Self-normalized Cramér type moderate deviations for stationary sequences and applications

Xiequan Fan, Ion Grama, Quansheng Liu and Qi-Man Shao

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

Université de Bretagne-Sud, LMBA, UMR CNRS 6205, Campus de Tohannic, 56017 Vannes, France.

Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen 518000, China.

Abstract

Let $(X_i)_{i\geq 1}$ be a stationary sequence. Denote $m = \lfloor n^\alpha \rfloor, 0 < \alpha < 1$, and $k = \lfloor n/m \rfloor$, where $\lfloor a \rfloor$ stands for the integer part of a. Set $S_j^\circ = \sum_{i=1}^m X_{m(j-1)+i}, 1 \leq j \leq k$, and $(V_k^\circ)^2 = \sum_{j=1}^k (S_j^\circ)^2$. We prove a Cramér type moderate deviation expansion for $\mathbb{P}(\sum_{j=1}^k S_j^\circ / V_k^\circ \geq x)$ as $n \to \infty$. Applications to mixing type sequences, contracting Markov chains, expanding maps and confidence intervals are discussed.

Keywords: moderate deviations; stationary processes; Cramér moderate deviations

2000 MSC: primary 60F10; 60G10; secondary 60E15

1. Introduction

Let $(X_i)_{i\geq 1}$ be a sequence of independent and identically distributed (i.i.d.) centered real random variables, that is $\mathbb{E}X_1 = 0$. Denote $S_n = \sum_{i=1}^n X_i$ the partial sums of $(X_i)_{i\geq 1}$ and $\sigma^2 = \mathbb{E}X_1^2$ the variance of X_1. Cramér [7] has established the following asymptotic moderate deviation expansion for the standardized sums: if $\mathbb{E}\exp\{c_0|X_1|\} < \infty$ for some constant $c_0 > 0$, termed Cramér's condition, then 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. Inequality (1.1) implies that

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

uniformly in the normal range $0 \leq x = o(n^{1/6})$. Notice that Cramér's condition is sufficient but not necessary for (1.2) to hold. Indeed, Linnik [25] proved that for $\alpha \in (0, \frac{1}{6}]$, formula (1.2) holds uniformly for $0 \leq x = o(n^{\alpha})$ as $n \to \infty$ if and only if $\mathbb{E} e^{|X_1|^{4\alpha/(2\alpha+1)}} < \infty$. Following the seminal work of Cramér, various moderate deviation expansions for standardized sums have been obtained by many authors, see, for instance, Petrov [28], Saulis and Statulevičius [36] and [15]. See also Račkauskas [29, 30], Grama [19], Grama and Haeusler [20] and [14] for martingales, and Wu and Zhao [38] and Cuny and Merlevède [9] for stationary processes.

For establishing moderate deviation expansions of type (1.2) with a range $0 \leq x = o(n^{\alpha})$, $\alpha > 0$, Linnik's condition is necessary. However, Linnik's condition becomes too restrictive if we only have finite moments of order $2 + \rho, \rho \in (0, 1]$. Although we still can establish (1.2) via (non-uniform) Berry-Esseen estimations (see Bikelis [3]), the range cannot be wider than $0 \leq x = O(\sqrt{\ln n})$, which is much more narrow than $0 \leq x = o(n^{\alpha})$. To overcome this limitation, instead of considering the standardized sums, one may consider the self-normalized sums, defined as follows:

$$W_n = S_n/V_n, \qquad \text{where } \ V_n^2 = \sum_{i=1}^{n} X_i^2.$$

One of the motivations to consider self-normalized sums is due to Student's t-statistic:

$$T_n = \sqrt{n}\, \overline{X}_n/\widehat{\sigma},$$

where

$$\overline{X}_n = \frac{S_n}{n} \qquad \text{and} \qquad \widehat{\sigma}^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X}_n)^2}{n-1}.$$

The Student's t-statistic T_n is linked to the the self-normalized sum W_n by the following formula: for all $x \geq 0$,

$$\mathbb{P}\Big(T_n \geq x\Big) = \mathbb{P}\Bigg(W_n \geq x\Big(\frac{n}{n + x^2 - 1}\Big)^{1/2}\Bigg),$$

see Chung [6]. So, an asymptotic bound on the tail probabilities for self-normalized sums implies an asymptotic bound on the tail probabilities for T_n. Shao [32] established self-normalized large and moderate deviation principles without any moment assumptions, and Shao [33] proved the following self-normalized Cramér type moderate deviations: if $\mathbb{E}|X_1|^{2+\rho} < \infty$ for some $\rho \in (0, 1]$, then

$$\frac{\mathbb{P}(W_n \geq x)}{1 - \Phi\,(x)} = 1 + o(1) \tag{1.3}$$

uniformly for $0 \leq x = o(n^{\rho/(4+2\rho)})$ as $n \to \infty$. The later result indicates that the normal range of x for (1.3) on self-normalized sums can be much wider than that for classical moderate deviation expansion (1.2) on sums of i.i.d. r.v.'s. The expansion (1.3) was further extended to independent but not necessarily identically distributed random variables by Jing, Shao and Wang [22]. Their result implies the following precise asymptotic normality under finite $(2+\rho)$-th moments:

$$\frac{\mathbb{P}(W_n \geq x)}{1 - \Phi\,(x)} = \exp\Big\{O(1)\frac{(1+x)^{2+\rho}}{n^{\rho/2}}\Big\}, \tag{1.4}$$

uniformly for $0 \leq x = o(\sqrt{n})$ as $n \to \infty$. Moderate deviation results of types (1.3) and (1.4) play an important role in statistical inference of means since in practice one usually does not know the variance σ^2. Even when the later can be estimated, it is still advisable to use self-normalized sums to obtain more precise results. Due to these significant advantages, the limit theory for self-normalized sums attracts more and more attention. Giné, Götze and Mason [21] gave a necessary and sufficient condition for the asymptotic normality of self-normalized partial sums. Csörgő, Szyszkowicz and Wang [8] established Donsker's theorem. For various moderate and large deviations results for self-normalized sums, we refer to, for instance, Jing, Shao and Wang [22], Liu, Shao and Wang [23], de la Peña, Lai and Shao [13], Shao and Wang [35] and Shao [34]. Dembo and Shao [11] and Liu and Shao [24] studied Hotelling's T^2-statistic.

The moderate deviation theory for self-normalized sums of independent random variables has been studied in depth. However, there are only a few results for dependent random variables. Chen, Shao, Wu and Xu [5] established self-normalized Cramér type moderate deviations for β-mixing sequences and functional dependent sequences (see Wu [37] for the definition of functional dependent sequences). Fan, Grama, Liu and Shao [16] gave two self-normalized Cramér type moderate deviation results for martingales. For a closely related topic, that is, exponential inequalities for self-normalized martingales, we refer to de la Peña [12] and Bercu and Touati [2]. The main purpose of this paper is to establish self-normalized Cramér type moderate deviations for general stationary sequences. We deduce also a self-normalized moderate deviation principle and a Berry-Esseen bound.

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

All over the paper, c and C, possibly enabled with indices (arguments), denote constants depending only on the previously introduced constants and on its indices (arguments). Their values may change on every occurence. For two positive real sequences $(a_n)_{i\geq 1}$ and $(b_n)_{i\geq 1}$, we write $a_n = O(b_n)$ if there exists a positive constant C such that $a_n \leq Cb_n$ holds for all large n, and $a_n = o(b_n)$ if $\lim_{n\to\infty} a_n/b_n = 0$. We also write $a_n \asymp b_n$ if $a_n = O(b_n)$ and $b_n = O(a_n)$, and $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$.

2. Main results

Assume that $(X_i)_{i\in\mathbb{Z}}$ is a stationary sequence of centered random variables, where $X_i = X_0 \circ T^i$ and $T : \Omega \mapsto \Omega$ is a bijective bimeasurable transformation preserving the probability $\mathbb{P}$ on $(\Omega, \mathcal{F})$. For a subfield $\mathcal{F}_0$ satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$, let $\mathcal{F}_i = T^{-i}(\mathcal{F}_0)$. Assume that X_0 is $\mathcal{F}_0$-measurable, so that the sequence $(X_i)_{i\in\mathbb{Z}}$ is adapted to the filtration $(\mathcal{F}_i)_{i\in\mathbb{Z}}$.

Denote by $\lfloor a \rfloor$ the integer part of the real a. Let $m \in [1, n]$ and $k = \lfloor n/m \rfloor$, where m may depend on n. Define

$$H_j = \{i : m(j-1) + 1 \leq i \leq mj\}, \quad 1 \leq j \leq k.$$

Consider the block sums $S_j^\circ = \sum_{i\in H_j} X_i$, and the block self-normalized sums

$$W_n^\circ = \frac{\sum_{j=1}^k S_j^\circ}{V_k^\circ}, \qquad \text{where} \quad (V_k^\circ)^2 = \sum_{j=1}^k (S_j^\circ)^2.$$

In particular, when $m = 1$, the block self-normalized sum W_n° becomes self-normalized sum W_n. We also denote the $\mathbb{L}^\infty$-norm of X by $\|X\|_\infty$, that is $\|X\|_\infty = \inf\{u : \mathbb{P}(|X| > u) = 0\}$. For any $1 \le m \le n$, set

$$\varepsilon_m = \frac{1}{n^{1/2} m^{1/\rho} \sigma^{2/\rho+1}} \Big\| \mathbb{E}[|S_m|^{2+\rho} | \mathcal{F}_0] \Big\|_\infty^{1/\rho}, \tag{2.1}$$

$$\gamma_m = \frac{1}{m^{1/2}\sigma} \sum_{j=1}^{\infty} \frac{1}{j^{3/2}} \Big\| \mathbb{E}[S_{mj} | \mathcal{F}_0] \Big\|_\infty \tag{2.2}$$

and

$$\delta_m^2 = \frac{1}{m\sigma^2} \Big\| \mathbb{E}[S_m | \mathcal{F}_0] \Big\|_\infty^2 + \Big\| \frac{1}{m\sigma^2} \mathbb{E}[S_m^2 | \mathcal{F}_0] - 1 \Big\|_\infty, \tag{2.3}$$

where ρ and σ are two positive constants. We are interested in the case where

$$\max\{\varepsilon_m,\, \gamma_m,\, \delta_m, m/n\} \to 0 \quad \text{as } n \to \infty. \tag{2.4}$$

We remark that $\delta_m \to 0$ implies that $\frac{1}{m} \sum_{i=1}^m \mathbb{E} S_m^2 \to \sigma^2$ as $n \to \infty$.

Remark 2.1. *Let us comment on condition (2.4).*

1. *If* $\big\| \mathbb{E}[|X_1|^{2+\rho} | \mathcal{F}_0] \big\|_\infty < \infty$, *then, by convexity, we have*

$$\|\mathbb{E}[|\frac{1}{m} S_m|^{2+\rho} | \mathcal{F}_0]\|_\infty \le \frac{1}{m} \sum_{i=1}^m \|\mathbb{E}[|X_i|^{2+\rho} | \mathcal{F}_0]\|_\infty \le \big\| \mathbb{E}[|X_1|^{2+\rho} | \mathcal{F}_0] \big\|_\infty$$

   and thus $\varepsilon_m = O(m^{1+1/\rho}/n^{1/2})$ *as* $n \to \infty$. *In particular, the claim holds provided that* X_1 *is bounded, that is* $\big\|X_1\big\|_\infty < \infty$.

2. *If* $\|X_1\|_\infty < \infty$ *and* $\delta_m \to 0$, *then we have*

$$\big\| \mathbb{E}[|S_m|^{2+\rho} | \mathcal{F}_0] \big\|_\infty \le m^\rho \big\|X_1\big\|_\infty^\rho \big\| \mathbb{E}[S_m^2 | \mathcal{F}_0] \big\|_\infty = O(m^{1+\rho}).$$

   Therefore, it holds $\varepsilon_m = O(m/n^{1/2})$ *as* $n \to \infty$.

