Electron. Commun. Probab. **20** (2015), no. 77, 1–10.
DOI: 10.1214/ECP.v20-4204
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Exponential inequalities for weighted sums of bounded random variables

Emmanuel Rio*

Abstract

In this paper we give new exponential inequalities for weighted sums of real-valued independent random variables bounded on the right. Our results are extensions of the results of Bennett (1962) to weighted sums

Keywords: Hoeffding's inequality ; Bernstein's inequality ; Bennett's inequality ; Deviation inequalities ; Weighted sums.
AMS MSC 2010: 60E15.
Submitted to ECP on March 25, 2015, final version accepted on October 25, 2015.

1 Introduction and previous results

In this paper, we are interested in the deviation on the right of weighted sums of independent random variables. We will assume throughout the paper that the random variables are bounded on the right and that the weights are positive. So, let $X_1, X_2, \ldots$ be a sequence of independent random variables satisfying the conditions below:

$$v_k := \operatorname{Var} X_k < \infty, \ \mathbb{E}(X_k) = 0 \ \text{ and } X_k \le 1 \ \text{ almost surely.} \tag{1.1}$$

Let then $(c_k)_{k>0}$ be a sequence of positive deterministic reals. The normalized weighted sums $(W_n)_{n>0}$ are defined by

$$W_n = V_n^{-1/2} \sum_{k=1}^{n} c_k X_k, \ \text{ where } V_n = \sum_{k=1}^{n} c_k^2 v_k. \tag{1.2}$$

We now recall some known results on random variables bounded on the right. Bennett (1962, page 42) proved that, for a centered random variable X with variance v bounded on the right by some positive constant c, the value of $\mathbb{E}(\exp(tX)$ is maximized for any positive t by the discrete distribution μ given by

$$\mu(\{c\}) = v/(c^2 + v) \ \text{ and } \mu(\{-v/c\}) = c^2/(c^2 + v). \tag{1.3}$$

When $c = 1$, Bennett's result ensures that, for any positive t,

$$\log \mathbb{E}(\exp(tX)) \le \ell_v(t) := \log\bigl(ve^t + e^{-vt}\bigr) - \log(1 + v). \tag{1.4}$$

Next Hoeffding (1963, Lemma 3, page 23) proved that, for any $t > 0$, the function $v \to \ell_v(t)$ is concave with respect to v. Hoeffding's lemma ensures that, for any $t > 0$,

$$\ell_{v_1}(t) + \ell_{v_2}(t) + \cdots + \ell_{v_n}(t) \le n\ell_v(t) \ \text{ with } v = (v_1 + v_2 + \cdots + v_n)/n. \tag{1.5}$$

*UMR 8100 CNRS, Laboratoire de Mathématiques de Versailles, France. E-mail: emmanuel.rio@uvsq.fr

Using the above results, Hoeffding (1963, Theorem 3, page 16) obtained the large deviations inequality

$$\mathbb{P}(X_1 + X_2 + \cdots + X_n \geq nu) \leq \exp(-n\ell_v^*(u)), \tag{1.6}$$

where $\ell_v^*(u) = +\infty$ for $u > 1$ and

$$\ell_v^*(u) = \sup_{t\geq 0}\big(ut - \ell_v(t)\big) = \Big(\frac{v+u}{v+1}\Big)\log\Big(1+\frac{u}{v}\Big) + \Big(\frac{1-u}{v+1}\Big)\log\big(1-u\big) \text{ for } u \in [0,1]. \tag{1.7}$$

Hoeffding (1963) also proved that, for any positive u,

$$\ell_v^*(u) \geq u^2/(2v) \text{ for } v \geq 1 \text{ and } \ell_v^*(u) \geq u^2 \log(1/v)/(1-v^2) \text{ for } v < 1. \tag{1.8}$$

Since $\ell_v = (\ell_v^*)^*$, the above lower bound implies that, for any positive t,

$$\ell_v(t) \leq \varphi(v)t^2/4 \text{ with } \varphi(v) = 2v \text{ for } v \geq 1 \text{ and } \varphi(v) = (1-v^2)/\log(1/v) \text{ for } v < 1. \tag{1.9}$$

The upper bound (1.9) was rediscovered much later by Kearns and Saul (1998). We refer to Bentkus (2002, 2003, 2004), Pinelis (2014), Fan, Grama and Liu (2015) and Bercu, Delyon and Rio (2015) for additional results concerning Hoeffding's type inequalities and exponential inequalities for sums or martingales.

