Journal of Mathematical Research with Applications
May, 2019, Vol. 39, No. 3, pp. 304–314
DOI:10.3770/j.issn:2095-2651.2019.03.010
Http://jmre.dlut.edu.cn

Self-Normalized Exponential Inequalities for Martingales

Shen WANG
Center for Applied Mathematics, Tianjin University, Tianjin 300072*, P. R. China*

Abstract We give an extension of Delyon's inequality for locally square integrable martingales. The result is very useful for establishing self-normalized exponential inequality for martingales. An application to linear regressions is discussed.

Keywords exponential inequalities; self-normalized processes; martingales

MR(2010) Subject Classification 60E15; 60F10

1. Introduction

Let $(\xi_i, \mathcal{F}_i)_{i=0,\dots,n}$ be a finite sequence of real-valued square integrable martingale differences defined on a probability space $(\Omega, \mathcal{F}, P)$, where $\xi_0 = 0$ and $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq \cdots \subseteq \mathcal{F}_n \subseteq \mathcal{F}$ are increasing σ-fields. So by definition, we have $\mathbb{E}[\xi_i|\mathcal{F}_{i-1}] = 0$, $i = 1, \dots, n$. Set $S_0 = 0$ and $S_k = \sum_{i=1}^k \xi_i$ for $k = 1, \dots, n$. Then $S = (S_k, \mathcal{F}_k)_{k=1,\dots,n}$ is a martingale. Let $[S]$ and $\langle S\rangle$ be, respectively, the squared variance and the conditional variance of the martingale S, that is

$$[S]_0 = \langle S\rangle_0 = 0,\ [S]_k = \sum_{i=1}^k \xi_i^2 \quad \text{and} \quad \langle S\rangle_k = \sum_{i=1}^k \mathbb{E}[\xi_i^2|\mathcal{F}_{i-1}], \quad k = 1, \dots, n.$$

Bercu and Touati [1] established the following exponential inequality. For any $x, v_n > 0$,

$$\mathbb{P}(|S_n| \geq x,\ [S]_n + \langle S\rangle_n \leq v_n^2) \leq 2\exp\{-\frac{x^2}{2v_n^2}\}.$$

Delyon [2] refined the inequality of Bercu and Touati, and gave the following result. For any $x, v_n > 0$,

$$\mathbb{P}(S_n \geq x,\ \frac{1}{3}[S]_n + \frac{2}{3}\langle S\rangle_n \leq v_n^2) \leq \exp\{-\frac{x^2}{2v_n^2}\}. \tag{1.1}$$

The proof of Deylon is based on the following fact that for any $t > 0$,

$$\mathbb{E}[\exp\{tS_n - \frac{t^2}{2}(\frac{1}{3}[S]_n + \frac{2}{3}\langle S\rangle_n)\}] \leq 1. \tag{1.2}$$

Combining inequality (1.2) and exponential Markov's inequality together, we have

$$\mathbb{P}(S_n \geq x,\ \frac{1}{3}[S]_n + \frac{2}{3}\langle S\rangle_n \leq v_n^2) \leq \exp\{tx - \frac{t^2}{2}v_n^2\}.$$

Taking $t = x/v^2$ in the r.h.s. of the last inequality, we obtain inequality (1.1). Inequality (1.2) is very useful. For instance, de la Peña and Pang [3] showed that such type inequalities are closely related to the exponential inequalities for self-normalized martingales.

Received July 21, 2018; Accepted December 26, 2018
E-mail address: sherry0wang@163.com

In this paper, we show that inequality (1.2) holds also when the constants $\frac{1}{6}$ and $\frac{1}{3}$ therein are replaced by certain positive constants, that is

$$\mathbb{E}[\exp\{tS_n - \frac{t^2}{2}(p[S]_n + p'\langle S\rangle_n)\}] \leq 1, \tag{1.3}$$

where p and p' are certain positive constants such that (2.1) holds. Inequality (1.3) is very useful. For instance, by exponential Markov's inequality, it immediately leads to the following result: For any $x, v_n > 0$,

$$\mathbb{P}(S_n \geq x,\ p[S]_n + p'\langle S\rangle_n \leq v_n^2) \leq \exp\{-\frac{x^2}{2v_n^2}\},$$

which can be regarded as a generalization of Delyon's inequality. See also (2.3) for a more stronger result. Moreover, following de la Peña and Pang [3] and Bercu and Touati [1], we show that inequality (1.3) can be applied to self-normalized exponential inequalities for martingales. In particular, we prove that if $\langle S\rangle_n \geq b^2\mathbb{E}S_n^2$ a.s. for small positive b, then for any $x > 0$,

$$\mathbb{P}(\frac{|S_n|}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \geq x) \leq C|\ln b|\exp\{-\frac{x^2}{6}\}.$$

The rest of the paper is organized as follows. Our main results are stated and discussed in Section 2. The application to linear regressions is given in Section 3. In Section 4, we prove our main results.

