To appear in *Statistics: A Journal of Theoretical and Applied Statistics*
Vol. 00, No. 00, Month 20XX, 1–15

Martingale inequalities of type Dzhaparidze and van Zanten

X. Fan[a,b,*], I. Grama[c] and Q. Liu[c]

[a] *Center for Applied Mathematics, Tianjin University, 300072 Tianjin, China;* [b] *Regularity Team, Inria, France;* [c] *Université de Bretagne-Sud, UMR 6205, LMBA, Vannes, France.*

(*v5.0 released February 2015*)

Freedman's inequality is a supermartingale counterpart to Bennett's inequality. This result shows that the tail probabilities of a supermartingale is controlled by the quadratic characteristic and a uniform upper bound for the supermartingale difference sequence. Replacing the quadratic characteristic by $\mathrm{H}_k^y := \sum_{i=1}^k \big(\mathbf{E}(\xi_i^2|\mathcal{F}_{i-1}) + \xi_i^2 \mathbf{1}_{\{|\xi_i|>y\}}\big)$, Dzhaparidze and van Zanten (*Stochastic Process. Appl.*, 2001) have extended Freedman's inequality to martingales with unbounded differences. In this paper, we prove that H_k^y can be refined to $\mathrm{G}_k^y := \sum_{i=1}^k \big(\mathbf{E}(\xi_i^2 \mathbf{1}_{\{\xi_i \le y\}}|\mathcal{F}_{i-1}) + \xi_i^2 \mathbf{1}_{\{\xi_i>y\}}\big)$. Moreover, we also establish two inequalities of type Dzhaparidze and van Zanten. These results extend Sason's inequality (*Statist. Probab. Lett.*, 2012) to the martingales with possibly unbounded differences and establish the connection between Sason's inequality and De la Peña's inequality (*Ann. Probab.*, 1999). An application to self-normalized deviations is given.

Keywords: Freedman's inequality; De la Peña's inequality; exponential inequalities; tail probabilities; martingales; self-normalized deviations

AMS Subject Classification: 60G42; 60E15

1. Introduction

Exponential inequalities for tail probabilities of sums of independent real-valued random variables and their extension to martingales have numerous important applications in probability theory and statistics. See, for instance, De la Peña and Pang [4], Bercu and Touati [2] and [8]. The classical Bennett inequality [1] gives a tail bound for sums of independent random variables with a bounded range. If $(\xi_i)_{i\geq 1}$ are zero-mean independent random variables, all bounded by some constant a so that $|\xi_i| \leq a$ for all i, then the sum $S_n = \sum_{i=1}^n \xi_i$ obeys the following Bennett inequality (see also Bernstein [3]): for all $x > 0$,

$$\mathbf{P}\Big(S_n \geq x\Big) \leq B_1(x,a,v) := \left(\frac{v^2}{xa+v^2}\right)^{\frac{x}{a}+\frac{v^2}{a^2}} e^{\frac{x}{a}} \quad (1)$$

$$\leq B_2(x,a,v) := \exp\left\{-\frac{x^2}{2(v^2+xa/3)}\right\}, \quad (2)$$

where v^2 is the variance of S_n.

*Corresponding author. Email: fanxiequan@hotmail.com

Freedman have extended Bennett's result to the case of discrete-time supermartingales with bounded jumps. Let $(\xi_i, \mathcal{F}_i)_{i=1,...,n}$ be a sequence of supermartingale differences. Denote by $S_k = \sum_{i=1}^k \xi_i$ and $\langle S\rangle_k = \sum_{i=1}^k \mathbf{E}(\xi_i^2|\mathcal{F}_{i-1})$. The well-known Freedman's inequality [9] for supermartingales states that: if $\xi_i \leq a$ for a positive constant a, then, for all $x, v > 0$,

$$\mathbf{P}\Big(S_k \geq x \text{ and } \langle S\rangle_k \leq v^2 \text{ for some } k \in [1,n]\Big) \leq B_1(x,a,v) \tag{3}$$

$$\leq B_2(x,a,v). \tag{4}$$

In particular, when $(\xi_i)_{i=1,...,n}$ are independent, the bounds (3) and (4) reduce to the bounds of Bennett [1] and Bernstein [3], respectively. Moreover, Freedman's inequality implies that the bounds (3) and (4) not only hold for S_n but even hold for the maximum of partial sums $\max_{1\leq k\leq n} S_k$.

Replacing the quadratic characteristic $\langle S\rangle_k$ by

$$\mathrm{H}_k^y := \sum_{i=1}^k \Big(\mathbf{E}(\xi_i^2|\mathcal{F}_{i-1}) + \xi_i^2 \mathbf{1}_{\{|\xi_i|>y\}}\Big),$$

Dzhaparidze and van Zanten [7] have established a generalization of Freedman's inequality for martingales with unbounded differences: for all $x, y, v > 0$,

$$\mathbf{P}\Big(S_k \geq x \text{ and } \mathrm{H}_k^y \leq v^2 \text{ for some } k \in [1,n]\Big) \leq B_1(x,y,v). \tag{5}$$

In particular, if $|\xi_i| \leq a$ for all i, it holds $\mathrm{H}_k^a = \langle S\rangle_k$, and then the inequality of Dzhaparidze and van Zanten (5) reduces to Freedman's inequality (3), implying inequality (4).

However, if (ξ_i) are not bounded from below, inequality (5) does not imply Freedman's inequality (3) in general. To fill this gap, we propose replacing the random variable H_k^y in inequality (5) by a smaller one G_k^y, where

$$\mathrm{G}_k^y = \sum_{i=1}^k \Big(\mathbf{E}(\xi_i^2 \mathbf{1}_{\{\xi_i \leq y\}}|\mathcal{F}_{i-1}) + \xi_i^2 \mathbf{1}_{\{\xi_i>y\}}\Big). \tag{6}$$

Our Theorem 2.1 states that, for all $x, y \geq 0$ and $v > 0$,

$$\mathbf{P}\left(S_k \geq x \text{ and } \mathrm{G}_k^y \leq v^2 \text{ for some } k \in [1,n]\right) \leq B_1(x,y,v). \tag{7}$$

Since $\mathrm{G}_k^y \leq \mathrm{H}_k^y$, inequality (7) implies the inequality of Dzhaparidze and van Zanten (5). Moreover, if $\xi_i \leq a$ for all $i \in [1,n]$ (may not be bounded from below), it holds $\mathrm{G}_k^a = \langle S\rangle_k$ for all $k \in [1,n]$, and then (7) also implies Freedman's inequality (3). This fills the gap.

