The Annals of Probability
1999, Vol. 27, No. 4, 2068–2088

AN EXPONENTIAL NONUNIFORM BERRY–ESSEEN BOUND FOR SELF-NORMALIZED SUMS

BY QIYING WANG AND BING-YI JING

University of Wollongong and Hong Kong University of Science and Technology

In this paper we shall derive exponential nonuniform Berry–Esseen bounds in the central limit theorem for self-normalized sums. We show that the size of the error can be reduced considerably by replacing the usual standardization by self-normalization. In particular, we establish the exponential bounds for the probability of the self-normalized sums under the condition that the third moment is finite, whereas an exponential moment assumption is required for the standardized sums. Applications to t-statistics and the probabilities of moderate deviations of self-normalized sums are also discussed.

1. Introduction. Let $X_1, \dots, X_n$ be independent nondegenerate random variables such that $EX_j = 0$ and $\mathrm{var}(X_j) = \sigma_j^2 < \infty$. Set

$$S_n = \sum_{j=1}^n X_j, \qquad B_n^2 = \sum_{j=1}^n \sigma_j^2, \qquad V_n^2 = \sum_{j=1}^n X_j^2.$$

Under appropriate conditions (e.g., the Lindeberg condition), it is well known that

$$\Delta_n(x) \equiv \left|P(S_n/B_n \le x) - \Phi(x)\right| \to 0 \quad \text{as } n \to \infty$$

uniformly in $x \in R$, where $\Phi(x)$ is the distribution function of the standard normal variable and R is the real line. Under some further conditions, the rate of convergence to normality is provided by the Berry–Esseen inequality

$$\Delta_n(x) \le \frac{A}{B_n^3} \sum_{j=1}^n E|X_j|^3, \tag{1.1}$$

which holds uniformly in $x \in R$, where $A > 0$ is an absolute constant; see Theorem 3 of Petrov [(1975), page 111]. Although the bound given by (1.1) is valid for all $x \in R$, it is only useful for values of x near the center of the distribution. The reason is that, for x sufficiently large, the difference $P(S_n/B_n \le x) - \Phi(x)$ becomes so close to 0 that the bound (1.1) is simply too crude to be of any use. One way to refine the Berry–Esseen bound is to reflect dependence on x as well as n. In this direction, under the assumption

Received January 1996; revised April 1999.
AMS 1991 *subject classifications*. Primary 60F10, 60F15; secondary 60G50, 62F03.
Key words and phrases. Berry–Esseen bound, self-normalized sum, t-statistics, moderate deviation, nonuniform estimate.

$E|X_j|^3 < \infty$, the bound (1.1) can be replaced by

$$\Delta_n(x) \le \frac{A}{(1+|x|)^3 B_n^3} \sum_{j=1}^n E|X_j|^3, \tag{1.2}$$

where $A > 0$ is an absolute constant; see Problem 23 of Petrov [(1975), page 132]. It should be noted that the power of $|x|$ in (1.2) is optimal under the assumed moment conditions; see Michel (1976).

The purpose of this paper is to develop nonuniform Berry–Esseen bounds for so-called self-normalized sums, S_n/V_n. In other words, we wish to obtain a bound for

$$\delta_n(x) \equiv \left|P(S_n/V_n \le x) - \Phi(x)\right|.$$

We shall see that the Berry–Esseen bound for the self-normalized sum S_n/V_n can be reduced considerably from the one in (1.2) for the standardized sum S_n/B_n. More precisely, we shall show that the polynomial bounds in x for the standardized sum in (1.2) can now be replaced by the exponential bounds for the self-normalized sum under only the finite third moment condition; see Theorems 2.1 and 2.2 in Section 2. These results are not only of interest in their own right but also should be useful in many applications.

The fact that the exponential bounds for the self-normalized sum hold under only the finite third moment condition is a seemingly surprising result since such results are usually available under the assumption that the moment generating function exists around the origin. We are not aware of any such results in the literature regarding the nonuniform Berry–Esseen bounds. The stark contrasts in their limiting behavior between the standardized sum and self-normalized sum have also been noted in other contexts; see, for instance, Logan, Mallow, Rice and Shepp (1973), Griffin and Kuelbs (1989, 1991) and Shao (1997), Giné, Götze and Mason (1997).

The rest of the paper is organized as follows. The main results will be given in Section 2. Applications to Student's t-statistics are discussed in Section 3. In Section 4, we shall introduce some lemmas, which will be needed in the proofs of the main results. Finally, the proofs of the main theorems are given in Section 5.

Before we end the section, let us introduce some further notation. Let $X_1, \dots, X_n$ be a sequence of independent nondegenerate random variables. Throughout the paper we shall assume without loss of generality that $EX_j = 0$ for $j = 1, \dots, n$. Furthermore, we let

$$L_{3n} = B_n^{-3} \sum_{j=1}^n E|X_j|^3.$$

In particular, if $X_1, \dots, X_n$ are independent and identically distributed (i.i.d.) random variables, then we shall denote

$$\sigma^2 = \mathrm{var}(X_1) \quad \text{and} \quad \beta_j = \sigma^{-j} E|X_1|^j \quad \text{for } j \ge 3.$$

Clearly, for i.i.d. random variables, we have $B_n^2 = n\sigma^2$ and $L_{3n} = n^{-1/2}\beta_3$, respectively. We shall use A to denote a generic absolute positive constant which may be different at each occurrence. Finally, we shall make regular reference to the following well-known inequality:

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x}-\frac{1}{x^3}\right)\exp\left(\frac{-x^2}{2}\right) \\ &\qquad \le 1-\Phi(x) \le \frac{1}{\sqrt{2\pi}x}\exp\left(\frac{-x^2}{2}\right) \quad \text{for } x>0. \end{aligned} \tag{1.3}$$

2. Main results. We now present exponential nonuniform Berry–Esseen bounds for the convergence rate of the distribution function of the self-normalized sums to normality. Since the results for symmetric random variables are better than those in general cases, for more clarity we shall deal with the two cases separately.

2.1. *Symmetric random variables.* The following theorem gives exponential nonuniform Berry–Esseen bounds for the self-normalized mean for symmetric random variables. Its proof is long and tedious, and hence will be postponed to Section 5.

THEOREM 2.1. *Let $X_1, \dots, X_n$ be independent symmetric random variables with $E|X_j|^3 < \infty$ for $1 \le j \le n$.*

(i) *If $|x| \le (5L_{3n}^{1/3})^{-1}$, we have*

$$\delta_n(x) \le A\left\{(1+x^2)L_{3n} + \sum_{j=1}^n P\big(|X_j| > B_n(6|x|)^{-1}\big)\right\}\exp\left(\frac{-x^2}{2}\right). \tag{2.1}$$

(ii) *If $|x| \ge (5L_{3n}^{1/3})^{-1}$, we have*

$$\delta_n(x) \le \left(1+\frac{1}{\sqrt{2\pi}|x|}\right)\exp\left(\frac{-x^2}{2}\right). \tag{2.2}$$

Theorem 2.1 can be illustrated more clearly if $X_1, \dots, X_n$ are i.i.d. random variables, in which $L_{3n} = n^{-1/2}\beta_3$. Therefore, the inequalities (2.1) and (2.2) give exponential nonuniform Berry–Esseen bounds, respectively, for the two regions $|x| \le cn^{1/6}$ and $|x| \ge cn^{1/6}$, where $c = (5\beta_3^{1/3})^{-1}$.

Under the assumption $E|X_j|^3 < \infty$, we can apply the Markov inequality to further simplify inequality (2.1) to $\delta_n(x) \le A(1+|x|^3)L_{3n}\exp(-x^2/2)$ for $|x| \le (5L_{3n}^{1/3})^{-1}$. Hence the following corollary follows immediately from Theorem 2.1.

COROLLARY 2.1. *Let $X_1, \dots, X_n$ be independent symmetric random variables with $E|X_j|^3 < \infty$ for $1 \le j \le n$. Then for all $n \ge 1$ and $x \in R$, we have* *that*

$$\delta_n(x) \le A \min\{(1+|x|^3)L_{3n}, 1\} \exp\left(\frac{-x^2}{2}\right).$$

Note that in Theorem 2.1, only the finite third moment condition is imposed. Under slightly stronger moment condition, Theorem 2.1 yields the following moderate deviation results.

