The Annals of Probability
1988, Vol. 16, No. 4, 1699–1720

A NONUNIFORM BOUND ON THE RATE OF CONVERGENCE IN THE MARTINGALE CENTRAL LIMIT THEOREM

BY ERICH HAEUSLER AND KONRAD JOOS

University of Munich

The main result of the present paper is a sharp nonuniform bound on the rate of convergence to normality in the central limit theorem for martingales having finite moments of order $2 + 2\delta$ for some $0 < \delta < \infty$. A nonuniform bound on the rate for convergence to mixtures of normal distributions is obtained as a consequence.

1. Introduction and results. Uniform bounds on the rate of convergence in martingale central limit theorems have been obtained by many authors, but nonuniform bounds seem to be available only through the work of Hall and Heyde (1980, 1981) and Bose (1986a, 1986b). The results of Hall and Heyde provide rates under basic conditions of the martingale central limit theorem, as the conditional form of Lindeberg's or Liapounov's condition, for instance, presupposing only that the convergence in probability normally appearing in these conditions is strengthened to convergence in an L_p-norm. Bose works under more stringent assumptions about the conditional moments of the underlying martingale difference array, but he establishes also moderate deviation results, which are of added interest. In the present paper the general setup as studied by Hall and Heyde (1980) in their Theorem 3.9 is considered, and we will concentrate on optimal nonuniform bounds without examining the problem of moderate deviations.

Throughout this paper let the real-valued random variables $X_1, \dots, X_n$ form a square integrable martingale difference sequence (mds for short) w.r.t. the σ-fields $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots \subset \mathscr{F}_n$, i.e., suppose that X_i is measurable w.r.t. $\mathscr{F}_i$ with $E(X_i^2) < \infty$ and $E(X_i | \mathscr{F}_{i-1}) = 0$ a.s. for $i = 1, \dots, n$. Set $S_n = \sum_{i=1}^n X_i$, and for $0 < \delta < \infty$ set

$$L_{n,2\delta} = \sum_{i=1}^{n} E\left(|X_i|^{2+2\delta}\right)$$

and

$$N_{n,2\delta} = E\left(\left|\sum_{i=1}^{n} E\left(X_i^2 | \mathscr{F}_{i-1}\right) - 1\right|^{1+\delta}\right).$$

Then one has the following uniform bound on the distance between the distribution function of S_n and the distribution function Φ of the standard normal distribution. For each $0 < \delta < \infty$ there exists a finite constant C_δ depending only

Received January 1987; revised August 1987.
AMS 1980 *subject classifications*. Primary 60F05, 60G42.
Key words and phrases. Martingale central limit theorem, rates of convergence, nonuniform bounds, mixtures of normal distributions.

on δ such that for each mds $X_1, \dots, X_n$,

$$D_n \equiv \sup_{x \in R} |P(S_n \le x) - \Phi(x)| \le C_\delta (L_{n,2\delta} + N_{n,2\delta})^{1/(3+2\delta)}. \tag{1.1}$$

Of course, this inequality is nontrivial only if $L_{n,2\delta}$ and $N_{n,2\delta}$ are finite, and it provides a rate of weak convergence of S_n to normality if $L_{n,2\delta} + N_{n,2\delta} \to 0$ as $n \to \infty$. Inequality (1.1) has been shown for $0 < \delta \le 1$ by Heyde and Brown (1970) by an application of the martingale version of the Skorohod embedding scheme and for $0 < \delta \le 1/2$ by Erickson, Quine and Weber (1979) who used the classical characteristic function technique. In Haeusler (1988) a version of Bolthausen's (1982) iterative method is developed to establish (1.1) for every $\delta > 0$, and an example is constructed demonstrating that it is asymptotically exact for every $\delta > 0$ if $L_{n,2\delta} + N_{n,2\delta} \to 0$ as $n \to \infty$. The main result of the present paper is the following nonuniform version of (1.1).

THEOREM 1. *For any $\delta > 0$ there exists a finite constant C_δ depending only on δ such that for each mds $X_1, \dots, X_n$ whenever $L_{n,2\delta} + N_{n,2\delta} \le 1$,*

$$|P(S_n \le x) - \Phi(x)| \le C_\delta (1 + |x|^{2+2\delta})^{-1} (L_{n,2\delta} + N_{n,2\delta})^{1/(3+2\delta)}, \tag{1.2}$$

for all $x \in R$.

Since (1.1) is asymptotically exact w.r.t. $L_{n,2\delta} + N_{n,2\delta}$, the same holds true for (1.2). From the theory of nonuniform rates of convergence in the case of independent $X_1, \dots, X_n$ it is well known that the $(1 + |x|^{2+2\delta})^{-1}$-factor is sharp for large $|x|$ if $X_1, \dots, X_n$ have finite moments of order $2 + 2\delta$; see, e.g., Remark 1 in Michel (1976).

Theorem 1 improves the x-factor inequality (3.75) of Theorem 3.9 of Hall and Heyde (1980), whereas in terms of $L_{n,2\delta} + N_{n,2\delta}$ the two bounds are the same. Observe, however, that Hall and Heyde assume $0 < \delta \le 1$, whereas (1.2) is valid for all $0 < \delta < \infty$. Hall and Heyde (1980) replace $L_{n,2\delta} + N_{n,2\delta}$ in inequality (3.75) of their Theorem 3.9 by $L_{n,2\delta} + M_{n,2\delta}$ with $M_{n,2\delta} = E(|\sum_{i=1}^n X_i^2 - 1|^{1+\delta})$ in order to obtain their estimate (3.76). For $0 < \delta \le 1$ such a replacement is also possible in (1.2) as a consequence of Burkholder's square function inequality, cf. (3.87) in Hall and Heyde (1980), but for $\delta > 1$ this argument does no longer work, and it is presently unclear whether this replacement is possible in (1.2) for $\delta > 1$, too. For the same reason Theorem 1 only improves the x-factor and for $\delta > 1$ also the L_n-term in inequality (3.91) of Theorem 3.10 in Hall and Heyde (1980), but not inequality (3.92) of the same theorem. The bound in (1.2) contains the same optimal x-factor as the bound in Theorem 3 of Bose (1986b) for all $0 < \delta < \infty$. For $0 < \delta \le 1/2$ Theorem 1 also gives Bose's n-factor under his conditions, and for $\delta > 1/2$ it yields a better one (notice that Bose writes δ where we use 2δ).

The proof of Theorem 1 will be given in Section 3, after some technical lemmas have been established in Section 2. Our method of proof is the "nonuniform" modification of the proof of (1.1) given in Haeusler (1988) and is therefore entirely different from the techniques of Hall and Heyde (1980, 1981), who employ the martingale version of the Skorohod embedding scheme, and of Bose (1986a, 1986b), who uses a conditional version of the classical exponential centering after truncation approach.

A nonuniform bound like (1.2) immediately entails bounds on the rate of convergence of moments and L_p-norms in the central limit theorem, whereas such results do not follow from uniform bounds like (1.1). For general information on the convergence of moments in the martingale central limit theorem the reader is referred to Hall (1978), and for results on rates in L_p-norms to Nakata (1976).

In Haeusler (1988) it was shown that a discretization procedure can be used to derive from (1.1) the corresponding uniform bound on the rate of convergence in the central limit theorem for locally square integrable martingales with continuous time. From the formulation and proof of Theorem 2 of that paper it is clear how by the same argument the nonuniform estimate (1.2) carries over to the continuous time case so that the details can be skipped here. Instead, we want to show how one can obtain from Theorem 1 a nonuniform bound for the rate of weak convergence of discrete time martingales to mixtures of normal distributions. For this, we have to introduce some notation. Let η be a random variable with $0 < \eta < \infty$, and for $m = 0, 1, \dots, n$ set

$$Q_{n,2\delta}(m) = E\left(\left|\sum_{i=1}^{m} X_i^2\right|^{1+\delta}\right) + \sum_{i=m+1}^{n} E\left(|X_i|^{2+2\delta}\right)$$
$$+E\left(\left|\sum_{i=m+1}^{n} E\left(X_i^2|\mathscr{G}_{i-1}\right) - \eta_m^2\right|^{1+\delta}\right) + E\left(|\eta^2 - \eta_m^2|^{(3+2\delta)/2}\right)$$

and

$$\tilde{Q}_{n,2\delta}(m) = E\left(V_m^{2+2\delta}\right) + L_{n,2\delta} + E\left(|V_n^2 - \eta^2|^{1+\delta}\right) + E\left(|\eta^2 - \eta_m^2|^{(3+2\delta)/2}\right),$$

where $\mathscr{G}_i = \sigma(X_1, \dots, X_i)$, with $\mathscr{G}_0$ being the trivial σ-field, $\eta_m^2 = E(\eta^2|\mathscr{G}_m)$ and $V_l^2 = \sum_{i=1}^{l} E(X_i^2|\mathscr{G}_{i-1})$ for $l = m, n$. Then we have

THEOREM 2. *Let $0 < \delta < \infty$ be fixed, and assume that $E(\eta^{3+2\delta}) < \infty$ and $E(\eta^{-3-2\delta}) < \infty$. There exists a finite constant C_δ depending only on δ such that for each mds $X_1, \dots, X_n$, all $x \in R$ and $m = 0, 1, \dots, n-1$, whenever $Q_{n,2\delta}(m) \le 1$,*

$$\left|P(S_n \le x) - E\left(\Phi(\eta^{-1}x)\right)\right| \le C_\delta C_\delta(\eta)\left(1 + |x|^{2+2\delta}\right)^{-1} Q_{n,2\delta}(m)^{1/(3+2\delta)}, \tag{1.3}$$

where

$$C_\delta(\eta) = \left(E\left(\eta^{3+2\delta}\right) + E\left(\eta^{-3-2\delta}\right)\right)^{(2+2\delta)/(3+2\delta)} \ge 2^{(2+2\delta)/(3+2\delta)}.$$

The result is also true if $Q_{n,2\delta}(m)$ is replaced by $\tilde{Q}_{n,2\delta}(m)$.

