Left deviation inequalities for suprema of empirical processes

Antoine Marchina*

September 17, 2018

Abstract

In this paper, we provide left deviation inequalities for suprema of unbounded empirical processes associated with independent and identically distributed random variables by means of martingale methods. This work complete the paper [10] in which the deviation on the right-hand side of the mean is studied.

1 Introduction and notations

In the recent paper [10], we prove new upper bounds on the deviation above the mean for suprema of empirical processes indexed by classes of unbounded functions. More specifically, let $X_1, \ldots, X_n$ be a finite sequence of independent random variables and identically distributed according to a law P, valued in some measurable space $(\mathcal{X}, \mathcal{F})$ and let $\mathscr{F}$ be a countable class of measurable functions from $\mathcal{X}$ to $\mathbb{R}$. We assume that for all $f \in \mathscr{F}$, $P(f) = 0$ and that $\mathscr{F}$ has a square integrable envelope function Φ, that is

$$|f| \leq \Phi \quad \text{for any } f \in \mathscr{F}, \text{ and } \quad \Phi \in \mathbb{L}^2. \tag{1.1}$$

We set

$$Z := \sup_{f \in \mathscr{F}} \sum_{k=1}^{n} f(X_k). \tag{1.2}$$

If $\Phi(X_1)$ admits a finite weak moment of order $\ell > 2$ (denoted by $\Lambda_\ell^+(\Phi(X_1))$), we establish the following Fuk-Nagaev type upper bound on the Conditionnal Value at Risk (CVaR for short, see Definition 1.1 below) of $Z - \mathbb{E}[Z]$:

$$\tilde{Q}_{Z-\mathbb{E}[Z]}(u) \leq \sqrt{2\log(1/u)}\left(\sigma\sqrt{n} + \sqrt{V_n}\right) + 3\, n^{1/\ell} \mu_\ell \, \Lambda_\ell^+(\Phi(X_1)) u^{-1/\ell}, \tag{1.3}$$

*Université Paris-Est, LAMA (UMR 8050), UPEM, CNRS, UPEC. E-mail: antoine.marchina@u-pem.fr

for any $u \in]0\,,1[$. Here $\sigma^2 := \sup_{f\in\mathscr{F}} P(f^2)$, $\mu_\ell := 2+\max(4/3, \ell/3)$ and V_n is an explicit corrective term. We recall that an upper bound on the CVaR leads to a deviation inequality (see Proposition 1.4 further). We stress out that there are few results in the literature concerning concentration inequalities for suprema of unbounded empirical processes whose main ones can be found in Boucheron, Bousquet, Lugosi and Massart [5], Adamczak [1] and Lederer and van de Geer [16, 9]. Their results provide a less accurate estimate compared to (1.3). In these papers, the deviation is around $(1+\eta)\mathbb{E}[Z]$ with $\eta > 0$ and/or the constants are nonexplicit or nonoptimal.

The purpose of this paper is to complete the analysis started in [10], by providing upper bounds for the left-hand side deviations, also in unbounded settings. We consider four different cases:

- $\sup_{f\in\mathscr{F}}(-f)_+(X_1)$ admits a weak moment of order $\ell > 2$.
- For all $f \in \mathscr{F}$, $f(X_1)$ admits a sub-Gamma tail on the left
- For all $f \in \mathscr{F}$, $f(X_1)$ admits a sub-Gaussian tail on the left.
- Suprema of randomized empirical processes.

It is worth noticing that the left-hand side deviations around the mean for Z heavily depend on the behavior of the left tails (under the distribution P) of the functions in the class $\mathscr{F}$. The behavior of right tails only takes part in a corrective term in the subgaussian coefficient in our inequalities. The main results, under the assumptions mentionned above, are stated respectively in Sections 2, 3, 4 and 5. All the proofs are postponed to Section 6.

In the rest of this section, we give notations and definitions which we will use all along the paper. Let us start with the classical notations $x_+ := \max(0, x)$ and $x_+^\alpha := (x_+)^\alpha$ for all real x and α. Next, we define the quantile function and the Conditional Value-at-Risk.

Definition 1.1. *Let X be a real-valued integrable random variable. Let the function Q_X be the càdlàg inverse of $x \mapsto \mathbb{P}(X > x)$. The Conditional Value-at-Risk is defined by*

$$\tilde{Q}_X(u) := u^{-1}\int_0^u Q_X(s)ds \quad \textit{for any } u \in]0\,,1]. \tag{1.4}$$

Let us now recall the definition of the Legendre transform of a convex function.

Definition 1.2. *Let* $\phi : [0, \infty[\to [0, \infty]$ *be a convex, nondecreasing and càdlàg function such that* $\phi(0) = 0$*. The Legendre transform* ϕ^* *of the function* ϕ *is defined by*

$$\phi^*(\lambda) := \sup\{\lambda t - \phi(t) : t > 0\} \quad \textit{for any } \lambda \geq 0. \tag{1.5}$$

The inverse function of ϕ^* admits the following variational expression (see, for instance, Rio [14, Lemma A.2]).

$$\phi^{*-1}(x) = \inf\{t^{-1}(\phi(t) + x) : t > 0\} \quad \text{for any } x \geq 0. \tag{1.6}$$

A particular function ϕ satisfying conditions of Definition 1.2 is the log-Laplace transform of a random variable:

Notation 1.3. *Let* X *be a real-valued integrable random variable with a finite Laplace transform on right neighborhood of* 0*. The log-Laplace transform of* X*, denoted by* ℓ_X*, is defined by*

$$\ell_X(t) := \log \mathbb{E}[\exp(tX)] \quad \textit{for any } t \geq 0. \tag{1.7}$$

The function Q_X and the CVaR satisfy the following elementary properties, which are given and proved in Pinelis [12, Theorem 3.4].

Proposition 1.4. *Let* X *and* Y *be real-valued and integrable random variables. Then, for any* $u \in]0, 1]$,

(i) $\mathbb{P}(X > Q_X(u)) \leq u$,

(ii) $Q_X(u) \leq \tilde{Q}_X(u)$,

(iii) $\tilde{Q}_{X+Y}(u) \leq \tilde{Q}_X(u) + \tilde{Q}_Y(u)$.

(iv) Assume that X *has a finite Laplace transform on a right neighborhood of* 0*. Then* $\tilde{Q}_X(u) \leq \ell_X^{*-1}(\log(1/u))$.

Remark 1.5. *Since we use different notations from those of Pinelis, let us mention that his notations* $Q_0(X; u)$*,* $Q_1(X; u)$ *and* $Q_\infty(X; u)$ *correspond respectively to* $Q_X(u)$*,* $\tilde{Q}_X(u)$ *and* $\ell_X^{*-1}(\log(1/u))$.