2. Main results

Let $p > 0$. Denote

$$f(x) = \frac{2(\exp\{x - \frac{p}{2}x^2\} - 1 - x)}{x^2}, \quad x \in \mathbb{R}.$$

Set

$$q = \sup_{x\in\mathbb{R}} f(x).$$

Notice that

$$\lim_{x\to 0} f(x) = 1 - p, \quad \lim_{x\to\pm\infty} f(x) = 0,$$

and $f(x)$ is continuous on $\mathbb{R}$. Thus, q always exists and satisfies $q \geq 1-p$. In particular, Delyon [2] proved that if $p = \frac{1}{3}$, then $q = \frac{2}{3}$. By the definition of q, we have for any constant $p' \geq q$,

$$\exp\{x - \frac{p}{2}x^2\} \leq 1 + x + \frac{p'}{2}x^2, \quad x \in \mathbb{R}. \tag{2.1}$$

For practical purposes, one can use the following table:

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	2	3
p'	3.56	0.98	0.71	0.62	0.58	0.55	0.54	0.53	0.52	0.52	0.51	0.51

Table 1 Value of p' under p

It is easy to see that when p is increasing, q is decreasing.

Our first result is the following exponential inequality for martingales.

Theorem 2.1 *Let $p > 0$. Assume $p' \geq q$. Then for any $t \in \mathbb{R}$,*

$$\mathbb{E}[\exp\{tS_n - \frac{t^2}{2}(p[S]_n + p'\langle S\rangle_n)\}] \leq 1, \tag{2.2}$$

and for any $x, v_n > 0$,

$$\mathbb{P}(S_k \geq x \text{ and } p[S]_k + p'\langle S\rangle_k \leq v_n^2 \text{ for some } k \in [1, n]) \leq \exp\{-\frac{x^2}{2v_n^2}\}. \tag{2.3}$$

In particular, the last inequality implies that for any $x, v_n > 0$,

$$\mathbb{P}(\max_{1\leq k\leq n} S_k \geq x,\ p[S]_n + p'\langle S\rangle_n \leq v_n^2) \leq \exp\{-\frac{x^2}{2v_n^2}\}.$$

Clearly, inequality (2.3) with $p = \frac{1}{3}$ and $p' = \frac{2}{3}$ implies inequality (1.1). Thus our inequality can be regarded as an extension of Delyon's inequality (1.1).

Inequality (2.2) is very useful for establishing exponential inequality for self-normalized martingales. Using (2.2), we obtain the following analogue of Bercu and Touati [1, Theorem 4.2].

Theorem 2.2 *Assume $p' \geq q$. Then for any $x > 0$,*

$$\mathbb{P}(\frac{S_n}{p[S]_n + p'\langle S\rangle_n} \geq x) \leq \inf_{l>1}(\mathbb{E}[\exp\{-(l-1)\frac{x^2}{2}(p[S]_n + p'\langle S\rangle_n)\}])^{\frac{1}{l}},$$

and, for all $y > 0$,

$$\mathbb{P}(\frac{S_n}{p[S]_n + p'\langle S\rangle_n} \geq x,\ p[S]_n + p'\langle S\rangle_n \geq y) \leq \exp\{-\frac{x^2 y}{2}\}.$$

Moreover, the same inequalities hold when S_n is replaced by $-S_n$.

Similarly, we have the following result.

Theorem 2.3 *Assume $p' \geq q$. Then for any $x, y > 0$,*

$$\mathbb{P}(\frac{S_n}{[S]_n} \geq x,\ (2-p)[S]_n \geq p'\langle S\rangle_n + y) \leq \exp\{-\frac{x^2 y}{2}\}. \tag{2.4}$$

Moreover, we also have

$$\mathbb{P}(\frac{S_n}{[S]_n} \geq x,\ p'\langle S\rangle_n \leq yp[S]_n) \leq \inf_{l>1}(\mathbb{E}[\exp\{-\frac{(l-1)x^2}{2(1+y)p}[S]_n\}])^{\frac{1}{l}}. \tag{2.5}$$

Moreover, the same inequalities hold when S_n is replaced by $-S_n$.

Clearly, when $p = p' = 1$, Theorem 2.3 reduces to Theorem 2.2 of Bercu and Touati [1] with $a = 0$ and $b = 1$.

Remark 2.4 According to the proof of Theorem 2.3, it is not hard to see that (2.4) and (2.5) also hold exchanging the roles of $\langle S\rangle_n$ and $[S]_n$, and exchanging the roles of p and p' at the same time. That is, the following two inequalities hold: for any $x, y > 0$,

$$\mathbb{P}(\frac{S_n}{\langle S\rangle_n} \geq x,\ (2-p')\langle S\rangle_n \geq p[S]_n + y) \leq \exp\{-\frac{x^2 y}{2}\}$$

and

$$\mathbb{P}(\frac{S_n}{\langle S\rangle_n} \geq x,\ p[S]_n \leq yp'\langle S\rangle_n) \leq \inf_{l>1}(\mathbb{E}[\exp\{-\frac{(l-1)x^2}{2(1+y)p'}\langle S\rangle_n\}])^{\frac{1}{l}}.$$

Denote s_n the standard variance of S_n. It is well known that

$$s_n^2 = \mathbb{E}S_n^2 = \sum_{i=1}^{n} \mathbb{E}\xi_i^2.$$

Inequality (2.2) implies the following exponential inequality for self-normalized martingales.