In Theorem 2.2, we give a generalization of (5) to the supermartingales with non-square-integrable differences. Write

$$\mathrm{G}_n(\beta) = \sum_{i=1}^n \Big(\mathbf{E}\big(|\xi_i|^\beta|\mathcal{F}_{i-1}\big) + |\xi_i|^\beta\Big)$$

for a constant $\beta \in (1, 2)$. Then, for all $x, v > 0$,

$$\mathbf{P}\left(\max_{1\leq k\leq n} S_k \geq x \text{ and } \mathrm{G}_n(\beta) \leq v^\beta\right) \leq \exp\left\{-C(\beta)\left(\frac{x}{v}\right)^{\frac{\beta}{\beta-1}}\right\}, \tag{8}$$

where $C(\beta) = \beta^{\frac{1}{1-\beta}}\left(1-\beta^{-1}\right)$.

Under the additional assumption of conditional symmetry, Sason [12] gave an improvements of Freedman's inequality (4). Sason proved that if $(\xi_i, \mathcal{F}_i)_{i\geq 1}$ is a sequence of conditionally symmetric martingale differences with $|\xi_i| \leq a$ for a positive constant a, then, for all $x, v > 0$,

$$\begin{aligned} &\mathbf{P}\left(S_k \geq x \text{ and } \langle S\rangle_k \leq v^2 \text{ for some } k \in [1, n]\right) \\ &\qquad \leq\ B_0(x, a, v) := \exp\left\{-\lambda x + \left(\frac{\cosh(\lambda a) - 1}{a^2}\right)v^2\right\} \end{aligned} \tag{9}$$

and $B_0(x, a, v) \leq B_1(x, a, v)$, where

$$\lambda = \frac{1}{a}\log\left(\sqrt{1+\frac{x^2a^2}{v^4}} + \frac{xa}{v^2}\right).$$

In the spirit of Dzhaparide and van Zanten [7], we establish the following generalization of Sason's inequality (9). Define

$$\mathrm{M}_k^y = \sum_{i=1}^{k}\Big(\mathbf{E}(\xi_i^2\mathbf{1}_{\{|\xi_i|\leq y\}}|\mathcal{F}_{i-1}) + \xi_i^2\mathbf{1}_{\{|\xi_i|>y\}}\Big). \tag{10}$$

Then, for all $x, v > 0$ and all $y \geq 0$,

$$\mathbf{P}\left(S_k \geq x \text{ and } \mathrm{M}_k^y \leq v^2 \text{ for some } k \in [1, n]\right)\ \leq\ B_0(x, y, v), \tag{11}$$

where, by convention,

$$B_0(x, 0, v) = \lim_{y\to 0+} B_0(x, y, v) = \exp\Big\{-\frac{x^2}{2\,v^2}\Big\}$$

applied when $y = 0$. If $|\xi_i| \leq a$ for all i, then $\mathrm{M}_k^a = \langle S\rangle_k$ for all k and (11) reduces to Sason's inequality (9). Notice that when $y = 0$, inequality (11) is known as De la Peña's inequality [5]. Hence, our bound establishes a connection between the inequalities of De la Peña and Sason.

The paper is organized as follows. We present our main results in Section 2 and the application to self-normalized deviations in Section 3, and devote to the proofs of the main results in Sections 4 - 6.

2. Main results

Assume that we are given a sequence of real-valued supermartingale differences $(\xi_i, \mathcal{F}_i)_{i=0,\ldots,n}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$, where $\xi_0 = 0$ and $\{\emptyset, \Omega\} =$

$\mathcal{F}_0 \subseteq ... \subseteq \mathcal{F}_n \subseteq \mathcal{F}$ are increasing σ-fields. So we have $\mathbf{E}(\xi_i|\mathcal{F}_{i-1}) \leq 0,\ i = 1, ..., n$, by definition. Set

$$S_k = \sum_{i=1}^{k} \xi_i, \quad k = 1, ..., n. \tag{12}$$

Then $S = (S_k, \mathcal{F}_k)_{k=1,...,n}$ is a supermartingale. Let $\langle S \rangle$ be the quadratic characteristic and $[S]$ be the squared variation of the supermartingale S :

$$\langle S \rangle_k = \sum_{i=1}^{k} \mathbf{E}(\xi_i^2|\mathcal{F}_{i-1}) \qquad \text{and} \qquad [S]_k = \sum_{i=1}^{k} \xi_i^2. \tag{13}$$

The following theorem strengthens the inequality of Dzhaparide and van Zanten [7].

Theorem 2.1 *Assume* $\mathbf{E}\xi_i^2 < \infty$ *for all* $i \in [1, n]$. *Then, for all* $x, y \geq 0$ *and* $v > 0$,

$$\begin{aligned} \mathbf{P}\left(S_k \geq x \text{ and } G_k^y \leq v^2 \text{ for some } k \in [1, n]\right) &\leq B_1(x, y, v) & (14)\\ &\leq B_2(x, y, v). & (15) \end{aligned}$$

where G_k^y *is defined by (6) and*

$$B_1(x, 0, v) = \lim_{y \to 0+} B_1(x, y, v) = B_2(x, 0, v)$$

applied when $y = 0$.