The proof of Theorem 2 will be given in Section 4. The result is similar to a part of Theorem 1 of Hall and Heyde (1981), but not directly comparable, since their L_{nm}-terms and our $Q_{n,2\delta}(m)$-terms are different. Observe that the complicated structure of these terms is due to the fact that weak convergence of martingales to mixtures of normal distributions requires either a measurability assumption on η or nested σ-fields for the underlying triangular array. Nevertheless, the bound (1.3) provides a proper rate of convergence, since under the corresponding conditions for weak convergence of martingales to mixtures of normal distributions one has $Q_{n,2\delta}(m) \to 0$ and $\tilde{Q}_{n,2\delta}(m) \to 0$ as $n \to \infty$ for an appropriate choice of $m = m(n)$; cf. Hall and Heyde's (1981) discussion concerning their L_{nm}-terms.

2. Notation and technical lemmas. Throughout Sections 2–4, the following conventions will be used to simplify the notation. The symbol C always denotes a generic finite absolute constant, whereas C_δ, C_p, $C_{\delta,p}$ and $C_{p,K}$ are always generic finite constants depending only on their indices. φ denotes the standard normal density. Equations, inequalities, etc., between random variables are always assumed to hold almost surely without explicit mention, especially when conditional expectations are involved.

For a random variable X and $0 \le x < \infty$ we set

$$D(X) = \sup\{|P(X \le y) - \Phi(y)|: y \in R\}$$

and

$$d(X,x) = \sup\{|P(X \le y) - \Phi(y)|: y \ge x\}.$$

Lemma 1. *For any $p > 0$ there exists a finite constant C_p such that for all $x > 0$ and random variables X and Y with $E(|X - Y|^p) \le 1$,*

$$|P(X \le x) - P(Y \le x)| \le C_p x^{-p} E(|X - Y|^p)^{1/(1+p)} + 2d(X, x/2).$$

Proof. For $x > 0$ and $0 < a \le x/2$ we have

$$\begin{aligned} P(X \le x) - P(Y \le x) &\le P(X \le x) - P(X \le x - a) + P(|X - Y| > a) \\ &\le \Phi(x) + d(X,x) - \Phi(x-a) \\ &\quad + d(X, x-a) + P(|X - Y| > a) \\ &\le \varphi(x-a)a + 2d(X, x/2) + a^{-p}E(|X - Y|^p) \end{aligned}$$

and a similar bound for $P(Y \le x) - P(X \le x)$ so that

$$|P(X \le x) - P(Y \le x)| \le \exp(-x^2/8)a + 2d(X, x/2) + a^{-p}E(|X - Y|^p).$$

For $a = xE(|X - Y|^p)^{1/(1+p)}/2$ we have $a \le x/2$ by assumption, and we may assume $a > 0$ since otherwise there is nothing to prove. Substituting a into the preceding inequality completes the proof of the lemma. □

The first part of the next lemma is a small extension of a part of Lemma 1 of Bolthausen (1982), and the second part is the corresponding "nonuniform" statement.

LEMMA 2. (i) *For every $p > 1$ there exists a finite constant C_p such that for all random variables X and Y,*

$$D(X) \le C_p\left(D(X+Y) + \|E(|Y|^p|X)\|_\infty^{1/p}\right).$$

(ii) *For every $p > 1$ and $0 < K < \infty$ there exists a finite constant $C_{p,K}$ such that for all $x > 0$ and random variables X and Y with $\|E(|Y|^p|X)\|_\infty \le K$,*

$$\begin{aligned} d(X,x) &\le d(X+Y, x/2) \\ &\quad + C_{p,K}x^{-p}\left(\|E(|Y|^p|X)\|_\infty^{1/p} + \|E(|Y|^p|X)\|_\infty + D(X)\right) \end{aligned} \tag{2.1}$$

and

$$\begin{aligned} d(X+Y,x) &\le d(X, x/2) \\ &\quad + C_{p,K}x^{-p}\left(\|E(|Y|^p|X)\|_\infty^{1/p} + \|E(|Y|^p|X)\|_\infty + D(X)\right). \end{aligned} \tag{2.2}$$

PROOF. We may assume $0 < \gamma = \|E(|Y|^p|X)\|_\infty < \infty$, since otherwise there is nothing to prove. By Bolthausen's (1982) arguments in the proof of his inequalities (2.1) and (2.2) with $(t-X)^{-2}$ replaced by $(t-X)^{-p}$ one obtains for every $t \in R$ and $a > 0$,

$$\begin{aligned} P(X+Y \le t) &\ge P(X \le t-a) \\ &\quad -\gamma\int_{-\infty}^{t-a}\varphi(u)(t-u)^{-p}\,du - 2\gamma D(X)a^{-p} \end{aligned} \tag{2.3}$$

and

$$\begin{aligned} P(X+Y \le t) &\le P(X \le t+a) \\ &\quad +\gamma\int_{t+a}^{\infty}\varphi(u)(u-t)^{-p}\,du + 2\gamma D(X)a^{-p}. \end{aligned} \tag{2.4}$$

Observing that the integrals in (2.3) and (2.4) are less than or equal to $a^{-p+1}/(p-1)$ and following Bolthausen's (1982) reasoning again, one gets for all $a > 0$,

$$D(X) \le D(X+Y) + a + \gamma a^{-p+1}/(p-1) + 2\gamma D(X)a^{-p}.$$

Taking $a = (4\gamma)^{1/p}$ yields the inequality in part (i) of the lemma. For the proof of part (ii) we replace t by $t+a$ in (2.3) and for $t, a > 0$ use the estimate

$$\begin{aligned} \int_{-\infty}^{t}\varphi(u)(t+a-u)^{-p}\,du &\le \exp(-t^2/8)\int_{t/2}^{t}(t+a-u)^{-p}\,du \\ &\quad +2^p t^{-p}\int_{-\infty}^{t/2}\varphi(u)\,du \\ &\le \exp(-t^2/8)a^{-p+1}/(p-1) + 2^p t^{-p} \end{aligned}$$

to deduce

$$\begin{aligned} P(X \le t) - \Phi(t) &\le P(X + Y \le t + a) - \Phi(t) \\ &\quad + \gamma \exp(-t^2/8) a^{-p+1}/(p-1) \\ &\quad + 2^p \gamma t^{-p} + 2\gamma D(X) a^{-p} \\ &\le d(X + Y, t + a) + \exp(-t^2/2) a \\ &\quad + \gamma \exp(-t^2/8) a^{-p+1}/(p-1) \\ &\quad + 2^p \gamma t^{-p} + 2\gamma D(X) a^{-p}. \end{aligned}$$

Replacing t by $t - a$ in (2.4) and using the estimate

$$\begin{aligned} \int_t^\infty \varphi(u)(u - t + a)^{-p}\, du &\le \exp(-t^2/2) \int_t^\infty (u - t + a)^{-p}\, du \\ &= \exp(-t^2/2) a^{-p+1}/(p-1), \end{aligned}$$

we obtain for $t > 0$ and $0 < a \le t/2$,

$$\begin{aligned} \Phi(t) - P(X \le t) &\le \Phi(t - a) - P(X + Y \le t - a) + \varphi(t - a) a \\ &\quad + \gamma \exp(-t^2/2) a^{-p+1}/(p-1) + 2\gamma D(X) a^{-p} \\ &\le d(X + Y, t/2) + a \exp(-t^2/8) \\ &\quad + \gamma \exp(-t^2/2) a^{-p+1}/(p-1) \\ &\quad + 2\gamma D(X) a^{-p}. \end{aligned}$$

Thus for $t > 0$ and $0 < a \le t/2$ we have

$$\begin{aligned} |P(X \le t) - \Phi(t)| &\le d(X + Y, t/2) + \exp(-t^2/8) a + 2^p \gamma t^{-p} \\ &\quad + \gamma \exp(-t^2/8) a^{-p+1}/(p-1) + 2\gamma D(X) a^{-p}. \end{aligned}$$

Taking $a = \gamma^{1/p} t/(2K^{1/p})$, we have $0 < a \le t/2$ by assumption and get

$$\begin{aligned} |P(X \le t) - \Phi(t)| &\le d(X + Y, t/2) + C_{p,K} t^{-p} (\gamma^{1/p} + \gamma + D(X)) \\ &\le d(X + Y, x/2) + C_{p,K} x^{-p} (\gamma^{1/p} + \gamma + D(X)), \end{aligned}$$

for all $t \ge x > 0$ since the right side is nonincreasing in t. This completes the proof of (2.1). Inequality (2.2) is proven by a similar argument. □

LEMMA 3. *For every $p > 1$ there exists a finite constant C_p such that for all $x \ge 2$ and $0 < \alpha \le 1$ and all random variables X,*

$$E\left(\min\left(1, |(x - X)/\alpha|^{-p}\right)\right) \le C_p\left(d(X, x/2) + x^{-p}(D(X) + \alpha)\right).$$

Proof. On account of $x \geq 2$ and $0 < \alpha \leq 1$,

$$
\begin{aligned}
&E\Big(\min\big(1, |(x-X)/\alpha|^{-p}\big)\Big) \\
&\quad \leq E\Big(|(x-X)/\alpha|^{-p} I(|x-X|/\alpha \geq x/2)\Big) + P\Big(|(x-X)/\alpha| < x/2\Big) \\
&\quad \leq p\int_{x/2}^{\infty} P(|x-X| < \alpha u) u^{-p-1}\, du + \Phi(x+\alpha x/2) \\
(2.5)\qquad &\qquad -\Phi(x-\alpha x/2) + 2d(X, x/2) \\
&\quad \leq p\int_{x/2}^{\infty} (\Phi(x+\alpha u) - \Phi(x-\alpha u)) u^{-p-1}\, du + 2^{p+1} D(X) x^{-p} \\
&\qquad + C_p x^{-p}\alpha + 2d(X, x/2).
\end{aligned}
$$