Let us now define the following class of distribution functions.

Notation 1.6. *Let* $q \in [0, 1]$*. Let* ψ *be a nonnegative random variable. We denote by* F_ψ *the distribution function of* ψ *and by* F_ψ^{-1} *the càdlàg inverse of* F_ψ*. Set* $b_{\psi,q} := F_\psi^{-1}(1 - q)$*. We denote by* $F_{\psi,q}$ *the distribution function defined by*

$$F_{\psi,q}(x) := (1 - q)\mathbb{1}_{0 \leq x < b_{\psi,q}} + F_\psi(x)\mathbb{1}_{x \geq b_{\psi,q}}. \tag{1.8}$$

Set for any $k = 1, \dots, n$,

$$E_k := \mathbb{E}\Big[\sup_{f\in\mathscr{F}} P_k(f)\Big], \tag{1.9}$$

and let ζ_k denote a random variable with distribution function $F_{2\phi(X_1),q_k}$, where q_k is the real in $[0\,,1]$ such that $\mathbb{E}[\zeta_k] = E_k$. Let us now define the quantity which appears in all of our inequalities in the corrective term in the subgaussian coefficient:

$$V_n := \sum_{k=1}^{n} \mathbb{E}[\zeta_k^2]. \tag{1.10}$$

The justification of the term "corrective" is given by this remark:

Remark 1.7. *If the class $\mathscr{F}$ satisfies the uniform law of large numbers, that is $\sup_{f\in\mathscr{F}} |P_n(f)|$ converges to 0 in probability, then E_n decreases to 0 (see, for instance, Section 2.4 of van der Vaart and Wellner [17]). Now, from the integrability of Φ^2 and (1.8), the convergence of E_n to 0 implies the convergence of $\mathbb{E}[\zeta_n^2]$ to 0, which ensures that $V_n = o(n)$ as n tends to infinity.*

2 Fuk-Nagaev type inequalities

In this Section we provide Fuk-Nagaev type inequalities which allow us to derive strong and weak moment inequalities. These left deviation inequalities are the exact counterparts of the right deviation inequalities stated in [10]. We recall that $\Phi := \sup_{f\in\mathscr{F}} |f|$ and we define also $\Phi^- := \sup_{f\in\mathscr{F}} (-f)_+$.

2.1 General inequalities

Theorem 2.1. *Let $x > 0$. For any $s > 0$, we have*

$$\mathbb{P}(Z \le \mathbb{E}[Z] - x) \le \Big(1 + \frac{x^2}{4sn\sigma^2}\Big)^{-s/2} + \exp\Big(-\frac{x^2}{8V_n}\Big) + n\mathbb{P}\Big(\Phi^-(X_1) \ge \frac{x}{2s}\Big),$$

where V_n is defined by (1.10).

Consequently, if $\Phi^-(X_1)$ admits an ℓ-th moment, we derive then by integrating the above inequality the following moment inequality:

Corollary 2.2. *Let $\ell \ge 2$. Assume that $\Phi^-(X_1)$ is $\mathbb{L}^\ell$-integrable. Then*

$$\|(\mathbb{E}[Z] - Z)_+\|_\ell \le 2\beta_\ell\, \ell^{1/\ell}\sqrt{\ell+1}\Big(\sigma\sqrt{n} + \delta_\ell\sqrt{V_n}\Big) + 2\, n^{1/\ell}(\ell+1)\|\Phi^-(X_1)\|_\ell,$$

where $\beta_\ell := \Big((\sqrt{\pi}/2)\,\Gamma(\ell/2)/\Gamma((\ell+1)/2)\Big)^{1/\ell}$, $\delta_\ell := \sqrt{2}/\alpha_\ell$,

and $\alpha_\ell := \Big(\sqrt{\pi}/\Gamma((\ell+1)/2)\Big)^{1/\ell}\sqrt{\ell+1}$.

Remark 2.3. *Note that* $\ell^{1/\ell} \le e^{1/e} \approx 1.4447$*, and* $0.8995 \le \beta_\ell \le 1$ *for all* $\ell \ge 2$*. Moreover,* $\delta_2 = 1/\sqrt{3}$ *and* $\ell \mapsto \delta_\ell$ *increases to* $1/\sqrt{e}$ *as* ℓ *tends to infinity.*

2.2 Under weak moments assumptions

We first introduce some more definitions and notations. For any real-valued integrable random variable X and any $r \ge 1$, let

$$\Lambda_r^+(X) := \sup_{t>0} t\,(\mathbb{P}(X > t))^{1/r}. \tag{2.1}$$

We say that X have a weak moment of order r if $\Lambda_r^+(|X|)$ is finite. Define also

$$\tilde{\Lambda}_r^+(X) := \sup_{u\in]0,1]} u^{(1/r)-1} \int_0^u Q_X(s)ds. \tag{2.2}$$

From the definition of Q_X, we have

$$\Lambda_r^+(X) = \sup_{u\in]0,1]} u^{1/r} Q_X(u). \tag{2.3}$$

Hence, we get that

$$\Lambda_r^+(X) \le \tilde{\Lambda}_r^+(X) \le \left(\tfrac{r}{r-1}\right)\Lambda_r^+(X). \tag{2.4}$$

Furthermore, from Proposition 1.4 (iii), $\tilde{\Lambda}_r^+(.)$ is subadditive. The main result of this section is the following Fuk-Nagaev type inequality for $\mathbb{E}[Z]-Z$.

Theorem 2.4. *Let* $\ell > 2$*. Assume that* $\Phi^-(X_1)$ *have a weak moment of order* ℓ*. Then for any* $u \in]0\,,1[$*,*

$$Q_{\mathbb{E}[Z]-Z}(u) \le \tilde{Q}_{\mathbb{E}[Z]-Z}(u) \tag{a}$$

$$\le \sqrt{2\log(1/u)}\Big(\sigma\sqrt{n} + \sqrt{V_n}\,\Big) + n^{1/\ell}\mu_\ell\,\Lambda_\ell^+(\Phi^-(X_1))u^{-1/\ell}, \tag{b}$$

where $\mu_\ell := 2 + \max(4/3, \ell/3)$ *and* V_n *is defined by* (1.10)*. Consequently,*

$$\mathbb{P}\Big(Z < \mathbb{E}[Z] - \sqrt{2\log(1/u)}\,\left(\sigma\sqrt{n} + \sqrt{V_n}\,\right) - n^{1/\ell}\mu_\ell\,\Lambda_\ell^+(\Phi^-(X_1))u^{-1/\ell}\Big) \le u. \tag{c}$$

