Cramér large deviation expansions for martingales under Bernstein's condition

Xiequan Fan, Ion Grama, Quansheng Liu

Cramér large deviation expansions for martingales under Bernstein's condition

Xiequan Fan*, Ion Grama and Quansheng Liu

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

Abstract

An expansion of large deviation probabilities for martingales is given, which extends the classical result due to Cramér to the case of martingale differences satisfying the conditional Bernstein condition. The upper bound of the range of validity and the remainder of our expansion is the same as in the Cramér result and therefore are optimal. Our result implies a moderate deviation principle for martingales.

Keywords: expansions of large deviations; Cramér type large deviations; large deviations; moderate deviations; exponential inequality; Bernstein's condition; central limit theorem

2000 MSC: Primary 60G42; 60F10; 60E15; Secondary 60F05

1. Introduction

Consider a sequence of independent and identically distributed (i.i.d.) centered real random variables $\xi_1, ..., \xi_n$ satisfying Cramér's condition $\mathbb{E} \exp\{c_0|\xi_1|\} < \infty$, for some constant $c_0 > 0$. Denote $\sigma^2 = \mathbb{E}\xi_1^2$ and $X_n = \sum_{i=1}^n \xi_i$. In 1938, Cramér [5] established an asymptotic expansion of the probabilities of large deviations of X_n, based on the powerful technique of conjugate distributions (see also Esscher [8]). The results of Cramér imply that, uniformly in $1 \leq x = o(n^{1/2})$,

$$\log \frac{\mathbb{P}(X_n > x\sigma\sqrt{n})}{1 - \Phi(x)} = O\Big(\frac{x^3}{\sqrt{n}}\Big) \quad \text{as} \quad n \to \infty, \tag{1}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\{-t^2/2\}dt$ is the standard normal distribution. Various large deviation expansions for sums of independent random variables have been obtained by many authors, see for instance Feller [10], Petrov [22], Rubin and Sethuraman [27], Statulevičius [29], Saulis and Statulevičius [28] and Bentkus and Račkauskas [1]. We refer to the book of Petrov

*Corresponding author.
E-mail: fanxiequan@hotmail.com (X. Fan), ion.grama@univ-ubs.fr (I. Grama), quansheng.liu@univ-ubs.fr (Q. Liu).

[23] and the references therein for a detailed account. Despite the fact that the case of sums of independent random variables is well studied, there are only a few results on expansions of type (1) for martingales: see Bose [3, 4], Račkauskas [24, 25, 26], Grama [13, 14] and Grama and Haeusler [15, 16]. It is also worth noting that limit theorems for large and moderate deviation principle for martingales have been proved by several authors, see e.g. Liptser and Pukhalskii [21], Gulinsky and Veretennikov [17], Gulinsky, Liptser and Lototskii [18], Gao [12], Dembo [6], Worms [30] and Djellout [7]. However, these theorems are less precise than large deviation expansions of type (1).

Let $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$ be a sequence of square integrable martingale differences defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\xi_0 = 0$ and $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq ... \subseteq \mathcal{F}_n \subseteq \mathcal{F}$. Denote $X_n = \sum_{i=1}^n \xi_i$. Assume that there exist absolute constants $H > 0$ and $N \geq 0$ such that $\max_i |\xi_i| \leq H$ and $|\sum_{i=1}^n \mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) - n| \leq N^2$. Here and hereafter, the equalities and inequalities between random variables are understood in the $\mathbb{P}$-almost sure sense. From the results in Grama and Haeusler [15], it follows that, for any constant $\alpha > 0$ and $\alpha\sqrt{\log n} \leq x = o\left(n^{1/6}\right)$,

$$\frac{\mathbb{P}\left(X_n > x\sqrt{n}\right)}{1 - \Phi\left(x\right)} = 1 + O\Big((H+N)\frac{x^3}{\sqrt{n}}\Big) \tag{2}$$

and, for any $0 \leq x = O\left(\sqrt{\log n}\right)$,

$$\frac{\mathbb{P}\left(X_n > x\sqrt{n}\right)}{1 - \Phi\left(x\right)} = 1 + O\Big((H+N)(1+x)\frac{\log n}{\sqrt{n}}\Big) \tag{3}$$

as $n \to \infty$ (se also [14, 16] for more results in the last range). In this paper we extend the expansions (2) and (3) to the case of martingale differences $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$ satisfying the conditional Bernstein condition,

$$|\mathbb{E}(\xi_i^k|\mathcal{F}_{i-1})| \leq \frac{1}{2}k!H^{k-2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}), \qquad \text{for } \ k \geq 3 \ \text{ and } \ 1 \leq i \leq n, \tag{4}$$

where H is a positive absolute constant. Note that in the i.i.d. case Bernstein's condition (4) is equivalent to Cramér's condition (see Section 8) and therefore (2) implies Cramér's expansion (1). It is worth stressing that the remainder in expansion (2) is of the same order as that in (1) in the stated range and therefore cannot be improved. As to the remainder in (3), from the rate of convergence result in Bolthausen [2] we conclude that it is also optimal.

Another objective of the paper is to find an asymptotic expansion of large deviation for martingales in a wider range than that of (2). From Theorems 2.1 and 2.2 of the paper it follows that, for any constant $\alpha > 0$ and $\alpha\sqrt{\log n} \leq x = o\left(n^{1/2}\right)$,

$$\log\frac{\mathbb{P}\left(X_n > x\sqrt{n}\right)}{1 - \Phi\left(x\right)} = O\Big(\frac{x^3}{\sqrt{n}}\Big) \ \text{ as } \ n \to \infty. \tag{5}$$

This improves the corresponding result in [15] where (5) has been established only in the range $x \in [\alpha\sqrt{\log n}, \alpha_1 n^{1/4}]$ for some absolute constant $\alpha_1 > 0$. The upper bound of the range and the remainder in expansion (5) cannot be improved since they are of the same order as in the Cramér's expansion (1).

The idea behind our approach is similar to that of Cramér for independent random variables with corresponding adaptations to the martingale case. We make use of the conjugate multiplicative martingale for changing the probability measure as proposed in Grama and Haeusler [15] (see also [9]). However, we refine [15] in two aspects. First, we relax the boundedness condition $|\xi_i| \leq L$, replacing it by Bernstein's condition (4). Secondly, we establish upper and lower bounds for the large deviation probabilities in the range $x \in [0, \alpha_1 n^{1/2})$ thus enlarging the range $x \in [0, \alpha_1 n^{1/4}]$ established in [15]. In the proof we make use of a rate of convergence result for martingales under the conjugate measure. It is established under the Bernstein condition (4), unlike [15] where it is established only for bounded martingale differences. As a consequence, we improve the result on the rate of convergence in the central limit theorem (CLT) due to Bolthausen [2] (see Theorem 3.1 below).

The paper is organized as follows. Our main results are stated and discussedin in Section 2. A rate of convergence in the CLT for martingales is given in Section 3. Section 4 contains auxiliary assertions used in the proofs of the main results. Proofs are deferred to Sections 5, 6 and 7. We clarify the relations among the conditions of Bernstein, Cramér and Sakhanenko in Section 8.

Throughout the paper, c and c_α, probably supplied with some indices, denote respectively a generic positive absolute constant and a generic positive constant depending only on α. Moreover, θ_i's stand for values satisfying $|\theta_i| \leq 1$.

2. Main results

2.1. Main theorems

Assume that we are given a sequence of martingale differences $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\xi_0 = 0$, $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq ... \subseteq \mathcal{F}_n \subseteq \mathcal{F}$ are increasing σ-fields and $(\xi_i)_{i=1,...,n}$ are allowed to depend on n. Set

$$X_0 = 0, \qquad X_k = \sum_{i=1}^{k} \xi_i, \quad k = 1, ..., n. \tag{6}$$

Let $\langle X \rangle$ be the quadratic characteristic of the martingale $X = (X_k, \mathcal{F}_k)_{k=0,...,n}$:

$$\langle X \rangle_0 = 0, \qquad \langle X \rangle_k = \sum_{i=1}^{k} \mathbb{E}(\xi_i^2 | \mathcal{F}_{i-1}), \quad k = 1, ..., n. \tag{7}$$

In the sequel we shall use the following conditions:

(A1) There exists a number $\epsilon \in (0, \frac{1}{2}]$ such that

$$|\mathbb{E}(\xi_i^k | \mathcal{F}_{i-1})| \leq \frac{1}{2} k! \epsilon^{k-2} \mathbb{E}(\xi_i^2 | \mathcal{F}_{i-1}), \qquad \text{for } k \geq 3 \text{ and } 1 \leq i \leq n;$$

(A2) There exists a number $\delta \in [0, \frac{1}{2}]$ such that $|\langle X\rangle_n - 1| \leq \delta^2$.

Note that in the case of normalized sums of i.i.d. random variables conditions (A1) and (A2) are satisfied with $\epsilon = \frac{1}{\sigma\sqrt{n}}$ and $\delta = 0$ (see conditions (A1$'$) and (A2$'$) below). In the case of martingales ϵ and δ usually depend on n such that $\epsilon = \epsilon_n \to 0$ and $\delta = \delta_n \to 0$.

The following two theorems give upper and lower bounds for large deviation probabilities.

Theorem 2.1. *Assume conditions (A1) and (A2). Then, for any constant $\alpha \in (0, 1)$ and all $0 \leq x \leq \alpha\,\epsilon^{-1}$, we have*

$$\frac{\mathbb{P}(X_n > x)}{1 - \Phi\,(x)} \leq \exp\left\{c_\alpha\Big(x^3\epsilon + x^2\delta^2\Big)\right\}\Big(\ 1 + c_\alpha\,(1 + x)\,(\epsilon\,|\log\epsilon| + \delta)\ \Big) \tag{8}$$

and

$$\frac{\mathbb{P}(X_n < -x)}{\Phi\,(-x)} \leq \exp\left\{c_\alpha\Big(x^3\epsilon + x^2\delta^2\Big)\right\}\Big(\ 1 + c_\alpha\,(1 + x)\,(\epsilon\,|\log\epsilon| + \delta)\ \Big), \tag{9}$$

where the constant c_α does not depend on $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$, n and x.

Theorem 2.2. *Assume conditions (A1) and (A2). Then there is an absolute constant $\alpha_0 > 0$ such that, for all $0 \leq x \leq \alpha_0\,\epsilon^{-1}$ and $\delta \leq \alpha_0$,*

$$\frac{\mathbb{P}(X_n > x)}{1 - \Phi\,(x)} \geq \exp\left\{ - c_{\alpha_0}\Big(x^3\epsilon + x^2\delta^2 + (1 + x)\,(\epsilon\,|\log\epsilon| + \delta)\ \Big)\right\} \tag{10}$$

and

$$\frac{\mathbb{P}(X_n < -x)}{\Phi\,(-x)} \geq \exp\left\{ - c_{\alpha_0}\Big(x^3\epsilon + x^2\delta^2 + (1 + x)\,(\epsilon\,|\log\epsilon| + \delta)\ \Big)\right\}, \tag{11}$$

where the constants α_0 and c_{α_0} do not depend on $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$, n and x.

Using the inequality $|e^x - 1| \leq e^\alpha|x|$ valid for $|x| \leq \alpha$, from Theorems 2.1 and 2.2, we obtain the following improvement of the main result of [15].

