J. Math. Anal. Appl. 486 (2020) 123902

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Cramér type moderate deviations for self-normalized ψ-mixing sequences

Xiequan Fan

Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

ARTICLE INFO

Article history:
Received 6 May 2019
Available online 28 January 2020
Submitted by V. Pozdnyakov

Keywords:
Cramér moderate deviations
Self-normalized processes
Studentized statistics
Relative error
Continued fraction expansions

ABSTRACT

Let $(\eta_i)_{i\geq 1}$ be a sequence of ψ-mixing random variables. Let $m = \lfloor n^\alpha \rfloor, 0 < \alpha < 1, k = \lfloor n/(2m) \rfloor$, and $Y_j = \sum_{i=1}^m \eta_{m(j-1)+i}, 1 \leq j \leq k$. Set $S_k^o = \sum_{j=1}^k Y_j$ and $[S^o]_k = \sum_{i=1}^k (Y_j)^2$. We prove a Cramér type moderate deviation expansion for $\mathbb{P}(S_k^o/\sqrt{[S^o]_k} \geq x)$ as $n \to \infty$. Our result is similar to the recent work of Chen et al. (2016) [4] where the authors established Cramér type moderate deviation expansions for β-mixing sequences. Comparing to the result of Chen et al., our results hold for mixing coefficients with polynomial decaying rate and wider ranges of validity.

1. Introduction

The study of the relative errors for Gaussian approximations can be traced back to Cramér [6]. Let $(X_i)_{i\geq 1}$ be a sequence of independent and identically distributed (i.i.d.) centered real random variables satisfying the condition $\mathbb{E}\exp\{c_0|X_1|\} < \infty$ for some constant $c_0 > 0$. Denote $\sigma^2 = \mathbb{E}X_1^2$ and $S_n = \sum_{i=1}^n X_i$. Cramér established the following asymptotic moderate deviation expansion on the tail probabilities of S_n: For all $0 \leq x = o(n^{1/2})$,

$$\left| \ln \frac{\mathbb{P}(S_n \geq x\sigma\sqrt{n})}{1-\Phi(x)} \right| = O(1)\frac{(1+x)^3}{\sqrt{n}} \quad \text{as} \quad n \to \infty, \tag{1.1}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x \exp\{-t^2/2\}dt$ is the standard normal distribution. In particular, inequality (1.1) implies that

$$\frac{\mathbb{P}(S_n \geq x\sigma\sqrt{n})}{1-\Phi(x)} = 1 + o(1) \tag{1.2}$$

E-mail address: fanxiequan@hotmail.com.

https://doi.org/10.1016/j.jmaa.2020.123902

uniformly for $0 \leq x = o(n^{1/6})$. Following the seminal work of Cramér, various moderate deviation expansions for standardized sums have been obtained by many authors (see, for instance, Petrov [22,23], Linnik [20], Saulis and Statulevičius [26] and [11,13]). See also Račkauskas [24,25], Grama [16], Grama and Haeusler [17] and Fan, Grama and Liu [12] for martingales.

To establish moderate deviation expansions type of (1.2) for $0 \leq x = o(n^{\alpha}), \alpha > 0$, we should assume that the random variables have finite moments of any order, see Linnik [20]. The last assumption becomes too restrictive if we only have finite moments of order $2+\delta, \delta \in (0,1]$. Thought we still can obtain (1.2) via Berry-Esseen estimations, the range cannot wider than $|x| = O(\sqrt{\ln n}), n \to \infty$. To overcome this shortcoming, a new type Cramér type moderate deviations (CMD), called self-normalized CMD, has been developed by Shao [27]. Instead of considering the moderate deviations for standardized sums $S_n/\sqrt{n\sigma^2}$, Shao [27] considered the moderate deviations for self-normalized sums $W_n := S_n/\sqrt{\sum_{i=1}^n X_i^2}$. Comparing to the standardized counterpart, the range of Gaussian approximation for self-normalized CMD can be much wider range than its counterpart for standardized sums under same finite moment conditions. Moreover, in practice one usually does not known the variance of S_n. Even the latter can be estimated, it is still advisable to use self-normalized CMD for more user-friendly. Due to these significant advantages, the study of CMD for self-normalized sums attracts more and more attentions. For more self-normalized CMD for independent random variables, we refer to, for instance, Jing, Shao and Wang [18] and Liu, Shao and Wang [21]. We also refer to de la Peña, Lai and Shao [9], Shao and Wang [29] and Shao [28] for recent developments in this area. For closely related results, see also de la Peña [8] and Bercu and Touati [2] for exponential inequalities for self-normalized martingales.

Thought self-normalized CMD for independent random variables has been well studied, there are only a few of results for weakly dependent random variables. One of the main results in this field is due to Chen et al. [4]. Let $(\eta_i)_{i\geq 1}$ be a (may be non-stationary) sequence of random variables. Set $\alpha \in (0,1)$. Let $m = \lfloor n^{\alpha} \rfloor$ and $k = \lfloor n/(2m) \rfloor$, where $\lfloor x \rfloor$ denote the integer part of x. Denote

$$Y_j = \sum_{i=1}^{m} \eta_{2m(j-1)+i}, \quad 1 \leq j \leq k.$$

Set

$$S_k^o = \sum_{j=1}^{k} Y_j \qquad \text{and} \qquad [S^o]_k = \sum_{j=1}^{k} (Y_j)^2.$$

Define the interlacing self-normalized sums as follows

$$W_n^o = S_k^o / \sqrt{[S^o]_k}. \tag{1.3}$$

Let $\mathcal{F}_j$ and $\mathcal{F}_{j+k}^{\infty}$ be σ-fields generated respectively by $(\eta_i)_{i\leq j}$ and $(\eta_i)_{i\geq j+k}$. The sequence of random variables $(\eta_i)_{i\geq 1}$ is called β-*mixing* if the mixing coefficient

$$\beta(n) := \sup_j \mathbb{E} \sup\{ \big| \mathbb{P}(A|\mathcal{F}_j) - \mathbb{P}(A) \big| : \ A \in \mathcal{F}_{j+n}^{\infty} \} \to 0 \qquad \text{as } n \to \infty, \tag{1.4}$$

see Doukhan [10]. Write

$$S_{l,s} = \sum_{i=l+1}^{l+s} \eta_i$$

the block sums of $(\eta_i)_{i\geq 1}$ for $l+1 \leq i \leq l+s$. Throughout the paper, denote c, probably supplied with some indices, a generic positive constant. Assume that $(\eta_i)_{i\geq 1}$ are centered, that is

$$\mathbb{E}\eta_i = 0 \quad \text{for all } i, \tag{1.5}$$

and that there exists a constant $\nu \in (0,1]$ such that

$$\mathbb{E}|\eta_i|^{2+\nu} \leq c_0^{2+\nu} \tag{1.6}$$

and

$$\mathbb{E}S_{l,s}^2 \geq c_1^2 s \qquad \text{for all } l \geq 0 \text{ and } s \geq 1. \tag{1.7}$$