Theorem 2.5 Assume $p' \geq q$. Then for any $b \in (0, 1/2]$ and $x \geq 0$,

$$\mathbb{P}\Big(\frac{|S_n|}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \geq x,\ \sqrt{p[S]_n + p'\langle S\rangle_n} \geq bs_n\Big) \leq C|\ln b| \exp\{-\frac{x^2}{6}\}, \tag{2.6}$$

where C is an absolute constant. In particular, if $p'\langle S\rangle_n \geq b^2 s_n^2 > 0$ a.s., then (2.6) implies that

$$\mathbb{P}\Big(\frac{|S_n|}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \geq x\Big) \leq C|\ln b| \exp\{-\frac{x^2}{6}\}.$$

Some exponential inequalities similar to (2.6) can be found in Liptser and Spokoiny [4], de la Peña and Pang [3]. In these papers, Liptser and Spokoiny [4] considered the probability of the event $\{\frac{S_n}{\sqrt{p[S]_n+p'\langle S\rangle_n}} \geq x,\ b \leq \sqrt{p[S]_n + p'\langle S\rangle_n} \leq bM\}$, while de la Peña and Pang [3] considered the tail probability of $\frac{|S_n|}{\sqrt{p[S]_n+p'\langle S\rangle_n+s_n^2}}$.

3. Application to linear regressions

The stochastic linear regression model is given by, for all $n \geq 0$,

$$X_{n+1} = \theta\phi_n + \varepsilon_{n+1}, \tag{3.1}$$

where X_n, ϕ_n and ε_n are the observation, the regression variable, and the driven noise, respectively. We assume that (ϕ_n) is a sequence of independent and identically distributed random variables, and that (ε_n) is a sequence of identically distributed random variables with mean zero and variance $\sigma^2 > 0$. Furthermore, we suppose that, for all $n \geq 0$, the random variable ε_{n+1} is independent of $\mathcal{F}_n$ where $\mathcal{F}_n = \sigma(\phi_0, \varepsilon_1, \ldots, \phi_{n-1}, \varepsilon_n)$. Our interest is to estimate the unknown parameter θ. The well-known least-squares estimator $\widehat{\theta_n}$ is given below

$$\widehat{\theta_n} = \frac{\sum_{k=1}^{n} \phi_{k-1} X_k}{\sum_{k=1}^{n} \phi_{k-1}^2}. \tag{3.2}$$

It immediately follows from (3.1) and (3.2) that

$$\widehat{\theta_n} - \theta_n = \sigma^2 \frac{S_n}{\langle S\rangle_n}, \tag{3.3}$$

where

$$S_n = \sum_{k=1}^{n} \phi_{k-1}\varepsilon_k \quad \text{and} \quad \langle S\rangle_n = \sigma^2 \sum_{k=1}^{n} \phi_{k-1}^2.$$

Let H and L be the cumulant generating functions of the sequences (ϕ_n^2) and (ε_n^2), respectively given, for all $t \in \mathbb{R}$, by

$$H(t) = \log \mathbb{E}[\exp\{t\phi_n^2\}] \quad \text{and} \quad L(t) = \log \mathbb{E}[\exp\{t\varepsilon_n^2\}].$$

When ϕ_n and ε_n are sub-Gaussian, the following exponential inequality on the convergence of $\widehat{\theta_n} - \theta_n$ has been established by Bercu and Touati [1]. Assume that L is finite on some interval $[0, c]$ with $c > 0$ and denote by I its Fenchel-Legendre transform on $[0, c]$,

$$I(x) = \sup_{0 \le t \le c} \{xt - L(t)\}.$$

Then, for all $n \ge 1$, $x > 0$ and $y > 0$, Bercu and Touati proved that

$$\mathbb{P}(|\widehat{\theta_n} - \theta| \ge x) \le 2 \inf_{l > 1} \exp\{\frac{n}{l} H(-\frac{(l-1)x^2}{2\sigma^2(1+y)})\} + \exp\{-nI(\frac{y\sigma^2}{n})\}. \tag{3.4}$$

Here, we would like to give a refinement of (3.4).

Theorem 3.1 *For all $n \ge 1$, $x > 0$, and $y > 0$, we have*

$$\mathbb{P}(|\widehat{\theta_n} - \theta| \ge x) \le 2 \inf_{l > 1} \exp\{\frac{n}{l} H(-\frac{(l-1)x^2}{2\sigma^2(1+y)p'})\} + \exp\{-nI(\frac{p' y\sigma^2}{pn})\}.$$

Clearly, when $p = p' = 1$, our inequality reduces to the inequality of Bercu and Touati.

4. Proofs of theorems

In the proof of Theorem 2.1, we make use of the following lemma.

