The Annals of Statistics
2019, Vol. 47, No. 3, 1616–1633
https://doi.org/10.1214/18-AOS1728

KHINCHINE'S THEOREM AND EDGEWORTH APPROXIMATIONS FOR WEIGHTED SUMS

BY SERGEY G. BOBKOV[1]

University of Minnesota and National Research University Higher School of Economics

Let F_n denote the distribution function of the normalized sum of n i.i.d. random variables. In this paper, polynomial rates of approximation of F_n by the corrected normal laws are considered in the model where the underlying distribution has a convolution structure. As a basic tool, the convergence part of Khinchine's theorem in metric theory of Diophantine approximations is extended to the class of product characteristic functions.

1. Introduction. Let $X, X_1, X_2, \dots$ be independent, identically distributed random variables (r.v.'s) with mean zero and variance σ^2 ($\sigma > 0$), and let $F_n(x) = \mathbb{P}\{Z_n \le x\}$ denote the distribution functions of the normalized sums

$$Z_n = \frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}.$$

The Edgeworth expansions are used to sharpen the standard $\frac{1}{\sqrt{n}}$–rate of approximation for F_n in the Berry–Esseen theorem, which is possible under certain assumptions on the distribution of X. To describe the simplest situation, first consider the Edgeworth correction of the third order

$$\Phi_3(x) = \Phi(x) - \frac{\alpha_3}{6\sigma^3\sqrt{n}}\left(x^2 - 1\right)\varphi(x), \qquad \alpha_3 = \mathbb{E}X^3 \quad (x \in \mathbb{R}),$$

where Φ stands for the standard normal distribution function with density $\varphi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$. Note that Φ_3 also depends on n, except for the case $\alpha_3 = 0$, when $\Phi_3 = \Phi$. It is known that, if the fourth moment $\mathbb{E}X^4$ is finite, and the characteristic function (c.f.) $f(t) = \mathbb{E}\, e^{itX}$ satisfies the Cramér condition $\limsup_{t\to\infty} |f(t)| < 1$, then the uniform deviations

$$\Delta_n^{(3)} = \sup_x \left|F_n(x) - \Phi_3(x)\right|$$

are at most of order $1/n$ (cf. [7, 8]). To reach this rate, the Cramér condition may not be removed in general, even under higher order moment assumptions. Nevertheless (alternatively), suppose that $X = \xi_0 + \alpha\xi_1$ for some independent r.v.'s ξ_k

Received October 2017; revised May 2018.
[1]Supported in part by the NSF Grant DMS-1612961 and the Russian Academic Excellence Project "5-100."
MSC2010 subject classifications. 60F05.
Key words and phrases. Central limit theorem, Edgeworth approximations.

with nondegenerate distributions, and write $\Delta_n^{(3)}(\alpha)$ so that to use α as parameter. This model appears naturally in the situation where it is known that several independent observations of X contain a certain univariate "noise" $\alpha\xi_1$. If so, how accurate is an application of the central limit theorem to the observed data? As it turns out, chances for an improved (corrected) normal approximation are rather high, although there is no confident criterion. Indeed, if say both ξ_0 and ξ_1 have a Bernoulli distribution, the order of magnitude of $\Delta_n^{(3)}(\alpha)$ may vary between $1/\sqrt{n}$ and $1/n$, depending on the arithmetic properties of the number α. However, in a typical situation, that is, for almost all $\alpha \in \mathbb{R}$ with respect to the Lebesgue measure, the order is actually $1/n$ modulo a logarithmic factor. This observation has been made in [5], and here we extend it to the case of arbitrary distributions participating in the convolution.

THEOREM 1.1. *If the r.v.'s ξ_0 and ξ_1 have mean zero with finite moments $\mathbb{E}\xi_k^4$, then, for any given $\delta > 0$, we have $\Delta_n^{(3)}(\alpha) = o(\frac{1}{n}(\log n)^{\frac{3}{2}+\delta})$ for almost all α.*