Corollary 2.1. *Assume conditions (A1) and (A2). Then there is an absolute constant $\alpha_0 > 0$ such that, for all $0 \leq x \leq \alpha_0\, \min\{(\epsilon\,|\log\epsilon|)^{-1}, \delta^{-1}\}$,*

$$\frac{\mathbb{P}(X_n > x)}{1 - \Phi\,(x)} = \exp\{\theta_1 c_{\alpha_0} x^3\epsilon\}\ \Big(\ 1 + \theta_2 c_{\alpha_0}(1 + x)(\epsilon\,|\log\epsilon| + \delta)\Big) \tag{12}$$

and

$$\frac{\mathbb{P}(X_n < -x)}{\Phi\,(-x)} = \exp\{\theta_3 c_{\alpha_0} x^3\epsilon\}\ \Big(\ 1 + \theta_4 c_{\alpha_0}(1 + x)(\epsilon\,|\log\epsilon| + \delta)\Big)\,, \tag{13}$$

where c_{α_0} does not depend on n, x but θ_i possibly depend on $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$, n and x.

For bounded martingale differences $|\xi_i| \leq \epsilon$ under condition (A2), Grama and Haeusler [15] proved the asymptotic expansions (12) and (13) for $x \in [0, \alpha_1 \min\{\epsilon^{-1/2}, \delta^{-1}\}]$ and some small absolute constant $\alpha_1 \in (0, \frac{1}{8}]$. Thus Corollary 2.1 extends the asymptotic expansions of [15] to a larger range $x \in [0, \alpha_0\, \min\{(\epsilon\,|\log\epsilon|)^{-1}, \delta^{-1}\})$ and non bounded martingale differences.

2.2. Remarks on the main theorems

Combining the inequalities (8) and (10), we conclude that under (A1) and (A2) there is an absolute constant $\alpha_0 > 0$ such that, for all $0 \leq x \leq \alpha_0 \, \epsilon^{-1}$ and $\delta \leq \alpha_0$,

$$\left| \log \frac{\mathbb{P}(X_n > x)}{1 - \Phi\left(x\right)} \right| \leq c_{\alpha_0} \Big(x^3 \epsilon + x^2 \delta^2 + (1+x) \left(\epsilon \left| \log \epsilon \right| + \delta \right) \Big). \tag{14}$$

We show that this result can be regarded as a refinement of the moderate deviation principle (MDP) in the framework where (A1) and (A2) hold. Assume that (A1) and (A2) are satisfied with $\epsilon = \epsilon_n \to 0$ and $\delta = \delta_n \to 0$ as $n \to \infty$. Let a_n be any sequence of real numbers satisfying $a_n \to \infty$ and $a_n \epsilon_n \to 0$ as $n \to \infty$. Then inequality (14) implies the MDP for $(X_n)_{n \geq 1}$ with the speed a_n and rate function $x^2/2$. Indeed, using the inequalities

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

we deduce that, for any $x \geq 0$,

$$\lim_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P}(X_n > a_n x) = -\frac{x^2}{2}.$$

By a similar argument, we also have, for any $x \geq 0$,

$$\lim_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P}(X_n < -a_n x) = -\frac{x^2}{2}.$$

The last two equalities are equivalent to the statement that: for each Borel set B,

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

where B^o and $\overline{B}$ denote the interior and the closure of B respectively, see Lemma 4.4 of [20]. Similar results can be found in Gao [12] for the martingale differences satisfying the conditional Cramér condition $||\mathbb{E}(\exp\{c_0|\xi_i|\}|\mathcal{F}_{i-1})||_\infty < \infty$.

To show that our results are sharp, assume that $\xi_i = \eta_i/\sqrt{n}$, where $(\eta_i, \mathcal{F}_i)_{i=1,...,n}$ is a sequence of martingale differences satisfying the following conditions:

(A1′) (*Bernstein's condition*) There exists a positive absolute constant H such that

$$|\mathbb{E}(\eta_i^k|\mathcal{F}_{i-1})| \leq \frac{1}{2} k! H^{k-2} \mathbb{E}(\eta_i^2|\mathcal{F}_{i-1}), \quad \text{for } k \geq 3 \ \text{ and } \ 1 \leq i \leq n;$$

(A2′) There exists an absolute constant $N \geq 0$ such that $|\sum_{i=1}^n \mathbb{E}(\eta_i^2|\mathcal{F}_{i-1}) - n| \leq N^2$.

These conditions are satisfied with some $H > 0$ and $N = 0$ if, for instance, $\eta_1, \eta_2, ..., \eta_n$ are i.i.d. random variables with finite exponential moments (see Section 8 for an explicit expression of the positive absolute constant H).

Corollary 2.2. *Assume conditions (A1′) and (A2′). Then there is an absolute constant* $\alpha_2 > 0$ *such that for any absolute constant* $\alpha_1 > 0$ *and all* $\alpha_1\sqrt{\log n} \leq x \leq \alpha_2 n^{1/2}$, *we have*

$$\log \frac{\mathbb{P}(\sum_{i=1}^n \eta_i > x\sqrt{n})}{1 - \Phi(x)} = O\left((H+N)\frac{x^3}{\sqrt{n}}\right) \tag{15}$$

and

$$\log \frac{\mathbb{P}(\sum_{i=1}^n \eta_i < -x\sqrt{n})}{\Phi(-x)} = O\left((H+N)\frac{x^3}{\sqrt{n}}\right) \tag{16}$$

as $n \to \infty$.

It is worth noting that the remainders of the expansions (15) and (16) are of the same order as in (1) and therefore are optimal.

Corollary 2.3. *Assume conditions (A1′) and (A2′). Then, for all* $0 \leq x = O\left(\sqrt{\log n}\right)$,

$$\frac{\mathbb{P}(\sum_{i=1}^n \eta_i > x\sqrt{n})}{1 - \Phi(x)} = 1 + O\left((H+N)(1+x)\frac{\log n}{\sqrt{n}}\right) \tag{17}$$

and

$$\frac{\mathbb{P}(\sum_{i=1}^n \eta_i < -x\sqrt{n})}{\Phi(-x)} = 1 + O\left((H+N)(1+x)\frac{\log n}{\sqrt{n}}\right) \tag{18}$$

as $n \to \infty$.

Notice that (17) extends expansion (3) proved in Grama and Haeusler [15] to the case of martingale differences satisfying the conditional Bernstein condition (A1′). The Remark 2.1 of [15] and the sharp rate of convergence in the CLT due to Bolthausen [2] hint that the remainders of the expansions (17) and (18) are sharp.

Corollary 2.4. *Assume conditions (A1′) and (A2′). Then, for any absolute constant* $\alpha > 0$ *and* $\alpha\sqrt{\log n} \leq x = o\left(n^{1/6}\right)$,

$$\frac{\mathbb{P}(\sum_{i=1}^n \eta_i > x\sqrt{n})}{1 - \Phi(x)} = 1 + O\Big((H+N)\frac{x^3}{\sqrt{n}}\Big) \tag{19}$$

and

$$\frac{\mathbb{P}(\sum_{i=1}^n \eta_i < -x\sqrt{n})}{\Phi(-x)} = 1 + O\Big((H+N)\frac{x^3}{\sqrt{n}}\Big) \tag{20}$$

as $n \to \infty$.

The remainders of the expansions (19) and (20) are of the same order as in (1) in the stated range and therefore cannot be improved.

Remark 2.1. *The results formulated above are proved under Bernstein's condition (A1′). But they are also valid under some equivalent conditions which are stated in Section 8.*

3. Rates of convergence in the CLT

Let $(\xi_i, \mathcal{F}_i)_{i=0,...,n}$ be a sequence of martingale differences satisfying condition (A1) and $X = (X_k, \mathcal{F}_k)_{k=0,...,n}$ be the corresponding martingale defined by (6). For any real λ satisfying $|\lambda| < \epsilon^{-1}$, consider the *exponential multiplicative martingale* $Z(\lambda) = (Z_k(\lambda), \mathcal{F}_k)_{k=0,...,n}$, where

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

For each $k = 1, ..., n$, the random variable $Z_k(\lambda)$ defines a probability density on $(\Omega, \mathcal{F}, \mathbb{P})$. This allows us to introduce, for $|\lambda| < \epsilon^{-1}$, the *conjugate probability measure* $\mathbb{P}_\lambda$ on $(\Omega, \mathcal{F})$ defined by

$$d\mathbb{P}_\lambda = Z_n(\lambda) d\mathbb{P}. \tag{21}$$

Denote by $\mathbb{E}_\lambda$ the expectation with respect to $\mathbb{P}_\lambda$. For all $i = 1, \ldots, n$, let

$$\eta_i(\lambda) = \xi_i - b_i(\lambda) \qquad \text{and} \qquad b_i(\lambda) = \mathbb{E}_\lambda(\xi_i|\mathcal{F}_{i-1}).$$

We thus obtain the well-known semimartingale decomposition:

$$X_k = Y_k(\lambda) + B_k(\lambda), \qquad k = 1, ..., n, \tag{22}$$

where $Y(\lambda) = (Y_k(\lambda), \mathcal{F}_k)_{k=1,...,n}$ is the *conjugate martingale* defined as

$$Y_k(\lambda) = \sum_{i=1}^{k} \eta_i(\lambda), \qquad k = 1, ..., n, \tag{23}$$

and $B(\lambda) = (B_k(\lambda), \mathcal{F}_k)_{k=1,...,n}$ is the *drift process* defined as

$$B_k(\lambda) = \sum_{i=1}^{k} b_i(\lambda), \qquad k = 1, ..., n.$$

In the proofs of Theorems 2.1 and 2.2, we make use of the following assertion, which gives us a rate of convergence in the central limit theorem for the conjugate martingale $Y(\lambda)$ under the probability measure $\mathbb{P}_\lambda$.

Lemma 3.1. *Assume conditions (A1) and (A2). Then, for all* $0 \leq \lambda < \epsilon^{-1}$,

$$\sup_x |\mathbb{P}_\lambda(Y_n(\lambda) \leq x) - \Phi(x)| \leq c\, (\lambda\, \epsilon + \epsilon\, |\log \epsilon| + \delta)\, .$$

If $\lambda = 0$, then $Y_n(\lambda) = X_n$ and $\mathbb{P}_\lambda = \mathbb{P}$. So Lemma 3.1 implies the following theorem.

Theorem 3.1. *Assume conditions (A1) and (A2). Then*

$$\sup_x |\mathbb{P}(X_n \leq x) - \Phi\,(x)\,| \leq c\, (\epsilon\, |\log \epsilon| + \delta). \tag{24}$$

Remark 3.1. *By inspecting the proof of Lemma 3.1, we can see that Theorem 3.1 holds true when condition (A1) is replaced by the following weaker one:*

(C1) *There exists a number $\epsilon \in (0, \frac{1}{2}]$ depending on n such that*

$$|\mathbb{E}(\xi_i^k|\mathcal{F}_{i-1})| \leq \epsilon^{k-2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}), \quad \textit{for } k = 3, 5 \textit{ and } 1 \leq i \leq n.$$

Remark 3.2. *Bolthausen (see Theorem 2 of [2]) showed that if $|\xi_i| \leq \epsilon$ and condition (A2) holds, then*

$$\sup_x |\mathbb{P}(X_n \leq x) - \Phi(x)| \leq c_1 (\epsilon^3 n \log n + \delta). \tag{25}$$

We note that Theorem 3.1 implies Bolthausen's inequality (25) under the less restrictive condition (A1). Indeed, by condition (A2), we have $3/4 \leq \langle X \rangle_n \leq n\epsilon^2$ and then $\epsilon \geq \sqrt{3/(4n)}$. For $\epsilon \leq 1/2$, it is easy to see that $\epsilon^3 n \log n \geq 3\,\epsilon |\log \epsilon|/4$. Thus, inequality (24) implies (25) with $c_1 = 4c/3$.

4. Auxiliary results

In this section, we establish some auxiliary lemmas which will be used in the proofs of Theorems 2.1 and 2.2. We first prove upper bounds for the conditional moments.