Corollary 2.5. *Let* $\ell > 2$*. Assume that* $\Phi^-(X_1)$ *have a weak moment of order* ℓ*. Then*

$$\Lambda_\ell^+(\mathbb{E}[Z] - Z) \le \tilde{\Lambda}_\ell^+(\mathbb{E}[Z] - Z) \tag{a}$$

$$\le \sqrt{(\ell/e)}\Big(\sigma\sqrt{n} + \sqrt{V_n}\,\Big) + n^{1/\ell}\mu_\ell\,\Lambda_\ell^+(\Phi^-(X_1)), \tag{b}$$

where $\mu_\ell := 2 + \max(4/3, \ell/3)$*.*

2.3 Application to power-type tail

Let $Y_1, \dots, Y_n$ be a finite sequence of nonnegative, independent and identically distributed random variables and let $X_1, \dots, X_n$ be a finite sequence of independent and identically distributed random variables with values in some measurable space $(\mathcal{X}, \mathcal{F})$ such that the two sequences are independent. Let P denote the common distribution of the X_k. Let $\mathscr{G}$ be a countable class of measurable functions from $\mathcal{X}$ into $[-1, 1]$ such that for all $g \in \mathscr{G}$,

$$P(g) = 0 \quad \text{and} \quad P(g^2) < \delta^2 \text{ for some } \delta \in]0, 1[. \tag{2.5}$$

Let G be a measurable envelope function of $\mathscr{G}$, that is

$$|g| \le G \text{ for any } g \in \mathscr{G}, \text{ and } G(x) \le 1 \text{ for all } x \in \mathcal{X}. \tag{2.6}$$

Let G^- be a measurable function such that

$$(-g)_+ \le G^- \text{ for any } g \in \mathscr{G}, \text{ and } G^-(x) \le 1 \text{ for all } x \in \mathcal{X}. \tag{2.7}$$

We suppose furthermore that for some constant $p > 2$,

$$\mathbb{P}(Y_1 > t) \le t^{-p} \quad \text{ for any } t > 0. \tag{2.8}$$

Define now

$$Z := \sup_{g \in \mathscr{G}} \sum_{k=1}^{n} Y_k \, g(X_k). \tag{2.9}$$

Deviations of Z above the mean are studied in [10] and more precise bounds for V_n in terms of the uniform entropy integral (using a result of van der Vaart and Wellner [18]) are given. Here, we prodive upper bounds for the left-hand side deviations. Since the proofs are rather equivalent, we refer the reader to [10] for the details.

Let us first recall the definitions of covering numbers and uniform entropy integral.

Definition 2.6 (Covering number and uniform entropy integral)**.** *The covering number $N(\epsilon, \mathscr{G}, Q)$ is the minimal number of balls of radius ϵ in $\mathbb{L}^2(Q)$ needed to cover the set $\mathscr{G}$. The uniform entropy integral is defined by*

$$J(\delta, \mathscr{G}) := \sup_Q \int_0^\delta \sqrt{1 + \log N\Big(\epsilon \|G\|_{Q,2}, \mathscr{G}, Q\Big)} \, d\epsilon.$$

Here, the supremum is taken over all finitely discrete probability distributions Q on $(\mathcal{X}, \mathcal{F})$ and $\|f\|_{Q,2}$ denotes the norm of a function f in $\mathbb{L}^2(Q)$.

In the following theorem, K denotes an universal constant which may change from line to line.

Theorem 2.7. *Let $p > 2$. Let Z be defined by* (2.9)*. Under conditions* (2.5) – (2.8)*, the following results hold:*
(i) If Y_1 is $\mathbb{L}^p$-integrable, then

$$\|(\mathbb{E}[Z] - Z)_+\|_p \leq 2\,\beta_p\, p^{1/p}\sqrt{p+1}\left(\sigma\sqrt{n}+K\sqrt{\frac{p}{p-2}}\left(n^{q/4}\sqrt{J(\delta,\mathscr{G})}+\sqrt{p}\,n^{1/p}\Big(\frac{J(\delta,\mathscr{G})}{\delta}\Big)^{1/q}\right)\right) + 2\,n^{1/p}(p+1)\|Y_1\|_p\,\|G^{-}(X_1)\|_p, \quad (a)$$

where $\beta_p = \Big((\sqrt{\pi}/2)\,\Gamma(p/2)\,/\Gamma((p+1)/2)\Big)^{1/p}$.
(ii) Moreover,

$$\tilde{\Lambda}_p^{+}(\mathbb{E}[Z] - Z) \leq \sqrt{(p/e)}\left(\sigma\sqrt{n} + K\sqrt{\frac{p}{p-2}}\left(n^{q/4}\sqrt{J(\delta,\mathscr{G})} + \sqrt{p}\,n^{1/p}\Big(\frac{J(\delta,\mathscr{G})}{\delta}\Big)^{1/q}\right)\right) + n^{1/p}\mu_p, \quad (b)$$

where $q = p/(p-1)$ and $\mu_p = 2 + \max(4/3, p/3)$.

3 Under sub-Gamma tails on the left assumptions

Here we assume that $f(X_1)$ admits a sub-Gamma tail on the left for all $f \in \mathscr{F}$.

To the best of our knowledge, there exists only one exponential inequality for the left-hand side deviations of suprema of empirical processes in the unbounded case (from one side) which is due to Klein [8]. More precisely, Klein provides an upper bound on the Laplace transform of $\mathbb{E}[Z] - Z$ which implies the following deviation inequality:

Theorem 3.1 (Theorem 1.1 (3) in [8])**.** *Assume that for all $f \in \mathscr{F}$, $f \leq 1$, and for all integer $p \geq 2$, $|P(f^p)| \leq \sigma^2 p!/2$. Then*

$$\mathbb{P}\left(Z < \mathbb{E}[Z] - \sqrt{2xv_n} - x - \frac{x\sqrt{2v_n}}{\sqrt{x} + \sqrt{v_n/2}}\right) \leq \exp(-x),$$

where $v_n = n\sigma^2 + 2\mathbb{E}[Z]$.