Since $\mathrm{G}_k^0 \leq \langle S \rangle_n + [S]_n$ for all $k \in [1, n]$, inequality (15) implies the following result: for all $x, v > 0$,

$$\mathbf{P}\left(\max_{1 \leq k \leq n} S_k \geq x \text{ and } \langle S \rangle_n + [S]_n \leq v^2\right) \leq B_1(x, 0, v). \tag{16}$$

This result slightly refines an earlier inequality of Bercu and Touati [2], where they have obtained the same bound on tail probabilities $\mathbf{P}\left(S_n \geq x \text{ and } \langle S \rangle_n + [S]_n \leq v^2\right)$. Thus the sum S_n has been strengthened to the maximum of partial sums $\max_{1 \leq k \leq n} S_k$. A similar refinement is applied to Delyon's inequality [6], where he has established the following result, for all $x, v > 0$,

$$\mathbf{P}\left(S_n \geq x \text{ and } \mathrm{G}_n^0 \leq v^2\right) \leq B_2(x, 0, v). \tag{17}$$

Consider the supermartingales with non-square-integrable differences. We have the following large deviation exponential bound, which can be regarded as a generalization of Delyon's inequality (17) or the inequality of Dzhaparide and van Zanten (5).

Denote by $x^+ = \max\{x, 0\}$ and $x^- = -\min\{x, 0\}$ the positive and negative parts of x, respectively.

Theorem 2.2 *Assume* $\mathbf{E}|\xi_i|^\beta < \infty$ *for a constant* $\beta \in (1, 2)$ *and all* $i \in [1, n]$. *Write*

$$G_k^0(\beta) = \sum_{i=1}^{k} \Big(\mathbf{E}\big((\xi_i^-)^\beta|\mathcal{F}_{i-1}\big) + (\xi_i^+)^\beta\Big), \qquad k \in [1, n].$$

Then, for all $x, v > 0$,

$$\mathbf{P}\Big(S_k \geq x \text{ and } G_k^0(\beta) \leq v^\beta \text{ for some } k \in [1, n]\Big) \leq \exp\left\{-C(\beta)\left(\frac{x}{v}\right)^{\frac{\beta}{\beta-1}}\right\}, \tag{18}$$

where

$$C(\beta) = \beta^{\frac{1}{1-\beta}}\left(1 - \beta^{-1}\right). \tag{19}$$

In particular, if $||\mathrm{G}_n(\beta)||_\infty = O(n)$ as $n \to \infty$, then, for any $x > 0$,

$$\mathbf{P}\left(\max_{1\leq k\leq n} S_k \geq nx\right) = O\Big(\exp\Big\{-n\, C_x(\beta)\Big\}\Big), \tag{20}$$

where $C_x(\beta) > 0$ does not depend on n.

It is also interesting to see that when β decreases to 1 in (18), the power $\frac{\beta}{\beta-1}$ is increasing to infinity and the corresponding constant $C(\beta)$ is decreasing to 0. This means the larger the power, the smaller the corresponding constant.

One calls $(\xi_i, \mathcal{F}_i)_{i\geq1}$ a sequence of *conditionally symmetric* martingale differences, if $\mathbf{E}(\xi_i > y|\mathcal{F}_{i-1}) = \mathbf{E}(\xi_i < -y|\mathcal{F}_{i-1})$ for all i and any $y \geq 0$. Motivated by the result of Dzhaparide and van Zanten [7], we give a generalization of Sason's inequality (9) to the martingales with unbounded differences.

THEOREM 2.3 *Assume that $(\xi_i, \mathcal{F}_i)_{i\geq1}$ is a sequence of conditionally symmetric martingale differences. Then, for all $x, v > 0$ and all $y \geq 0$,*

$$\mathbf{P}\left(S_k \geq x \text{ and } M_k^y \leq v^2 \text{ for some } k \in [1, n]\right) \leq B_0(x, y, v), \tag{21}$$

where M_k^y is defined by (10).

If the martingale differences are bounded $|\xi_i| \leq a$ for a positive constant a, then $\mathrm{M}_k^a = \langle S\rangle_k$ and inequality (21) with $y = a$ reduces to Sason's inequality (9). As pointed out by Sason [12], inequality (21) is the best possible that can be obtained from Chernoff's inequality $\mathbf{P}\left(S_n \geq x\right) \leq \inf_{\lambda\geq0} \mathbf{E}e^{\lambda(S_n - x)}$ under the present assumption in a certain sense. Indeed, if $(\xi_i)_{i\geq1}$ are i.i.d. random variables and satisfy the following distribution

$$\mathbf{P}(\xi_i = y) = \mathbf{P}(\xi_i = -y) = \frac{v^2}{2ny^2} \quad \text{and} \quad \mathbf{P}(\xi_i = 0) = 1 - \frac{v^2}{ny^2}, \tag{22}$$

then the bound (21) equals to $\lim_{n\to\infty} \inf_{\lambda\geq0} \mathbf{E}e^{\lambda(S_n - x)}$.

Since $\lim_{y\to0} \lambda = \frac{x}{v^2}$ and $\mathrm{G}_k^0 = [S]_k$, inequality (21) reduces to De la Peña's inequality [5] as $y \to 0$: for all $x, v > 0$,

$$\mathbf{P}\left(S_k \geq x \text{ and } [S]_k \leq v^2 \text{ for some } k \in \mathbb{N}\right) \leq B_0(x, 0, v). \tag{23}$$

Thus inequality (21) establishes the connection between the inequalities of De la Peña and Sason.

3. Application to self-normalized deviations

As an application of Theorem 2.2, consider the self-normalized deviations for independent random variables.

THEOREM 3.1 *Assume that* $(\xi_i)_{i=1,...,n}$ *is a sequence of independent and symmetric random variables. Denote by*

$$V_n(\beta) = \Big(\sum_{i=1}^n |\xi_i|^\beta \Big)^{1/\beta}$$

for a constant $\beta \in (1, 2]$. *Then, for all* $x > 0$,

$$\mathbf{P}\left(\max_{1 \le k \le n} S_k / V_n(\beta) \ge x \right) \le \exp\left\{ - \widetilde{C}(\beta) x^{\frac{\beta}{\beta-1}} \right\}, \tag{24}$$

where

$$\widetilde{C}(\beta) = \left(\frac{\beta}{2}\right)^{\frac{1}{1-\beta}} \left(1 - \beta^{-1}\right). \tag{25}$$

Proof of Theorem 3.1. Assume that $(\xi_i)_{i=1,...,n}$ are independent and symmetric. Set

$$\mathcal{F}_i = \sigma \left\{ \xi_k, k \le i, |\xi_j|, 1 \le j \le n \right\}.$$