The statement admits a generalization with refinement of approximation in the model of the multivariate "noise," that is, for the class of r.v.'s represented as the weighted sum

$$X = \xi_0 + \alpha_1\xi_1 + \cdots + \alpha_m\xi_m \qquad (m = 1, 2, \dots) \tag{1.1}$$

of independent r.v.'s ξ_k having fixed nondegenerate distributions. Namely, for almost all coefficients α_k, the rate may be improved to $n^{-\frac{m+1}{2}}$ modulo a logarithmic factor, if we replace Φ_3 with an Edgeworth correction of a suitable order. To this aim, assuming that $\mathbb{E}|\xi_k|^{m+2} < \infty$ for all $k \le m$ (so that $\mathbb{E}|X|^{m+2} < \infty$), introduce the function of bounded variation

$$\Phi_{m+2}(x) = \Phi(x) - \varphi(x) \sum_{j=1}^{m} n^{-\frac{j}{2}} Q_j(x), \qquad x \in \mathbb{R}, \tag{1.2}$$

which may also be viewed as a polynomial in $1/\sqrt{n}$ of degree at most m. Here, the polynomials in the sum are defined to be

$$Q_j(x) = \sum \frac{1}{k_1! \dots k_m!} \left(\frac{\gamma_3}{3!}\right)^{k_1} \dots \left(\frac{\gamma_{m+2}}{(m+2)!}\right)^{k_m} \sigma^{-k} H_{k-1}(x),$$

where $\gamma_r = i^{-r}(\log f)^{(r)}(0)$ are the cumulants of X, and the summation is running over all integers $k_1, \dots, k_m \ge 0$ such that $k_1 + 2k_2 + \cdots + mk_m = j$, with $k = 3k_1 + \cdots + (m+2)k_m$. As usual, $H_{k-1}(x)$ denotes the Chebyshev–Hermite polynomial with leading term x^{k-1}.

Following Esseen [7], (1.2) defines the Edgeworth expansion for F_n of order $m + 2$. It is constructed in such a way that the first $m + 2$ moments of Φ_{m+2} treated as a signed measure coincide with the first $m + 2$ moments of F_n. In the case $\gamma_3 = \cdots = \gamma_{m+1} = 0$, that is, when the first $m+1$ moments of X coincide with those of a standard normal r.v. Z, (1.2) is simplified to

$$\Phi_{m+2}(x) = \Phi(x) - \frac{\gamma_{m+2}}{(m+2)!} H_{m+1}(x)\varphi(x)\, n^{-\frac{m}{2}}$$

with $\gamma_{m+2} = \mathbb{E}X^{m+2} - \mathbb{E}Z^{m+2}$. For a detail exposition of Edgeworth expansions, we refer to [2, 8] and a recent survey [4].

Note that the Edgeworth expansion is well defined under the moment assumptions regardless of the convolution structure of the distribution of X. Collecting the coefficients in (1.1) in a vector $\alpha = (\alpha_1, \dots, \alpha_m)$, put

$$\Delta_n^{(m+2)}(\alpha) = \sup_x \big|F_n(x) - \Phi_{m+2}(x)\big|.$$

THEOREM 1.2. *Suppose that the r.v.'s ξ_k in (1.1) have mean zero and finite moments $\mathbb{E}|\xi_k|^{m+3}$ for some integer $m \geq 1$. Then, for any given $\delta > 0$, for almost all $\alpha \in \mathbb{R}^m$,*

$$\Delta_n^{(m+2)}(\alpha) = o\big(n^{-\frac{m+1}{2}} (\log n)^{\frac{m}{2}+1+\delta}\big). \tag{1.3}$$

As easy to check, if ξ_k have a symmetric Bernoulli distribution, and the numbers $1, \alpha_1, \dots, \alpha_m$ are linearly independent over the field of rationals, then $\Delta_n^{(m+2)}(\alpha) \geq cn^{-\frac{m+1}{2}}$ with some constant $c > 0$ not depending on n. Hence, the power of n in the o-term of (1.3) may not be improved. On the other hand, the power of the logarithmic term may be sharpened on average.

THEOREM 1.3. *Under the same assumptions as in Theorem 1.2,*

$$\int_{(-1,1)^m} \Delta_n^{(m+2)}(\alpha)\, d\alpha = O\big(n^{-\frac{m+1}{2}} \log n\big). \tag{1.4}$$

When n is large, $\Delta_n^{(m+2)}(\alpha)$ is thus of order $n^{-\frac{m+1}{2}} \log n$ for all α from a large part of the cube $(-1,1)^m$. The proofs of (1.3)–(1.4) use the Berry–Esseen bound, while (1.3) also involves a rather interesting property that the c.f. f for the r.v. X in (1.1) is properly bounded away from 1 at infinity.

THEOREM 1.4. *Suppose that the r.v.'s ξ_k have mean zero and finite moments $\mathbb{E}|\xi_k|^3$. Given a nonincreasing function $\varepsilon(t) > 0$ in $t > 0$, such that*

$$\sum_{q=1}^{\infty} \varepsilon(q)^{\frac{m}{2}} < \infty, \tag{1.5}$$

for almost all $\alpha \in \mathbb{R}^m$, we have $|f(t)| \leq 1 - \varepsilon(t)$ for all t large enough.