Lemma 4.1. *Assume condition (A1). Then*

$$|\mathbb{E}(\xi_i^k|\mathcal{F}_{i-1})| \leq 6k!\epsilon^k, \quad \textit{for } k \geq 2,$$

and

$$\mathbb{E}(|\xi_i|^k|\mathcal{F}_{i-1}) \leq k!\epsilon^{k-2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}), \quad \textit{for } k \geq 2.$$

Proof. By Jensen's inequality and condition (A1),

$$\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1})^2 \leq \mathbb{E}(\xi_i^4|\mathcal{F}_{i-1}) \leq 12\epsilon^2\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}),$$

from which we get

$$\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) \leq 12\epsilon^2.$$

We obtain the first assertion. Again by condition (A1), for $k \geq 3$,

$$|\mathbb{E}(\xi_i^k|\mathcal{F}_{i-1})| \leq \frac{1}{2}k!\epsilon^{k-2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) \leq 6k!\epsilon^k.$$

If k is even, the second assertion holds obviously. If $k = 2l+1$, $l \geq 1$, is odd, by Hölder's inequality and condition (A1), it follows that

$$\begin{aligned}
\mathbb{E}\left(|\xi_i|^{2l+1}|\mathcal{F}_{i-1}\right) &\leq \mathbb{E}\left(|\xi_i|^l|\xi_i|^{l+1}|\mathcal{F}_{i-1}\right) \leq \sqrt{\mathbb{E}\left(\xi_i^{2l}|\mathcal{F}_{i-1}\right)\mathbb{E}\left(\xi_i^{2(l+1)}|\mathcal{F}_{i-1}\right)} \\
&\leq \frac{1}{2}\sqrt{(2l)!(2l+2)!}\,\epsilon^{2l-1}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) \\
&\leq (2l+1)!\epsilon^{2l-1}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}).
\end{aligned}$$

This completes the proof of Lemma 4.1. $\square$

The following lemma establishes a two sided bound for the drift process $B_n(\lambda)$.

Lemma 4.2. *Assume conditions (A1) and (A2). Then for any constant* $\alpha \in (0,1)$ *and all* $0 \le \lambda \le \alpha\,\epsilon^{-1}$,

$$|B_n(\lambda) - \lambda| \le \lambda\delta^2 + c_\alpha \lambda^2 \epsilon. \tag{26}$$

Proof. By the relation between $\mathbb{E}$ and $\mathbb{E}_\lambda$ on $\mathcal{F}_i$, we have

$$b_i(\lambda) = \frac{\mathbb{E}(\xi_i e^{\lambda\xi_i}|\mathcal{F}_{i-1})}{\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1})}, \qquad i = 1, ..., n.$$

Jensen's inequality and $\mathbb{E}(\xi_i|\mathcal{F}_{i-1}) = 0$ imply that $\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1}) \ge 1$. Since

$$\mathbb{E}(\xi_i e^{\lambda\xi_i}|\mathcal{F}_{i-1}) = \mathbb{E}\left(\xi_i(e^{\lambda\xi_i} - 1)|\mathcal{F}_{i-1}\right) \ge 0, \quad \text{for } \lambda \ge 0,$$

by Taylor's expansion for e^x, we find that

$$\begin{aligned} B_n(\lambda) &\le \sum_{i=1}^n \mathbb{E}(\xi_i e^{\lambda\xi_i}|\mathcal{F}_{i-1}) \\ &= \sum_{i=1}^n \mathbb{E}\left(\xi_i(e^{\lambda\xi_i} - 1)|\mathcal{F}_{i-1}\right) \\ &= \lambda\langle X\rangle_n + \sum_{i=1}^n \sum_{k=2}^{+\infty} \mathbb{E}\left(\frac{\xi_i(\lambda\xi_i)^k}{k!}\bigg|\mathcal{F}_{i-1}\right). \end{aligned} \tag{27}$$

Using condition (A1), we obtain, for any constant $\alpha \in (0,1)$ and all $0 \le \lambda \le \alpha\,\epsilon^{-1}$,

$$\begin{aligned} \sum_{i=1}^n \sum_{k=2}^{+\infty} \left(\frac{\xi_i(\lambda\xi_i)^k}{k!}\bigg|\mathcal{F}_{i-1}\right) &\le \sum_{i=1}^n \sum_{k=2}^{+\infty} |\mathbb{E}\left(\xi_i^{k+1}|\mathcal{F}_{i-1}\right)| \frac{\lambda^k}{k!} \\ &\le \frac{1}{2}\lambda^2\epsilon\langle X\rangle_n \sum_{k=2}^{+\infty}(k+1)(\lambda\epsilon)^{k-2} \\ &\le c_\alpha\,\lambda^2\epsilon\langle X\rangle_n. \end{aligned} \tag{28}$$

Using condition (A2), we get $\langle X\rangle_n \le 2$ and, for any constant $\alpha \in (0,1)$ and all $0 \le \lambda \le \alpha\,\epsilon^{-1}$,

$$\sum_{i=1}^n \sum_{k=2}^{+\infty} \left|\mathbb{E}\left(\frac{\xi_i(\lambda\xi_i)^k}{k!}\bigg|\mathcal{F}_{i-1}\right)\right| \le 2\,c_\alpha\lambda^2\epsilon. \tag{29}$$

Condition (A2) together with (27) and (29) imply the upper bound of $B_n(\lambda)$: for any constant $\alpha \in (0,1)$ and all $0 \le \lambda \le \alpha\,\epsilon^{-1}$,

$$B_n(\lambda) \le \lambda + \lambda\delta^2 + 2\,c_\alpha\,\lambda^2\epsilon.$$

Using Lemma 4.1, we have, for any constant $\alpha \in (0,1)$ and all $0 \leq \lambda \leq \alpha\, \epsilon^{-1}$,

$$\begin{aligned}
\mathbb{E}\left(e^{\lambda \xi_i} | \mathcal{F}_{i-1}\right) &\leq 1 + \sum_{k=2}^{+\infty} \left| \mathbb{E}\left(\frac{(\lambda \xi_i)^k}{k!} \middle| \mathcal{F}_{i-1}\right) \right| \\
&\leq 1 + 6 \sum_{k=2}^{+\infty} (\lambda \epsilon)^k \\
&\leq 1 + c_{1,\alpha} \, (\lambda \epsilon)^2. \qquad (30)
\end{aligned}$$

This inequality together with condition (A2) and (29) imply the lower bound of $B_n(\lambda)$: for any constant $\alpha \in (0,1)$ and all $0 \leq \lambda \leq \alpha\, \epsilon^{-1}$,

$$\begin{aligned}
B_n(\lambda) &\geq \left(\sum_{i=1}^n \mathbb{E}(\xi_i e^{\lambda \xi_i} | \mathcal{F}_{i-1}) \right) \Big(1 + c_{1,\alpha} \, (\lambda \epsilon)^2 \Big)^{-1} \\
&\geq \left(\lambda \langle X \rangle_n - \sum_{i=1}^n \sum_{k=2}^{+\infty} \left| \mathbb{E}\left(\frac{\xi_i (\lambda \xi_i)^k}{k!} \middle| \mathcal{F}_{i-1}\right) \right| \right) \Big(1 + c_{1,\alpha} \, (\lambda \epsilon)^2 \Big)^{-1} \\
&\geq \Big(\lambda - \lambda \delta^2 - 2\, c_\alpha \lambda^2 \epsilon \Big) \Big(1 + c_{1,\alpha} \, (\lambda \epsilon)^2 \Big)^{-1} \\
&\geq \lambda - \lambda \delta^2 - (2\, c_\alpha + \alpha\, c_{1,\alpha}) \, \lambda^2 \epsilon,
\end{aligned}$$

where the last line follows from the following inequality, for any constant $\alpha \in (0,1)$ and all $0 \leq \lambda \leq \alpha\, \epsilon^{-1}$,

$$\begin{aligned}
\lambda - \lambda \delta^2 - 2\, c_\alpha \lambda^2 \epsilon &\geq \lambda - \lambda \delta^2 - (2\, c_\alpha + \alpha\, c_{1,\alpha}) \lambda^2 \epsilon + c_{1,\alpha} \lambda^3 \epsilon^2 \\
&\geq \Big(\lambda - \lambda \delta^2 - (2\, c_\alpha + \alpha\, c_{1,\alpha}) \, \lambda^2 \epsilon \Big) \Big(1 + c_{1,\alpha} (\lambda \epsilon)^2 \Big).
\end{aligned}$$

The proof of Lemma 4.2 is finished. □

Now, consider the predictable cumulant process $\Psi(\lambda) = (\Psi_k(\lambda), \mathcal{F}_k)_{k=0,\ldots,n}$ related with the martingale X as follows:

$$\Psi_k(\lambda) = \sum_{i=1}^k \log \mathbb{E}\left(e^{\lambda \xi_i} | \mathcal{F}_{i-1}\right). \qquad (31)$$

We establish a two sided bound for the process $\Psi(\lambda)$.

Lemma 4.3. *Assume conditions (A1) and (A2). Then, for any constant $\alpha \in (0,1)$ and all $0 \leq \lambda \leq \alpha\, \epsilon^{-1}$,*

$$\left| \Psi_n(\lambda) - \frac{\lambda^2}{2} \right| \leq c_\alpha \lambda^3 \epsilon + \frac{\lambda^2 \delta^2}{2}.$$

Proof. Since $\mathbb{E}(\xi_i|\mathcal{F}_{i-1})=0$, it is easy to see that

$$\Psi_n(\lambda)=\sum_{i=1}^n\Big(\log\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1})-\lambda\mathbb{E}(\xi_i|\mathcal{F}_{i-1})-\frac{\lambda^2}{2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1})\Big)+\frac{\lambda^2}{2}\left\langle X\right\rangle_n.$$

Using a two-term Taylor's expansion of $\log(1+x), x\geq 0$, we obtain

$$\begin{aligned}\Psi_n(\lambda)-\frac{\lambda^2}{2}\left\langle X\right\rangle_n &= \sum_{i=1}^n\Big(\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1})-1-\lambda\mathbb{E}(\xi_i|\mathcal{F}_{i-1})-\frac{\lambda^2}{2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1})\Big)\\ &\quad+\frac{1}{2\Big(1+|\theta|\left(\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1})-1\right)\Big)}\sum_{i=1}^n\Big(\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1})-1\Big)^2.\end{aligned}$$

Since $\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1})\geq 1$, we find that

$$\begin{aligned}\left|\Psi_n(\lambda)-\frac{\lambda^2}{2}\left\langle X\right\rangle_n\right| &\leq \sum_{i=1}^n\left|\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1})-1-\lambda\mathbb{E}(\xi_i|\mathcal{F}_{i-1})-\frac{\lambda^2}{2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1})\right|\\ &\quad+\frac{1}{2}\sum_{i=1}^n\Big(\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1})-1\Big)^2\\ &\leq \sum_{i=1}^n\sum_{k=3}^{+\infty}\frac{\lambda^k}{k!}|\mathbb{E}(\xi_i^k|\mathcal{F}_{i-1})|+\frac{1}{2}\sum_{i=1}^n\left(\sum_{k=2}^{+\infty}\frac{\lambda^k}{k!}|\mathbb{E}(\xi_i^k|\mathcal{F}_{i-1})|\right)^2.\end{aligned}$$

In the same way as in the proof of (28), using condition (A1) and the inequality $\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1})\leq 12\,\epsilon^2$ (cf. Lemma 4.1), we have, for any constant $\alpha\in(0,1)$ and all $0\leq\lambda\leq\alpha\,\epsilon^{-1}$,

$$\left|\Psi_n(\lambda)-\frac{\lambda^2}{2}\left\langle X\right\rangle_n\right|\leq c_\alpha\lambda^3\epsilon\langle X\rangle_n.$$

