A BERRY–ESSEEN BOUND FOR STUDENT'S STATISTIC IN THE NON-I.I.D. CASE

V. Bentkus[1] M. Bloznelis[1"] F. Götze[1"]

Abstract. We establish a Berry–Esséen bound for Student's statistic for independent (non-identically) distributed random variables. The bound is very general and sharp. In particular, it implies an estimate like to the classical Berry–Esséen bound. In the i.i.d. case it yields as well the best known sufficient conditions for the Central Limit Theorem for studentized sums. For non-i.i.d. random variables the bound shows that the Lindeberg condition is sufficient for the Central Limit Theorem for studentized sums.

1. Introduction and Results

Let $X_1, \dots, X_N$ be independent random variables. Let

$$\mathbf{t} = \overline{X}/\hat{\sigma}$$

denote the Student statistic, where

$$\overline{X} = N^{-1}(X_1 + \cdots + X_N), \qquad \hat{\sigma}^2 = N^{-1}\sum\nolimits_{i=1}^N (X_i - \overline{X})^2.$$

We shall investigate the rate of the normal approximation of $\sqrt{N}\,\mathbf{t}$ when the random variables $X_1, \dots, X_N$ are non-identically distributed. Define

$$\delta_N = \sup\nolimits_x \big|\, \mathbf{P}\,(\sqrt{N}\,\mathbf{t} < x) - \Phi(x)\,\big|,$$

where $\Phi(x)$ is the standard normal distribution function. Denote

$$B_N^2 = \sum\nolimits_{i=1}^N \operatorname{Var} X_i \qquad \text{and} \qquad L_N = B_N^{-3}\sum\nolimits_{i=1}^N \mathbf{E}\,|X_i|^3.$$

1991 *Mathematics Subject Classification.* Primary 62E20; secondary 60F05.

Key words and phrases. Student's statistic, Berry–Esséen bound, non-identically distributed random variables, convergence rate, Central Limit Theorem, self-normalized sums.

Theorem 1.1. *Let $X_1, \dots, X_N$ have mean zero and $B_N > 0$. Then there exists an absolute constant $c > 0$ such that*

$$\delta_N \le c L_N. \tag{1.1}$$

Inequality (1.1) is an extension to the case of studentized sums of the classical bound by Esséen (1945)

$$\sup_x \; \mathbf{P}\,(X_1 + \cdots + X_N < x B_N) - \Phi(x) \; \le c L_N.$$

Theorem 1.1 is a consequence of a general result formulated as Theorem 1.2.

For a given number $V > 0$, define the truncated random variables

$$Z_i = V^{-1} X_i \, \mathbb{I}\{X_i^2 \le V^2\}, \qquad \text{for } 1 \le i \le N, \tag{1.2}$$

where $\mathbb{I}\{A\}$ denotes the indicator function of the event A. Denote

$$M^2 = \sum\nolimits_{i=1}^N \operatorname{Var} Z_i, \qquad \varPi = \sum\nolimits_{i=1}^N \mathbf{P}\left(X_i^2 > V^2\right), \qquad \varLambda = M^{-3} \sum\nolimits_{i=1}^N \mathbf{E}\,|Z_i|^3$$

and

$$\varUpsilon_0 = M^{-1} \; \sum\nolimits_{i=1}^N \mathbf{E}\, Z_i \; , \quad \varUpsilon_1 = M^{-1} \sum\nolimits_{i=1}^N |\,\mathbf{E}\, Z_i|, \quad \varUpsilon_2 = M^{-2} \sum\nolimits_{i=1}^N (\,\mathbf{E}\, Z_i)^2.$$

Theorem 1.2. *Let $X_1, \dots, X_N$ be arbitrary independent random variables. Then there exists an absolute constant $c > 0$ such that, for any $V > 0$,*

$$\delta_N \le c\varPi + c\varUpsilon_0 + c\varUpsilon_2 + c\varLambda. \tag{1.3}$$

Example 3.1 in the Appendix shows that the term $\varUpsilon_2$ in the bound (1.3) is optimal.

Corollary 1.1. *Let independent random variables $X_1, \dots, X_N$ be arbitrary. Then there exists an absolute constant $c > 0$ such that, for any $V > 0$,*

$$\delta_N \le c\varPi + c\varUpsilon_1 + c\varLambda.$$

Let a denote the largest solution of the equation

$$a^2 = \mathbf{E}\, X_1^2 \, \mathbb{I}\{X_1^2 \le a^2 N\}. \tag{1.4}$$

Theorem 1.2 yields the following Berry–Esséen bound for the Student statistic in the i.i.d. case.

Corollary 1.2 (Bentkus and Götze (1994))**.** *Let $X_1, \dots, X_N$ be i.i.d. random variables. Assume that the largest solution a of the equation* (1.4) *is positive. There exists an absolute constant $c > 0$ such that*

$$\delta_N \le c N \, \mathbf{P} \, (X_1^2 \ge a^2 N) + c N \, | \, \mathbf{E} \, Z_1 | + c N \, \mathbf{E} \, |Z_1|^3, \tag{1.5}$$

where

$$Z_1 = \left(a^2 N \right)^{-1/2} X_1 \, \mathbb{I} \, \{ X_1^2 \le a^2 N \}.$$

The explicit bounds of Theorems 1.1 and 1.2 and Corollaries 1.1 and 1.2 are applicable for any independent random variables $X_1, \dots, X_N$. The random variables may depend on any parameters (for instance, on N). An application of a chosen bound reduces to estimation of moments or tails of random variables X_i and to summation over $1 \le i \le N$ of expressions obtained since the bounds depend additively on a separate summand. The bounds are indeed sharp and a lot of consequences may be derived, for example, the Central Limit Theorem for the Student test under the most precise known conditions follows (see the discussion at the bottom of the section).

All the results remain valid if instead of the Student statistic $\sqrt{N} \, \mathbf{t}$ one considers the selfnormalized sums defined as

$$S = \frac{X_1 + \dots + X_N}{\sqrt{X_1^2 + \dots + X_N^2}} \, .$$

In particular, Theorems 1.1, 1.2 and Corollaries 1.1–1.3 hold with δ_N replaced by $\sup_x \big| \, \mathbf{P} \, (S < x) - \Phi(x) \big|$. Indeed, (see the argument in the proof of Lemma 2.3 below)

$$\sup_x \big| \, \mathbf{P} \, (\sqrt{N} \, \mathbf{t} < x) - \, \mathbf{P} \, (S < x) \big| \le c \mathit{\Pi} + c \mathit{\Upsilon}_0 + c \mathit{\Upsilon}_2 + c \mathit{\Lambda}.$$

The convergence of selfnormalized sums was investigated by a number of authors, see, for instance, Efron (1969), Logan, Mallows, Rice and Shepp (1973), LePage, Woodroofe and Zinn (1981), Griffin and Kuelbs (1989, 1991), Hahn, Kuelbs and Weiner (1990), Griffin and Mason (1991).

Convergence rates and Edgeworth expansions for Student's and related statistics were considered by Chung (1946), Chibisov (1980), Helmers and van Zwet (1982), van Zwet (1984), Helmers (1985), Slavova (1985), Bhattacharya and Ghosh (1986), Hall (1987, 1988), Praskova (1989), Friedrich (1989), Bhattacharya and Denker (1990), Bentkus, Götze and van Zwet (1994), Bentkus and Götze (1994), etc.

For i.i.d. $X_1, \dots, X_N$, the rate $O(N^{-1/2})$ under the 3-rd moment condition was obtained by Slavova (1985). The bound $\delta_N \le c N^{-1/2} \beta_3 / \sigma^3$, where $\beta_s = \mathbf{E} \, |X_1|^s$ and $\sigma^2 = \mathbf{E} \, X_1^2$, follows from a general result of Bentkus and Götze (1994). A related result $O(N^{-1/2})$ under the additional condition: there exists $y \in \mathbf{R}$ such that

$$\mathbf{E} \, p(y) \big(1 - p(y) \big) > 0, \qquad \text{where} \quad p(y) = \mathbf{P} \left\{ X - y \quad |X - y| \right\},$$

was proved by Hall (1988) (see Bai and Shepp (1994) for a discussion of such condition).

Most of the papers mentioned above deal with the case of independent and *identically distributed* random variables. Friedrich (1989) constructed a Berry–Esséen bound for a general class of statistics without the i.i.d. assumption. In the i.i.d. case the result of Friedrich (1989) yields $\delta_N = O(N^{-1/2})$ provided $\beta_{10/3}$ is finite, whereas our result implies $\delta_N \le c N^{-1/2} \beta_3 / \sigma^3$.

In the remaining part of the Introduction we shall discuss sufficient conditions for the Central Limit Theorem (CLT) for studentized sums in i.i.d. and non-i.i.d. cases. We say that a sequence of studentized sums satisfies the CLT if it converges in distribution to the standard normal distribution. We are going to apply the bounds in the particular case when observations $X_1, \dots, X_N$ are taken from a (fixed) infinite sequence $X_1, X_2, \dots$ of independent random variables and $\mathbf{t} = \mathbf{t}(X_1, \dots, X_N)$. By $\xi, \xi_1, \xi_2, \dots$ we shall denote i.i.d standard normal variables.

The CLT for i.i.d. random variables. Let $X_1, X_2, \dots$ be i.i.d. Then the right-hand side of (1.5) tends to zero as $N \to \infty$ if and only if X_1 belongs to the domain of attraction of the standard normal law, see Bentkus and Götze (1994). Thus (1.5) yields the weakest known sufficient condition for the Central Limit Theorem for Student's statistic (Maller (1981), Csörgő and Mason (1987)). There is a conjecture (communicated for us by D.M. Mason and seemingly difficult to prove) that the condition is necessary as well.

The CLT for non-i.i.d. random variables. Let $X_1, X_2, \dots$ denote a sequence of independent random variables with mean zero and finite variances $\sigma_i^2 = \mathbf{E}\, X_i^2 < \infty$. It is well known (Feller (1971)) that if X_i, $i \ge 1$, have mean zero then the Lindeberg condition

$$\forall\, \varepsilon > 0 \qquad\qquad B_N^{-2} \sum\nolimits_{i=1}^{N} \mathbf{E}\, X_i^2\, \mathbb{I}\, \{X_i^2 \ge \varepsilon B_N^2\} \to 0 \tag{1.6}$$

is sufficient for

$$\mathcal{L}\ \ B_N^{-1}(X_1 + \dots + X_N) \ \ \to N(0,1), \tag{1.7}$$

and it is necessary for (1.7) provided that

$$\max\nolimits_{1 \le i \le N}\ \sigma_i^2 / B_N^2 \to 0. \tag{1.8}$$

Corollary 1.3. *Assume that $X_1, X_2, \dots$ have mean zero and satisfy the Lindeberg condition* (1.6). *Then $\delta_N \to 0$, that is, $\mathcal{L}(\sqrt{N}\,\mathbf{t})$ converge to the standard normal distribution.*

Remark 1.1. Corollary 1.3 tells us that the Lindeberg condition is sufficient for

$$\mathcal{L}(\sqrt{N}\,\mathbf{t}) \to N(0,1). \tag{1.9}$$

It is interesting to note that the Lindeberg condition (1.6) is not necessary for (1.9) even if (1.8) holds. More specifically, there exists a sequence (see Example 1.1 below) of centered independent random variables which satisfies (1.8) and (1.9) and which fails to satisfy (1.6).

Example 1.1 (cf. Feller (1971), p. 521). Let $\tau_1, \tau_2, \ldots$ be a sequence of independent symmetric random variables such that

$$\mathbf{P}\left(|\tau_i| = i^2\right) = 1 - \mathbf{P}\left(|\tau_i| = 0\right) = i^{-2} \qquad \text{for all} \quad i \geq 1.$$

Assume that i.i.d. standard normal random variables $\xi_1, \xi_2, \ldots$ are independent of $\tau_1, \tau_2, \ldots$. It is easy to see that independent centered random variables $X_i = \tau_i + \xi_i$ satisfy (1.7) and (1.9) and fail to satisfy the Lindeberg condition. In this case

$$B_N^{-1}(X_1 + \cdots + X_N) \to_P 0.$$

Notice as well that in this particular case (1.9) follows from the Borel–Cantelli lemma (or Corollary 1.1 with $V = V(N) = \sqrt{N}$).