3. *Assume* $\left\|\mathbb{E}[|S_m|^{2+\rho}|\mathcal{F}_0]\right\|_\infty = O(m^{1+\rho/2})$ *as* $m \to \infty$*. Then it is easy to see that* $\varepsilon_m = O(\sqrt{m/n})$ *as* $n \to \infty$*. In particular, if* $(X_i, \mathcal{F}_i)_{i\in\mathbb{Z}}$ *is a martingale difference sequence satisfying* $\left\|\mathbb{E}[|X_1|^{2+\rho}|\mathcal{F}_0]\right\|_\infty < \infty$*, then, by Theorem 2.1 of Rio [31], it is easy to see that*

$$(\mathbb{E}[|S_m|^{2+\rho}|\mathcal{F}_0])^{2/(2+\rho)} \le (1+\rho)\sum_{k=1}^{m}(\mathbb{E}[|X_i|^{2+\rho}|\mathcal{F}_0])^{2/(2+\rho)} \le (1+\rho)\left\|\mathbb{E}[|X_1|^{2+\rho}|\mathcal{F}_0]\right\|_\infty^{2/(2+\rho)} m$$

a.s., which leads to

$$\left\|\mathbb{E}[|S_m|^{2+\rho}|\mathcal{F}_0]\right\|_\infty = O(m^{1+\rho/2}) \quad \text{and} \quad \varepsilon_m = O(\sqrt{m/n})$$

as $n \to \infty$*, and*

$$\gamma_m = 0 \quad \text{and} \quad \delta_m^2 = \left\|\frac{1}{m\sigma^2}\sum_{i=1}^{m}\mathbb{E}[X_i^2|\mathcal{F}_0] - 1\right\|_\infty.$$

4. *Dedecker et al. [10] introduced the following two conditions for stationary sequences:*

(A1) *The following sum is finite:*

$$\sum_{n=1}^{\infty}\frac{1}{n^{3/2}}\left\|\mathbb{E}[S_n|\mathcal{F}_0]\right\|_\infty < \infty. \tag{2.5}$$

(A2) *There exists a positive constant* σ *such that*

$$\lim_{n\to\infty}\left\|\frac{1}{n}\mathbb{E}[S_n^2|\mathcal{F}_0] - \sigma^2\right\|_\infty = 0. \tag{2.6}$$

Clearly, under conditions (A1) and (A2), by Lemma 29 of Dedecker et al. [10], it holds that $\max\{\gamma_m, \delta_m\} \to 0$ *for any sequence* $m = m(n)$ *such that* $m \to \infty$ *and* $m/n \to 0$ *as* $n \to \infty$*.*

For any sequence of small positive numbers $(\varepsilon_m)_{m\ge1}$, let $\widehat{\varepsilon}_m(x,\rho)$ be a function of ε_m, x and ρ defined as follows

$$\widehat{\varepsilon}_m(x,\rho) = \frac{\varepsilon_m^{\rho(2-\rho)/4}}{1 + x^{\rho(2+\rho)/4}}. \tag{2.7}$$

The following theorem gives a self-normalized Cramér type moderate deviation result for stationary sequences.

Theorem 2.1. *Assume that there exists* $\rho \in (0,1]$ *such that* $\max\{\varepsilon_m, \gamma_m, \delta_m, m/n\} \to 0$ *as* $n \to \infty$*.*

[i] *If $\rho \in (0,1)$, then there exists an absolute constant $\alpha_\rho > 0$ such that for all $0 \leq x \leq \alpha_\rho \min\{\varepsilon_m^{-1},\, \sqrt{n/m}\}$,*

$$\left| \ln \frac{\mathbb{P}(W_n^\circ \geq x)}{1-\Phi\left(x\right)} \right| \ \leq \ C_\rho \bigg(x^{2+\rho}\varepsilon_m^\rho + x^2 \Big(\delta_m^2 + \gamma_m |\ln \gamma_m| + \frac{m}{n}\Big) \\ + \ (1+x)\Big(\delta_m + \gamma_m|\ln \gamma_m| + \varepsilon_m^\rho + \widehat{\varepsilon}_m(x,\rho) + \sqrt{\frac{m}{n}}\Big)\bigg),$$

where C_ρ depends only on ρ.

[ii] *If $\rho = 1$, then there exists an absolute constant $\alpha > 0$ such that for all $0 \leq x \leq \alpha \min\{\varepsilon_m^{-1},\, \sqrt{n/m}\}$,*

$$\left| \ln \frac{\mathbb{P}(W_n^\circ \geq x)}{1-\Phi\left(x\right)} \right| \ \leq \ C \bigg(x^3\varepsilon_m + x^2 \Big(\delta_m^2 + \gamma_m |\ln \gamma_m| + \frac{m}{n}\Big) \\ + \ (1+x)\Big(\delta_m + \gamma_m|\ln \gamma_m| + \varepsilon_m|\ln \varepsilon_m| + \widehat{\varepsilon}_m(x,1) + \sqrt{\frac{m}{n}}\Big)\bigg).$$

In particular, the last two inequalities imply that, for any $\rho \in (0,1]$,

$$\frac{\mathbb{P}(W_n^\circ \geq x)}{1-\Phi\left(x\right)} = 1 + o(1) \tag{2.8}$$

uniformly for $0 \leq x = o\big(\min\big\{\varepsilon_m^{-\rho/(2+\rho)},\, \delta_m^{-1}\,, (\gamma_m|\ln \gamma_m|)^{-1/2}, \sqrt{n/m}\big\}\big)$ as $n \to \infty$. Moreover, the same results hold with $\dfrac{\mathbb{P}(W_n^\circ \geq x)}{1-\Phi\left(x\right)}$ replacing by $\dfrac{\mathbb{P}(W_n^\circ \leq -x)}{\Phi\left(-x\right)}$.

Remark 2.2. *Let us comment on the results of Theorem 2.1.*

1. *The range of validity of (2.8) can be very large. For instance, if $\big\|\mathbb{E}[|S_n|^{2+\rho}|\mathcal{F}_0]\big\|_\infty = O(n^{1+\rho/2})$, $\big\|\mathbb{E}[S_n|\mathcal{F}_0]\big\|_\infty = O(1)$, and $\big\|\frac{1}{n}\mathbb{E}[S_n^2|\mathcal{F}_0] - \sigma^2\big\|_\infty = O\big(\frac{1}{n}\big)$ as $n \to \infty$, then*

   $$\varepsilon_m = O(\sqrt{m/n}), \qquad \gamma_m, \delta_m = O(\sqrt{1/m}).$$

   With $m = \lfloor n^{2\rho/(2+3\rho)} \rfloor$, equality (2.8) holds uniformly for $0 \leq x = o(n^{\rho/(4+6\rho)}/\sqrt{\ln n})$ as $n \to \infty$. The last range coincides with the classical range, up to a term $\sqrt{\ln n}$, when applied for block self-normalized sums of i.i.d. random variables, that is $0 \leq x = o(k^{\rho/(4+2\rho)})$. See Remark 1 of Shao [33].

2. *If $(X_i, \mathcal{F}_i)_{i\in\mathbb{Z}}$ is a martingale difference sequence satisfying $\|\mathbb{E}[|X_1|^{2+\rho}|\mathcal{F}_0]\|_\infty < \infty$, then Theorem 2.1 gives a block self-normalized Cramér type moderate deviation result, with*

$$\varepsilon_m = O\big(\sqrt{m/n}\big), \quad \gamma_m = 0 \quad and \quad \delta_m^2 = \Big\| \frac{1}{m\sigma^2}\sum_{i=1}^m \mathbb{E}[X_i^2|\mathcal{F}_0] - 1 \Big\|_\infty$$

as $n \to \infty$, which extends the main result of Fan et al. [16] to block self-normalized martingales. Furthermore, if $\|\mathbb{E}[X_i^2|\mathcal{F}_0] - \sigma^2\|_\infty \le Ci^{-\theta}$ for some positive constants C and θ, then we have

$$\delta_m^2 = \begin{cases} O(m^{-1}), & \text{if } \theta > 1, \\ O(m^{-1}\ln m), & \text{if } \theta = 1, \\ O(m^{-\theta}), & \text{if } \theta \in (0,1). \end{cases}$$

Taking

$$m = \begin{cases} \lfloor n^{\rho/(2+2\rho)} \rfloor, & \text{if } \theta \ge 1, \\ \lfloor n^{\rho/(\rho+\theta(2+\rho))} \rfloor, & \text{if } \theta \in (0,1), \end{cases}$$

we have the following results:

[i] *If $\rho \in (0,1)$, then (2.8) holds for $0 \le x = o(n^{\theta\rho/(2\rho+2\theta(2+\rho))})$.*

[ii] *If $\rho = 1$, then (2.8) holds for $0 \le x = o(n^{\rho/(4+4\rho)}/\ln n)$.*

[iii] *If $\rho > 1$, then (2.8) holds for $0 \le x = o(n^{\rho/(4+4\rho)})$.*

3. *Besides block self-normalized sums, we can also consider the interlacing self-normalized sums. Let $\alpha \in (0,1)$ and $m = \lfloor n^\alpha \rfloor$, $k = \lfloor n/(2m) \rfloor$ (instead of $\lfloor n/m \rfloor$ considered before) and*

$$B_j = \Big\{ i : 2m(j-1)+1 \le i \le 2mj - m \Big\}, \qquad 1 \le j \le k.$$

Let $Y_j^* = \sum_{l\in B_j} X_l$, $(V_k^*)^2 = \sum_{j=1}^k (Y_j^*)^2$ and write

$$I_n^* = \frac{\sum_{j=1}^k Y_j^*}{V_k^*}$$

for the interlacing self-normalized sum. Clearly, Theorem 2.1 also holds for interlacing self-normalized sums I_n^*, with $\mathbb{E}[\,\cdot\,|\mathcal{F}_0]$ and W_n° replaced respectively by $\mathbb{E}[\,\cdot\,|\mathcal{F}_{-m}]$ and I_n^*. Such type of results for β-mixing and some functional dependent sequences have been considered by Chen et al. [5].

The following self-normalized moderate deviation principle (MDP) result is a consequence of Theorem 2.1.

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

$$- \inf_{x\in B^o} \frac{x^2}{2} \leq \liminf_{n\to\infty} a_n^2 \ln \mathbb{P}\Big(a_n W_n^o \in B\Big) \leq \limsup_{n\to\infty} a_n^2 \ln \mathbb{P}\Big(a_n W_n^o \in B\Big) \leq - \inf_{x\in \overline{B}} \frac{x^2}{2}\,, \tag{2.9}$$

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

In the i.i.d. case, W_n^o is a self-normalized sum of k i.i.d. random variables. According to the classical result of Jing, Shao and Wang [22], the MDP holds for $0 \leq x = o(k^{1/2})$. Since $k = \lfloor n/m \rfloor$, the last range reads also as $0 \leq x = o(\sqrt{n/m})$. Notice that ε_m^{-1} is of order $\sqrt{n/m}$. Thus, the convergence rate of a_n in the last corollary cannot be improved even for i.i.d. random variables.

Theorem 2.1 also implies the following self-normalized Berry-Esseen bound for stationary sequences.

Corollary 2.2. *Assume the condition of Theorem 2.1. Then, for $\rho \in (0, 1]$,*

$$\sup_x \Big|\mathbb{P}(W_n^\circ \leq x) - \Phi\,(x)\,\Big| \leq C_\rho \Big(\delta_m + \gamma_m |\ln \gamma_m| + \varepsilon_m^{\rho(2-\rho)/4} + \sqrt{\frac{m}{n}}\Big),$$

where C_ρ depends only on ρ.

3. Applications

In this section, we present some applications of our results, including ϕ-mixing type sequences, contracting Markov chains, expanding maps and confidence intervals.

3.1. ϕ-mixing type sequences

Let Y be a random variable with values in a Polish space $\mathcal{Y}$. If $\mathcal{M}$ is a σ-field, the ϕ-mixing coefficient between $\mathcal{M}$ and $\sigma(Y)$ is defined by

$$\phi(\mathcal{M}, \sigma(Y)) = \sup_{A\in\mathfrak{B}(\mathcal{Y})} \Big\|\mathbb{P}_{Y|\mathcal{M}}(A) - \mathbb{P}_Y(A)\Big\|_\infty. \tag{3.1}$$

For a sequence of random variables $(X_i)_{i\in\mathbb{Z}}$ and a positive integer m, denote

$$\phi_m(n) = \sup_{i_m > ... > i_1 \geq n} \phi(\mathcal{F}_0, \sigma(X_{i_1}, ..., X_{i_m})),$$

and let $\phi(k) = \lim_{m\to\infty} \phi_m(k)$ be the usual ϕ-mixing coefficient. Under the following condition

$$\sum_{k\geq 1} k^{-1/2} \phi_1(k) < \infty \quad \text{ and } \quad \lim_{k\to\infty} \phi_2(k) = 0, \tag{3.2}$$

Dedecker et al. [10] obtained a MDP for standardized sums of bounded ϕ-mixing random variables. See also Gao [18] for an earlier version of MDP under the condition $\sum_{k\geq 1}\phi(k)<\infty$ which is stronger than (3.2). Denote

$$\begin{aligned}\eta_{1,n} &= \sup_{k\geq n}\|\mathbb{E}[X_k|\mathcal{F}_0]\|_\infty,\\ \eta_{2,n} &= \sup_{k,l\geq n}\|\mathbb{E}[X_kX_l|\mathcal{F}_0]-\mathbb{E}[X_kX_l]\|_\infty.\end{aligned}$$

Clearly, when the random variable X_0 is bounded, it holds that $\eta_{1,n}=O(\phi_1(n))$ and $\eta_{2,n}=O(\phi_2(n))$ as $n\to\infty$.