In particular, $\frac{1}{1-|f(t)|} = o(t^{\frac{2}{m}}(\log t)^{\frac{2}{m}+\delta})$ as $t \to \infty$, for any fixed $\delta > 0$; cf. Corollary 5.1. This relation is what is needed for the proof of (1.3).

Being specialized to the case of Bernoulli summands ξ_k, this assertion is equivalent to the "convergence" part of the following Khinchine's theorem: Under the condition (1.5), for almost all $\alpha \in \mathbb{R}^m$, the system of m Diophantine inequalities

$$\left|\alpha_k - \frac{p_k}{q}\right| < \frac{1}{q}\sqrt{\varepsilon(q)} \qquad (1 \le k \le m) \tag{1.6}$$

has only finitely many rational solutions p_k/q. In this sense, Theorem 1.4 may be viewed as a natural extension of this result from integer numbers to the class of probability distributions with product c.f.'s.

The derivation of Theorem 1.4 occupies Sections 3–4, with a preliminary reminder of one general bound on c.f.'s. Its relationship with Diophantine inequalities is explained in Section 5. In Section 6, we state the Berry–Esseen inequality, when it is specialized to the Edgeworth corrections, and in Sections 7–8, there have been done final steps of the proof of Theorems 1.2–1.3 (under the more general assumption $\mathbb{E}|\xi_k|^{s+1}$ with $3 \le s \le m+2$).

2. Esseen's upper bound on characteristic functions. Let ξ be a r.v. with distribution function $F(x) = \mathbb{P}\{\xi \le x\}$, $x \in \mathbb{R}$, and c.f.

$$v(t) = \mathbb{E}\, e^{it\xi} = \int_{-\infty}^{\infty} e^{itx}\, dF(x), \qquad t \in \mathbb{R}.$$

Put $\beta_s = \mathbb{E}\,|\xi|^s$ and, until Proposition 3.4 below, let $\mathrm{Var}(\xi) = 1$. Assuming that β_3 is finite, we are going to see that $|v(t)|$ is well bounded away from 1 on a "significant part" of the real line. We will use the following observation due to Esseen [7] (page 94, Lemma 1).

LEMMA 2.1. *Putting $g = |v|^2$, for all points $t_0, t \in \mathbb{R}$ with $|t - t_0| = r$,*

$$|v(t)|^2 \le |v(t_0)|^2 + g'(t_0)(t-t_0) - r^2\Big(1 - 6\big(1 - |v(t_0)|^2\big)^{1/3}\beta_3^{2/3} - \frac{4r}{3}\beta_3\Big). \tag{2.1}$$

PROOF. For completeness, let us remind the argument. First, let v be real-valued, that is, let F be symmetric about the origin (as measure), in which case $v(t) = \int_{-\infty}^{\infty} \cos(tx)\, dF(x)$. For a moment, we do not require that $\mathrm{Var}(\xi) = 1$. Expanding the function $t \to \cos(tx)$ near t_0 according to Taylor's formula, we have the representation

$$\cos(tx) = \cos(t_0 x) - (t-t_0)x\sin(t_0 x) - \frac{(t-t_0)^2}{2}x^2 + \frac{(t-t_0)^2}{2}x^2\big(1 - \cos(t_0 x)\big) + \theta\frac{(t-t_0)^3}{6}x^3\sin(tx), \qquad |\theta| \le 1.$$

Hence, after integration over $dF(x)$, we get

$$v(t) = v(t_0) + (t-t_0)v'(t_0) - \frac{(t-t_0)^2}{2}\beta_2 + \theta_0 \frac{(t-t_0)^2}{2} J + \theta_1 \frac{|t-t_0|^3}{6}\beta_3$$

with some $\theta_0 \in [0,1]$ and $\theta_1 \in [-1,1]$, where $J = \int_{-\infty}^{\infty} x^2(1-\cos(t_0 x))\,dF(x)$. Splitting the integration into the regions $|x| \le \lambda$ and $|x| > \lambda$ leads to

$$\begin{aligned} J &\le \lambda^2 \int_{|x|\le\lambda} \bigl(1-\cos(t_0x)\bigr)\,dF(x) + 2\int_{|x|>\lambda} x^2\,dF(x) \\ &\le \lambda^2 \int_{-\infty}^{\infty} \bigl(1-\cos(t_0x)\bigr)\,dF(x) + \frac{2}{\lambda}\int_{|x|>\lambda} |x|^3\,dF(x) \\ &\le \lambda^2\bigl(1-v(t_0)\bigr) + \frac{2}{\lambda}\beta_3. \end{aligned}$$