Lemma 4.1 *Assume $p' \ge q$. Then for any $t \in \mathbb{R}$,*

$$\mathbb{E}[\exp\{t\xi_i - \frac{p}{2} t^2 \xi_i^2\} | \mathcal{F}_{i-1}] \le \exp\{\frac{p'}{2} t^2 \mathbb{E}[\xi_i^2 | \mathcal{F}_{i-1}]\}.$$

Proof By inequality (2.1), we have

$$\exp\{t\xi_i - \frac{p}{2} t^2 \xi_i^2\} \le 1 + t\xi_i + \frac{p'}{2} t^2 \xi_i^2.$$

Taking conditional expectation on both sides of the last inequality, we get

$$\mathbb{E}[\exp\{t\xi_i - \frac{p}{2} t^2 \xi_i^2\} | \mathcal{F}_{i-1}] \le 1 + \frac{p'}{2} t^2 \mathbb{E}[\xi_i^2 | \mathcal{F}_{i-1}].$$

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

$$\mathbb{E}[\exp\{t\xi_i - \frac{p}{2} t^2 \xi_i^2\} | \mathcal{F}_{i-1}] \le \exp\{\frac{p'}{2} t^2 \mathbb{E}[\xi_i^2 | \mathcal{F}_{i-1}]\},$$

which gives the desired inequality. □

From the last lemma, we get

Lemma 4.2 *Assume $p' \ge q$. For any $t \in \mathbb{R}$, denote*

$$V_n(t) = \exp\{tS_n - \frac{t^2}{2}(p[S]_n + p'\langle S \rangle_n)\}.$$

Then $(V_i(t), \mathcal{F}_i)_{i=1,\dots,n}$ is a positive supermartingale with $\mathbb{E}[V_n(t)] \le 1$.

Proof For all $t \in \mathbb{R}$ and $n \ge 0$, we have

$$V_n(t) = V_{n-1}(t) \exp\{t\xi_n - \frac{p}{2} t^2 (p\xi_n^2 + p'\mathbb{E}[\xi_n^2 | \mathcal{F}_{n-1}])\}.$$

Hence, by Lemma 4.1, we deduce that for all $t \in \mathbb{R}$,

$$\begin{aligned}\mathbb{E}[V_n(t)|\mathcal{F}_{n-1}] &= V_{n-1}(t)\mathbb{E}[\exp\{t\xi_n - \frac{t^2}{2}(p\xi_n^2 + p'\mathbb{E}[\xi_n^2|\mathcal{F}_{n-1}])\}|\mathcal{F}_{n-1}] \\ &\le V_{n-1}(t),\end{aligned}$$

which means $(V_i(t), \mathcal{F}_i)_{i=0,\dots,n}$ is a positive supermartingale. Moreover, there holds

$$\mathbb{E}[V_n(t)] \le \mathbb{E}[V_{n-1}(t)] \le \cdots \le \mathbb{E}[V_1(t)] \le 1.$$

This completes the proof of lemma. □

Proof of Theorem 2.1 We follow the argument of Fan et al. [5]. For any nonnegative number λ, define the exponential multiplicative martingale $(Z_k(\lambda), \mathcal{F}_k)_{k=0,\dots,n}$, where

$$Z_k(\lambda) = \prod_{i=1}^{k} \frac{e^{\lambda\xi_i - \frac{p}{2}\lambda^2\xi_i^2}}{\mathbb{E}[e^{\lambda\xi_i - \frac{p}{2}\lambda^2\xi_i^2}|\mathcal{F}_{i-1}]}, \quad Z_0(\lambda) = 1, \ \lambda \ge 0.$$

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

$$Z_{T\wedge k}(\lambda) = \prod_{i=1}^{T\wedge k} \frac{e^{\lambda\xi_i - \frac{p}{2}\lambda^2\xi_i^2}}{\mathbb{E}[e^{\lambda\xi_i - \frac{p}{2}\lambda^2\xi_i^2}|\mathcal{F}_{i-1}]}, \quad Z_0(\lambda) = 1, \ \lambda \ge 0.$$

Thus the random variable $Z_{T\wedge k}(\lambda)$ is a probability density on $(\Omega, \mathcal{F}, \mathbb{P})$, i.e.,

$$\int Z_{T\wedge k}(\lambda)\mathrm{d}\mathbb{P} = \mathbb{E}[Z_{T\wedge k}(\lambda)] = 1.$$

Define the conjugate probability measure $\mathbb{P}_\lambda$ on $(\Omega, \mathcal{F})$ as follows

$$\mathrm{d}\mathbb{P}_\lambda = Z_{T\wedge n}(\lambda)\mathrm{d}\mathbb{P}, \tag{4.1}$$

and denote by $\mathbb{E}_\lambda$ the expectation with respect to $\mathbb{P}_\lambda$. For any $x, v_n > 0$, define the stopping time

$$T(x, v_n) = \min\{k \in [1, n] : \ S_k \ge x \text{ and } \ p[S]_k + p'\langle S\rangle_k \le v_n^2\},$$

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

$$\mathbf{1}_{\{S_k \ge x \text{ and } p[S]_k + p'\langle S\rangle_k \le v_n^2 \text{ for some } k\in[1,n]\}} = \sum_{k=1}^{n} \mathbf{1}_{\{T(x,v_n)=k\}}.$$

