e^{-I(t)}.$$

Now $I(u) - I(t) \le (u-t)$ by Lemma 4.5. Consequently, letting $\varepsilon \to 0$ in (4.24) and applying (4.25), we get

$$(4.26) \qquad \Lambda(t) \le -\mathbb{E}(Z)te^{-I(t)} + \tfrac{1}{2}V_n te^{-t}\int_0^t u^{-2}(e^{2u}-1)(1+(u-1)e^u)e^u\,du.$$

To bound up L_{-Z}, we then apply the Bennett bound $L_i(t) \le V_n(e^t - t - 1)$ together with Lemma 4.1. This yields the proposition below.

Proposition 4.2. *Let the function J be defined by*

$$J(t) = \frac{1}{2}\int_0^t u^{-2}(e^{2u}-1)(1+(u-1)e^u)e^u\,du$$

and I be the function defined in (4.22). For any t in $[0,t_0]$,

$$L_{-Z}(t) + t\mathbb{E}(Z) \le t\mathbb{E}(Z)(1-e^{-I(t)}) + V_n(te^{-t}J(t) + e^t - t - 1).$$

To obtain Theorem 1.2(a) for t in $[0,t_0]$, we bound up the functions appearing in Proposition 4.2 via Lemma 4.6 below, proved in section 5.

Lemma 4.6. *For any t in $[0,t_0]$,*

$$(a) \qquad te^{-t}J(t) + e^t - t - 1 \le \tfrac{1}{9}(\exp(3t) - 3t - 1),$$
$$(b) \qquad t(1 - e^{-I(t)}) \le \tfrac{2}{9}(\exp(3t) - 3t - 1).$$

Next, proceeding as in Klein (2002), we prove (a) for t in $[t_0, +\infty[$. For sake of brievity, set $E = \mathbb{E}(Z)$. By Lemma 4.1, for any positive t,

$$L_{-Z}(t) + tE \leq tE + \sup_i L_i(t) \leq tE + V_n(e^t - t - 1) \leq v\max(t/2, e^t - t - 1).$$

Now, let t_1 be the unique positive solution of the equation $e^t - t - 1 = t/2$. t_1 belongs to $[0.76, 0.77]$, whence $t_1 > t_0$ (note that $t_0 \in [0.46, 0.47]$). If $t \geq t_1$ then $t/2 \leq e^t - t - 1$. In that case

$$L_{-Z}(t) + tE \leq v(e^t - t - 1) \leq (v/9)(e^{3t} - 3t - 1),$$

which proves (a) for $t \geq t_1$.

If t belongs to $[t_0, t_1]$, from the convexity of L_{-Z} we have

$$L_{-Z}(t) + tE \leq \tfrac{1}{9}v(e^{3t_0} - 3t_0 - 1)\frac{t_1 - t}{t_1 - t_0} + \tfrac{1}{2}vt_1\frac{t - t_0}{t_1 - t_0} \leq \tfrac{1}{9}v(e^{3t} - 3t - 1),$$

which completes the proof of (a) ■

To prove Theorem 1.2(c), we note that

(4.27) $$\tfrac{1}{9}\left(e^{3t} - 3t - 1\right) \leq \frac{t^2}{2 - 2t}$$

(cf. Rio (2000), p. 152). Theorem 1.2(c) follows from (4.27) via the usual Cramér-Chernoff calculation.

5. Technical tools.

In this section, we prove Lemmas 4.5 and 4.6.

Proof of Lemma 4.5. By definition of ψ and φ,

$$\varphi(t) = t\psi(t) + \psi(t)\log\cosh(t) \geq t,$$

since $\psi(t) \geq 1$ for any nonnegative t. Next

$$t\exp(2t) - (t^2/2) - \varphi(t) = \psi(t)(t\tanh(t) - \log\cosh(t) - (e^{2t} + 1)^{-1}t^2),$$

so that Lemma 4.5 holds if $p(t) := t\tanh(t) - \log\cosh(t) - t^2/(e^{2t} + 1) \geq 0$ for t in $[0, t_0]$. Now $p(0) = 0$ and

$$2\cosh^2(t)p'(t) = 2t - t(e^{-2t} + 1) - t^2 = t(1 - t - e^{-2t}).$$

Since $\exp(-2t) \leq 1 - t$ for t in $[0, 1/2]$, the above identity ensures that $p'(t) \geq 0$ on $[0, t_0]$ (recall that $t_0 < 1/2$), which implies Lemma 4.5.