COROLLARY 2.2. *Let $X_1, \dots, X_n$ be i.i.d. symmetric random variables with $E|X_1|^{7/2} < \infty$. Then for any sequence t_n satisfying $t_n \to \infty$ and $t_n/n^{1/6} \to 0$ as $n \to \infty$, we have*

$$\frac{P(S_n/V_n > t_n)}{1-\Phi(t_n)} \to 1 \quad \text{and} \quad \frac{P(S_n/V_n \le -t_n)}{\Phi(-t_n)} \to 1. \tag{2.3}$$

PROOF. Under the i.i.d. assumption, we have $B_n^2 = n\sigma^2$ and $L_{3n} = n^{-1/2}\beta_3$. In view of the relationship $1-\Phi(x) \sim x^{-1}\exp(-x^2/2)/\sqrt{2\pi}$ for $x \to \infty$, we can apply Theorem 2.1(i) to get

$$\begin{aligned}
&\left|\frac{P(S_n/V_n > t_n)}{1-\Phi(t_n)} - 1\right| \\
&\quad \sim \sqrt{2\pi}t_n \exp\left(t_n^2/2\right)\left|P\left(S_n/V_n > t_n\right) - (1-\Phi(t_n))\right| \\
&\quad \le A\beta_3(t_n + t_n^3)n^{-1/2} + At_n nP(|X_1| > (n\sigma^2)^{1/2}(6t_n)^{-1}) \\
&\quad \le A\beta_3(t_n + t_n^3)n^{-1/2} + At_n nE|X_1|^{7/2}t_n^{7/2}(n\sigma^2)^{-7/4} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(by Markov's inequality)} \\
&\quad = o(1) \quad \text{as } t_n = o(n^{1/6}).
\end{aligned}$$

The second part of Corollary 2.2 follows by replacing X_j by $-X_j$ in the above proof. This completes the proof. □

REMARK 2.1. Similar results to those in Corollary 2.2 hold for the standardized means under much stronger conditions. For instance, from Linnik (1962), it is well known that for the sequence t_n as in Corollary 2.2,

$$\frac{P\left(S_n/\sqrt{n\sigma^2} > t_n\right)}{1-\Phi(t_n)} \to 1 \quad \text{and} \quad \frac{P\left(S_n/\sqrt{n\sigma^2} \le -t_n\right)}{\Phi(-t_n)} \to 1$$

hold only when $E(\exp(s|X_1|^{1/2})) < \infty$ for some $s > 0$, which is a very stringent condition in many applications since it requires that all moments of X_1 exist.

REMARK 2.2. Q. M. Shao pointed out in private communications that Corollary 2.2 above was also obtained in his unpublished manuscript (1998) under weaker moment condition $E|X_1|^3 < \infty$ and without the symmetry constraint.

2.2. *Nonsymmetric random variables.* Now we shall investigate the more general case where the random variables $X_1, \dots, X_n$ are not necessarily symmetric. It turns out that we can still get a nonuniform exponential Berry–Esseen bound for the self-normalized sum although this bound is rougher than the one given in Theorem 2.1 for symmetric random variables.

THEOREM 2.2. *Let $X_1, \dots, X_n$ be i.i.d. random variables with $E|X_1|^3 < \infty$. Then for any $0 < \varepsilon < 1/2$, there exist constants $0 < \eta_0 < 1$ and $A(\varepsilon)$ (depending only on ε) such that for all $|x| \le \eta_0\sqrt{n}$,*

$$\begin{aligned}\delta_n(x) &\le A(\varepsilon)\big[\beta_3 n^{-1/2} + \sigma^{-2}E\big(X_1^2 I(|X_1| > \sigma n^{3/8}(2|x|)^{-1})\big)\big] \\ &\quad \times \exp\left(\frac{-(1-\varepsilon)x^2}{2}\right).\end{aligned} \tag{2.4}$$

Under slightly stronger moment condition than the existence of the third moment, we can derive the next corollary from Theorem 2.2.

COROLLARY 2.3. *Let $X_1, \dots, X_n$ be i.i.d. random variables with $E|X_1|^{10/3} < \infty$. Then there exists an absolute constant $0 < \eta < 1$ such that*

$$\delta_n(x) \le \frac{A\beta_{10/3}}{\sqrt{n}} \exp\left(\frac{-\eta x^2}{2}\right) \quad \text{for } x \in R \text{ and } n \ge 1. \tag{2.5}$$

PROOF. Since $E|X_1|^{10/3} < \infty$, it is easy to see that

$$\sigma^{-2}E\big(X^2 I(|X| > \sigma n^{3/8}(2|x|)^{-1}\big) \le \sigma^{-10/3}E|X|^{10/3} n^{-1/2}(2|x|)^{4/3}.$$

Then from Theorem 2.2, for a special $\varepsilon = 1/4$, there exist absolute positive constants $\eta_0 < 1$ and A such that for any $|x| \le \eta_0\sqrt{n}$, and any η_1 satisfying $0 < \eta_1 < 1 - \varepsilon$, we have

$$\begin{aligned}\delta_n(x) &\le A\big(\beta_3 n^{-1/2} + \beta_{10/3}(2|x|)^{4/3} n^{-1/2}\big) \exp\left(\frac{-(1-\varepsilon)x^2}{2}\right) \\ &\le A\beta_{10/3} n^{-1/2} \exp\left(\frac{-\eta_1 x^2}{2}\right),\end{aligned} \tag{2.6}$$

where we have used the inequality $\beta_3 \le \beta_{10/3}$. Therefore (2.5) holds for $|x| \le \eta_0\sqrt{n}$.

We now investigate the case for $|x| \ge \eta_0\sqrt{n}$. First, if $\eta_0\sqrt{n} \le |x| \le \sqrt{n}$, then applying (2.6) and (1.3) we have that for any $0 < \eta_2 < 1$ (such as $\eta_2 = 1/2$),

$$\begin{aligned}P\left(\frac{S_n}{V_n} > |x|\right) &\le P\left(\frac{S_n}{V_n} > \eta_2\eta_0\sqrt{n}\right) \\ &\le 1 - \Phi\big(\eta_2\eta_0\sqrt{n}\big) + \frac{A\beta_{10/3}}{\sqrt{n}} \exp\left(-\frac{1}{2}\eta_1\big(\eta_2\eta_0\sqrt{n}\big)^2\right) \\ &\le A\beta_{10/3} n^{-1/2} \exp\left(\frac{-\eta x^2}{2}\right),\end{aligned}$$

where $0 < \eta \equiv \eta_1 \eta_2^2 \eta_0^2 < 1$. On the other hand, if $|x| \geq \sqrt{n}$, we have that $P(S_n > |x| V_n) = 0$ since $S_n^2 \leq n V_n^2$ by Cauchy–Schwarz inequality. Hence for $|x| \geq \sqrt{n}$, by applying (1.3) again, we get

$$\delta_n(x) \leq P\big(S_n/V_n > |x|\big) + 1 - \Phi(|x|) \leq A\beta_{10/3} n^{-1/2} \exp\left(\frac{-\eta x^2}{2}\right).$$

The proof of Corollary 2.3 is thus complete. □

REMARK 2.3. In this section, we have established exponential nonuniform Berry–Esseen bounds for the self-normalized sums under moment conditions. This is in stark contrast with Berry–Esseen theorems for the standardized sums, where exponential bounds are only available under the exponential moment condition (cf. Remark 2.1). One reason why only moment conditions are sufficient for the self-normalized sums is that large values of X_i play no role in the tail probability behavior of $P(S_n/V_n \leq x)$ since they appear in both numerator and denominator and effectively cancel each other's influences.

3. An application to Student's *t*-statistics. Let $X_1, \ldots, X_n$ be independent random variables with $EX_i = 0$ for $1 \leq i \leq n$. Consider Student's t-statistic T_n defined by

$$T_n = \sqrt{n}\bar{X}/\hat{\sigma},$$

where $\bar{X}_n = S_n/n$ and $\hat{\sigma}^2 = \sum_{j=1}^n (X_j - \bar{X}_n)^2/(n-1)$. It is well known [see Efron (1969)] that for $x \geq 0$,

$$P(T_n \geq x) = P\left(\frac{S_n}{V_n} \geq x\left(\frac{n}{n+x^2-1}\right)^{1/2}\right). \tag{3.1}$$

With the help of (3.1), the following nonuniform bounds and Cramér moderate deviations for the t-statistics can be derived from those results presented in Section 2.