Since ξ_i is symmetric, it is easy to see that

$$\mathbf{E}\Big(\frac{\xi_i}{\mathrm{V}_n(\beta)} \Big| \mathcal{F}_{i-1} \Big) = \mathbf{E}\Big(\xi_i \Big| |\xi_i| \Big) \frac{1}{\mathrm{V}_n(\beta)} = 0.$$

Therefore, $\big(\xi_i / \mathrm{V}_n(\beta), \mathcal{F}_i\big)_{i=1,...,n}$ is a sequence of martingale differences. Notice that

$$\begin{aligned} \frac{1}{\mathrm{V}_n(\beta)^\beta} \sum_{i=1}^k \Big(\mathbf{E}\big((\xi_i^-)^\beta | \mathcal{F}_{i-1} \big) + (\xi_i^+)^\beta \Big) &\le \frac{1}{\mathrm{V}_n(\beta)^\beta} \sum_{i=1}^n \Big(\mathbf{E}\big(|\xi_i|^\beta | \mathcal{F}_{i-1} \big) + |\xi_i|^\beta \Big) \\ &= \frac{2}{\mathrm{V}_n(\beta)^\beta} \sum_{i=1}^n |\xi_i|^\beta = 2. \end{aligned}$$

Applying Theorem 2.2 to $\big(\xi_i / \mathrm{V}_n(\beta), \mathcal{F}_i\big)_{i=1,...,n}$, we obtain (24). ■

The power $x^{\frac{\beta}{\beta-1}}$ in (24) is the best possible for x in the moderate deviation and large deviation ranges. Indeed, Jing, Liang and Zhou [10] have obtained the following self-normalized moderate deviation result. Assume that

$$\mathbf{P}(\xi_i \ge x) = \mathbf{P}(\xi_i \le -x) \sim \frac{c}{x^\alpha} h_i(x), \qquad x \to \infty,$$

where $\alpha \in (0, 2)$, $c > 0$ and $h_i(x)$'s are slowly varying at ∞. Under certain conditions on the tail probabilities of ξ_i (cf. Theorem 2.3 of [10] for details), for $x_n \to \infty$ and

$x_n = o(n^{(\beta-1)/\beta})$ and $\beta > \max\{1, \alpha\}$, the limit exists

$$\lim_{n\to\infty} x_n^{-\frac{\beta}{\beta-1}} \log \mathbf{P}\left(S_n/\mathrm{V}_n(\beta) \geq x_n\right) = -(\beta-1)C_\alpha(\beta), \tag{26}$$

where $C_\alpha(\beta)$ is a positive constant depending on α and β. Equality (26) suggests that the power $x^{\frac{\beta}{\beta-1}}$ in (24) is the best possible for moderate x's. See also Shao [13] for self-normalized large deviation result.

4. Proof of Theorem 2.1

Assume $(\xi_i, \mathcal{F}_i)_{i=0,\dots,n}$ a sequence of square integrable supermartingale differences. For any nonnegative numbers y and λ, define the exponential multiplicative martingale $Z(\lambda) = (Z_k(\lambda), \mathcal{F}_k)_{k=0,\dots,n}$, where

$$Z_k(\lambda) = \prod_{i=1}^{k} \frac{\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{\xi_i > y\}}\right\}}{\mathbf{E}\left(\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{\xi_i > y\}}\right\} | \mathcal{F}_{i-1}\right)}, \qquad Z_0(\lambda) = 1.$$

If T is a stopping time, then $Z_{T\wedge k}(\lambda)$, $\lambda > 0$, is also a martingale, where

$$Z_{T\wedge k}(\lambda) = \prod_{i=1}^{T\wedge k} \frac{\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{\xi_i > y\}}\right\}}{\mathbf{E}\left(\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{\xi_i > y\}}\right\} | \mathcal{F}_{i-1}\right)}, \qquad Z_0(\lambda) = 1.$$

Then for any nonnegative number λ, we have the following conjugate probability measure $\mathbf{P}_\lambda$ on $(\Omega, \mathcal{F})$:

$$d\mathbf{P}_\lambda = Z_{T\wedge n}(\lambda) d\mathbf{P}. \tag{27}$$

Lemma 4.1 *For all $y \geq 0$ and all $\lambda > 0$, it holds*

$$\mathbf{E}\left(\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{\xi_i > y\}}\right\} \middle| \mathcal{F}_{i-1}\right) \leq \exp\left\{ \left(\frac{e^{\lambda y} - 1 - \lambda y}{y^2}\right) \mathbf{E}(\xi_i^2 \mathbf{1}_{\{\xi_i \leq y\}} | \mathcal{F}_{i-1}) \right\},$$

where, by convention, $\frac{e^{\lambda y}-1-\lambda y}{y^2} = \frac{\lambda^2}{2}$ applied when $y = 0$.

Proof. Let $y \geq 0$. If $\xi_i \leq y$, since the function

$$g(x) = \frac{e^x - 1 - x}{x^2}$$

is increasing in $x \in \mathbf{R}$ (by convention $g(0) = 1/2$), we have, for all $\lambda > 0$,

$$\frac{e^{\lambda\xi_i} - 1 - \lambda\xi_i}{(\lambda\xi_i)^2} \leq \frac{e^{\lambda y} - 1 - \lambda y}{(\lambda y)^2}. \tag{28}$$

If $\xi_i > y$, since $\exp\left\{x - \frac{1}{2}x^2\right\} \leq 1 + x$ for all $x \geq 0$, it follows that, for all $\lambda > 0$,

$$\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\right\} \leq 1 + \lambda\xi_i. \tag{29}$$

Combining (28) and (29) together, we find that, for all $y \geq 0$ and all $\lambda > 0$,

$$\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{\xi_i > y\}}\right\} \leq 1 + \lambda\xi_i + \left(\frac{e^{\lambda y} - 1 - \lambda y}{y^2}\right)\xi_i^2\mathbf{1}_{\{\xi_i \leq y\}}.$$