By Theorem 4.1 of Shao and Yu [30], it is known that condition (1.6) usually implies the following condition: there exists a constant $\rho \in (0,1]$ such that

$$\mathbb{E}|S_{l,s}|^{2+\rho} \leq s^{1+\rho/2} c_2^{2+\rho} \qquad \text{for all } l \geq 0 \text{ and } s \geq 1, \tag{1.8}$$

provided that the mixing coefficient has a polynomially decaying rate as $n \to \infty$. In (1.8), it is usually that $\rho < \nu$. Assume conditions (1.5)-(1.7). Assume also that there exist positive constants a_1, a_2 and τ such that

$$\beta(n) \leq a_1 e^{-a_2 n^{\tau}}.$$

Using m-dependent approximation, Chen et al. [4] proved that for any positive $\rho < \nu$,

$$\left| \ln \frac{\mathbb{P}(W_n^o \geq x)}{1 - \Phi(x)} \right| \leq c_\rho \left(\frac{(1+x)^{2+\rho}}{n^{(1-\alpha)\rho/2}} \right) \tag{1.9}$$

uniform for $0 \leq x = o(\min\{n^{(1-\alpha)/2}, n^{\alpha\tau/2}\})$, where c_ρ depends only on c_0, c_1, ρ, a_1, a_2 and τ. In particular, it implies that

$$\frac{\mathbb{P}(W_n^o \geq x)}{1 - \Phi(x)} = 1 + o(1) \tag{1.10}$$

uniformly for $0 \leq x = o(\min\{n^{(1-\alpha)\rho/(4+2\rho)}, n^{\alpha\tau/2}\})$. Equality (1.10) implies that the tail probabilities of W_n^o can be uniformly approximated by the standard normal distribution for moderate x's. Such type of results play an important role in statistical inference of means, see Section 5 of Chen et al. [4] for applications. Inspiring the proof of Chen et al. [4], it is easy to see that (1.9) remains valid when the conditions (1.5)-(1.7) are replaced by the slightly more general conditions (1.5), (1.7) and (1.8).

In this paper, we are interested to extend the results of Chen et al. [4] to ψ-mixing sequences, with conditions (1.5), (1.7) and (1.8). By Proposition 1 in Doukhan [10], it is known that ψ-mixing usually implies β-mixing. However, the ranges of our results do not depend on the mixing coefficients. Indeed, our ranges of validity for (1.9) and (1.10) are respectively $0 \leq x = o(n^{(1-\alpha)/2})$ and $0 \leq x = o(n^{(1-\alpha)\rho/(4+2\rho)})$ as $n \to \infty$, which are the best possible even $(\eta_i)_{i\geq 1}$ are independent. Moreover, we show that (1.10) remains true if ψ-mixing coefficient $\psi(n)$ decays in a polynomial decaying rate, in contrast to β-mixing sequences which does not share this property. For methodology, our approach is based on martingale approximation and self-normalized Cramér type moderate deviations for martingales due to Fan et al. [15].

The paper is organized as follows. Our main results are stated and discussed in Section 2. Applications and simulation study are given in Section 3. Proofs of results are deferred to Section 4.

2. Main results

Recall that $\mathcal{F}_j$ and $\mathcal{F}_{j+k}^{\infty}$ be σ-fields generated respectively by $(\eta_i)_{i\le j}$ and $(\eta_i)_{i\ge j+k}$. We say that $(\eta_i)_{i\ge 1}$ is *ψ-mixing* if the mixing coefficient

$$\psi(n) := \sup_j \sup_A \{ \big|\mathbb{P}(A|\mathcal{F}_j) - \mathbb{P}(A)\big| / \mathbb{P}(A) : \ A \in \mathcal{F}_{j+n}^{\infty} \} \to 0 \qquad \text{as } n \to \infty, \tag{2.11}$$

see Doukhan [10]. It is known that continued fraction expansions of irrational numbers and certain Gibbs-Markov dynamical systems are ψ-mixing, see Bazarova, Berkes and Horváth [1] and Denker and Kabluchko [7] respectively.

Our main result is the following self-normalized Cramér type moderate deviations for ψ-mixing sequences.

Theorem 2.1. *Assume conditions (1.5), (1.7) and (1.8). Set $\alpha \in (0,1)$. Let $m = \lfloor n^{\alpha} \rfloor$ and $k = \lfloor n/(2m) \rfloor$ be respectively the integers part of n^{α} and $n/(2m)$. Denote*

$$\delta_n^2 = m\psi^2(m) + k\psi(m)$$

and

$$\gamma_n = k^{1/2}\psi^{1/2}(m) + n\psi(m).$$

Assume also that $\delta_n, \gamma_n \to 0$ as $n \to \infty$.

[i] *If $\rho \in (0,1)$, then for all $0 \le x = o(n^{(1-\alpha)/2})$,*

$$\left| \ln \frac{\mathbb{P}(W_n^o \ge x)}{1 - \Phi(x)} \right| \le c_\rho \left(\frac{x^{2+\rho}}{n^{(1-\alpha)\rho/2}} + x^2\delta_n^2 + (1+x)\big(\frac{1}{n^{(1-\alpha)\rho(2-\rho)/8}(1 + x^{\rho(2+\rho)/4})} + \gamma_n \big) \right), \tag{2.12}$$

where c_ρ depends only on c_1, c_2 and ρ.

[ii] *If $\rho = 1$, then for all $0 \le x = o(n^{(1-\alpha)/2})$,*

$$\left| \ln \frac{\mathbb{P}(W_n^o \ge x)}{1 - \Phi(x)} \right| \le c \left(\frac{x^3}{n^{(1-\alpha)/2}} + x^2\delta_n^2 + (1+x)\big(\frac{1}{n^{(1-\alpha)/8}(1 + x^{3/4})} + \frac{\ln n}{n^{(1-\alpha)/2}} + \gamma_n \big) \right), \tag{2.13}$$

where c depends only on c_1 and c_2.

Notice that in the i.i.d. case, W_n^o is self-normalized sums of k i.i.d. random variables, that is $(Y_i)_{1\le i\le k}$. According to the classical result of Jing, Shao and Wang [18], Cramér type moderate deviations holds for $0 \le x = o(k^{1/2})$. Since the last range is equivalent to the range $0 \le x = o(n^{(1-\alpha)/2})$, the ranges of validity for (2.12) and (2.13) coincide with the case of i.i.d., and, therefore, it is the best possible.

The following MDP result is a consequence of the last theorem.

Corollary 2.1. *Assume the conditions of Theorem 2.1. Let a_n be any sequence of real numbers satisfying $a_n \to \infty$ and $a_n/n^{(1-\alpha)/2} \to 0$ as $n \to \infty$. Then for each Borel set $B \subset \mathbb{R}$,*

$$
\begin{aligned}
-\inf_{x\in B^o}\frac{x^2}{2} &\le \liminf_{n\to\infty}\frac{1}{a_n^2}\ln\mathbb{P}\left(\frac{1}{a_n}W_n^o\in B\right)\\
&\le \limsup_{n\to\infty}\frac{1}{a_n^2}\ln\mathbb{P}\left(\frac{1}{a_n}W_n^o\in B\right)\le -\inf_{x\in\overline{B}}\frac{x^2}{2},
\end{aligned}
$$

where B^o and $\overline{B}$ denote the interior and the closure of B, respectively.

If $\psi(n)=O\big(n^{-(1+\rho)/\alpha}\big)$, then $\delta_n^2=o\big(n^{-(1-\alpha)\rho/2}\big)$ and $\gamma_n=o\big(n^{-(1-\alpha)\rho/2}\big)$. The following corollary is nonetheless worthy to state.

Corollary 2.2. *Assume conditions (1.5), (1.7) and (1.8). Set $\alpha\in(0,1)$. Assume also that*

$$\psi(n)=O\big(n^{-(1+\rho)/\alpha}\big)$$

as $n\to\infty$.