THEOREM 3.1. *Let $X_1, \ldots, X_n$ be independent symmetric random variables with $E|X_j|^3 < \infty$ for $1 \leq j \leq n$. Then we have that for all $n \geq 2$ and $x \in R$,*

$$|P\big(T_n \leq x\big) - \Phi(x)| \leq A \min\{(1+|x|^3)L_{3n}, 1\} \exp\left(-\frac{nx^2}{2(n+x^2-1)}\right).$$

THEOREM 3.2. *Let $X_1, \ldots, X_n$ be i.i.d. random variables with $E|X|^{10/3} < \infty$. Then there exist absolute constant $0 < \eta < 1/2$ such that for all $n \geq 2$ and $x \in R$,*

$$\big|P\big(T_n \leq x\big) - \Phi(x)\big| \leq \frac{A\beta_{10/3}}{\sqrt{n}} \exp\left(-\frac{\eta n x^2}{n+x^2-1}\right).$$

THEOREM 3.3. *Let $X_1, \dots, X_n$ be i.i.d. symmetric random variables with $E|X|^{7/2} < \infty$. Then for all t_n such that $t_n \to \infty$ and $t_n/n^{1/6} \to 0$ as $n \to \infty$, we have*

$$\frac{P(T_n > t_n)}{1 - \Phi(t_n)} \to 1 \quad \text{and} \quad \frac{P(T_n \le -t_n)}{\Phi(-t_n)} \to 1. \tag{3.2}$$

REMARK 3.1. For the Cramér type large deviations of t-statistics, Vandemaele and Veraverbeke (1985) established the following results: if $E|X|^p \le K^p p^{rp}$, for all integer numbers $p \ge 2$ (which implies that $E(\exp(t|X|^{1/r})) < \infty$, for some $t > 0$), then (3.2) holds, uniformly in the range $0 \le x \le o(n^\alpha)$ with

$$\alpha = \begin{cases} 1/6, & \text{if } r \ge 2, \\ (4+4r)^{-1}, & \text{if } 2/3 \le r \le 2, \\ (8r-2)^{-1}, & \text{if } 0 \le r \le 2/3. \end{cases}$$

From Theorem 3.3, we see that for symmetric random variables, the result (3.2) still holds if we replace the exponential moment condition by the third moment condition.

In the remainder of this section, we shall give a proof of Theorem 3.1. The proofs of Theorems 3.2 and 3.3 are similar but simpler than that of Theorem 3.1 and hence omitted here.

PROOF OF THEOREM 3.1. Without loss of generality, assume $x \ge 0$. For $0 \le x \le 1$, the Berry–Esseen bound was given by Bentkus and Götze (1996). So we only need to show the theorem for the case $x > 1$ below. Write $a = n^{1/2}/(n + x^2 - 1)^{1/2}$. It is easy to see that $0 < a < 1$, $ax > 1$ and

$$|ax - x| = \frac{(a^2-1)x}{a+1} = \frac{(x^2-1)x}{[n + (x^2-1)](a+1)} \le \frac{x^3}{n} \le A(1+x^3)L_{3n},$$

where we have used $L_{3n} \ge 1/n$ by Jensen's and Hölder's inequalities. Then applying the mean-value theorem, we get

$$\begin{aligned} \big|\Phi(ax) - \Phi(x)\big| &= |\phi(x_0)(ax - x)| \le \phi(ax)|ax - x| \\ &\le A(1+x^3)L_{3n} \exp\left(\frac{-a^2x^2}{2}\right), \end{aligned} \tag{3.3}$$

where $\phi(x) = \Phi'(x)$ and $ax \le x_0 \le x$. On the other hand, by using the inequality (1.3), we get

$$\begin{aligned} \big|\Phi(ax) - \Phi(x)\big| &\le \frac{\phi(ax)}{ax} + \frac{\phi(x)}{x} \le \phi(ax)\left(\frac{1}{ax} + \frac{1}{x}\right) \\ &\le A \exp\left(\frac{-a^2x^2}{2}\right). \end{aligned} \tag{3.4}$$

Then it follows from (3.1), (3.3), (3.4) and Corollary 2.1 that

$$\begin{aligned}\big|P(T_n \le x) - \Phi(x)\big| &\le \big|P(S_n/V_n \le ax) - \Phi(ax)\big| + \big|\Phi(ax) - \Phi(x)\big| \\ &\le A \min\{(1+x^3)L_{3n}, 1\} \exp\left(\frac{-a^2x^2}{2}\right),\end{aligned}$$

which completes the proof. □

4. Some preliminary lemmas. In this section, we shall provide some lemmas which will be needed in the proofs of the main results. These lemmas are also of interest in their own right.

LEMMA 4.1. *Let $X_1, \dots, X_n$ be independent random variables with $E|X_j|^3 < \infty$.*

(i) *For $n \ge 1$ and $x > 0$ satisfying $(1+x^3)L_{3n} \le 1/125$, we have*

$$\begin{aligned}P\big(S_n > x(V_n^2 + B_n^2)/(2B_n)\big) &= \big(1-\Phi(x)\big)\exp\left(r_{1n}(x)\right) \\ &\quad + \exp\left(\frac{-x^2}{2}\right) r_{2n}(x),\end{aligned} \tag{4.1}$$

where $|r_{1n}(x)| \le 14x^3 L_{3n}$ and $|r_{2n}(x)| \le A(1+x^2)L_{3n}\exp(14x^3L_{3n})$.

(ii) *For $n \ge 1$ and $x \ge 1$ satisfying $x^3 L_{3n} \le 1/125$, we have*

$$P\big(S_n > x(V_n^2 + B_n^2)/(2B_n)\big) = \big(1-\Phi(x)\big)\big(1 + r_{3n}(x)\big), \tag{4.2}$$

where $|r_{3n}(x)| \le Ax(x^2+1)L_{3n}\exp(14x^3L_{3n})$.

PROOF. (i) First we note that the left-hand side of (4.1) can be rewritten as

$$P\big(S_n > x(V_n^2 + B_n^2)/(2B_n)\big) = P\left(\sum_{j=1}^{n} \eta_j > xB_n\right), \tag{4.3}$$

where

$$h = \frac{x}{B_n}, \qquad \eta_j = X_j - \frac{h}{2}(X_j^2 - \sigma_j^2).$$

Note that $\eta_1, \dots, \eta_n$ are independent random variables and that $E\exp(h\eta_j)$ always exists for $x > 0$. The rest of the proof is based on the conjugate method [first introduced by Esscher (1932)] which is a very useful tool in deriving large deviation probabilities; see also Petrov [(1975), page 221] or Feller [(1971), page 549]. To employ the method, let $\xi_1, \dots, \xi_n$ be independent random variables with ξ_j having distribution function $V_j(u)$ defined by

$$V_j(u) = E\big(\exp(h\eta_j)I(\eta_j \le u)\big)\big/E\exp(h\eta_j) \quad \text{for } j = 1, \dots, n.$$

Also define $M_n^2(h) = \sum_{j=1}^n \operatorname{Var}(\xi_j)$ and

$$G_n(t) = P\left(\frac{\sum_{j=1}^n(\xi_j - E\xi_j)}{M_n(h)} \le t\right), \qquad R_n(h) = \frac{xB_n - \sum_{j=1}^n E\xi_j}{M_n(h)}.$$