Using the definition of Φ and Fubini's theorem, we can rewrite the integral as

$$
\begin{aligned}
&p\int_{-\infty}^{x-\alpha x/2}\int_{(x-v)/\alpha}^{\infty} u^{-p-1}\, du\, \varphi(v)\, dv + p\int_{x-\alpha x/2}^{x+\alpha x/2}\int_{x/2}^{\infty} u^{-p-1}\, du\, \varphi(v)\, dv \\
&\quad + p\int_{x+\alpha x/2}^{\infty}\int_{(v-x)/\alpha}^{\infty} u^{-p-1}\, du\, \varphi(v)\, dv \\
&\quad \equiv \mathrm{I} + \mathrm{II} + \mathrm{III}.
\end{aligned}
$$

Then we have

$$
\begin{aligned}
\mathrm{I} &= \int_{-\infty}^{x/2} ((x-v)/\alpha)^{-p} \varphi(v)\, dv + \int_{x/2}^{x-\alpha x/2} ((x-v)/\alpha)^{-p} \varphi(v)\, dv \\
&\leq (2\alpha)^p x^{-p} \int_{-\infty}^{x/2} \varphi(v)\, dv + \alpha^p \exp(-x^2/8) \int_{x/2}^{x-\alpha x/2} (x-v)^{-p}\, dv \\
&\leq 2^p \alpha^p x^{-p} + 2^{p-1}\alpha^p x^{-p+1} \exp(-x^2/8)/(p-1), \\
\mathrm{II} &= 2^p x^{-p} \int_{x-\alpha x/2}^{x+\alpha x/2} \varphi(u)\, du \leq 2^p x^{-p+1} \alpha \exp(-x^2/8), \\
\mathrm{III} &= \int_{x+\alpha x/2}^{\infty} ((v-x)/\alpha)^{-p} \varphi(v)\, dv \leq 2^{p-1} \alpha x^{-p+1} \exp(-x^2/2)/(p-1).
\end{aligned}
$$

Thus we can bound all the summands on the right side of (2.5) by the right side of the inequality stated in the lemma. □

3. Proof of Theorem 1. Throughout the proof we shall be using the same notation as in the proof of Theorem 1 in Haeusler (1988). We shall have to extend the given sequence $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots \subset \mathscr{F}_n$ to an infinite sequence $\mathscr{F}_i$, $i \geq 0$, of σ-fields. For $i \geq 1$ and a square integrable random variable X we shall use the abbreviation $\sigma_i^2(X) = E(X^2 | \mathscr{F}_{i-1})$.

First, we shall prove the assertion under the additional assumption that

$$
\sum_{i=1}^{n} \sigma_i^2(X_i) = 1 \qquad (3.1)
$$

holds. For any $0 < \beta \le 1$ we define a mds $Y_1, \dots, Y_n$ by

$$Y_i = X_i I(|X_i| \le \beta^{1/2}/2) - E\big(X_i I(|X_i| \le \beta^{1/2}/2) | \mathscr{F}_{i-1}\big)$$

and set $S_n' = \sum_{i=1}^n Y_i$. Then

$$E\big(|S_n - S_n'|^2\big) \le \sum_{i=1}^n E\big(X_i^2 I(|X_i| > \beta^{1/2}/2)\big) \le 4^\delta \beta^{-\delta} L_{n,2\delta},$$

so that by Lemma 1 for all $x > 0$ provided that $4^\delta \beta^{-\delta} L_{n,2\delta} \le 1$,

$$\begin{aligned} |P(S_n \le x) - \Phi(x)| &\le |P(S_n \le x) - P(S_n' \le x)| \\ &\quad + |P(S_n' \le x) - \Phi(x)| \\ &\le C_\delta\big(x^{-2}\beta^{-\delta/3} L_{n,2\delta}^{1/3} + d(S_n', x/2)\big). \end{aligned} \tag{3.2}$$

Let $Y_{n+1}, Y_{n+2}, \dots$ be independent random variables with $P(Y_i = \beta^{1/2}) = 1/2 = P(Y_i = -\beta^{1/2})$ for all i, which are independent of $\mathscr{F}_n$. For $i \ge n+1$ set $\mathscr{F}_i = \sigma(\mathscr{F}_n, Y_{n+1}, \dots, Y_i)$. Then Y_i, $i \ge 1$, is a mds w.r.t. $\mathscr{F}_i$, $i \ge 0$. Observe that the random variable

$$\tau = \max\left\{ l \ge 1 \colon \sum_{i=1}^l \sigma_i^2(Y_i) \le 1 \right\}$$

is a stopping time w.r.t. $\mathscr{F}_i$, $i \ge 0$, for which we have $n \le \tau \le n + [\gamma]$, where $[\cdot]$ denotes the integer part and $\gamma = \beta^{-1}$ for notational convenience. For $i = 1, \dots, k \equiv n + [\gamma] + 1$ we set

$$Z_i = Y_i I(i \le \tau) + \left(\left(1 - \sum_{j=1}^\tau \sigma_j^2(Y_j)\right) \Big/ \beta \right)^{1/2} Y_i I(i = \tau + 1)$$

and obtain a mds w.r.t. $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots \subset \mathscr{F}_k$. Writing from now on

$$\sigma_i^2 = \sigma_i^2(Z_i), \quad \text{for } i = 1, \dots, k, \quad \text{and} \quad S_k'' = \sum_{i=1}^k Z_i,$$

we have

$$\sum_{i=1}^k \sigma_i^2 = 1, \tag{3.3}$$

$$|Z_i| \le \beta^{1/2}, \qquad i = 1, \dots, k, \tag{3.4}$$

and, on account of $Z_i = Y_i$ for $i = 1, \dots, n$,

$$E\big(|S_n' - S_k''|^2\big) = \sum_{i=n+1}^k E\big(Z_i^2\big) \le 2 \cdot 4^\delta \beta^{-\delta} L_{n,2\delta}. \tag{3.5}$$

Thus by Lemma 1 for all $x > 0$ provided that $2 \cdot 4^\delta \beta^{-\delta} L_{n,2\delta} \le 1$,

$$|P(S_n' \le x) - \Phi(x)| \le C_\delta\big(x^{-2}\beta^{-\delta/3} L_{n,2\delta}^{1/3} + d(S_k'', x/2)\big).$$

Since the right side is nonincreasing in x, this inequality implies

$$d(S_n', x) \le C_\delta\big(x^{-2}\beta^{-\delta/3}L_{n,2\delta}^{1/3} + d(S_k'', x/2)\big). \tag{3.6}$$

Combining (3.2) and (3.6), we obtain for all $x > 0$ provided that $2 \cdot 4^\delta\beta^{-\delta}L_{n,2\delta} \le 1$,

$$|P(S_n \le x) - \Phi(x)| \le C_\delta\big(x^{-2}\beta^{-\delta/3}L_{n,2\delta}^{1/3} + d(S_k'', x/4)\big). \tag{3.7}$$

Let $N_1, \dots, N_k$ be standard normal random variables and let ξ be a normal random variable with mean 0 and variance 3β such that $\mathscr{F}_k$, $N_1, \dots, N_k$ and ξ are independent. Then $N_k'' = \sum_{i=1}^k \sigma_i N_i$ is a standard normal random variable because of (3.3). Fix any $p > 1$. Independence of $\mathscr{G} \equiv \sigma(\mathscr{F}_k, N_1, \dots, N_k)$ and ξ implies $E(|\xi|^p|\mathscr{G}) = (3\beta)^{p/2}E(|N_1|^p)$ so that

$$\|E(|\xi|^p|\mathscr{G})\|_\infty = C_p\beta^{p/2} \le C_p\beta^{1/2} \le C_p < \infty, \tag{3.8}$$

where the inequalities hold since $0 < \beta \le 1$. On account of (3.8) we obtain from inequality (2.1) in Lemma 2 for all $x > 0$,

$$\begin{aligned} d(S_k'', x/4) &\le d(S_k'' + \xi, x/8) + C_p x^{-p}\big(\beta^{1/2} + \beta^{p/2} + D(S_k'')\big) \\ &\le \sup\big\{|P(S_k'' + \xi \le t) - P(N_k'' + \xi \le t)|: t \ge x/8\big\} \\ &\quad + d(N_k'' + \xi, x/8) + C_p x^{-p}\big(\beta^{1/2} + D(S_k'')\big). \end{aligned} \tag{3.9}$$

From (3.4) and the inequality in (3.5) we infer

$$\sum_{i=1}^k E\big(|Z_i|^{2+2\delta}\big) \le C_\delta \sum_{i=1}^n E\big(|X_i|^{2+2\delta}\big) + \beta^\delta \sum_{i=n+1}^k E\big(Z_i^2\big) \le C_\delta L_{n,2\delta},$$

hence by Theorem 1 in Haeusler (1988), taking (3.3) into account,

$$D(S_k'') \le C_\delta L_{n,2\delta}^{1/(3+2\delta)}. \tag{3.10}$$

From (3.8) and inequality (2.2) in Lemma 2 we conclude for all $x > 0$,

$$\begin{aligned} d(N_k'' + \xi, x/8) &\le d(N_k'', x/16) + C_p x^{-p}\big(\beta^{1/2} + \beta^{p/2} + D(N_k'')\big) \\ &\le C_p x^{-p}\beta^{1/2}, \end{aligned} \tag{3.11}$$

since $d(N_k'', x/16) = 0 = D(N_k'')$ in view of the fact that N_k'' is standard normal.