The objective here is to relax the boundedness assumption on the right. Our main result is the following:

Theorem 3.2. *Assume that there exists a positive constant c such that for any $f \in \mathscr{F}$ and any $t \in]0\,, c[$,*

$$\log P(e^{-tf}) \leq \frac{\sigma^2 t^2}{2(1-ct)}. \tag{3.1}$$

Then for any $u \in]0\,,1[$,

$$\begin{aligned} Q_{\mathbb{E}[Z]-Z}(u) &\leq \tilde{Q}_{\mathbb{E}[Z]-Z}(u) & (a)\\ &\leq \sqrt{2\log(1/u)}\Big(\sigma\sqrt{n} + \sqrt{V_n}\Big) + c\log(1/u). & (b)\end{aligned}$$

Consequently, for any $x > 0$,

$$\mathbb{P}\Big(Z < \mathbb{E}[Z] - \sqrt{2x}\Big(\sigma\sqrt{n} + \sqrt{V_n}\Big) - cx\Big) \leq \exp(-x). \qquad (c)$$

Remark 3.3. *If the elements of $\mathscr{F}$ satisfy Bernstein's moment conditions on the left, then the hypothesis* (3.1) *above is satisfied. Precisely, assume that there exists a positive constant c such that, for any integer $p \geq 3$, and all $f \in \mathscr{F}$,*

$$P((-f)_+^p) \leq \frac{p!c^{p-2}}{2}\sigma^2. \tag{3.2}$$

Then, since $P(f) = 0$ for all $f \in \mathscr{F}$, $\log P(e^{-tf}) \leq \frac{\sigma^2 t^2}{2(1-ct)}$. *We refer the reader to Bercu, Delyon and Rio [4] (see the proof of their Theorem 2.1) for a proof.*

4 Under sub-Gaussian tails on the left assumptions

Here we assume that $f(X_1)$ admits a sub-Gaussian tail on the left for all $f \in \mathscr{F}$. We show in the following theorem that $Z - \mathbb{E}[Z]$ is sub-Gaussian on the left tail.

Theorem 4.1. *Assume that*

$$C(\mathscr{F}) := \sup_{t>0}\sup_{f\in\mathscr{F}} \frac{2}{t^2}\log P(e^{-tf}) < \infty. \tag{4.1}$$

Then for any $u \in]0\,,1[$,

$$\begin{aligned} Q_{\mathbb{E}[Z]-Z}(u) &\leq \tilde{Q}_{\mathbb{E}[Z]-Z}(u) && (a) \\ &\leq \sqrt{2\log(1/u)}\Big(\sqrt{nC(\mathscr{F})} + \sqrt{V_n}\Big). && (b) \end{aligned}$$

Consequently, for any $x > 0$,

$$\mathbb{P}(Z < \mathbb{E}[Z] - x) \leq \exp\left(- \frac{x^2}{2\Big(\sqrt{nC(\mathscr{F})} + \sqrt{V_n}\Big)^2} \right). \qquad (c)$$

4.1 Application to classes of nonnegative functions

Let $X_1, \ldots, X_n$ be a finite sequence of independent random variables and identically distributed according to a law P. Let $\mathscr{F}$ be a countable class of functions from $\mathcal{X}$ to $[0, +\infty[$ and define

$$\mathscr{F}_0 := \{f - P(f) : f \in \mathscr{F}\}. \qquad (4.2)$$

Define now

$$Z := \sup_{g \in \mathscr{F}_0} \sum_{k=1}^{n} g(X_k). \qquad (4.3)$$

Remark 4.2. *Equivalently, one can consider the case of classes of bounded from below functions which are centered under* P.

Let us define

$$\sigma^2 := \sup_{f \in \mathscr{F}} \mathrm{Var}_P(f) := \sup_{f \in \mathscr{F}} (P(f^2) - P(f)^2). \qquad (4.4)$$

Proposition 4.3. *One has*

$$C(\mathscr{F}_0) \leq \sigma^2 + \frac{1}{6} \sup_{f \in \mathscr{F}} P(f)^2. \qquad (a)$$

Consequently, for any $x < 0$,

$$\mathbb{P}(Z < \mathbb{E}[Z] - x) \leq \exp\left(- \frac{x^2}{2\Big(\sqrt{n(\sigma^2 + \frac{1}{6}\sup_{f \in \mathscr{F}} P(f)^2)} + \sqrt{V_n}\Big)^2} \right), \qquad (b)$$

where V_n *is defined by* (1.10) *with* $\Phi := \sup_{g \in \mathscr{F}_0} |g| = \sup_{f \in \mathscr{F}} |f - P(f)|$.

5 Suprema of randomized empirical processes

Let $X_1, \ldots, X_n$ be a finite sequence of independent random variables and identically distributed according to a law P, valued in some measurable space $(\mathcal{X}, \mathcal{F})$ with common distribution P. Let $Y_1, \ldots, Y_n$ be a finite sequence of independent, identically distributed and centered real-valued random variables with finite Laplace transform on a neighborhood of 0. And we assume that the two sequences are independent. Let $\mathscr{G}$ be a countable class of measurable functions from $\mathcal{X}$ to $[0\,,1]$ and define

$$Z := \sup_{g\in\mathscr{G}} \sum_{k=1}^{n} Y_k g(X_k). \tag{5.1}$$

The deviations of Z above the mean when the Y_k are independent standard Gaussian random variables are studied in Marchina [11, Section 2.7]. Let us define $v := \mathbb{E}[\sup_{g\in\mathscr{G}} g^2(X_1)] \leq 1$ and let γ_v be the function defined on $]0\,,\infty[$ by

$$\gamma_v(x) := x\sqrt{2}/\sqrt{\log(1 + v^{-1}(e^x - 1))}. \tag{5.2}$$

Note that $\gamma_v(x) = \sqrt{2vx}(1+O(x))$ as x goes to 0 and $\gamma_v(x) \sim \sqrt{2x}$ as x goes to infinity. We have for any $x > 0$,

$$\mathbb{P}\Big(Z - \mathbb{E}[Z] > n\gamma_v\Big(\frac{x}{n}\Big)\Big) \leq \exp(-x). \tag{5.3}$$

This result is based on a slighlty different use of martingale methods and comparison inequalities which in fact allow to examine general separately convex functions of independent random variables. However, we provide in [11] only right deviation inequalities and it seems not that easy to generalize to obtain left deviation inequalities. Therefore, the objective here is to fill this gap. In what follows, we establish a result in a more general setting that Gaussian case. When we specify to the Gaussian case, we obtain a result similar to (5.3) with a better variance term v but with an additionnal corrective term.