Combining this inequality with condition (A2), we get, for any constant $\alpha\in(0,1)$ and all $0\leq\lambda\leq\alpha\,\epsilon^{-1}$,

$$\left|\Psi_n(\lambda)-\frac{\lambda^2}{2}\right|\leq 2\,c_\alpha\lambda^3\epsilon+\frac{\lambda^2\delta^2}{2},$$

which completes the proof of Lemma 4.3. □

5. Proof of Theorem 2.1

For $0\leq x<1$, the assertion follows from Theorem 3.1. It remains to prove Theorem 2.1 for any $\alpha\in(0,1)$ and all $1\leq x\leq\alpha\,\epsilon^{-1}$. Changing the probability measure according to (21), we have, for all $0\leq\lambda<\epsilon^{-1}$,

$$\begin{aligned}\mathbb{P}(X_n>x) &= \mathbb{E}_\lambda\left(Z_n(\lambda)^{-1}\mathbf{1}_{\{X_n>x\}}\right)\\ &= \mathbb{E}_\lambda\left(\exp\left\{-\lambda X_n+\Psi_n(\lambda)\right\}\mathbf{1}_{\{X_n>x\}}\right)\\ &= \mathbb{E}_\lambda\left(\exp\left\{-\lambda Y_n(\lambda)-\lambda B_n(\lambda)+\Psi_n(\lambda)\right\}\mathbf{1}_{\{Y_n(\lambda)+B_n(\lambda)>x\}}\right).\end{aligned}\qquad(32)$$

Let $\overline{\lambda} = \overline{\lambda}(x)$ be the largest solution of the equation

$$\lambda + \lambda \delta^2 + c_\alpha \lambda^2 \epsilon = x, \tag{33}$$

where c_α is given by inequality (26). The definition of $\overline{\lambda}$ implies that there exist $c_{\alpha,0}, c_{\alpha,1} > 0$ such that, for all $1 \leq x \leq \alpha\, \epsilon^{-1}$,

$$c_{\alpha,0}\, x \leq \overline{\lambda} = \frac{2x}{\sqrt{(1+\delta^2)^2 + 4c_\alpha x \epsilon} + 1 + \delta^2} \leq x \tag{34}$$

and

$$\overline{\lambda} = x - c_{\alpha,1} |\theta| (x^2 \epsilon + x\delta^2) \in [c_{\alpha,0}, \alpha\, \epsilon^{-1}\,]. \tag{35}$$

From (32), using Lemmas 4.2, 4.3 and equality (33), we obtain, for all $1 \leq x \leq \alpha\, \epsilon^{-1}$,

$$\mathbb{P}(X_n > x) \leq e^{c_{\alpha,2}\, (\overline{\lambda}^3 \epsilon + \overline{\lambda}^2 \delta^2) - \overline{\lambda}^2/2} \mathbb{E}_{\overline{\lambda}} \left(e^{-\overline{\lambda} Y_n(\overline{\lambda})} \mathbf{1}_{\{Y_n(\overline{\lambda})>0\}} \right). \tag{36}$$

It is easy to see that

$$\mathbb{E}_{\overline{\lambda}} \left(e^{-\overline{\lambda} Y_n(\overline{\lambda})} \mathbf{1}_{\{Y_n(\overline{\lambda})>0\}} \right) = \int_0^\infty \overline{\lambda} e^{-\overline{\lambda} y} \mathbb{P}_{\overline{\lambda}}(0 < Y_n(\overline{\lambda}) \leq y) dy. \tag{37}$$

Similarly, for a standard gaussian random variable $\mathcal{N}$, we have

$$\mathbb{E} \left(e^{-\overline{\lambda} \mathcal{N}} \mathbf{1}_{\{\mathcal{N}>0\}} \right) = \int_0^\infty \overline{\lambda} e^{-\overline{\lambda} y} \mathbb{P}(0 < \mathcal{N} \leq y) dy. \tag{38}$$

From (37) and (38), it follows

$$\left| \mathbb{E}_{\overline{\lambda}} \left(e^{-\overline{\lambda} Y_n(\overline{\lambda})} \mathbf{1}_{\{Y_n(\overline{\lambda})>0\}} \right) - \mathbb{E} \left(e^{-\overline{\lambda} \mathcal{N}} \mathbf{1}_{\{\mathcal{N}>0\}} \right) \right| \leq 2 \sup_y \left| \mathbb{P}_{\overline{\lambda}}(Y_n(\overline{\lambda}) \leq y) - \Phi(y) \right|.$$

Using Lemma 3.1, we obtain the following bound: for all $1 \leq x \leq \alpha\, \epsilon^{-1}$,

$$\left| \mathbb{E}_{\overline{\lambda}} \left(e^{-\overline{\lambda} Y_n(\overline{\lambda})} \mathbf{1}_{\{Y_n(\overline{\lambda})>0\}} \right) - \mathbb{E} \left(e^{-\overline{\lambda} \mathcal{N}} \mathbf{1}_{\{\mathcal{N}>0\}} \right) \right| \leq c \left(\overline{\lambda} \epsilon + \epsilon \left| \log \epsilon \right| + \delta \right). \tag{39}$$

From (36) and (39) we find that, for all $1 \leq x \leq \alpha\, \epsilon^{-1}$,

$$\mathbb{P}(X_n > x) \leq e^{c_{\alpha,2}\, (\overline{\lambda}^3 \epsilon + \overline{\lambda}^2 \delta^2) - \overline{\lambda}^2/2} \Bigg(\mathbb{E}\Big(e^{-\overline{\lambda} \mathcal{N}} \mathbf{1}_{\{\mathcal{N}>0\}} \Big) + c \left(\overline{\lambda} \epsilon + \epsilon \left| \log \epsilon \right| + \delta \right) \Bigg).$$

Since

$$e^{-\lambda^2/2} \mathbb{E} \left(e^{-\lambda \mathcal{N}} \mathbf{1}_{\{\mathcal{N}>0\}} \right) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(y+\lambda)^2/2} dy = 1 - \Phi\left(\lambda\right) \tag{40}$$

and, for all $\lambda \geq c_{\alpha,0}$,

$$1 - \Phi\left(\lambda\right) \geq \frac{1}{\sqrt{2\pi}(1+\lambda)}\ e^{-\lambda^2/2} \geq \frac{c_{\alpha,0}}{\sqrt{2\pi}(1+c_{\alpha,0})}\frac{1}{\lambda}e^{-\lambda^2/2} \tag{41}$$

(see Feller [11]), we obtain the following upper bound on tail probabilities: for all $1 \leq x \leq \alpha\,\epsilon^{-1}$,

$$\frac{\mathbb{P}(X_n > x)}{1-\Phi\left(\overline{\lambda}\right)} \quad \leq \quad e^{c_{\alpha,2}\,(\overline{\lambda}^3\epsilon + \overline{\lambda}^2\delta^2)}\left(1 + c_{\alpha,3}\,(\,\overline{\lambda}^2\epsilon + \overline{\lambda}\epsilon\,|\log\epsilon| + \overline{\lambda}\delta\,)\right). \tag{42}$$

Next, we would like to compare $1-\Phi(\overline{\lambda})$ with $1-\Phi(x)$. By (34), (35) and (41), we get

$$\begin{aligned} 1 \leq \frac{\int_{\overline{\lambda}}^{\infty}\exp\{-t^2/2\}dt}{\int_{x}^{\infty}\exp\{-t^2/2\}dt} \quad &= \quad 1 + \frac{\int_{\overline{\lambda}}^{x}\exp\{-t^2/2\}dt}{\int_{x}^{\infty}\exp\{-t^2/2\}dt} \\ &\leq \quad 1 + c_{\alpha,4}x(x-\overline{\lambda})\exp\{(x^2-\overline{\lambda}^2)/2\} \\ &\leq \quad \exp\{c_{\alpha,5}\,(x^3\epsilon + x^2\delta^2)\}. \end{aligned} \tag{43}$$

So, we find that

$$1-\Phi\left(\overline{\lambda}\right) = \left(1-\Phi(x)\right)\exp\left\{|\theta_1|c_{\alpha,5}\,(x^3\epsilon + x^2\delta^2)\right\}. \tag{44}$$

Implementing (44) in (42) and using (34), we obtain, for all $1 \leq x \leq \alpha\,\epsilon^{-1}$,

$$\begin{aligned} \frac{\mathbb{P}(X_n > x)}{1-\Phi\left(x\right)} \quad &\leq \quad \exp\{c_{\alpha,6}(x^3\epsilon + x^2\delta^2)\}\left(1 + c_{\alpha,7}\left(x^2\epsilon + x\epsilon\,|\log\epsilon| + x\delta\right)\right) \\ &\leq \quad \exp\{c_{\alpha,6}(x^3\epsilon + x^2\delta^2)\}\left(1 + c_{\alpha,7}\,x^2\epsilon\right)\left(1 + c_{\alpha,7}\,x\,(\epsilon\,|\log\epsilon| + \delta)\right) \\ &\leq \quad \exp\{c_{\alpha,8}(x^3\epsilon + x^2\delta^2)\}\left(1 + c_{\alpha,7}\,x\,(\epsilon\,|\log\epsilon| + \delta)\right). \end{aligned}$$

Taking $c_\alpha = \max\{c_{\alpha,7}, c_{\alpha,8}\}$, we prove the first assertion of Theorem 2.1. The second assertion follows from the first one applied to the martingale $(-X_k)_{k=0,\ldots,n}$.

6. Proof of Theorem 2.2

For $0 \leq x < 1$, the assertion follows from Theorem 3.1. It remains to prove Theorem 2.2 for $1 \leq x \leq \alpha_0\epsilon^{-1}$, where $\alpha_0 > 0$ is an absolute constant. Let $\underline{\lambda} = \underline{\lambda}(x)$ be the smallest solution of the equation

$$\lambda - \lambda\delta^2 - c_{1/2}\lambda^2\epsilon = x, \tag{45}$$

where c_α is given by inequality (26). The definition of $\underline{\lambda}$ implies that, for all $1 \leq x \leq 0.01c_{1/2}^{-1}\epsilon^{-1}$, it holds

$$x \leq \underline{\lambda} = \frac{2x}{1-\delta^2+\sqrt{(1-\delta^2)^2 - 4c_{1/2}x\epsilon}} \leq 2\,x \tag{46}$$

and

$$\underline{\lambda} = x + c_0|\theta|(x^2\epsilon + x\delta^2) \in [1, 0.02\, c_{1/2}^{-1}\epsilon^{-1}]. \tag{47}$$

From (32), using Lemmas 4.2, 4.3 and equality (45), we obtain, for all $1 \leq x \leq 0.01 c_{1/2}^{-1}\epsilon^{-1}$,

$$\mathbb{P}(X_n > x) \geq e^{-c_1\,(\underline{\lambda}^3\epsilon + \underline{\lambda}^2\delta^2) - \underline{\lambda}^2/2}\mathbb{E}_{\underline{\lambda}}\left(e^{-\underline{\lambda}Y_n(\underline{\lambda})}\mathbf{1}_{\{Y_n(\underline{\lambda})>0\}}\right). \tag{48}$$

In the subsequent we distinguish two cases. First, let $1 \leq \underline{\lambda} \leq \alpha_1 \min\{\epsilon^{-1/2}, \delta^{-1}\}$, where $\alpha_1 > 0$ is a small absolute constant whose value will be given later. Note that inequality (39) can be established with $\overline{\lambda}$ replaced by $\underline{\lambda}$, which, in turn, implies

$$\mathbb{P}(X_n > x) \geq e^{-c_1\,(\underline{\lambda}^3\epsilon + \underline{\lambda}^2\delta^2) - \underline{\lambda}^2/2}\Big(\mathbb{E}\left(e^{-\underline{\lambda}\mathcal{N}}\mathbf{1}_{\{\mathcal{N}>0\}}\right) - c_2\left(\underline{\lambda}\epsilon + \epsilon\,|\log\epsilon| + \delta\right)\Big).$$

By (40) and (41), we obtain the following lower bound on tail probabilities:

$$\frac{\mathbb{P}(X_n > x)}{1 - \Phi\left(\underline{\lambda}\right)} \;\geq\; e^{-c_1\,(\underline{\lambda}^3\epsilon + \underline{\lambda}^2\delta^2)}\left(1 - c_2\left(\underline{\lambda}^2\epsilon + \underline{\lambda}\epsilon\,|\log\epsilon| + \underline{\lambda}\delta\right)\right). \tag{49}$$

Taking $\alpha_1 = (8c_2)^{-1}$, we deduce that, for all $1 \leq \underline{\lambda} \leq \alpha_1 \min\{\epsilon^{-1/2}, \delta^{-1}\}$,

$$1 - c_2\left(\underline{\lambda}^2\epsilon + \underline{\lambda}\epsilon\,|\log\epsilon| + \underline{\lambda}\delta\right) \;\geq\; \exp\left\{-2c_2\left(\underline{\lambda}^2\epsilon + \underline{\lambda}\epsilon\,|\log\epsilon| + \underline{\lambda}\delta\right)\right\}. \tag{50}$$

Implementing (50) in (49), we obtain

$$\frac{\mathbb{P}(X_n > x)}{1 - \Phi\left(\underline{\lambda}\right)} \;\geq\; \exp\left\{ - c_3\left(\underline{\lambda}^3\epsilon + \underline{\lambda}\epsilon\,|\log\epsilon| + \underline{\lambda}\delta + \underline{\lambda}^2\delta^2\right)\right\} \tag{51}$$

which is valid for all $1 \leq \underline{\lambda} \leq \alpha_1 \min\{\epsilon^{-1/2}, \delta^{-1}\}$.