Then using the conjugate method, integration by parts and the indentity $\int_0^\infty \exp(-sx)\,d\Phi(x) = \exp(s^2/2)(1-\Phi(s))$, we have

$$
\begin{aligned}
&P\Big(\sum_{j=1}^n \eta_j > xB_n\Big)\\
&\quad= \Big(\prod_{j=1}^n E\,\exp(h\eta_j)\Big)\int_{xB_n}^\infty e^{-hu}\,dP\Big(\sum_{j=1}^n \xi_j \le u\Big),\\
&\quad= \Big(\prod_{j=1}^n E\exp(h\eta_j)\Big)\int_0^\infty \exp(-hxB_n - hM_n(h)v)\,dG_n\big(v+R_n(h)\big),\\
&\qquad\quad [\text{by a change of variable } u = xB_n + vM_n(h)]\\
&\quad= \Big(\prod_{j=1}^n E\,\exp(h\eta_j)\Big)\exp(-x^2)\\
&\qquad\times\Big(\int_0^\infty \exp\big(-hM_n(h)v\big)\,d\big(G_n\big(v+R_n(h)\big)-\Phi(v)\big)\\
&\qquad\qquad + \int_0^\infty \exp(-hM_n(h)v)\,d\Phi(v)\Big)\\
&\quad= I_0(h)\exp(-x^2)\Big(\exp\Big(\frac{x^2}{2}\Big)(1-\Phi(x)) + I_1(h)+I_2(h)+I_3(h)\Big),
\end{aligned}
\tag{4.4}
$$

where

$$
\begin{aligned}
I_0(h) &= \prod_{j=1}^n E\,\exp(h\eta_j),\\
I_1(h) &= \int_0^\infty \exp(-hM_n(h)v)\,d\big(G_n(v+R_n(h)) - \Phi(v+R_n(h))\big),\\
I_2(h) &= \int_0^\infty \exp(-hM_n(h)v)\,d\big(\Phi(v+R_n(h)) - \Phi(v)\big),\\
I_3(h) &= \int_0^\infty \big(\exp(-hM_n(h)v) - \exp(-xv)\big)\,d\Phi(v).
\end{aligned}
$$

To estimate $I_i(h)$ for $0 \le i \le 3$, we need to establish some inequalities first. From the assumption that $(1+x^3)L_{3n} \le 1/125$ and Jensen's inequality, we get

$$
\sigma_j^3 \le E|X_j|^3, \qquad \sigma_j h \le \big(x^3B_n^{-3}E|X_j|^3\big)^{1/3} \le 1/5, \tag{4.5}
$$

$$
h\eta_j = -\tfrac{1}{2}h^2(X_j - h^{-1})^2 + \tfrac{1}{2} + \tfrac{1}{2}\sigma_j^2h^2 \le 13/25. \tag{4.6}
$$

Since $EX_j = 0$ and $E|X_j|^3 < \infty$, it is easy to derive the following inequalities:

$$
\big|E\big(\eta_j I(|X_j| \le h^{-1})\big)\big| \le 2h^2E\big(|X_j|^3 I(|X_j| > h^{-1})\big), \tag{4.7}
$$

$$\left|E\left(\eta_j^2 I(|X_j| \le h^{-1}) - \sigma_j^2\right)\right| \le \tfrac{3}{2} h\left(E|X_j|^3 + h\sigma_j^4\right), \tag{4.8}$$

$$E\left(|\eta_j|^3 I(|X_j| \le h^{-1})\right) \le 6E\left(|X_j|^3 I(|X_j| \le h^{-1})\right) + 2h^3\sigma_j^6. \tag{4.9}$$

From the elementary inequality $|e^x - 1 - x - x^2/2| \le |x|^3 e^x/6$ for any $x \in R$, we have

$$\begin{aligned} E\exp(h\eta_j) &= E\left(\exp(h\eta_j) I\left(|X_j| \le h^{-1}\right)\right) + E\left(\exp(h\eta_j) I\left(|X_j| > h^{-1}\right)\right) \\ &= E\Big(\left(1 + h\eta_j + \tfrac{1}{2}(h\eta_j)^2\right) I\left(|X_j| \le h^{-1}\right)\Big) \\ &\quad + E\left(\exp(h\eta_j) I\left(|X_j| > h^{-1}\right)\right) \\ &\quad + E\Big(\left(\exp(h\eta_j) - 1 - h\eta_j - \tfrac{1}{2}(h\eta_j)^2\right) I\left(|X_j| \le h^{-1}\right)\Big) \\ &= 1 + \tfrac{1}{2}h^2\sigma_j^2 + l_{1j}(h) \\ &= \exp\left(\tfrac{1}{2}h^2\sigma_j^2 + l_{2j}(h)\right), \end{aligned} \tag{4.10}$$

$$\left(E\exp(h\eta_j)\right)^{-1} = 1 - \tfrac{1}{2}h^2\sigma_j^2 + l_{3j}(h), \tag{4.11}$$

where, from (4.7)–(4.9) and noting that $\exp(\zeta h\eta_j) \le 2$ for $0 \le \zeta \le 1$,

$$\begin{aligned} |l_{1j}(h)| &\le h\left|E\left(\eta_j I(|X_j| \le h^{-1})\right)\right| + \tfrac{1}{2}h^2\left|E\left(\eta_j^2 I(|X_j| \le h^{-1})\right) - \sigma_j^2\right| \\ &\quad + \tfrac{1}{3}h^3 E\left(|\eta_j|^3 I(|X_j| \le h^{-1})\right) + 3P(|X_j| > h^{-1}) \\ &\le 7h^3 E|X_j|^3 \quad (\le 1/16), \\ |l_{2j}(h)| &\le 2|l_{1j}(h)| \le 14h^3 E|X_j|^3, \\ |l_{3j}(h)| &\le 2|l_{1j}(h)| \le 14h^3 E|X_j|^3. \end{aligned}$$

Similarly, by noting that (4.6) implies that $|h\eta_j|^k \exp(h\eta_j) \le e$ for $k = 1, 2, 3$, we have that

$$\left|E\left(\eta_j \exp(h\eta_j)\right) - h\sigma_j^2\right| \le 16h^2 E|X_j|^3, \tag{4.12}$$

$$\left|E\left(\eta_j^2 \exp(h\eta_j)\right) - \sigma_j^2\right| \le 30hE|X_j|^3, \tag{4.13}$$

$$E\left(|\eta_j|^3 \exp(h\eta_j)\right) \le 30E|X_j|^3. \tag{4.14}$$

It follows from (4.5)–(4.14) that

$$E\xi_j = E\left(\eta_j \exp(h\eta_j)\right)/E\exp(h\eta_j) = h\sigma_j^2 + l_{4j}(h), \tag{4.15}$$

$$\operatorname{Var}(\xi_j) = E\left(\eta_j^2 \exp(h\eta_j)\right)/\left(E\exp(h\eta_j)\right)^2 - \left(E\xi_j\right)^2 = \sigma_j^2 + l_{5j}(h), \tag{4.16}$$

$$E|\xi_j|^3 = E\left(|\eta_j|^3 \exp(h\eta_j)\right)/E\exp(h\eta_j) \le 34E|X_j|^3, \tag{4.17}$$

where

$$
\begin{aligned}
|l_{4j}(h)| &\le \big|\big(1/E\,\exp(h\eta_j)-1\big)E\big(\eta_j\,\exp(h\eta_j)\big)\big| + \big|E\big(\eta_j\,\exp(h\eta_j)\big)-h\sigma_j^2\big| \\
&\le \big|\big(l_{3j}(h)-\tfrac{1}{2}h^2\sigma_j^2\big)\big(h\sigma_j^2+16h^2E|X_j|^3\big)\big| + 16h^2E|X_j|^3 \\
&\le 22h^4\sigma_j^2E|X_j|^3 + 14\times 16h^5(E|X_j|^3)^2 + 16h^2E|X_j|^3 + \tfrac{1}{2}h^3\sigma_j^4 \\
&\le 20h^2E|X_j|^3; \\
|l_{5j}(h)| &\le (E\xi_j)^2 + \big|\big(1/E\,\exp(h\eta_j)-1\big)E\big(\eta_j^2\,\exp(h\eta_j)\big)\big| \\
&\qquad + \big|\big(E\eta_j^2\exp(h\eta_j)\big)-\sigma_j^2\big| \\
&\le 5hE|X_j|^3 + \Big|\Big(l_{3j}(h)-\tfrac{1}{2}h^2\sigma_j^2\big)\big(\sigma_j^2+30hE|X_j|^3\big)\Big| + 30hE|X_j|^3 \\
&\le 41hE|X_j|^3.
\end{aligned}
$$