From Theorem 2.1 we obtain the following self-normalized Cramér type moderate deviation expansion with depending structure defined by $\eta_{1,n}$ and $\eta_{2,n}$.

Proposition 3.1. *Assume that* $\|X_0\|_\infty<\infty$,

$$\sigma^2:=\sum_{k=-\infty}^{\infty}\mathbb{E}[X_0X_k]>0 \quad \textit{and} \quad \max_{i=1,2}\{\eta_{i,n}\}=O(n^{-\beta}),\quad n\to\infty,$$

for some constant $\beta>1$.

[i] *If* $\beta\geq 3/2$, *then (2.8) with* $m=\lfloor n^{2/7}\rfloor$ *holds uniformly for* $0\leq x=o(n^{1/14}/\sqrt{\ln n})$ *as* $n\to\infty$.

[ii] *If* $\beta\in(1,3/2)$, *then (2.8) with* $m=\lfloor n^{1/(3\beta-1)}\rfloor$ *holds uniformly for* $0\leq x=o(n^{(\beta-1)/(6\beta-2)})$ *as* $n\to\infty$.

[iii] *Assume* $m:=m(n)\to\infty$ *and* $n^{1/2}/m\to\infty$ *as* $n\to\infty$. *Let* $(a_n)_{n\geq1}$ *be any sequence of real numbers such that* $a_n\to0$ *and* $a_nn^{1/2}/m\to\infty$ *as* $n\to\infty$. *Then (2.9) holds.*

By point 3 of Remark 2.1, if $\mathbb{E}|S_n|^{2+\rho}=O(n^{1+\rho/2})$ for some $\rho>0$, then point [iii] of Proposition 3.1 can be further improved. Indeed, in this case, (2.9) holds for any $m\to\infty$, and any sequence of real numbers $(a_n)_{n\geq1}$ such that $a_n\to0$ and $a_n\sqrt{n/m}\to\infty$ as $n\to\infty$.

3.2. Contracting Markov chains

Let $(Y_n)_{n\geq0}$ be a stationary Markov chain of bounded random variables with invariant measure μ and transition kernel K. Denote by $\|\cdot\|_{\infty,\mu}$ the essential norm with respect to μ. Let Λ_1 be the set of 1-Lipschitz functions. Assume that the Markov chain satisfies the following condition:

(B) There exist two constants $C>0$ and $\rho\in(0,1)$ such that

$$\sup_{g\in\Lambda_1}\|K^n(g)-\mu(g)\|_{\infty,\mu}\leq C\rho^n$$

and for any $g,g'\in\Lambda_1$ and any $m\geq0$,

$$\lim_{n\to\infty}\Big\|K^n\big(g'K^m(g)\big)-\mu\big(g'K^m(g)\big)\Big\|_{\infty,\mu}=0.$$

Denote by $\mathcal{L}$ the class of functions $f : \mathbf{R} \mapsto \mathbf{R}$ such that

$$|f(x) - f(y)| \leq h(|x - y|), \tag{3.3}$$

where h is a concave and non-decreasing function satisfying

$$\int_0^1 \frac{h(t)}{t\sqrt{|\ln t|}} dt < \infty, \tag{3.4}$$

see [10]. Clearly, inequality (3.4) holds if $h(t) \leq c|\ln(t)|^{-\gamma}$ for some constants $c > 0$ and $\gamma > 1/2$. In particular, $\mathcal{L}$ contains the class of α-Hölder continuous functions from $[0, 1]$ to $\mathbf{R}$, where $\alpha \in (0, 1]$.

Dedecker et al. [10] proved a MDP for the sequence

$$X_n = f(Y_n) - \mu(f) \tag{3.5}$$

under the condition that the function f belongs to the class $\mathcal{L}$. The following proposition gives an extension of the MDP to *self-normalized* sums $W_n^\circ = \frac{\sum_{j=1}^k S_j^\circ}{V_k^\circ}$, where $S_j^\circ = \sum_{i \in H_j} X_i$ and $(V_k^\circ)^2 = \sum_{j=1}^k (S_j^\circ)^2$.

Proposition 3.2. *Assume that the stationary Markov chain $(Y_n)_{n\geq 0}$ satisfies condition (B), and let X_n be defined by (3.5), with f belonging to $\mathcal{L}$. Assume $m := m(n) \to \infty$ and $n^{1/2}/m \to \infty$ as $n \to \infty$. Let a_n be any sequence of real numbers such that $a_n \to 0$ and $a_n n^{1/2}/m \to \infty$ as $n \to \infty$. If*

$$\sigma^2 := \sigma^2(f) = \mu\big((f - \mu(f))^2\big) + 2 \sum_{n \geq 1} \mu\big(K^n(f)\ (f - \mu(f))\big) > 0,$$

then (2.9) holds.

Proof. By Lemma 15 of Dedecker et al. [10], it is easy to see that X_1 is bounded: $\|X_1\|_{\infty,\mu} \leq h(C\rho)$ with h defined by (3.3). Then by point 2 of Remark 2.1, we have $\varepsilon_m = O(m/n^{1/2})$ as $n \to \infty$. The conditions of Proposition 3.2 imply the conditions (2.5) and (2.6): see the proof of Proposition 14 in Dedecker et al. [10]. Hence, by point 4 of Remark 2.1, the conditions of Proposition 3.2 imply the conditions of Corollary 2.1, thus Proposition 3.2 follows. □

Furthermore, assume that the Markov chain satisfies the following condition which is stronger than condition (B).

(C) There exist two constants $C > 0$ and $\rho \in (0, 1)$ such that

$$\sup_{g \in \Lambda_1} \|K^n(g) - \mu(g)\|_{\infty,\mu} \leq C\rho^n$$

and for any $m \geq 0$,

$$\sup_{g, g' \in \Lambda_1} \Big\|K^n\big(g' K^m(g)\big) - \mu\big(g' K^m(g)\big)\Big\|_{\infty,\mu} \leq C\rho^n.$$

Then we have the following self-normalized Cramér type moderate deviation expansion.

Proposition 3.3. *Assume that the stationary Markov chain $(Y_n)_{n\geq 0}$ satisfies condition (C), and let X_n be defined by (3.5). Assume $f \in \mathcal{L}$,*

$$\sigma^2 := \sigma^2(f) = \mu\Big((f-\mu(f))^2\Big) + 2\sum_{n>0} \mu\Big(K^n(f)\cdot(f-\mu(f))\Big) > 0$$

and, for some constant $\beta > 1$,

$$h(\rho^n) = O(n^{-\beta}), \qquad n \to \infty, \tag{3.6}$$

where h is defined by (3.3).

[i] *If $\beta \geq 3/2$, then (2.8) with $m = \lfloor n^{2/7} \rfloor$ holds uniformly for $0 \leq x = o(n^{1/14}/\sqrt{\ln n})$ as $n \to \infty$.*

[ii] *If $\beta \in (1, 3/2)$, then (2.8) with $m = \lfloor n^{1/(3\beta-1)} \rfloor$ holds uniformly for $0 \leq x = o(n^{(\beta-1)/(6\beta-2)})$ as $n \to \infty$.*

Notice that if $g(t) \leq c|\ln(t)|^{-\beta}$ for some constants $c > 0$ and $\beta > 1$, then (3.6) is satisfied.
Proof. From the proof of Propositions 14 of [10], it is easy to see that

$$\max_{i=1,2}\{\eta_{i,n}\} = O\big(h(C\rho^n)\big),$$

where C is given by condition (C) and h is defined by (3.3). Notice that $C\rho^n \leq \rho^{n/2}$ for n large enough. Hence, Proposition 3.3 is a simple consequence of Proposition 3.1. □

3.3. Expanding maps

Dedecker et al. [10] have obtained the MDP for expanding maps. Here we show that our results can also be applied to expanding maps for getting self-normalized MDP and Cramér type moderate deviations.

Let T be a map from $[0,1]$ to $[0,1]$ preserving a probability μ on $[0,1]$, and denote

$$X_k = f \circ T^{n-k+1} - \mu(f),$$

for any function $f \in L^2([0,1],\mu)$. Let $W_n^\circ = \frac{\sum_{j=1}^k S_j^\circ}{V_k^\circ}$, where $S_j^\circ = \sum_{i\in H_j} X_i$ and $(V_k^\circ)^2 = \sum_{j=1}^k (S_j^\circ)^2$. Denote by $\mathcal{BV}$ the class of bounded variation functions from $[0,1]$ to $\mathbf{R}$. For any $f \in \mathcal{BV}$, denote by $\|df\|$ the total variation norm of the measure df : $\|df\| = \sup\{\int g df, \|g\|_\infty \leq 1\}$. A Markov kernel K is said to be $\mathcal{BV}$-contracting if there exist two constants $k > 0$ and $\rho \in [0,1)$ such that

$$\|dK^n(f)\| \leq k\rho^n \|df\|. \tag{3.7}$$

Define the Perron-Frobenius operator K from $L^2([0,1],\mu)$ to $L^2([0,1],\mu)$ via the equality

$$\int_0^1 (Kh)(x)f(x)\mu(dx) = \int_0^1 h(x)(f\circ T)(x)\mu(dx). \tag{3.8}$$

The map T is said to be $\mathcal{BV}$-contracting if its Perron-Frobenius operator is $\mathcal{BV}$-contracting. We have the following corollary for the self-normalized sum W_n°.

Proposition 3.4. *Assume that T is $\mathcal{BV}$-contracting, $f \in \mathcal{BV}$ and $\sigma^2 := \mu\big((f-\mu(f))^2\big) + 2\sum_{n\geq 2}\mu\big(f\circ T^n\cdot(f-\mu(f))\big) > 0$.*

[i] *Let $m = \lfloor n^{2/7}\rfloor$. Equality (2.8) holds uniformly for $0 \leq x = o(n^{1/14}/\sqrt{\ln n})$ as $n\to\infty$.*

[ii] *Assume $m := m(n) \to \infty$ and $n^{1/2}/m \to \infty$ as $n\to\infty$. Let (a_n) be any sequence of real numbers such that $a_n \to 0$ and $a_n n^{1/2}/m \to \infty$ as $n\to\infty$. Then (2.9) holds.*

Proof. Let $(Y_i)_{i\geq 1}$ be the Markov chain with transition kernel K and invariant measure μ in the stationary regime. Using equality (3.8), it is easy to see that $(Y_0,...,Y_n)$ is distributed as $(T^{n+1},...,T)$. Assume that $f\in\mathcal{BV}$. Since K is $\mathcal{BV}$-contracting, by the proof of Corollary 18 of [10], we have

$$\|\mathbb{E}[X_k|Y_0]\|_\infty \leq C\rho^k\|df\|$$

and, for any $l > k \geq 0$,

$$\|\mathbb{E}[X_kX_l|Y_0] - \mathbb{E}[X_kX_l]\|_\infty \leq C(1+C)\rho^k\|df\|^2.$$

By an argument similar to the proof of Proposition 3.1, Proposition 3.4 follows. □

3.4. Application to confidence intervals

Consider the problem of constructing confidence intervals for the mean value μ of the stationary sequence $(\zeta_i)_{i\geq 1}$. Let $X_i = \zeta_i - \mu, i\geq 1$. Assume that $(X_i)_{i\geq 1}$ satisfies the conditions (2.1)-(2.4). Let

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

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

Proposition 3.5. *Let $\kappa_n \in (0,1)$. Assume that $\kappa_n \to 0$ and*

$$\big|\ln\kappa_n\big| = o\Big(\min\big\{\varepsilon_m^{-2},\, n/m\big\}\Big), \quad n\to\infty. \tag{3.9}$$

Let $\Delta_n = \frac{\sqrt{2|\ln(\kappa_n/2)|}}{km}\sqrt{\sum_{j=1}^k (Y_j - \overline{Y}_j)^2}$. Then $[A_n, B_n]$ with

$$A_n = \frac{\sum_{j=1}^k Y_j}{km} - \Delta_n, \quad B_n = \frac{\sum_{j=1}^k Y_j}{km} + \Delta_n,$$

is a $1-\kappa_n$ confidence interval for μ, for n large enough.