To derive a bound for the supremum on the right side of (3.9), we shall show for all $x \ge 2$ that

$$\begin{aligned} &|P(S_k'' + \xi \le x) - P(N_k'' + \xi \le x)| \\ &\quad \le C_{\delta,p} L_{n,2\delta}^{1/(2+2\delta)}\Big\{d(S_k'', x/4)^{(1+2\delta)/(2+2\delta)}\beta^{-1/2} + x^{-p(1+2\delta)/(2+2\delta)} \\ &\qquad \times \Big(L_{n,2\delta}^{(1+2\delta)/((2+2\delta)(3+2\delta))}\beta^{-1/2} + \beta^{-1/(4+4\delta)}\Big)\Big\}. \end{aligned} \tag{3.12}$$

For the proof of (3.12) we fix $x \geq 2$ and set $U_m = \Sigma_{i=1}^{m-1} Z_i$, $\lambda_m^2 = \Sigma_{i=m+1}^{k} \sigma_i^2 + 3\beta$ and $T_m = \lambda_m^{-1}(x - U_m)$. Since null sets do not affect distributional properties, we may and do assume w.l.o.g. that $\mathscr{F}_0$ contains all P-null sets. Then λ_m^2 is $\mathscr{F}_{m-1}$-measurable because of (3.3), and together with the mds property of $Z_1, \dots, Z_k$ this fact enables one to obtain the crucial estimate

$$\begin{aligned} &|P(S_k'' + \xi \leq x) - P(N_k'' + \xi \leq x)| \\ (3.13)\qquad &\leq \tfrac{1}{6} \sum_{m=1}^{k} E\left(|\varphi''(T_m - \theta_m \lambda_m^{-1} Z_m)| \lambda_m^{-3} |Z_m|^3\right) \\ &\quad + \tfrac{1}{6} \sum_{m=1}^{k} E\left(|\varphi''(T_m - \tilde{\theta}_m \lambda_m^{-1} \sigma_m N_m)| \lambda_m^{-3} \sigma_m^3 |N_m|^3\right) \equiv \tfrac{1}{6}\mathrm{I} + \tfrac{1}{6}\mathrm{II}, \end{aligned}$$

where $0 \leq \theta_m, \tilde{\theta}_m \leq 1$; cf. the arguments leading from (4.2)–(4.4) in Bolthausen (1982).

To apply Bolthausen's (1982) method for deriving bounds on I and II, we introduce stopping times $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_{[\gamma]} \leq \tau_{[\gamma]+1} = k$ defined by

$$\tau_j = \inf\left\{ l \geq 1 : \sum_{i=1}^{l} \sigma_i^2 \geq j\beta \right\}, \qquad j = 1, \dots, [\gamma].$$

Then for $j = 1, \dots, [\gamma] + 1$ and $m = 1, \dots, k$ on the event $\{\tau_{j-1} < m \leq \tau_j\}$,

$$\lambda_m^2 \leq 1 - (j-1)\beta + 3\beta \equiv \bar{\lambda}_j^2, \tag{3.14}$$

and

$$\lambda_m^2 \geq 1 - (j+1)\beta + 3\beta \equiv \underline{\lambda}_j^2 \geq \beta > 0. \tag{3.15}$$

Setting $R_m = \Sigma_{i=\tau_{j-1}+1}^{m-1} Z_i$ and $A_m = \{|R_m| \leq |x - U_{\tau_{j-1}+1}|/2\}$, we use the stopping times τ_j, (3.15) and $\|\varphi''\|_\infty \leq 1$ to write

$$\begin{aligned} \mathrm{I} &= \sum_{j=1}^{[\gamma]+1} E\left(\sum_{m=\tau_{j-1}+1}^{\tau_j} |\varphi''(T_m - \theta_m \lambda_m^{-1} Z_m)| \lambda_m^{-3} |Z_m|^3 \right) \\ (3.16)\qquad &\leq \sum_{j=1}^{[\gamma]+1} \underline{\lambda}_j^{-3} \left\{ E\left(\sum_{m=\tau_{j-1}+1}^{\tau_j} |\varphi''(T_m - \theta_m \lambda_m^{-1} Z_m)| \, |Z_m|^3 I(A_m) \right) \right. \\ &\qquad \left. + E\left(\sum_{m=\tau_{j-1}+1}^{\tau_j} |Z_m|^3 I(A_m^c) \right) \right\} \\ &\equiv \sum_{j=1}^{[\gamma]+1} \underline{\lambda}_j^{-3} \{ \mathrm{I}_{j,1} + \mathrm{I}_{j,2} \}. \end{aligned}$$

The following estimate is crucial for producing a bound on $\mathrm{I}_{j,1} + \mathrm{I}_{j,2}$. For each $j = 1, \ldots, [\gamma] + 1$ and all $x \geq 2$ we have

$$E\Big(\min\Big(1, \big|\bar{\lambda}_j^{-1}\big(x - U_{\tau_{j-1}+1}\big)\big|^{-p}\Big)\Big) \leq C_{\delta,p}\Big(d(S_k'', x/4) + x^{-p}\Big(L_{n,2\delta}^{1/(3+2\delta)} + \bar{\lambda}_j\Big)\Big). \tag{3.17}$$

For the proof of (3.17) notice from (3.14) that $\bar{\lambda}_j^2 \leq 1 + 3\beta \leq 4$, hence

$$E\Big(\min\Big(1, \big|\bar{\lambda}_j^{-1}\big(x - U_{\tau_{j-1}+1}\big)\big|^{-p}\Big)\Big) \leq 2^p E\Big(\min\Big(1, \big|\big(\bar{\lambda}_j/2\big)^{-1}\big(x - U_{\tau_{j-1}+1}\big)\big|^{-p}\Big)\Big),$$

which for all $x \geq 2$ by Lemma 3 is less than or equal to

$$C_p\Big(d\big(U_{\tau_{j-1}+1}, x/2\big) + x^{-p}\Big(D\big(U_{\tau_{j-1}+1}\big) + \bar{\lambda}_j/2\Big)\Big).$$

Let $\mathscr{F}(\tau_{j-1})$ denote the σ-field of all events known at time τ_{j-1}. Applying the conditional form of a convex function inequality for martingales, cf., e.g., Theorem 2.11 in Hall and Heyde (1980), we obtain

$$E\left(\max_{\tau_{j-1} < l \leq k}\left|\sum_{i=\tau_{j-1}+1}^{l} Z_i\right|^p \Big| \mathscr{F}(\tau_{j-1})\right) \leq C_p E\left(\left(\sum_{i=\tau_{j-1}+1}^{k} \sigma_i^2\right)^{p/2} + \max_{\tau_{j-1}+1 \leq i \leq k} |Z_i|^p \Big| \mathscr{F}(\tau_{j-1})\right),$$

which by (3.3), (3.4) and the definition of τ_{j-1} is less than or equal to

$$C_p\big((1 - (j-1)\beta)^{p/2} + \beta^{p/2}\big) \leq C_p \bar{\lambda}_j^p.$$

Thus

$$\left\| E\left(\max_{\tau_{j-1} < l \leq k}\left|\sum_{i=\tau_{j-1}+1}^{l} Z_i\right|^p \Big| U_{\tau_{j-1}+1}\right)\right\|_\infty \leq C_p \bar{\lambda}_j^p \leq C_p < \infty. \tag{3.18}$$

Consequently, we conclude from Lemma 2(i) and (3.10)

$$D\big(U_{\tau_{j-1}+1}\big) \leq D(S_k'') + C_p \bar{\lambda}_j \leq C_{\delta,p}\Big(L_{n,2\delta}^{1/(3+2\delta)} + \bar{\lambda}_j\Big)$$

and from this result and (2.1) in Lemma 2(ii) for all $x > 0$ (observe that $\bar{\lambda}_j \leq 2$)

$$d\big(U_{\tau_{j-1}+1}, x/2\big) \leq d(S_k'', x/4) + C_{\delta,p} x^{-p}\Big(\bar{\lambda}_j + L_{n,2\delta}^{1/(3+2\delta)}\Big).$$

This completes the proof of (3.17).

Now we consider $\mathrm{I}_{j,1}$ for fixed j. Let the function $\psi\colon R \to [0,\infty)$ be defined by $\psi(x) = \sup\{|\varphi''(y)|\colon |y| \geq (|x|/2) - 1\}$. Then we have as in the proof of Theorem 1 in Haeusler (1988) (notice that $\psi^{(2+2\delta)/(1+2\delta)} \leq \psi$ since $\|\varphi''\|_\infty \leq 1$)

$$\mathrm{I}_{j,1} \leq C_\delta \beta E\Big(\psi\big(\bar{\lambda}_j^{-1}(x - U_{\tau_{j-1}+1})\big)\Big)^{(1+2\delta)/(2+2\delta)} \times E\Big(\max_{\tau_{j-1}<m\leq\tau_j} |Z_m|^{2+2\delta}\Big)^{1/(2+2\delta)}.$$

Obviously, $\psi(x) \leq C_p \min(1, |x|^{-p})$ for all $x \in R$ so that by (3.17)

$$\mathrm{I}_{j,1} \leq C_{\delta,p}\beta\Big(d(S_k'', x/4) + x^{-p}\big(L_{n,2\delta}^{1/(3+2\delta)} + \bar{\lambda}_j\big)\Big)^{(1+2\delta)/(2+2\delta)} \times E\Big(\max_{\tau_{j-1}<m\leq\tau_j} |Z_m|^{2+2\delta}\Big)^{1/(2+2\delta)}. \tag{3.19}$$

For $\mathrm{I}_{j,2}$ we have as in the proof of Theorem 1 in Haeusler (1988)

$$\mathrm{I}_{j,2} \leq C_\delta \beta P(B_j)^{(1+2\delta)/(2+2\delta)} E\Bigg(\sum_{m=\tau_{j-1}+1}^{\tau_j} |Z_m|^{2+2\delta}\Bigg)^{1/(2+2\delta)},$$

where

$$B_j = \Bigg\{\max_{\tau_{j-1}<l\leq\tau_j}\Bigg|\sum_{m=\tau_{j-1}+1}^{l} Z_m\Bigg| > |x - U_{\tau_{j-1}+1}|/4\Bigg\}.$$