Let us now turn to the general case of distincts weights. Let $c = \max(c_1, c_2, \ldots, c_n)$. Applying Inequality (1.6), which is Hoeffding's Theorem 3, to the random variables $X_k' = (c_k/c)X_k$ with $u = x\sqrt{V_n}\,/nc$, one can obtain that

$$\mathbb{P}(W_n \geq x) \leq \exp\big(-n\ell^*_{(V_n/nc^2)}(x\sqrt{V_n}\,/nc)\big). \tag{1.10}$$

Let then the function g be defined by

$$g(u) = (1+u)\log(1+u) - u \text{ for any } u \geq 0. \tag{1.11}$$

Since $\ell_v^*(u) \geq vg(x/v)$, Inequality (1.10) implies the inequality of Bennett (1962):

$$\mathbb{P}(W_n \geq x) \leq \exp\big(-c^{-2}V_n g(V_n^{-1/2}cx)\big). \tag{1.12}$$

Now, on the one hand, if $V_n \geq nc^2$, (1.10) and the first part of (1.8) ensure that

$$\mathbb{P}(W_n \geq x) \leq \exp\big(-x^2/2\big), \tag{1.13}$$

for any positive x, and, one the other hand, if $v_k \geq 1$ for every k in $[1,n]$, then, by (1.9),

$$\log \mathbb{E}\big(\exp(tW_n)\big) \leq \frac{1}{2V_n}\sum_{k=1}^n v_k c_k^2 t^2 = \frac{1}{2}t^2, \tag{1.14}$$

which also implies (1.13). If $v_k < 1$ for some k in $[1,n]$ and $V_n < nc^2$, the situation becomes more intricate. Using (1.9), one can obtain the Kearns-Saul type inequality

$$\mathbb{P}(W_n \geq x) \leq \exp\Big(-\frac{V_n x^2}{\sum_{k=1}^n c_k^2\varphi(v_k)}\Big) \tag{1.15}$$

(see Bercu, Delyon and Rio (2015) for more details). However, in this inequality, the denominator may be much larger than $2V_n$. Next, using (1.12) and the lower bound

$$g(x) \geq x^2/(2v + 2x/3), \tag{1.16}$$

one can obtain the Bernstein inequality

$$\mathbb{P}(W_n \geq x) \leq \exp\Big(-\frac{x^2}{2\big(1+cx/(3\sqrt{V_n}\,\big)}\Big). \tag{1.17}$$

In the above inequality, the first order term is exact. However the second order term may be very large, due to the fact that $c = \max(c_1, c_2, \dots, c_n)$, which limits drastically the accuracy of this inequality in some cases.

In this paper, we will obtain inequalities with new second order terms. The main idea of the paper is that, for some adequate function γ,

$$\ell_v(t) \leq v(t^2/2) + \gamma(v)(t^3/6). \tag{1.18}$$

Combining this bound with (1.4), we will obtain upper bounds on the Laplace transform of W_n which will allow us to get new exponential inequalities. In Section 2, we explain our method and we give exponential inequalities with upper bounds depending on the above function γ. Section 3 is devoted to upper bounds on γ for large values of v and Section 4 is devoted to upper bounds on γ for small values of v. These upper bounds on the function γ will allow us to show that, in the case of weighted sums of independent and bounded random variables, our method provides more efficient inequalities than the inequalities of Bennett (1962), Hoeffding (1963) and Kearns and Saul (1998) for intermediate values of the deviation, under adequate conditions on the weigths c_k and the variances v_k (*see Remark 4.2*). Finally, in Section 5, we compare our results with the previous results on an example.

2 The main inequality

In this Section, we explain how Inequality (1.18) can be used to obtain new exponential inequalities. The estimation of the function γ appearing here is carried out in Sections 3 and 4. Let us now state our main result.

Theorem 2.1. *Let the random variable W_n be defined by (1.2). Define the function $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ by*

$$\gamma(v) = 0 \text{ if } v \geq 1 \text{ and } \gamma(v) = 6 \sup_{t>0} t^{-3}(\ell_v(t) - vt^2/2) \text{ if } v < 1. \tag{2.1}$$

For any positive t,

$$\log \mathbb{E}\big(\exp(tW_n)\big) \leq (t^2/2) + A_{3,n}(t^3/6), \text{ where } A_{3,n} = V_n^{-3/2} \sum_{k=1}^n c_k^3 \gamma(v_k). \tag{2.2}$$

Consequently, for any positive x,

$$\mathbb{P}\big(W_n \geq x\big) \leq \exp\left(-\frac{(1+2A_{3,n}x)^{3/2} - 1 - 3A_{3,n}x}{3A_{3,n}^2}\right) \tag{2.3}$$

$$\leq \exp\Big(-\frac{g\big(A_{3,n}x\big)}{A_{3,n}^2}\Big) \leq \exp\Big(-\frac{x^2}{2(1+xA_{3,n}/3)}\Big), \tag{2.4}$$

where $g(x) = (1+x)\log(1+x) - x$. Furthermore, for any positive x,

$$\mathbb{P}\Big(W_n > x\left(1 + A_{3,n}(x/2)\ \right)^{1/3}\Big) \leq \exp\big(-x^2/2\big). \tag{2.5}$$

Remark 2.1. We will prove in Section 3 that $\gamma(v)$ is finite for any positive v. Now assume that, for any positive k, the random variables X_k have the variance v, for some

$v < 1$. Suppose that the sequence (c_k) does not belong to $\ell^2(\mathbb{N})$ and that, however, this sequence belongs to $\ell^3(\mathbb{N})$. Then $\lim_{n\to\infty} A_{3,n}\sqrt{V_n} = 0$, which shows that the corrective term $xA_{3,n}/n$ in (2.4) is much smaller than the term $cx/\sqrt{V_n}$ appearing in (1.17).