Proof. It is well 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 [6]. The last equality and Theorem 2.1 together implies that

$$\frac{\mathbb{P}(T_n \geq x)}{1-\Phi\,(x)} = \exp\Big\{o(1)(1+x)^2\Big\} \tag{3.10}$$

uniformly for $0 \leq x = o\big(\min\big\{\varepsilon_m^{-1},\, \sqrt{n/m}\big\}\big)$. Let $F(x) = 1-(1-\Phi\,(x))\exp\{o(1)(1+x)^2\}$. Notice that

$$1-\Phi\,(x_n) \to \frac{1}{x_n\sqrt{2\pi}}e^{-x_n^2/2} = \exp\bigg\{-\frac{x_n^2}{2}\Big(1+\frac{2}{x_n^2}\ln(x_n\sqrt{2\pi})\Big)\bigg\},\ \ x_n \to \infty.$$

Thus the upper $(\kappa_n/2)$-th quantile of the distribution function F satisfies

$$F^{-1}(\kappa_n/2) \to \sqrt{2|\ln(\kappa_n/2)|}, \quad n \to \infty,$$

which, by (3.9), is of order $o\big(\min\big\{\varepsilon_m^{-1},\, \sqrt{n/m}\big\}\big)$. Then applying (3.10) to T_n, we complete the proof of Proposition 3.5. □

By (3.9), a good choice of the size m is such that $R_n := \min\big\{\varepsilon_m^{-2},\, n/m\big\}$ is large enough, so that κ_n can be small enough. A suitable choice is $m = \lfloor \ln n \rfloor$; then, by Remark 2.1, we have

$$R_n = \begin{cases} \frac{n}{\lfloor \ln n\rfloor^2}, & \text{if } \|X_1\|_\infty < \infty, \\ \frac{n}{\lfloor \ln n\rfloor^{2+2/\rho}}, & \text{if } \big\|\mathbb{E}[|X_1|^{2+\rho}|\mathcal{F}_0]\big\|_\infty < \infty, \\ \frac{n}{\lfloor \ln n\rfloor}, & \text{if } \big\|\mathbb{E}[|S_m|^{2+\rho}|\mathcal{F}_0]\big\|_\infty = O(m^{1+\rho/2}). \end{cases}$$

Proposition 3.5 uses a condition on the $\mathbb{L}^\infty$-norm. We should mention that Hannan's central limit theorem (cf. Hannan [17]) holds under the condition on the $\mathbb{L}^2$-norm. Accordingly, a confidential interval for linear regression can be obtained via Hannan's theorem (cf. Caron and Dede [4]), but with larger risk probability; the risk probability can be significantly improved, using Cramér type moderate deviations of Wu and Zhao [38] and Cuny and Merlevède [9] on stationary sequences. Notice that the results of [38] and [9] also hold when X_i has finite p-th moments with $p > 2$. See also Chen et al. [5] for self-normalized Cramér type moderate deviations for β-mixing sequences and functional dependent sequences.

4. Proofs of Theorems

The proofs of our results are mainly based on the following lemmas which give some exponential deviation inequalities for the partial sums of dependent random variables.

4.1. Preliminary lemmas

Assume on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we are given a sequence of martingale differences $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$, where $\xi_0 = 0$, $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq ... \subseteq \mathcal{F}_n \subseteq \mathcal{F}$ are increasing σ-fields. Define

$$M_0 = 0, \qquad M_k = \sum_{i=1}^{k} \xi_i, \quad k = 1, ..., n. \tag{4.1}$$

Let $[M]_n$ and $\langle M \rangle_n$ be respectively the squared variance and the conditional variance of the martingale $M = (M_k, \mathcal{F}_k)_{k=0,...,n}$, that is

$$[M]_0 = 0, \quad [M]_k = \sum_{i=1}^{k} \xi_i^2, \quad \langle M \rangle_0 = 0, \quad \langle M \rangle_k = \sum_{i=1}^{k} \mathbb{E}[\xi_i^2|\mathcal{F}_{i-1}], \quad k = 1, ..., n. \tag{4.2}$$

Assume the following two conditions:

(C1) There exist $\epsilon_n \in (0, \frac{1}{2}]$ and $\rho \in (0, +\infty)$ such that

$$\mathbb{E}[|\xi_i|^{2+\rho}|\mathcal{F}_{i-1}] \leq \epsilon_n^\rho \, \mathbb{E}[\xi_i^2|\mathcal{F}_{i-1}], \quad 1 \leq i \leq n.$$

(C2) There exists $\iota_n \in [0, \frac{1}{2}]$ such that $\| \langle M \rangle_n - 1\|_\infty \leq \iota_n^2$.

In many situations we have $\epsilon_n, \iota_n \to 0$ as $n \to \infty$. In the case of sums of i.i.d. random variables with finite $(2+\rho)$-th moments, conditions (A1) and (A2) are satisfied with $\iota_n = 0$ and $\epsilon_n = O(1/\sqrt{n})$ as $n \to \infty$.

Define the self-normalized martingale

$$W_n = \frac{M_n}{\sqrt{[M]_n}}, \quad n \geq 1. \tag{4.3}$$

Define $\widehat{\epsilon}_m(x, \rho)$ in the same way as in (2.7) but with ε_m replaced by ϵ_m. The proof of Theorem 2.1 is based on the following technical lemma which gives a Cramér type moderate deviation expansion for self-normalized martingales.

Lemma 4.1. *Assume conditions (C1) and (C2).*

[i] *If $\rho \in (0, 1)$, then there exists an absolute constant $\alpha_{\rho,0} > 0$ such that for all $0 \leq x \leq \alpha_{\rho,0}\epsilon_n^{-1}$,*

$$\left| \ln \frac{\mathbb{P}(W_n \geq x)}{1 - \Phi(x)} \right| \leq c_\rho \bigg(x^{2+\rho}\epsilon_n^\rho + x^2 \iota_n^2 + (1+x)\big(\iota_n + \epsilon_n^\rho + \widehat{\epsilon}_n(x, \rho)\big) \bigg).$$

[ii] *If $\rho = 1$, then there exists an absolute constant $\alpha_0 > 0$ such that for all $0 \leq x \leq \alpha_0\epsilon_n^{-1}$,*

$$\left| \ln \frac{\mathbb{P}(W_n \geq x)}{1 - \Phi(x)} \right| \leq c \bigg(x^3\epsilon_n + x^2 \iota_n^2 + (1+x)\big(\iota_n + \epsilon_n|\ln \epsilon_n| + \widehat{\epsilon}_n(x, 1)\big) \bigg).$$

Moreover, the two above inequalities remain valid with $\frac{\mathbb{P}(W_n \leq -x)}{\Phi(-x)}$ *instead of* $\frac{\mathbb{P}(W_n \geq x)}{1-\Phi(x)}$.

Proof. The points [i] and [ii] follows by Corollary 2.3 of Fan *et al.* [16].

Remark 4.1. *Notice that in Fan* et al. *[16], the range for Lemma 4.1 is* $0 \leq x = o(\epsilon_n^{-1})$*. However, the proof of Fan* et al. *[16] can be applied with no changes to extend the range to* $0 \leq x \leq \alpha_{\rho,0}\epsilon_n^{-1}$*, where* $\alpha_{\rho,0}$ *is a sufficiently small positive constant.*

Denote by $x^+ = \max\{x, 0\}$ the positive part of x.

Lemma 4.2. *Assume that* $\xi_i \geq -a$ *a.s. for all* $i \in [1, n]$*. Write*

$$\mathrm{H}_n(\beta) = \sum_{i=1}^{n} \Big(\mathbb{E}\big[(\xi_i^+)^\beta | \mathcal{F}_{i-1}\big] + a^\beta \Big), \qquad \beta \in (1, 2].$$

Then for all $x, v > 0$,

$$\mathbb{P}\left(S_n \leq -x, \quad \mathrm{H}_n(\beta) \leq v^\beta\right) \leq \exp\left\{-\frac{1}{2}C(\beta)\left(\frac{x}{v}\right)^{\frac{\beta}{\beta-1}}\right\}, \tag{4.4}$$

where $C(\beta) = \beta^{-1/(\beta-1)} - \beta^{-\beta/(\beta-1)} > 0$ *and* $\beta \in (1, 2]$.

Proof. Let $\beta \in (1, 2]$. Using the inequality

$$e^{-x} \leq 1 - x + x^\beta \quad \text{for} \quad x \geq 0,$$

we have, for all $i \in [1, n]$ and all $t > 0$,

$$\begin{aligned}
\mathbb{E}\big[e^{-t(\xi_i + a)}\big|\mathcal{F}_{i-1}\big] &\leq 1 - \mathbb{E}\big[t(\xi_i + a)\big|\mathcal{F}_{i-1}\big] + \mathbb{E}\big[t^\beta(\xi_i + a)^\beta\big|\mathcal{F}_{i-1}\big] \\
&\leq 1 - ta + 2^{\beta-1}t^\beta\big(\mathbb{E}\big[(\xi_i^+)^\beta|\mathcal{F}_{i-1}\big] + a^\beta\big) \\
&\leq \exp\{-ta + 2^{\beta-1}t^\beta\big(\mathbb{E}\big[(\xi_i^+)^\beta|\mathcal{F}_{i-1}\big] + a^\beta\big)\}.
\end{aligned}$$

Therefore, for all $x, t, v > 0$,

$$\begin{aligned}
&\mathbb{P}\left(S_n \leq -x, \quad \mathrm{H}_n(\beta) \leq v^\beta\right) \\
&\leq \mathbb{E}\Big[\exp\Big\{-tx - t\sum_{i=1}^{n}(\xi_i + a) + tna + 2^{\beta-1}t^\beta\mathrm{H}_n(\beta) - 2^{\beta-1}t^\beta\mathrm{H}_n(\beta)\Big\}\mathbf{1}_{\{\mathrm{H}_n(\beta)\leq v^\beta\}}\Big] \\
&\leq e^{-tx+tna+2^{\beta-1}t^\beta v^\beta}\mathbb{E}\Big[\exp\Big\{-t\sum_{i=1}^{n}(\xi_i + a) - 2^{\beta-1}t^\beta\mathrm{H}_n(\beta)\Big\}\Big] \\
&\leq e^{-tx+tna+2^{\beta-1}t^\beta v^\beta}\mathbb{E}\Big[\exp\Big\{-t\sum_{i=1}^{n-1}(\xi_i + a) - 2^{\beta-1}t^\beta\mathrm{H}_n(\beta)\Big\}\mathbb{E}\big[e^{-t(\xi_n+a)}\big|\mathcal{F}_{n-1}\big]\Big] \\
&\leq e^{-tx+t(n-1)a+2^{\beta-1}t^\beta v^\beta}\mathbb{E}\Big[\exp\Big\{-t\sum_{i=1}^{n-1}(\xi_i + a) - 2^{\beta-1}t^\beta\mathrm{H}_{n-1}(\beta)\Big\}\Big] \\
&\leq e^{-tx+2^{\beta-1}t^\beta v^\beta}.
\end{aligned}$$

Taking $t = \frac{1}{2}\big(\frac{x}{\beta v^\beta}\big)^{1/(\beta-1)}$ yields the desired inequality. $\square$

The following exponential inequality of Peligrad et al. [27] (cf. Proposition 2 therein) plays an important role in the proof of Theorem 2.1.

Lemma 4.3. *Let $(X_i)_{i\in\mathbb{Z}}$ be a sequence of random variables adapted to the filtration $(\mathcal{F}_i)_{i\in\mathbb{Z}}$. Then, for all $x \geq 0$,*

$$\mathbb{P}\bigg(\max_{1\leq i\leq n}|S_i| \geq x\bigg) \leq 4\sqrt{e}\exp\bigg\{-\frac{x^2}{2n(\|X_1\|_\infty + 80\sum_{j=1}^n j^{-3/2}\|\mathbb{E}[S_j|\mathcal{F}_0]\|_\infty)^2}\bigg\}. \tag{4.5}$$

The last lemma shows that the tail probability of $\max_{1\leq i\leq n}|S_i|$ has a sub-Gaussian decay rate. In the proof of Theorem 2.1, we apply it to estimate the tail probabilities for the drift of a stationary sequence.