Theorem 5.1. *Let Z be defined by* (5.1)*. Assume that* $\mathrm{Var}(Y_1) = 1$ *and that there exists a nondecreasing convex function R from* $[0\,,\infty[$ *to* $[1\,,\infty]$ *such that*

$$\mathbb{E}[\exp(-tY_1)] \leq R(t) \textit{ for any } t \geq 0, \tag{5.4}$$

$$R(0) = 1, \quad \textit{and } \tilde{R} : t \mapsto R(\sqrt{t}\,) \textit{ is convex.} \tag{5.5}$$

Define

$$\sigma^2 := \sup_{g\in\mathscr{G}} \mathrm{Var}(Y_1 g(X_1)) = \sup_{g\in\mathscr{G}} P(g^2), \tag{5.6}$$

and let the function $\ell_{\sigma^2,R}$ defined for any $t \geq 0$ by

$$\ell_{\sigma^2,R}(t) := \log(1 + \sigma^2(R(t) - 1)). \tag{5.7}$$

We denote by G an envelope function of $\mathscr{G}$, that is $g(x) \leq G(x) \leq 1$ for any $x \in \mathcal{X}$ and any $g \in \mathscr{G}$. Then, for any $x \geq 0$,

$$\mathbb{P}\Big(Z < \mathbb{E}[Z] - n\,\ell^{*-1}_{\sigma^2,R}\Big(\frac{x}{n}\Big) - \sqrt{2xV_n}\Big) \leq \exp(-x), \tag{a}$$

where V_n is defined by (1.10) *with $\Phi(x,y) = |y|G(x)$.*
Consequently, for any $x \geq 0$,

$$\mathbb{P}\Big(Z < \mathbb{E}[Z] - n\,\gamma_{\sigma^2,R}\Big(\frac{x}{n}\Big) - \sqrt{2xV_n}\Big) \leq \exp(-x), \tag{b}$$

where $\gamma_{v,R}$, $v \geq 0$, is the function defined for any $x \geq 0$ by

$$\gamma_{v,R}(x) = \frac{2x}{R^{-1}(1 + v^{-1}(e^x - 1))}. \tag{5.8}$$

Remark 5.2. *If there exists a nondecreasing convex function ℓ such that $\ell_{-Y_1}(t) \leq \ell(t)$ for any $t \geq 0$, $\ell(0) = 0$ and $t \mapsto \ell(\sqrt{t})$ is convex, then $R(t) = e^{\ell(t)}$ satisfies hypotheses of the above theorem.*

Remark 5.3. *Note that if $R(t) = \exp(t^2/2)$ (i.e. the Y_k are sub-Gaussian), and $\sigma^2 = 1$, then $\ell^{*-1}_{1,R} = \gamma_{1,R}$.*

Example 5.4 (Sub-Gaussian case)**.** *Here we assume that $R(t) := \exp(t^2/2)$. Clearly, this function R satisfies hypotheses of Theorem 5.1. Notice now that $\gamma_{\sigma^2,R} = \gamma_{\sigma^2}$ where the function γ_{σ^2} is defined by* (5.2)*. Theorem 5.1 leads then to the following inequality*

$$\mathbb{P}\Big(Z < \mathbb{E}[Z] - n\,\gamma_{\sigma^2}\Big(\frac{x}{n}\Big) - \sqrt{2xV_n}\Big) \leq \exp(-x). \tag{5.9}$$

Note that the variance term σ^2 in the inequality above is better than the variance term $v = \mathbb{E}[\sup_{g\in\mathscr{G}} g^2(X_1)]$ in Inequality (5.3)*.*
However, methods developed in [11] allow us to consider the non-identically distributed case, that is $Z = \sup_{g\in\mathscr{G}} \sum_{k=1}^n a_k Y_k\, g(X_k)$ where $a_1, \ldots, a_n$ are a sequence of positive reals.

Example 5.5 (Sub-Gamma case)**.** *Here we assume that*

$$R(t) := \exp\Big(\frac{t^2}{2(1-t)}\Big),$$

which satisfies the hypotheses of Theorem 5.1. Then

$$\mathbb{P}\Big(Z < \mathbb{E}[Z] - n\,\gamma_{\sigma^2,L}\Big(\frac{x}{n}\Big) - \sqrt{2xV_n}\Big) \leq \exp(-x), \tag{5.10}$$

where

$$\gamma_{\sigma^2,R} := \frac{2x}{-z_x + \sqrt{z_x^2 + 2z_x}} \text{ with } z_x := \log(1 + \sigma^{-2}(e^x - 1)). \tag{5.11}$$

Here, note that $\gamma_{\sigma^2,R} = (\sqrt{2\sigma^2 x} + x)(1 + O(x))$ *as* x *goes to* 0*, and* $\gamma_{\sigma^2,R} \sim 2x$ *as* x *tends to infinity.*

Example 5.6 (Symmetric exponential case)**.** *Here we assume that* $R(t) := 1/(1 - t^2)$*, which is equivalent to suppose that* $Y_1 \overset{law}{=} \varepsilon W$*, where* ε *is a Rademacher random variable,* W *is an exponential random variable with parameter* 1 *and the two random variables are independent. We can easily verify that* R *satisfies the hypotheses of Theorem 5.1. Then*

$$\mathbb{P}\Big(Z < \mathbb{E}[Z] - n\,\gamma_{\sigma^2,R}\Big(\frac{x}{n}\Big) - \sqrt{2xV_n}\Big) \leq \exp(-x), \tag{5.12}$$

where

$$\gamma_{\sigma^2,R} := \frac{2x\sqrt{1 + \sigma^{-2}(e^x - 1)}}{\sqrt{\sigma^{-2}(e^x - 1)}}. \tag{5.13}$$

Here, note that $\gamma_{\sigma^2,R} = 2\sqrt{\sigma^2 x} + O(x)$ *as* x *goes to* 0 *and* $\gamma_{\sigma^2,R} \sim 2x$ *as* x *tends to infinity.*

6 Proofs

The starting point of the proofs is a martingale decomposition of $\mathbb{E}[Z] - Z$ which we briefly recall (we refer the reader to [10] for more details). First by virtue of the monotone convergence theorem, we can suppose that $\mathscr{F}$ is a finite class of functions. Set $\mathcal{F}_0 := \{\varnothing, \Omega\}$ and for all $k = 1, \ldots, n$, $\mathcal{F}_k := \sigma(X_1, \ldots, X_k)$ and $\mathcal{F}_n^k := \sigma(X_1, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n)$. Let $\mathbb{E}_k$ (respectively $\mathbb{E}_n^k$) denote the conditional expectation operator associated with $\mathcal{F}_k$ (resp. $\mathcal{F}_n^k$). Set also $Z^{(k)} := \sup\{nP_n(f) - f(X_k) : f \in \mathscr{F}\}$ and $Z_k := \mathbb{E}_k[Z]$. Let us number the functions of the class $\mathscr{F}$ and consider the random variables

$$\tau := \inf\{i > 0 : nP_n(f_i) = Z\} \text{ and } \tau_k := \inf\{i > 0 : nP_n(f_i) - f_i(X_k) = Z^{(k)}\}.$$