Thus we conclude that in case of finite variance and centered and asymptotically negligible summands in sense of (1.8) the studentization of the sum may lead to better results than normalization by B_N. However, in general a different situation may arise. In the following example we construct a sequence $X_1, X_2, \ldots$ of independent (non-centered) random variables such that (1.7) holds and $\mathcal{L}\big(\sqrt{N}\,\mathbf{t}\big) \nrightarrow N(0,1)$.

Example 1.2. Define

$$X_i = i^{-1/2}\xi_i + (-1)^i i^{-p}, \qquad p > 0. \tag{1.10}$$

It is easy to see that $B_N^2 \sim \ln(N+2)$. Furthermore, $X_1 + \cdots + X_N$ is distributed as $B_N \xi_1 + \sum_{i=1}^N (-1)^i i^{-p}$. Therefore

$$\mathcal{L}\ \ (X_1 + \cdots + X_N)/\sqrt{\ln(N+2)}\ \ \to \mathcal{L}(\xi)$$

converges to the standard normal distribution.

Let $V^2 = V^2(N) = \log(N+2)$. In what follows we show that, for the sequence (1.10),

$$\mathcal{L}\big(\sqrt{N}\,\mathbf{t}\big) \to N(0,1) \tag{1.11}$$

if and only if

$$\sum\nolimits_{i=1}^{N} (\,\mathbf{E}\, Z_i)^2 \to 0, \tag{1.12}$$

where r.v. Z_i are defined in (1.2). A calculation shows that all terms other than $\sum_{i=1}^N (\,\mathbf{E}\, Z_i)^2$ in the right-hand side of (1.3) tend to 0 as $N \to \infty$, for each $p > 0$.

Hence (1.12) implies (1.11). Notice that (1.12) holds if and only if $p > 1/2$. Therefore, in order to prove the converse implication, (1.11) $\Rightarrow$ (1.12) it suffices to show that

$$\sqrt{N}\,\mathbf{t} \xrightarrow{P} 0, \qquad \text{for} \quad 0 < p < 1/2, \tag{1.13}$$

and that

$$\mathcal{L}\big(\sqrt{N}\,\mathbf{t}\big) \to N(0, 1/2), \qquad \text{for} \quad p = 1/2. \tag{1.14}$$

Proofs of (1.13) and (1.14) are given in Appendix.

Remark 1.2. Example 1.2 demonstrates that the limiting behavior of the distribution of studentized sums crucially depends on $\sum_{i=1}^N (\mathbf{E}\, Z_i)^2$. This observation can be generalized. Let $X_1, X_2, \ldots$ be a sequence of independent random variables. Assume that, for a sequence of positive numbers $V(N)$,

$$\forall \varepsilon > 0, \qquad \max_{1 \le k \le N} \mathbf{P}\ (|X_k| > \varepsilon V(N)) \to 0 \tag{1.15}$$

and

$$\mathcal{L}\big((X_1 + \cdots + X_N)/V(N)\big) \to N(0,1). \tag{1.16}$$

Then (see Petrov (1975)), for each $\varepsilon > 0$,

$$\sum_{i=1}^N \mathbf{P}\left(X_i^2 \ge \varepsilon^2 V^2(N)\right) \to 0, \quad \sum_{i=1}^N \mathbf{E}\, Z_i \to 0, \quad M \to 1, \quad \sum_{i=1}^N \mathbf{E}\, |Z_i - \mathbf{E}\, Z_i|^3 \to 0$$

as $N \to \infty$. Hence if (1.15) and (1.16) hold then (1.12) implies (1.11). A simple analysis shows that the converse implication (1.11) $\Rightarrow$ (1.12) is also true provided (1.15) and (1.16) hold. A similar effect was noticed by Robbins (1948). He considered the distribution of the Student statistic based on the independent normal observations $X_i \sim N(\mu_i, \sigma^2)$, $\sigma^2 > 0$, with $\mu_i = \mathbf{E}\, X_i$ such that $\sum_{i=1}^N \mu_i = 0$. In particular, Robbins (1948) showed that in this case the distribution function depends only on N and $\lambda = \sum_{i=1}^N \mu_i^2/(2\sigma^2)$. In his results λ plays a role similar to that of $\sum_{i=1}^N (\mathbf{E}\, Z_i)^2$ in our considerations.

2. Proofs

The proofs are close to those given in Bentkus and Götze (1994). By means of truncation we replace the Student statistic $\mathbf{t} = \mathbf{t}(X_1, \ldots, X_N)$ by a smooth function $S = S(X_1, \ldots, X_N)$ of the sample (see Lemma 2.3 below). Then we apply Esséen's smoothing inequality. When estimating the difference between the characteristic functions $\varphi(t) = \exp\{-t^2/2\}$ and $f(t) = \mathbf{E}\, \exp\{itS\}$ we split f into a conditional product of two characteristic functions (different for different values of t) of conditionally independent random variables. This type of approach has been used by Helmers and van Zwet (1982), van Zwet (1984) and was extended in Götze and van Zwet (1992), Bentkus, Götze and van Zwet (1994) to the case of general symmetric statistics of i.i.d. random variables. The crucial step of the proof in the non-identically distributed case is to bound the characteristic function of the sum (which may contain, for example, only one summand) by a characteristic function which splits into a product. For this purpose we use a randomization by means of Bernoulli random variables.

The Section is organized as follows. First we shall derive Corollary 1.1 from Theorem 1.2. Then we shall show that Corollary 1.1 implies Theorem 1.1, Corollary 1.2 and Corollary 1.3. The proof of these implications is simple. The real problem presents the proof of the main result—Theorem 1.2. Lemmas 2.1–2.4 are needed for the proof of Theorem 1.2. In particular, the preparatory Lemma 2.4 may serve a simplified version of the proof of Theorem 1.2: the result of the lemma supplies a sufficiently precise bound for the difference of the characteristic functions for Fourier frequencies $|t| \le 1$; for frequencies $|t| \ge 1$ much more elaborated techniques is needed in order to ensure sufficiently fast decrease (as $|t| \to \infty$) of the bound.

Proof of Corollary 1.1. We shall derive the Corollary from Theorem 1.2. Let us consider two incompatible cases: a) there exists $1 \le i \le N$ such that $M \le |\mathbf{E}\, Z_i|$; b) for all $1 \le i \le N$, the inverse inequality $|\mathbf{E}\, Z_i| < M$ holds. In the case a) the Corollary follows from the trivial bound $\delta_N \le 1$ since the inequality $M \le |\mathbf{E}\, Z_i|$ implies $\delta_N \le 1 \le M^{-1}|\mathbf{E}\, Z_i| \le \varUpsilon_1$. In the case b) we have $M^{-1}|\mathbf{E}\, Z_i| \le 1$, for all $1 \le i \le N$. Consequently, $\varUpsilon_2 \le \varUpsilon_1$. Obviously $\varUpsilon_0 \le \varUpsilon_1$, and the result follows from the bound of Theorem 1.2.

Proof of Theorem 1.1. If $L_N \ge 1/4$, then the desired bound follows from the trivial estimate $\delta_N \le 1$. Therefore we shall assume that $L_N \le 1/4$. Let us apply the bound $\delta_N \le c\varPi + c\varUpsilon_1 + c\varLambda$ of Corollary 1.1. Choose $V = B_N$. By the Chebyshev inequality, $\varPi \le L_N$. Obviously $\varLambda \le M^{-3} L_N$. Recall that $\mathbf{E}\, X_i = 0$, and therefore we have $|\mathbf{E}\, Z_i| = B_N^{-1}|\mathbf{E} X_i \mathbb{I}\{X_i^2 > B_N^2\}|$. Thus, by the Chebyshev inequality, we obtain $\varUpsilon_1 \le M^{-1} L_N$. To conclude the proof of the Theorem, it suffices to show that $M^2 \ge 1/2$. We have

$$\operatorname{Var} Z_i = \mathbf{E}\, Z_i^2 - (\mathbf{E}\, Z_i)^2 \ge B_N^{-2}\, \mathbf{E}\, X_i^2 - B_N^{-2}\, \mathbf{E}\, X_i^2\, \mathbb{I}\{X_i^2 \ge B_N^2\} - |\mathbf{E}\, Z_i|$$

since $|Z_i| \le 1$. Applying the Chebyshev inequality and summing over i, we obtain

$$M^2 \ge 1 - L_N - \sum\nolimits_{i=1}^N |\mathbf{E}\, Z_i| \ge 1 - 2L_N \ge 1/2$$

since, again by the Chebyshev inequality, $\sum_{i=1}^N |\mathbf{E}\, Z_i| \le L_N$, and $L_N \le 1/4$.

Proof of Corollary 1.2. Choose $V^2 = a^2 N$ and apply Corollary 1.1. Due to the i.i.d. assumption, we have

$$\delta_N \le c N \mathbf{P} (X_1^2 \ge a^2 N) + c N M^{-1} | \mathbf{E} Z_1| + c N M^{-3} \mathbf{E} |Z_1|^3,$$

and it remains to show that $M^2 \ge 1/2$. We can assume that $N | \mathbf{E} Z_1| \le 1/2$ since otherwise the trivial bound $\delta_N \le 1$ implies the result. Therefore, due to the choice of a and $|Z_i| \le 1$, we have

$$M^2 = N \mathbf{E} Z_i^2 - N (\mathbf{E} Z_i)^2 = 1 - N (\mathbf{E} Z_i)^2 \ge 1 - N | \mathbf{E} Z_i| \ge 1/2.$$

Proof of Corollary 1.3. We have to show that $\delta_N \to 0$ as $N \to \infty$. Let us choose $V = B_N$, and apply the bound $\delta_N \le c\varPi + c\varUpsilon_1 + c\varLambda$ of Corollary 1.1. Then it is sufficient to prove that $\varPi \to 0$, $\varUpsilon_1 \to 0$ and $\varLambda \to 0$.

Denote

$$\lambda_\varepsilon = B_N^{-2} \sum\nolimits_{i=1}^N \mathbf{E} X_i^2 \,\mathbb{I}\, \{X_i^2 \ge \varepsilon B_N^2\}.$$

Then, for any $\varepsilon > 0$, the Lindeberg condition (1.6) yields $\lambda_\varepsilon \to 0$, as $N \to \infty$.

The Chebysvev inequality implies $\varPi \le \lambda_1 \to 0$.

Using $\mathbf{E} X_i = 0$, the inequality $| \mathbf{E} Z_i| \le 1$ and the Chebyshev inequality, we have

$$\operatorname{Var} Z_i = \mathbf{E} Z_i^2 - (\mathbf{E} Z_i)^2 \ge B_N^{-2} \mathbf{E} X_i^2 - 2 B_N^{-2} \mathbf{E} X_i^2 \,\mathbb{I}\, \{X_i^2 \ge B_N^2\},$$

whence, summing over i, we obtain $M^2 \ge 1 - 2\lambda_1$. Consequently, $M \ge 1/2$, for sufficiently large N.

To bound $\varUpsilon_1$, use $\mathbf{E} X_i = 0$ and the Chebyshev inequality. Then $\varUpsilon_1 \le M^{-1} \lambda_1$, and $\varUpsilon_1 \to 0$ since $M \ge 1/2$ and $\lambda_1 \to 0$.

It remains to show that $\varLambda \to 0$. For given $0 < \varepsilon < 1$, we have

$$\begin{aligned} \mathbf{E} |Z_i|^3 &= B_N^{-3} \mathbf{E} |X_i|^3 \,\mathbb{I}\, \{X_i^2 \le \varepsilon^2 B_N^2\} + B_N^{-3} \mathbf{E} |X_i|^3 \,\mathbb{I}\, \{\varepsilon^2 B_N^2 \le X_i^2 \le B_N^2\} \\ &\le \varepsilon B_N^{-2} \mathbf{E} X_i^2 + B_N^{-2} \mathbf{E} X_i^2 \,\mathbb{I}\, \{X_i^2 \ge \varepsilon^2 B_N^2\}. \end{aligned}$$

Multiplying by M^{-3} and summing over i, we get $\varLambda \le \varepsilon M^{-3} + \lambda_\varepsilon M^{-3}$. Recall that $M \ge 1/2$, for sufficiently large N. Therefore we have $0 \le \limsup_{N\to\infty} \varLambda \le 8\varepsilon$, and $\lim_{N\to\infty} \varLambda = 0$ follows, since $\varepsilon > 0$ is arbitrary.