[i] *If $\rho\in(0,1)$, then for all $0\le x=o(n^{(1-\alpha)/2})$,*

$$\left|\ln\frac{\mathbb{P}(W_n^o\ge x)}{1-\Phi(x)}\right|\le c_\rho\left(\frac{x^{2+\rho}}{n^{(1-\alpha)\rho/2}}+\frac{1+x}{n^{(1-\alpha)\rho(2-\rho)/8}(1+x^{\rho(2+\rho)/4})}\right),\tag{2.14}$$

where c_ρ depends only on c_1,c_2 and ρ.

[ii] *If $\rho=1$, then for all $0\le x=o(n^{(1-\alpha)/2})$,*

$$\left|\ln\frac{\mathbb{P}(W_n^o\ge x)}{1-\Phi(x)}\right|\le c\left(\frac{x^3}{n^{(1-\alpha)/2}}+(1+x)\big(\frac{1}{n^{(1-\alpha)/8}(1+x^{3/4})}+\frac{\ln n}{n^{(1-\alpha)/2}}\big)\right),\tag{2.15}$$

where c depends only on c_1 and c_2.

In particular, (2.14) and (2.15) together implies that for $\rho\in(0,1]$,

$$\frac{\mathbb{P}(W_n^o>x)}{1-\Phi(x)}=1+o(1)\tag{2.16}$$

uniformly for $0\le x=o(n^{(1-\alpha)\rho/(4+2\rho)})$.

Chen et al. [4] (see Section 3 therein) showed that if β-mixing coefficient $\beta(n)$ decays only polynomial slowly, then (2.16) is not valid at $x=(C\ln n)^{1/2}$ for sufficiently large constant C. However, Theorem 2.1 shows that the range of validity of (2.16) can be much wider when β-mixing is replaced by ψ-mixing.

Recall that in the i.i.d. case, W_n^o is self-normalized sums of k i.i.d. random variables. By Remark 2 of Shao [27], the range of validity for (2.16) is also the best possible.

Remark 2.1. Notice that if $(\eta_i)_{i\ge1}$ satisfies conditions (1.5), (1.7) and (1.8), then $(-\eta_i)_{i\ge1}$ also satisfies the same conditions. Thus the assertions in Theorem 2.1 and Corollary 2.2 remain valid when $\dfrac{\mathbb{P}(W_n^o\ge x)}{1-\Phi(x)}$ is replaced by $\dfrac{\mathbb{P}(W_n^o\le -x)}{\Phi(-x)}$.

3. Applications

3.1. Application to simultaneous confidence intervals

Consider the problem of constructing simultaneous confidence intervals for the mean value μ of the random variables $(\zeta_i)_{i\geq 1}$. Assume that $(\zeta_i - \mu)_{i\geq 1}$ satisfies the conditions (1.5), (1.7) and (1.8). Let

$$T_n = \frac{\sum_{j=1}^{k}(Y_j - m\mu)}{\sqrt{\sum_{j=1}^{k}(Y_j - \overline{Y}_j)^2}},$$

where $m = \lfloor n^\alpha \rfloor, k = \lfloor n/(2m) \rfloor$, $Y_j = \sum_{i=1}^{m} \zeta_{2m(j-1)+i}$, $1 \leq j \leq k$, and $\overline{Y}_j = k^{-1}\sum_{j=1}^{k} Y_j$.

Corollary 3.1. *Let $\delta_n \in (0,1)$. Assume that*

$$\big|\ln \delta_n\big| = o\big(n^{(1-\alpha)\rho/(2+\rho)}\big). \tag{3.17}$$

If $\psi(n) = O\big(n^{-(1+\rho)/\alpha}\big)$, $n\to\infty$, then

$$\frac{\sum_{j=1}^{k} Y_j}{km} \pm \frac{\Phi^{-1}(1-\delta_n/2)}{km}\sqrt{\sum_{j=1}^{k}(Y_j - \overline{Y}_j)^2}$$

is $1-\delta_n$ conservative simultaneous confidence intervals for μ.

Proof. It is known that for all $x \geq 0$,

$$\mathbb{P}\Big(T_n \geq x\Big) = \mathbb{P}\left(\frac{\sum_{j=1}^{k}(Y_j - m\mu)}{\sqrt{\sum_{j=1}^{k}(Y_j - m\mu)^2}} \geq x\Big(\frac{k}{k-1}\Big)^{1/2}\Big(\frac{k}{k+x^2-1}\Big)^{1/2}\right),$$

see Chung [5]. The last equality and (2.16) together implies that

$$\frac{\mathbb{P}(T_n \geq x)}{1-\Phi(x)} = 1 + o(1) \tag{3.18}$$

uniformly for $0 \leq x = o(n^{(1-\alpha)\rho/(4+2\rho)})$. Clearly, the upper $(\delta_n/2)$th quartile of a standard normal distribution $\Phi^{-1}(1-\delta_n/2)$ satisfies

$$\Phi^{-1}(1-\delta_n/2) = O(\sqrt{|\ln \delta_n|}),$$

which, by (3.17), is of order $o(n^{(1-\alpha)\rho/(4+2\rho)})$. Then applying the last equality to T_n, we complete the proof of Corollary 3.1. □

Similar results in statistical inference for hight-dimensional time series can be found in Chen et al. [4], where the authors have established simultaneous confidence intervals for functional dependence sequences.

3.2. Application to continued fraction and simulation study

One of the well known example of ψ-mixing sequences is called continued fraction expansions of irrational numbers on $(0,1)$. For an irrational number $x \in (0,1)$, let

$$a_1(x) = \lfloor 1/x \rfloor,\ a_{n+1}(x) = a_1(x \circ T^n), \qquad n \geq 1,$$

be the continued fraction expansion of x, where T is defined by $T(x) = 1/x - \lfloor 1/x \rfloor$, that is the fractional part of $1/x$. It is easy to see that

$$x = \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \cfrac{1}{a_3(x) + \cfrac{1}{\cdots}}}}$$

The sequence $(a_n(x))_{n\geq 1}$ with respect to the uniform measure in $(0,1)$ is ψ-mixing. Indeed, Lévy [19] proved that

$$\psi(n) = \sup_j \sup_A \{ \big|\mathbb{P}(A|\mathcal{F}_j) - \mathbb{P}(A)\big| / \mathbb{P}(A) : \ A \in \mathcal{F}_{j+n}^{\infty} \} \leq C e^{-\lambda n} \tag{3.19}$$

with positive absolute constants C and λ, where $\mathcal{F}_1^j$ and $\mathcal{F}_{j+n}^{\infty}$ be σ-fields generated respectively by $(a_i(x))_{1\leq i\leq j}$ and $(a_i(x))_{i\geq j+n}$. Denote by

$$\mathbb{G}(E) = \frac{1}{\ln 2} \int_E \frac{1}{1+x} dx,$$

the Gauss measure on the class of Borel subsets $\mathcal{B}$ of $(0,1)$. It is known that (cf. Billingsley [3]) T is an ergodic transformation preserving the Gauss measure and thus $(a_n(x))_{n\geq 1}$ is a stationary ergodic sequence with respect to the probability space $((0,1), \mathcal{B}, \mathbb{G})$. Clearly, the set $\{a_1 = k\}$ is the interval $(1/(k+1), 1/k]$ and thus

$$\mathbb{G}(\{a_1 = k\}) = \frac{1}{\ln 2} \int_{1/(k+1)}^{1/k} \frac{1}{1+x} dx = \frac{1}{\ln 2} \ln\Big(1 + \frac{1}{k(k+2)}\Big).$$

Hence, by the ergodic theorem we have for any function $F : \mathbb{N} \to \mathbb{R}$, it holds

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} F(a_k(x)) = \frac{1}{\ln 2} \sum_{j=1}^{\infty} F(j) \ln\Big(1 + \frac{1}{j(j+2)}\Big) \quad \text{a.e.} \tag{3.20}$$

whenever the series on the right hand side converges absolutely. Recently, Bazarova, Berkes and Horváth [1] gave a central limit theorem for $(a_n(x))_{n\geq 1}$. Next, we give self-normalized Cramér type moderate deviations.