Remark 2.2. We will prove in Section 4 that $\gamma(v) \gg v$ for small values of v. Consequently, for constant values of c_i and small values of v_i, the quantity $A_{3,n}$ is larger than $V_n^{-1/2}$. In that case, (1.12) is more efficient than (2.4) for small values of x.

Proof of Theorem 2.1. We start by proving (2.2). From (1.4) and the independence of the random variables X_k, for any positive s,

$$\log \mathbb{E}\big(\exp(sV_n^{1/2}W_n)\big) \le \sum_{k=1}^{n} \ell_{v_k}(c_k s). \tag{2.6}$$

If $v_k \ge 1$, then, by (1.9), $\ell_{v_k}(c_k s) \le v_k c_k^2 (s^2/2)$. If $v_k < 1$, it follows from the definition of γ that

$$\ell_{v_k}(c_k s) \le v_k c_k^2 (s^2/2) + c_k^3 \gamma(v_k)(s^3/6). \tag{2.7}$$

Hence, for any positive s,

$$\log \mathbb{E}\big(\exp(sV_n^{1/2}W_n)\big) \le \frac{s^2}{2}\Big(\sum_{k=1}^{n} c_k^2 v_k\Big) + \frac{s^3}{6}\Big(\sum_{k=1}^{n} c_k^3 \gamma(v_k)\Big). \tag{2.8}$$

Now, setting $s = tV_n^{-1/2}$ in the above inequality, we get (2.2).

We now prove (2.3), (2.4) and (2.5). Let $h(t) = (t^2/2) + (t^3/6)$. The proofs are based on the calculation of the Legendre dual h^* of h and on some upper bound for the inverse function of h^*.

Lemma 2.2. *Let the function h be defined by $h(t) = (t^2/2) + (t^3/6)$ for any nonnegative t. Then, for any positive x,*

$$h^*(x) = \big((1+2x)^{3/2} - 1 - 3x\big)/3 \ge (1+x)\log(1+x) - x \ge x^2/(2+2x/3) \tag{2.9}$$

and

$$h^{*-1}(x) \le \sqrt{2x}\,\big(1+\sqrt{x/2}\,\big)^{1/3} \le \sqrt{2x} + x/3. \tag{2.10}$$

Proof of Lemma 2.2. Let t_x be the positive solution of the equation $h'(t) = x$. Then $t_x = \sqrt{1+2x} - 1$, whence

$$h^*(x) = xt_x - h(t_x) = \big((1+2x)^{3/2} - 1 - 3x\big)/3$$

after straightforward computations. Now $h''(x) = (1+2x)^{-1/2} \ge (1+x)^{-1} \ge (1+x/3)^{-3}$. Integrating two times these inequalities, we obtain the two lower bounds in (2.9).

We now prove the second part of Lemma 2.2. From the inversion formula for h^* given in Rio (2000, p. 159),

$$h^{*-1}(x) = \inf\{s^{-1}(h(s)+x) : s > 0\}. \tag{2.11}$$

Let $s_x = \sqrt{2x}\,\big(1+\sqrt{x/2}\,\big)^{-1/3}$. According to (2.11),

$$h^{*-1}(x) \le s_x^{-1}(h(s_x)+x). \tag{2.12}$$

Now set $u_x = \big(1+\sqrt{x/2}\,\big)^{1/3}$. Then $\sqrt{x/2} = u_x^3 - 1$, from which $s_x = 2u_x^{-1}(u_x^3-1)$ and

$$s_x^{-1}(h(s_x)+x) = u_x(u_x^3-1)\big(1 + u_x^{-2} + \tfrac{2}{3}(1-u_x^{-3})\big). \tag{2.13}$$

Now, applying the elementary inequality $2(1-a^3) \le 3(1-a^2)$, valid for $a \ge 0$, to the last term in the above equation, we get that

$$s_x^{-1}(h(s_x)+x) \le 2u_x(u_x^3-1) = \sqrt{2x}\,\big(1+\sqrt{x/2}\,\big)^{1/3} \le \sqrt{2x} + x/3, \tag{2.14}$$

which ends up the proof of Lemma 2.2. □

We now complete the proof of Theorem 2.1. Let the function h be defined as in Lemma 2.2 and define the functions h_A for $A > 0$ by $h_A(t) = A^{-2}h(At)$. From (2.2),

$$\log \mathbb{E}\big(\exp(tW_n)\big) \leq h_A(t), \quad \text{with } A = A_{3,n}. \tag{2.15}$$