Then by the assumption $(1+x^3)L_{3n} \le 1/125$, we can get

$$
M_n^2(h) = B_n^2 + \sum_{j=1}^n l_{5j}(h) > \tfrac{2}{3}B_n^2. \tag{4.18}
$$

We are now ready to estimate $I_j(h)$, $0 \le j \le 3$. For $I_0(h)$, we use (4.10) to get

$$
I_0(h) = \exp\Bigg(\frac{1}{2}h^2B_n^2 + \sum_{j=1}^n l_{2j}(h)\Bigg) = \exp\left(\frac{x^2}{2}\right)\exp\Bigg(\sum_{j=1}^n l_{2j}(h)\Bigg). \tag{4.19}
$$

By (4.15)–(4.18), the Berry–Esseen bound and Taylor expansion, we have

$$
I_1(h) \le \sup_x |G_n(v)-\Phi(v)| \le \frac{A}{M_n^3(h)}\sum_{j=1}^n E|\xi_j - E\xi_j|^3 \le AL_{3n}, \tag{4.20}
$$

$$
I_2(h) \le \sup_x |\Phi(v+R_n(h))-\Phi(v)| \le \frac{A}{M_n(h)}\sum_{j=1}^n |l_{4j}(h)| \le Ax^2L_{3n}. \tag{4.21}
$$

By applying the mean value estimate to $I_3(h)$ [see Petrov (1975), page 227], we get

$$
\begin{aligned}
|I_3(h)| &\le \frac{1}{x}\left|\frac{M_n(h)}{B_n}-1\right|\max\left\{1, \frac{B_n^2}{M_n^2(h)}\right\} \\
&\le \frac{3}{2x}\left|\frac{M_n^2(h)-B_n^2}{B_n(M_n(h)+B_n)}\right| \\
&\le AL_{3n}.
\end{aligned} \tag{4.22}
$$

It then follows from the relationships (4.3)–(4.4) and the estimates (4.19)–(4.22) that

$$
\begin{aligned}
&P\big(S_n > x(V_n^2 + B_n^2)/(2B_n)\big) \\
&\quad = \exp\left(\frac{-x^2}{2}\right)\exp\big(r_{1n}(x)\big)\left(\exp\left(\frac{x^2}{2}\right)(1-\Phi(x)) + \sum_{j=1}^{3} I_j(h)\right) \qquad (4.23)\\
&\quad = (1-\Phi(x))\exp\big(r_{1n}(x)\big) + \exp\left(\frac{-x^2}{2}\right) r_{2n}(x),
\end{aligned}
$$

where $r_{1n}(x) = \sum_{j=1}^{n} l_{2j}(h)$, and $r_{2n}(x) = \exp(r_{1n}(x))\sum_{j=1}^{3} I_j(h)$ satisfying

$$
|r_{1n}(x)| \le \sum_{j=1}^{n} |l_{2j}(h)| \le 14x^3 L_{3n},
$$

$$
|r_{2n}(x)| \le \exp\big(|r_{1n}(x)|\big)\sum_{j=1}^{3} |I_j(h)| \le A(1+x^2)L_{3n}\exp\big(14x^3 L_{3n}\big).
$$

We thus proved (4.1).

(ii) Tracing the proof in part (i) above, it is easy to see that (4.1) also holds for $n \ge 1$ and $x \ge 1$ satisfying $x^3 L_{3n} \le 1/125$. Therefore, the proof of (4.2) follows from (4.1) by using the inequalities $e^t \le 1 + te^t$ for $t > 0$ and (1.3). □

LEMMA 4.2. *Let $\varepsilon_1, \dots, \varepsilon_n$ be i.i.d. Rademacher random variables, that is, $P(\varepsilon_j = \pm 1) = 1/2$. Then for any $x \ge 1$ and any sequence $a_1, \dots, a_n$ satisfying $|a_j| \le B_n/(6x)$ and $\sum_{j=1}^{n} a_j^2 > \frac{4}{9}B_n^2$, we have*

$$
\begin{aligned}
&P\left(\sum_{j=1}^{n} a_j\varepsilon_j > x\left(\sum_{j=1}^{n} a_j^2\right)^{1/2}\right) \qquad (4.24)\\
&\qquad \le (1-\Phi(x))\big[1 + Ax(1+x^2)L_{3n}^{*}\exp(2x^3 L_{3n}^{*})\big],
\end{aligned}
$$

where $L_{3n}^{*} = B_n^{-3}\sum_{j=1}^{n} |a_j|^3$.

PROOF. The proof of the lemma follows very similar lines to those of Lemma 4.1, so we shall only give an outline here. Let

$$
\eta_j^{*} = a_j\varepsilon_j, \qquad B_n^{*2} = \sum_{j=1}^{n} a_j^2, \qquad h^{*} = x/B_n^{*}.
$$

Then we can rewrite the left-hand side of (4.24) as

$$
P\left(\sum_{j=1}^{n} a_j\varepsilon_j > x\left(\sum_{j=1}^{n} a_j^2\right)^{1/2}\right) = P\left(\sum_{j=1}^{n} \eta_j^{*} > xB_n^{*}\right).
$$

Now under the assumptions of the lemma, it is easy to see that

$$E\eta_j^* = 0, \qquad E\eta_j^{*2} = a_j^2, \qquad E|\eta_j^*|^3 = |a_j|^3, \\ |h^*\eta_j^*| = |h^* a_j| \le 1/4. \tag{4.25}$$

So

$$E\,\exp(h^*\eta_j^*) = 1 + \tfrac{1}{2}h^{*2}a_j^2 + l_{1j}^*(h^*) = \exp\bigl(\tfrac{1}{2}h^{*2}a_j^2 + l_{2j}^*(h^*)\bigr), \tag{4.26}$$

$$\bigl(E\,\exp(h^*\eta_j^*)\bigr)^{-1} = 1 - \frac{h^{*2}a_j^2}{2} + l_{3j}^*(h^*) \le \frac{17}{16}, \tag{4.27}$$

where

$$\bigl|l_{1j}^*(h^*)\bigr| \le \tfrac{1}{6}E\bigl(\bigl|h^*\eta_j^*\bigr|^3\exp(|h^*\eta_j^*|)\bigr) \le \tfrac{1}{4}h^{*3}|a_j|^3 \quad (< 1/256), \tag{4.28}$$

$$\bigl|l_{2j}^*(h^*)\bigr| \le 2\bigl|l_{1j}^*(h^*)\bigr|, \tag{4.29}$$

$$\bigl|l_{3j}^*(h^*)\bigr| \le 2\bigl|l_{2j}^*(h^*)\bigr|. \tag{4.30}$$

To use the conjugate method, let $\xi_1^*, \dots, \xi_n^*$ be independent random variables with ξ_j^* having distribution function $V_j^*(u)$ defined by $V_j^*(u) = E(\exp(h\eta_j^*) I(\eta_j^* \le u))/E\exp(h\eta_j^*)$ for $j = 1, \dots, n$. Then in view of (4.25)–(4.29), we get that

$$\bigl|E\bigl(\eta_j^*\exp(h^*\eta_j^*)\bigr) - h^*a_j^2\bigr| \le h^{*2}|a_j|^3 \le |h^*|a_j^2/4, \\ \bigl|E\bigl(\eta_j^{*2}\exp(h^*\eta_j^*)\bigr) - a_j^2\bigr| \le 2h^*|a_j|^3 \le a_j^2/2,$$

which implies that (recall $|h^*a_j| \le 1/4$)

$$\begin{aligned} |E\xi_j^*| &= \bigl|E\bigl(\eta_j^*\exp(h^*\eta_j^*)\bigr)\bigr|/\bigl|E\bigl(\exp(h^*\eta_j^*)\bigr)\bigr| \le \frac{85}{64}|h^*|a_j^2 \le |a_j|/3, \\ \bigl|\operatorname{var}(\xi_j^*) - a_j^2\bigr| &\le \bigl|\bigl(E\exp(h^*\eta_j^*)\bigr)^{-1} - 1\bigr|\bigl|E\bigl(\eta_j^{*2}\exp(h^*\eta_j^*)\bigr)\bigr| \\ &\quad + \bigl|E\bigl(\eta_j^{*2}\exp(h^*\eta_j^*)\bigr) - a_j^2\bigr| + E\xi_j^{*2} \\ &\le \frac{9}{16}h^{*2}a_j^4 + \frac{1}{2}a_j^2 + \frac{1}{9}a_j^2 \le \frac{3a_j^2}{4}. \end{aligned}$$