If $v(t_0) < 1$, the last expression is minimized when $\lambda^3 = \beta_3/(1-v(t_0))$ in which case it equals $3(1-v(t_0))^{1/3}\beta_3^{2/3}$. If $v(t_0) = 1$, then $J = 0$. Hence

$$\begin{aligned} v(t) = v(t_0) + (t-t_0)v'(t_0) - \frac{(t-t_0)^2}{2}\beta_2 \\ + \theta_0' \frac{(t-t_0)^2}{2} 3\bigl(1-v(t_0)\bigr)^{1/3}\beta_3^{2/3} + \theta_1\frac{|t-t_0|^3}{6}\beta_3, \end{aligned}$$

so,

$$v(t) \le v(t_0) + v'(t_0)(t-t_0) - \frac{r^2}{2}\left(\beta_2 - 3\bigl(1-v(t_0)\bigr)^{1/3}\beta_3^{2/3} - \frac{r}{3}\beta_3\right). \tag{2.2}$$

Finally, one may apply (2.2) to the c.f. $g(t) = |v(t)|^2 = \mathbb{E}\,e^{it\eta}$, where $\eta = \xi - \xi'$ with ξ' being an independent copy of ξ. In that case, $\beta_2(\eta) = 2$, while by Jensen's inequality, $\beta_3(\eta) \le 8\beta_3$. □

COROLLARY 2.2. *If $|v|^2$ has a local minimum at the point t_0, then*

$$\bigl|v(t_0)\bigr|^2 \le 1 - \frac{1}{216\,\beta_3^2}. \tag{2.3}$$

Indeed, the derivative of $g = |v|^2$ is vanishing at t_0 so that, by (2.1),

$$\bigl|v(t)\bigr|^2 \le \bigl|v(t_0)\bigr|^2 - r^2\left(1 - 6\bigl(1-\bigl|v(t_0)\bigr|^2\bigr)^{1/3}\beta_3^{2/3} - \frac{4r}{3}\beta_3\right), \tag{2.4}$$

where $|t - t_0| = r$. Hence

$$1 - 6\bigl(1-\bigl|v(t_0)\bigr|^2\bigr)^{1/3}\beta_3^{2/3} - \frac{4r}{3}\beta_3 \le 0$$

for all $r > 0$ small enough. Letting $r \to 0$, we arrive at (2.3).

Below we will be more interested in local maxima. If t_0 is a point of local maximum of $g = |v|^2$, then $g'(t_0) = 0$, so that again we obtain (2.4). If $|v(t_0)|^2 \geq 1 - \varepsilon$, $\varepsilon \in (0, 1)$, this inequality implies

$$|v(t)|^2 \leq |v(t_0)|^2 - r^2\Big(1 - 6\,\varepsilon^{1/3}\beta_3^{2/3} - \frac{4r}{3}\beta_3\Big).$$

To further simplify, one may impose the conditions $6\varepsilon^{1/3}\beta_3^{2/3} \leq \frac{1}{2}$ and $\frac{4r}{3}\beta_3 \leq \frac{1}{3}$, under which the expression in brackets $\geq \frac{1}{6}$. Then we arrive at the following.

COROLLARY 2.3. *Given* $0 < \varepsilon \leq \frac{1}{12^3\beta_3^2}$, *suppose that* $|v|^2$ *has at the point* t_0 *a local maximum with* $|v(t_0)|^2 \geq 1 - \varepsilon$. *Then*

$$|v(t)|^2 \leq |v(t_0)|^2 - \frac{1}{6}|t - t_0|^2 \qquad \text{for } |t - t_0| \leq \frac{1}{4\beta_3}. \tag{2.5}$$

In particular, if $|v(t)|^2 \geq 1 - \varepsilon$ *on some finite interval* $[a, b]$ *containing* t_0, *then for all* $t \in [a, b]$,

$$|t - t_0| \leq \frac{1}{4\beta_3} \quad \Rightarrow \quad |t - t_0| \leq \sqrt{6\varepsilon}. \tag{2.6}$$

The last conclusion follows from (2.5), by using $|v(t_0)|^2 \leq 1$. Note that, by Corollary 2.2, no point in $[a, b]$ may be a point of local minimum of $|v|^2$.

Corollary 2.3 has the following consequence. If the distance from t_0 to one of the endpoints, say a, is greater than or equal to $\frac{1}{4\beta_3}$, then $t = t_0 - \frac{1}{4\beta_3} \in [a, b]$, and we conclude that $\frac{1}{4\beta_3} \leq \sqrt{6\varepsilon}$, that is, $\varepsilon \geq \frac{1}{96\,\beta_3^2}$. But this contradicts to the assumption on ε. Therefore, the distance from t_0 to the endpoints must be smaller than $\frac{1}{4\beta_3}$. Choosing $t = a$ and $t = b$ in (2.6), we thus obtain the following.