It remains to prove Theorem 1.2. Introduce the notation

$$\begin{aligned} Y_i &= M^{-1}(Z_i - \mathbf{E} Z_i), && 1 \le i \le N, \\ Y &= Y_1 + \dots + Y_N, && \\ \eta &= \eta_1 + \dots + \eta_N, && \text{where } \eta_i = Y_i^2 - \mathbf{E} Y_i^2. \end{aligned}$$

The random variables $Y_1, \dots, Y_N$ (resp., $\eta_1, \dots, \eta_N$) have means zero and are independent. Furthermore,

$$\sum_{i=1}^{N} \mathbf{E}\, Y_i^2 = 1, \qquad \sum_{i=1}^{N} \mathbf{E}\, |Y_i|^3 \le 8\,\varLambda.$$

We shall assume throughout the rest of the paper that for a small absolute constant $c_0 > 0$,

$$\varLambda \le c_0, \qquad \varUpsilon_2 \le c_0, \qquad \varUpsilon_0 \le c_0, \tag{2.1}$$

since otherwise the result follows from the obvious estimate $\delta_N \le 1$.

Let $\alpha = (\alpha_1, \dots, \alpha_N)$ denote a sequence of i.i.d. Bernoulli random variables independent of $Y_1, \dots, Y_N$ and such that

$$\mathbf{P}\,(\alpha_1 = 1) = 1 - \mathbf{P}\,(\alpha_1 = 0) = m \qquad \text{with some } \ m \le 1.$$

Lemma 2.1. *Assume that* (2.1) *holds. Denote*

$$\varkappa_i = \alpha_i \left(\eta_i + M^{-1} Y_i\, \mathbf{E}\, Z_i \right), \qquad \text{for } \ 1 \le i \le N.$$

Then there exists an absolute constant $c > 0$ such that

$$\mathbf{E} \Big| \sum\nolimits_{i=1}^{N} Y_i\, \varkappa_i \Big| \le c\, m\, \varLambda, \qquad \mathbf{E} \Big| \sum\nolimits_{i=1}^{N} \varkappa_i \Big|^{3/2} \le c\, m\, \varLambda.$$

The U-statistic

$$U = \sum_{\{1 \le i, j \le N,\ i \ne j\}} \alpha_i\, Y_i\, \varkappa_j \qquad \textit{satisfies} \quad \mathbf{E}\, |U|^{3/2} \le c m^{7/4}\, \varLambda.$$

Proof of Lemma 2.1 is given in the Appendix.

Lemma 2.2. *Let β be a random variable with finite third moment. Then*

$$\big|\, \mathbf{E}\, \exp\{it\beta\} \big|^2 \le 1 - t^2\, \mathbf{E}\, \beta^2 + \frac{4\,|t|^3}{3}\, \mathbf{E}\, |\beta|^3. \tag{2.2}$$

Let α_1 be a Bernoulli random variable with $\mathbf{P}\,(\alpha_1 = 1) = 1 - \mathbf{P}\,(\alpha_1 = 0) = m$. *Then*

$$\mathbf{E}\ \ 1 - \alpha_1 t^2\, \mathbf{E}\, \beta^2 + \alpha_1\, \frac{4\,|t|^3}{3}\, \mathbf{E}\, |\beta|^3\ \ ^{1/2} \le \exp\Big\{ - \frac{m\, t^2}{2}\, \mathbf{E}\, \beta^2 + \frac{2\, m\, |t|^3}{3}\, \mathbf{E}\, |\beta|^3 \Big\}. \tag{2.3}$$

Proof of Lemma 2.2. Proof of (2.2) is easy and can be found in Petrov (1975). Formula (2.3) is a simple consequence of Jensen's inequality.

Let $g : R \to R$ denote a function which is infinitely many times differentiable with bounded derivatives and such that

$$\frac{1}{8} \le g(x) \le 2, \ \text{ for } \ x \in R, \qquad \text{and} \qquad g(x) = \frac{1}{\sqrt{|x|}}\,, \ \text{ for } \ \frac{1}{2} \le x \le \frac{7}{4}\,.$$

Define the statistic $S = Y g(1 + \eta + 2A + \varUpsilon_2)$, where $A = M^{-1} \sum_{i=1}^{N} Y_i\, \mathbf{E}\, Z_i$.

Lemma 2.3. *Assume that* (2.1) *holds. There exists an absolute constant $c > 0$ such that*

$$\delta_N \le \sup_x \big|\, \mathbf{P}\,(S < x) - \Phi(x)\big| + c\varPi + c\varUpsilon_0 + c\varLambda. \tag{2.4}$$

Proof of Lemma 2.3 is given in the Appendix.

Recall that by $\xi, \xi_1, \xi_2, \dots$ we denote i.i.d. standard normal r.v. Define

$$\psi_i = (\,\mathbf{E}\,Y_i^2)^{1/2}\,\xi_i, \qquad\qquad 1 \le i \le N.$$

In what follows θ and $\theta_1, \theta_2, \dots$ denote generic real numbers such that $|\theta| \le 1$ and $|\theta_i| \le 1$. Furhtermore, $\vartheta, \vartheta_1, \vartheta_2, \dots$ denote i.i.d. uniformly distributed in $[0,1]$ random variables. For a vector valued differentiable function H, the mean value formula may be written as

$$H(b) - H(a) = \mathbf{E}\,H'\big(a + \vartheta\,(b-a)\big)\,(b-a), \qquad a, b \in \mathbf{R}.$$

We shall use the notation

$$\mathbf{E}_\zeta\, F(\zeta,\beta) = \mathbf{E}\,\big(F(\zeta,\beta)\,\big|\,\beta\big), \qquad \mathbf{E}^\zeta\, F(\zeta,\beta) = \mathbf{E}\,\big(F(\zeta,\beta)\,\big|\,\zeta\big),$$

for conditional expectations. For $Q = q_1 + \dots + q_N$, we shall denote $Q^i = Q - q_i$. Similarly, $Q^{i,j} = Q - q_i - q_j$.

Lemma 2.4. *Assume that* (2.1) *holds. Let* $H : \mathbf{R} \to \mathbf{C}$ *denote a bounded infinitely many times differentiable function with bounded derivatives. Then, for* $1 \le k \le N$,

$$\mathbf{E}\,H\big((Y_1 + \dots + Y_k)\,g(1 + \varkappa + \varUpsilon_2)\big) - \mathbf{E}\,H(\psi_1 + \dots + \psi_k) \;\le c\varGamma(\varLambda + \varUpsilon_2), \tag{2.5}$$

where

$$\varkappa = \sum\nolimits_{i=1}^{k} \varkappa_i, \qquad \varkappa_i = \eta_i + 2M^{-1}Y_i\,\mathbf{E}\,Z_i, \qquad 1 \le i \le N,$$

and

$$\varGamma = \;|H|_\infty + 1\;\; |H'|_\infty^{3/4} + |H'|_\infty + |H''|_\infty + |H'''|_\infty, \qquad |H|_\infty = \sup_x |H(x)|.$$

Proof of Lemma 2.4. We shall prove the lemma for $k = N$ only. Then the sums in (2.5) are equal to $\varUpsilon_2$ and $\varLambda$ respectively, $\mathcal{L}(\psi_1 + \dots + \psi_N) = \mathcal{L}(\xi)$, and it suffices to prove that

$$\begin{aligned} \mathbf{E}\,H\big(Y\,g(1 + \varkappa + \varUpsilon_2)\big) - \mathbf{E}\,H(Y\,g(1 + \varkappa) \;&\le c\varGamma\varUpsilon_2, \\ \big|\,\mathbf{E}\,H(Y\,g(1 + \varkappa)\big) - \mathbf{E}\,H(Y)\big| &\le c\varGamma\varLambda \end{aligned}$$

and

$$\big|\,\mathbf{E}\,H(Y) - \,\mathbf{E}\,H(\xi)\big| \le c\,\varGamma\,\varLambda.$$

Expanding g in powers of $\varUpsilon_2$ and then expanding H we obtain the first inequality. The proof of the third inequality is easy and does not differ from that of given in Bentkus et al. (1990) where the i.i.d. case was considered.

Let us prove the second inequality. Expanding g in powers of $\varkappa$ we have

$$Y\,g(1+\varkappa) = Y\,g(1) + Y\,\varkappa\,g'(1+\theta_1\varkappa) = Y + D\,g'(1+\theta_1\,\varkappa) + U\,g'(1+\theta_1\,\varkappa).$$

Here we split $Y\,\varkappa$ into the sum $D+U$, where

$$D = \sum\nolimits_{i=1}^{N} Y_i\,\varkappa_i, \qquad\qquad U = \sum\nolimits_{1\le i,j\le N,\ i\ne j} Y_i\,\varkappa_j.$$

Expanding H in powers of $D\,g'(1+\theta_1\,\varkappa)$ we obtain

$$\mathbf{E}\,H\big(Y\,g(1+\varkappa)\big) = \,\mathbf{E}\,H(Y + U\,g'(1+\theta_1\,\varkappa)) + R$$

with $|R| \le c\,\varGamma\;\mathbf{E}\;|D|$. An application of Lemma 2.1 gives $\,\mathbf{E}\,|R| \le c\,\varGamma\,\varLambda$. Expanding

$$U\,g'(1+\theta_1\,\varkappa) = U\,g'(1) + U\,\theta_1\,\varkappa\,g''(1+\theta\,\theta_1\,\varkappa)$$

and applying the inequalities

$$\big|H(s) - H(t)\big| \le |H'|_\infty\,|t-s| \le \quad |H'|_\infty\,|t-s| \quad ^{3/4},$$

for $|H'|_\infty\,|t-s| \le 1$, and

$$\big|H(s) - H(t)\big| \le |H|_\infty \le |H|_\infty\,\big(|H'|_\infty\,|t-s|\big)^{3/4},$$

for $|H'|_\infty\,|t-s| > 1$, we get

$$\mathbf{E}\,H(Y + U\,g'(1+\theta_1\,\varkappa)) = \,\mathbf{E}\,H(Y + U\,g'(1)) + R$$

with

$$|R| \le c\,\varGamma\;\mathbf{E}\,\big(|U|\,|\varkappa|\big)^{3/4} \le c\,\varGamma \quad \mathbf{E}\,|U|^{3/2} + \,\mathbf{E}\,|\varkappa|^{3/2} \quad \le c\,\varGamma\,\varLambda.$$

In the last step we applied the inequality $a\,b \le a^2 + b^2$ with $a = |U|^{3/4}$ and $b = |\varkappa|^{3/4}$ and Lemma 2.1. Since H', H'' are uniformly bounded, we can write

$$\mathbf{E}\,H(Y + U\,g'(1)) = \,\mathbf{E}\,H(Y) + \,\mathbf{E}\,H'(Y)\,U\,g'(1) + R,$$

with

$$|R| \le c\,\Gamma\,\mathbf{E}\,|U|^{3/2} \le c\,\Gamma\,\Lambda.$$

It remains to estimate

$$\mathbf{E}\,H'(Y)\,U = \mathbf{E}\sum\nolimits_{1\le i,j\le N,\ i\ne j} Y_i\,\varkappa_j\,H'(Y).$$

Recall that Y^i denotes the sum Y without summand Y_i. The Taylor expansion of $H'(Y) = H'(Y^i + Y_i)$ in powers of Y_i and $\mathbf{E}\,Y_i = 0$ yield

$$\mathbf{E}\,Y_i\varkappa_j\,H'(Y^i + Y_i) = \mathbf{E}\,Y_i\,\varkappa_j\left[H''(Y^i)\,Y_i + H'''(Y^i + \vartheta_2\vartheta_1 Y_i)\,\vartheta_1\,Y_i^2\right],$$

where ϑ_1 and ϑ_2 are mutually independent (and independent of $Y_1,\dots,Y_N$) random variables uniformly distributed in $[0,1]$. Thus $\mathbf{E}\,H'(Y)\,U = R_1 + R_2$ where

$$R_1 = \mathbf{E}\sum_{1\le i,j\le N,\ i\ne j} Y_i^2\,\varkappa_j\,H''(Y^i),$$

and

$$|R_2| \le c\,\Gamma\sum_{1\le i,j\le N}\mathbf{E}\,|Y_i|^3\,\mathbf{E}\,|\varkappa_j| \le c\,\Gamma\,\Lambda.$$

Here we estimated

$$|\varkappa_j| \le |\eta_j| + \left|Y_j\,M^{-1}\,\mathbf{E}\,Z_j\right|, \qquad \left|Y_j\,M^{-1}\,\mathbf{E}\,Z_j\right| \le Y_j^2 + \left(M^{-1}\,\mathbf{E}\,Z_j\right)^2$$

and

$$\sum\nolimits_{j=1}^N \mathbf{E}\,|\varkappa_j| \le c\sum\nolimits_{j=1}^N \mathbf{E}\,(Y_j)^2 + \sum\nolimits_{j=1}^N\left(\mathbf{E}\,Y_j^2 + (M^{-1}\,\mathbf{E}\,Z_j)^2\right) \le c.$$

Let us estimate $|R_1|$. Expanding

$$H''(Y^i) = H''(Y^{i,j} + Y_j) = H''(Y^{i,j}) + Y_j\,H'''(Y^{i,j} + \vartheta Y_j)$$

and using $\mathbf{E}\,\varkappa_j = 0$ we have

$$\mathbf{E}\,Y_i^2\,\varkappa_j\,H''(Y^i) = \mathbf{E}\,Y_i^2\,\varkappa_j\,Y_j\,H'''(Y^{i,j} + \vartheta Y_j).$$

Hence

$$|R_1| \le c\,\Gamma\sum_{i=1}^N \mathbf{E}\,Y_i^2\sum_{j=1}^N \mathbf{E}\,|\varkappa_j\,Y_j| \le c\,\Gamma\sum_{j=1}^N \mathbf{E}\,|\varkappa_j\,Y_j| \le c\,\Gamma\,\Lambda$$

Thus proving the lemma.