4.2. Proof of Theorem 2.1

Define

$$D_j^\circ = S_j^\circ - \mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}], \quad 1\leq j\leq k.$$

Then $(D_j^\circ, \mathcal{F}_{(j-1)m})_{1\leq j\leq k}$ is a stationary sequence of martingale differences. Clearly,

$$\mathbb{E}[(D_j^\circ)^2|\mathcal{F}_{(j-1)m}] = \mathbb{E}[(S_j^\circ)^2|\mathcal{F}_{(j-1)m}] - (\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}])^2.$$

By stationarity and the fact that $k = \lfloor n/m\rfloor$, it follows that

$$\frac{1}{n}\bigg\|\sum_{j=1}^k(\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}])^2\bigg\|_\infty \leq \frac{1}{m}\bigg\|\mathbb{E}[S_m|\mathcal{F}_0]\bigg\|_\infty^2,$$

and that

$$\begin{aligned}\bigg\|\frac{1}{n}\sum_{j=1}^k\mathbb{E}[(S_j^\circ)^2|\mathcal{F}_{(j-1)m}] - \sigma^2\bigg\|_\infty &\leq \frac{1}{n}\sum_{j=1}^k\bigg\|\mathbb{E}[(S_j^\circ)^2|\mathcal{F}_{(j-1)m}] - m\sigma^2\bigg\|_\infty + \frac{n-mk}{n}\sigma^2 \\ &\leq \bigg\|\frac{1}{m}\mathbb{E}[S_m^2|\mathcal{F}_0] - \sigma^2\bigg\|_\infty + \frac{m}{n}\sigma^2. \end{aligned}\tag{4.6}$$

Consequently, we have

$$\begin{aligned}&\bigg\|\frac{1}{n\sigma^2}\sum_{j=1}^k\mathbb{E}[(D_j^\circ)^2|\mathcal{F}_{(j-1)m}] - 1\bigg\|_\infty \\ &\leq \bigg\|\frac{1}{n\sigma^2}\sum_{j=1}^k\mathbb{E}[(S_j^\circ)^2|\mathcal{F}_{(j-1)m}] - 1\bigg\|_\infty + \frac{1}{n\sigma^2}\bigg\|\sum_{j=1}^k(\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}])^2\bigg\|_\infty \\ &\leq \bigg\|\frac{1}{m\sigma^2}\mathbb{E}[S_m^2|\mathcal{F}_0] - 1\bigg\|_\infty + \frac{m}{n} + \frac{1}{m\sigma^2}\bigg\|\mathbb{E}[S_m|\mathcal{F}_0]\bigg\|_\infty^2 \\ &= \delta_m^2 + \frac{m}{n}. \end{aligned}\tag{4.7}$$

Since $\delta_m \to 0$ as $n \to \infty$, it follows that

$$\Big\|\frac{1}{m}\mathbb{E}[(D_j^\circ)^2|\mathcal{F}_{(j-1)m}]\Big\|_\infty = \Big\|\frac{1}{m}\mathbb{E}[S_m^2|\mathcal{F}_0]-\frac{1}{m}(\mathbb{E}[S_m|\mathcal{F}_0])^2\Big\|_\infty \sim \sigma^2, \quad n\to\infty. \tag{4.8}$$

Using the inequality

$$|x-y|^{2+\rho} \leq 2^{1+\rho}(|x|^{2+\rho}+|y|^{2+\rho}), \tag{4.9}$$

by (4.8) and stationarity, we deduce that

$$\begin{aligned}
\mathbb{E}[|D_j^\circ/(n^{1/2}\sigma)|^{2+\rho}|\mathcal{F}_{(j-1)m}] &\leq (n\sigma^2)^{-1-\rho/2}2^{2+\rho}\mathbb{E}[|S_j^\circ|^{2+\rho}|\mathcal{F}_{(j-1)m}] \\
&\leq \frac{2^{2+\rho}}{n^{\rho/2}\sigma^\rho}\Big\|\frac{\mathbb{E}[|S_j^\circ|^{2+\rho}|\mathcal{F}_{(j-1)m}]}{\mathbb{E}[(D_j^\circ)^2|\mathcal{F}_{(j-1)m}]}\Big\|_\infty \mathbb{E}[(D_j^\circ/(n^{1/2}\sigma))^2|\mathcal{F}_{(j-1)m}] \\
&\leq C_{\rho,0}\frac{1}{n^{\rho/2}m\sigma^{2+\rho}}\Big\|\mathbb{E}[|S_m|^{2+\rho}|\mathcal{F}_0]\Big\|_\infty \mathbb{E}[(D_j^\circ/(n^{1/2}\sigma))^2|\mathcal{F}_{(j-1)m}] \\
&= C_{\rho,0}\,\varepsilon_m^\rho \mathbb{E}[(D_j^\circ/(n^{1/2}\sigma))^2|\mathcal{F}_{(j-1)m}].
\end{aligned} \tag{4.10}$$

We first prove Theorem 2.1 for $\rho\in(0,1)$. Set $\xi_j = D_j^\circ/(n^{1/2}\sigma)$, and denote $M_k=\sum_{j=1}^k \xi_j$. Then, by (4.7) and (4.10), conditions (C1) and (C2) are satisfied with $n=k$, $\epsilon_n^\rho = C_{\rho,0}\varepsilon_m^\rho$ and $\iota_n^2 = \delta_m^2+\frac{m}{n}$. By Lemma 4.1, there exists a constant $\alpha_{\rho,0}>0$ such that for all $0\leq x\leq \alpha_{\rho,0}\varepsilon_m^{-1}$,

$$\begin{aligned}
&\left|\ln\frac{\mathbb{P}(M_k/\sqrt{[M]_k}\geq x)}{1-\Phi\left(x\right)}\right| \\
&\qquad \leq c_\rho\Bigg(x^{2+\rho}\varepsilon_m^\rho + x^2(\delta_m^2+\frac{m}{n}) + (1+x)\Big(\delta_m+\sqrt{\frac{m}{n}}+\varepsilon_m^\rho+\widehat{\varepsilon}_m(x,\rho)\Big)\Bigg).
\end{aligned} \tag{4.11}$$

Notice that, by Cauchy-Schwarz's inequality,

$$\begin{aligned}
\Big\|\frac{[M]_k}{(V_k^\circ)^2/(n\sigma^2)}-1\Big\|_\infty &= \Big\|\frac{2}{(V_k^\circ)^2}\sum_{j=1}^k S_j^\circ\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}] + \frac{1}{(V_k^\circ)^2}\sum_{j=1}^k\Big(\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}]\Big)^2\Big\|_\infty \\
&\leq \Big\|\frac{2}{(V_k^\circ)^2}\sum_{j=1}^k S_j^\circ\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}]\Big\|_\infty + \sum_{j=1}^k\Big\|\frac{1}{(V_k^\circ)^2}(\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}])^2\Big\|_\infty \\
&\leq \Big\|\frac{2}{(V_k^\circ)^2}\sum_{j=1}^k\big(\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}]\big)^2\Big\|_\infty^{1/2} + \sum_{j=1}^k\Big\|\frac{1}{V_k^\circ}|\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}]|\Big\|_\infty^2.
\end{aligned}$$

By stationarity and the fact that $\delta_m \to 0$, when $(V_k^\circ)^2 \geq \frac{1}{2}n\sigma^2$, we have

$$\begin{aligned}
\Big\| \frac{[M]_k}{(V_k^\circ)^2/(n\sigma^2)} - 1 \Big\|_\infty &\leq \frac{2\sqrt{2}}{\sqrt{n}\sigma} \Big\| \sum_{j=1}^k \big(\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}]\big)^2 \Big\|_\infty^{1/2} + \frac{2}{n\sigma^2} \sum_{j=1}^k \Big\| \mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}] \Big\|_\infty^2 \\
&\leq \frac{2\sqrt{2k}}{\sqrt{n}\sigma} \Big\| \mathbb{E}[S_m|\mathcal{F}_0] \Big\|_\infty + \frac{2}{m\sigma^2} \Big\| \mathbb{E}[S_m|\mathcal{F}_0] \Big\|_\infty^2 \\
&\leq \frac{6}{\sqrt{m}\sigma} \Big\| \mathbb{E}[S_m|\mathcal{F}_0] \Big\|_\infty =: \kappa_m.
\end{aligned}$$

Clearly, $\delta_m \to 0$ as $n \to \infty$ implies that $\kappa_m \to 0$ as $n \to \infty$. Thus, the last inequality implies that

$$\frac{V_k^\circ}{\sqrt{n}\sigma} \geq \sqrt{\frac{[M]_k}{1+\kappa_m}} \geq \sqrt{[M]_k(1-\kappa_m)}.$$

Recall that $W_n^\circ = \frac{\sum_{j=1}^k S_j^\circ}{V_k^\circ} = \frac{\sum_{j=1}^k S_j^\circ/(n^{1/2}\sigma)}{V_k^\circ/(n^{1/2}\sigma)}$. It is easy to see that, for all $x \geq 0$,

$$\begin{aligned}
\mathbb{P}\bigg(W_n^\circ \geq x,\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2} \bigg) &\leq \mathbb{P}\bigg(\frac{\sum_{j=1}^k S_j^\circ/(n^{1/2}\sigma)}{\sqrt{[M]_k}} \geq x\sqrt{1-\kappa_m},\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2} \bigg) \\
&\leq \mathbb{P}\bigg(\frac{M_k}{\sqrt{[M]_k}} \geq x(1-\gamma_m|\ln\gamma_m|)\sqrt{1-\kappa_m},\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2} \bigg) \\
&\quad + \mathbb{P}\bigg(\frac{1}{n^{1/2}\sigma} \sum_{j=1}^k \mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}] \geq x\gamma_m|\ln\gamma_m|\sqrt{1-\kappa_m} \bigg) \\
&\leq \mathbb{P}\bigg(\frac{M_k}{\sqrt{[M]_k}} \geq x(1-\gamma_m|\ln\gamma_m|)\sqrt{1-\kappa_m} \bigg) \\
&\quad + \mathbb{P}\bigg(\frac{1}{n^{1/2}\sigma} \sum_{j=1}^k \mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}] \geq x\gamma_m|\ln\gamma_m|\sqrt{1-\kappa_m} \bigg) \\
&=: I_1(x) + I_2(x). \qquad (4.12)
\end{aligned}$$

We proceed to estimate $I_1(x)$ and $I_2(x)$. First, we deal with $I_1(x)$. From (4.11), we have, for all $0 \leq x \leq \alpha_{\rho,0}\varepsilon_m^{-1}$,

$$\begin{aligned}
&\frac{I_1(x)}{1-\Phi\big(x(1-\gamma_m|\ln\gamma_m|)\sqrt{1-\kappa_m}\big)} \\
&\qquad \leq \exp\bigg\{ c'_\rho \bigg(x^{2+\rho}\varepsilon_m^\rho + x^2(\delta_m^2 + \frac{m}{n}) + (1+x)\Big(\delta_m + \sqrt{\frac{m}{n}} + \varepsilon_m^\rho + \widehat{\varepsilon}_m(x,\rho)\Big) \bigg) \bigg\}.
\end{aligned}$$

Using the following inequalities

$$\frac{1}{\sqrt{2\pi}(1+x)}e^{-x^2/2} \leq 1-\Phi(x) \leq \frac{1}{\sqrt{\pi}(1+x)}e^{-x^2/2}, \quad x \geq 0, \tag{4.13}$$

we deduce that, for all $x \geq 0$ and $0 \leq \varepsilon \leq 1$,

$$\begin{aligned} \frac{1-\Phi\left(x\sqrt{1-\varepsilon}\right)}{1-\Phi\left(x\right)} &= 1+\frac{\int_{x\sqrt{1-\varepsilon}}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt}{1-\Phi\left(x\right)} \\ &\leq 1+\frac{\frac{1}{\sqrt{2\pi}}e^{-x^2(1-\varepsilon)/2}x\varepsilon}{\frac{1}{\sqrt{2\pi}(1+x)}e^{-x^2/2}} \\ &\leq 1+C(1+x^2)\varepsilon e^{x^2\varepsilon/2} \\ &\leq \exp\Big\{C(1+x^2)\varepsilon\Big\}. \end{aligned} \tag{4.14}$$

Notice that $(1-\gamma_m|\ln\gamma_m|)\sqrt{1-\kappa_m} \geq \sqrt{1-2(\gamma_m|\ln\gamma_m|+\kappa_m)}$. Using inequality (4.14) and the fact that $\kappa_m \leq 6\gamma_m \leq 6\gamma_m|\ln\gamma_m|$, we obtain, for all $0 \leq x \leq \alpha_{\rho,0}\varepsilon_m^{-1}$,

$$\begin{aligned} \frac{I_1(x)}{1-\Phi\left(x\right)} &= \frac{I_1(x)}{1-\Phi\left(x(1-\gamma_m|\ln\gamma_m|)\sqrt{1-\kappa_m}\right)}\frac{1-\Phi\left(x(1-\gamma_m|\ln\gamma_m|)\sqrt{1-\kappa_m}\right)}{1-\Phi\left(x\right)} \\ &\leq \exp\Bigg\{C_\rho'\bigg(x^{2+\rho}\varepsilon_m^\rho+x^2\Big(\delta_m^2+\frac{m}{n}+\gamma_m|\ln\gamma_m|+\kappa_m\Big) \\ &\qquad +(1+x)\Big(\delta_m+\sqrt{\frac{m}{n}}+\varepsilon_m^\rho+\widehat{\varepsilon}_m(x,\rho)+\gamma_m|\ln\gamma_m|+\kappa_m\Big)\bigg)\Bigg\} \\ &\leq \exp\Bigg\{C_\rho''\bigg(x^{2+\rho}\varepsilon_m^\rho+x^2\Big(\delta_m^2+\gamma_m|\ln\gamma_m|+\frac{m}{n}\Big) \\ &\qquad +(1+x)\Big(\delta_m+\sqrt{\frac{m}{n}}+\varepsilon_m^\rho+\gamma_m|\ln\gamma_m|+\widehat{\varepsilon}_m(x,\rho)\Big)\bigg)\Bigg\}, \end{aligned} \tag{4.15}$$

which gives the suitable bound for $I_1(x)$.