Taking conditional expectations on both sides of the last inequality, we deduce that

$$\mathbf{E}\left(\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{\xi_i > y\}}\right\}\middle|\mathcal{F}_{i-1}\right) \leq 1 + \left(\frac{e^{\lambda y} - 1 - \lambda y}{y^2}\right)\mathbf{E}(\xi_i^2\mathbf{1}_{\{\xi_i \leq y\}}|\mathcal{F}_{i-1}). \tag{30}$$

Using the inequality $1 + x \leq e^x$ in the right-hand side of (30), we obtain the desired inequality. ■

Proof of Theorem 2.1. For any $x, v > 0$ and any $y \geq 0$, define the stopping time T:

$$T(x, y, v) = \min\{k \in [1, n] : S_k \geq x \text{ and } G_k^y \leq v^2\},$$

with the convention that $\min \emptyset = 0$. Then it follows that

$$\mathbf{1}_{\{S_k \geq x \text{ and } G_k^y \leq v^2 \text{ for some } k \in [1,n]\}} = \sum_{k=1}^{n} \mathbf{1}_{\{T(x,y,v)=k\}}.$$

Denote by $\mathbf{E}_\lambda$ the expectation with respect to the conjugate probability measure $\mathbf{P}_\lambda$. Using the change of probability measure (27), we have, for all $x, \lambda, v > 0$ and all $y \geq 0$,

$$\begin{aligned}
&\mathbf{P}(S_k \geq x \text{ and } G_k^y \leq v^2 \text{ for some } k \in [1, n]) \\
= \;& \mathbf{E}_\lambda\left(Z_{T\wedge n}(\lambda)^{-1}\mathbf{1}_{\{S_k \geq x \text{ and } G_k^y \leq v^2 \text{ for some } k \in [1,n]\}}\right) \\
= \;& \sum_{k=1}^{n} \mathbf{E}_\lambda\left(\exp\left\{-\lambda S_k + \frac{\lambda^2}{2}[S]_k(y) + \widetilde{\Psi}_k(\lambda)\right\}\mathbf{1}_{\{T(x,y,v)=k\}}\right),
\end{aligned} \tag{31}$$

where

$$[S]_k(y) = \sum_{i=1}^{k} \xi_i^2\mathbf{1}_{\{\xi_i > y\}}$$

and

$$\widetilde{\Psi}_k(\lambda) = \sum_{i=1}^{k} \log \mathbf{E}\left(\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{\xi_i > y\}}\right\}\middle|\mathcal{F}_{i-1}\right).$$

Since the function $g(x)$ in increasing in x and $g(0) = 1/2$, we have

$$\frac{\lambda^2}{2} \leq \frac{e^{\lambda y} - 1 - \lambda y}{y^2} \qquad \text{for all } \ y, \lambda > 0.$$

Hence, by Lemma 4.1 and the last inequality,

$$\begin{aligned} & \mathbf{P}(S_k \geq x \text{ and } G_k^y \leq v^2 \text{ for some } k \in [1, n]) \\ \leq & \sum_{k=1}^{n} \mathbf{E}_\lambda \left(\exp\left\{ -\lambda S_k + \frac{\lambda^2}{2}[S]_k(y) + \left(\frac{e^{\lambda y} - 1 - \lambda y}{y^2} \right) \langle S \rangle_k(y) \right\} \mathbf{1}_{\{T(x,y,v)=k\}} \right) \\ \leq & \sum_{k=1}^{n} \mathbf{E}_\lambda \left(\exp\left\{ -\lambda S_k + \left(\frac{e^{\lambda y} - 1 - \lambda y}{y^2} \right) G_k^y \right\} \mathbf{1}_{\{T(x,y,v)=k\}} \right), \qquad (32) \end{aligned}$$

where $\langle S \rangle_k(y) = \sum_{i=1}^{k} \mathbf{E}(\xi_i^2 \mathbf{1}_{\{\xi_i \leq y\}} | \mathcal{F}_{i-1})$. Therefore, by the fact $S_k \geq x$ and $G_k^y \leq v^2$ on the set $\{T(x, y, v) = k\}$, inequality (32) implies that, for all $x, \lambda, v > 0$ and all $y \geq 0$,

$$\begin{aligned} & \mathbf{P}(S_k \geq x \text{ and } G_k^y \leq v^2 \text{ for some } k \in [1, n]) \\ \leq & \sum_{k=1}^{n} \mathbf{E}_\lambda \left(\exp\left\{ -\lambda x + \left(\frac{e^{\lambda y} - 1 - \lambda y}{y^2} \right) v^2 \right\} \mathbf{1}_{\{T(x,y,v)=k\}} \right) \\ \leq & \exp\left\{ -\lambda x + \left(\frac{e^{\lambda y} - 1 - \lambda y}{y^2} \right) v^2 \right\}. \qquad (33) \end{aligned}$$

The last inequality attains its minimum at

$$\lambda = \lambda(x) = \frac{1}{y} \log\left(1 + \frac{xy}{v^2}\right).$$

Substituting $\lambda = \lambda(x)$ in (33), we obtain (14). Using the inequality

$$e^t - 1 - t \leq \frac{t^2}{2(1 - \frac{t}{3})}, \qquad t \geq 0,$$

we get, for all $x, v > 0$ and all $y \geq 0$,

$$\begin{aligned} \inf_{\lambda > 0} \exp\left\{ -\lambda x + \left(\frac{e^{\lambda y} - 1 - \lambda y}{y^2} \right) v^2 \right\} & \leq \inf_{\lambda > 0} \exp\left\{ -\lambda x + \frac{\lambda^2 v^2}{2(1 - \frac{\lambda y}{3})} \right\} \\ & \leq B_2(x, y, v). \end{aligned}$$