Tracing the proof of Lemma 4.1, we can get [cf. (4.23)]

$$P\biggl(\sum_{j=1}^n \eta_j^* > xB_n^*\biggr) = (1 - \Phi(x))\exp(r_{1n}^*(x)) + \exp\biggl(\frac{-x^2}{2}\biggr)r_{2n}^*(x),$$

where $|r_{1n}^*(x)| \le \sum_{j=1}^n |l_{2j}^*(h)| \le 2x^3L_{3n}^*$ and $|r_{2n}^*(x)| \le A(1 + x^2)L_{3n}^* \exp(2x^3L_{3n}^*)$ by noting $B_n^* \ge \frac{2}{3}B_n$. Then (4.24) follows easily from (1.3). The proof is thus complete. □

LEMMA 4.3. (i) *Let $X_1, \dots, X_n$ be independent symmetric random variables. Then for any $x \geq 0$ and $n \geq 1$, we have*

$$P(S_n > xV_n) \leq \exp\left(\frac{-x^2}{2}\right).$$

(ii) *Let $X_1, \dots, X_n$ be i.i.d. random variables with finite variance. Then for arbitrary $0 < \varepsilon_1 < 1/2$, there exist $0 < \eta < 1$, $x_0 > 1$ and n_0 such that for any $n \geq n_0$ and $x_0 < x \leq \eta\sqrt{n}$,*

$$P(S_n > xV_n) \leq \exp\left(-\frac{(1-\varepsilon_1)x^2}{2}\right). \tag{4.31}$$

PROOF. (i) We assume without loss of generality that $X_1, \dots, X_n$ are defined on a probability space $(\Omega, \mathscr{F}, P)$ which also supports a sequence of independent Rademacher random variables $\varepsilon_1, \dots, \varepsilon_n$ independent of $X_1, \dots, X_n$ In view of the symmetry of X_j and independence of X_j and ε_j, we have that

$$\begin{aligned} P(S_n > xV_n) &= P\left(\sum_{j=1}^n X_j\varepsilon_j > xV_n\right) \\ &= \int \cdots \int P\left(\sum_{j=1}^n x_j\varepsilon_j > x\left(\sum_{j=1}^n x_j^2\right)^{1/2}\right) dF_1(x_1)\cdots dF_n(x_n), \end{aligned}$$

where $F_j(x_j)$ is the distribution function of X_j. Now by applying Lemma 2.1 in Griffin and Kuelbs (1991), we get the desired result.

(ii) For inequality (4.31), see Remark 4.1 in Shao (1997).

LEMMA 4.4. *Let $X_1, \dots, X_n$ be independent random variables. Then for any $x \geq 1$, $y \geq 0$ and $1 \leq k \leq n$, we have*

$$\begin{aligned} &P\big(S_n > xV_n, |X_k| > y\big) \\ &\quad \leq P(|X_k| > y) P\left(\sum_{\substack{j=1\\ j\neq k}}^n X_j > (x^2-1)^{1/2}\left(\sum_{\substack{j=1\\ j\neq k}}^n X_j^2\right)^{1/2}\right). \end{aligned}$$

PROOF. From the following well-known fact, for any positive numbers $s > 0, t > 0$,

$$st = \inf_{b>0} \frac{1}{2b}(s^2 + t^2b^2),$$

we get that

$$
\begin{aligned}
&P(S_n > xV_n, |X_k| > y) \\
(4.32)\qquad &= P\biggl(S_n > \inf_{b>0} \frac{x}{2\sqrt{n}b}(V_n^2 + nb^2),\ |X_k| > y\biggr) \\
&= \biggl(\sup_{b>0} \sum_{j=1}^{n} \biggl[bX_j - \frac{x}{2\sqrt{n}}(X_j^2 + b^2)\biggr] > 0,\ |X_k| > y\biggr).
\end{aligned}
$$

In view of

$$
\begin{aligned}
bX_k - \frac{x}{2\sqrt{n}}(X_k^2 + b^2) &= -\frac{x}{2\sqrt{n}}\biggl(X_k - \frac{\sqrt{n}b}{x}\biggr)^2 - \frac{xb^2}{2\sqrt{n}} + \frac{\sqrt{n}b^2}{2x} \\
&\le \frac{b^2}{2}\biggl(\frac{\sqrt{n}}{x} - \frac{x}{\sqrt{n}}\biggr),
\end{aligned}
$$

it follows from (4.32) that

$$
\begin{aligned}
&P(S_n > xV_n, |X_k| > y) \\
&\quad \le P\biggl(\sup_{b>0}\biggl(\sum_{\substack{j=1\\ j\ne k}}^{n}\biggl[bX_j - \frac{x}{2\sqrt{n}}(X_j^2 + b^2)\biggr] + \frac{b^2}{2}\biggl(\frac{\sqrt{n}}{x} - \frac{x}{\sqrt{n}}\biggr)\biggr) \\
&\qquad\qquad > 0,\ |X_k| > y\biggr) \\
(4.33)\qquad &\quad = P\biggl(\sum_{\substack{j=1\\ j\ne k}}^{n} X_j > \inf_{b>0} \frac{x}{2b\sqrt{n}}\biggl[\sum_{\substack{j=1\\ j\ne k}}^{n} X_j^2 + n\biggl(1 - \frac{1}{x^2}\biggr)b^2\biggr],\ |X_k| > y\biggr) \\
&\quad = P\biggl(\sum_{\substack{j=1\\ j\ne k}}^{n} X_j > (x^2-1)^{1/2}\biggl(\sum_{\substack{j=1\\ j\ne k}}^{n} X_j^2\biggr)^{1/2},\ |X_k| > y\biggr) \\
&\quad = P(|X_k| > y)P\biggl(\sum_{\substack{j=1\\ j\ne k}}^{n} X_j > (x^2-1)^{1/2}\biggl(\sum_{\substack{j=1\\ j\ne k}}^{n} X_j^2\biggr)^{1/2}\biggr).
\end{aligned}
$$

The proof is thus complete. □

5. Proof of main results. In this section, we shall prove the two theorems in Section 2. Without loss of generality we assume that $x > 0$. The case for $x < 0$ can be obtained easily by replacing X_j by $-X_j$ in the proofs. As for the case $x = 0$, the assertions follow from the classical results.

PROOF OF THEOREM 2.1. (i) For $0 < x \le 1$, then (2.1) was shown by Bentkus, Bloznelis and Götze (1996). Therefore, it suffices to show that (2.1) also holds for x in the range

$$1 \le x \le (5L_{3n}^{1/3})^{-1}.$$

From the elementary inequality $2B_nV_n \le V_n^2 + B_n^2$ and Lemma 4.1(ii), we have that

$$\begin{aligned} P(S_n > xV_n) &\ge P\big(2B_nS_n > x(V_n^2 + B_n^2)\big) \\ &\ge \big(1-\Phi(x)\big)\big(1 - Ax(x^2+1)L_{3n}\big), \end{aligned}$$

which, together with (1.3), implies that

$$P\big(S_n \le xV_n\big) - \Phi(x) \le \frac{A}{\sqrt{2\pi}}(x^2+1)L_{3n} \exp\left(\frac{-x^2}{2}\right). \tag{5.1}$$

Therefore, to prove (2.1), it suffices to show that, for $1 \le x \le (5L_{3n}^{1/3})^{-1}$,

$$\begin{aligned} &P(S_n > xV_n) \\ &\quad \le 1-\Phi(x) + A\left((1+x^2)L_{3n} + \sum_{j=1}^n P(|X_j| > B_n/(6x))\right)\exp\left(\frac{-x^2}{2}\right). \end{aligned} \tag{5.2}$$

To show (5.2), let us define

$$Y_j = X_j I\big(|X_j| \le B_n/(6x)\big), \qquad S_n^* = \sum_{j=1}^n Y_j, \qquad V_n^{*2} = \sum_{j=1}^n Y_j^2.$$