Proof of Theorem 1.2. Without loss of generality we shall assume that (2.1) is fulfilled. An application of Lemma 2.3 reduces the proof to the verification of the inequality

$$\sup_x \Big| \mathbf{P}\big(S < x\big) - \mathbf{P}\big(\xi < x\big) \Big| \le c\Lambda + c\Upsilon_2, \tag{2.6}$$

assuming that (2.1) holds.

In order to prove (2.6) let us apply the Esséen inequality for characteristic functions. Write

$$\begin{aligned} f(t) &= \mathbf{E}\,\exp\{itS\} = \mathbf{E}\,\exp\{itY\,g(1+\eta+2A+\Upsilon_2)\}, \\ \varphi(t) &= \mathbf{E}\,\exp\{it\xi\} = \exp\{-t^2/2\}. \end{aligned}$$

Estimating

$$\int_{|t|\le C_1} \big| f(t) - \varphi(t) \big|\, dt/|t| \le c\Lambda + c\Upsilon_2$$

by Lemma 2.4, we see that (2.6) is a consequence of

$$\int_{C_1 \le |t| \le T} |f(t) - \varphi(t)|/|t| dt \le c\Lambda + c\Upsilon_2, \tag{2.7}$$

where $T = c_1/(\sum_{i=1}^N \mathbf{E}\,|Y_i|^3)$. Here we may choose the absolute constant C_1 sufficiently large, and the absolute constant c_1 sufficiently small. We may assume that the interval (C_1, T) is non-empty since otherwise the theorem follows from (2.1).

Define the function

$$m(t) = C_2\,\frac{\ln t}{t^2} < \frac{1}{2}, \quad \text{for} \quad C_1 \le |t| \le T,$$

where C_2 is a sufficiently large absolute constant. Throughout the proof we shall write $h \simeq g$ if

$$\int_{C_1 \le |t| \le T} \big| h(t) - g(t) \big|\, dt/|t| \le c\Lambda + c\Upsilon_2.$$

In particular, (2.7) is equivalent to $f \simeq \varphi$. We shall prove the inequality (2.7) in two steps. In the first step using randomization by means of Bernoulli random variables with the parameter $m(t)$ we shall split f into a product of two conditionally independent characteristic functions. The first characteristic function will account for the contribution of the $m(t)$-th part of the sum $Y_1 + \cdots + Y_N$ and will ensure the convergence of the integral, cf. Bentkus and Götze (1994).

Step 1. Let $(\overline{Y}_1, \ldots, \overline{Y}_N)$ denote an independent copy of $(Y_1, \ldots, Y_N)$. Recall that $\alpha_1, \ldots, \alpha_N$ denotes a sequence of i.i.d. Bernoulli random variables, assume that

$\mathbf{P}(\alpha_1 = 1) = 1 - \mathbf{P}(\alpha_1 = 0) = m(t)$ and assume that all these random variables are independent. It is easy to verify that

$$f(t) = \mathbf{E} \exp\big\{ it\,(X+Z)\,g(1+\gamma+\rho)\big\},$$

where

$$\begin{aligned}
X &= \alpha_1 Y_1 + \dots + \alpha_N Y_N, \quad Z = (1-\alpha_1)\overline{Y}_1 + \dots + (1-\alpha_N)\overline{Y}_N,\\
\gamma &= \sum_{i=1}^N \alpha_i \gamma_i \quad \text{with} \quad \gamma_i = \eta_i + 2M^{-1} Y_i \,\mathbf{E}\, Z_i,\\
\rho &= \sum_{i=1}^N (1-\alpha_i)\rho_i + \varUpsilon_2 \quad \text{with} \quad \rho_i = \bar\eta_i + 2M^{-1}\overline{Y}_i \,\mathbf{E}\, Z_i,\\
\bar\eta_i &= \overline{Y}_i^2 - \mathbf{E}\,\overline{Y}_i^2, \quad 1 \le i \le N.
\end{aligned}$$

Let us show that

$$f(t) \simeq f_2(t) = \mathbf{E} \exp\{itW\},$$

where

$$W = X\,g(1+\rho) + U\,g'(1+\rho) + Z\,g(1+\rho) + Z\,\gamma\,g'(1+\rho)$$

and

$$U = \sum_{\{1 \le j,k \le N,\ j \ne k\}} \alpha_j Y_j\, \alpha_k\, \gamma_k.$$

Denote $D = \sum_{i=1}^N \alpha_i Y_i \gamma_i$. Expanding $g(1+\gamma+\rho)$ in powers of γ we have

$$\begin{aligned}
(X+Z)\,g(1+\gamma+\rho) &= X\,g(1+\rho) + X\,\gamma\,g'(1+\rho+\theta_1\gamma)\\
&\quad + Z\,g(1+\rho) + Z\,g'(1+\rho)\gamma + Z\,g''(1+\rho+\theta_2\,\gamma)\,\theta_3\,\gamma^2;\\
X\,\gamma\,g'(1+\rho+\theta_1\,\gamma) &= (D+U)\,g'(1+\rho+\theta_1\,\gamma)\\
&= D\,g'(1+\rho+\theta_1\,\gamma) + U\,g'(1+\rho) + U\,\theta_1\,\gamma\,g''(1+\rho+\theta_4\,\gamma),
\end{aligned}$$

Expanding now the exponent in powers of $it\,D\,g'(1+\rho+\theta_1\,\gamma)$ we obtain

$$f(t) = f_1(t) + R, \qquad f_1(t) = \mathbf{E} \exp\{itW + itr\},$$

with

$$r = Z\,g''(1+\rho+\theta_2\,\gamma)\theta_3\,\gamma^2 + U\,\theta_1\,\gamma\,g''(1+\rho+\theta_4\,\gamma)$$

and

$$|R| \le c|t|\,\mathbf{E}\,|D| \le c|t|\,m(t)\varLambda.$$

In the last estimate we applied Lemma 2.1. Since the function $u \to \exp\{iu\}$ and it's derivatives are uniformly bounded there exists an absolute constant $c > 0$ such that

$$f_1(t) = f_2(t) + R$$

and $|R| \le c|t|^{3/4}\, \mathbf{E}\,|r|^{3/4}$. The inequality $ab \le a^2 + b^2$ with $a = |U|^{3/4}$ and $b = |\gamma|^{3/4}$ implies

$$\begin{aligned} \mathbf{E}\,|r|^{3/4} &\le c\,\mathbf{E}\,|\gamma|^{3/2}\;\mathbf{E}_{\overline{Y}}|Z|^{3/4} + c\,\mathbf{E}\,|U\gamma|^{3/4} \\ &\le c\,\mathbf{E}\,|\gamma|^{3/2} + c\,\mathbf{E}\,|U|^{3/2} \\ &\le cm(t)\Lambda + cm^{7/4}(t)\Lambda. \end{aligned}$$

In the last step we estimated $\mathbf{E}_{\overline{Y}}|Z|^{3/4} \le c$ uniformly over α and applied lemma 2.1. We have

$$\big| f(t) - f_2(t) \big| \le c(|t| + |t|^{3/4})m(t)\Lambda.$$

Hence $f \simeq f_2$ because the factor $m(t)$ ensures the convergence of the integral with respect to the measure $dt/|t|$ as $|t| \to \infty$.

Since the function $u \to \exp\{iu\}$ and it's derivatives are bounded there exists an absolute constant $c > 0$ such that

$$\big| f_2(t) - f_3(t) - f_4(t) \big| \le R,$$

where

$$\begin{aligned} f_3(t) &= \mathbf{E}\,\exp\{it\,[X\,g(1+\rho) + Z\,g(1+\rho) + Z\,\gamma\,g'(1+\rho)]\}, \\ f_4(t) &= \mathbf{E}\,\exp\Big\{it\,[X\,g(1+\rho) + Z\,g(1+\rho) + Z\,\gamma\,g'(1+\rho)]\Big\}\,itU\,g'(1+\rho) \end{aligned}$$

and $|R| \le c|t|^{3/2}\,\mathbf{E}\,|U|^{3/2}$. Applying Lemma 2.1 we get

$$|R| \le c|t|^{3/2} m^{7/4}(t)\Lambda.$$

Hence $f_2 \simeq f_3 + f_4$. Furthermore, denoting

$$f_5(t) = \mathbf{E}\,\exp\Big\{it\,[X\,g(1+\rho) + Z\,g(1+\rho)]\Big\}\;itU\,g'(1+\rho),$$

we have

$$\big| f_4(t) - f_5(t) \big| \le |t|^{3/2}\,R, \qquad \text{with} \qquad R = \mathbf{E}\,|U|\,|Z\,\gamma|^{1/2}.$$

Estimating $\mathbf{E}_{\overline{Y}}|Z|^{1/2} \le \big(\mathbf{E}_{\overline{Y}}|Z|^2\big)^{1/4} \le c$ uniformly with respect to α we obtain

$$\begin{aligned} R &= \mathbf{E}\,|U|\,|Z\gamma|^{1/2} \le c\,\mathbf{E}\,|U|\,|\gamma|^{1/2}\,\mathbf{E}_{\overline{Y}}|Z|^{1/2} \\ &\le c\,\mathbf{E}\,|U|\,|\gamma|^{1/2} \le c(\,\mathbf{E}\,|U|^{3/2} + \mathbf{E}\,|\gamma|^{3/2}) \le cm(t)\Lambda. \end{aligned}$$

In the last step we used the inequality $a^{2/3}b^{1/3} \le a+b$ with $a^{2/3}=|U|$ and $b^{1/3}=|\gamma|^{1/2}$ and applied Lemma 2.1. Since the integral with respect to the measure $|t|^{3/2}m(t)\,\dfrac{dt}{|t|}$ converges, we conclude that $f_4 \simeq f_5$. We rewrite $f_5(t)$ in a more convenient form such that some independent random variables get separated. Write

$$P_{k,l} := \exp\{it\,[X^{k,l}\,g(1+\rho)+Z\,g(1+\rho)]\}g'(1+\rho), \quad 1\le k,l\le N, \quad k\ne l.$$

Recall that $X^{k,l}$ denotes the sum $\alpha_1 Y_1+\dots+\alpha_N Y_N$ with the summands $\alpha_k Y_k$ and $\alpha_l Y_l$ removed. Then

$$f_5(t) =$$
$$it\,\mathbf{E}_{\alpha} \sum_{\{1\le k,l\le N,\ k\ne l\}} \mathbf{E}^{\alpha}\alpha_k Y_k \exp\Big\{it\alpha_k Y_k g(1+\rho)\Big\}\,\alpha_l\gamma_l \exp\{it\alpha_l Y_l g(1+\rho)\}\times P_{k,l}.$$