By the change of measure (4.1), we deduce that for any $x, v_n > 0$ and $\lambda \ge 0$,

$$\begin{aligned}&\mathbb{P}(S_k \ge x \text{ and } p[S]_k + p'\langle S\rangle_k \le v_n^2 \text{ for some } k \in [1, n]) \\ &= \mathbb{E}_\lambda[Z_{T\wedge n}(\lambda)^{-1}\mathbf{1}_{\{S_k \ge x \text{ and } p[S]_k + p'\langle S\rangle_k \le v_n^2 \text{ for some } k\in[1,n]\}}] \\ &= \sum_{k=1}^{n} \mathbb{E}_\lambda[\exp\{-\lambda S_{T\wedge n} + \Psi_{T\wedge n}(\lambda)\}\mathbf{1}_{\{T(x,v_n)=k\}}] \\ &= \sum_{k=1}^{n} \mathbb{E}_\lambda[\exp\{-\lambda S_k + \Psi_k(\lambda)\}\mathbf{1}_{\{T(x,v_n)=k\}}] \\ &\le \sum_{k=1}^{n} \mathbb{E}_\lambda[\exp\{-\lambda x + \Psi_k(\lambda)\}\mathbf{1}_{\{T(x,v_n)=k\}}],\end{aligned}$$

where

$$\Psi_k(\lambda) = \sum_{i=1}^{k} \log \mathbb{E}[e^{\lambda\xi_i - \frac{p}{2}\lambda^2\xi_i^2}|\mathcal{F}_{i-1}], \quad 0 \le k \le n.$$

By Lemma 4.1, it follows that

$$\begin{aligned}\Psi_k(\lambda) &\le \sum_{i=1}^{k} \frac{\lambda^2}{2}(p\xi_i^2 + p'\mathbb{E}[\xi_i^2|\mathcal{F}_{i-1}]) \\ &= \frac{\lambda^2}{2}(p[S]_k + p'\langle S\rangle_k).\end{aligned}$$

Using the fact that $\Psi_k(\lambda) \le \frac{\lambda^2}{2}v_n^2$ on the set $\{T(x, v_n) = k\}$, we deduce that for any $x, v_n > 0$,

$$\begin{aligned}&\mathbb{P}(S_k \ge x \text{ and } p[S]_k + p'\langle S\rangle_k \le v_n^2 \text{ for some } k \in [1, n]) \\ &\le \exp\{-\lambda x + \frac{\lambda^2}{2}v_n^2\}\mathbb{E}_\lambda[\sum_{k=1}^{n} \mathbf{1}_{\{T(x,v_n)=k\}}] \\ &= \exp\{-\lambda x + \frac{\lambda^2}{2}v_n^2\}\mathbb{E}_\lambda[\mathbf{1}_{\{S_k \ge x \text{ and } p[S]_k + p'\langle S\rangle_k \le v_n^2 \text{ for some } k\in[1,n]\}}] \\ &\le \exp\{-\lambda x + \frac{\lambda^2}{2}v_n^2\}. \end{aligned} \tag{4.2}$$

The bound (4.2) attains its minimum at

$$\lambda = \lambda(x) = \frac{x}{v_n^2},$$

Substituting $\lambda = \lambda(x)$ into (4.2), we obtain for any $x, v_n > 0$,

$$\mathbb{P}(S_k \ge x \text{ and } p[S]_k + p'\langle S\rangle_k \le v_n^2 \text{ for some } k \in [1, n]) \le \{-\frac{x^2}{2v_n^2}\}.$$

This completes the proof of theorem. □

Proof of Theorem 2.2 We follow the method of Bercu and Touati [1]. For all $t \in \mathbb{R}$, recall

$$V_n(t) = \exp\{tS_n - \frac{t^2}{2}(p[S]_n + p'\langle S\rangle_n)\}.$$

It follows from Lemma 4.2 that for all $t \ge 0$, $\mathbb{E}[V_n(t)] \le 1$. For all $x > 0$, let $A_n = \{S_n \ge x(p[S]_n + p'\langle S\rangle_n)\}$. Set $\frac{1}{\tilde p} + \frac{1}{\tilde q} = 1$, $\tilde p, \tilde q > 1$. By Hölder's inequality, we have for all $t > 0$,

$$\begin{aligned}\mathbb{P}(A_n) &= \mathbb{P}(S_n \ge x(p[S]_n + p'\langle S\rangle_n)) \\ &\le \mathbb{E}[\exp\{\frac{t}{\tilde q}(S_n - x(p[S]_n + p'\langle S\rangle_n))\}\mathbf{1}_{A_n}] \\ &= \mathbb{E}[\exp\{\frac{t}{\tilde q}(S_n - \frac{t}{2}(p[S]_n + p'\langle S\rangle_n))\}\exp\{\frac{t}{2\tilde q}(t-2x)(p[S]_n + p'\langle S\rangle_n)\}\mathbf{1}_{A_n}] \\ &= (\mathbb{E}[\exp\{\frac{\tilde p t}{2\tilde q}(t-2x)(p[S]_n + p'\langle S\rangle_n)\}\mathbf{1}_{A_n}])^{\frac{1}{\tilde p}}(\mathbb{E}[V_n(t)])^{\frac{1}{\tilde q}} \\ &\le (\mathbb{E}[\exp\{\frac{\tilde p t}{2\tilde q}(t-2x)(p[S]_n + p'\langle S\rangle_n)\}])^{\frac{1}{\tilde p}}. \end{aligned} \tag{4.3}$$