Therefore

$$P(B_j) \leq E\Bigg\{\min\Bigg(1, \big(|x - U_{\tau_{j-1}+1}|/4\big)^{-p} \max_{\tau_{j-1}<l\leq\tau_j}\Bigg|\sum_{m=\tau_{j-1}+1}^{l} Z_m\Bigg|^p\Bigg)\Bigg\}$$

$$\leq C_p E\Bigg\{\min\Bigg(1, |x - U_{\tau_{j-1}+1}|^{-p} E\Bigg(\max_{\tau_{j-1}<l\leq\tau_j}\Bigg|\sum_{m=\tau_{j-1}+1}^{l} Z_m\Bigg|^p \Bigg| U_{\tau_{j-1}+1}\Bigg)\Bigg)\Bigg\},$$

which in view of (3.18) and (3.17) is less than or equal to

$$C_p E\Big(\min\big(1, |\bar{\lambda}_j^{-1}(x - U_{\tau_{j-1}+1})|^{-p}\big)\Big) \leq C_{\delta,p}\Big(d(S_k'', x/4) + x^{-p}\big(L_{n,2\delta}^{1/(3+2\delta)} + \bar{\lambda}_j\big)\Big).$$

Consequently, we have

$$\mathrm{I}_{j,2} \leq C_{\delta,p}\beta\Big(d(S_k'', x/4) + x^{-p}\big(L_{n,2\delta}^{1/(3+2\delta)} + \bar{\lambda}_j\big)\Big)^{(1+2\delta)/(2+2\delta)} \times E\Bigg(\sum_{m=\tau_{j-1}+1}^{\tau_j} |Z_m|^{2+2\delta}\Bigg)^{1/(2+2\delta)}. \tag{3.20}$$

Substituting (3.19) and (3.20) into the right side of (3.16) and employing the reasoning leading from (2.26) to (2.27) in Haeusler (1988), we obtain

$$\mathrm{I} \le C_{\delta,p}\beta \sum_{j=1}^{[\gamma]+1} \underline{\lambda}_j^{-3}\Big\{d(S_k'', x/4) + x^{-p}\big(L_{n,2\delta}^{1/(3+2\delta)} + \bar{\lambda}_j\big)\Big\}^{(1+2\delta)/(2+2\delta)}$$
$$\times E\left(\sum_{m=\tau_{j-1}+1}^{\tau_j} |Z_m|^{2+2\delta}\right)^{1/(2+2\delta)}$$
$$\le C_{\delta,p}\beta L_{n,2\delta}^{1/(2+2\delta)}\Big\{d(S_k'', x/4)^{(1+2\delta)/(2+2\delta)}\beta^{-3/2} + x^{-p(1+2\delta)/(2+2\delta)}$$
$$\times\Big(L_{n,2\delta}^{(1+2\delta)/((2+2\delta)(3+2\delta))}\beta^{-3/2} + \beta^{(-5-4\delta)/(4+4\delta)}\Big)\Big\},$$

and this bound equals the terms on the right side of (3.12). Therefore, on account of (3.13), the proof of (3.12) will be complete if we establish the same bound for II. This can be done exactly as in Haeusler (1988) by the same reduction to the already solved problem of bounding $\mathrm{I}_{j,1}$ and $\mathrm{I}_{j,2}$. One only has to note that for the function $h(x) = E(|N_1|^3 I(8|N_1| > |x|))$, $x \in R$, used to bound the term $\mathrm{II}_{j,3}$ in Haeusler (1988) one has

$$h(x)^{(2+2\delta)/(1+2\delta)} \le C_{\delta,p}\min(1, |x|^{-p}),$$

for all $x \in R$ so that (3.17) can be applied to produce the appropriate bound for $\mathrm{II}_{j,3}$ also in the present situation of nonuniform estimates. We omit the details of these straightforward modifications of Haeusler (1988) here, which complete the proof of (3.12).

Since the right side of (3.12) is nonincreasing in x we conclude from (3.9)–(3.12) that for all $x \ge 16$,

$$\begin{aligned} d(S_k'', x/4) \le C_{\delta,p}\Big\{&L_{n,2\delta}^{1/(2+2\delta)} d(S_k'', x/32)^{(1+2\delta)/(2+2\delta)}\beta^{-1/2} \\ &+ x^{-p(1+2\delta)/(2+2\delta)}\Big(L_{n,2\delta}^{2/(3+2\delta)}\beta^{-1/2} + L_{n,2\delta}^{1/(2+2\delta)}\beta^{-1/(4+4\delta)}\Big) \\ &+ x^{-p}\beta^{1/2} + x^{-p}L_{n,2\delta}^{1/(3+2\delta)}\Big\}. \end{aligned} \tag{3.21}$$

Arguments similar to those that established (3.7) can be used to prove

$$d(S_k'', x) \le C_\delta\big(x^{-2}\beta^{-\delta/3}L_{n,2\delta}^{1/3} + d(S_n, x/4)\big),$$

for all $x > 0$ provided that $2 \cdot 4^\delta\beta^{-\delta}L_{n,2\delta} \le 1$. Applying this inequality to the right side of (3.21) and substituting the resulting bound for $d(S_k'', x/4)$ into (3.7), we obtain for all $x \ge 16$ provided that $2 \cdot 4^\delta\beta^{-\delta}L_{n,2\delta} \le 1$,

$$\begin{aligned} &|P(S_n \le x) - \Phi(x)| \\ &\quad\le C_{\delta,p}\Big\{d(S_n, x/128)^{(1+2\delta)/(2+2\delta)}L_{n,2\delta}^{1/(2+2\delta)}\beta^{-1/2} \\ &\qquad+ \big(x^{-2}\beta^{-\delta/3}L_{n,2\delta}^{1/3}\big)^{(1+2\delta)/(2+2\delta)}L_{n,2\delta}^{1/(2+2\delta)}\beta^{-1/2} \\ &\qquad+ x^{-p(1+2\delta)/(2+2\delta)}\Big(L_{n,2\delta}^{2/(3+2\delta)}\beta^{-1/2} + L_{n,2\delta}^{1/(2+2\delta)}\beta^{-1/(4+4\delta)}\Big) \\ &\qquad+ x^{-p}\beta^{1/2} + x^{-p}L_{n,2\delta}^{1/(3+2\delta)} + x^{-2}\beta^{-\delta/3}L_{n,2\delta}^{1/3}\Big\}. \end{aligned} \tag{3.22}$$

Now we consider an $x \geq x_\delta \equiv \max(16, (2 \cdot 4^\delta)^{1/(6\delta)})$ and assume that $0 < \beta \equiv x^6 L_{n,2\delta}^{2/(3+2\delta)} \leq 1$. Then we have $2 \cdot 4^\delta \beta^{-\delta} L_{n,2\delta} \leq 1$ (recall $0 < L_{n,2\delta} \leq 1$ by assumption) so that we can substitute β into the right side of (3.22). Choosing at the same time a large enough, but fixed $p = p_\delta > 1$, this gives

$$\begin{aligned} &|P(S_n \leq x) - \Phi(x)| \\ &\quad \leq C_\delta\Big(d(S_n, x/128)^{(1+2\delta)/(2+2\delta)} x^{-3} L_{n,2\delta}^{1/((2+2\delta)(3+2\delta))} + x^{-2-2\delta} L_{n,2\delta}^{1/(3+2\delta)}\Big). \end{aligned} \tag{3.23}$$

Next we consider the second case $1 < x^6 L_{n,2\delta}^{2/(3+2\delta)}$. Setting

$$X_i' = X_i I(|X_i| \leq 1) - E\big(X_i I(|X_i| \leq 1) | \mathscr{F}_{i-1}\big) \quad \text{and} \quad X_i'' = X_i - X_i',$$

for $i = 1, \dots, n$, we have

$$\begin{aligned} |P(S_n \leq x) - \Phi(x)| &\leq P\left(\left|\sum_{i=1}^n (X_i' + X_i'')\right| > x\right) + 1 - \Phi(x) \\ &\leq P\left(\left|\sum_{i=1}^n X_i'\right| > x/2\right) + P\left(\left|\sum_{i=1}^n X_i''\right| > x/2\right) \\ &\quad + x^{-1}\exp(-x^2/2). \end{aligned} \tag{3.24}$$

Applying the convex function inequality in Theorem 2.11 of Hall and Heyde (1980) to the mds $X_1', \dots, X_n'$, we get

$$P\left(\left|\sum_{i=1}^n X_i'\right| > x/2\right) \leq C_\delta x^{-8-2\delta}\left\{E\left(\left|\sum_{i=1}^n \sigma_i^2(X_i')\right|^{4+\delta}\right) + E\Big(\max_{1\leq i\leq n} |X_i'|^{8+2\delta}\Big)\right\},$$

so that on account of $|X_i'| \leq 2$, $\sigma_i^2(X_i') \leq \sigma_i^2(X_i)$ and assumption (3.1)

$$P\left(\left|\sum_{i=1}^n X_i'\right| > x/2\right) \leq C_\delta x^{-8-2\delta} \leq C_\delta x^{-2-2\delta} L_{n,2\delta}^{1/(3+2\delta)}, \tag{3.25}$$

where the last inequality follows from $1 < x^6 L_{n,2\delta}^{2/(3+2\delta)}$. Moreover, by an application of Theorem 2.11 in Hall and Heyde (1980) to the mds $X_1'', \dots, X_n''$,

$$P\left(\left|\sum_{i=1}^n X_i''\right| > x/2\right) \leq C_\delta x^{-2-2\delta}\left\{E\left(\left|\sum_{i=1}^n \sigma_i^2(X_i'')\right|^{1+\delta}\right) + E\Big(\max_{1\leq i\leq n} |X_i''|^{2+2\delta}\Big)\right\}.$$