We notice that

$$f_{\tau_k}(X_k) \le Z - Z^{(k)} \le f_\tau(X_k). \tag{6.1}$$

Set now $\xi_k := \mathbb{E}_k[f_{\tau_k}(X_k)]$, $r_k := Z - Z^{(k)} - \xi_k$ and $\varepsilon_k := \mathbb{E}_k[f_\tau(X_k)] - \xi_k$. Projecting then (6.1) on $\mathcal{F}_k$ leads to

$$\xi_k \le \xi_k + r_k \le \xi_k + \varepsilon_k. \tag{6.2}$$

Since $\mathbb{E}_n^k[f_{\tau_k}(X_k)] = P(f_{\tau_k})$, the centering assumption on the elements of $\mathscr{F}$ ensures that

$$Z_k - Z_{k-1} = \xi_k + r_k - \mathbb{E}_{k-1}[r_k]. \tag{6.3}$$

Thus we get the following decomposition:

$$\mathbb{E}[Z] - Z = \Xi_n^o + R_n^o, \tag{6.4}$$

where

$$\Xi_n^o := \sum_{k=1}^n \xi_k^o, \quad \xi_k^o := -\xi_k \ \text{ and } \ R_n^o := \sum_{k=1}^n (\mathbb{E}_{k-1}[r_k] - r_k).$$

Now the martingale R_n^o is sub-Gaussian, as shown by the lemma below.

Lemma 6.1. *For any $t \ge 0$, we have*

$$\ell_{R_n^o}(t) := \log \mathbb{E}[\exp(tR_n^o)] \le \frac{t^2 V_n}{2}, \tag{a}$$

where V_n is defined by (1.10)*. Consequently, for any $u \in]0\,,1[$,*

$$\tilde{Q}_{R_n^o}(u) \le \ell_{R_n^o}^{*-1}(\log(1/u)) \le \sqrt{2\, V_n \log(1/u)}. \tag{b}$$

Proof of Lemma 6.1. We start by showing that

$$\mathbb{E}_{k-1}[r_k] - r_k \le E_{n-k+1} \ \text{ and } \ \mathbb{E}_{k-1}[(\mathbb{E}_{k-1}[r_k] - r_k)^2] \le \mathbb{E}[\zeta_{n-k+1}^2]. \tag{6.5}$$

These bounds are stated and proved in Marchina [10]. We briefly recall below main arguments. First, note that $r_k \ge 0$ by (6.2). Next, a property of exchangeability of variables (see Lemma 3.10 in [10]) shows that $\mathbb{E}_{k-1}[r_k] \le E_{n-k+1}$. Then we get the first bound of (6.5). The second bound follows from a comparison inequality of generalized moments due to Bentkus [3, Lemma 1] (see also Lemma 2.1 in Marchina [11]). Indeed, since (6.2) implies $0 \le r_k \le 2\Phi(X_k)$, we have $\mathbb{E}_{k-1}[\varphi(r_k)] \le \mathbb{E}[\varphi(\zeta_{n-k+1})]$ for any real-valued, convex and differentiable function φ such that $\lim_{x\to-\infty} \varphi(x) = 0$. Then, since $x \mapsto x_+^2$ is such a function and $r_{k+} = r_k$, we get

$$\mathbb{E}_{k-1}[(\mathbb{E}_{k-1}[r_k] - r_k)^2] \le \mathbb{E}_{k-1}[r_{k+}^2] \le \mathbb{E}[\zeta_{n-k+1}^2]. \tag{6.6}$$

We recall that ζ_k denotes a random variable with distribution function $F_{2\Phi(X_1),q_k}$ such that $\mathbb{E}[\zeta_k] = E_k$. Thus we have $\max(\mathbb{E}[\zeta_k^2], E_k^2) = \mathbb{E}[\zeta_k^2]$. Applying now the bound for martingales with differences bounded from above proved by Bentkus [2] (see his Inequality (2.16)), we get for any $t \geq 0$,

$$\mathbb{E}[\exp(tR_n^o)] \leq \exp(t^2 V_n/2), \tag{6.7}$$

which then gives (a). Now, it is a classical calculation that

$$\inf_{t>0} \left\{ \frac{1}{t} \Big(\frac{t^2 V_n}{2} + x \Big) \right\} = \sqrt{2\, V_n\, x}, \tag{6.8}$$

where the infimum is given by the optimal value $t = \sqrt{2\, x/V_n}$. Finally, Proposition 1.4 (iv) ends the proof of (b). □

Now, in view of (6.4), it remains us to bound up the CVaR (or the log-Laplace) of Ξ_n^o according to the assumptions made in different sections.

6.1 Proofs of Section 2

Proofs in this section are almost identical to that of the article [10]. Thus we only give main arguments.

Proof of Theorem 2.1. The proof is similar to the proof of Theorem 3.2 in [10]. First, note that one derive from Lemma 6.1 (a) and the usual Cramér-Chernoff calculation that for any $x \geq 0$,

$$\mathbb{P}(R_n^o \geq x) \leq \exp\Big(- \frac{x^2}{2V_n} \Big). \tag{6.9}$$

We bound up the log-Laplace of Ξ_n^o thanks to a Fuk-Nagaev type inequality for martingales obtained by Courbot [6]. Hence, we first need to control the quadratic variation. Since τ_k is $\mathcal{F}_n^k$-measurable, the centering assumption on the elements of $\mathscr{F}$ yields that

$$\mathbb{E}_{k-1}[\xi_k^{o2}] = \mathbb{E}_{k-1}[\xi_k^2] \leq \mathbb{E}_{k-1}\mathbb{E}_n^k[f_{\tau_k}^2(X_k)] = \mathbb{E}_{k-1}P(f_{\tau_k}^2) \leq \sigma^2, \tag{6.10}$$

where the first inequality follows from the conditionnal Jensen inequality. Then, since $\xi_k^o \leq \Phi(X_k)$ and $(1+x)\log(1+x) - x \geq x\log(1+x)/2$, Theorem 1 in Courbot [6] leads to the following inequality:

$$\mathbb{P}(\Xi_{n+}^o \geq x) \leq \Big(1 + \frac{x^2}{sn\sigma^2}\Big)^{-s/2} + n\mathbb{P}\Big(\Phi(X_1) > \frac{x}{s}\Big), \tag{6.11}$$

for any $x > 0$ and any $s > 0$. Furthermore, $\mathbb{E}[Z] - Z = \Xi_n^o + R_n^o \leq \Xi_{n+}^o + R_n^o$, which yields that for any $x \geq 0$,

$$\mathbb{P}(Z < \mathbb{E}[Z] - x) \leq \inf_{t \in [0,1]} \{\mathbb{P}(\Xi_{n+}^o \geq tx) + \mathbb{P}(R_n^o \geq (1-t)x)\}. \tag{6.12}$$