Expanding the exponents in powers of $it\alpha_k Y_k\,g(1+\rho)$ and $it\alpha_l Y_l\,g(1+\rho)$ we have

$$\begin{aligned}
|f_5(t)| &\le c|t|\,\mathbf{E}_{\alpha} \sum_{\{1\le k,l\le N,\ k\ne l\}} \alpha_k\,c|t|\,\mathbf{E}\,Y_k^2\,\alpha_l\,|t|\,\mathbf{E}\,|\gamma_l Y_l| \\
&\le c|t|^3\,m^2(t) \sum_{i=1}^{N} \mathbf{E}\,|\gamma_l Y_l| \\
&= c|t|^3\,m^2(t)\,\Lambda.
\end{aligned}$$

Here we used the fact that Y_k, γ_l and $P_{k,l}$ are independent, for $k\ne l$, $|P_{k,l}|\le c$ and $\mathbf{E}\,Y_k=0$, $\mathbf{E}\,\gamma_l=0$. Hence $f_5\simeq 0$ and we conclude that $f\simeq f_3$. We shall show that $f_3\simeq f_6+f_7$, where

$$\begin{aligned}
f_6(t) &= \mathbf{E}\,\exp\big\{it\,[X\,g(1+\rho)+Z\,g(1+\rho)]\big\},\\
f_7(t) &= \mathbf{E}\,it\,Z\,\gamma\,g'(1+\rho)\,\exp\{it\,[X\,g(1+\rho)+Z\,g(1+\rho)]\}.
\end{aligned}$$

Using the inequality
$$|\exp\{ia\}-1-ia| \le |a|^{3/2}$$
with $a=t\,Z\,\gamma\,g'(1+\rho)$ we obtain $f_3(t)=f_6(t)+f_7(t)+|t|^{3/2}\,R$, where

$$|R| \le c\,\mathbf{E}\,|\gamma|^{3/2}\,\mathbf{E}_{\overline{Y}}|Z|^{3/2} \le c\,\mathbf{E}\,|\gamma|^{3/2} \le c\,m(t)\Lambda.$$

Thus $f_3\simeq f_6+f_7$. Next we will prove that $f_7\simeq 0$. Introduce for brevity

$$T_j := \mathbf{E}_{X^j}\exp\{it\,X^j\,g(1+\rho)\}, \quad Q := \exp\Big\{it\,Z\,g(1+\rho)\Big\}\,Z\,g'(1+\rho).$$

Then

$$f_7(t) = it\,\mathbf{E}\sum_{j=1}^{N}\alpha_j\,\gamma_j\,\exp\{it\,\alpha_j\,Y_j\,g(1+\rho)\}\,T_j\,Q.$$

Expanding the exponent and estimating $|\alpha_j| \le 1$ we have

$$|f_7(t)| \le c|t|^2\,\mathbf{E}_{\alpha}\sum_{i=1}^{N}\mathbf{E}_{Y_j}\,|\gamma_j\,Y_j|\,\mathbf{E}^{\alpha}|T_j|\,|Q|.$$

Using inequality (2.2) with $\beta = Y_l\,\alpha_l\,g(1+\rho)$ and with expectation taken with respect to Y_l and using the fact that $1/8 \le g \le 2$ we get

$$|T_j| \le T_j' := \prod_{1\le l\le N,\ l\ne j}(1 - \frac{1}{64}\,t^2\alpha_l\,\mathbf{E}\,Y_l^2 + \frac{32}{3}\,|t|^3\alpha_l\,\mathbf{E}\,|Y_l|^3)^{1/2}.$$

Using the simple bound $\mathbf{E}_{\overline{Y}}\,|Q| \le c$ and inequality (2.3) for $\mathbf{E}_{\alpha}\,T_j'$ we get

$$\mathbf{E}\,|Q|\,T_j' \le c\,\mathbf{E}\,T_j' \le c\,\exp\{-W/2\} \le c\,\exp\{-10^{-3}t^2\,m(t)\},$$

where

$$W = \frac{t^2\,m(t)}{64}\sum_{\{1\le l\le N,\ l\ne j\}}\mathbf{E}\,Y_l^2 - \frac{32}{3}\,|t|^3\,m(t)\sum_{i=1}^{N}\mathbf{E}\,|Y_l|^3 \ge 500^{-1}\,t^2\,m(t).$$

The last inequality is true for $C_1 \le |t| \le T$, provided that $\mathbf{E}\,|Y_j|^2 \le 1/2$ which is satisfied due to (2.1) (otherwise $\mathbf{E}\,|Y_i|^3 \ge (\,\mathbf{E}\,|Y_i|^2)^{3/2} \ge c_0$). Thus we have

$$|f_7(t)| \le c|t|^2\,\exp\{-C_3\,t^2\,m(t)\}\sum_{i=1}^{N}\mathbf{E}\,|\gamma_i\,Y_i| \le c|t|^2\,\exp\{-C_3\,t^2\,m(t)\}\,\Lambda,$$

where C_3 is a sufficiently large absolute positive constant. Hence, $f_7 \simeq 0$.

It remains to prove that $f_6 \simeq \varphi$. Introduce the sequence of i.i.d Bernoulli random variables

$$\tau = (\tau_1,\dots,\tau_N) \qquad \text{such that} \qquad \mathbf{P}\,(\tau_1 = 1) = \mathbf{P}\,(\tau_1 = 0) = 1/2.$$

Let $(Y_1',\dots,Y_N')$ be independent copy of $(Y_1,\dots,Y_N)$. Assume that

$$\tau,\quad \alpha,\quad (Y_1,\dots,Y_N),\quad (\overline{Y}_1,\dots,\overline{Y}_N) \qquad \text{and} \qquad (Y_1',\dots,Y_N')$$

are independent. Denote

$$X(\tau) = \alpha_1 \tau_1 Y_1 + \cdots + \alpha_N \tau_N Y_N$$

and

$$X(1-\tau) = \alpha_1 (1-\tau_1) Y_1' + \cdots + \alpha_N (1-\tau_N) Y_N'.$$

Observe that $X(\tau)$ and $X(1-\tau)$ are conditionally independent given α and τ.

We have

$$f_6(t) = \mathbf{E} \exp\big\{ it Z g(1+\rho) \big\} \exp\Big\{ it \big(X(\tau) + X(1-\tau) \big) g(1+\rho) \Big\}.$$

Write

$$Q_0 := \exp\{it Z g(1+\rho)\}, \qquad Q_1 := \exp\{it X(\tau) g(1+\rho)\},$$
$$Q_2 := \exp\{it X(1-\tau) g(1+\rho)\}.$$

Expanding $g(1+\rho)$ in powers of ρ we obtain

$$g(1+\rho) = g(1) + g'(1)\rho + \theta_1 g''(1+\theta_2\rho)\rho^2.$$

Expanding the exponents

$$\exp\big\{ it X(\tau) \theta_1 g''(1+\theta_2\rho)\rho^2 \big\} = 1 + r_1 \quad \text{with} \quad |r_1| \le c|t X(\tau)\rho^2|^{3/4},$$
$$\exp\big\{ -2^{-1} it X(\tau)\rho \big\} = 1 - 2^{-1} it X(\tau)\rho + r_2 \quad \text{with} \quad |r_2| \le c|t X(\tau)\rho|^{3/2}$$

we get

$$\begin{aligned} Q_1 &= \exp\big\{ it X(\tau) (1 + g'(1)\rho + \theta_1 g''(1+\theta_2\rho)\rho^2) \big\} \\ &= \exp\big\{ it X(\tau) \big\} \exp\big\{ -2^{-1} it X(\tau)\rho \big\} \exp\big\{ it X(\tau) \theta_1 g''(1+\theta_2\rho)\rho^2 \big\} \\ &= F(\tau) + P(\tau) + R(\tau), \end{aligned}$$

where

$$F(\tau) = \exp\big\{ it X(\tau) \big\}, \qquad P(\tau) = -2^{-1} it X(\tau)\rho \exp\big\{ it X(\tau) \big\}$$

and

$$\big| R(\tau) \big| \le c|t X(\tau)|^{3/4} |\rho|^{3/2} + c|t X(\tau)\rho|^{3/2}.$$

Similarly,

$$Q_2 = F(1-\tau) + P(1-\tau) + R(1-\tau).$$

On the other hand one may write

$$Q_2 = \exp\Big\{ it\, X(1-\tau)\,\big(1 + g'(1+\theta_1\,\rho)\,\rho\big)\Big\} = F(1-\tau) + R_0(1-\tau),$$

where

$$|R_0(1-\tau)| \le c|t\,X(1-\tau)\,\rho|^{1/2}.$$

We have

$$\begin{aligned} f_6 &= \mathbf{E}\,Q_0\,\big(F(\tau)+P(\tau)+R(\tau)\big)\,Q_2 \\ &= \mathbf{E}\,Q_0\,F(\tau)\,\big(F(1-\tau)+P(1-\tau)+R(1-\tau)\big) \\ &+ \mathbf{E}\,Q_0\,P(\tau)\,\big(F(1-\tau)+R_0(1-\tau)\big) + \,\mathbf{E}\,Q_0\,R(\tau)\,Q_2. \end{aligned}$$

In what follows we shall prove that

$$f_6 \simeq \,\mathbf{E}\,Q_0\,F(\tau)\,F(1-\tau). \tag{2.8}$$

First, let us prove that $\mathbf{E}\,Q_0\,R(\tau)\,Q_2 \simeq 0$. The random variables $R(\tau)$ and Q_2 are conditionally independent given $\alpha, \tau, \overline{Y}$ and $|Q_0| \le 1$. Hence

$$|\,\mathbf{E}\,Q_0\,R(\tau)\,Q_2| \le |\,\mathbf{E}\,Q_0\;\mathbf{E}_{\,Y} R(\tau)\;\mathbf{E}_{\,Y'} Q_2| \le \,\mathbf{E}\,\mathbf{E}_Y\,|R(\tau)|\,|\mathbf{E}_{Y'}\,Q_2|.$$

Proceeding in the same way as above while estimating T_j, we get

$$|\mathbf{E}_{Y'}\,Q_2| \le \prod_{l=1}^{N}\Big(1 - \frac{1}{64}\;t^2\,(1-\tau_l)\,\alpha_l\;\mathbf{E}\,(Y_l')^2 + \frac{32}{3}\;|t|^3\,(1-\tau_l)\,\alpha_l\;\mathbf{E}\,|Y_l'|^3\Big)^{1/2}.$$

The estimate

$$\mathbf{E}_Y\,\big|\,R(\tau)\big| \le c\,(|t|^{3/4} + |t|^{3/2})\,|\rho|^{3/2}$$

which holds uniformly with respect to τ together with integration with respect to τ yields

$$\mathbf{E}_\tau\,\mathbf{E}_Y\,\big|\,R(\tau)\big|\,\big|\,\mathbf{E}_{Y'}\,Q_2\,\big| \le c\,(|t|^{3/4} + |t|^{3/2})\,|\rho|^{3/2}\prod_{l=1}^{N} T_l, \tag{2.9}$$

where

$$T_l = (1 - \frac{1}{128}\,t^2\alpha_l\,\mathbf{E}\,(Y_l')^2 + \frac{16}{3}\,|t|^3\alpha_l\,\mathbf{E}\,|Y_l'|^3)^{1/2}, \qquad 1 \le l \le N.$$

Now we may integrate the product obtained in the right-hand side of (2.9) firstly with respect to $\overline{Y}$ and afterwards with respect to α. Applying Lemma 2.1 and (2.1) we get

$$\mathbf{E}_{\overline{Y}}\,|\rho|^{3/2} \le 2^{3/2}\,\mathbf{E}\;\;\sum\nolimits_{i=1}^{N}\rho_i\;\;^{3/2} + 2^{3/2}\,\Upsilon_2^{3/2} \le c\,(\Lambda + \Upsilon_2).$$

Hence,

$$\mathbf{E}\,(|t|^{3/4}+|t|^{3/2})\,|\rho|^{3/2}\prod_{l=1}^{N}T_l\le c\,(|t|^{3/4}+|t|^{3/2})\,\mathbf{E}_\alpha\,(\varLambda+\varUpsilon_2)\prod_{l=1}^{N}T_l.$$

The expectation $\mathbf{E}_\alpha\prod\ldots$ does not exceed

$$\exp\{-\frac{1}{256}\,m(t)\,t^2+\frac{8}{3}\,m(t)\,|t|^3\sum_{i=1}^{N}\mathbf{E}\,|Y_i|^3\}\le\exp\{-C_3\,m(t)\,t^2\},$$

for $C_1\le|t|\le T$. We conclude that $\mathbf{E}\,Q_0\,R(\tau)\,Q_2\simeq 0$. Proceeding in a similar way we get $\mathbf{E}\,Q_0\,F(\tau)\,R(1-\tau)\simeq 0$.