Now we deal with $I_2(x)$. By Lemma 4.3, the definition of γ_m (cf. (2.2)) and the fact that $\gamma_m \to 0$, we derive that, for all $x \geq 0$,

$$\begin{aligned} I_2(x) &\leq 4\sqrt{e}\exp\left\{-\frac{n\sigma^2x^2\gamma_m^2(\ln\gamma_m)^2(1-\kappa_m)}{2k(\left\|\mathbb{E}[S_m|\mathcal{F}_0]\right\|_\infty+80\sum_{j=1}^{k}j^{-3/2}\|\mathbb{E}[S_{jm}|\mathcal{F}_0]\|_\infty)^2}\right\} \\ &\leq 4\sqrt{e}\exp\Big\{-C_0x^2(\ln\gamma_m)^2\Big\}. \end{aligned} \tag{4.16}$$

From the last inequality, using (4.13), we deduce that, for all $x \geq 1$,

$$\begin{aligned} \frac{I_2(x)}{1-\Phi\left(x\right)} &\leq C_1(1+x)\exp\Big\{-C_0 x^2(\ln\gamma_m)^2+\frac{1}{2}x^2\Big\} \\ &\leq C_2\,(1+x)\gamma_m|\ln\gamma_m|, \end{aligned} \tag{4.17}$$

which gives the suitable bound for $I_2(x)$. Thus, from (4.12), for all $x \geq 1$,

$$\begin{aligned} &\frac{\mathbb{P}\Big(W_n^\circ \geq x,\, \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big)}{1-\Phi\left(x\right)} \leq \frac{I_1(x)+I_2(x)}{1-\Phi\left(x\right)} \\ &\leq \exp\bigg\{c''_\rho\Big(x^{2+\rho}\varepsilon_m^\rho + x^2\Big(\delta_m^2+\gamma_m|\ln\gamma_m|+\frac{m}{n}\Big) \\ &\qquad\qquad +(1+x)\Big(\delta_m+\sqrt{\frac{m}{n}}+\varepsilon_m^\rho+\gamma_m|\ln\gamma_m|+\widehat{\varepsilon}_m(x,\rho)\Big)\Big)\bigg\}. \end{aligned} \tag{4.18}$$

Clearly, we have

$$\begin{aligned} \mathbb{P}\Big((V_k^\circ)^2 < \frac{1}{2}n\sigma^2\Big) &= \mathbb{P}\bigg(\sum_{j=1}^k\Big((S_j^\circ)^2-\mathbb{E}[(S_j^\circ)^2|\mathcal{F}_{(j-1)m}]\Big) < \frac{1}{2}n\sigma^2-\sum_{j=1}^k\mathbb{E}[(S_j^\circ)^2|\mathcal{F}_{(j-1)m}]\bigg) \\ &\leq \mathbb{P}\bigg(\sum_{j=1}^k\Big((S_j^\circ)^2-\mathbb{E}[(S_j^\circ)^2|\mathcal{F}_{(j-1)m}]\Big) < -\frac{1}{4}n\sigma^2\bigg), \end{aligned} \tag{4.19}$$

where the last line follows by (4.6) and the fact that $\delta_m \to 0$ and $m/n \to 0$. Denote

$$\eta_j = \Big(\frac{S_j^\circ}{\sigma\sqrt{n}}\Big)^2 - \mathbb{E}\Big[\Big(\frac{S_j^\circ}{\sigma\sqrt{n}}\Big)^2\Big|\mathcal{F}_{(j-1)m}\Big].$$

Then, by (4.9) and stationarity, it is easy to see that

$$\sum_{j=1}^k\Big\|\mathbb{E}[|\eta_j|^{(2+\rho)/2}|\mathcal{F}_{(j-1)m}]\Big\|_\infty \leq 2^{1+\rho}\sum_{j=1}^k\Big\|\mathbb{E}\Big[\Big|\frac{S_j^\circ}{\sigma\sqrt{n}}\Big|^{2+\rho}\Big|\mathcal{F}_{(j-1)m}\Big]\Big\|_\infty \leq 2^{1+\rho}\varepsilon_m^\rho$$

and that, for some positive constant c,

$$\eta_i \geq -\frac{1}{n\sigma^2}\big\|\mathbb{E}\big[S_m^2\big|\mathcal{F}_0\big]\big\|_\infty \geq -\frac{m}{n}c \quad \text{ a.s.},$$

where the last inequality follows by the fact that $\delta_m \to 0$ as $n \to \infty$. From (4.19), using Lemma 4.2 with $a = \frac{m}{n}c$ and $\beta = (2+\rho)/2$, we have

$$\mathbb{P}\Big(\frac{(V_k^\circ)^2}{n\sigma^2} < \frac{1}{2}\Big) \leq \exp\Big\{-C(\rho)\Big(\frac{1}{\varepsilon_m^2}+\frac{n}{m}\Big)\Big\}, \tag{4.20}$$

where $C(\rho) > 0$ depends only on ρ. Notice that, by (4.13), it holds, for small enough $\alpha_{\rho,0} > 0$ and all $1 \le x \le \alpha_{\rho,0} \min\{\varepsilon_n^{-1}, \sqrt{n/m}\}$,

$$\frac{1}{1-\Phi(x)} \exp\left\{ - C(\rho)\Big(\frac{1}{\varepsilon_m^2} + \frac{n}{m}\Big)\right\} \le \sqrt{2\pi}(1+x)\Big(\sqrt{\frac{m}{n}} + \varepsilon_m^\rho\Big). \tag{4.21}$$

Then, by (4.18), (4.20) and (4.21), we obtain, for all $1 \le x \le \alpha_{\rho,0} \min\{\varepsilon_n^{-1}, \sqrt{n/m}\}$,

$$\begin{aligned}
\frac{\mathbb{P}\Big(W_n^\circ \ge x\Big)}{1-\Phi(x)} &\le \frac{\mathbb{P}\Big(W_n^\circ \ge x,\ \frac{(V_k^\circ)^2}{n\sigma^2} \ge \frac{1}{2}\Big)}{1-\Phi(x)} + \frac{\mathbb{P}\Big(\frac{(V_k^\circ)^2}{n\sigma^2} < \frac{1}{2}\Big)}{1-\Phi(x)} \\
&\le \exp\Bigg\{ c_\rho''\Bigg(x^{2+\rho}\varepsilon_m^\rho + x^2\Big(\delta_m^2 + \gamma_m|\ln\gamma_m| + \frac{m}{n}\Big) \\
&\qquad\qquad + (1+x)\Big(\delta_m + \sqrt{\frac{m}{n}} + \varepsilon_m^\rho + \gamma_m|\ln\gamma_m| + \widehat{\varepsilon}_m(x,\rho)\Big)\Bigg)\Bigg\} \\
&\quad + \frac{1}{1-\Phi(x)} \exp\left\{ - C(\rho)\Big(\frac{1}{\varepsilon_m^2} + \frac{n}{m}\Big)\right\} \\
&\le \exp\Bigg\{ c_\rho'''\Bigg(x^{2+\rho}\varepsilon_m^\rho + x^2\Big(\delta_m^2 + \gamma_m|\ln\gamma_m| + \frac{m}{n}\Big) \\
&\qquad\qquad + (1+x)\Big(\delta_m + \sqrt{\frac{m}{n}} + \varepsilon_m^\rho + \gamma_m|\ln\gamma_m| + \widehat{\varepsilon}_m(x,\rho)\Big)\Bigg)\Bigg\}.
\end{aligned}$$

From the last inequality, we get, for all $1 \le x \le \alpha_\rho \min\{\varepsilon_m^{-1}, \sqrt{n/m}\}$,

$$\begin{aligned}
\ln \frac{\mathbb{P}(W_n^\circ \ge x)}{1-\Phi(x)} \ \le\ & c_\rho'''\Bigg(x^{2+\rho}\varepsilon_m^\rho + x^2\Big(\delta_m^2 + \gamma_m|\ln\gamma_m| + \frac{m}{n}\Big) \\
& + (1+x)\Big(\delta_m + \sqrt{\frac{m}{n}} + \varepsilon_m^\rho + \gamma_m|\ln\gamma_m| + \widehat{\varepsilon}_m(x,\rho)\Big)\Bigg),
\end{aligned} \tag{4.22}$$

which gives the upper bound of $\ln \frac{\mathbb{P}(W_n^\circ \ge x)}{1-\Phi(x)}$ for $\rho \in (0,1)$. The proof of the lower bound of $\ln \frac{\mathbb{P}(W_n^\circ \ge x)}{1-\Phi(x)}$, $1 \le x \le \alpha_{\rho,0} \min\{\varepsilon_n^{-1}, \sqrt{n/m}\}$, is similar to the proof of (4.22), but, instead of using (4.12), we use the following inequalities: for all $x \geq 0$,

$$
\begin{aligned}
\mathbb{P}\Big(W_n^\circ \geq x\Big) \ &\geq \ \mathbb{P}\Big(W_n^\circ \geq x,\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big) \\
&\geq \ \mathbb{P}\Big(\frac{\sum_{j=1}^{k} S_j^\circ/(n^{1/2}\sigma)}{\sqrt{[M]_k}} \geq x\sqrt{1+\kappa_m},\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big) \\
&\geq \ \mathbb{P}\Big(\frac{M_k}{\sqrt{[M]_k}} \geq x(1+\gamma_m|\ln\gamma_m|)\sqrt{1+\kappa_m},\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big) \\
&\quad\ - \mathbb{P}\Big(\frac{1}{n^{1/2}\sigma}\sum_{j=1}^{k}\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}] \geq x\gamma_m|\ln\gamma_m|\sqrt{1+\kappa_m},\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big) \\
&\geq \ \mathbb{P}\Big(\frac{M_k}{\sqrt{[M]_k}} \geq x(1+\gamma_m|\ln\gamma_m|)\sqrt{1+\kappa_m}\Big) - \mathbb{P}\Big(\frac{(V_k^\circ)^2}{n\sigma^2} < \frac{1}{2}\Big) \\
&\quad\ - \mathbb{P}\Big(\frac{1}{n^{1/2}\sigma}\sum_{j=1}^{k}\mathbb{E}[S_j^\circ|\mathcal{F}_{(j-1)m}] \geq x\gamma_m|\ln\gamma_m|\sqrt{1+\kappa_m}\Big) \\
&=: \ P_1(x) - P_2 - P_3(x).
\end{aligned}
\tag{4.23}
$$

By an argument similar to that of (4.15), we deduce that, for all $0 \leq x \leq \alpha_{\rho,0}\varepsilon_m^{-1}$,

$$
\begin{aligned}
\frac{P_1(x)}{1-\Phi(x)} \geq \exp\bigg\{ - c_\rho\Big(& x^{2+\rho}\varepsilon_m^\rho + x^2(\delta_m^2 + \gamma_m|\ln\gamma_m| + \frac{m}{n}) \\
& +(1+x)\Big(\delta_m + \sqrt{\frac{m}{n}} + \varepsilon_m^\rho + \gamma_m|\ln\gamma_m| + \widehat{\varepsilon}_m(x,\rho)\Big)\Big)\bigg\}.
\end{aligned}
\tag{4.24}
$$

By (4.20), we have, for small enough $\alpha_{\rho,0} > 0$ and all $0 \leq x \leq \alpha_{\rho,0}\min\{\varepsilon_n^{-1},\ \sqrt{n/m}\}$,

$$
\begin{aligned}
\frac{P_2}{1-\Phi\,(x)} \ &\leq \ \sqrt{2\pi}(1+x)\exp\Big\{ - C(\rho)\Big(\frac{1}{\varepsilon_m^2} + \frac{n}{m}\Big) + \frac{1}{2}x^2\Big\} \\
&\leq \ C_{\rho,3}\Big(\sqrt{\frac{m}{n}} + \varepsilon_m^\rho\Big)\exp\Big\{ - \frac{1}{2}x^2\Big\}.
\end{aligned}
\tag{4.25}
$$

By an argument similar to that of (4.17), we get, for all $x \geq 1$,

$$
\begin{aligned}
\frac{P_3(x)}{1-\Phi\,(x)} \ &\leq \ C_1(1+x)\exp\Big\{ - C_0x^2(\ln\gamma_m)^2 + \frac{1}{2}x^2\Big\} \\
&\leq \ C_4\,\gamma_m|\ln\gamma_m|\exp\Big\{ - \frac{1}{2}x^2\Big\}.
\end{aligned}
\tag{4.26}
$$

Combining the inequalities (4.23)-(4.26) together, we obtain, for all $1 \leq x \leq \alpha_{\rho,0} \min\{\varepsilon_n^{-1},\ \sqrt{n/m}\}$,

$$\begin{aligned}\ln \frac{\mathbb{P}(W_n^\circ \geq x)}{1-\Phi\left(x\right)} &\geq -C_\rho\Big(x^{2+\rho}\varepsilon_m^\rho + x^2\Big(\delta_m^2 + \gamma_m|\ln\gamma_m| + \frac{m}{n}\Big)\\ &\qquad + (1+x)\Big(\delta_m + \sqrt{\frac{m}{n}} + \varepsilon_m^\rho + \gamma_m|\ln\gamma_m| + \widehat{\varepsilon}_m(x,\rho)\Big)\Big).\end{aligned}$$

This completes the proof of Theorem 2.1 for all $1 \leq x \leq \alpha_{\rho,0} \min\{\varepsilon_n^{-1},\ \sqrt{n/m}\}$.