Since the optimization in t in the right-hand side is difficult to calculate, we take $t = 1/2$ in the sequel. Finally, combining (6.12), (6.11) and (6.9) ends the proof of Theorem 2.1. □

Proof of Corollary 2.2. The proof relies on the following equality: for any $\ell \geq 1$,

$$\mathbb{E}[(\mathbb{E}[Z] - Z)_+^\ell] = \ell \int_0^\infty \mathbb{P}(\mathbb{E}[Z] - Z \geq x) x^{\ell-1} dx. \tag{6.13}$$

After calulation of this integral, we put $s = \ell + 1$ and we use the sub-additivity of $x \mapsto x^\ell$ to conclude the proof. □

Proof of Theorem 2.4. The proof is similar to the proof of Theorem 3.3 in [10]. Inequality (a) is the point (ii) of Propostion 1.4 and Inequality (c) follows immediately from (b) by the point (i) of the same Proposition 1.4. Let us prove (b). Since $\xi_k^o \leq \Phi(X_k)$, we have

$$C_\ell^w(\Xi_n^o) \leq n^{1/\ell} \Lambda_\ell^+(\Phi(X_1)), \tag{6.14}$$

where

$$C_\ell^w(\Xi_n^o) := \left\| \sup_{t>0} \left(t^\ell \sum_{k=1}^n \mathbb{P}(\xi_{k+}^o > t \mid \mathcal{F}_{k-1}) \right) \right\|_\infty^{1/\ell}.$$

Now, recalling (6.10), the Fuk-Nagaev type inequality for martingale obtained by Rio [15] (see his Theorem 4.1) yields

$$\tilde{Q}_{\Xi_n^o}(u) \leq \sigma \sqrt{2 \log(1/u)} + n^{1/\ell} \mu_\ell \, \Lambda_\ell^+(\Phi(X_1)) u^{-1/\ell}. \tag{6.15}$$

Combining Lemma 6.1, (6.14) and the point (iii) of Proposition 1.4 implies Inequality (b) of the theorem and thus concludes the proof. □

Proof of Corollary 2.5. Inequality (a) follows from (2.4). To prove (b) we proceed exactly as in Rio [13, Theorem 5.1] (see also the proof of (b) of Corollary 3.7 in [10]). □

6.2 Proof of Section 3

Proof of Theorem 3.2. As in Theorem 2.4, we only have to prove (b). Moreover we already saw in the proof of Theorem 2.1 that

$$\langle \Xi^o \rangle_n := \sum_{k=1}^{n} \mathbb{E}_{k-1}[\xi_k^{o2}] \le n\sigma^2. \tag{6.16}$$

Furthermore, as in (6.10), since τ_k is $\mathcal{F}_n^k$-measurable, the conditionnal Jensen inequality implies that for any $t \ge 0$,

$$\mathbb{E}_{k-1}[\exp(t\xi_k^o)] \le \mathbb{E}_{k-1}\mathbb{E}_n^k[\exp(-tf_{\tau_k}(X_k))] = \mathbb{E}_{k-1}[P(\exp(-tf_{\tau_k}))]. \tag{6.17}$$

Whence, assumption (3.1) on the elements of $\mathscr{F}$ and an immediate induction on n give that

$$\log \mathbb{E}[\exp(t\, \Xi_n^o)] \le n\, \frac{\sigma^2 t^2}{2(1-ct)}. \tag{6.18}$$

Now, it is a classical calculation that

$$\inf_{t \in]0,1/c[} \left(\frac{\sigma^2 t}{2(1-tc)} + \frac{x}{t} \right) = cx + \sqrt{2x\sigma^2}, \tag{6.19}$$

where the infimum is given by the optimal value $t = \sqrt{2x}/(\sqrt{\sigma^2} + c\sqrt{2x})$. Recalling (1.6), one conclude that for any $u \in]0\,,1[$,

$$\tilde{Q}_{\Xi_n}(u) \le c\log(1/u) + \sigma\sqrt{2n\log(1/u)}. \tag{6.20}$$

Finally, combining Lemma 6.1, (6.20) and the point (iii) of Proposition 1.4 implies inequality (b) of the theorem and completes the proof. □

6.3 Proofs of Section 4

Proof of Theorem 4.1. As previously mentioned, we only have to prove (b). By reasoning in the same way as (6.17), the assumptions on the elements of $\mathscr{F}$ allow us to derive that

$$\log \mathbb{E}[\exp(t\Xi_n^o)] \le n\frac{t^2}{2} C(\mathscr{F}), \tag{6.21}$$

for any $t \ge 0$. Therefore, the same conclusion as in the proof of Theorem 3.2 yields that

$$\tilde{Q}_{\Xi_n^o}(u) \le \sqrt{2nC(\mathscr{F})\log(1/u)}, \tag{6.22}$$

for any $u \in]0\,,1[$, which associated to Lemma 6.1 and Proposition 1.4 give (b) and end the proof. □

Proof of Proposition 4.3. Both inequality (c) of Theorem 4.1 and the upper bound on $C(\mathscr{F}_0)$ (a) imply (b). Hence we only have to prove the upper bound on $C(\mathscr{F}_0)$. Let $g = f - P(f) \in \mathscr{F}_0$. Note that $-g = P(f) - f \leq P(f)$ and $\mathrm{Var}_P(g) = \mathrm{Var}_P(f) \leq \sigma^2$, where σ^2 is defined by (4.4). Then Lemma 2.36 in Bercu, Delyon and Rio [4] implies that

$$C(\mathscr{F}_0) \leq \frac{1}{2} \sup_{f\in\mathscr{F}} \Big\{ P(f)^2 \, \varphi\Big(\frac{\sigma^2}{P(f)^2}\Big) \Big\}, \tag{6.23}$$

where φ is the function defined by

$$\varphi(x) = \begin{cases} \dfrac{1-x^2}{|\log(x)|} & \text{if} \quad x < 1 \\ \quad 2x & \text{if} \quad x \geq 1. \end{cases} \tag{6.24}$$

The following lemma provides an upper bound for the function φ. We refer the reader to Lemma 2.37 and the proof of Theorem 2.33 in [4] for a proof.