Letting $\mathbb{E}$ denote expectation with respect to $\mathbb{G}$, by (3.20), we have $\mathbb{E}a_1(x) = \infty$ and $\mathbb{E}(a_1(x))^{\alpha} < \infty$ for any $\alpha \in (0,1)$. Consider the self-normalized moderate deviation for the random variables $(\zeta_i)_{i\geq 1}$, where $\zeta_i = \sqrt[3]{a_i(x)}$ for any i. Then $\mathbb{E}\zeta_1^{2+\rho} < \infty$ for any $\rho \in (0,1)$ and

$$\mu := \mathbb{E}\zeta_i = \frac{1}{\ln 2} \sum_{j=1}^{\infty} j^{1/3} \ln\Big(1 + \frac{1}{j(j+2)}\Big). \tag{3.21}$$

Let

$$W_n^o = \frac{\sum_{j=1}^k (Y_j - m\mu)}{\sqrt{\sum_{j=1}^k (Y_j - m\mu)^2}},$$

where $m = \lfloor n^\alpha \rfloor, k = \lfloor n/(2m) \rfloor$, $Y_j = \sum_{i=1}^m \zeta_{2m(j-1)+i}$, $1 \le j \le k$. By (2.16), we have the following result.

Corollary 3.2. *Set $\alpha \in (0,1)$. Then for any $\rho \in (0,1)$,*

$$\frac{\mathbb{P}(W_n^o \ge t)}{1 - \Phi(t)} = 1 + o(1) \tag{3.22}$$

uniformly for $0 \le t = o(n^{(1-\alpha)\rho/(4+2\rho)})$.

Next, we give a simulation study for the last corollary. We let $n = 30$, $m = 1, 2, 3, 4$ and consider 13 levels of $t : t = 0, .1, .2, ..., 1.0, 1.2, 1.4$. Let x be the discrete uniform distribution random variable, with possible values $\pi/10000, 2\pi/10000, ..., 3182\pi/10000$. Since π is an irrational number, x are irrational numbers. In W_n^o, we take

$$\mu = \frac{1}{\ln 2} \sum_{j=1}^{300} j^{1/3} \ln\Big(1 + \frac{1}{j(j+2)}\Big). \tag{3.23}$$

Then $\mathbb{P}(W_n^o \ge t) \approx \#(W_n^o : W_n^o \ge t)/3182$. The following table shows the simulate rations $\frac{\mathbb{P}(W_n^o \ge t)}{1-\Phi(t)}$. From the table, we see that the interlacing self-normalized sums (that is $m = 2, 3, 4$) has a better performance than self-normalized sums (that is $m = 1$) when x close to 0. When x moves away from 0, the reverse is true.

m	$t=0$	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0	1.2	1.4
1	1.11	1.13	1.15	1.16	1.17	1.16	1.11	1.08	1.03	0.96	0.90	0.75	0.53
2	1.01	1.02	1.02	1.02	1.02	1.01	1.00	0.99	0.94	0.88	0.78	0.57	0.42
3	1.00	1.03	1.04	1.07	1.06	1.06	1.04	1.01	0.98	0.92	0.85	0.67	0.48
4	1.01	1.00	0.99	0.96	0.94	0.89	0.82	0.74	0.67	0.56	0.46	0.29	0.13

4. Proofs

To shorten notations, for two real positive sequences $(a_n)_{i\ge1}$ and $(b_n)_{i\ge1}$, write $a_n \preceq b_n$ if there exists a positive constant C such that $a_n \le Cb_n$ holds for all large n, $a_n \succeq b_n$ if $b_n \preceq a_n$, and $a_n \asymp b_n$ if $a_n \preceq b_n$ and $b_n \preceq a_n$.

4.1. Preliminary lemmas

Let $(X_i, \mathcal{F}_i)_{i=0,...,n}$ be a sequence of martingale differences defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Set

$$S_0 = 0, \qquad S_k = \sum_{i=1}^k X_i, \quad k = 1, ..., n. \tag{4.24}$$

Then $(S_k, \mathcal{F}_k)_{k=0,...,n}$ is a martingale. Denote $B_n^2 = \sum_{i=1}^n \mathbb{E}X_i^2$ the variance of S_n. We assume the following conditions:

(A1) There exists $\varsigma_n \in [0, \frac{1}{4}]$ such that

$$\Big| \sum_{i=1}^{n} \mathbb{E}[X_i^2|\mathcal{F}_{i-1}] - B_n^2 \Big| \leq \varsigma_n^2 B_n^2;$$

(A2) There exist $\rho \in (0, 1]$ and $\tau_n \in (0, \frac{1}{4}]$ such that

$$\mathbb{E}[|X_i|^{2+\rho}|\mathcal{F}_{i-1}] \leq (\tau_n B_n)^\rho \, \mathbb{E}[X_i^2|\mathcal{F}_{i-1}], \quad 1 \leq i \leq n.$$

In practice, we usually have $\varsigma_n, \tau_n \to 0$ as $n \to \infty$. In the case of sums of i.i.d. random variables with finite $(2+\rho)$th moments, then it holds $B_n \asymp \sqrt{n}$, and thus conditions (A1) and (A2) are satisfied with $\varsigma_n = 0$ and $\tau_n = O(1/\sqrt{n})$ as $n \to \infty$.

Define the self-normalized martingales

$$W_n = \frac{S_n}{\sqrt{\sum_{i=1}^{n} X_i^2}}. \tag{4.25}$$

The proof of Theorem 2.1 is based on the following technical lemma due to Fan et al. [15] (see Corollary 2.3 therein), which gives a Cramér type moderate deviation expansion for self-normalized martingales.

Lemma 4.1. *Assume conditions (A1) and (A2). Denote*

$$\widehat{\tau}_n(x, \rho) = \frac{\tau_n^{\rho(2-\rho)/4}}{1 + x^{\rho(2+\rho)/4}}. \tag{4.26}$$

[i] *If $\rho \in (0, 1)$, then for all $0 \leq x = o(\tau_n^{-1})$,*

$$\left| \ln \frac{\mathbb{P}(W_n \geq x)}{1 - \Phi(x)} \right| \leq c_\rho \bigg(x^{2+\rho} \tau_n^\rho + x^2 \varsigma_n^2 + (1 + x)\Big(\varsigma_n + \widehat{\tau}_n(x, \rho)\Big) \bigg),$$

where c_ρ depends only on ρ.

[ii] *If $\rho = 1$, then for all $0 \leq x = o(\gamma_n^{-1})$,*

$$\left| \ln \frac{\mathbb{P}(W_n \geq x)}{1 - \Phi(x)} \right| \leq c \bigg(x^3 \tau_n + x^2 \varsigma_n^2 + (1 + x)\Big(\varsigma_n + \tau_n |\ln \tau_n| + \widehat{\tau}_n(x, 1)\Big) \bigg),$$

where c is a constant.