Let us prove that $\mathbf{E}\,Q_0\,P(\tau)\,R_0(1-\tau)\simeq 0$. Expanding the exponent in powers of $it\,\alpha_j\,\tau_j\,Y_j$ we get

$$\big|\mathbf{E}_Y\,P(\tau)\big|=\ 2^{-1}\,t\,\rho\,\mathbf{E}_Y\ \sum_{j=1}^{N}\alpha_j\,\tau_j\,Y_j\ \exp\big\{\,it\,\alpha_j\,\tau_j\,Y_j+it\,X^j(\tau)\,\big\}\ \le c\,|t|\,|\rho|\sum_{j=1}^{N}G_j\,H_j,$$

where $G_j=|t|\ \mathbf{E}\,Y_j^2$, and

$$H_j:=\ \mathbf{E}_Y\ \exp\big\{\,it\,X^j(\tau)\,\big\}\ \le\prod_{1\le l\le N,l\ne j}\ 1-t^2\,\alpha_l\,\tau_l\,\mathbf{E}\,Y_l^2+\frac{4}{3}\,|t|^3\,\alpha_l\,\tau_l\,\mathbf{E}\,|Y_l|^3\ ^{1/2}.$$

An application of Lemma 2.2 leads to

$$\begin{aligned}\mathbf{E}_{\alpha,\tau}\,H_j&\le\exp\Big\{-4^{-1}\,t^2\,m(t)\sum_{1\le l\le N,\ l\ne j}\mathbf{E}\,Y_l^2+3^{-1}\,|t|^3\,m(t)\,\sum_{l=1}^{N}\mathbf{E}\,|Y_l|^3\Big\}\\&\le\exp\{-C_3\,m(t)\,t^2\},\qquad\qquad(2.10)\end{aligned}$$

for $C_1\le|t|\le T$. Here we assumed that $\sum_{1\le l\le N,\ l\ne j}\mathbf{E}\,Y_l^2>1/2$, since otherwise the theorem follows from (2.1). Estimating

$$|Q_0|\le 1,\quad \mathbf{E}_{Y'}\,\big|\,X(1-\tau)\,\big|^{1/2}\le c,\quad \mathbf{E}_{\overline{Y}}\,|\rho|^{3/2}\le c\varLambda+c\varUpsilon_2,\quad \sum_{j=1}^{N}G_j\le c|t|.$$

and using (2.10) we get

$$\begin{aligned}\big|\,\mathbf{E}\,Q_0\,P(\tau)\,R_0(1-\tau)\,\big|&=\mathbf{E}_{\alpha,\tau,\overline{Y}}\,|Q_0|\,\big|\mathbf{E}_Y\,P(\tau)\big|\,\mathbf{E}_{Y'}\,\big|\,R_0(1-\tau)\,\big|\\&\le c\,|t|^{3/2}\sum_{j=1}^{N}G_j\,\mathbf{E}_{\alpha,\tau}\,H_j\,\mathbf{E}_{\overline{Y}}\,|\rho|^{3/2}\\&\le c\,|t|^{3/2}\,\exp\{-C_3\,m(t)\,t^2\}\,(\varLambda+\varUpsilon_2).\end{aligned}$$

Thus we conclude that $\mathbf{E}\, Q_0\, P(\tau)\, R_0(1-\tau) \simeq 0$.

Let us prove that $\mathbf{E}\, Q_0\, P(\tau)\, F(1-\tau) \simeq 0$. Expanding the exponent as above we get

$$\mathbf{E}_Y\, X(\tau)\, \exp\big\{ it\, X(\tau) \big\} \;\le c \sum\nolimits_{j=1}^{N} G_j\, H_j.$$

Observe that

$$\big|\, \mathbf{E}_{Y'}\, F(1-\tau) \big| \le \;\; \mathbf{E}_{Y'}\, \exp\big\{ it\, X^j(1-\tau) \big\} \;, \quad \text{for all} \quad j.$$

Hence, $\;\mathbf{E}^{\tau}\, X(\tau)\, \exp\big\{ it X(\tau) \big\}\, F(1-\tau)\;$ is bounded from above by

$$\begin{aligned} \sum_{j=1}^{N} G_j\, \mathbf{E}^{\tau} \;\; \mathbf{E}_{Y,Y'}\, \exp\big\{ it X^j(\tau) + it X^j(1-\tau) \big\} \; &\le \sum_{j=1}^{N} G_j\, \exp\{-C_3\, m(t)\, t^2\} \\ &\le c\, |t|\, \exp\{-C_3\, m(t)\, t^2\}, \end{aligned}$$

for $C_1 \le |t| \le T$. The relation $\mathbf{E}\, Q_0\, P(\tau)\, F(1-\tau) \simeq 0$ now follows provided that the following estimate holds

$$|\, \mathbf{E}_{\overline{Y}}\, Q_0\, \rho| \le c\, \big(1 + |t|^{1/2} + |t|^{3/4} + |t| \big)\, (\varLambda + \varUpsilon_2) \tag{2.11}$$

uniformly in α. Let us prove (2.11). Define

$$\hat\rho := \sum\nolimits_{l=1}^{N} \hat\rho_l, \qquad\qquad \text{where} \quad \hat\rho_l = (1-\alpha)\, \rho_l.$$

We have

$$\mathbf{E}_{\overline{Y}}\, \rho Q_0 = \mathbf{E}_{\overline{Y}}\, \hat\rho Q_0 + R, \qquad\qquad |R| \le |\, \mathbf{E}_{\overline{Y}}\, \varUpsilon_2\, Q_0| \le \varUpsilon_2.$$

Write $g_0(u) := g(u + \varUpsilon_2)$ and

$$\begin{aligned} \hat\rho\, Q_0 :&= \sum_{j=1}^{N} \hat\rho_j\, \exp\big\{ it\, Z\, g_0(1+\hat\rho) \big\} \\ &= \sum_{j=1}^{N} \hat\rho_j\, \exp\big\{ it\, (1-\alpha_j)\, \overline{Y}_j\, g_0(1+\hat\rho) \big\}\, \exp\big\{ it Z^j\, g_0(1+\hat\rho) \big\}. \end{aligned}$$

Expanding the first exponent we get

$$\hat\rho Q_0 = f_8 + R,$$

where

$$f_8(t)=\sum\nolimits_{j=1}^{N}\hat\rho_j\,\exp\bigl\{it\,Z^j\,g_0(1+\hat\rho)\bigr\}$$

and

$$\mathbf{E}_{\overline{Y}}\,|R|\le c|t|\,\sum\nolimits_{j=1}^{N}\mathbf{E}_{\overline{Y}}\,|\hat\rho_j\,\overline{Y}_j|\le c|t|\,\varLambda,$$

see the proof of the first inequality of lemma 2.1. Expanding

$$g_0(1+\hat\rho)=g_0(1)+g_0'(1+\theta\hat\rho)\,(\hat\rho^j+\hat\rho_j)$$

and then expanding the exponent in f_8 as follows

$$\exp\bigl\{it\,Z^j\,g_0(1+\hat\rho)\bigr\}=\exp\Bigl\{it\,Z^j\,g_0(1)+it\,Z^j\,g_0'(1+\theta\,\hat\rho)\,\hat\rho^j\Bigr\}+r,$$

with $|r|\le c\,|t\,Z^j\,\hat\rho_j|^{1/2}$, we get $f_8=f_9+R$. Here

$$f_9=\sum\nolimits_{j=1}^{N}\hat\rho_j\,\exp\bigl\{it\,Z^j g_0(1)+it\,Z^j\,g_0'(1+\theta\,\hat\rho)\,\hat\rho^j\bigr\}$$

and

$$|\,\mathbf{E}_{\overline{Y}}\,R|\le c\,|t|^{1/2}\sum_{j=1}^{N}\mathbf{E}_{\overline{Y}}\,|\hat\rho_j|^{3/2}\,|Z^j|^{1/2}\le c\,|t|^{1/2}\sum_{j=1}^{N}\mathbf{E}_{\overline{Y}}\,|\hat\rho_j|^{3/2}\le c\,|t|^{1/2}\,\varLambda.$$

Here we used the independence of $\hat\rho_j$ and Z^j as well as bounds

$$\mathbf{E}_{\overline{Y}}\,|Z^j|^{1/2}\le\bigl(\mathbf{E}_{\overline{Y}}\,|Z^j|^2\bigr)^{1/4}\le c\qquad\text{and}\qquad\sum\nolimits_{i=1}^{N}\mathbf{E}_{\overline{Y}}\,|\hat\rho_j|^{3/2}\le c\,\varLambda.$$

For the last inequality see the proof of lemma 2.1. Consider the second summand in the argument of the exponent in f_9. Let $D^{(j)}$ denote the diagonal part of the product $Z^j\,\hat\rho^j$, that is,

$$D^{(j)}=\sum_{\{1\le k\le N,\ k\ne j\}}(1-\alpha_k)\,\overline{Y}_k\,\hat\rho_k$$

and let $U^{(j)}$ denote the rest, i.e., $U^{(j)}+D^{(j)}=Z^j\,\hat\rho^j$. Expanding the exponent in powers of $it\,D^{(j)}\,g_0'(1+\theta\,\hat\rho)$ we obtain $f_9=f_{10}+R$, where

$$f_{10}(t)=\sum\nolimits_{j=1}^{N}\hat\rho_j\,\exp\bigl\{it\,Z^j+it\,U^{(j)}\,g_0'(1+\theta\hat\rho)\bigr\}$$

and

$$\mathbf{E}_{\overline{Y}}\,|R| \le c\,|t|\,\mathbf{E}_{\overline{Y}}\,\sum_{j=1}^{N}|\hat\rho_j|\sum_{\{1\le l\le N, l\ne j\}}|\overline{Y}_l\,\hat\rho_l| \le c|t|\,\mathbf{E}_{\overline{Y}}\,\sum_{j=1}^{N}|\rho_j|\,\sum_{l=1}^{N}\mathbf{E}_{\overline{Y}}\,|\overline{Y}_l\,\rho_l|.$$

Estimating $\mathbf{E}_{\,\overline{Y}}\sum_{j=1}^{N}|\rho_j| \le c$ as it was done in proof of Lemma 2.4 but with $\varkappa_j$ instead of ρ_j and estimating $\sum_{l=1}^{N}\mathbf{E}_{\overline{Y}}\,|\overline{Y}_l\,\rho_l| \le c\varLambda$, we get $\mathbf{E}_{\overline{Y}}\,|R| \le c|t|\varLambda$.