Next, we consider the case of $\alpha_1 \min\{\epsilon^{-1/2}, \delta^{-1}\} \leq \underline{\lambda} \leq \alpha_0\epsilon^{-1}$ and $\delta \leq \alpha_0$. Let $K \geq 1$ be an absolute constant, whose exact value will be chosen later. It is easy to see that

$$\begin{aligned}\mathbb{E}_{\underline{\lambda}}\left(e^{-\underline{\lambda}Y_n(\underline{\lambda})}\mathbf{1}_{\{Y_n(\underline{\lambda})>0\}}\right) \;&\geq\; \mathbb{E}_{\underline{\lambda}}\Big(e^{-\underline{\lambda}Y_n(\underline{\lambda})}\mathbf{1}_{\{0<Y_n(\underline{\lambda})\leq K\gamma\}}\Big)\\ &\geq\; e^{-\underline{\lambda}K\gamma}\mathbb{P}_{\underline{\lambda}}\Big(0 < Y_n(\underline{\lambda}) \leq K\gamma\Big),\end{aligned} \tag{52}$$

where $\gamma = \underline{\lambda}\epsilon + \epsilon|\log\epsilon| + \delta \leq 4\alpha_0^{1/2}$, if $\alpha_0 \leq 1$. From Lemma 3.1, we have

$$\begin{aligned}\mathbb{P}_{\underline{\lambda}}\Big(0 < Y_n(\underline{\lambda}) \leq K\gamma\Big) \;&\geq\; \mathbb{P}\Big(0 < \mathcal{N} \leq K\gamma\Big) - c_5\gamma\\ &\geq\; K\gamma e^{-K^2\gamma^2/2} - c_5\gamma\\ &\geq\; \Big(Ke^{-8K^2\alpha_0} - c_5\Big)\,\gamma.\end{aligned}$$

Taking $\alpha_0 = 1/(16K^2)$, we find that

$$\mathbb{P}_{\underline{\lambda}}\Big(0 < Y_n(\underline{\lambda}) \leq K\gamma\Big) \geq \left(\frac{1}{2}K - c_5\right)\gamma.$$

Letting $K \geq 8c_5$, it follows that

$$\mathbb{P}_{\underline{\lambda}}\Big(0 < Y_n(\underline{\lambda}) \leq K\gamma\Big) \geq \frac{3}{8}K\gamma \geq \frac{3}{8}K\frac{\max\{\underline{\lambda}^2\epsilon, \underline{\lambda}\delta\}}{\underline{\lambda}}.$$

Choosing $K = \max\Big\{8c_5, \frac{8\alpha_1^{-2}}{3\sqrt{\pi}}\Big\}$ and taking into account that $\alpha_1 \min\{\epsilon^{-1/2}, \delta^{-1}\} \leq \underline{\lambda} \leq \alpha_0\epsilon^{-1}$, we deduce that

$$\mathbb{P}_{\underline{\lambda}}\Big(0 < Y_n(\underline{\lambda}) \leq K\gamma\Big) \geq \frac{1}{\sqrt{\pi}\underline{\lambda}}.$$

Since the inequality $\frac{1}{\sqrt{\pi}\lambda}e^{-\lambda^2/2} \geq 1 - \Phi(\lambda)$ is valid for $\lambda \geq 1$ (see Feller [11]), it follows that, for all $\alpha_1 \min\{\epsilon^{-1/2}, \delta^{-1}\} \leq \underline{\lambda} \leq \alpha_0\epsilon^{-1}$,

$$\mathbb{P}_{\underline{\lambda}}\Big(0 < Y_n(\underline{\lambda}) \leq K\gamma\Big) \geq \Big(1 - \Phi(\underline{\lambda})\Big)e^{\underline{\lambda}^2/2}. \tag{53}$$

From (48), (52) and (53), we obtain

$$\frac{\mathbb{P}(X_n > x)}{1 - \Phi(\underline{\lambda})} \;\geq\; \exp\Big\{-c_{\alpha_0,6}\left(\underline{\lambda}^3\epsilon + \underline{\lambda}\epsilon|\log\epsilon| + \underline{\lambda}\delta + \underline{\lambda}^2\delta^2\right)\Big\} \tag{54}$$

which is valid for all $\alpha_1 \min\{\epsilon^{-1/2}, \delta^{-1}\} \leq \underline{\lambda} \leq \alpha_0\epsilon^{-1}$.

Putting (51) and (54) together, we obtain, for all $1 \leq \underline{\lambda} \leq \alpha_0\epsilon^{-1}$ and $\delta \leq \alpha_0$,

$$\frac{\mathbb{P}(X_n > x)}{1 - \Phi(\underline{\lambda})} \;\geq\; \exp\Big\{-c_{\alpha_0,7}\left(\underline{\lambda}^3\epsilon + \underline{\lambda}\epsilon\,|\log\epsilon| + \underline{\lambda}\delta + \underline{\lambda}^2\delta^2\right)\Big\}. \tag{55}$$

As in the proof of Theorem 2.1, we now compare $1-\Phi(\underline{\lambda})$ with $1-\Phi(x)$. By a similar argument as in (43), we have

$$1 - \Phi(\underline{\lambda}) = \Big(1 - \Phi(x)\Big)\exp\left\{-|\theta|c_3\left(x^3\epsilon + x^2\delta^2\right)\right\}. \tag{56}$$

Combining (46), (55) and (56), we obtain, for all $1 \leq x \leq \alpha_0\epsilon^{-1}$ and $\delta \leq \alpha_0$,

$$\frac{\mathbb{P}(X_n > x)}{1 - \Phi(x)} \geq \exp\Big\{-c_{\alpha_0,8}\left(x^3\epsilon + x\epsilon\,|\log\epsilon| + x\delta + x^2\delta^2\right)\Big\} \tag{57}$$

which gives the first conclusion of Theorem 2.2. The second conclusion follows from the first one applied to the martingale $(-X_k)_{k=0,\dots,n}$. □

7. Proof of Lemma 3.1

The proof of Lemma 3.1 is a refinement of Lemma 3.3 of Grama and Haeusler [15] where it is assumed that $|\eta_i| \leq 2\epsilon$, which is a particular case of condition (A1). Compared to the case where η_i are bounded, the main challenge of our proof comes from the control of I_1 defined in (64) below.

In this section, α denotes a positive absolute number satisfying $\alpha \in (0,1)$, ϑ denotes a real number satisfying $0 \leq \vartheta \leq 1$, which is different from θ, and $\varphi(t)$ denotes the density function of the standard normal distribution. For the sake of simplicity, we also denote $Y(\lambda), Y_n(\lambda)$ and $\eta(\lambda)$ by Y, Y_n and η, respectively. We want to obtain a rate of convergence in the central limit theorem for the conjugate martingale $Y = (Y_k, \mathcal{F}_k)_{k=1,...,n}$, where $Y_k = \sum_{i=1}^k \eta_i$. Denote the quadratic characteristic of the conjugate martingale Y by $\langle Y \rangle_k = \sum_{i\leq k} \mathbb{E}_\lambda(\eta_i^2|\mathcal{F}_{i-1})$, and set $\Delta \langle Y \rangle_k = \mathbb{E}_\lambda(\eta_k^2|\mathcal{F}_{k-1})$. It is easy to see that, for $k = 1, ..., n$,

$$\begin{aligned}\Delta \langle Y \rangle_k &= \mathbb{E}_\lambda\left((\xi_k - b_k(\lambda))^2|\mathcal{F}_{k-1}\right) \\ &= \frac{\mathbb{E}(\xi_k^2 e^{\lambda\xi_k}|\mathcal{F}_{k-1})}{\mathbb{E}(e^{\lambda\xi_k}|\mathcal{F}_{k-1})} - \frac{\mathbb{E}(\xi_k e^{\lambda\xi_k}|\mathcal{F}_{k-1})^2}{\mathbb{E}(e^{\lambda\xi_k}|\mathcal{F}_{k-1})^2}. \qquad (58)\end{aligned}$$

Since $\mathbb{E}(e^{\lambda\xi_i}|\mathcal{F}_{i-1}) \geq 1$ and $|\eta_i|^k \leq 2^{k-1}(|\xi_i|^k + \mathbb{E}_\lambda(|\xi_i||\mathcal{F}_{i-1})^k)$, using condition (A1) and Lemma 4.1, we obtain, for all $k \geq 3$ and all $0 \leq \lambda \leq \frac{1}{4}\epsilon^{-1}$,

$$\begin{aligned}\mathbb{E}_\lambda\left(|\eta_i|^k|\mathcal{F}_{i-1}\right) &\leq 2^{k-1}\mathbb{E}_\lambda\left(|\xi_i|^k + \mathbb{E}_\lambda(|\xi_i||\mathcal{F}_{i-1})^k|\mathcal{F}_{i-1}\right) \\ &\leq 2^k\mathbb{E}_\lambda\left(|\xi_i|^k|\mathcal{F}_{i-1}\right) \\ &\leq 2^k\mathbb{E}\left(|\xi_i|^k\exp\{|\lambda\xi_i|\}|\mathcal{F}_{i-1}\right) \\ &\leq c2^k k!\epsilon^{k-2}\mathbb{E}\left(\xi_i^2|\mathcal{F}_{i-1}\right).\end{aligned}$$