In view of Lemma 4.3(i) and Lemma 4.4, we have that for any $x \ge 1$,

$$\begin{aligned} P(S_n > xV_n) - P(S_n^* > xV_n^*) &\le \sum_{k=1}^n P\big(S_n > xV_n, |X_k| > B_n/(6x)\big) \\ &\le \sum_{k=1}^n P\big(|X_k| > B_n/(6x)\big)\exp\left(\frac{-(x^2-1)}{2}\right), \\ &\le e\sum_{k=1}^n P\big(|X_k| > B_n/(6x)\big)\exp\left(\frac{-x^2}{2}\right). \end{aligned} \tag{5.3}$$

We shall next place a bound for the term $P(S_n^* > xV_n^*)$ above. As in the proof of Lemma 4.3, we assume that $\{Y_j, j \ge 1\}$ are defined on a probability space $(\Omega, \mathscr{F}, P)$ which also supports a sequence of independent Rademacher random variables $\{\varepsilon_j, j \ge 1\}$ independent of the initial sequence $\{Y_j, j \ge 1\}$. In view of symmetry of X_j, we have that

$$\begin{aligned} P(S_n^* > xV_n^*) &= P\left(\sum_{j=1}^n Y_j\varepsilon_j > xV_n^*\right) \\ &\le P\left(\sum_{j=1}^n Y_j\varepsilon_j > xV_n^*, V_n^{*2} > \frac{4}{9}B_n^2\right) \\ &\quad + P\left(V_n^{*2} \le \frac{4}{9}B_n^2\right). \end{aligned} \tag{5.4}$$

Let us investigate the first term in (5.4). Denote $F_j(x)$ to be the distribution function of X_j for $j \ge 1$. Using the inequality $e^t \le 1 + te^t$ for any $t > 0$, we get for $1 \le i \le n$,

$$\begin{aligned}
\prod_{\substack{j=1\\ j\ne i}}^{n} E\bigl(\exp(2x^3 B_n^{-3}|Y_j|^3)\bigr) &\le \prod_{\substack{j=1\\ j\ne i}}^{n} \bigl(1 + 2x^3 B_n^{-3} E|Y_j|^3 e^{2/125}\bigr)\\
&\le \prod_{\substack{j=1\\ j\ne i}}^{n} \exp\bigl(4x^3 B_n^{-3} E|Y_j|^3\bigr)\\
&\le \exp\bigl(4x^3 L_{3n}\bigr)\\
&\le 2.
\end{aligned}$$

From this and Lemma 4.2, it follows that

$$\begin{aligned}
&P\left(\sum_{j=1}^{n} Y_j\varepsilon_j > xV_n^*,\ V_n^{*2} > \frac{4}{9}B_n^2\right)\\
&\quad = \int\cdots\int_{\substack{\sum_{j=1}^n y_j^2 > \frac{4}{9}B_n^2\\ |y_j|\le B_n/(6x),\, j=1,\dots,n}} P\left(\sum_{j=1}^{n} y_j\varepsilon_j > x\left(\sum_{j=1}^{n} y_j^2\right)^{1/2}\right)\\
&\qquad\qquad \times dF_1(y_1)\cdots dF_n(y_n)\\
&\quad \le (1-\Phi(x))\int\cdots\int_{\substack{|y_j|\le B_n/(6x),\\ j=1,\dots,n}}\\
&\qquad\qquad \times\left[1 + Ax(1+x^2)B_n^{-3}\sum_{i=1}^{n}|y_i|^3 \exp\left(2x^3B_n^{-3}\sum_{j=1}^{n}|y_j|^3\right)\right]\\
&\qquad\qquad \times dF_1(y_1)\cdots dF_n(y_n)\\
&\quad \le (1-\Phi(x))\Bigg[1 + Ax(1+x^2)B_n^{-3}\\
&\qquad\qquad \times \sum_{i=1}^{n} E\left(|Y_i|^3 \exp\left(2x^3B_n^{-3}\sum_{j=1}^{n}|Y_j|^3\right)\right)\Bigg]\\
&\quad \le (1-\Phi(x))\Bigg[1 + Ax(1+x^2)B_n^{-3}\\
&\qquad\qquad \times \sum_{i=1}^{n}\left\{E|Y_i|^3\left(\prod_{\substack{j=1\\ j\ne i}}^{n} E\exp\left(2x^3B_n^{-3}|Y_j|^3\right)\right)\right\}\Bigg]\\
&\quad \le 1-\Phi(x) + A(1+x^2)L_{3n}\exp\left(\frac{-x^2}{2}\right).
\end{aligned} \tag{5.5}$$

We now look at the second term in (5.4). Note that

$$\sum_{j=1}^{n} E\left(X_j^2 I\left\{|X_j| > \frac{B_n}{(6x)}\right\}\right) \le \frac{6x}{B_n} \sum_{j=1}^{n} E|X_j|^3 \le \frac{6}{125} B_n^2,$$

$$\sum_{j=1}^{n} EY_j^2 = B_n^2 - \sum_{j=1}^{n} E\big(X_j^2 I\{|X_j| > B_n/(6x)\}\big).$$

Then for any $t > 0$, we have

$$\begin{aligned}
P\left(V_n^{*2} \le \frac{4}{9} B_n^2\right) &= P\left(\sum_{j=1}^{n} (EY_j^2 - Y_j^2) > \frac{5}{9} B_n^2 - \sum_{j=1}^{n} E\big(X_j^2 I(|X_j| > B_n/(6x))\big)\right) \\
&\le P\left(\sum_{j=1}^{n} (EY_j^2 - Y_j^2) > \frac{1}{2} B_n^2\right) \\
&\le e^{-t/2} \prod_{j=1}^{n} E \exp\big(tB_n^{-2}(EY_j^2 - Y_j^2)\big) \\
&\le e^{-t/2} \prod_{j=1}^{n} \left(1 + \frac{1}{2} t^2 B_n^{-4} \operatorname{var}(Y_j^2) \exp\left(\frac{tx^{-2}}{36}\right)\right) \\
&\le e^{-t/2} \prod_{j=1}^{n} \exp\left(\frac{t^2 E|X_j|^3}{6xB_n^3} \exp\left(\frac{tx^{-2}}{36}\right)\right) \\
&\le e^{-t/2} \exp\left(\frac{t^2}{6x} L_{3n} \exp\left(\frac{tx^{-2}}{36}\right)\right),
\end{aligned}$$

where in the second last inequality we have used the inequalities that $1 + |x| \le e^{|x|}$ and $\operatorname{var}(Y_j^2) \le EY_j^4 \le E|X_j|^3 B_n/(6x)$. In particular, if we choose $t = 4x^2\big(1 + x^{-2} \log L_{3n}^{-1/2}\big)$, then we have

$$P\left(V_n^{*2} \le \frac{4}{9} B_n^2\right) \le AL_{3n} \exp\left(\frac{-x^2}{2}\right). \tag{5.6}$$

Hence, the inequality (5.2) follows from (5.3)–(5.6). The proof of Theorem 2.1(i) is thus complete.

(ii) The proof of the second part of Theorem 2.1, (2.2), follows from (1.3) and Lemma 4.3. □

PROOF OF THEOREM 2.2. For an arbitrary $0 < \varepsilon < 1/2$, by applying Lemma 4.3(ii) with $\varepsilon_1 = \varepsilon/2$, there exist $0 < \eta_0 < 1$, $x_0 > 1$ and n_0 (only depending on ε) such that for any $x_0 < x \le \eta_0\sqrt{n}$ and $n \ge n_0$,

$$P(S_n > xV_n) \le \exp\bigl(-(1-\varepsilon/2)x^2/2\bigr). \tag{5.7}$$

Since n_0 only depends on ε, (5.7) implies that there exists $A(\varepsilon)$ (only depending on ε) such that for any $x_0 < x \le \eta_0\sqrt{n}$,

$$P(S_n > xV_n) \le A\exp\bigl(-(1-\varepsilon/2)x^2/2\bigr). \tag{5.8}$$

In view of (5.8) and (1.3), it follows that for any x such that $x_0 < x \le \eta_0\sqrt{n}$ and $x \ge (5\beta_3^{1/3})^{-1}n^{1/6}$,

$$\begin{aligned}\delta_n(x) &= \bigl|P(S_n > xV_n) - (1-\Phi(x))\bigr| \\ &\le A(\varepsilon)x^3 n^{-1/2}\beta_3 \exp\bigl(-(1-\varepsilon/2)x^2/2\bigr) \\ &\le A(\varepsilon)n^{-1/2}\beta_3 \exp\biggl(\frac{-(1-\varepsilon)x^2}{2}\biggr).\end{aligned} \tag{5.9}$$