Thus (14) implies (15). This completes the proof of Theorem 2.1. ∎

5. Proof of Theorem 2.2

Assume $\mathbf{E}|\xi_i|^\beta < \infty$ for a constant $\beta \in (1, 2)$ and for all $i \in [1, n]$. For any nonnegative numbers λ, define the exponential multiplicative martingale $Z(\lambda) = (Z_k(\lambda), \mathcal{F}_k)_{k=0,\ldots,n}$, where

$$Z_k(\lambda) = \prod_{i=1}^{k} \frac{\exp\{\lambda\xi_i - (\lambda\xi_i^+)^\beta\}}{\mathbf{E}\left(\exp\{\lambda\xi_i - (\lambda\xi_i^+)^\beta\} | \mathcal{F}_{i-1}\right)}, \qquad Z_0(\lambda) = 1.$$

If T is a stopping time, then $Z_{T\wedge k}(\lambda)$, $\lambda \geq 0$, is also a martingale, where

$$Z_{T\wedge k}(\lambda) = \prod_{i=1}^{T\wedge k} \frac{\exp\{\lambda\xi_i - (\lambda\xi_i^+)^\beta\}}{\mathbf{E}\left(\exp\{\lambda\xi_i - (\lambda\xi_i^+)^\beta\} | \mathcal{F}_{i-1}\right)}, \qquad Z_0(\lambda) = 1.$$

Then for any nonnegative number λ, we introduce the following conjugate probability measure $\mathbf{P}_\lambda$ on $(\Omega, \mathcal{F})$:

$$d\mathbf{P}_\lambda = Z_{T\wedge n}(\lambda) d\mathbf{P}. \tag{34}$$

LEMMA 5.1 *If* $\mathbf{E}|\xi_i|^\beta < \infty$ *for a constant* $\beta \in (1, 2)$*, then, for all* $\lambda > 0$*,*

$$\mathbf{E}\Big(\exp\Big\{\lambda\xi_i - \lambda^\beta (\xi_i^+)^\beta\Big\} \Big| \mathcal{F}_{i-1}\Big) \leq \exp\Big\{\lambda^\beta \mathbf{E}((\xi_i^-)^\beta | \mathcal{F}_{i-1})\Big\}.$$

Proof. It is easy to see that, for all $x \in \mathbf{R}$ and $\beta \in (1, 2)$,

$$\exp\Big\{x - (x^+)^\beta\Big\} \leq 1 + x + (x^-)^\beta.$$

With $x = \lambda\xi_i$, we easily obtain, for all $\lambda \geq 0$,

$$\exp\Big\{\lambda\xi_i - (\lambda\xi_i^+)^\beta\Big\} \leq 1 + \lambda\xi_i + (\lambda\xi_i^-)^\beta. \tag{35}$$

Taking conditional expectations on both sides of (35), we get

$$\mathbf{E}\Big(\exp\Big\{\lambda\xi_i - \lambda^\beta (\xi_i^+)^\beta\Big\} \Big| \mathcal{F}_{i-1}\Big) \leq 1 + \lambda^\beta \mathbf{E}((\xi_i^-)^\beta | \mathcal{F}_{i-1}).$$

Using the inequality $1 + x \leq e^x$, we obtain the desired inequality. ■

Proof of Theorem 2.2. For given $x, v > 0$, define the stopping time T :

$$T = \min\{k \in [1, n] : S_k \geq x \text{ and } \mathrm{G}_k^0(\beta) \leq v^\beta\},$$

with the convention that $\min \emptyset = 0$. Then we have

$$\mathbf{1}_{\{S_k \geq x \text{ and } G_k^0(\beta) \leq v^\beta \text{ for some } k \in [1,n]\}} = \sum_{k=1}^{n} \mathbf{1}_{\{T=k\}}.$$

Denote by $\mathbf{E}_\lambda$ the expectation with respect to the conjugate probability measure (34).

Using the change of measure (34), we get, for all $x, \lambda, v > 0$,

$$\begin{aligned} & \mathbf{P}(S_k \geq x \text{ and } \mathrm{G}_k^0(\beta) \leq v^\beta \text{ for some } k \in [1,n]) \\ = \ & \mathbf{E}_\lambda\Big(Z_{T\wedge n}(\lambda)^{-1}\mathbf{1}_{\{S_k\geq x \text{ and } G_k^0(\beta)\leq v^2 \text{ for some } k\in[1,n]\}}\Big) \\ = \ & \sum_{k=1}^n \mathbf{E}_\lambda\Big(\exp\{-\lambda S_k + \lambda^\beta \sum_{i=1}^k (\xi_i^+)^\beta + \widehat{\Psi}_k(\lambda)\}\mathbf{1}_{\{T=k\}}\Big), \end{aligned} \tag{36}$$

where

$$\widehat{\Psi}_k(\lambda) = \sum_{i=1}^k \log \mathbf{E}\exp\Big\{\lambda\xi_i - (\lambda\xi_i^+)^\beta\Big\}.$$

From inequality (36), by Lemma 5.1, it follows that, for all $x, \lambda, v > 0$,

$$\begin{aligned} & \mathbf{P}(S_k \geq x \text{ and } \mathrm{G}_k^0(\beta) \leq v^\beta \text{ for some } k \in [1,n]) \\ & \qquad \leq \sum_{k=1}^n \mathbf{E}_\lambda\Big(\exp\{-\lambda S_k + \lambda^\beta \mathrm{G}_k^0(\beta)\}\mathbf{1}_{\{T=k\}}\Big). \end{aligned}$$

Since $S_k \geq x$ and $\mathrm{G}_k^0(\beta) \leq v^\beta$ on the set $\{T = k\}$, we obtain, for all $x, \lambda, v > 0$,

$$\mathbf{P}(S_k \geq x \text{ and } \mathrm{G}_k^0(\beta) \leq v^\beta \text{ for some } k \in [1,n]) \leq \exp\{-\lambda x + \lambda^\beta v^\beta\}. \tag{37}$$

The last inequality attains its minimum at

$$\lambda = \lambda(x) = \left(\frac{x}{\beta v^\beta}\right)^{\frac{1}{\beta-1}}.$$