Now, using Lemma 2.2, it is readily checked that

$$h_A^*(x) = A^{-2}h^*(Ax) = \frac{(1+2Ax)^{3/2} - 1 - 3Ax}{3A^2} \tag{2.16}$$

and

$$h_A^{*-1}(x^2/2) = A^{-1}h^{*-1}(A^2x^2/2) \leq x\left(1 + (Ax/2)\right)^{1/3}. \tag{2.17}$$

The three above facts imply (2.3) and (2.5). The upper bound (2.4) follows from (2.9). □

3 Upper bound on γ : large values of v

In this section, we give an upper bound on the function γ, which is exact if the variances v_k are in the interval $[2-\sqrt{3}, 1]$. We now state the main result of this section.

Proposition 3.1. *Define the function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ by*

$$\psi(v) = 0 \text{ if } v \geq 1,\ \psi(v) = v(1-v) \text{ if } v \in (2-\sqrt{3}, 1) \text{ and } \psi(v) = \frac{(1+v)^3}{6\sqrt{3}} \text{ if } v \leq 2-\sqrt{3}.$$

Then $\gamma(v) \leq \psi(v)$ for any positive v.

Remark 3.1. From the definition of $A_{3,n}$,

$$A_{3,n}\sqrt{V_n} = \Big(\sum_{k=1}^n c_k^2 v_k\Big)^{-1}\Big(\sum_{k=1}^n c_k^3\gamma(v_k)\Big) \leq \sup_{k\in[1,n]} \frac{c_k\gamma(v_k)}{v_k}.$$

Now, if $v \geq 0.145$, then $\psi(v) \leq v$, which ensures that $\gamma(v) \leq v$. Thus, if $v_k \geq 0.145$ for any k in $[1, n]$, then $A_{3,n} \leq cV_n^{-1/2}$, where $c = \max(c_1, c_2, \dots, c_n)$. Noting that $x \to x^{-2}g(x)$ is nonincreasing, one infers that

$$A_{3,n}^{-2}g\big(A_{3,n}x\big) \geq c^{-2}V_n g\big(V_n^{-1/2}cx\big),$$

provided that $v_k \geq 0.145$ for any k in $[1, n]$. Then Inequality (2.4) of Theorem 2.1 is more efficient than (1.12) for any choice of the weights c_k.

Remark 3.2. Let μ_v be the discrete distribution given by $\mu_v(\{1\}) = v/(1+v)$ and $\mu_v(\{-v\}) = 1/(1+v)$. If X_k has the distribution μ_{v_k} for any positive k, then

$$\log \mathbb{E}\big(\exp(tW_n)\big) = \sum_{k=1}^n \ell_{v_k}(c_k V_n^{-1/2}t) = \frac{t^2}{2} + \frac{t^3}{6V_n^{3/2}}\Big(\sum_{k=1}^n c_k^3 v_k(1-v_k)\Big) + \mathcal{O}(t^4).$$

Consequently, if v_k belongs to $[2-\sqrt{3}, 1]$ for any positive k, then the upper bound (2.2) is exactly the expansion at order three of the logarithm of the Laplace transform of W_n. In that case, Theorem 2.1 provides the optimal second order term.

Proof of Proposition 3.1. If $v \geq 1$, then, by (1.9), $\ell_v(t) \leq v(t^2/2)$, which implies the result. We now prove the proposition in the case $v < 1$. Let $a(t) = ve^{(1+v)t}$. With this notation,

$$\ell_v'(t) = \frac{a(t)-v}{a(t)+1},\ \ell_v''(t) = \frac{(1+v)^2 a(t)}{(a(t)+1)^2},\ \ell_v^{(3)}(t) = \frac{(1+v)^3 a(t)(1-a(t))}{(a(t)+1)^3} \tag{3.1}$$

and

$$\ell_v^{(4)}(t) = \frac{(1+v)^4 a(t)(1-4a(t)) + (a(t))^2)}{(a(t)+1)^4}. \tag{3.2}$$

From (3.1), if $a(t) \geq 1$, which is equivalent to $t \geq t_0 := \log(1/v)/(1+v)$, then $\ell_v^{(3)}(t) \leq 0$. Recall that $a(t) \geq v$. Hence, in order to find the maximum of $\ell_v^{(3)}$, it is enough to study the sign of $\ell_v^{(4)}$ for $a(t)$ in $[v, 1]$. Now

$$1 - 4a(t) + (a(t))^2 = (a(t) - 2 + \sqrt{3})(a(t) - 2 - \sqrt{3}). \tag{3.3}$$

If $v \geq 2 - \sqrt{3}$, then $\ell_v^{(4)}(t) \leq 0$ for t in $[0, t_0]$. In that case

$$\ell_v^{(3)}(t) \leq \ell_v^{(3)}(0) = v(1-v) \text{ for any } t \in [0, t_0]. \tag{3.4}$$