Next we prove for any $x_0 < x \le \eta_0\sqrt{n}$ and $x \le (5\beta_3^{1/3})^{-1}n^{1/6}$,

$$\begin{aligned}\delta_n(x) \le{}& A\bigl(n^{-1/2}\beta_3 + \sigma^{-2}E\bigl(X_1^2 I(|X_1| > \sigma n^{3/8}(2|x|)^{-1})\bigr)\bigr) \\ &\times \exp\biggl(\frac{-(1-\varepsilon)x^2}{2}\biggr).\end{aligned} \tag{5.10}$$

We note that (5.1) does not depend on the symmetry of X_j. Thus in order to prove (5.10), it remains to show that

$$\begin{aligned}&P(S_n > xV_n) \\ &\quad\le 1-\Phi(x) + A\bigl(n^{-1/2}\beta_3 + \sigma^{-2}E\bigl(X_1^2 I(|X_1| > \sigma n^{3/8}(2|x|)^{-1})\bigr)\bigr) \\ &\qquad \times \exp\left(\tfrac{-(1-\varepsilon)x^2}{2}\right).\end{aligned} \tag{5.11}$$

Let

$$Z_j = \sigma^{-1}X_j I\bigl(|X_j| \le \sigma\sqrt{n}(2x)^{-1}\bigr), \qquad Z_j^* = \sigma^{-1}X_j I\bigl(|X_j| \le \sigma n^{3/8}(2|x|)^{-1}\bigr),$$

$$S_{1n} = \sum_{j=1}^n Z_j, \qquad V_{1n}^2 = \sum_{j=1}^n Z_j^2, \qquad V_{2n}^2 = \sum_{j=1}^n Z_j^{*2},$$

$$\eta_{1j} = Z_j - \tfrac{1}{2}xn^{-1/2}(Z_j^{*2} - EZ_1^{*2}), \qquad g_n(x) = n^{-1/2} + \tfrac{1}{2}\bigl(1 - EZ_1^{*2}\bigr).$$

From the inequality $(1+y)^{1/2} \geq 1 + y/2 - y^2$ for any $y \geq -1$ and noting $V_{1n} \geq V_{2n}$, we have that

$$\begin{aligned}
P(S_{1n} > xV_{1n}) &\leq P(S_{1n} > xV_{2n}) \\
&= P\big(S_{1n} > x\sqrt{n}\big(1 + (n^{-1}V_{2n}^2 - 1)\big)^{1/2}\big) \\
&\leq P\Big(S_{1n} > x\sqrt{n}\big(1 + \tfrac{1}{2}\big(n^{-1}V_{2n}^2 - 1\big) - \big(n^{-1}V_{2n}^2 - 1\big)^2\big)\Big) \\
&= P\bigg(\sum_{j=1}^{n} \eta_{1j} > x\sqrt{n}\big(1 + \tfrac{1}{2}(EZ_1^{*2} - 1) - \big(n^{-1}V_{2n}^2 - 1\big)^2\big)\bigg) \qquad (5.12) \\
&\leq P\bigg(\sum_{j=1}^{n} \eta_{1j} > x\sqrt{n}(1 - g_n(x))\bigg) \\
&\quad + P\big(\big|n^{-1}V_{2n}^2 - 1\big| > n^{-1/4}\big).
\end{aligned}$$

Similarly to the proof of Lemma 4.1, we can show that (noting $|\eta_{1j}| \leq \sqrt{n}/x$)

$$P\bigg(\sum_{j=1}^{n} \eta_{1j} > x\sqrt{n}(1 - g_n(x))\bigg) \leq 1 - \Phi(x) + r_{5n}(x)\exp\bigg(\frac{-x^2}{2}\bigg),$$

where

$$\begin{aligned}
|r_{5n}(x)| &\leq Ax^3 n^{-1/2}\beta_3 + Ax|g_n(x)| \\
&\leq Ax^3\big(n^{-1/2}\beta_3 + \sigma^{-2}E\big(X_1^2 I(|X_1| > \sigma n^{3/8}(2|x|)^{-1})\big)\big).
\end{aligned}$$

Similarly to the proof of (5.6), we can give a bound to the second term in (5.12),

$$P\big(\big|n^{-1}V_{2n}^2 - 1\big| \geq n^{-1/4}\big) \leq An^{-1/2}\beta_3 \exp\bigg(\frac{-x^2}{2}\bigg). \qquad (5.13)$$

From this, we can show that, similarly to the proof of (5.3),

$$P\big(S_n > xV_n\big) - P\big(S_{1n} > xV_{1n}\big) \leq An^{-1/2}\beta_3 \exp\bigg(\frac{-(1-\varepsilon)x^2}{2}\bigg). \qquad (5.14)$$

Then, inequality (5.11) follows from (5.12)–(5.14).

Finally, for any $|x| \leq x_0$, we wish to show that

$$\begin{aligned}
\delta_n(x) &\leq A(\varepsilon)\big(n^{-1/2}\beta_3 + \sigma^{-2}E\big(X_1^2 I(|X_1| > \sigma n^{3/8}(2|x|)^{-1})\big)\big) \\
&\quad \times \exp\bigg(\frac{-(1-\varepsilon)x^2}{2}\bigg). \qquad (5.15)
\end{aligned}$$

However, (5.15) follows from Theorem 1.2 of Bentkus and Götze (1996) since x_0 only depends on ε.

Combining (5.10), (5.10) and (5.15), we have completed the proof of Theorem 2.2. $\square$

Acknowledgments. The authors thank the referees and an Associate Editor for their valuable comments, which led to a much improved version of the paper. They also wish to thank Qi-Man Shao for providing an unpublished manuscript and useful comments about the paper.

REFERENCES

BENTKUS, V., BLOZNELIS, M. and GÖTZE, F. (1996). A Berry–Esseen bound for Student's statistic in the non-i.i.d. case. *J. Theoret. Probab.* **9** 765–796.

BENTKUS, V. and GÖTZE, F. (1996). The Berry–Esseen bound for Student's statistic. *Ann. Probab.* **24** 491–501.

EFRON, B. (1969). Student's t-test under symmetry conditions. *J. Amer. Statist. Assoc.* **64** 1278–1302.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, 2nd ed. Wiley, New York.

GINÉ, E., GÖTZE, F. and MASON, D. M. (1997). When is the Student t-statistic asymptotically standard normal? *Ann. Probab.* **25** 1514–1531.

GRIFFIN, P. S. and KUELBS, J. D. (1989). Self-normalized laws of the iterated logarithm. *Ann. Probab.* **17** 1571–1601.

GRIFFIN, P. S. and KUELBS, J. D. (1991). Some extensions of the laws of the iterated logarithm via self-normalized. *Ann. Probab.* **19** 380–395.

LINNIK, YU V. (1962). Limit theorems for sums of independent random variables, taking account of large deviations. *Theory Probab. Appl.* **N2** 115–129.

LOGAN, B. F., MALLOWS, C. L., RICE, S. O. and SHEPP, L. A. (1973). Limit distributions of self-normalized sums. *Ann. Probab.* **1** 788–809.

MICHEL, R. (1976). Nonuniform central limit bounds with applications to probabilities of deviations. *Ann. Probab.* **4** 102–106.

PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, Berlin.

SHAO, Q. M. (1997). Self-normalized large deviations. *Ann. Probab.* **25** 285–328.

SHAO, Q. M. (1998). Cramér type large deviation for Student's t statistic. Unpublished manuscript.

VANDEMAELE, M. and VERAVERBEKE, N. (1985). Cramér type large deviations for studentized U-Statistics. *Metrika* **32** 165–180.

SCHOOL OF MATHEMATICS
AND APPLIED STATISTICS
UNIVERSITY OF WOLLONGONG
NSW 2522
AUSTRALIA
E-MAIL: qiying@uow.edu.au

DEPARTMENT OF MATHEMATICS
HONG KONG UNIVERSITY
OF SCIENCE AND TECHNOLOGY
CLEAR WATER BAY, KOWLOON
HONG KONG
E-MAIL: majing@uxmail.ust.hk