17

Proof of Lemma 4.6. We start by proving (a). Clearly (a) holds if

(5.1) $\alpha(t) = \frac{1}{9}e^t(e^{3t} - 3t - 1) + e^t(1 + t - e^t) - J(t) \geq 0$

for any t in $[0, 4]$. (with the convention $\alpha(0) = 0$). The function α is analytic on the real line. To prove (5.1), we then note that $\alpha^{(i)}(0) = 0$ for $i = 1, 2$. Consequently (a) holds if, for t in $[0, 4]$,

(5.2) $\alpha^{(3)}(t) = e^{3t}(-t + (19/3)) - 4e^{2t} + e^t(-3t - 5) + (8/3) > 0.$

Now (5.2) holds if $\alpha^{(4)}(t) > 0$, since $\alpha^{(3)}(0) > 0$. Next

$$\beta(t) := e^{-t}\alpha^{(4)}(t) = 3e^{2t}(-2t + 11) - 8e^t - 3$$

satisfies $\beta(0) > 0$ and, for t in $[0, 4]$,

$$\beta'(t) = 12e^{2t}(5 - t) - 8e^t > e^t(12e^t - 8) > 0,$$

which ensures that $\beta(t) > 0$ for t in $[0, 4]$. Hence Lemma 4.6(a) holds.

To prove (b), we apply Lemma 4.5 to bound up the function $I(t)$. This gives

$$I(t) \leq \int_0^t (e^{2u} - u/2)du = \frac{e^{2t} - 1}{2} - \frac{t^2}{4}.$$

Now, recall $t_0 \leq 1/2$. For t in $[0, 1/2]$, expanding $\exp(2t)$ in entire series yields

$$(\exp(2t) - 1)/2 = t + t^2 + 4t^3 \sum_{k \geq 3} \frac{1}{k!}(2t)^{k-3} \leq t + t^2 + 4t^3 \sum_{k \geq 3} \frac{1}{k!}.$$

Hence, for $t \leq 1/2$,

(5.3) $I(t) \leq t + \frac{3}{4}t^2 + (4e - 10)t^3 \leq t + \frac{3}{4}t^2 + \frac{7}{8}t^3 =: \gamma(t).$

From (5.3), Lemma 4.6(b) holds if

$$d(t) = \frac{2}{9}(e^{3t} - 3t - 1) - t + t\exp(-\gamma(t)) \geq 0.$$

Now $d(0) = d'(0) = 0$ and

$$d''(t) = 2e^{-\gamma(t)}\left(e^{4t + (3/4)t^2 + (7/8)t^3} - 1 - \frac{7}{4}t - \frac{15}{4}t^2 + \frac{15}{4}t^3 + \frac{63}{16}t^4 + \frac{441}{128}t^5\right).$$

Since

$$e^{4t + (3/4)t^2 + (7/8)t^3} \geq e^{4t} \geq 1 + 4t + 8t^2,$$

18

we have $d''(t) > 0$ for any positive t. Consequently $d(t) \geq 0$, which implies Lemma 4.6(b).

References

Bobkov, S.: Some extremal properties of the Bernoulli distribution. Theory probab. appl. **41**, 748-755 (1996)

Bousquet, O.: Concentration inequalities for sub-additive functions using the entropy method. To appear in the proceedings of the conference on stochastic inequalities 2002. Eds Giné, Houdré and Nualart.

Klein, T.: Une inégalité de concentration à gauche pour les processus empiriques. C.R. Acad. Sci. Paris, Série I, **334**, 495-500 (2002).

Ledoux, M.: On Talagrand's deviation inequalities for product measures. European series in applied and industrial mathematics, probability and statistics **1**, 63-87 (1996).

Ledoux, M., Talagrand, M.: Probability in Banach spaces. Isoperimetry and Processes. Berlin: Springer (1991).

Massart, P.: About the constants in Talagrand's concentration inequalities for empirical processes. Annals probab. **28**, 863-884, (2000).

Panchenko, D.: A note on Talagrand's concentration inequality. Electron. commun. probab. **6**, 55-65 (2001).

Panchenko, D.: Symmetrization approach to concentration inequalities for empirical processes. Annals probab. **31**, 2068-2081 (2003).

Rio, E.: Théorie asymptotique des processus aléatoires faiblement dépendants. Ed. J.M. Ghidaglia et X. Guyon. Mathématiques et Applications 31. Springer. (2000).

Rio, E.: Inégalités de concentration pour les processus empiriques de classes de parties. Probab. theory relat. fields **119**, 163-175, (2001).

Rio, E.: Une inégalité de Bennett pour les maxima de processus empiriques. Annales de l'institut Henri Poincaré, probabilités et statistiques **38**, 1053-1057 (2002).

Talagrand, M.: New concentration inequalities in product spaces. Invent. math. **126**, 503-563 (1996).