Using Taylor's expansion for e^x and Lemma 1, we have, for all $0 \leq \lambda \leq \frac{1}{4}\epsilon^{-1}$,

$$\begin{aligned}|\Delta \langle Y \rangle_k - \Delta \langle X \rangle_k| &\leq \left|\frac{\mathbb{E}(\xi_k^2 e^{\lambda\xi_k}|\mathcal{F}_{k-1})}{\mathbb{E}(e^{\lambda\xi_k}|\mathcal{F}_{k-1})} - \mathbb{E}(\xi_k^2|\mathcal{F}_{k-1})\right| + \left|\frac{\mathbb{E}(\xi_k e^{\lambda\xi_k}|\mathcal{F}_{k-1})^2}{\mathbb{E}(e^{\lambda\xi_k}|\mathcal{F}_{k-1})^2}\right| \\ &\leq \left|\mathbb{E}(\xi_k^2 e^{\lambda\xi_k}|\mathcal{F}_{k-1}) - \mathbb{E}(\xi_k^2|\mathcal{F}_{k-1})\mathbb{E}(e^{\lambda\xi_k}|\mathcal{F}_{k-1})\right| \\ &\quad + \mathbb{E}(\xi_k e^{\lambda\xi_k}|\mathcal{F}_{k-1})^2 \\ &\leq \sum_{l=1}^{\infty}|\mathbb{E}(\xi_k^{l+2}|\mathcal{F}_{k-1})|\frac{\lambda^l}{l!} + \Delta\langle X\rangle_k \sum_{l=1}^{\infty}|\mathbb{E}(\xi_k^{l}|\mathcal{F}_{k-1})|\frac{\lambda^l}{l!} \\ &\quad + \left(\sum_{l=1}^{\infty}|\mathbb{E}(\xi_k^{l+1}|\mathcal{F}_{k-1})|\frac{\lambda^l}{l!}\right)^2 \\ &\leq c\lambda\epsilon\,\Delta\langle X\rangle_k. \qquad (59)\end{aligned}$$

Therefore,

$$|\langle Y\rangle_n - 1| \leq |\langle Y\rangle_n - \langle X\rangle_n| + |\langle X\rangle_n - 1| \leq c\lambda\epsilon\langle X\rangle_n + \delta^2.$$

Thus the martingale Y satisfies the following conditions (analogous to conditions (A1) and (A2)): for all $0 \leq \lambda \leq \frac{1}{4}\epsilon^{-1}$,

(B1) $\mathbb{E}_\lambda(|\eta_i|^k|\mathcal{F}_{i-1}) \leq c_k\epsilon^{k-2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}), \quad 5 \geq k \geq 3;$

(B2) $|\langle Y\rangle_n - 1| \leq c(\lambda\epsilon + \delta^2).$

We first prove Lemma 3.1 for $1 \leq \lambda < \epsilon^{-1}$. Without loss of generality, we can assume that $1 \leq \lambda \leq \frac{1}{4}\epsilon^{-1}$, otherwise we take $c \geq 4$ in the assertion of the lemma. Set $T = 1 + \delta^2$ and introduce a modification of the quadratic characteristic $\langle X\rangle$ as follows:

$$V_k = \langle X\rangle_k \mathbf{1}_{\{k<n\}} + T\mathbf{1}_{\{k=n\}}. \tag{60}$$

Note that $V_0 = 0$, $V_n = T$ and that $(V_k, \mathcal{F}_k)_{k=0,...,n}$ is a predictable process. Set $\gamma = \lambda\epsilon+\delta$, where $\lambda \in [1, \epsilon^{-1})$. Let $c_* \geq 4$ be a "free" absolute constant, whose exact value will be chosen later. Consider the non-increasing discrete time predictable process $A_k = c_*^2\gamma^2 + T - V_k$, $k = 1, ..., n$. For any fixed $u \in \mathbb{R}$ and any $x \in \mathbb{R}$ and $y > 0$, set for brevity,

$$\Phi_u(x, y) = \Phi\left((u - x)/\sqrt{y}\right). \tag{61}$$

In the proof we make use of the following two assertions, which can be found in Bolthausen's paper [2].

Lemma 7.1. *[2] Let X and Y be random variables. Then*

$$\sup_u |\mathbb{P}(X \leq u) - \Phi(u)| \leq c_1 \sup_u |\mathbb{P}(X + Y \leq u) - \Phi(u)| + c_2 \left\|\mathbb{E}\left(Y^2|X\right)\right\|_\infty^{1/2}.$$

Lemma 7.2. *[2] Let $G(x)$ be an integrable function of bounded variation, X be a random variable and a, $b > 0$ are real numbers. Then*

$$\mathbb{E}\, G\left(\frac{X + a}{b}\right) \leq c_1 \sup_u |\mathbb{P}(X \leq u) - \Phi(u)| + c_2\, b.$$

Let $\mathcal{N}_{c_*^2\gamma^2} = \mathcal{N}(0, c_*\gamma)$ be a normal random variable independent of Y_n. Using a well-known smoothing procedure (which employs Lemma 7.1), we get

$$\begin{aligned}
\sup_u |\mathbb{P}_\lambda(Y_n \leq u) - \Phi(u)| &\leq c_1 \sup_u |\mathbb{E}_\lambda\Phi_u(Y_n, A_n) - \Phi(u)| + c_2\gamma \\
&\leq c_1 \sup_u |\mathbb{E}_\lambda\Phi_u(Y_n, A_n) - \mathbb{E}_\lambda\Phi_u(Y_0, A_0)| \\
&\quad + c_1 \sup_u |\mathbb{E}_\lambda\Phi_u(Y_0, A_0) - \Phi(u)| + c_2\gamma \\
&= c_1 \sup_u |\mathbb{E}_\lambda\Phi_u(Y_n, A_n) - \mathbb{E}_\lambda\Phi_u(Y_0, A_0)| \\
&\quad + c_1 \sup_u \left|\Phi\left(\frac{u}{\sqrt{c_*^2\gamma^2 + T}}\right) - \Phi(u)\right| + c_2\gamma \\
&\leq c_1 \sup_u |\mathbb{E}_\lambda\Phi_u(Y_n, A_n) - \mathbb{E}_\lambda\Phi_u(Y_0, A_0)| + c_3\gamma, \qquad (62)
\end{aligned}$$

where

$$\mathbb{E}_\lambda \Phi_u(Y_n, A_n) = \mathbb{P}_\lambda(Y_n + \mathcal{N}_{c_*^2\gamma^2} \le u) \ \text{ and } \ \mathbb{E}_\lambda \Phi_u(Y_0, A_0) = \mathbb{P}_\lambda(\mathcal{N}_{c_*^2\gamma^2+T} \le u).$$

By simple telescoping, we find that

$$\mathbb{E}_\lambda \Phi_u(Y_n, A_n) - \mathbb{E}_\lambda \Phi_u(Y_0, A_0) = \mathbb{E}_\lambda \sum_{k=1}^{n} \Big(\Phi_u(Y_k, A_k) - \Phi_u(Y_{k-1}, A_{k-1}) \Big).$$

From this, taking into account that $(\eta_i, \mathcal{F}_i)_{i=0,\dots,n}$ is a $\mathbb{P}_\lambda$-martingale and that

$$\frac{\partial^2}{\partial x^2} \Phi_u(x, y) = 2 \frac{\partial}{\partial y} \Phi_u(x, y),$$

we obtain

$$\mathbb{E}_\lambda \Phi_u(Y_n, A_n) - \mathbb{E}_\lambda \Phi_u(Y_0, A_0) = I_1 + I_2 - I_3, \tag{63}$$

where

$$\begin{aligned} I_1 &= \mathbb{E}_\lambda \sum_{k=1}^{n} \bigg(\Phi_u(Y_k, A_k) - \Phi_u(Y_{k-1}, A_k) \\ &\qquad\qquad - \frac{\partial}{\partial x} \Phi_u(Y_{k-1}, A_k) \eta_k - \frac{1}{2} \frac{\partial^2}{\partial x^2} \Phi_u(Y_{k-1}, A_k) \eta_k^2 \bigg), & (64) \\ I_2 &= \frac{1}{2} \mathbb{E}_\lambda \sum_{k=1}^{n} \frac{\partial^2}{\partial x^2} \Phi_u(Y_{k-1}, A_k) \Big(\Delta \langle Y \rangle_k - \Delta V_k \Big), & (65) \end{aligned}$$

$$I_3 = \mathbb{E}_\lambda \sum_{k=1}^{n} \bigg(\Phi_u(Y_{k-1}, A_{k-1}) - \Phi_u(Y_{k-1}, A_k) - \frac{\partial}{\partial y} \Phi_u(Y_{k-1}, A_k) \Delta V_k \bigg). \tag{66}$$

We now give estimates of I_1, I_2 and I_3. To shorten notations, set

$$T_{k-1} = (u - Y_{k-1}) / \sqrt{A_k}.$$

a) *Control of* I_1. Using a three-term Taylor's expansion, we have

$$I_1 = -\mathbb{E}_\lambda \sum_{k=1}^{n} \frac{1}{6 A_k^{3/2}} \varphi'' \left(T_{k-1} - \frac{\vartheta_k \eta_k}{\sqrt{A_k}} \right) \eta_k^3 \,. \tag{67}$$

In order to bound $\varphi''(\cdot)$ we distinguish two cases as follows.

Case 1: $|\eta_k / \sqrt{A_k}| \le |T_{k-1}|/2$. In this case, by the inequality $\varphi''(t) \le \varphi(t)(1 + t^2)$, it follows

$$\begin{aligned} \left| \varphi'' \left(T_{k-1} - \frac{\vartheta_k \eta_k}{\sqrt{A_k}} \right) \right| &\le \varphi \left(T_{k-1} - \frac{\vartheta_k \eta_k}{\sqrt{A_k}} \right) \left(1 + \left(T_{k-1} - \frac{\vartheta_k \eta_k}{\sqrt{A_k}} \right)^2 \right) \\ &\le \sup_{|t - T_{k-1}| \le |T_{k-1}|/2} \varphi(t)(1 + t^2) \\ &\le \varphi(T_{k-1}/2)(1 + 4 T_{k-1}^2). \end{aligned}$$

Define $g_1(t) = \sup_{|t-z|\le 3} f_1(z)$, where $f_1(t) = \varphi(t/2)(1+4t^2)$. It is easy to see that $g_1(t)$ is a symmetric integrable function of bounded variation, non-increasing in $t \ge 0$. Therefore,

$$\left|\varphi''\left(T_{k-1} - \frac{\vartheta_k \eta_k}{\sqrt{A_k}}\right)\right| \mathbf{1}_{\left\{|\eta_k/\sqrt{A_k}|\le |T_{k-1}|/2\right\}} \le g_1(T_{k-1}). \tag{68}$$

Case 2: $|\eta_k/\sqrt{A_k}| > |T_{k-1}|/2$. Since $|\varphi''(t)| \le 2$, it follows that

$$\left|\varphi''\left(T_{k-1} - \frac{\vartheta_k \eta_k}{\sqrt{A_k}}\right)\right| \mathbf{1}_{\left\{|\eta_k/\sqrt{A_k}| > |T_{k-1}|/2\right\}} \le 2\left(\mathbf{1}_{\{|T_{k-1}|<2\}} + \frac{4\eta_k^2}{T_{k-1}^2 A_k}\mathbf{1}_{\{|T_{k-1}|\ge 2\}}\right). \tag{69}$$

Now we bound the conditional expectation of $|\eta_k|^k$. Using condition (B1), we have

$$\mathbb{E}_\lambda(|\eta_k|^3|\mathcal{F}_{k-1}) \le c\,\epsilon \Delta\langle X\rangle_k \quad \text{and} \quad \mathbb{E}_\lambda(|\eta_k|^5|\mathcal{F}_{k-1}) \le c\,\epsilon^3 \Delta\langle X\rangle_k,$$

where $\Delta\langle X\rangle_k = \langle X\rangle_k - \langle X\rangle_{k-1}$. From the definition of the process V (see (60)), it follows that $\Delta\left\langle X\right\rangle_k \le \Delta V_k = V_k - V_{k-1}$,

$$\mathbb{E}_\lambda(|\eta_k|^3|\mathcal{F}_{k-1}) \le c\,\Delta V_k\,\epsilon \quad \text{and} \quad \mathbb{E}_\lambda(|\eta_k|^5|\mathcal{F}_{k-1}) \le c\,\Delta V_k\,\epsilon^3. \tag{70}$$