Since $\ell_v^{(3)}(t) \leq 0$ for $t \geq t_0$, it implies that $\ell_v^{(3)}(t) \leq v(1-v)$ for any $t \geq 0$. Integrating three times this inequality, we then get Proposition 3.1 in the case $v \geq 2 - \sqrt{3}$.

If $v < 2 - \sqrt{3}$, then the equation $\ell_v^{(4)}(t) = 0$ has an unique solution t_1 in $[0, t_0]$. More precisely $t_1 = (\log(2-\sqrt{3}) - \log v)/(1+v)$ and $a(t_1) = 2 - \sqrt{3}$. In that case $\ell_v^{(3)}$ takes its maximum at point t_1. Since $\ell_v^{(3)}(t_1) = (1+v)^3/(6\sqrt{3})$, integrating three times this inequality, we get Proposition 3.1 in the case $v < 2 - \sqrt{3}$, which completes the proof. □

4 Upper bound on γ : small values of v

From Proposition 3.1, we know that $\gamma(v) \leq \psi(v)$. Nevertheless the function ψ does not decrease to 0 as v tends to 0. Hence it seems clear that the bounds of Section 3 can be improved for small values of v. This is done in the proposition below, which gives the exact order of magnitude of γ.

Proposition 4.1. *Let the function g_v be defined, for positive values of t, by*

$$g_v(t) = t^{-3}(\ell_v(t) - vt^2/2). \tag{4.1}$$

Then, for any $v \leq 1/25$,

$$\gamma(v) \leq \beta(v), \quad \textit{where} \quad \beta(v) = 12 g_v\Big(\frac{2|\log v|}{1+v}\Big) = \frac{3(1+v)(1-v^2+2v\log v)}{2\,|\log v\,|^2}$$

and $\gamma(v) \geq \frac{1}{2}\beta(v)$ for any $v < 1$.

Proof. From the definitions of γ and g_v,

$$\gamma(v) = 6 \sup_{t>0} g_v(t). \tag{4.2}$$

Let $t_0 = \log(1/v)/(1+v)$. Then $\frac{1}{2}\beta(v) = 6g_v(2t_0)$, which implies the second part of Proposition 4.1. The main tool for proving the first part of Proposition 4.1 is the lemma below, which will allow us to localize the maximum of g_v and to bound up the maximal value.

Lemma 4.2. *Let the function f_v be defined by $f_v(t) = tg_v(t)$. Suppose that $v \leq 1/25$. Let $t_0 = \log(1/v)/(1+v)$. Then f_v reaches its maximum at point $2t_0$ and g_v reaches its global maximum at some point t_c in the interval $(t_0, 2t_0)$.*

Proof of Lemma 4.2. The first assertion is due to Kearns and Saul (1998). Below we give a proof for the sake of completeness (see Berend and Kontorovich (2013) for an other proof). By definition of f_v, $\lim_{t\downarrow 0} f_v(t) = 0$ and

$$t^3 f_v'(t) = t\ell_v'(t) - 2\ell_v(t) := \eta(t).$$

In order to study the sign of η, we note that $\eta(0) = \eta'(0) = 0$. Next $\eta''(t) = t\ell_v^{(3)}(t)$. Hence, from (2.6), $\eta''(t) > 0$ for t in $(0, t_0)$ and $\eta''(t) < 0$ for $t > t_0$, which means that η is convex on $[0, t_0]$ and concave on $[t_0, +\infty)$. Since $\eta(0) = \eta'(0) = 0$, η is increasing and convex on $(0, t_0)$. Since η is concave on $[t_0, +\infty)$, it follows that η has at most one zero on (t_0, ∞) and that this zero is the unique maximum of f_v. Now, noting that

$$\ell_v(t) = \log(1 + a(t)) - \log(1 + v) - vt \tag{4.3}$$

and $a(2t_0) = 1/v$, we get that

$$\eta(2t_0) = 2t_0(1 - v) - 2\log(1/v) + 4vt_0 = 2t_0(1 + v) - 2\log(1/v) = 0.$$

Consequently f_v has an unique maximum at point $2t_0$, which proves the first assertion. Furthermore f_v is increasing on $(0, 2t_0)$ and decreasing on $(2t_0, \infty)$.