Let us consider the second summand in the argument of the exponent in f_{10}. We may write

$$U^{(j)}\,g_0'(1+\theta\,\hat\rho) = U^{(j)}\,g_0'(1) + U^{(j)}\,\hat\rho\theta\,g_0''(1+\theta_1\,\theta\,\hat\rho), \quad \text{and} \quad U^{(j)}\,\hat\rho = U^{(j)}\,\hat\rho^{j} + U^{(j)}\,\hat\rho_j.$$

Applying simple inequality $|\exp\{i\,(a+b)\}-1| \le c\left(|a|^{1/2}+|b|^{3/4}\right)$, where $c>0$ is an absolute constant, we have

$$\exp\bigl\{\,it\,Z^j + it\,U^{(j)}\,g_0'(1+\theta\,\hat\rho)\,\bigr\} = \exp\bigl\{\,it\,Z^j + it\,U^{(j)}\,g_0'(1)\,\bigr\} + r_1 + r_2,$$

where

$$|r_1| \le |t\,U^{(j)}\,\hat\rho_j|^{1/2}, \qquad\qquad |r_2| \le |t\,U^{(j)}\,\hat\rho^{j}|^{3/4}.$$

Hence,

$$f_{10} = f_{11} + R,$$

where

$$f_{11}(t) = \sum_{j=1}^{N}\hat\rho_j\,\exp\bigl\{\,it\,Z^j + it\,U^{(j)}\,g_0'(1)\,\bigr\},$$

$$|R| \le \sum_{j=1}^{N}|\hat\rho_j|\,\left(|t\,U^{(j)}\,\hat\rho_j|^{1/2} + |t\,U^{(j)}\,\hat\rho^{j}|^{3/4}\right).$$

Observe that $\mathbf{E}_{\overline{Y}}\,f_{11} = 0$, since $\hat\rho_j$ and $(Z^j, U^{(j)})$ are independent and $\mathbf{E}_{\overline{Y}}\,\hat\rho_j = 0$. Furthermore, estimating

$$\mathbf{E}_{\overline{Y}_j}\,|\rho_j|^{3/2}\,\mathbf{E}_{\overline{Y}}\,|U^{(j)}|^{1/2} \le c\,\mathbf{E}_{\overline{Y}_j}\,|\rho_j|^{3/2},$$

$$\mathbf{E}_{\overline{Y}_j}\,|\rho_j|\,\mathbf{E}_{\overline{Y}}\,|U^{(j)}\,\hat\rho^{j}|^{3/4} \le c\,\mathbf{E}_{\overline{Y}_j}\,|\rho_j|\,\mathbf{E}_{\overline{Y}}\,\left(|U^{(j)}|^{3/2} + |\hat\rho^{j}|^{3/2}\right),$$

$$\textstyle\sum_{j=1}^{N}\mathbf{E}\,|\rho_j| \le c$$

and using Lemma 2.1 combined with (2.1) we obtain

$$\mathbf{E}_{\overline{Y}}\,|R| \le c\left(|t|^{1/2} + |t|^{3/4}\right)\varLambda.$$

Thus (2.11) is proved. We have $\mathbf{E}\, Q_0\, P(\tau)\, F(1-\tau) \simeq 0$. Similarly, we prove that $\mathbf{E}\; Q_0 F(\tau)\, P(1-\tau) \simeq 0$. We arrive at (2.8).

Step 2. Observe that $\mathbf{E}\; Q_0\, F(\tau)\, F(1-\tau) = \mathbf{E}\, Q_0\, \exp\{it\, X\} =: f_{12}(t)$. We shall show that $f_{12} \simeq \varphi$. Recall that $\psi_i, 1 \le i \le N$, denote independent centered Gaussian random variables with variances $\mathbf{E}\, \psi_i^2 = \mathbf{E}\, Y_i^2, \quad 1 \le i \le N$. Given α, let us apply Lemma 2.4 conditionally. We get

$$\mathbf{E}_{\psi_1} \dots \mathbf{E}_{\psi_N}\, \exp\big\{ it \sum\nolimits_{i=1}^{N} (1-\alpha_i)\, \psi_i \big\} - \mathbf{E}_{\overline{Y}}\, Q_0 \;\; \le c|t|^3 (\varLambda + \varUpsilon_2).$$

Thus, using

$$\mathbf{E}_\alpha \, \big|\, \mathbf{E}_Y\, \exp\{it\, X\} \big| \le \exp\{-C_3\, m(t)\, t^2\},$$

we obtain

$$f_{12} \simeq \mathbf{E}\, \exp\Big\{ it \sum\nolimits_{i=1}^{N} (1-\alpha_i)\, \psi_i \Big\}\, \exp\{it\, X\}.$$

Furthermore,

$$f_{12} \simeq \mathbf{E}\, \exp\Big\{ it \sum\nolimits_{i=1}^{N} (1-\alpha_i)\, \psi_i \Big\}\, \exp\Big\{ it \sum\nolimits_{i=1}^{N} \alpha_i\, \psi_i' \Big\} = \exp\{-t^2/2\}.$$

Here $\psi_1', \dots, \psi_N'$ denote independent copies of $\psi_i, \quad 1 \le i \le N$. Thus Theorem 1.2 is proved.

3. Appendix

Proof of Lemma 2.1. Let us prove the first inequality of the Lemma. Recall that $\varkappa_i = \alpha_i \big(\eta_i + M^{-1}\, Y_i\, \mathbf{E}\, Z_i \big)$. By the triangle inequality, we have

$$\mathbf{E} \;\; \sum\nolimits_{i=1}^{N} Y_i\, \varkappa_i \;\; \le m \;\sum\nolimits_{i=1}^{N} \mathbf{E}\, |Y_i\, \eta_i| + m \;\sum\nolimits_{i=1}^{N} \mathbf{E}\, Y_i^2\, |M^{-1}\, \mathbf{E}\, Z_i|$$

since $\mathbf{E}\, \alpha_i = m$. Obviously

$$\mathbf{E}\, |Y_i\, \eta_i| \le 2\, \mathbf{E}\, |Y_i|^3 \le c\, M^{-3}\, \mathbf{E}\, |Z_i|^3.$$

Similarly $\mathbf{E}\, Y_i^2 \le c\, M^{-2}\, \mathbf{E}\, Z_i^2$. Consequently, an application of the Hölder inequality yields $\mathbf{E}\, Y_i^2 |M^{-1}\, \mathbf{E}\, Z_i| \le c\, M^{-3}\, \mathbf{E}\, |Z_i|^3$. Summing over i we derive the desired bound $\mathbf{E} \;\; \sum_{i=1}^{N} Y_i\, \varkappa_i \;\; \le c\, m\, \varLambda$.

To prove the second and the third inequalities of the Lemma, we shall apply the following well known bound. Assume that $T_0 = 0, T_1, \dots, T_N$ is a martingale-difference sequence, that is, $\mathbf{E}\, (T_j \,|\, T_1, \dots, T_{j-1}) = 0$. Then

$$\mathbf{E} \;\; \sum\nolimits_{i=1}^{N} T_i \;\;^{p} \le c(p) \sum\nolimits_{i=1}^{N} \mathbf{E}\; |T_i|^p\,, \qquad \text{for} \quad 1 \le p \le 2. \tag{3.1}$$

To prove (3.1) define $f(u) = |u|^p$. Then $|f'(s) - f'(t)| \le c|t-s|^{p-1}$, for $1 < p \le 2$. Writing $S_j = T_1 + \dots + T_j$ and expanding into the Taylor series, we have

$$\mathbf{E}\, f(S_{j+1}) = \mathbf{E}\, f(S_j) + \mathbf{E}\, f'(S_j)\, T_{j+1} + R,$$

where

$$\mathbf{E}\,|R| \le c\, \mathbf{E}\,|f'(S_j + \theta T_{j+1}) - f'(S_j)|\, |T_{j+1}| \le c\, \mathbf{E}\,|T_{j+1}|^p.$$

Thus, it follows that $\big| \mathbf{E}\, f(S_{j+1}) - \mathbf{E}\, f(S_j) \big| \le c\, \mathbf{E}\,|T_{j+1}|^p$, since $\mathbf{E}\, f'(S_j)\, T_{j+1} = 0$. Applying the last inequality iteratively, we derive (3.1).

Let us prove the second inequality of the Lemma. Random variables $\varkappa_1, \dots, \varkappa_N$ are independent and have mean zero. Thus, by (3.1),

$$\sum\nolimits_{i=1}^N \varkappa_i \quad ^{3/2} \le c \sum\nolimits_{i=1}^N \mathbf{E}\,|\varkappa_i|^{3/2} \le c m\, \varLambda, \tag{3.2}$$

since, similarly to the proof of the first inequality, $\mathbf{E}\,|\varkappa_i|^{3/2} \le cm\, \mathbf{E}\,|Z_i|^3$.

For the proof of the third inequality $\mathbf{E}\,|U|^{3/2} \le c m^{7/4} \varLambda$ of the Lemma, write $U = W + W'$, where

$$W = \sum\nolimits_{1 \le i < j \le N} \alpha_i Y_i \varkappa_j, \qquad W' = \sum\nolimits_{1 \le j < i \le N} \alpha_i Y_i \varkappa_j.$$

It is sufficient to estimate $\mathbf{E}\,|W|^{3/2}$ and $\mathbf{E}\,|W'|^{3/2}$ since $|U|^{3/2} \le 2\,|W|^{3/2} + 2\,|W'|^{3/2}$. Let us estimate $\mathbf{E}\,|W|^{3/2}$. The estimation of $\mathbf{E}\,|W'|^{3/2}$ is similar. Split the sum W as follows

$$W = T_2 + \dots + T_N, \quad \text{where} \quad T_j = \varkappa_j\, (\alpha_1 Y_1 + \dots + \alpha_{j-1} Y_{j-1}), \quad 2 \le j \le N.$$

By (3.1), we obtain

$$\mathbf{E}\,|W|^{3/2} \le c \sum\nolimits_{j=2}^N \mathbf{E}\,|T_j|^{3/2}. \tag{3.3}$$

Furthermore

$$\begin{aligned} \mathbf{E}\,|T_j|^{3/2} &= \mathbf{E}\,|\varkappa_j|^{3/2}\, \mathbf{E}\,|\alpha_1 Y_1 + \dots + \alpha_{j-1} Y_{j-1}|^{3/2} \\ &\le \mathbf{E}\,|\varkappa_j|^{3/2} \left(\mathbf{E}\, (\alpha_1 Y_1 + \dots + \alpha_{j-1} Y_{j-1})^2 \right)^{3/4} \\ &\le c m^{3/4}\, \mathbf{E}\,|\varkappa_j|^{3/2}. \end{aligned}$$

Thus (3.3) together with (3.2) implies $\mathbf{E}\,|W|^{3/2} \le c m^{7/4} \varLambda$, that completes the proof of the Lemma.

Proof of Lemma 2.3. Recall that

$$S = Y g (1 + \eta + 2A + \varUpsilon_2), \qquad \text{where } A = M^{-1} \sum\nolimits_{i=1}^N Y_i\, \mathbf{E}\, Z_i.$$

Introduce the statistic $\mathbf{t}_Z = \mathbf{t}(Z_1, \dots, Z_N)$ based on observations $Z_1, \dots, Z_N$. Furhermore, denote $S' = (Y + C)\, g(1 + W)$, where

$$W = \eta + 2A + \Upsilon_2 + N^{-1}(Y + C)^2 \qquad \text{and} \qquad C = M^{-1} \sum\nolimits_{i=1}^{N} \mathbf{E}\, Z_i.$$

To prove the Lemma, it is sufficient to show that

$$\Big| \mathbf{P}\left(\sqrt{N}\,\mathbf{t} < x\right) - \mathbf{P}\left(\sqrt{N}\,\mathbf{t}_Z < x\right) \Big| \le 2\varPi, \tag{3.4}$$

$$\Big| \mathbf{P}\left(\sqrt{N}\,\mathbf{t}_Z < x\right) - \mathbf{P}\left(S' < x\right) \Big| \le c\varLambda, \tag{3.5}$$

$$\Big| \mathbf{P}\left(S' < x\right) - \Phi(x) \Big| \le \sup_x \Big| \mathbf{P}\left(S < x\right) - \Phi(x) \Big| + c\Upsilon_0 + c\varLambda. \tag{3.6}$$

To prove (3.4) notice that the event $\mathbf{t}(X_1, \dots, X_N) \ne \mathbf{t}(Z_1, \dots, Z_N)$ has probability less than $\varPi = \sum_{i=1}^{N} \mathbf{P}\,(X_i^2 > V^2)$.