Observe that $\sigma_i^2(X_i'') \leq E(X_i^2 I(|X_i| > 1) | \mathscr{F}_{i-1}) \leq \sigma_i^2(X_i)$ so that $\sum_{i=1}^n \sigma_i^2(X_i'') \leq 1$ by assumption (3.1). Hence

$$\begin{aligned} P\left(\left|\sum_{i=1}^n X_i''\right| > x/2\right) &\leq C_\delta x^{-2-2\delta}\left\{E\left(\sum_{i=1}^n \sigma_i^2(X_i'')\right) + \sum_{i=1}^n E\big(|X_i''|^{2+2\delta}\big)\right\} \\ &\leq C_\delta x^{-2-2\delta}\left\{\sum_{i=1}^n E\big(X_i^2 I(|X_i| > 1)\big) + L_{n,2\delta}\right\} \\ &\leq C_\delta x^{-2-2\delta} L_{n,2\delta} \leq C_\delta x^{-2-2\delta} L_{n,2\delta}^{1/(3+2\delta)}. \end{aligned} \tag{3.26}$$

From $1 < x^6 L_{n,2\delta}^{2/(3+2\delta)}$ it follows that

$$x^{-1}\exp(-x^2/2) \le x^2\exp(-x^2/2)L_{n,2\delta}^{1/(3+2\delta)} \le C_\delta x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)},$$

and combining this result with (3.24)–(3.26), we arrive at

$$|P(S_n \le x) - \Phi(x)| \le C_\delta x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}.$$

Thus we see that inequality (3.23) is true for all $x \ge x_\delta$. The right side of (3.23) is nonincreasing in x so that (3.23) implies for $x \ge x_\delta$,

$$d(S_n, x) \le C_\delta\Big\{d(S_n, x/128)^{(1+2\delta)/(2+2\delta)}x^{-3}L_{n,2\delta}^{1/((2+2\delta)(3+2\delta))} + x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}\Big\}. \tag{3.27}$$

If $d(S_n, x/128) \le x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}$, then utilizing this inequality in (3.27) yields $d(S_n, x) \le C_\delta x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}$. If $d(S_n, x/128) > x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}$, then we have $x^{-1}L_{n,2\delta}^{1/((2+2\delta)(3+2\delta))} < d(S_n, x/128)^{1/(2+2\delta)}$, and (3.27) now implies

$$d(S_n, x) \le C_\delta\big(x^{-2}d(S_n, x/128) + x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}\big).$$

Hence for all $x \ge x_\delta$,

$$d(S_n, x) \le C_\delta^*\big(x^{-2}d(S_n, x/128) + x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}\big), \tag{3.28}$$

where C_δ^* is some finite constant depending only on δ. For $x \ge x_\delta^* \equiv \max(x_\delta, 128^{1+\delta}(2C_\delta^*)^{1/2})$ we have $x^{-2} \le 128^{-2-2\delta}(2C_\delta^*)^{-1}$, and (3.28) yields for all $x \ge x_\delta^*$,

$$d(S_n, x) \le (1/2)128^{-2-2\delta}d(S_n, x/128) + C_\delta^* x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}. \tag{3.29}$$

For fixed $x \ge x_\delta^*$ we set

$$h = \min\big\{m \in \{0,1,2,\dots\}: d(S_n, x/128^m) \ge 128^{-2-2\delta}d\big(S_n, x/128^{m+1}\big) \text{ or } x/128^m \le x_\delta^*\big\}$$

so that

$$d(S_n, x) \le 128^{-h(2+2\delta)}d\big(S_n, x/128^h\big). \tag{3.30}$$

If $x/128^h \le x_\delta^*$, then we have by (3.30) and (1.1)

$$d(S_n, x) \le (x_\delta^*/x)^{2+2\delta}D(S_n) \le C_\delta x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}.$$

If $s \equiv x/128^h > x_\delta^*$, then (3.29) holds with x replaced by s, and by definition of h we must have $d(S_n, s) \ge 128^{-2-2\delta}d(S_n, s/128)$ so that

$$d(S_n, s) \le \tfrac{1}{2}d(S_n, s) + C_\delta^* s^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)},$$

i.e.,

$$d\big(S_n, x/128^h\big) \le 2C_\delta^* 128^{h(2+2\delta)}x^{-2-2\delta}L_{n,2\delta}^{1/(3+2\delta)}.$$

Combined with (3.30), the last inequality gives $d(S_n, x) \le 2C_\delta^* x^{-2-2\delta} L_{n,2\delta}^{1/(3+2\delta)}$, and we have shown that (1.2) is true for all $x \ge x_\delta^*$. For $0 \le x < x_\delta^*$ it follows trivially from (1.1), and for $-\infty < x < 0$ it is easily obtained by an application of the just proven result to the mds $-X_1, \dots, -X_n$. This completes the proof of (1.2) under the additional assumption (3.1). To remove this assumption, one proceeds similarly as in the proof of the main result in Haeusler (1984), using Theorem 2.11 in Hall and Heyde (1980) and Lemma 1. We skip the details of these considerations, but remark that for $x \in R$ and $L_{n,2\delta} + N_{n,2\delta} > 1$ this reasoning also yields the inequality

$$|P(S_n \le x) - \Phi(x)| \le C_\delta \big(1 + |x|^{2+2\delta}\big)^{-1} \big(L_{n,2\delta} + N_{n,2\delta}\big). \qquad (3.31) \quad \square$$

4. Proof of Theorem 2. Let $m \in \{0, 1, \dots, n-1\}$ be fixed such that $Q_{n,2\delta}(m) \le 1$. Set $S_m = \sum_{i=1}^m X_i$, and observe that $\eta_m^2 > 0$ a.s. The first steps in the proof are borrowed from the proof of Theorem 1 of Hall and Heyde (1981). From their inequality (10) we deduce for $m = 0, 1, \dots, n-1$, all $x \in R$ and $a > 0$,

$$\begin{aligned} &\big|P(S_n \le x) - E\big(\Phi(\eta^{-1}x)\big)\big| \\ &\quad\le \max_{\beta=\pm 1} \big\{\big|P(S_n - S_m \le x + \beta a) - E\big(\Phi\big(\eta_m^{-1}(x + \beta a)\big)\big)\big|\big\} \\ &\qquad + \max_{\beta=\pm 1} \big\{\big|E\big(\Phi\big(\eta_m^{-1}(x + \beta a)\big)\big) - E\big(\Phi\big(\eta^{-1}(x + \beta a)\big)\big)\big|\big\} \qquad (4.1) \\ &\qquad + \max_{\beta=\pm 1} \big\{\big|E\big(\Phi\big(\eta^{-1}(x + \beta a)\big)\big) - E\big(\Phi\big(\eta^{-1}x\big)\big)\big|\big\} + P(|S_m| > a) \\ &\quad\equiv \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}. \end{aligned}$$

Burkholder's inequality implies

$$\mathrm{IV} \le C_\delta a^{-2-2\delta} E\left(\left|\sum_{i=1}^m X_i^2\right|^{1+\delta}\right) \le C_\delta a^{-2-2\delta} Q_{n,2\delta}(m), \qquad (4.2)$$

and with $m(a, x) = \min\{|u|: x - a \le u \le x + a\}$ one has for $x \in R$ and $\beta = \pm 1$,

$$\begin{aligned} \big|\Phi\big(\eta^{-1}(x + \beta a)\big) - \Phi\big(\eta^{-1}x\big)\big| &\le \eta^{-1} a \exp\big(-\eta^{-2} m(a, x)^2/2\big) \\ &\le C_\delta \Big(1 + \big(\eta^{-1} m(a, x)\big)^{3+2\delta}\Big)^{-1} \eta^{-1} a \\ &\le C_\delta a \eta^{2+2\delta} \big(1 + m(a, x)^{3+2\delta}\big)^{-1} I(\eta > 1) \\ &\quad + C_\delta a \eta^{-1} \big(1 + m(a, x)^{3+2\delta}\big)^{-1} I(0 < \eta \le 1), \end{aligned}$$

hence

$$\mathrm{III} \le C_\delta C_\delta(\eta) a \big(1 + m(a, x)^{3+2\delta}\big)^{-1}. \qquad (4.3)$$

To produce a bound for II, we consider first a fixed $t \in R$ and write

$$\begin{aligned}&|\Phi(\eta_m^{-1}t) - \Phi(\eta^{-1}t)|\\&\quad\le \big(\exp(-\eta_m^{-2}t^2/2) + \exp(-\eta^{-2}t^2/2)\big)|t|\,|\eta^{-1} - \eta_m^{-1}|\\&\quad= |\eta^2 - \eta_m^2|\big(\eta\eta_m(\eta + \eta_m)\big)^{-1}\big(1 + |t|^{2+2\delta}\big)^{-1}\\&\qquad\times\big(|t| + |t|^{3+2\delta}\big)\big(\exp(-\eta_m^{-2}t^2/2) + \exp(-\eta^{-2}t^2/2)\big)\end{aligned}$$

and use $\exp(-At^2/2) \le (\alpha/A)^{\alpha/2}\exp(-\alpha/2)$ for $t \in R$ and $0 < \alpha, A < \infty$ to obtain

$$\begin{aligned}&|\Phi(\eta_m^{-1}t) - \Phi(\eta^{-1}t)|\\&\quad\le C_\delta\big(1 + |t|^{2+2\delta}\big)^{-1}|\eta^2 - \eta_m^2|\big(\eta^{-2} + \eta_m^{-2} + \eta^{-1}\eta_m^{1+2\delta} + \eta_m^{-1}\eta^{1+2\delta}\big).\end{aligned}$$