For the case $0 \leq x \leq 1$, instead of (4.12), we make use of the following estimations:

$$\begin{aligned}\mathbb{P}\Big(W_n^\circ \geq x,\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big) &\leq \mathbb{P}\Big(\frac{\sum_{j=1}^k S_j^\circ/(n^{1/2}\sigma)}{\sqrt{[M]_k}} \geq x\sqrt{1-\kappa_m},\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big)\\ &\leq \mathbb{P}\Big(\frac{M_k}{\sqrt{[M]_k}} \geq (x-\gamma_m|\ln\gamma_m|)\sqrt{1-\kappa_m},\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big) + I_2(1)\\ &\leq \mathbb{P}\Big(\frac{M_k}{\sqrt{[M]_k}} \geq (x-\gamma_m|\ln\gamma_m|)\sqrt{1-\kappa_m}\Big) + I_2(1)\\ &=: \widetilde{I}_1(x) + I_2(1).\end{aligned}$$

By an argument similar to the case of $1 \leq x \leq \alpha_{\rho,0} \min\{\varepsilon_n^{-1},\ \sqrt{n/m}\}$, we obtain the upper bound of $\ln \frac{\mathbb{P}(W_n^\circ \geq x)}{1-\Phi\left(x\right)}$ for all $0 \leq x \leq 1$. To prove the lower bound of $\ln \frac{\mathbb{P}(W_n^\circ \geq x)}{1-\Phi\left(x\right)}, 0 \leq x \leq 1$, instead of (4.12), we should use the following estimations:

$$\begin{aligned}\mathbb{P}\Big(W_n^\circ \geq x\Big) &\geq \mathbb{P}\Big(W_n^\circ \geq x,\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big)\\ &\geq \mathbb{P}\Big(\frac{\sum_{j=1}^k S_j^\circ/(n^{1/2}\sigma)}{\sqrt{[M]_k}} \geq x\sqrt{1+\kappa_m},\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big)\\ &\geq \mathbb{P}\Big(\frac{M_k}{\sqrt{[M]_k}} \geq (x+\gamma_m|\ln\gamma_m|)\sqrt{1+\kappa_m},\ \frac{(V_k^\circ)^2}{n\sigma^2} \geq \frac{1}{2}\Big) - I_2(1)\\ &\geq \mathbb{P}\Big(\frac{M_k}{\sqrt{[M]_k}} \geq (x+\gamma_m|\ln\gamma_m|)\sqrt{1+\kappa_m}\Big) - \mathbb{P}\Big(\frac{(V_k^\circ)^2}{n\sigma^2} < \frac{1}{2}\Big) - I_2(1).\end{aligned}$$

Again by an argument similar to the case of $1 \leq x \leq \alpha_{\rho,0} \min\{\varepsilon_n^{-1},\ \sqrt{n/m}\}$, we get the lower bound of $\ln \frac{\mathbb{P}(W_n^\circ \geq x)}{1-\Phi\left(x\right)}$ for all $0 \leq x \leq 1$. This completes the proof of Theorem 2.1 for $\rho \in (0,1)$.

For $\rho = 1$, the proof of Theorem 2.1 is similar to the case of $\rho \in (0,1)$, where the term $\varepsilon_m|\ln\varepsilon_m|$ comes from point [ii] of Lemma 4.1 with $\epsilon_n = C_{1,0}\varepsilon_m$. Notice that if $(X_i)_{i\in\mathbb{Z}}$ satisfies the condition of Theorem 2.1, then $(-X_i)_{i\in\mathbb{Z}}$ also satisfies the same condition. Thus the assertions in Theorem 2.1 remain valid when $\frac{\mathbb{P}(W_n^\circ \geq x)}{1-\Phi\left(x\right)}$ is replaced by $\frac{\mathbb{P}(W_n^\circ \leq -x)}{\Phi\left(-x\right)}$, $x \geq 0$.

4.3. Proof of Corollary 2.1

First, we prove that

$$\limsup_{n\to\infty} a_n^2 \ln \mathbb{P}\Big(a_n W_n^\circ \in B \Big) \le - \inf_{x \in \overline{B}} \frac{x^2}{2}. \tag{4.27}$$

For any given Borel set $B \subset \mathbb{R}$, let $x_0 = \inf_{x\in B}|x| \ge \inf_{x\in\overline{B}}|x|$. By Theorem 2.1, we deduce that

$$\begin{aligned}
&\mathbb{P}\Big(a_n W_n^\circ \in B \Big)\\
&\le \mathbb{P}\Big(W_n^\circ \ge \frac{x_0}{a_n} \Big) + \mathbb{P}\Big(W_n^\circ \le -\frac{x_0}{a_n} \Big)\\
&\le 2\Big(1 - \Phi\Big(\frac{x_0}{a_n}\Big)\Big) \exp\Bigg\{ C\Big((\frac{x_0}{a_n})^{2+\rho}\varepsilon_m^\rho + \Big(\frac{x_0}{a_n}\Big)^2 \Big(\delta_m^2 + \gamma_m|\ln\gamma_m| + \frac{m}{n}\Big)\\
&\qquad\qquad + (1 + \frac{x_0}{a_n})\Big(\delta_m + \gamma_m|\ln\gamma_m| + \varepsilon_m^{\rho/4} + \sqrt{\frac{m}{n}}\Big)\Big)\Bigg\}.
\end{aligned}$$

Notice that $a_n \to 0$ and $a_n \min\{\varepsilon_m^{-1}, \sqrt{n/m}\} \to \infty$ as $n\to\infty$. Using (4.13) and (2.4), we deduce that

$$\limsup_{n\to\infty} a_n^2 \ln \mathbb{P}\Big(a_n W_n^\circ \in B \Big) \ \le\ -\frac{x_0^2}{2} \ \le\ -\inf_{x\in\overline{B}} \frac{x^2}{2},$$

which gives (4.27).

Next, we prove that

$$\liminf_{n\to\infty} a_n^2 \ln \mathbb{P}\Big(a_n W_n^\circ \in B \Big) \ge - \inf_{x \in B^o} \frac{x^2}{2}. \tag{4.28}$$

Without loss of generality, we assume that $B^o \ne \emptyset$, otherwise (4.28) holds obviously, since in this case the infimum of a function over an empty set is equal to ∞ by convention. For any given $\varepsilon_1 > 0$, there exists an $x_0 \in B^o$ such that

$$0 < \frac{x_0^2}{2} \le \inf_{x\in B^o} \frac{x^2}{2} + \varepsilon_1. \tag{4.29}$$

We only consider the case when $x_0 > 0$, the case $x_0 < 0$ being proved in the same way. Since B^o is an open set, for $x_0 \in B^o$ and small enough $\varepsilon_2 \in (0, x_0)$ it holds $(x_0 - \varepsilon_2, x_0 + \varepsilon_2] \subset B$. Clearly, $x_0 \ge \inf_{x\in\overline{B}} x$. It is easy to see that

$$\begin{aligned}
\mathbb{P}\Big(a_n W_n^\circ \in B \Big) \ &\ge\ \mathbb{P}\Big(W_n^\circ \in (a_n^{-1}(x_0-\varepsilon_2), a_n^{-1}(x_0+\varepsilon_2)] \Big)\\
&\ge\ \mathbb{P}\Big(W_n^\circ > a_n^{-1}(x_0-\varepsilon_2) \Big) - \mathbb{P}\Big(W_n^\circ > a_n^{-1}(x_0+\varepsilon_2) \Big).
\end{aligned}$$

By Theorem 2.1, we have

$$\lim_{n\to\infty} \frac{\mathbb{P}\Big(W_n^o > a_n^{-1}(x_0+\varepsilon_2)\Big)}{\mathbb{P}\Big(W_n^o > a_n^{-1}(x_0-\varepsilon_2)\Big)} = 0.$$

Again, by Theorem 2.1, (4.13) and (2.4), it follows that

$$\liminf_{n\to\infty} a_n^2 \ln \mathbb{P}\Big(a_n W_n^\circ \in B\Big) \ge \liminf_{n\to\infty} a_n^2 \ln \frac{1}{2}\mathbb{P}\Big(W_n^\circ > a_n^{-1}(x_0-\varepsilon_2)\Big) = -\frac{1}{2}(x_0-\varepsilon_2)^2.$$

Letting $\varepsilon_2 \to 0$, we obtain

$$\liminf_{n\to\infty} a_n^2 \ln \mathbb{P}\Big(a_n W_n^\circ \in B\Big) \;\ge\; -\frac{x_0^2}{2} \;\ge\; -\inf_{x\in B^o}\frac{x^2}{2} - \varepsilon_1.$$

Since $\varepsilon_1 > 0$ can be arbitrarily small, we get (4.28). The proof of Corollary 2.1 is complete.

4.4. Proof of Corollary 2.2

We only need to consider the case where $\max\{\gamma_m, \varepsilon_m, \delta_m, m/n\} \le 1/10$. Otherwise, Corollary 2.2 holds obviously by choosing C_ρ large enough. Denote

$$\kappa_n = \min\{\gamma_m^{-1/4},\, \varepsilon_m^{-\rho(2-\rho)/8}, \delta_m^{-1/4}, (m/n)^{-1/4}\}.$$

It is easy to see that

$$\begin{aligned}
\sup_x \Big|\mathbb{P}(W_n^\circ \le x) - \Phi(x)\Big| \;&\le\; \sup_{|x|\le\kappa_n}\Big|\mathbb{P}(W_n^\circ \le x) - \Phi(x)\Big| + \sup_{|x|>\kappa_n}\Big|\mathbb{P}(W_n^\circ \le x) - \Phi(x)\Big| \\
&=\; \sup_{|x|\le\kappa_n}\Big|\mathbb{P}(W_n^\circ \le x) - \Phi(x)\Big| \\
&\quad + \sup_{x<-\kappa_n} \mathbb{P}(W_n^\circ \le x) + \sup_{x<-\kappa_n} \Phi(x) \\
&\quad + \sup_{x>\kappa_n} \mathbb{P}(W_n^\circ > x) + \sup_{x>\kappa_n} (1-\Phi(x)). \qquad (4.30)
\end{aligned}$$

Notice that

$$\sup_{|x|\le\kappa_n} \{\varepsilon_m^\rho |\ln \varepsilon_m|, \widehat{\varepsilon}_m(x,\rho)\} = \varepsilon_m^{\rho(2-\rho)/4}.$$

By Theorem 2.1 and the inequality $|e^x - 1| \le |x|e^{|x|}$, we have

$$\begin{aligned}
&\sup_{|x|\le\kappa_n}\Big|\mathbb{P}(W_n^\circ \le x) - \Phi(x)\Big| \\
&\le \sup_{|x|\le\kappa_n}\Big(1-\Phi(|x|)\Big)\Big|e^{C_\rho\left(x^{2+\rho}\varepsilon_m^\rho + x^2\left(\delta_m^2 + \gamma_m|\ln\gamma_m| + \frac{m}{n}\right) \;+\; (1+x)\left(\delta_m + \gamma_m|\ln\gamma_m| + \varepsilon_m^{\rho(2-\rho)/4} + \sqrt{\frac{m}{n}}\right)\right)} - 1\Big| \\
&\le C_{\rho,1}\Big(\delta_m + \gamma_m|\ln\gamma_m| + \varepsilon_m^{\rho(2-\rho)/4} + \sqrt{\frac{m}{n}}\Big). \qquad (4.31)
\end{aligned}$$

From the last inequality, we get

$$\begin{aligned}\sup_{x<-\kappa_n}\mathbb{P}(W_n^\circ\le x) &= \mathbb{P}(W_n^\circ\le -\kappa_n)\\ &\le C_{\rho,1}\Big(\delta_m+\gamma_m|\ln\gamma_m|+\varepsilon_m^{\rho(2-\rho)/4}+\sqrt{\frac{m}{n}}\Big)+\Phi\left(-\kappa_n\right)\\ &\le C_{\rho,2}\Big(\delta_m+\gamma_m|\ln\gamma_m|+\varepsilon_m^{\rho(2-\rho)/4}+\sqrt{\frac{m}{n}}\Big). \qquad (4.32)\end{aligned}$$

Similarly, we have

$$\sup_{x>\kappa_n}\mathbb{P}(W_n^\circ> x) \le C_{\rho,3}\Big(\delta_m+\gamma_m|\ln\gamma_m|+\varepsilon_m^{\rho(2-\rho)/4}+\sqrt{\frac{m}{n}}\Big). \qquad (4.33)$$

Clearly, it holds that

$$\sup_{x>\kappa_n}\left(1-\Phi\left(x\right)\right)=\sup_{x<-\kappa_n}\Phi\left(x\right)=\Phi\left(-\kappa_n\right)\le C_{\rho,4}\Big(\delta_m+\gamma_m|\ln\gamma_m|+\varepsilon_m^{\rho(2-\rho)/4}+\sqrt{\frac{m}{n}}\Big). \quad (4.34)$$

Combining the inequalities (4.30)-(4.34) together, we obtain the desired inequality.