We now prove the second assertion. By definition of g_v, $\lim_{t\downarrow 0} g_v(t) = v(1 - v)/6$ and

$$t^4 g_v'(t) = t\ell_v'(t) - 3\ell_v(t) + v(t^2/2) := \delta(t).$$

In order to study the sign of δ, we note that $\delta(0) = \delta'(0) = \delta''(0) = 0$. Next

$$\delta^{(3)}(t) = t\ell_v^{(4)}(t) = \frac{(1 + v)^4 t a(t)(a(t) - 2 + \sqrt{3})(a(t) - 2 - \sqrt{3})}{(a(t) + 1)^4}$$

by (3.2) and (3.3). Now, let t_1 and t_2 by defined respectively by $a(t_1) = 2 - \sqrt{3}$ and $a(t_2) = 2 + \sqrt{3}$. Then $t_1 < t_0 < t_2$ and, from the above equation $\delta^{(3)}$ is positive on $(0, t_1)$, negative on (t_1, t_2) and positive on (t_2, ∞). Since $\delta(0) = \delta'(0) = \delta''(0) = 0$, it implies that δ is increasing and convex on $[0, t_1]$. Furthermore δ'' has at most two zeros, which implies that δ has at most two zeros. Now, recall that $\eta(2t_0) = 0$. Therefore

$$\delta(2t_0) = \eta(2t_0) - \ell_v(2t_0) + 2vt_0^2 = -\ell_v(2t_0) + 2vt_0^2.$$

Now ℓ_v'' is increasing on $[0, t_0]$, decreasing on $[t_0, 2t_0]$ and $\ell_v''(0) = \ell_v''(2t_0) = v$. Consequently $\ell_v(2t_0) > 2vt_0^2$. Thus $\delta(2t_0) < 0$. It follows that δ has an unique zero t_c in $(0, 2t_0)$. If $\delta(t_0) > 0$, then t_c belongs to $(t_0, 2t_0)$ and is the maximum of g_v on $[0, 2t_0]$.

$$\delta(t_0) = \tfrac{1}{2}t_0(1 - v) + 3\log(1 + v) - 3\log 2 + 3vt_0 + \tfrac{1}{2}vt_0^2.$$

If $(1/v) \geq 25$, then $t_0 \geq 3.08$ and $\log(1 + v) \geq 0.98v$, whence

$$\delta(t_0) \geq 0.50\, t_0 + 4.04\, vt_0 + 2.94\, v - 2.08. \tag{4.4}$$

Now $t_0 \geq (1 - v)\log(1/v) \geq 0.96\log(1/v)$. Therefore

$$\begin{aligned}\delta(t_0) &\geq 0.50(1 - v)\log(1/v) + 3.87v\log(1/v) + 2.94v - 2.08\\ &\geq 0.50\log(1/v) + 3.37v\log(1/v) + 2.94v - 2.08.\end{aligned}$$

Now, $\log(1/v) \geq \log 25 \geq 3.21$, whence

$$\delta(t_0) \geq 0.50\log(1/v) + 13.75v - 2.08 \tag{4.5}$$

The above lower bound takes its minimum at $v = 2/55$ and the value of this minimum is strictly positive. Hence $\delta(t_0) > 0$ for any v in $(0, 1/25]$, which proves that δ has an unique zero t_c in $(t_0, 2t_0)$. Moreover t_c is the unique maximum of g_v on $[0, 2t_0]$.

It remains to prove that $g_v(t) < g_v(t_c)$ for $t \geq 2t_0$. Recall that $f_v = tg_v$ is positive at point $2t_0$ and decreasing on $[2t_0, \infty)$. Furthermore $\lim_{t\uparrow\infty} f_v(t) = -v/2$. Hence f_v has an unique zero t_3 on $(2t_0, \infty)$. Clearly f_v is nonnegative on $[2t_0, t_3]$ and negative on (t_3, ∞). If $t > t_3$ then $g_v(t) < 0 < g_v(t_c)$. Now, if t belongs to $[2t_0, t_3]$, then $g_v = t^{-1} f_v$ is the product of two decreasing nonnegative functions, whence $g_v(t) \leq g_v(2t_0) < g_v(t_c)$, which ends up the proof of Lemma 4.2. □

We now complete the proof of Proposition 4.1. From the proof of the first part of Lemma 4.2, we know that f_v is increasing on $[t_0, 2t_0]$. Hence $f_v(t_c) \le f_v(2t_0)$, which is equivalent to $t_c g_v(t_c) \le 2t_0 g_v(2t_0)$. It follows that $g_v(t_c) \le 2t_0 t_c^{-1} g_v(2t_0) \le 2g_v(2t_0)$. Finally