Hence by Hölder's inequality, with $A_m = \{|\eta^2 - \eta_m^2| \le 1\}$,

$$\begin{aligned}&E\big(|\Phi(\eta_m^{-1}t) - \Phi(\eta^{-1}t)|I(A_m)\big)\\&\quad\le C_\delta\big(1 + |t|^{2+2\delta}\big)^{-1}E\big(|\eta^2 - \eta_m^2|^{3+2\delta}I(A_m)\big)^{1/(3+2\delta)}\\&\qquad\times E\Big(\big(\eta^{-2} + \eta_m^{-2} + \eta^{-1}\eta_m^{1+2\delta} + \eta_m^{-1}\eta^{1+2\delta}\big)^{(3+2\delta)/(2+2\delta)}\Big)^{(2+2\delta)/(3+2\delta)}.\end{aligned}$$

Notice that the first expectation on the right side is smaller than $Q_{n,2\delta}(m)$. The function $x^{(-3-2\delta)/2}$ is convex for $0 < x < \infty$ so that by Jensen's inequality $\eta_m^{-3-2\delta} = E(\eta^2|\mathscr{F}_m)^{(-3-2\delta)/2} \le E(\eta^{-3-2\delta}|\mathscr{F}_m)$, hence $E(\eta_m^{-3-2\delta}) \le E(\eta^{-3-2\delta})$. Using this result, Hölder's inequality, $E(\eta^{3+2\delta}) + E(\eta^{-3-2\delta}) \ge 2$ and the elementary estimate $xy \le p^{-1}x^p + q^{-1}y^q$ for $x, y \ge 0$ and $p^{-1} + q^{-1} = 1$, it is a matter of straightforward computations to verify that the last factor on the right side of the preceding inequality is smaller than $C_\delta(\eta)$. Consequently,

$$\begin{aligned}&E\big(|\Phi(\eta_m^{-1}t) - \Phi(\eta^{-1}t)|I(A_m)\big)\\&\quad\le C_\delta C_\delta(\eta)\big(1 + |t|^{2+2\delta}\big)^{-1}Q_{n,2\delta}(m)^{1/(3+2\delta)}.\end{aligned}\tag{4.4}$$

For $0 \le t \le 1$ we have

$$\begin{aligned}&|\Phi(\eta_m^{-1}t) - \Phi(\eta^{-1}t)|I(A_m^c)\\&\quad\le 2\big(1 + |t|^{2+2\delta}\big)^{-1}\big(\eta^{2+2\delta} + \eta_m^{2+2\delta}\big)I(A_m^c),\end{aligned}\tag{4.5}$$

since $\eta^{2+2\delta} + \eta_m^{2+2\delta} > 1$ on the event A_m^c. For $t > 1$ we write

$$|\Phi(\eta_m^{-1}t) - \Phi(\eta^{-1}t)|I(A_m^c) \le \big(|1 - \Phi(\eta_m^{-1}t)| + |1 - \Phi(\eta^{-1}t)|\big)I(A_m^c)$$

and consider the two cases $\eta_m^{-1}t \le 1$ and $1 < \eta_m^{-1}t$ separately, which easily leads to the same bound as before. The same reasoning applies with η instead of η_m so that inequality (4.5), with the factor 2 replaced by C_δ, holds for all $t \ge 0$ and therefore, by symmetry of Φ, for all $t \in R$. Integrating and applying Hölder's inequality, we obtain

$$\begin{aligned}
&E\Big(\big|\Phi\big(\eta_m^{-1}t\big) - \Phi\big(\eta^{-1}t\big)\big|I(A_m^c)\Big) \\
&\quad\le C_\delta\big(1 + |t|^{2+2\delta}\big)^{-1}\Big(E\big(\eta_m^{3+2\delta}\big) \\
&\qquad\qquad + E\big(\eta^{3+2\delta}\big)\Big)^{(2+2\delta)/(3+2\delta)} P\big(|\eta^2 - \eta_m^2| > 1\big)^{1/(3+2\delta)} \\
&\quad\le C_\delta C_\delta(\eta)\big(1 + |t|^{2+2\delta}\big)^{-1} Q_{n,2\delta}(m)^{1/(3+2\delta)}.
\end{aligned}$$

Combining this estimate with (4.4), we arrive at

$$\big|E\big(\Phi\big(\eta_m^{-1}t\big)\big) - E\big(\Phi\big(\eta^{-1}t\big)\big)\big| \le C_\delta C_\delta(\eta)\big(1 + |t|^{2+2\delta}\big)^{-1} Q_{n,2\delta}(m)^{1/(3+2\delta)},$$

hence for all $x \in R$ and $a > 0$,

$$\text{II} \le C_\delta C_\delta(\eta)\Big(\max_{\beta = \pm 1}\big(1 + |x + \beta a|^{2+2\delta}\big)^{-1}\Big) Q_{n,2\delta}(m)^{1/(3+2\delta)}. \tag{4.6}$$

It remains to estimate the term I on the right side of (4.1). For this, we consider again a fixed $t \in R$ first and write

$$\begin{aligned}
&\big|P(S_n - S_m \le t) - E\big(\Phi\big(\eta_m^{-1}t\big)\big)\big| \\
&\qquad \le E\Big(\big|P(S_n - S_m \le t|G_m) - \Phi\big(\eta_m^{-1}t\big)\big|\Big).
\end{aligned} \tag{4.7}$$

Let the random variables $Y_{m+1}, \dots, Y_n$ have the distribution of $X_{m+1}, \dots, X_n$ conditional on $\mathscr{G}_m$, and set $\mathscr{H}_i = \sigma(Y_{m+1}, \dots, Y_i)$ for $i = m + 1, \dots, n$, with $\mathscr{H}_m$ being the trivial σ-field. The conditional probability measure and expectation operator will be denoted by P_m and E_m, respectively. Since $X_{m+1}, \dots, X_n$ is a mds and η_m is measurable w.r.t. $\mathscr{G}_m$, the random variables $\eta_m^{-1}(\omega)Y_{m+1}, \dots, \eta_m^{-1}(\omega)Y_n$ constitute a mds under P_m w.r.t. $\mathscr{H}_m \subset \mathscr{H}_{m+1} \subset \cdots \subset \mathscr{H}_n$ for P-almost all fixed ω in the basic probability space $(\Omega, \mathscr{F}, P)$. Thus we can apply Theorem 1 together with inequality (3.31) to obtain P-a.s.,

$$\begin{aligned}
&\big|P(S_n - S_m \le t|\mathscr{G}_m) - \Phi\big(\eta_m^{-1}t\big)\big| \\
&\quad = \left|P_m\left(\sum_{i=m+1}^{n} \eta_m^{-1}Y_i \le \eta_m^{-1}t\right) - \Phi\big(\eta_m^{-1}t\big)\right| \\
&\quad \le C_\delta\big(1 + |\eta_m^{-1}t|^{2+2\delta}\big)^{-1} \\
&\qquad \times\Big(\big(L_{n,2\delta}(m) + N_{n,2\delta}(m)\big)^{1/(3+2\delta)} + L_{n,2\delta}(m) + N_{n,2\delta}(m)\Big),
\end{aligned} \tag{4.8}$$

where

$$L_{n,2\delta}(m) = \sum_{i=m+1}^{n} E_m\big(|\eta_m^{-1}Y_i|^{2+2\delta}\big)$$

and

$$N_{n,2\delta}(m) = E_m\left(\left|\sum_{i=m+1}^{n} E_m\left(\eta_m^{-2}Y_i^2|\mathscr{H}_{i-1}\right) - 1\right|^{1+\delta}\right).$$

We shall exploit inequality (4.8) for all $t \in R$ with $|t| \geq 1$. Observe that η_m is a constant w.r.t. E_m so that for $i = m+1,\dots,n$, P-a.s.

$$E_m\left(|\eta_m^{-1}Y_i|^{2+2\delta}\right) = \eta_m^{-2-2\delta}E_m\left(|Y_i|^{2+2\delta}\right) = \eta_m^{-2-2\delta}E\left(|X_i|^{2+2\delta}|\mathscr{G}_m\right).$$

Therefore, we have

$$\begin{aligned}(4.9)\qquad &\left(1 + |\eta_m^{-1}t|^{2+2\delta}\right)^{-1}L_{n,2\delta}(m)^{1/(3+2\delta)}\\ &\quad= C_\delta\left(\eta_m^{2+2\delta} + |t|^{2+2\delta}\right)^{-1}\eta_m^{2+2\delta-(2+2\delta)/(3+2\delta)}\\ &\qquad\times\left(\sum_{i=m+1}^{n} E\left(|X_i|^{2+2\delta}|\mathscr{G}_m\right)\right)^{1/(3+2\delta)}\end{aligned}$$

Integration and application of Hölder's inequality yield

$$\begin{aligned}(4.10)\qquad &E\left(\left(1 + |\eta_m^{-1}t|^{2+2\delta}\right)^{-1}L_{n,2\delta}(m)^{1/(3+2\delta)}\right)\\ &\quad\leq C_\delta\left(1 + |t|^{2+2\delta}\right)^{-1}E\left(\max\left(1, \eta_m^{2+2\delta}\right)\right)^{(2+2\delta)/(3+2\delta)}\\ &\qquad\times\left(\sum_{i=m+1}^{n} E\left(|X_i|^{2+2\delta}\right)\right)^{1/(3+2\delta)}\\ &\quad\leq C_\delta C_\delta(\eta)\left(1 + |t|^{2+2\delta}\right)^{-1}Q_{n,2\delta}(m)^{1/(3+2\delta)}.\end{aligned}$$

From

$$\left(1 + |\eta_m^{-1}t|^{2+2\delta}\right)^{-1}L_{n,2\delta}(m) \leq C_\delta\left(1 + |t|^{2+2\delta}\right)^{-1}\sum_{i=m+1}^{n} E\left(|X_i|^{2+2\delta}|\mathscr{G}_m\right),$$

we obtain [recall $Q_{n,2\delta}(m) \leq 1$ by assumption]

$$\begin{aligned}(4.11)\qquad &E\left(\left(1 + |\eta_m^{-1}t|^{2+2\delta}\right)^{-1}L_{n,2\delta}(m)\right)\\ &\quad\leq C_\delta\left(1 + |t|^{2+2\delta}\right)^{-1}Q_{n,2\delta}(m)^{1/(3+2\delta)}.\end{aligned}$$