Consequently, as $\tilde p/\tilde q = \tilde p - 1$, we can deduce from (4.3) and the particular choice $t = x$ that

$$\mathbb{P}(A_n) \le \inf_{\tilde p > 1}(\mathbb{E}[\exp\{-(\tilde p - 1)\frac{x^2}{2}(p[S]_n + p'\langle S\rangle_n)\}])^{\frac{1}{\tilde p}},$$

which gives the first desired inequality. Furthermore, for all $x, y > 0$, let $B_n = \{S_n \geq x(p[S]_n + p'\langle S\rangle_n),\ p[S]_n + p'\langle S\rangle_n \geq y\}$. By Cauchy-Schwarz's inequality, it follows that for all $0 < t < 2x$,

$$\begin{aligned}\mathbb{P}(B_n) &\leq \mathbb{E}[\exp\{\frac{t}{2}S_n - \frac{t^2}{4}(p[S]_n + p'\langle S\rangle_n)\}\exp\{\frac{t}{4}(t-2x)(p[S]_n + p'\langle S\rangle_n)\}\mathbf{1}_{B_n}] \\ &\leq \exp\{\frac{ty}{4}(t-2x)\}\mathbb{E}[V_n(t)^{\frac{1}{2}}\mathbf{1}_{B_n}].\end{aligned}$$

Again by Cauchy-Schwarz's inequality, we have

$$\mathbb{E}[V_n(t)^{\frac{1}{2}}\mathbf{1}_{B_n}] \leq \sqrt{\mathbb{P}(B_n)}.$$

Therefore, we have

$$\sqrt{\mathbb{P}(B_n)} \leq \exp\{\frac{ty}{4}(t-2x)\}.$$

Choosing the value $t = x$, we obtain

$$\mathbb{P}(B_n) \leq \exp\{-\frac{x^2y}{2}\},$$

which gives the second desired inequality. □

Proof of Theorem 2.3 Denote

$$Z_n = p[S]_n + p'\langle S\rangle_n.$$

For all $x, y > 0$, let

$$C_n = \{S_n \geq x[S]_n, (2-p)[S]_n \geq p'\langle S\rangle_n + y\}.$$

Set $\frac{1}{\tilde{p}} + \frac{1}{\tilde{q}} = 1$, $\tilde{p}, \tilde{q} > 1$. By Hölder's inequality, we have the following inequality. For all $t > 0$,

$$\begin{aligned}\mathbb{P}(C_n) &\leq \mathbb{E}[\exp\{\frac{t}{\tilde{q}}S_n - \frac{tx}{\tilde{q}}[S]_n\}\mathbf{1}_{C_n}] \\ &= \mathbb{E}[\exp\{\frac{t}{\tilde{q}}S_n - \frac{t^2}{2\tilde{q}}Z_n\}\exp\{\frac{t^2}{2\tilde{q}}Z_n - \frac{tx}{\tilde{q}}[S]_n\}\mathbf{1}_{C_n}] \\ &\leq (\mathbb{E}[\exp\{\frac{t^2\tilde{p}}{2\tilde{q}}(p[S]_n + p'\langle S\rangle_n) - \frac{\tilde{p}tx}{\tilde{q}}[S]_n\}\mathbf{1}_{C_n}])^{\frac{1}{\tilde{p}}}.\end{aligned} \tag{4.4}$$

Consequently, as $\frac{\tilde{p}}{\tilde{q}} = \tilde{p} - 1$, we can deduce from (4.4) and the particular choice $t = x$ that

$$\begin{aligned}\mathbb{P}(C_n) &\leq \inf_{\tilde{p}>1}\exp\{-\frac{x^2y(\tilde{p}-1)}{2\tilde{p}}\} \\ &= \exp\{-\frac{x^2y}{2}\},\end{aligned}$$

which gives the first desired inequality. Furthermore, for all $x, y > 0$, let

$$D_n = \{S_n \geq x[S]_n,\ p'\langle S\rangle_n \leq yp[S]_n\}.$$

By Hölder's inequality, we have for all $t > 0$,

$$\begin{aligned}\mathbb{P}(D_n) &\leq \mathbb{E}[\exp\{\frac{t}{\tilde{q}}S_n - \frac{tx}{\tilde{q}}[S]_n\}\mathbf{1}_{D_n}] \\ &= \mathbb{E}[\exp\{\frac{t}{\tilde{q}}S_n - \frac{t^2}{2\tilde{q}}Z_n\}\exp\{\frac{t^2}{2\tilde{q}}Z_n - \frac{tx}{\tilde{q}}[S]_n\}\mathbf{1}_{D_n}]\end{aligned}$$

$$\le (\mathbb{E}[\exp\{\frac{t\tilde{p}}{2\tilde{q}}(tp[S]_n + tp'\langle S\rangle_n - 2x[S]_n)\}\mathbf{1}_{D_n}])^{\frac{1}{\tilde{p}}}$$
$$\le (\mathbb{E}[\exp\{\frac{t\tilde{p}}{2\tilde{q}}(tp + typ - 2x)[S]_n\}])^{\frac{1}{\tilde{p}}}. \tag{4.5}$$

Therefore, as $\frac{\tilde{p}}{\tilde{q}} = \tilde{p}-1$, we can deduce from (4.5) and the particular choice $t = x/((1+y)p)$ that

$$\mathbb{P}(D_n) \le \inf_{\tilde{p}>1}(\mathbb{E}[\exp\{-(\tilde{p}-1)\frac{x^2}{2(1+y)p}[S]_n\}])^{\frac{1}{\tilde{p}}}.$$

This completes the proof of Theorem 2.3. □

Proof of Theorem 2.5 The proof of Theorem 2.5 is based on the following two lemmas. Following the main result of de la Peña and Pang [3], we obtain the following exponential inequality for self-normalized martingales.