Thus, from (68), we obtain

$$\mathbb{E}_\lambda\left(\left|\varphi''\left(T_{k-1} - \frac{\vartheta_k \eta_k}{\sqrt{A_k}}\right)\eta_k^3\right| \mathbf{1}_{\left\{|\eta_k/\sqrt{A_k}|\le |T_{k-1}|/2\right\}}\middle|\mathcal{F}_{k-1}\right) \le c_4\, g_1(T_{k-1})\Delta V_k\,\epsilon. \tag{71}$$

From (69), by (70) and the inequality $\frac{\epsilon^2}{A_k} \ge c_*^{-2}$, we find

$$\mathbb{E}_\lambda\left(\left|\varphi''\left(T_{k-1} - \frac{\vartheta_k \eta_k}{\sqrt{A_k}}\right)\eta_k^3\right| \mathbf{1}_{\left\{|\eta_k/\sqrt{A_k}| > |T_{k-1}|/2\right\}}\middle|\mathcal{F}_{k-1}\right) \le g_2(T_{k-1})\Delta V_k\,\epsilon, \tag{72}$$

where $g_2(t) = 2\,c(\mathbf{1}_{\{|t|<2\}} + 4\frac{1}{t^2}\mathbf{1}_{\{|t|\ge 2\}})$. Set $G(t) = c_4\, g_1(t) + g_2(t)$. Then $G(t)$ is a symmetric integrable function of bounded variation, non-increasing in $t \ge 0$. Returning to (67), by (71) and (72), we get

$$|I_1| \le \mathbb{E}_\lambda\left[\sum_{k=1}^n \frac{1}{6A_k^{3/2}}\mathbb{E}_\lambda\left(\left|\varphi''\left(T_{k-1} - \frac{\vartheta_k \eta_k}{\sqrt{A_k}}\right)\eta_k^3\right|\middle|\mathcal{F}_{k-1}\right)\right] \le J_1, \tag{73}$$

where

$$J_1 = c\,\epsilon\,\mathbb{E}_\lambda \sum_{k=1}^n \frac{1}{A_k^{3/2}} G\left(T_{k-1}\right)\Delta V_k\,. \tag{74}$$

Let us introduce the time change τ_t as follows: for any real $t \in [0, T]$,

$$\tau_t = \min\{k \le n : V_k > t\}, \quad \text{where} \quad \min\emptyset = n. \tag{75}$$

It is clear that, for any $t \in [0,T]$, the stopping time τ_t is predictable. Let $(\sigma_k)_{k=1,\dots,n+1}$ be the increasing sequence of moments when the increasing stepwise function τ_t, $t \in [0,T]$, has jumps. It is clear that $\Delta V_k = \int_{[\sigma_k,\sigma_{k+1})} dt$ and that $k = \tau_t$, for $t \in [\sigma_k, \sigma_{k+1})$. Since $\tau_T = n$, we have

$$\begin{aligned}
\sum_{k=1}^{n} \frac{1}{A_k^{3/2}} G\left(T_{k-1}\right) \Delta V_k &= \sum_{k=1}^{n} \int_{[\sigma_k,\sigma_{k+1})} \frac{1}{A_{\tau_t}^{3/2}} G\left(T_{\tau_t-1}\right) dt \\
&= \int_0^T \frac{1}{A_{\tau_t}^{3/2}} G\left(T_{\tau_t-1}\right) dt.
\end{aligned}$$

Set, for brevity, $a_t = c_*^2\gamma^2 + T - t$. Since $\Delta V_{\tau_t} \leq 12\gamma^2$, we see that

$$t \leq V_{\tau_t} \leq V_{\tau_t-1} + \Delta V_{\tau_t} \leq t + 12\gamma^2, \quad t \in [0,T]. \tag{76}$$

Taking into account that $c_* \geq 4$, we have

$$\frac{1}{4}\, a_t \leq A_{\tau_t} = c_*^2\gamma^2 + T - V_{\tau_t} \leq a_t, \quad t \in [0,T]. \tag{77}$$

Since $G(z)$ is symmetric and is non-increasing in $z \geq 0$, the last bound implies that

$$J_1 \leq c\,\epsilon \int_0^T \frac{1}{a_t^{3/2}} \mathbb{E}_\lambda G\left(\frac{u - Y_{\tau_t-1}}{a_t^{1/2}}\right) dt. \tag{78}$$

By Lemma 7.2, it is easy to see that

$$\mathbb{E}_\lambda G\left(\frac{u - Y_{\tau_t-1}}{{a_t}^{1/2}}\right) \leq c_1 \sup_z \left|\mathbb{P}_\lambda(Y_{\tau_t-1} \leq z) - \Phi(z)\right| + c_2\sqrt{a_t}. \tag{79}$$

Since $V_{\tau_t-1} = V_{\tau_t} - \Delta V_{\tau_t}$, $V_{\tau_t} \geq t$ (cf. (76)) and $\Delta V_{\tau_t} \leq 12\gamma^2$, we get

$$V_n - V_{\tau_t-1} \leq V_n - V_{\tau_t} + \Delta V_{\tau_t} \leq 12\gamma^2 + T - t \leq a_t. \tag{80}$$

Thus

$$\begin{aligned}
\mathbb{E}_\lambda\left((Y_n - Y_{\tau_t-1})^2 | \mathcal{F}_{\tau_t-1}\right) &= \mathbb{E}_\lambda\left(\sum_{k=\tau_t}^{n} \mathbb{E}_\lambda(\eta_k^2|\mathcal{F}_{k-1}) \middle| \mathcal{F}_{\tau_t-1}\right) \\
&\leq c\,\mathbb{E}_\lambda\left(\sum_{k=\tau_t}^{n} \Delta \left\langle X\right\rangle_k \middle| \mathcal{F}_{\tau_t-1}\right) \\
&= c\,\mathbb{E}_\lambda\left(\left\langle X\right\rangle_n - \left\langle X\right\rangle_{\tau_t-1} | \mathcal{F}_{\tau_t-1}\right) \\
&\leq c\,\mathbb{E}_\lambda\left(V_n - V_{\tau_t-1} | \mathcal{F}_{\tau_t-1}\right) \\
&\leq c\,a_t.
\end{aligned}$$

Then, by Lemma 7.1, we find that, for any $t \in [0, T]$,

$$\sup_z |\mathbb{P}_\lambda(Y_{\tau_t - 1} \le z) - \Phi(z)| \le c_1 \sup_z |\mathbb{P}_\lambda(Y_n \le z) - \Phi(z)| + c_2 \sqrt{a_t}. \tag{81}$$

From (78), (79) and (81), we obtain

$$J_1 \le c_1 \, \epsilon \int_0^T \frac{dt}{a_t^{3/2}} \sup_z |\mathbb{P}_\lambda(Y_n \le z) - \Phi(z)| + c_2 \, \epsilon \int_0^T \frac{dt}{a_t}. \tag{82}$$

By elementary computations, we see that (since $\lambda \ge 1$)

$$\int_0^T \frac{dt}{a_t^{3/2}} \le \frac{c}{c_* \lambda \epsilon} \le \frac{c}{c_* \epsilon} \qquad \text{and} \qquad \int_0^T \frac{dt}{a_t} \le c \, |\log \epsilon| \, . \tag{83}$$

Then

$$|I_1| \le J_1 \le \frac{c}{c_*} \sup_z |\mathbb{P}(Y_n \le z) - \Phi(z)| + c_2 \, \epsilon \, |\log \epsilon| \, . \tag{84}$$

b) *Control of* I_2. Set $\widetilde{G}(z) = \sup_{|v| \le 2} \psi(z + v)$, where $\psi(z) = \varphi(z)(1 + z^2)^{3/2}$. Then $\widetilde{G}(z)$ is a symmetric integrable function of bounded variation, non-increasing in $t \ge 0$. Since $\Delta A_k = -\Delta V_k$, we have $|I_2| \le I_{2,1} + I_{2,2}$, where

$$\begin{aligned} I_{2,1} &= \mathbb{E}_\lambda \sum_{k=1}^n \frac{1}{2A_k} \left| \varphi'\left(T_{k-1}\right) \left(\Delta V_k - \Delta \left\langle X \right\rangle_k\right) \right| , \\ I_{2,2} &= \mathbb{E}_\lambda \sum_{k=1}^n \frac{1}{2A_k} \left| \varphi'\left(T_{k-1}\right) \left(\Delta \left\langle Y \right\rangle_k - \Delta \left\langle X \right\rangle_k\right) \right| . \end{aligned}$$

We first deal with $I_{2,1}$. Since $|\varphi'(z)| \le \psi(z) \le \widetilde{G}(z)$, for any real z, we have

$$\left| \varphi'\left(T_{k-1}\right) \right| \le \widetilde{G}\left(T_{k-1}\right). \tag{85}$$

Note that $0 \le \Delta V_k - \Delta \left\langle X \right\rangle_k \le 2\delta^2 \mathbf{1}_{\{k=n\}}$, $A_n = c_*^2 \gamma^2$ and $c_* \ge 4$. Then, using (85), we get the estimations

$$I_{2,1} \le \frac{c_2}{c_*} \, \mathbb{E}_\lambda \widetilde{G}\left(T_{n-1}\right),$$

and, by (79) with $G = \widetilde{G}$ and (81) with $t = T$,

$$|I_{2,1}| \le \frac{c_1}{c_*} \sup_z |\mathbb{P}_\lambda(Y_n \le z) - \Phi(z)| + c_2 \gamma.$$

We next consider $I_{2,2}$. By (59), we easily obtain the bound

$$\left| \Delta \left\langle Y \right\rangle_k - \Delta \left\langle X \right\rangle_k \right| \le c \lambda \epsilon \Delta \left\langle X \right\rangle_k \le c \lambda \epsilon \Delta V_k.$$

With this bound, we get

$$|I_{2,2}| \le c\lambda\epsilon\,\mathbb{E}_\lambda \sum_{k=1}^n \frac{1}{2A_k}\,|\varphi'\left(T_{k-1}\right)|\,\Delta V_k.$$

Since $|\varphi'(z)| \le \psi(z) \le \widetilde{G}(z)$, the right-hand side can be bounded exactly in the same way as J_1 in (74), with A_k replacing $A_k^{3/2}$. What we get is (cf. (82))

$$|I_{2,2}| \le c_1\lambda\epsilon \int_0^T \frac{dt}{a_t}\sup_z |\mathbb{P}_\lambda(Y_n \le z) - \Phi(z)| + c_2\lambda\epsilon \int_0^T \frac{dt}{a_t^{1/2}}.$$

By elementary computations, we see that

$$\int_0^T \frac{dt}{a_t^{1/2}} \le \int_0^T \frac{dt}{\sqrt{T-t}} \le c_2,$$

and, taking into account that $a_t \ge c_*^2\gamma^2$,

$$\int_0^T \frac{dt}{a_t} \le \frac{c_1}{c_*\lambda\epsilon}\int_0^T \frac{dt}{a_t^{1/2}} \le \frac{c_2}{c_*\lambda\epsilon}.$$

Then

$$|I_{2,2}| \le \frac{c_1}{c_*}\sup_z |\mathbb{P}_\lambda(Y_n \le z) - \Phi(z)| + c_2\lambda\epsilon.$$

Collecting the bounds for $I_{2,1}$ and $I_{2,2}$, we get

$$|I_2| \le \frac{c_1}{c_*}\sup_z |\mathbb{P}_\lambda(Y_n \le z) - \Phi(z)| + c_2\gamma. \tag{86}$$

c) *Control of* I_3. By Taylor's expansion,

$$I_3 = \frac{1}{8}\,\mathbb{E}_\lambda \sum_{k=1}^n \frac{1}{(A_k - \vartheta_k\Delta A_k)^2}\varphi'''\left(\frac{u - Y_{k-1}}{\sqrt{A_k - \vartheta_k\Delta A_k}}\right)\Delta A_k^2.$$