Substituting $\lambda = \lambda(x)$ in (37), we get the desired inequality. ∎

6. Proof of Theorem 2.3

Assume that $(\xi_i, \mathcal{F}_i)_{i=0,\dots,n}$ is a sequence of martingale differences. For any nonnegative numbers y and λ, define the exponential multiplicative martingale $Z(\lambda) = (Z_k(\lambda), \mathcal{F}_k)_{k=0,\dots,n}$, where

$$Z_k(\lambda) = \prod_{i=1}^k \frac{\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{|\xi_i|>y\}}\right\}}{\mathbf{E}\left(\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{|\xi_i|>y\}}\right\}|\mathcal{F}_{i-1}\right)}, \qquad Z_0(\lambda) = 1.$$

If T is a stopping time, then $Z_{T\wedge k}(\lambda)$, $\lambda > 0$, is also a martingale, where

$$Z_{T\wedge k}(\lambda) = \prod_{i=1}^{T\wedge k} \frac{\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{|\xi_i|>y\}}\right\}}{\mathbf{E}\left(\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{|\xi_i|>y\}}\right\}|\mathcal{F}_{i-1}\right)}, \qquad Z_0(\lambda) = 1.$$

Then for any nonnegative number λ, we have the following conjugate probability measure $\mathbf{P}_\lambda$ on $(\Omega, \mathcal{F})$:

$$d\mathbf{P}_\lambda = Z_{T\wedge n}(\lambda) d\mathbf{P}. \tag{38}$$

LEMMA 6.1 *Assume that* $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$ *is a sequence of conditionally symmetric martingale differences. For all* $\lambda, y \geq 0$, *it holds*

$$\mathbf{E}\left(\exp\left\{ \lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{|\xi_i|>y\}} \right\} \middle| \mathcal{F}_{i-1} \right) \leq \exp\left\{ \left(\frac{\cosh(\lambda y) - 1}{y^2} \right) \mathbf{E}(\xi_i^2 \mathbf{1}_{\{|\xi_i|\leq y\}} | \mathcal{F}_{i-1}) \right\},$$

where by convention $\frac{\cosh(\lambda y)-1}{y^2} = \frac{\lambda^2}{2}$ *when* $y = 0$.

Proof. Let $y \geq 0$. When $|\xi_i| \leq y$, it follows that $\xi_i^{2k} \leq y^{2k-2}\xi_i^2$ and that

$$\cosh(\lambda\xi_i) \leq 1 + \frac{\xi_i^2}{y^2} \sum_{k=1}^{\infty} \frac{(\lambda y)^{2k}}{(2k)!} = 1 + \frac{\xi_i^2}{y^2} \Big(\cosh(\lambda y) - 1 \Big). \tag{39}$$

When $|\xi_i| > y$, since $\cosh(x) \leq \exp\left\{\frac{1}{2}x^2\right\}$ for all $x \in \mathbf{R}$, it follows that, for all $\lambda > 0$,

$$\Big(\cosh(\lambda\xi_i) \Big) \exp\left\{ -\frac{1}{2}(\lambda\xi_i)^2 \right\} \leq 1. \tag{40}$$

Combining (39) and (40) together, we find that, for all $\lambda, y \geq 0$,

$$\Big(\cosh(\lambda\xi_i) \Big) \exp\left\{ -\frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{|\xi_i|>y\}} \right\} \leq 1 + \lambda\xi_i + \left(\frac{e^{\lambda y} - 1 - \lambda y}{y^2} \right) \xi_i^2 \mathbf{1}_{\{|\xi_i|\leq y\}}.$$

Taking conditional expectations on both sides of the last inequality, we have, for all $\lambda, y \geq 0$,

$$\begin{aligned} \mathbf{E}\left(\Big(\cosh(\lambda\xi_i) \Big) \exp\left\{ -\frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{|\xi_i|>y\}} \right\} \middle| \mathcal{F}_{i-1} \right) & \\ \leq 1 + \left(\frac{\cosh(\lambda y) - 1}{y^2} \right) \mathbf{E}(\xi_i^2 \mathbf{1}_{\{|\xi_i|\leq y\}} | \mathcal{F}_{i-1}). & \end{aligned} \tag{41}$$

Since $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$ are conditionally symmetric, it holds

$$\begin{aligned} \mathbf{E}\left(\Big(\exp\{\lambda\xi_i\} \Big) \exp\left\{ -\frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{|\xi_i|>y\}} \right\} \middle| \mathcal{F}_{i-1} \right) & \\ = \mathbf{E}\left(\Big(\exp\{-\lambda\xi_i\} \Big) \exp\left\{ -\frac{1}{2}(\lambda\xi_i)^2 \mathbf{1}_{\{|\xi_i|>y\}} \right\} \middle| \mathcal{F}_{i-1} \right). & \end{aligned} \tag{42}$$

Note that $\cosh(\lambda\xi_i) = \frac{1}{2}(\exp\{\lambda\xi_i\} + \exp\{-\lambda\xi_i\})$. Hence (42) implies that

$$\mathbf{E}\left(\Big(\cosh(\lambda\xi_i)\Big)\exp\left\{-\frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{|\xi_i|>y\}}\right\}\bigg|\mathcal{F}_{i-1}\right)$$
$$= \mathbf{E}\left(\Big(\exp\{\lambda\xi_i\}\Big)\exp\left\{-\frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{|\xi_i|>y\}}\right\}\bigg|\mathcal{F}_{i-1}\right). \quad (43)$$

Combining (41) and (43) together, we obtain

$$\mathbf{E}\left(\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{|\xi_i|>y\}}\right\}\bigg|\mathcal{F}_{i-1}\right) \le 1 + \left(\frac{\cosh(\lambda y)-1}{y^2}\right)\mathbf{E}(\xi_i^2\ \mathbf{1}_{\{|\xi_i|\le y\}}|\mathcal{F}_{i-1}).$$

Using the inequality $1 + x \le e^x$, we obtain the desired inequality. ■

Proof of Theorem 2.3. For any $y \ge 0$ and any $x, v > 0$, define the stopping time T:

$$T(x,y,v) = \min\{k \in [1,n] : S_k \ge x \text{ and } G_k^y \le v^2\},$$

with the convention that $\min\emptyset = 0$. Then it follows that

$$\mathbf{1}_{\{S_k\ge x \text{ and } G_k^y\le v^2 \text{ for some } k\in[1,n]\}} = \sum_{k=1}^n \mathbf{1}_{\{T(x,y,v)=k\}}.$$