$$g_v(2t_0) = \frac{(1+v)(1-v^2+2v\log v)}{8(\log v)^2},$$

which completes the proof of Proposition 4.1. □

Remark 4.1. Note that $\beta(v) \sim (3/2)|\log v|^{-2}$ as v tends to 0. Hence the functions β and γ decrease slowly to 0 as v tends to 0. Therefore the function β is less than the function ψ only for small values of v. Indeed one can prove that $\psi(v) \le \beta(v)$ for $v \ge v_0 = 3.62 \times 10^{-2}$ and $\beta(v) \le \psi(v)$ for $v \le v_0$. Since $v_0 < 1/25$, Propositions 3.1 and 4.1 ensure that

$$\tfrac{1}{2}\beta(v) \le \gamma(v) \le (\beta \wedge \psi)(v) \le \beta(v) \quad \text{for any positive } v, \tag{4.6}$$

with the convention that $\beta(v) = \psi(v) = 0$ if $v \ge 1$. Thus the results of Theorem 2.1 hold true with $B_{3,n}$ instead of $A_{3,n}$, where

$$B_{3,n} = V_n^{-3/2} \sum_{k=1}^{n} c_k^3 \min(\psi(v_k), \beta(v_k)).$$

Remark 4.2. Suppose that $c_1 = 1$ and $c_k \le 1$ for any positive k. Then (1.10) yields

$$-V_n^{-1} \log \mathbb{P}(W_n \ge \sqrt{V_n}\, x) \ge (n/V_n)\ell^*_{V_n/n}(xV_n/n).$$

Assume furthermore that $\sum_k c_k^2 v_k = +\infty$ and $\lim_k c_k^2 v_k = 0$. Then $\lim_n V_n = \infty$ and $\lim_n (V_n/n) = 0$, which implies that the above lower bound converges to $g(x)$. In that case, either (1.10) or (1.11) yield the asymptotic result

$$-\limsup_{n\to\infty} V_n^{-1} \log \mathbb{P}(W_n \ge \sqrt{V_n}\, x) \ge g(x) = (1+x)\log(1+x) - x. \tag{4.7}$$

Now, let

$$R_n = A_{3,n}\sqrt{V_n} = \Big(\sum_{k=1}^{n} c_k^2 v_k\Big)^{-1} \Big(\sum_{k=1}^{n} c_k^3 \gamma(v_k)\Big).$$

Inequality (2.3) of Theorem 2.1 yields

$$-V_n^{-1} \log \mathbb{P}(W_n \ge \sqrt{V_n}\, x) \ge x^2 r(R_n x), \quad \text{where} \quad r(y) = \frac{(1+2y)^{3/2} - 1 - 3y}{3y^2}. \tag{4.8}$$

Note that r is decreasing and that $r(0) := \lim_{y \downarrow 0} r(y) = (1/2)$. Suppose now that $\lim_{k\to\infty}(c_k \gamma(v_k)/v_k) = 0$. From (4.6), this condition is equivalent to the condition $\lim_{k\to\infty}(c_k \beta(v_k)/v_k) = 0$. Then the Toeplitz lemma ensures that $\lim_n R_n = 0$ and (4.8) yields

$$-\limsup_{n\to\infty} V_n^{-1} \log \mathbb{P}(W_n \ge \sqrt{V_n}\, x) \ge x^2 r(0) = x^2/2 > g(x), \tag{4.9}$$

which improves on (4.7). For example, if $v_k = v < 1$ for any positive k, $\sum_k c_k^2 = \infty$ and $\lim_k c_k = 0$, then $\lim_n R_n = 0$ and (4.9) holds true.

Remark 4.3. If $R_n \le a$, then $r(R_n x) \ge R(ax)$ and consequently

$$-V_n^{-1} \log \mathbb{P}(W_n \ge \sqrt{V_n}\, x) \ge x^2 r(ax) > a^{-2} g(ax).$$

When $a \le 1$, $a^{-2} g(ax) \ge g(x)$. In that case, Inequality (2.3) is more efficient than the Bennett Inequality.

5 An example

Throughout this section, we compare our results with the previous results on some example. We consider a triangular array $(X_{k,m})$ of independent centered random variables such that

$$\operatorname{Var} X_{k,m} = 1/(m+1) \text{ and } X_{k,m} \le 1 \text{ a.s. for any } (k,m) \in \mathbb{N}^* \times \mathbb{N}^*. \tag{5.1}$$

Define the centered and normalized random variables T_m by

$$T_m = X_{1,m} + \frac{1}{m} \sum_{k=2}^{m^3+1} X_{k,m} = \sum_{k=1}^{n} \xi_{k,m}, \tag{5.2}$$

where $n = m^3 + 1$, $\xi_{1,m} = X_{1,m}$ and $\xi_{k,m} = m^{-1} X_{k,m}$ for $k \ge 2$.