Since $|\Delta A_k| = \Delta V_k \le 12\gamma^2$ and $c_* \ge 4$, we have

$$A_k \le A_k - \vartheta_k\Delta A_k \le c_*^2\gamma^2 + T - V_k + 12\gamma^2 \le 2A_k. \tag{87}$$

Using (87) and the inequalities $|\varphi'''(z)| \le \psi(z) \le \widetilde{G}(z)$, we obtain

$$|I_3| \le c\gamma^2\mathbb{E}_\lambda \sum_{k=1}^n \frac{1}{A_k^2}\,\widetilde{G}\left(\frac{T_{k-1}}{\sqrt{2}}\right)\Delta V_k.$$

Proceeding in the same way as for estimating J_1 in (74), we get

$$|I_3| \leq \frac{c_1}{c_*} \sup_z |\mathbb{P}_\lambda(Y_n \leq z) - \Phi(z)| + c_2\gamma. \tag{88}$$

We are now in a position to end the proof of Lemma 3.1. From (63), using (84), (86) and (88), we find

$$|\mathbb{E}_\lambda \Phi_u(Y_n, A_n) - \mathbb{E}_\lambda \Phi_u(Y_0, A_0)| \leq \frac{c_1}{c_*} \sup_z |\mathbb{P}_\lambda(Y_n \leq z) - \Phi(z)| + c_2(\lambda\epsilon + \epsilon\,|\log \epsilon| + \delta).$$

Implementing the last bound in (62), we come to

$$\sup_z |\mathbb{P}_\lambda(Y_n \leq z) - \Phi(z)| \leq \frac{c_1}{c_*} \sup_z |\mathbb{P}_\lambda(Y_n \leq z) - \Phi(z)| + c_2(\lambda\epsilon + \epsilon\,|\log \epsilon| + \delta),$$

from which, choosing $c_* = \max\{2c_1, 4\}$, we get

$$\sup_z |\mathbb{P}_\lambda(Y_n \leq z) - \Phi(z)| \leq 2c_2(\lambda\epsilon + \epsilon\,|\log \epsilon| + \delta), \tag{89}$$

which proves Lemma 3.1 for $1 \leq \lambda < \epsilon^{-1}$.

For $0 \leq \lambda < 1$, we can prove Lemma 3.1 similarly by taking $\gamma = \epsilon |\log \epsilon| + \delta$. We only need to note that in this case, instead of (83),

$$\int_0^T \frac{dt}{a_t^{3/2}} \leq \frac{c}{c_* \epsilon |\ln \epsilon|} \qquad \text{and} \qquad \int_0^T \frac{dt}{a_t} \leq c\,|\log \epsilon|\,. \tag{90}$$

8. Equivalent conditions

In the following we give several equivalent conditions to the Bernstein condition (A1′). In the independent case equivalent conditions can be found in Saulis and Statulevičius [28]. For the convenience of the readers and motivated by the fact that in [28] the conditions are rather different from those used here, we decided to include independent proofs.

Proposition 8.1. *The following three conditions are equivalent:*
(I) *Bernstein's condition (A1′).*
(II) (Sakhanenko's condition) *There exists some positive absolute constant K such that*

$$K\,\mathbb{E}(|\eta_i|^3 \exp\{K|\eta_i|\}|\mathcal{F}_{i-1}) \leq \mathbb{E}(\eta_i^2|\mathcal{F}_{i-1}), \quad \text{for } 1 \leq i \leq n.$$

(III) *There exists some positive absolute constant ρ such that*

$$\mathbb{E}(|\eta_i|^k|\mathcal{F}_{i-1}) \leq \frac{1}{2} k! \rho^{k-2} \mathbb{E}(\eta_i^2|\mathcal{F}_{i-1}), \quad \text{for } k \geq 3 \ \text{ and } \ 1 \leq i \leq n.$$

Proof. First we prove that (I) implies (II). Let $t \in (0,1)$. By condition (I) and Lemma 4.1, we find that

$$\begin{aligned} E(|\eta_i|^3 e^{tH^{-1}|\eta_i|}|\mathcal{F}_{i-1}) &= \sum_{k=0}^{\infty} \frac{(tH^{-1})^k}{k!} E(|\eta_i|^{k+3}|\mathcal{F}_{i-1}) \\ &\leq \sum_{k=0}^{\infty} \frac{(tH^{-1})^k}{k!} (k+3)! H^{k+1} E(\eta_i^2|\mathcal{F}_{i-1}) \\ &\leq HE(\eta_i^2|\mathcal{F}_{i-1}) \sum_{k=0}^{\infty} \frac{(k+3)!}{k!}\, t^k \\ &=: f(t) HE(\eta_i^2|\mathcal{F}_{i-1}). \end{aligned} \tag{91}$$

Since $g(t) = tf(t)$ is a continuous function in $[0, \frac{1}{2}]$ and satisfies $g(0) = 0$ and $g(\frac{1}{2}) \geq 3$, there exists $t_0 \in (0, \frac{1}{2})$ such that $g(t_0) = 1$. Taking $K = t_0 H^{-1}$, we obtain condition (II) from (91).

Next we show that (II) implies (III). By the elementary inequality $x^k \leq k!\, e^x$, for $k \geq 0$ and $x \geq 0$, it follows that, for $k \geq 3$,

$$\begin{aligned} \mathbb{E}(|\eta_i|^k|\mathcal{F}_{i-1}) &= \mathbb{E}(|\eta_i|^3 K^{3-k} |K\eta_i|^{k-3}|\mathcal{F}_{i-1}) \\ &\leq (k-3)!\, K^{3-k} \mathbb{E}(|\eta_i|^3 \exp\{|K\eta_i|\}|\mathcal{F}_{i-1}). \end{aligned}$$

Using condition (II), for $k \geq 3$,

$$\mathbb{E}(|\eta_i|^k|\mathcal{F}_{i-1}) \leq (k-3)!\, K^{2-k} \mathbb{E}(\eta_i^2|\mathcal{F}_{i-1}) \leq \frac{1}{2} k! \rho^{k-2} \mathbb{E}(\eta_i^2|\mathcal{F}_{i-1}),$$

where $\rho = \frac{1}{K}$, which proves condition (III).

It is obvious that (III) implies (I) with $H = \rho$. □

Proposition 8.2. *If $\eta_1, ..., \eta_n$ are i.i.d., then Bernstein's condition, Cramér's condition and Sakhanenko's condition are all equivalent.*

Proof. According to Theorem 8.1, we only need to prove that Cramér's condition and Bernstein's condition are equivalent. We can assume that, a.s., $\eta_1 \neq 0$.

First, from (30), we find that Bernstein's condition (A1′) implies Cramér's condition:

$$\mathbb{E} e^{\frac{1}{2} H^{-1} \eta_1} < \infty.$$

Second, we show that Cramér's condition, i.e. $\mathbb{E} e^{c_0^{-1}|\eta_1|} := c_1 < \infty$, implies Bernstein's condition (A1′). By the inequality $x^k \leq k!\, e^x$, for $k \geq 0$ and $x \geq 0$, it follows that

$$|\mathbb{E}\eta_1^k| \leq c_0^k\, \mathbb{E}|c_0^{-1}\eta_1|^k \leq k!\, c_0^k\, \mathbb{E} e^{c_0^{-1}|\eta_1|} \;\; = \;\; k!\, c_0^k\, c_1.$$

Then, it is easy to see that, for $k \geq 3$,

$$|\mathbb{E}\eta_1^k| \leq \frac{1}{2}\, k!\, c_0^{k-2} \frac{2c_0^2 c_1}{\sigma^2}\, \mathbb{E}\eta_1^2 \leq \frac{1}{2}\, k!\, H^{k-2}\, \mathbb{E}\eta_1^2,$$

where $\sigma^2 = \mathbb{E}\eta_1^2$ and $H = \max\left\{c_0, \frac{2c_0^3 c_1}{\sigma^2}\right\}$, which proves that condition (A1′) is satisfied. □

Acknowledgements

We would like to thank the two referees for their helpful remarks and suggestions. The work has been partially supported by the National Natural Science Foundation of China (Grant no. 11101039 and Grant no. 11171044), and Hunan Provincial Natural Science Foundation of China (Grant No.11JJ2001). Fan was partially supported by the Post-graduate Study Abroad Program sponsored by China Scholarship Council.

[1] Bentkus, V. and Račkauskas, A., 1990. On probabilities of large deviations in Banach spaces. *Probab. Theory Relat. Fields*, **86**, 131–154.

[2] Bolthausen, E., 1982. Exact convergence rates in some martingale central limit theorems. *Ann. Probab.*, **10**, 672–688.

[3] Bose, A., 1986a. Certain non-uniform rates of convergence to normality for martingale differences. *J. Statist. Plann. Inference*, **14**, 155–167.

[4] Bose, A., 1986b. Certain non-uniform rates of convergence to normality for a restricted class of martingales. *Stochastics*, **16**, 279–294.

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

[6] Dembo, A., 1996. Moderate deviations for martingales with bounded jumps. *Electron. Comm. Proba.*, **1**, 11–17.

[7] Djellout, H., 2002. Moderate deviations for martingale differences and applications to $\phi-$mixing sequences. *Stochastic Stochastic Rep.*, **73**, 37–63.

[8] Esscher, F., 1932. On the probability function in the collective theory of risk. *Skand. Aktuarie Tidskr.*, **15**, 175–195.

[9] Fan, X., Grama, I. and Liu, Q., 2012. Hoeffding's inequality for supermartingales. *Stochastic Process. Appl.*, **122**, 3545–3559.

[10] Feller, W., 1943. Generalization of a probability limit theorem of Cramér. *Trans. Amer. Math. Soc.*, 361–372.

[11] Feller, W., 1971. *An introduction to probability theory and its applications*, J. Wiley and Sons.

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

[13] Grama, I., 1995. The probabilities of large deviations for semimartingales. *Stochastic Stochastic Rep.*, **54**, 1–19.

[14] Grama, I., 1997. On moderate deviations for martingales. *Ann. Probab.*, **25**, 152–184.

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

[16] Grama, I. and Haeusler, E., 2006. An asymptotic expansion for probabilities of moderate deviations for multivariate martingales. *J. Theoret. Probab.*, **19**, 1–44.

[17] Gulinsky, O. V. and Veretennikov, A. Yu., 1993. *Large deviations for discrete-time processes with averaging.* VSP, Utrecht.

[18] Gulinsky, O. V., Liptser, R. S. and Lototskii, S. V., 1994. Large deviations for unbounded additive functionals of Markov processes with discrete time. *J. Appl. Math. Stochastic Anal.*, **7**.

[19] Hall, P. and Heyde, C. C., 1980. *Martingale Limit Theory and Its Application.* Academic Press.

[20] Huang, C., 2012. Moments, moderate and large deviations for a branching process in a random environment. *Stochastic Process. Appl.*, **122**, 522–545.

[21] Liptser, R.S. and Pukhalskii, A. A., 1992. Limit theorems on large deviations for semimartingales. *Stochastic Stochastic Rep.*, **38**, 201–249.

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

[23] Petrov, V. V., 1975. *Sums of Independent Random Variables.* Springer-Verlag. Berlin.

[24] Račkauskas, A., 1990. On probabilities of large deviations for martingales. *Liet. Mat. Rink.*, **30**, 784–795.

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

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

[27] Rubin, H. and Sethuraman, J., 1965. Probabilities of moderate deviations. *Sankhyā Ser. A*, **37**, 325–346.

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

[29] Statulevičius, V. A., 1966. On large deviations. *Probab. Theory Relat. Fields*, **6**, 133–144.

[30] Worms, J., 2001. Moderate deviations for some dependent variables, part I: Martingales. *Math. Methods Statist.*, **10**, 38–72.