Let us prove (3.5). It is easy to verify that

$$\sqrt{N}\,\mathbf{t}(Z_1, \dots, Z_N) = \frac{Y + C}{\sqrt{1 + W}}\,.$$

The function $g(u) = 1/\sqrt{u}$, for $1/2 \le u \le 7/8$. Therefore the event

$$\sqrt{N}\,\mathbf{t}(Z_1, \dots, Z_N) \ne (Y + C) g\big(1 + W\big)$$

has probability less than $\mathbf{P}\,\big\{\big|W\big| > 1/4\big\}$. Thus, it suffices to show that $\mathbf{P}\,\big\{\big|W\big| > 1/4\big\} \le c\varLambda$. Notice that

$$|W| \le |\eta| + 2\,|A| + \Upsilon_2 + 2\,N^{-1}\,Y^2 + 2\,N^{-1}\,C^2.$$

By (2.1), we have $\Upsilon_2 + 2\,N^{-1}\,C^2 \le 1/8$. Therefore

$$\mathbf{P}\,\big\{\big|W\big| > 1/4\big\} \le \mathbf{P}\left(|\eta| > 1/24\right) + \mathbf{P}\left(2\,|A| > 1/24\right) + \mathbf{P}\left(2\,N^{-1}\,Y^2 > 1/24\right).$$

We have

$$\mathbf{P}\left(|\eta| > 1/24\right) \le c\,\mathbf{E}\,|\eta|^{3/2} \le c \sum\nolimits_{i=1}^{N} \mathbf{E}\,|Y_i|^3 \le c\varLambda.$$

Similarly $\mathbf{P}\left(2\,|A| > 1/24\right) \le c\,\mathbf{E}\,A^{3/2} \le c\varLambda$. To conclude the proof of (3.5) notice that

$$\frac{1}{\sqrt{N}} \le \sum_{i=1}^{N} \mathbf{E}\,|Y_i|^3 \le \varLambda \qquad \text{since} \qquad \sum_{i=1}^{N} \mathbf{E}\,Y_i^2 = 1, \tag{3.7}$$

and therefore

$$\mathbf{P}\,(2\,N^{-1}\,Y^2 > 1/24) \le c\,N^{-1}\,\mathbf{E}\,Y^2 = c\,N^{-1} \le c\,\Lambda.$$

It remains to prove (3.6). Expanding g in powers of $N^{-1}(Y+C)^2$ we obtain

$$S' = Y\,g\,(1+W) + C\,g\,(1+W) = Y\,g(1+\eta+2\,A+\Upsilon_2) + R = S + R,$$

where

$$|R| \le R_1 + R_2, \qquad R_1 = c\,|C| = c\,\Upsilon_0, \qquad R_2 = c\,N^{-1}\,|Y|\,(Y+C)^2.$$

Writing $\Phi(x) = \mathbf{P}\big(\xi < x\big)$, where ξ is a standard normal random variable, we have

$$\sup_x \ \mathbf{P}\big(S' < x\big) - \Phi(x) \ \le \sup_x \ \mathbf{P}\big(S < x\big) - \Phi(x) \ + I_1 + I_2$$

with

$$I_1 = \sup_x \mathbf{P}\big(\xi \in [x, x + 2R_1 + 2N^{-1/2}]\big), \qquad I_2 = \mathbf{P}\big(R_2 \ge N^{-1/2}\big).$$

Chebyshev's inequality and (2.1) give

$$I_1 = \mathbf{P}\big(R_2 > N^{-1/2}\big) \le c N^{-1/2}\big(\mathbf{E}\,|Y|\,C^2 + \mathbf{E}\,|Y|^3\big) \le c\,N^{-1/2}.$$

Estimating

$$I_2 = \sup_x \mathbf{P}\big(\xi \in [x, x + 2R_1 + 2\,N^{-1/2}]\big) \le c\,R_1 + c\,N^{-1/2} = c\,\Upsilon_0 + c\,N^{-1/2},$$

and using (3.7) we obtain (3.6), that completes the proof of the Lemma.

Proof of (1.13). For $X_1, X_2, \dots$ defined by (1.10) we have

$$X_i^2 = i^{-1}\,\xi_i^2 + 2\,(-1)^i\,i^{-1/2-p}\,\xi_i + i^{-2p}.$$

Therefore, for $0 < p < 1/2$,

$$\mathbf{P}\left(X_1^2 + \dots + X_N^2 > N^{1-2p}\right) \to 1$$

since

$$X_1^2 + \dots + X_N^2 \ge 2\sum\nolimits_{i=1}^N (-1)^i\,i^{-1/2-p}\,\xi_i + \sum\nolimits_{i=1}^N i^{-2p},$$

the first series in the right-hand side converges a.s. by Kolmogorov's Three Series Theorem, and

$$\sum\nolimits_{i=1}^{N} i^{-2p} > c(p)\, N^{1-2p}, \qquad C(p) > 1, \qquad N \geq 2.$$

Now it follows that

$$\sqrt{N}\,\mathbf{t} \to_P 0,$$

since asymptotically $(X_1 + \cdots + X_N)/\sqrt{\ln(N+2)}$ is standard normal. That proves (1.13).

Proof of (1.14). In this case $p = 1/2$ and in order to prove (1.14) it suffices to show that

$$T := B_N^{-2}\,(X_1^2 + \cdots + X_N^2) \xrightarrow{P} 2, \qquad \text{where } \; B_N^2 = \sum\nolimits_{i=1}^{N} i^{-1}.$$

For any $\varepsilon > 0$, we have

$$\begin{aligned}\mathbf{P}\left(|T-2| > 2\varepsilon\right) &= \mathbf{P}\Big(B_N^{-2} \; \sum_{i=1}^{N} i^{-1}\,(\xi_i^2 - 1) + 2\sum_{i=1}^{N}(-1)^i i^{-1}\,\xi_i \; > 2\varepsilon\Big) \\ &\leq \varepsilon^{-2}\, B_N^{-4}\,\mathbf{E}\Big(\sum_{i=1}^{N} i^{-1}\,(\xi_i^2 - 1)\Big)^2 + \mathbf{P}\Big(\; \sum_{i=1}^{N}(-1)^i i^{-1}\,\xi_i \; > \varepsilon B_N^2\Big).\end{aligned}$$

In the last step we applied the Chebyshev inequality. A simple calculation shows that the first summand tends to zero as $N \to 0$. The second summand tends to zero because the series $\sum_{i=1}^{N}(-i)^{-1}\,\xi_i$ converges a.s. by Kolmogorov's Three Series Theorem. Hence (1.14) is proved.

Example 3.1. The term $\varUpsilon_2 = M^{-2}\sum_{i=1}^{N}(\,\mathbf{E}\,Z_i)^2$ in the bound (1.3) is optimal in the sense that it can not be replaced by

$$\varGamma_\delta := \sum\nolimits_{i=1}^{N}\big|\,M^{-1}\,\mathbf{E}\,Z_i\,\big|^{2+\delta} \qquad \text{or by} \qquad \varUpsilon_2^{1+\delta}, \tag{3.8}$$

with some $\delta > 0$. Indeed, fix $0 < \varepsilon < 1$ and introduce the sequence $X_1, X_2, \ldots$ defined as $X_i = \xi_i + (-1)^i\,\varepsilon$, where ξ_i denote i.i.d. standard normal random variables. Choose $V(N) = \sqrt{N}$. A simple calculation shows that $\lim_{N\to\infty} \varUpsilon_2 = \varepsilon^2$, and that all other summands in the right hand side of (1.3) tend to zero as $N \to \infty$. On the other hand we have

$$\mathbf{P}\,(\sqrt{N}\,\mathbf{t} < 1) = \mathbf{P}\left(\xi < \sqrt{1+\varepsilon^2+R} + \varepsilon\left(1-(-1)^n\right)/(2\sqrt{N})\right), \tag{3.9}$$

where R is a random variable such that $R \xrightarrow{P} 0$ as $N \to \infty$. It follows from (3.9) that there exists an absolute positive constant c such that

$$\liminf_{N\to\infty} \delta_N \geq \liminf_{N\to\infty} \quad \mathbf{P}\left(\sqrt{N}\,\mathbf{t} < 1\right) - \mathbf{P}\left(\xi < 1\right) \ \geq c\,\varepsilon^2. \tag{3.10}$$

The relations (3.9) and (3.10) contradict to any estimate of the type (1.3) with terms (3.8) instead of $\varUpsilon_2$ since

$$\lim_{N\to\infty} \varGamma_\delta = 0 \qquad \text{and} \qquad \lim_{N\to\infty} \varUpsilon_2^{1+\delta} = \varepsilon^{2+2\delta}.$$

References

Bai, Z. D. and Shepp, L. A., *A note on the conditional distribution of X when $|X - y|$ is given*, Statistics & Probability Letters **19** (1994), 217–219.

Bentkus, V. and Götze, F., *The Berry–Esséen bound for Student's statistic* (1994), SFB 343 Preprint 94-033 (to appear in Ann. Probab.), Universität Bielefeld.

Bentkus, V., Götze, F., Paulauskas, V. and Račkauskas, A., *The accuracy of Gaussian approximation in Banach spaces*, SFB 343 Preprint 90–100, Universität Bielefeld (In Russian: Itogi Nauki i Tehniki, ser. Sovr. Probl. Matem., Moskva, VINITI **81** (1991), 39–139; to appear in English in Encyclopedia of Mathematical Sciences, Springer), 1990.

Bentkus, V., Götze, F. and van Zwet, W., *An Edgeworth expansion for symmetric statistics*, SFB 343 Preprint 94-020, Universität Bielefeld, 1994.

Bhattacharya, R.N. and Denker, M., *Asymptotic statistics*, Birkhäuser, Basel–Boston–Berlin, 1990.

Bhattacharya, R.N. and Ghosh, J.K., *On the validity of the formal Edgeworth expansion*, Ann. Statist. **6** (1986), 434–451.

Chibisov, D.M., *Asymptotic expansion for the distribution of a statistic admitting a stochastic expansion I*, Theor. Probab. Appl. **25** (1980), 732–744.

Chung, K.L., *The approximate distribution of Student's statistic*, Ann. Math. Stat. **17** (1946), 447-465.

Csörgő, S. and Mason, D.M., *Approximations of weighted empirical processes with applications to extreme, trimmed and self-normalized sums*, Proc. of the First World Congress of the Bernoulli Soc., Tashkent, USSR, vol. 2, VNU Science Publishers, Utrecht, 1987, pp. 811–819.

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

Esséen, C.–G., *Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law*, Acta Math. **77** (1945), 1–125.

Feller, W., *An introduction to probability theory and its applications. Vol. II, 2nd Ed.*, Wiley, New York, 1971.

Friedrich, K.O., *A Berry–Esséen bound for functions of independent random variables*, Ann. Statist. **17** (1989), 170–183.

Götze, F. and van Zwet, W., *Edgeworth expansions for asymptotically linear statistics*, SFB 343 Preprint 91–034, Universität Bielefeld, 1991, revised version 1992.

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

Griffin, P.S. and Kuelbs, J.D., *Some extensions of the LIL via self-normalizations*, Ann. Probab. **19** (1991), 380–395.

Griffin, P.S. and Mason, D.M., *On the asymptotic normality of self-normalized sums*, Math. Proc. Camb. Phil. Soc. **109** (1991), 597–610.

Hahn, M.G., Kuelbs, J. and Weiner, D.C., *The asymptotic joint distribution of self-normalized censored sums and sums of squares*, Ann. Probab. **18** (1990), 1284–1341.

Hall, P., *Edgeworth expansion for Student's t statistic under minimal moment conditions*, Ann. Probab. **15** (1987), 920–931.

Hall, P., *On the effect of random norming on the rate of convergence in the Central Limit Theorem*, Ann. Probab. **16** (1988), 1265–1280.

Helmers, R., *The Berry–Esséen bound for studentized U-statistics*, Canad. J. Stat. **13** (1985), 79–82.

Helmers, R. and van Zwet, W., *The Berry–Esséen bound for U-statistics*, Stat. Dec. Theory Rel. Top. (S.S. Gupta and J.O. Berger, eds.), vol. 1, Academic Press, New York, 1982, pp. 497–512.

LePage, R., Woodroofe, M. and Zinn, J., *Convergence to a stable distribution via order statistics*, Ann. Probab. **9** (1981), 624–632.

Logan, B., Mallows, C., Rice, S. and Shepp, L., *Limit distributions of self-normalized sums*, Ann. Probab. **1** (1973), 788–809.

Maller, A., *A theorem on products of random variables, with application to regression*, Austral. J. Statist. **23** (1981), 177–185.

Petrov, V.V., *Sums of independent random variables*, Springer, New York (1975).

Praskova, Z., *Sampling from a finite set of random variables: the Berry–Esséen bound for the Studentized mean*, Proceedings of the Fourth Prague Symposium on Asymptotic Statistics, Charles Univ., Prague, 1989, pp. 67–82.

Robbins, H., *The distribution of Student's* **t** *when the population means are unequal*, Ann. Statist. **19** (1948), 406–410.

Slavova, V.V., *On the Berry–Esséen bound for Student's statistic*, LNM **1155** (1985), 335–390.

van Zwet, W.R., *A Berry–Esséen bound for symmetric statistics*, Z. Wahrsch. verw. Gebiete **66** (1984), 425–440.