To deal with the remaining two summands on the right side of (4.8), we fix a version $f_i(x_1,\dots,x_{i-1})$ of $E(X_i^2|(X_1,\dots,X_{i-1}) = (x_1,\dots,x_{i-1}))$. Employing standard measure theoretic arguments, it can be shown that for $i = m+1,\dots,n$ and P-almost all ω the random variable $f_i(X_1(\omega),\dots,X_m(\omega),Y_{m+1},\dots,Y_{i-1})$ is a version of $E_m(Y_i^2|\mathscr{H}_{i-1})$, and this entails P-a.s.

$$\begin{aligned}(4.12)\qquad &E_m\left(\left|\sum_{i=m+1}^{n} E_m\left(Y_i^2|\mathscr{H}_{i-1}\right) - \eta_m^2\right|^{1+\delta}\right)\\ &\quad= E\left(\left|\sum_{i=m+1}^{n} E\left(X_i^2|\mathscr{G}_{i-1}\right) - \eta_m^2\right|^{1+\delta}|\mathscr{G}_m\right).\end{aligned}$$

Similarly, as in (4.9) we have

$$\begin{aligned}&\left(1+|\eta_m^{-1}t|^{2+2\delta}\right)^{-1}N_{n,2\delta}(m)^{1/(3+2\delta)}\\&\quad\le C_\delta\left(1+|t|^{2+2\delta}\right)^{-1}\eta_m^{2+2\delta-(2+2\delta)/(3+2\delta)}\\&\quad\times E_m\left(\left|\sum_{i=m+1}^{n}E_m\left(Y_i^2|\mathscr{H}_{i-1}\right)-\eta_m^2\right|^{1+\delta}\right)^{1/(3+2\delta)},\end{aligned}$$

hence by integration and application of Hölder's inequality together with (4.12)

$$\begin{aligned}&E\left(\left(1+|\eta_m^{-1}t|^{2+2\delta}\right)^{-1}N_{n,2\delta}(m)^{1/(3+2\delta)}\right)\\&\quad\le C_\delta\left(1+|t|^{2+2\delta}\right)^{-1}E\left(\eta_m^{2+2\delta}\right)^{(2+2\delta)/(3+2\delta)}\\&\quad\times E\left(\left|\sum_{i=m+1}^{n}E\left(X_i^2|\mathscr{G}_{i-1}\right)-\eta_m^2\right|^{1+\delta}\right)^{1/(3+2\delta)}\\&\quad\le C_\delta C_\delta(\eta)\left(1+|t|^{2+2\delta}\right)^{-1}Q_{n,2\delta}(m)^{1/(3+2\delta)}.\end{aligned}\tag{4.13}$$

Finally, analogously to (4.11),

$$\begin{aligned}&E\left(\left(1+|\eta_m^{-1}t|^{2+2\delta}\right)^{-1}N_{n,2\delta}(m)\right)\\&\quad\le C_\delta\left(1+|t|^{2+2\delta}\right)^{-1}Q_{n,2\delta}(m)^{1/(3+2\delta)}.\end{aligned}\tag{4.14}$$

Combining the estimates (4.7), (4.8), (4.10), (4.11), (4.13) and (4.14), we obtain for $t \in R$ with $|t| \ge 1$,

$$\begin{aligned}&\left|P(S_n-S_m\le t)-E\left(\Phi\left(\eta_m^{-1}t\right)\right)\right|\\&\quad\le C_\delta C_\delta(\eta)\left(1+|t|^{2+2\delta}\right)^{-1}Q_{n,2\delta}(m)^{1/(3+2\delta)}.\end{aligned}\tag{4.15}$$

The uniform estimate (1.1) implies for all $t \in R$,

$$\left|P(S_n-S_m\le t|\mathscr{G}_m)-\Phi\left(\eta_m^{-1}t\right)\right|\le C_\delta\left(L_{n,2\delta}(m)+N_{n,2\delta}(m)\right)^{1/(3+2\delta)},$$

and integrating this inequality combined with arguments similar to those used previously leads to

$$\left|P(S_n-S_m\le t)-E\left(\Phi\left(\eta_m^{-1}t\right)\right)\right|\le C_\delta C_\delta(\eta)Q_{n,2\delta}(m)^{1/(3+2\delta)},$$

which means that (4.15) holds for all $t \in R$. Consequently, for all $x \in R$ and $a > 0$,

$$\mathrm{I}\le C_\delta C_\delta(\eta)\left(\max_{\beta=\pm1}\left(1+|x+\beta a|^{2+2\delta}\right)^{-1}\right)Q_{n,2\delta}(m)^{1/(3+2\delta)}.\tag{4.16}$$

Substituting (4.2), (4.3), (4.6) and (4.16) into the right side of (4.1), we obtain for all $x \in R$ and $a > 0$,

$$\begin{aligned} &\left|P(S_n \le x) - E\big(\Phi(\eta^{-1}x)\big)\right| \\ &\quad \le C_\delta C_\delta(\eta)\bigg(a^{-2-2\delta}Q_{n,2\delta}(m) + a\big(1 + m(a,x)^{3+2\delta}\big)^{-1} \\ &\qquad + \max_{\beta=\pm 1}\big(1 + |x + \beta a|^{2+2\delta}\big)^{-1}Q_{n,2\delta}(m)^{1/(3+2\delta)}\bigg). \end{aligned} \tag{4.17}$$

For $x \in R$ with $|x| \ge 1$ we set $a = Q_{n,2\delta}(m)^{1/(3+2\delta)}|x|/2$ so that $0 < a \le |x|/2$ on account of $Q_{n,2\delta}(m) \le 1$, hence $|x \pm a| \ge |x|/2$ and $m(a,x) \ge |x|/2$, and substituting a into (4.17) yields (1.3). To finish the proof, it suffices again to derive a uniform bound. Observe that (4.17) implies

$$\begin{aligned} &\left|P(S_n \le x) - E\big(\Phi(\eta^{-1}x)\big)\right| \\ &\quad \le C_\delta C_\delta(\eta)\big(a^{-2-2\delta}Q_{n,2\delta}(m) + a + Q_{n,2\delta}(m)^{1/(3+2\delta)}\big), \end{aligned}$$

for all $x \in R$ and $a > 0$. Taking $a = Q_{n,2\delta}(m)^{1/(3+2\delta)}$ completes the proof of (1.3). To see that $Q_{n,2\delta}(m)$ can be replaced by $\tilde{Q}_{n,2\delta}(m)$, notice that $E(|\sum_{i=1}^m X_i^2|^{1+\delta})$ enters the bound only through inequality (4.2). Applying Theorem 2.11 in Hall and Heyde (1980) yields

$$\mathrm{IV} \le C_\delta a^{-2-2\delta}\left(E\big(V_m^{2+2\delta}\big) + \sum_{i=1}^m E\big(|X_i|^{2+2\delta}\big)\right) \le C_\delta a^{-2-2\delta}\tilde{Q}_{n,2\delta}(m),$$

instead of (4.2). Similarly, $Q_{n,2\delta}(m)$ can be replaced by $\tilde{Q}_{n,2\delta}(m)$ on the right side of (4.6) and (4.16). This is clear for (4.6), and for (4.16) observe

$$\begin{aligned} &E\left(\left|\sum_{i=m+1}^n E\big(X_i^2|\mathscr{G}_{i-1}\big) - \eta_m^2\right|^{1+\delta}\right) \\ &\quad \le C_\delta\big(E\big(V_m^{2+2\delta}\big) + E\big(|V_n^2 - \eta^2|^{1+\delta}\big) + E\big(|\eta^2 - \eta_m^2|^{1+\delta}\big)\big) \le C_\delta\tilde{Q}_{n,2\delta}(m), \end{aligned}$$

if $\tilde{Q}_{n,2\delta}(m) \le 1$. Consequently, (4.17) holds for all $x \in R$ and $a > 0$ with $\tilde{Q}_{n,2\delta}(m)$ instead of $Q_{n,2\delta}(m)$, and this implies the statement about $\tilde{Q}_{n,2\delta}(m)$ in Theorem 2 as (4.17) implies (1.3). □

Acknowledgments. The authors are indebted to the Editor and the referee for their comments and suggestions, which led to an improvement of the presentation. The first author also thanks Professors V. Yu. Bentkus and A. Račkauskas for a useful discussion.

REFERENCES

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

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

BOSE, A. (1986b). Certain non-uniform rates of convergence to normality for a restricted class of martingales. *Stochastics* **16** 279–294.

ERICKSON, R. V., QUINE, M. P. and WEBER, N. C. (1979). Explicit bounds for the departure from normality of sums of dependent random variables. *Acta Math. Acad. Sci. Hungar.* **34** 27–32.

HAEUSLER, E. (1984). A note on the rate of convergence in the martingale central limit theorem. *Ann. Probab.* **12** 635–639.

HAEUSLER, E. (1988). On the rate of convergence in the central limit theorem for martingales with discrete and continuous time. *Ann. Probab.* **16** 275–299.

HALL, P. (1978). The convergence of moments in the martingale central limit theorem. *Z. Wahrsch. verw. Gebiete* **44** 253–260.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application.* Academic, New York.

HALL, P. and HEYDE, C. C. (1981). Rates of convergence in the martingale central limit theorem. *Ann. Probab.* **9** 395–404.

HEYDE, C. C. and BROWN, B. M. (1970). On the departure from normality of a certain class of martingales. *Ann. Math. Statist.* **41** 2161–2165.

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

NAKATA, T. (1976). On the rate of convergence in mean central limit theorem for martingale differences. *Rep. Statist. Appl. Res. Un. Japan. Sci. Engrs.* **23** 10–15.

MATHEMATICAL INSTITUTE
UNIVERSITY OF MUNICH
THERESIENSTRASSE 39
D-8000 MUNICH 2
WEST GERMANY