The following lemma is useful in the proof of Theorem 2.1, see Theorem 2.2 of Fan et al. [14]. Denote $x^+ = \max\{x, 0\}$ and $x^- = (-x)^+$ the positive and negative parts of x, respectively.

Lemma 4.2. *Assume that $\mathbb{E}|X_i|^\beta < \infty$ for a constant $\beta \in (1, 2]$ and all $i \in [1, n]$. Write*

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

Then for all $x, v > 0$,

$$\mathbb{P}\left(S_k \geq x \textit{ and } G_k^0(\beta) \leq v^\beta \textit{ for some } k \in [1, n]\right) \leq \exp\left\{ -C(\beta) \left(\frac{x}{v}\right)^{\beta/(\beta-1)} \right\}, \tag{4.27}$$

where

$$C(\beta) = \beta^{\frac{1}{1-\beta}}\left(1-\beta^{-1}\right). \tag{4.28}$$

In the proof of Theorem 2.1, we also make use of the following lemma which can be found in Theorem 3 of Doukhan [10].

Lemma 4.3. *Suppose that X and Y are random variables which are $\mathcal{F}_{j+n}^{\infty}$- and $\mathcal{F}_j$-measurable, respectively, and that $\mathbb{E}|X| < \infty$, $\mathbb{E}|Y| < \infty$. Then*

$$\Big|\mathbb{E}XY - \mathbb{E}X\mathbb{E}Y\Big| \le \psi(n)\,\mathbb{E}|X|\,\mathbb{E}|Y|.$$

Moreover, since $\mathbb{E}|X| \le (\mathbb{E}|X|^2)^{1/2}$, it holds

$$\Big|\mathbb{E}XY - \mathbb{E}X\mathbb{E}Y\Big| \le \psi(n)\,(\mathbb{E}X^2)^{1/2}\,(\mathbb{E}Y^2)^{1/2}$$

provided that $\mathbb{E}X^2 < \infty$ and $\mathbb{E}Y^2 < \infty$.

4.2. Proof of Theorem 2.1

We first prove Theorem 2.1 for $\rho \in (0,1)$. Denote by $\mathcal{F}_l = \sigma\{\eta_i, 1 \le i \le 2ml - m\}$. Then Y_j is $\mathcal{F}_j$-measurable. Since $\mathbb{E}\eta_i = 0$ for all i, by the definition of mixing coefficient (2.11), it is easy to see that for $1 \le j \le k$,

$$\begin{aligned}
\Big|\mathbb{E}[Y_j|\mathcal{F}_{j-1}]\Big| &= \Big|\sum_{i=1}^{m}\Big(\mathbb{E}[\eta_{2m(j-1)+i}|\mathcal{F}_{j-1}] - \mathbb{E}\eta_{2m(j-1)+i}\Big)\Big| \\
&\le \sum_{i=1}^{m}\psi(m)\mathbb{E}|\eta_{2m(j-1)+i}| \\
&\le \sum_{i=1}^{m}\psi(m)(\mathbb{E}|\eta_{2m(j-1)+i}|^{2+\rho})^{1/(2+\rho)} \\
&\le m\psi(m)c_2,
\end{aligned} \tag{4.29}$$

where the last inequality follows by condition (1.8) with $s = 1$. Thus

$$\Big|\sum_{j=1}^{k}\mathbb{E}[Y_j|\mathcal{F}_{j-1}]\Big| \le km\psi(m)c_2 \le n\psi(m)c_2.$$

By condition (1.8) and the inequality

$$(x+y)^p \le 2^{p-1}(x^p+y^p) \qquad \text{for } x, y \ge 0 \text{ and } p \ge 1,$$

we have

$$\begin{aligned}
\mathbb{E}[|Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^{2+\rho}|\mathcal{F}_{j-1}] &\le 2^{1+\rho}\mathbb{E}[|Y_j|^{2+\rho} + |\mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^{2+\rho}|\mathcal{F}_{j-1}] \\
&\le 2^{2+\rho}\mathbb{E}[|Y_j|^{2+\rho}|\mathcal{F}_{j-1}] \\
&\le 2^{2+\rho}(1+\psi(m))\mathbb{E}|Y_j|^{2+\rho} \\
&\le 2^{2+\rho}(1+\psi(m))m^{1+\rho/2}c_2^{2+\rho}.
\end{aligned} \tag{4.30}$$

The last inequality implies that

$$\begin{aligned}\mathbb{E}[|Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^2|\mathcal{F}_{j-1}] &\leq (\mathbb{E}[|Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^{2+\rho}|\mathcal{F}_{j-1}])^{2/(2+\rho)} \\ &\leq 2^2(1+\psi(m))^{2/(2+\rho)} m c_2^2 \\ &\leq 2^2(1+\psi(m)) m c_2^2. \end{aligned} \tag{4.31}$$

Similarly, by (1.8) and the assumption $\delta_n \to 0$ as $n \to \infty$, it holds

$$\begin{aligned}\mathbb{E}[|Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^2|\mathcal{F}_{j-1}] &= \mathbb{E}[Y_j^2|\mathcal{F}_{j-1}] - |\mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^2 \\ &\geq (1-\psi(m))\mathbb{E}Y_j^2 - (m\psi(m)c_2)^2 \\ &\succeq \frac{1}{2}(1-\psi(m)) m c_1^2. \end{aligned} \tag{4.32}$$

Combining (4.30)-(4.32), we deduce that

$$\mathbb{E}[|Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^{2+\rho}|\mathcal{F}_{j-1}] \preceq m^{\rho/2}\mathbb{E}[|Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^2|\mathcal{F}_{j-1}],$$
$$\sum_{j=1}^{k} \mathbb{E}[|Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^2|\mathcal{F}_{j-1}] \asymp n$$

and, by Lemma 4.3 and (4.29),

$$\begin{aligned}&\Big|\sum_{j=1}^{k} \mathbb{E}[|Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^2|\mathcal{F}_{j-1}] - \mathbb{E}S_n^2\Big| \\ &\leq \Big|\sum_{j=1}^{k} \mathbb{E}[|Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]|^2|\mathcal{F}_{j-1}] - \sum_{j=1}^{k}\mathbb{E}Y_j^2\Big| + \Big|\mathbb{E}S_n^2 - \sum_{j=1}^{k}\mathbb{E}Y_j^2\Big| \\ &\leq \sum_{j=1}^{k}\Big|\mathbb{E}[Y_j^2|\mathcal{F}_{j-1}] - \mathbb{E}Y_j^2\Big| + \sum_{j=1}^{k}\Big|\mathbb{E}[Y_j|\mathcal{F}_{j-1}]\Big|^2 + 2\sum_{j=1}^{k}\sum_{l=1}^{j-1}\Big|\mathbb{E}Y_jY_l\Big| \\ &\leq k\psi(m)\mathbb{E}Y_j^2 + k(m\psi(m)c_2)^2 + 2\psi(m)\sum_{j=1}^{k}\sum_{l=1}^{j-1}\mathbb{E}|Y_j|\,\mathbb{E}|Y_l| \\ &\leq 2n\psi(m)c_2^2 + nm\psi^2(m)c_2^2 + 2\psi(m)\sum_{j=1}^{k}\sum_{l=1}^{j-1}\sqrt{\mathbb{E}Y_j^2}\,\sqrt{\mathbb{E}Y_l^2} \\ &\leq 2n\psi(m)c_2^2 + nm\psi^2(m)c_2^2 + 2nk\psi(m)c_2^2 \\ &\leq nm\psi^2(m)c_2^2 + 4nk\psi(m)c_2^2. \end{aligned}$$