Lemma 6.2. *For any $x \geq 0$, $\varphi(x) \leq 2x + \frac{1}{3}(1-x)_+$.*

Therefore combining this bound with (6.23) one has

$$C(\mathscr{F}_0) \leq \sigma^2 + \frac{1}{6} \sup_{f\in\mathscr{F}} \{(P(f)^2 - \sigma^2)_+\} \leq \sigma^2 + \frac{1}{6} \sup_{f\in\mathscr{F}} P(f)^2, \tag{6.25}$$

which is exactly (a) of Proposition 4.3. □

6.4 Proof of Section 5

Proof of Theorem 5.1. As previously, by an induction on n, to bound up $\log \mathbb{E}[\exp(t\,\Xi_n^o)]$ we only have to bound up the conditionnal Laplace of the increments ξ_k^o. Let us denote by η_k the random variable $\mathbb{E}_k[g_{\tau_k}(X_k)]$. Then, we write $\xi_k^o = -\eta_k Y_k$. Now, since η_k is $\mathcal{F}_{k-1} \bigvee \sigma(X_k)$-measurable, Y_k and η_k are independent and one has for any $t \geq 0$,

$$\begin{aligned} \mathbb{E}_{k-1}[\exp(t\,\xi_k^o)] &= \mathbb{E}_{k-1}[\exp(-(t\eta_k)Y_k)] \\ &\leq \mathbb{E}_{k-1}[R(t\eta_k)] \\ &= \mathbb{E}_{k-1}[\tilde{R}(t^2\eta_k^2)]. \end{aligned} \tag{6.26}$$

Observe now that $0 \leq \eta_k^2 \leq 1$ and $\mathbb{E}_{k-1}[\eta_k^2] \leq \sigma^2$ by a similar reasoning to (6.10) by means of the conditionnal Jensen's inequality. Therefore, applying (conditionally to $\mathcal{F}_{k-1}$) a classical convex comparison due to Hoeffding [7] (see his inequalities (4.1)–(4.2)) leads to

$$\mathbb{E}_{k-1}[\varphi(\eta_k^2)] \leq \mathbb{E}[\varphi(\theta_{\sigma^2})] \quad \text{for any convex function } \varphi : \mathbb{R} \to \mathbb{R}, \tag{6.27}$$

where θ_{σ^2} is a Bernoulli random variable with parameter σ^2, independent of all other random variables. Now, since we have assumed that the function $\tilde{R}$ is convex, (6.26) becomes

$$\begin{aligned}\mathbb{E}_{k-1}[\exp(t\,\xi_k^o)] \leq \mathbb{E}_{k-1}[\tilde{R}(t^2\eta_k^2)] &\leq \mathbb{E}[\tilde{R}(t^2\theta_{\sigma^2})] \\ &= \mathbb{E}[R(t\,\theta_{\sigma^2})] \\ &= 1 + \sigma^2(R(t) - 1).\end{aligned} \tag{6.28}$$

Thus, as mentionned at the beginning of the proof, we derive by an induction on n that for any $t \geq 0$

$$\log \mathbb{E}[\exp(t\,\Xi_n^o)] \leq n \log(1 + \sigma^2(R(t) - 1)) = n\,\ell_{\sigma^2,R}(t). \tag{6.29}$$

Combining now (6.29), Lemma 6.1 and Proposition 1.4 (*iv*) and (*iii*), concludes the proof of Inequality (*a*) of Theorem. Next, to prove (*b*), we show that $\ell^{*-1}_{\sigma^2,R} \leq \gamma_{\sigma^2,R}$. First, using the variational formula (1.5), one has

$$\ell^{*-1}_{\sigma^2,R}(x) = \inf_{t>0}\{t^{-1}(\ell_{\sigma^2,R} + x)\}. \tag{6.30}$$

Then, putting $t_x := R^{-1}(1 + \sigma^{-2}(e^x - 1))$ in (6.30) concludes the proof. Note that t_x is optimal in the standard Gaussian case, that is $R(t) = \exp(t^2/2)$ and $\sigma^2 = 1$. □

References

[1] R. Adamczak. A tail inequality for suprema of unbounded empirical processes with applications to Markov chains. *Electron. J. Probab.*, 13:no. 34, 1000–1034, 2008.

[2] V. Bentkus. An inequality for tail probabilities of martingales with differences bounded from one side. *Journal of Theoretical Probability*, 16(1):161–173, 2003.

[3] V. Bentkus. An extension of the Hoeffding inequality to unbounded random variables. *Lithuanian Mathematical Journal*, 48(2):137–157, 2008.

[4] B. Bercu, B. Delyon, and E. Rio. *Concentration Inequalities for Sums and Martingales.* SpringerBriefs in Mathematics. Springer, 2015.

[5] S. Boucheron, O. Bousquet, G. Lugosi, and P. Massart. Moment inequalities for functions of independent random variables. *Ann. Probab.*, 33(2):514–560, 03 2005.

[6] B. Courbot. Rates of convergence in the functional CLT for martingales. *Academie des Sciences Paris Comptes Rendus Serie Sciences Mathematiques*, 328:509–513, March 1999.

[7] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[8] T. Klein. Une inégalité de concentration à gauche pour les processus empiriques. *C. R. Math. Acad. Sci. Paris*, 334(6):501–504, 2002.

[9] J. Lederer and S. van de Geer. New concentration inequalities for suprema of empirical processes. *Bernoulli*, 20(4):2020–2038, 11 2014.

[10] A. Marchina. Concentration inequalities for suprema of unbounded empirical processes. Preprint on <hal-01545101>, 2017.

[11] A. Marchina. Concentration inequalities for separately convex functions. *Bernoulli*, 24(4A):2906–2933, 11 2018.

[12] I. Pinelis. An optimal three-way stable and monotonic spectrum of bounds on quantiles: a spectrum of coherent measures of financial risk and economic inequality. *Risks*, 2(3):349–392, 2014.

[13] E. Rio. About the constants in the Fuk-Nagaev inequalities. *Electron. Commun. Probab.*, 22:12 pp., 2017.

[14] E. Rio. *Asymptotic theory of weakly dependent random processes*, volume 80 of *Probability Theory and Stochastic Modelling*. Springer, Berlin, 2017. Translated from the 2000 French edition [MR2117923].

[15] E. Rio. New deviation inequalities for martingales with bounded increments. *Stochastic Processes and their Applications*, 127(5):1637 – 1648, 2017.

[16] S. van de Geer and J. Lederer. The Bernstein-Orlicz norm and deviation inequalities. *Probab. Theory Related Fields*, 157(1-2):225–250, 2013.

[17] A.W. van der Vaart and J.A. Wellner. *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Series in Statistics. Springer New York, 1996.

[18] A.W. van der Vaart and J.A. Wellner. A local maximal inequality under uniform entropy. *Electron. J. Statist.*, 5:192–203, 2011.