Inequality (2.3) of Theorem 2.1 and (4.6) yield

$$\mathbb{P}(T_m \ge x) \le \exp\big(-b_m^{-2} h^*(b_m x)\big), \text{ where } b_m = 2(\beta \wedge \psi)(1/(m+1)) \tag{5.3}$$

and h^* is the dual function given in (2.9). The Kearns-Saul type inequality (1.15) yields

$$\mathbb{P}(T_m \ge x) \le \exp\Big(-\frac{x^2(m+1)\log(m+1)}{m(m+2)}\Big) \tag{5.4}$$

and Hoeffding's inequality - *Inequality (1.10)* - gives

$$\mathbb{P}(T_m \ge x) \le \exp\big(-n\ell^*_{(1/n)}(x/n)\big). \tag{5.5}$$

Now, let Q denote the tail function of a standard normal random variable. If furthermore $|X_{k,m}| \le 1$ a.s. for any (k,m), the Berry-Esseen type estimates of Shevtsova (2013) yield

$$\mathbb{P}(T_m \ge x) \le Q(x) + \Delta_n \text{ with } \Delta_n = 0,3057\,(L_n + \tau_n), \tag{5.6}$$

$L_n = \sum_{k=1}^{n} \mathbb{E}\big(|\xi_{k,m}|^3\big) \le 2(m+1)^{-1}$ and $\tau_n = \sum_{k=1}^{n} \big(\operatorname{Var} \xi_{k,m}\big)^{3/2} \le 2(m+1)^{-3/2}$.

Below I give the numerical values of the above upper bounds for $x = x_0 = 3$ and $m = 2$, $m = 3$, $m = 4$, $m = 8$, $m = 24$, $m = 99$, $m = 9999$.

Ineq.	m=2	m = 3	m=4	m=8	m=24	m=99	m=9999
(5.3)	0,0330	0,0297	0,0276	0,0242	0,0219	0,0176	0,0129
(5.4)	0,0245	0,0359	0,0489	0,1081	0,3133	0,6607	0,9917
(5.5)	0,0607	0,0729	0,0761	0,0782	0,0785	0,0785	0,0785
(5.6)	0,3228	0,2306	0,1783	0,0919	0,0307	0,0081	0,00141

One can see here that Inequality (5.3) is more efficient than (5.4), (5.5) and (5.6) for m in $[3, 24]$, which corresponds to n in $[28, 13825]$. For $m = 2$, the Kearns-Saul inequality is more efficient. For large values of m, the Berry-Esseen type estimates provide better results.

Concerning the asymptotic behavior of these inequalities as m tends to ∞,

$$\lim_m \exp\big(-b_m^{-2} h^*(b_m x_0)\big) = e^{-x_0^2/2} = 0,0111, \ \lim_m \exp\Big(-\frac{x_0^2(m+1)\log(m+1)}{m(m+2)}\Big) = 1$$

and

$$\lim_m e^{-n\ell^*_{(1/n)}(x_0/n)} = e^{-g(x_0)} = 0,0785, \ \lim_m (Q(x_0) + \Delta_n) = Q(x_0) = 0,00135.$$

References

[1] Bennett, G. Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57**, 33-45 (1962).

[2] Bentkus, V. A remark on the inequalities of Bernstein, Prokhorov, Bennett, Hoeffding, and Talagrand. *Lithuanian Math. J.* **42**, no. 3, 262-269 (2002). MR-1947624

[3] Bentkus, V. An inequality for tail probabilities of martingales with differences bounded from one side. *J. Theoret. Probab.* **16**, no. 2, 161-173 (2003). MR-1956826

[4] Bentkus, V. On Hoeffding's inequalities. *Ann. Probab.* **32**, no. 2, 1650-1673 (2004). MR-2060313

[5] Bercu, B., Delyon, B. and Rio, E. Concentration inequalities for sums and martingales. SpringerBriefs (in preparation). 2015

[6] Berend, D. and Kontorovich, A. On the concentration of the missing mass. *Electron. Commun. Probab.* **18**, no. 3, 1-7 (2013). MR-3011530

[7] Fan, X., Grama, I. and Liu, Q. Exponential inequalities for martingales with applications *Electron. J. Probab.* **20**, no. 1, 1-22 (2015). MR-3311214

[8] Hoeffding, W. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13-30 (1963). MR-0144363

[9] Kearns, M. J. and Saul, L. K. Large deviation methods for approximate probabilistic inference. Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, San-Francisco, 311-319 (1998)

[10] Pinelis, I. On the Bennett-Hoeffding inequality. *Ann. Inst. Henri Poincaré Probab. Stat.* **50**, 15-27 (2014). MR-3161520

[11] Rio, E. Théorie asymptotique des processus aléatoires faiblement dépendants (French). *Mathématiques & Applications* **31**. Springer (2000) MR-2117923

[12] Shevtsova, I. G. On the absolute constants in the Berry-Esseen inequality and its structural and nonuniform improvements. *Inform. Primen.* **7**, no. 1, 124-125. http://mi.mathnet.ru/ia252 (2013)