Denote $\mathbf{E}_\lambda$ the expectation with respect to $\mathbf{P}_\lambda$. Using the change of probability measure (38), we have, for all $x, \lambda, v > 0$ and all $y \ge 0$,

$$\begin{aligned} & \mathbf{P}(S_k \ge x \text{ and } G_k^y \le v^2 \text{ for some } k \in [1,n]) \\ = \ & \mathbf{E}_\lambda\left(Z_{T\wedge n}(\lambda)^{-1}\mathbf{1}_{\{S_k\ge x \text{ and } G_k^y\le v^2 \text{ for some } k\in[1,n]\}}\right) \\ = \ & \sum_{k=1}^n \mathbf{E}_\lambda\left(\exp\left\{-\lambda S_k + \frac{\lambda^2}{2}[S]_k(y) + \breve{\Psi}_k(\lambda)\right\}\mathbf{1}_{\{T(x,y,v)=k\}}\right), \end{aligned} \quad (44)$$

where $[S]_k(y) = \sum_{i=1}^k \xi_i^2\mathbf{1}_{\{|\xi_i|>y\}}$ and

$$\breve{\Psi}_k(\lambda) = \sum_{i=1}^k \log\mathbf{E}\left(\exp\left\{\lambda\xi_i - \frac{1}{2}(\lambda\xi_i)^2\mathbf{1}_{\{|\xi_i|>y\}}\right\}\bigg|\mathcal{F}_{i-1}\right).$$

Since the function $f(y) = \frac{\cosh(\lambda y)-1}{y^2}$ is increasing in $y \ge 0$ and $f(0) = \lambda^2/2$, we have

$$\frac{\lambda^2}{2} \le \frac{\cosh(\lambda y)-1}{y^2} \quad \text{for all} \ \ y, \lambda > 0.$$

By Lemma 4.1, we find that

$$\widetilde{\Psi}_k(\lambda) \le \left(\frac{\cosh(\lambda y)-1}{y^2}\right)\langle S\rangle_k(y),$$

where $\langle S\rangle_k(y) = \sum_{i=1}^k \mathbf{E}(\xi_i^2 \mathbf{1}_{\{|\xi_i|>y\}}|\mathcal{F}_{i-1})$. Then, from equality (44), it follows that, for all $x, \lambda, v > 0$ and all $y \geq 0$,

$$\begin{aligned} & \mathbf{P}(S_k \geq x \text{ and } G_k^y \leq v^2 \text{ for some } k \in [1, n]) \\ \leq & \sum_{k=1}^n \mathbf{E}_\lambda \left(\exp\left\{ -\lambda S_k + \frac{\lambda^2}{2}[S]_k(y) + \left(\frac{\cosh(\lambda y) - 1}{y^2}\right) \langle S\rangle_k(y) \right\} \mathbf{1}_{\{T=k\}} \right) \\ \leq & \sum_{k=1}^n \mathbf{E}_\lambda \left(\exp\left\{ -\lambda S_k + \left(\frac{\cosh(\lambda y) - 1}{y^2}\right) G_k^y \right\} \mathbf{1}_{\{T=k\}} \right). \end{aligned} \tag{45}$$

By the fact $S_k \geq x$ and $G_k^y \leq v^2$ on the set $\{T(x, y, v) = k\}$, inequality (45) implies that, for all $x, \lambda, v > 0$ and all $y \geq 0$,

$$\begin{aligned} & \mathbf{P}(S_k \geq x \text{ and } G_k^y \leq v^2 \text{ for some } k \in [1, n]) \\ \leq & \exp\left\{ -\lambda x + \left(\frac{\cosh(\lambda y) - 1}{y^2}\right) v^2 \right\}. \end{aligned} \tag{46}$$

The last inequality attains its minimum at

$$\lambda = \lambda(x) = \frac{1}{y} \log\left(\sqrt{1 + \frac{x^2 y^2}{v^4}} + \frac{xy}{v^2} \right).$$

Substituting $\lambda = \lambda(x)$ in (46), we obtain the desired inequality. This completes the proof of Theorem 2.3. ■

Acknowledgements

We would like to thank Igal Sason for his helpful remarks and suggestions.

References

[1] Bennett, G., 1962. Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57**, No. 297, 33–45.

[2] Bercu, B. and Touati, A., 2008. Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.* **18**(5): 1848–1869.

[3] Bernstein, S. N., 1927. *Theorem of Probability.* Moscow.

[4] De la Peña, V. H., Pang, G., 2009. Exponential inequalities for self-normalized processes with applications. *Electron. Commun. Probab.* **14**: 372–381.

[5] De la Peña, V. H., 1999. A general class of exponential inequalities for martingales and ratios. *Ann. Probab.* **27**, No. 1, 537–564.

[6] Delyon, B., 2009. Exponential inequaties for sums of weakly dependent variables. *Electronic J. Probab.* **14**, No. 28, 752–779.

[7] Dzhaparidze, K. and van Zanten, J. H., 2001. On Bernstein-type inequalities for martingales. *Stochastic Process. Appl.* **93**, 109–117.

[8] Fan, X., Grama, I. and Liu, Q., 2015. Exponential inequalities for martingales with applications. *Electron. J. Probab.* **20**, no. 1, 1–22.

[9] Freedman, D. A., 1975. On tail probabilities for martingales. *Ann. Probab.* **3**, No. 1, 100–118.

[10] Jing, B. Y., Liang, H. Y. and Zhou, W., 2012. Self-normalized moderate devioations for independent random variables. *Sci. China Math.* **55**(11), 2297–2315.

[11] Raginsky, M. and Sason, I., 2013. Concentration of measure inequalities in information theory, communications, and coding. *Foundations and Trends ® in Communications and Information Theory* **10**, no. 1-2, 1–246.

[12] Sason, I., 2013. Tightened exponential bounds for discrete-time conditionally symmetric martingales with bounded jumps. *Statist. Probab. Lett.* **83**, 1928–1936.

[13] Shao, Q. M., 1997. Self-normalized large deviations. *Ann. Probab.* **25**(1): 285–328.