Lemma 4.3 *Assume $p' \ge q$. Then for any $x > 0$,*

$$\mathbb{P}(\frac{|S_n|}{\sqrt{\frac{3}{2}(p[S]_n + p'\langle S\rangle_n + s_n^2)}} \ge x) \le (\frac{3}{2})^{\frac{1}{3}} x^{-\frac{2}{3}} \exp\{-\frac{x^2}{2}\}.$$

In particular, it implies that for any $x > 0$,

$$\mathbb{P}(\frac{|S_n|}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \ge x,\ p[S]_n + p'\langle S\rangle_n \ge s_n^2) \le C\exp\{-\frac{x^2}{6}\} \tag{4.6}$$

for some positive absolute constant C.

Proof By Lemma 4.2, we have the following result

$$\mathbb{E}\Big[\exp\{tS_n - \frac{t^2}{2}(p[S]_n + p'\langle S\rangle_n)\}\Big] \le 1.$$

Put $A = S_n$ and $B^2 = p[S]_n + p'\langle S\rangle_n$. Then Lemma 4.3 follows by the main result of de la Peña and Pang [3]. □

We have the following new exponential inequality for self-normalized martingales.

Lemma 4.4 *Assume $p' \ge q$. Then for any $b > 0$, $M \ge 1$ and $x \ge 0$,*

$$\mathbb{P}(\frac{S_n}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \ge x,\ b \le \sqrt{p[S]_n + p'\langle S\rangle_n} \le bM)$$
$$\le \sqrt{e}(1 + 2(1+x)\ln M)\exp\{-\frac{x^2}{2}\}. \tag{4.7}$$

Proof We follow the method of Liptser and Spokoiny [4]. Given $a > 1$, introduce the geometric series $b_k = ba^k$ and define random events

$$A_k = \{\frac{S_n}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \ge x,\ b_k \le \sqrt{p[S]_n + p'\langle S\rangle_n} < b_{k+1}\},\quad k = 0,1,\ldots,K,$$

where K stands for the integer part of $\log_a M$. Clearly, it holds

$$\{\frac{S_n}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \ge x,\ b \le \sqrt{p[S]_n + p'\langle S\rangle_n} \le bM\} \subseteq \bigcup_{k=0}^{K} A_k. \tag{4.8}$$

For any positive t, inequality (2.2) implies that

$$\mathbb{E}[\exp\{tS_n - \frac{t^2}{2}(p[S]_n + p'\langle S\rangle_n)\Big\}\mathbf{1}_{A_k}] \leq 1.$$

Next, taking $t_k = x/b_k$, for any $x > 0$, we obtain

$$\begin{aligned}1 &\geq \mathbb{E}[\exp\{\frac{x}{b_k}S_n - \frac{x^2}{2b_k^2}(p[S]_n + p'\langle S\rangle_n)\}\mathbf{1}_{A_k}] \\ &\geq \mathbb{E}[\exp\{\frac{x^2}{b_k}\sqrt{p[S]_n + p'\langle S\rangle_n} - \frac{x^2}{2b_k^2}(p[S]_n + p'\langle S\rangle_n)\}\mathbf{1}_{A_k}] \\ &\geq \mathbb{E}[\exp\{\inf_{b_k \leq c \leq b_{k+1}}(\frac{x^2 c}{b_k} - \frac{x^2c^2}{2b_k^2})\}\mathbf{1}_{A_k}].\end{aligned}$$

Since "$\inf_{b_k\leq c\leq b_{k+1}}$" is attained at the point $c = b_{k+1} = ab_k$, we end up with

$$\mathbb{P}(A_k) \leq \exp\{-x^2(a - \frac{a^2}{2})\}.$$

Using (4.8) and the fact that $K \leq \log_a M$, we get

$$\begin{aligned}&\mathbb{P}(\frac{S_n}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \geq x,\ b \leq \sqrt{p[S]_n + p'\langle S\rangle_n} \leq bM) \leq \sum_{k=0}^{K}\mathbb{P}(A_k) \\ &\leq (1 + \log_a M)\exp\{-x^2(a - \frac{a^2}{2})\}.\end{aligned}$$

Finally, since the left-hand side of this inequality does not depend on a, we may pick a to make the right-hand side possibly small. This leads to the choice $a = 1 + \frac{1}{1+x}$, so that

$$x^2(a - \frac{a^2}{2}) = x^2\{1 + \frac{1}{1+x} - \frac{1}{2}(1 + \frac{1}{1+x})^2\} \geq \frac{1}{2}(x^2 - 1).$$

Since also $\log(1+\frac{1}{1+x}) \geq \frac{1}{2(1+x)}$ for $x \geq 0$, we obtain $\log_a M \leq 2(1+x)\ln M$ and (4.7) follows. □