4.5. Proof of Proposition 3.1

We only need to show that the quantities γ_m and δ_m can be dominated via the quantities $\eta_{1,n}$ and $\eta_{2,n}$. By the definition of γ_m, it is easy to see that

$$\gamma_m \le \frac{1}{m^{1/2}\sigma}\sum_{j=1}^{\infty}\frac{1}{j^{3/2}}\Big(\sum_{i=1}^{mj}\eta_{1,i}\Big).$$

Thus, when $\eta_{1,n}=O(n^{-\beta})$ for some $\beta>1$, it holds

$$\gamma_m=O(1/m^{1/2}).$$

Next, we give an estimation for δ_m. It is obvious that

$$\|\mathbb{E}[S_m|\mathcal{F}_0]\|_\infty\le\sum_{i=1}^{m}\eta_{1,i}$$

and

$$\Big\|\frac{1}{m\sigma^2}\mathbb{E}[S_m^2|\mathcal{F}_0]-1\Big\|_\infty\le\frac{1}{m\sigma_n^2}\Big(\|\mathbb{E}[S_m^2|\mathcal{F}_0]-\mathbb{E}[S_m^2]\|_\infty+|\mathbb{E}[S_m^2]-m\sigma^2|\Big).$$

Clearly, it holds

$$\|\mathbb{E}[S_m^2|\mathcal{F}_0]-\mathbb{E}[S_m^2]\|_\infty \le \sum_{i=1}^m \|\mathbb{E}[X_i^2|\mathcal{F}_0]-\mathbb{E}[X_i^2]\|_\infty + 2\sum_{i=1}^{m-1}\sum_{j=i+1}^m \|\mathbb{E}[X_iX_j|\mathcal{F}_0]-\mathbb{E}[X_iX_j]\|_\infty.$$

Splitting the last sum as

$$\sum_{1\le i\le m/2}\sum_{i+1\le j\le 2i} + \sum_{1\le i\le m/2}\sum_{2i+1\le j\le m} + \sum_{m/2\le i\le m-1}\sum_{i+1\le j\le m},$$

by the condition $\max_{i=1,2}\{\eta_{i,n}\} = O(n^{-\beta})$, we infer that

$$\|\mathbb{E}[S_m^2|\mathcal{F}_0]-\mathbb{E}[S_m^2]\|_\infty \le C_1\bigg(\sum_{i=1}^m i^{-\beta} + \sum_{1\le i\le m/2} i\eta_{2,i} + \|X_0\|_\infty \sum_{1\le i\le m/2}\sum_{j\ge i}\eta_{1,j} + m\sum_{i\ge m/2}\eta_{2,i}\bigg),$$

Notice that

$$|\mathbb{E}[X_iX_0]| = |\mathbb{E}[X_0\mathbb{E}[X_i|\mathcal{F}_0]]| \le \|X_0\|_\infty\|\mathbb{E}[X_i|\mathcal{F}_0]\|_\infty.$$

By $\eta_{1,n} = O(n^{-\beta}), \beta > 1$, it is easy to see that

$$\begin{aligned}
|\mathbb{E}[S_m^2]-m\sigma^2| &\le \sum_{i=1}^m \Big|\sum_{j=1}^m \mathbb{E}[X_iX_j] - \sum_{j=-\infty}^{\infty}\mathbb{E}[X_iX_j]\Big| \\
&= \sum_{i=1}^m \Big|\sum_{j=-\infty}^{0} \mathbb{E}[X_iX_j] + \sum_{j=m+1}^{\infty}\mathbb{E}[X_iX_j]\Big| \\
&\le \|X_0\|_\infty \sum_{i=1}^m \Big(\sum_{j=-\infty}^{-i} O(|j|^{-\beta}) + \sum_{j=m+1-i}^{\infty} O(j^{-\beta})\Big) \\
&\le C_2\sum_{i=1}^m i^{-\beta} \\
&\le C_3\,.
\end{aligned}$$

Hence, it holds

$$\delta_m^2 \le \frac{C_1}{m\sigma_n^2}\bigg[\Big(\sum_{i=1}^m \eta_{1,i}\Big)^2 + \sum_{i=1}^m i^{-\beta} + \sum_{1\le i\le m/2} i\eta_{2,i} + \|X_0\|_\infty \sum_{1\le i\le m/2}\sum_{j\ge i}\eta_{1,j} + m\sum_{i\ge m/2}\eta_{2,i} + C_4\bigg].$$

Then, taking into account that $\max_{i=1,2}\{\eta_{i,n}\} = O(n^{-\beta})$, we have

$$\delta_m = \begin{cases} O(m^{-(\beta-1)/2}), & \text{if } \beta \in (1,2), \\ O(m^{-1/2}\sqrt{\ln m}), & \text{if } \beta = 2, \\ O(m^{-1/2}), & \text{if } \beta > 2. \end{cases}$$

By point 2 of Remark 2.1, we have $\varepsilon_m = O(m/n^{1/2})$. If $\beta \geq 3/2$, then equality (2.8) with $m = \lfloor n^{2/7} \rfloor$ holds uniformly for $0 \leq x = o(n^{1/14}/\sqrt{\ln n})$ as $n \to \infty$. If $\beta \in (1, 3/2)$, then equality (2.8) with $m = \lfloor n^{1/(3\beta-1)} \rfloor$ holds uniformly for $0 \leq x = o(n^{(\beta-1)/(6\beta-2)})$ as $n \to \infty$. This completes the proof of points [i] and [ii].

To prove [iii], notice that $m := m(n) \to \infty$ and $n^{1/2}/m \to \infty$ imply $\varepsilon_m, \gamma_m, \delta_m \to 0$ as $n \to \infty$. Then, point [iii] follows by Corollary 2.1. □

Acknowledgements

The authors are deeply indebted to the editor and the anonymous referee for their helpful comments. The work has been supported by the National Natural Science Foundation of China (Grant nos. 11601375, 11971063, 11571052 and 11731012). The work has also benefited from the support of the Centre Henri Lebesgue (CHL, ANR-11-LABX-0020-01).

References

[1] Bentkus, V., Götze, F. (1996). The Berry-Esseen bound for Student's statistic. *Ann. Probab.* **24**(1): 491–501.

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

[3] Bikelis, A. (1966). Estimates of the remainder in the central limit theorem, *Litovsk. Mat. Sb.* **6**(3): 323–46.

[4] Caron, E., Dede, S. (2018). Asymptotic distribution of the least squares estimators for linear models with dependent errors: regular designs. *Math. Methods Statist.* **27**(4): 268–293.

[5] Chen, X., Shao, Q.M., Wu, W.B., Xu, L. (2016). Self-normalized Cramér-type moderate deviations under dependence. *Ann. Statist.* **44**(4): 1593–1617.

[6] Chung, K.L. (1946). The approximate distribution of Student's statistic. *Ann. Math. Statist.* **17**(4): 447–465.

[7] Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actualite's Sci. Indust.* **736**: 5–23.

[8] Csörgő, M., Szyszkowicz, B., Wang, Q. (2003). Donsker's theorem for self-normalized partial sums processes. *Ann. Probab.* **31**(3): 1228–1240.

[9] Cuny, C., Merlevède, F. (2014). On martingale approximations and the quenched weak inviariance principle. *Ann. Probab.* **42**(2): 760–793.

[10] Dedecker, J., Merlevède, F., Peligrad, M., Utev, S. (2009). Moderate deviations for stationary sequences of bounded random variables. *Ann. Inst. H. Poincaré Probab. Statist.* **45**(2): 453–476.

[11] Dembo, A., Shao, Q.M. (2006). Large and moderate deviations for Hotelling's T^2-statistics. *Electron. Comm. Proba.* **11**: 149–159.

[12] de la Peña, V.H. (1999). A general class of exponential inequalities for martingales and ratios. *Ann. Probab.* **27**(1): 537–564.

[13] de la Peña, V.H., Lai, T.L., Shao, Q.M. (2009). *Self-normalized Processes: Theory and Statistical Applications.* Springer Series in Probability and its Applications. Springer-Verlag. New York.

[14] Fan X, Grama I, Liu Q. (2013). Cramér large deviation expansions for martingales under Bernstein's condition. *Stochastic Process. Appl.* **123**(11): 3919–3942.

[15] Fan, X. (2017). Sharp large deviation results for sums of bounded from above random variables. *Sci. China Math.* **60**(12): 2465–2480.

[16] Fan, X., Grama, I., Liu, Q., Shao, Q.M. (2018). Self-normalized Cramér type moderate deviations for martingales. *Bernoulli* **25**(4A), 2793–2823.

[17] Hannan, E.J. (1973). Central limit theorems for time series regression. *Probability Theory and Related Fields* **26**, 157–170.

[18] Gao, F.Q. (1996). Moderate deviations for martingales and mixing random processes. *Stochastic Process. Appl.* **61**: 263–275.

[19] Grama, I. (1997). On moderate deviations for martingales. *Ann. Probab.* **25**: 152–184.

[20] Grama, I., Haeusler, E. (2000). Large deviations for martingales via Cramér's method. *Stochastic Process. Appl.* **85**: 279–293.

[21] Giné, E., Götze, F., Mason, D.M. (1997). When is the Student t-statistic asymptotically standard normal? *Ann. Probab.* **25**(3): 1514–1531.

[22] Jing, B.Y., Shao, Q.M., Wang, Q. (2003). Self-normalized Cramér-type large deviations for independent random variables. *Ann. Probab.* **31**(4): 2167–2215.

[23] Liu, W., Shao, Q.M., Wang, Q. (2013). Self-normalized Cramér type moderate deviations for the maximum of sums. *Bernoulli* **19**(3): 1006–1027.

[24] Liu, W., Shao, Q.M. (2013). A Cramér moderate deviation theorem for Hotelling's T^2-statistic with applications to global tests. *Ann. Statist.* **41**(1): 296–322.

[25] Linnik, Y.V. (1961). 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, pp. 289-306). Univ of California Press.

[26] Novak, S.Y. (2011). Extreme value methods with applications to finance. London: Chapman & Hall/CRC Press.

[27] Peligrad, M., Utev, S., Wu, W.B. (2007). A maximal L_p-inequality for stationary sequecens and its applications. *Proc. Amer. Math. Soc.* **135**: 541–550.

[28] Petrov, V.V. (1954). A generalization of Cramér's limit theorem. *Uspekhi Math. Nauk* **9**: 195–202.

[29] Račkauskas, A. (1995). Large deviations for martingales with some applications. *Acta Appl. Math.* **38**: 109–129.

[30] Račkauskas, A. (1997). Limit theorems for large deviations probabilites of certain quadratic forms. *Lithuanian Math. J.* **37**: 402–415.

[31] Rio, E. (2009). Moment inequalities for sums of dependent random variables under projective condition. *J. Theor. Probab.* **22**: 146–163.

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

[33] Shao, Q.M. (1999). A Cramér type large deviation result for Student's $t-$statistic. *J. Theor. Probab.* **12**(2): 385–398.

[34] Shao, Q.M. (2018). On necessary and sufficient conditions for the self-normalized central limit theorem. *Sci. China Math.* **61**(10): 1741–1748.

[35] Shao, Q.M., Wang, Q.Y. (2013). Self-normalized limit theorems: A survey. *Probab. Surv.* **10**: 69–93.

[36] Saulis, L. and Statulevičius, V.A. (1978). *Limit theorems for large deviations.* Kluwer Academic Publishers.

[37] Wu, W.B. (2005). Nonlinear system theorey: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102**: 14150–14154.

[38] Wu, W.B., Zhao, Z. (2008). Moderate deviations for stationary processes. *Statist. Sinica* **18**: 769–782.