Denote by

$$\delta_n^2 = m\psi^2(m) + k\psi(m).$$

Taking $X_j = Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]$, we find that the conditions (A1) and (A2) are satisfied with $B_n^2 = \mathbb{E}S_n^2 \asymp n$, $\varsigma_n \asymp \delta_n$ and $\tau_n \asymp \sqrt{m/n} \asymp n^{-(1-\alpha)/2}$. Applying Lemma 4.1 to

$$W_n := \frac{\sum_{j=1}^{k}(Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}])}{\sqrt{\sum_{j=1}^{k}(Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}])^2}},$$

we have for all $0 \le x = o(n^{(1-\alpha)/2})$,

$$\left| \ln \frac{\mathbb{P}(W_n \ge x)}{1-\Phi(x)} \right| \le c_\rho \left(\frac{x^{2+\rho}}{n^{(1-\alpha)\rho/2}} + x^2\delta_n^2 + (1+x)\Big(\frac{1}{n^{(1-\alpha)\rho(2-\rho)/8}(1+x^{\rho(2+\rho)/4})} + \delta_n \Big) \right). \tag{4.33}$$

Notice that, by Cauchy-Schwarz's inequality,

$$\begin{aligned} \Big| \sum_{j=1}^k \big(Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]\big)^2 - \sum_{j=1}^k Y_j^2 \Big| &\le 2\sum_{j=1}^k |Y_j \mathbb{E}[Y_j|\mathcal{F}_{j-1}]| + \sum_{j=1}^k (\mathbb{E}[Y_j|\mathcal{F}_{j-1}])^2 \\ &\le 2m\psi(m)c_2 \sum_{j=1}^k |Y_j| + \sum_{j=1}^k \big(m\psi(m)c_2\big)^2 \\ &\le 2k^{1/2}m\psi(m)c_2 \Big(\sum_{j=1}^k Y_j^2\Big)^{1/2} + km^2\psi^2(m)c_2^2. \end{aligned} \tag{4.34}$$

When $\sum_{j=1}^k Y_j^2 \ge 1/4$, both sides of the last inequality divided by $\sum_{j=1}^k Y_j^2$, we get

$$\left| \frac{\sum_{j=1}^k \big(Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]\big)^2}{\sum_{j=1}^k Y_j^2} - 1 \right| \le 4k^{1/2}m\psi(m)c_2 + 4km^2\psi^2(m)c_2^2. \tag{4.35}$$

By assumption $\gamma_n \to 0$, we have $n\psi(m) \to 0$ which leads to $k^{1/2}m\psi(m) \to 0$ as $n \to \infty$. Thus (4.35) implies that

$$\left| \frac{\sum_{j=1}^k Y_j^2}{\sum_{j=1}^k \big(Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]\big)^2} - 1 \right| \le C_0 k^{1/2}m\psi(m), \tag{4.36}$$

where C_0 is a positive constant. Hence, by the last inequality and the fact $k^{1/2}m\psi(m) \to 0$, when $\sum_{j=1}^k Y_j^2 \ge 1/4$, it holds

$$\begin{aligned} \Big| W_n - W_n^o \Big| &= \frac{1}{\sqrt{\sum_{j=1}^k Y_j^2}} \left| \sum_{j=1}^k (Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]) \sqrt{\frac{\sum_{j=1}^k Y_j^2}{\sum_{j=1}^k \big(Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]\big)^2}} - \sum_{j=1}^k Y_j \right| \\ &\le \frac{1}{\sqrt{\sum_{j=1}^k Y_j^2}} \Bigg(\Big| \sum_{j=1}^k (Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]) - \sum_{j=1}^k Y_j \Big| + \\ &\quad + C_0 k^{1/2} m\psi(m) \Big| \sum_{j=1}^k (Y_j - \mathbb{E}[Y_j|\mathcal{F}_{j-1}]) \Big| \Bigg) \\ &\le 3\sum_{j=1}^k \Big| \mathbb{E}[Y_j|\mathcal{F}_{j-1}] \Big| + C_0 k^{1/2} m\psi(m) \frac{\sum_{j=1}^k |Y_j|}{\sqrt{\sum_{j=1}^k Y_j^2}}. \end{aligned}$$

Using Cauchy-Schwarz's inequality, we have $\sum_{j=1}^k |Y_j| \le k^{1/2}\sqrt{\sum_{j=1}^k Y_j^2}$. Thus, by (4.29) and the last inequality,

$$\begin{aligned} \Big| W_n - W_n^o \Big| &\le 3km\psi(m)c_2 + C_0 km\psi(m) \\ &\le (3c_2 + C_0) n\psi(m). \end{aligned}$$

Hence, when $\sum_{j=1}^k Y_j^2 \geq 1/4$, we have

$$\Big|W_n - W_n^o\Big| \leq C_1 \varepsilon_n, \tag{4.37}$$

where C_1 is a positive constant and

$$\varepsilon_n = n\psi(m).$$

Clearly, by assumption $\gamma_n \to 0$, we have $\varepsilon_n \to 0$ as $n \to \infty$. Without loss of generality, we may assume that $\sum_{j=1}^k \mathbb{E}Y_j^2 = n$; otherwise, we may consider $(\eta_i/\sqrt{\sum_{j=1}^k \mathbb{E}Y_j^2/n})_{1\leq i\leq n}$ instead of $(\eta_i)_{1\leq i\leq n}$. Then it follows that

$$\begin{aligned}
\mathbb{P}\left(\sum_{j=1}^k Y_j^2 < \frac{1}{4}\right) &\leq \mathbb{P}\left(\sum_{j=1}^k (Y_j^2 - \mathbb{E}[Y_j^2|\mathcal{F}_{j-1}]) < \frac{1}{4} - \sum_{j=1}^k \mathbb{E}[Y_j^2|\mathcal{F}_{j-1}]\right) \\
&\leq \mathbb{P}\left(\sum_{j=1}^k (Y_j^2 - \mathbb{E}[Y_j^2|\mathcal{F}_{j-1}]) < \frac{1}{4} - (1-\psi(m))\sum_{j=1}^k \mathbb{E}Y_j^2\right) \\
&\leq \mathbb{P}\left(\sum_{j=1}^k Y_j^2 - \mathbb{E}[Y_j^2|\mathcal{F}_{j-1}] < -\frac{1}{2}n\right).
\end{aligned} \tag{4.38}$$

Denote $Z_j = \mathbb{E}[Y_j^2|\mathcal{F}_{j-1}] - Y_j^2$. Notice that

$$Z_j \leq \mathbb{E}[Y_j^2|\mathcal{F}_{j-1}] \leq (1-\psi(m))\mathbb{E}Y_j^2 \asymp m,$$