Now we are in position to prove Theorem 2.5. Let $s_n^2 = b^2M^2$. Then $M = s_n/b$. Inequality (4.6) can be rewritten as

$$\mathbb{P}(\frac{S_n}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \geq x,\ \sqrt{p[S]_n + p'\langle S\rangle_n} \geq bM) \leq C\exp\{-\frac{x^2}{6}\}, \tag{4.9}$$

where C is some positive absolute constant. Therefore, combining the inequalities (4.7) and (4.9), we obtain

$$\begin{aligned}&\mathbb{P}(\frac{S_n}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \geq x,\ \sqrt{p[S]_n + p'\langle S\rangle_n} \geq b) \\ &\leq \sqrt{e}(1 + 2(1+x)\ln\frac{s_n}{b})\exp\{-\frac{x^2}{2}\} + C_1\exp\{-\frac{x^2}{6}\} \\ &\leq (\sqrt{e}(1 + 2(1+x)\ln\frac{s_n}{b})\exp\{-\frac{x^2}{3}\} + C_1)\exp\{-\frac{x^2}{6}\}.\end{aligned}$$

Here, taking $b = ts_n$, $0 < t < \frac{1}{2}$, we have

$$\begin{aligned}&\mathbb{P}(\frac{S_n}{\sqrt{p[S]_n + p'\langle S\rangle_n}} \geq x,\ \sqrt{p[S]_n + p'\langle S\rangle_n} \geq ts_n) \\ &\leq (\sqrt{e}(1 + 2(1+x)|\ln t|)\exp\{-\frac{x^2}{3}\} + C_1)\exp\{-\frac{x^2}{6}\}\end{aligned}$$

$$\le (\sqrt{e}(1+2(1+x))\exp\{-\frac{x^2}{3}\}+C_1|\ln 2|)|\ln t|\exp\{-\frac{x^2}{6}\}.$$

Let $C=\sup_{x\ge 0}(\sqrt{e}(1+2(1+x))\exp\{-\frac{x^2}{3}\}+C_2)$. Then, we obtain

$$\mathbb{P}(\frac{S_n}{\sqrt{p[S]_n+p'\langle S\rangle_n}}\ge x,\ \sqrt{p[S]_n+p'\langle S\rangle_n}\ge ts_n)$$
$$\le C|\ln t|\exp\{-\frac{x^2}{6}\},$$

which gives the desired result. □

Proof of Theorem 3.1 From (3.3), for all $n\ge 1$, $x>0$, and $y>0$, we get

$$\begin{aligned}\mathbb{P}(|\widehat{\theta_n}-\theta|\ge x)&=\mathbb{P}(|S_n|\ge \frac{x}{\sigma^2}\langle S\rangle_n)\\&\le \mathbb{P}(|S_n|\ge \frac{x}{\sigma^2}\langle S\rangle_n, p[S]_n\le yp'\langle S\rangle_n)+\mathbb{P}(p[S]_n\ge yp'\langle S\rangle_n).\end{aligned}$$

By Remark 2.4, it follows that

$$\begin{aligned}\mathbb{P}(|S_n|\ge \frac{x}{\sigma^2}\langle S\rangle_n, p[S]_n\le yp'\langle S\rangle_n)&\le 2\inf_{\tilde{p}>1}(\mathbb{E}[\exp\{-(\tilde{p}-1)\frac{x^2}{2\sigma^4(1+y)p'}\langle S\rangle_n\}])^{\frac{1}{\tilde{p}}}\\&=2\inf_{\tilde{p}>1}\exp\{\frac{n}{\tilde{p}}H(-\frac{(\tilde{p}-1)x^2}{2\sigma^2(1+y)p'})\}.\end{aligned}$$

In addition, for all $y>0$ and $0\le t\le c$,

$$\begin{aligned}\mathbb{P}(p[S]_n\ge yp'\langle S\rangle_n)&=\mathbb{P}(p\sum_{k=1}^n\phi_{k-1}^2\varepsilon_k^2\ge p'y\sigma^2\sum_{k=1}^n\phi_{k-1}^2)\\&\le \mathbb{P}(\sum_{k=1}^n\varepsilon_k^2\ge \frac{p'}{p}y\sigma^2)\le \exp\{-\frac{p'}{p}y\sigma^2 t\}\mathbb{E}\exp\{t\sum_{k=1}^n\varepsilon_k^2\}\\&\le \exp\{-\frac{p'}{p}y\sigma^2t+nL(t)\}\\&\le \exp\{-nI(\frac{p'y\sigma^2}{pn})\}.\end{aligned}$$

Then, from the above discussions, we obtain the desired result. □

Acknowledgements We thank the referees for their time and comments.

References

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

[2] B. DELYON. *Exponential inequalities for sums of weakly dependent variables.* Electronic J. Probab., 2009, **14**(28): 752–779.

[3] V. H. DE LA PEÑA, Guodong PANG. *Exponential inequalities for self-normalized processes with applications.* Electron. Commun. Probab., 2009, **14**: 372–381.

[4] R. LIPTSER, V. SPOKOINY. *Deviation probability bound for martingales with applications to statistical estimation.* Statist. Probab. Lett., 2000, **46**(4): 347–357.

[5] Xiequan FAN, I. GRAMA, Quansheng LIU. *Hoeffding's inequality for supermartingales.* Stochastic Process. Appl., 2012, **122**(10): 3545–3559.