By an argument similar to the proof of (4.30), we have

$$\mathbb{E}[|Z_j|^{1+\rho/2}|\mathcal{F}_{j-1}] \preceq m^{1+\rho/2}.$$

The last two inequalities implies that

$$\sum_{i=1}^k \Big(\mathbb{E}\big[(Z_i^-)^{1+\rho/2}|\mathcal{F}_{i-1}\big] + (Z_j^+)^{1+\rho/2}\Big) \leq C_2 k m^{1+\rho/2},$$

where C_2 is a positive constant. Applying Lemma 4.2 to $(Z_j)_{1\leq j\leq k}$ with $\beta = 1+\rho/2, x = n/2$ and $v^\beta = C_2 k m^\beta$, we get, from (4.38),

$$\begin{aligned}
\mathbb{P}\left(\sum_{j=1}^k Y_j^2 < \frac{1}{4}\right) &\leq \mathbb{P}\left(\sum_{j=1}^k Z_j > \frac{1}{2}n\right) \\
&\leq \exp\Big\{-C(\rho)\, n^{1-\alpha}\Big\},
\end{aligned} \tag{4.39}$$

where $C(\rho)$ is a positive constant. For the last inequality, we make use of the fact that $m = \lfloor n^\alpha \rfloor$, $k = \lfloor n/(2m) \rfloor$ and

$$\Big(\frac{x}{v}\Big)^{\beta/(\beta-1)} = \Big(\frac{n/2}{C_2^{1/\beta}k^{1/\beta}m}\Big)^{\beta/(\beta-1)} \asymp \Big(\frac{n}{n^{(1-\alpha)/\beta}n^\alpha}\Big)^{\beta/(\beta-1)} = n^{1-\alpha}.$$

Using (4.37), we obtain the following upper bound for the relative error of normal approximation: for all $0 \leq x = o(n^{(1-\alpha)/2})$,

$$\begin{aligned}
\frac{\mathbb{P}(W_n^o \geq x)}{1-\Phi(x)} &= \frac{\mathbb{P}(W_n^o \geq x, \sum_{j=1}^k Y_j^2 \geq 1/4) + \mathbb{P}(W_n^o \geq x, \sum_{j=1}^k Y_j^2 < 1/4)}{1-\Phi(x)} \\
&\leq \frac{\mathbb{P}(W_n \geq x - C_1\varepsilon_n, \sum_{j=1}^k Y_j^2 \geq 1/4) + \mathbb{P}(\sum_{j=1}^k Y_j^2 < 1/4)}{1-\Phi(x)} \\
&\leq \frac{\mathbb{P}(W_n \geq x - C_1\varepsilon_n)}{1-\Phi(x - C_1\varepsilon_n)} \frac{1-\Phi(x - C_1\varepsilon_n)}{1-\Phi(x)} + \frac{\mathbb{P}(\sum_{j=1}^k Y_j^2 < 1/4)}{1-\Phi(x)}.
\end{aligned}$$

Using the inequality

$$1-\Phi(x) \geq \frac{1}{\sqrt{2\pi}(1+x)} e^{-x^2/2} \quad \text{for all } x \geq 0$$

(cf. (1.6) of [11]), we deduce that for all $x \geq 0$ and $\varepsilon_n = o(1)$,

$$\begin{aligned}
\frac{1-\Phi(x - C_1\varepsilon_n)}{1-\Phi(x)} &= 1 + \frac{\int_{x-C_1\varepsilon_n}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt}{1-\Phi(x)} \\
&\leq 1 + \frac{\frac{1}{\sqrt{2\pi}} e^{-(x-C_1\varepsilon_n)^2/2} C_1\varepsilon_n}{\frac{1}{\sqrt{2\pi}(1+x)} e^{-x^2/2}} \\
&\leq 1 + C_1(1+x)\varepsilon_n e^{C_1 x \varepsilon_n} \\
&\leq e^{C_1(1+x)\varepsilon_n}. \qquad (4.40)
\end{aligned}$$

By (4.33), (4.39) and (4.40), we have for all $0 \leq x = o(n^{(1-\alpha)/2})$,

$$\begin{aligned}
&\frac{\mathbb{P}(W_n^o \geq x)}{1-\Phi(x)} \\
&\leq \exp\left\{ C_\rho \left(\frac{x^{2+\rho}}{n^{(1-\alpha)\rho/2}} + x^2\delta_n^2 + (1+x)\Big(\frac{1}{n^{(1-\alpha)\rho(2-\rho)/8}(1+x^{\rho(2+\rho)/4})} + \delta_n + \varepsilon_n \Big) \right) \right\} \\
&\quad + \frac{1}{1-\Phi(x)} \exp\Big\{ - C(\rho)\, n^{1-\alpha} \Big\} \\
&\leq \exp\left\{ C'_\rho \left(\frac{x^{2+\rho}}{n^{(1-\alpha)\rho/2}} + x^2\delta_n^2 + (1+x)\Big(\frac{1}{n^{(1-\alpha)\rho(2-\rho)/8}(1+x^{\rho(2+\rho)/4})} + \gamma_n \Big) \right) \right\},
\end{aligned}$$

where

$$\gamma_n = \delta_n + \varepsilon_n \asymp k^{1/2}\psi^{1/2}(m) + n\psi(m).$$

Similar, we have the following lower bound for the relative error of normal approximation: for all $0 \leq x = o(n^{(1-\alpha)/2})$,

$$\begin{aligned}
&\frac{\mathbb{P}(W_n^o \geq x)}{1-\Phi(x)} \\
&\geq \exp\left\{ - C'_\rho \left(\frac{x^{2+\rho}}{n^{(1-\alpha)\rho/2}} + x^2\delta_n^2 + (1+x)\Big(\frac{1}{n^{(1-\alpha)\rho(2-\rho)/8}(1+x^{\rho(2+\rho)/4})} + \gamma_n \Big) \right) \right\}.
\end{aligned}$$

Combining the upper and lower bounds of $\frac{\mathbb{P}(W_n^o \geq x)}{1-\Phi(x)}$ together, we complete the proof of Theorem 2.1 for $\rho \in (0,1)$.

For the case $\rho = 1$, the assertion of Theorem 2.1 follows by a similar argument, but with δ_n replaced by $\frac{\ln n}{n^{(1-\alpha)/2}}+\delta_n$ in (4.33) and accordingly in the subsequent statements. This completes the proof of Theorem 2.1.

4.3. Proof of Corollary 2.1

In the proof of Corollary 2.1, we will make use of the following well-known inequalities:

$$\frac{1}{\sqrt{2\pi}(1+x)}e^{-x^2/2} \leq 1-\Phi(x) \leq \frac{1}{\sqrt{\pi}(1+x)}e^{-x^2/2}, \qquad x \geq 0. \tag{4.41}$$

First, we show that

$$\limsup_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\left(\frac{1}{a_n} W_n^o \in B\right) \leq - \inf_{x\in \overline{B}} \frac{x^2}{2}. \tag{4.42}$$

When $B = \emptyset$, the last inequality is obvious. So, we assume that $B \neq \emptyset$. For a given Borel set $B \subset \mathbb{R}$, let $x_0 = \inf_{x\in B} |x|$. Clearly, we have $x_0 \geq \inf_{x\in \overline{B}} |x|$. Therefore, by Theorem 2.1,

$$\begin{aligned}\mathbb{P}\left(\frac{1}{a_n} W_n^o \in B\right) &\leq \mathbb{P}\Big(\big|W_n\big| \geq a_n x_0\Big) \\ &\leq 2\Big(1-\Phi\left(a_n x_0\right)\Big)\exp\left\{c_\rho \left(\frac{(a_n x_0)^{2+\rho}}{n^{(1-\alpha)\rho/2}} + \left(a_n x_0\right)^2 \delta_n^2\right.\right. \\ &\qquad \left.\left. +(1+a_n x_0)\Big(\frac{1}{n^{(1-\alpha)\rho(2-\rho)/8}(1+(a_n x_0)^{\rho(2+\rho)/4})} + \gamma_n\Big)\right)\right\}.\end{aligned}$$

Using (4.41), we get

$$\limsup_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\left(\frac{1}{a_n} W_n^o \in B\right) \;\leq\; -\frac{x_0^2}{2} \;\leq\; -\inf_{x\in \overline{B}} \frac{x^2}{2},$$

which gives (4.42).

Next, we show that

$$\liminf_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\left(\frac{1}{a_n} W_n^o \in B\right) \geq - \inf_{x\in B^o} \frac{x^2}{2}. \tag{4.43}$$

When $B^o = \emptyset$, the last inequality is obvious. So, we assume that $B^o \neq \emptyset$. For any given $\varepsilon_1 > 0$, there exists an $x_0 \in B^o$, such that

$$0 < \frac{x_0^2}{2} \leq \inf_{x\in B^o} \frac{x^2}{2} + \varepsilon_1.$$

Without loss of generality, we assume that $x_0 > 0$. For $x_0 \in B^o$ and all small enough $\varepsilon_2 \in (0, x_0)$, it holds $(x_0 - \varepsilon_2, x_0 + \varepsilon_2] \subset B$. Obviously, we have

$$\begin{aligned}\mathbb{P}\left(\frac{1}{a_n} W_n^o \in B\right) &\geq \mathbb{P}\Big(W_n^o \in (a_n(x_0-\varepsilon_2), a_n(x_0+\varepsilon_2)]\Big) \\ &= \mathbb{P}\Big(W_n^o \geq a_n(x_0-\varepsilon_2)\Big) - \mathbb{P}\Big(W_n^o \geq a_n(x_0+\varepsilon_2)\Big).\end{aligned}$$

By Theorem 2.1, it is easy to see that

$$\lim_{n\to\infty} \frac{\mathbb{P}\Big(W_n^o \geq a_n(x_0+\varepsilon_2)\Big)}{\mathbb{P}\Big(W_n^o \geq a_n(x_0-\varepsilon_2)\Big)} = 0.$$

Then, by (4.41), it follows that

$$\liminf_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\left(\frac{1}{a_n} W_n^o \in B\right) \geq -\frac{1}{2}(x_0-\varepsilon_2)^2.$$

Now, letting $\varepsilon_2 \to 0$, we have

$$\liminf_{n\to\infty} \frac{1}{a_n^2} \ln \mathbb{P}\left(\frac{1}{a_n} W_n^o \in B\right) \ \geq \ -\frac{x_0^2}{2} \ \geq \ -\inf_{x\in B^o} \frac{x^2}{2} - \varepsilon_1.$$

Because ε_1 can be arbitrarily small, we get (4.43). Combining (4.42) and (4.43) together, we complete the proof of Corollary 2.1.

Acknowledgments

The author would like to thank editor and reviewer for their helpful comments. This work has been partially supported by the National Natural Science Foundation of China (Grant no. 11601375).

References

[1] A. Bazarova, I. Berkes, L. Horváth, On the extremal theory of continued fractions, J. Theor. Probab. 29 (1) (2016) 248–266.
[2] B. Bercu, A. Touati, Exponential inequalities for self-normalized martingales with applications, Ann. Appl. Probab. 18 (5) (2008) 1848–1869.
[3] P. Billingsley, Ergodic Theory and Information, Wiley, 1965.
[4] X. Chen, Q.M. Shao, W.B. Wu, L. Xu, Self-normalized Cramér-type moderate deviations under dependence, Ann. Stat. 44 (4) (2016) 1593–1617.
[5] K.L. Chung, The approximate distribution of Student's statistic, Ann. Math. Stat. 17 (4) (1946) 447–465.
[6] H. Cramér, Sur un nouveau théorème-limite de la théorie des probabilités, Actual. Sci. Ind. 736 (1938) 5–23.
[7] M. Denker, Z. Kabluchko, An Erdös-Rényi law for mixing processes, Probab. Math. Stat. 27 (1) (2007) 139–149.
[8] V.H. de la Peña, A general class of exponential inequalities for martingales and ratios, Ann. Probab. 27 (1) (1999) 537–564.
[9] V.H. de la Peña, T.L. Lai, Q.M. Shao, Self-Normalized Processes: Theory and Statistical Applications, Springer Series in Probability and Its Applications, Springer-Verlag, New York, 2009.
[10] P. Doukhan, Mixing: Properties and Examples, Springer-Verlag, New York, 1994.
[11] X. Fan, Sharp large deviation results for sums of bounded from above random variables, Sci. China Math. 60 (12) (2017) 2465–2480.
[12] X. Fan, I. Grama, Q. Liu, Cramér large deviation expansions for martingales under Bernstein's condition, Stoch. Process. Appl. 123 (2013) 3919–3942.
[13] X. Fan, I. Grama, Q. Liu, Sharp large deviation results for sums of independent random variables, Sci. China Math. 58 (9) (2015) 1939–1958.
[14] X. Fan, I. Grama, Q. Liu, Martingale inequalities of type Dzhaparidze and van Zanten, Statistics 51 (6) (2017) 1200–1213.
[15] X. Fan, I. Grama, Q. Liu, Q.M. Shao, Self-normalized Cramér type moderate deviations for martingales, Bernoulli 25 (4A) (2019) 2793–2823.
[16] I. Grama, On moderate deviations for martingales, Ann. Probab. 25 (1997) 152–184.
[17] I. Grama, E. Haeusler, Large deviations for martingales via Cramér's method, Stoch. Process. Appl. 85 (2000) 279–293.
[18] B.Y. Jing, Q.M. Shao, Q. Wang, Self-normalized Cramér-type large deviations for independent random variables, Ann. Probab. 31 (4) (2003) 2167–2215.
[19] P. Lévy, Sur les lois de probabilité dont dépendent les quotients complets et incomplets d'une fraction continue, Bull. Soc. Math. Fr. 57 (1929) 178–194.
[20] Y.V. Linnik, On the probability of large deviations for the sums of independent variables, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 2, Univ. of California Press, 1961, pp. 289–306.
[21] W. Liu, Q.M. Shao, Q. Wang, Self-normalized Cramér type moderate deviations for the maximum of sums, Bernoulli 19 (3) (2013) 1006–1027.
[22] V.V. Petrov, A generalization of Cramér's limit theorem, Usp. Mat. Nauk 9 (1954) 195–202.
[23] V.V. Petrov, Sums of Independent Random Variables, Springer-Verlag, Berlin, 1975.
[24] A. Račkauskas, On probabilities of large deviations for martingales, Liet. Mat. Rink. 30 (1990) 784–795.
[25] A. Račkauskas, Large deviations for martingales with some applications, Acta Appl. Math. 38 (1995) 109–129.

[26] L. Saulis, V.A. Statulevičius, Limit Theorems for Large Deviations, Kluwer Academic Publishers, 1978.
[27] Q.M. Shao, A Cramér type large deviation result for Student's t-statistic, J. Theor. Probab. 12 (2) (1999) 385–398.
[28] Q.M. Shao, On necessary and sufficient conditions for the self-normalized central limit theorem, Sci. China Math. 61 (10) (2018) 1741–1748.
[29] Q.M. Shao, Q.Y. Wang, Self-normalized limit theorems: a survey, Probab. Surv. 10 (2013) 69–93.
[30] Q.M. Shao, H. Yu, Weighted weak convergence for empirical processes of dependent sequence, Ann. Probab. 24 (1996) 2098–2127.