COROLLARY 2.4. *Suppose that* $|v(t)|^2 \geq 1 - \varepsilon$ *on some interval* $[a, b]$ *containing a point* t_0 *of local maximum of* $|v(t)|$, *where* $0 < \varepsilon \leq \frac{1}{12^3\beta_3^2}$. *Then* $|a - t_0| \leq \sqrt{6\varepsilon}$ *and* $|b - t_0| \leq \sqrt{6\varepsilon}$. *In particular,* $|a - b| \leq 2\sqrt{6\varepsilon}$.

3. Behavior above fixed levels and curves. The statement of Corollary 2.4 may be refined in terms of the open sets

$$U_\varepsilon = U_\varepsilon(v) = \{t \in \mathbb{R} : |v(t)|^2 > 1 - \varepsilon\}.$$

PROPOSITION 3.1. *If* $0 < \varepsilon \leq 1/(12^3\beta_3^2)$, *the set* U_ε *may be decomposed into finitely or countably many intervals* (a_n, b_n) *which do not touch each other and satisfy* $|a_n - b_n| \leq 2\sqrt{6\varepsilon}$.

One of these intervals must contain the origin, and in addition $|v(a_n)|^2 = |v(b_n)|^2 = 1 - \varepsilon$ for all n. The property that these intervals do not touch each other follows from Corollary 2.2. Indeed, in case $a_n = b_{n'}$, necessarily a_n must be a point of local minimum of $|v|^2$. But then we would have

$$|v(a_n)|^2 \le 1 - \frac{1}{216\,\beta_3^2} < 1 - \varepsilon,$$

which contradicts to the assumption on ε. Also, by Corollary 2.2, for any finite $[a, b] \subset (a_n, b_n)$, we have $|a - b| \le 2\sqrt{6\varepsilon}$, so the interval (a_n, b_n) must be finite, and $|a_n - b_n| \le 2\sqrt{6\varepsilon}$ as well.

In the sequel, we denote by $\operatorname{diam}(A)$ the diameter of a set $A \subset \mathbb{R}$, assigning the value zero in case A is empty.

PROPOSITION 3.2. *Let $0 < \varepsilon \le 1/(12^3\beta_3^2)$. For any interval $I \subset \mathbb{R}$ of length $|I| \le \frac{1}{6\beta_3}$,*

$$\operatorname{diam}(U_\varepsilon \cap I) \le 2\sqrt{6\varepsilon}. \tag{3.1}$$

PROOF. Using the intervals from Proposition 3.1 and assuming that $U_\varepsilon \cap I$ is nonempty, one may pick t' in this set and choose n such that $t' \in (a_n, b_n)$. Since $|v(a_n)|^2 = |v(b_n)|^2 = 1 - \varepsilon$, there is a point $t_n \in (a_n, b_n)$ of local maximum of $|v|^2$. By Corollary 2.4, $|a_n - t_n| \le \sqrt{6\varepsilon}$ and $|b_n - t_n| \le \sqrt{6\varepsilon}$, so $|t' - t_n| \le \sqrt{6\varepsilon}$ as well. Moreover, by Corollary 2.3, $|v(t)|^2 \le 1 - \varepsilon$ on the set

$$\sqrt{6\varepsilon} \le |t - t_n| \le \frac{1}{4\beta_3}.$$

But the interval $|t - t_n| \le \frac{1}{4\beta_3}$ contains I. Indeed, since $t' \in I$ and $|t' - t_n| \le \sqrt{6\varepsilon}$, we only need to check that $\sqrt{6\varepsilon} + |I| \le \frac{1}{4\beta_3}$. By the assumption on $|I|$, the latter follows from $\sqrt{6\varepsilon} \le \frac{1}{12\beta_3}$, that is, from $\varepsilon \le \frac{1}{6\cdot 12^2\beta_3^2}$. This holds by the assumption on ε. Thus, $U_\varepsilon \cap I$ is contained in the interval $|t - t_n| \le \sqrt{6\varepsilon}$. □

COROLLARY 3.3. *Given $0 < \varepsilon \le 1$, for all $T \ge \frac{1}{6\beta_3}$,*

$$\frac{1}{2T}\operatorname{mes}\{t \in [-T, T] : |v(t)|^2 \ge 1 - \varepsilon\} \le 24\,\beta_3\sqrt{6\varepsilon}. \tag{3.2}$$

PROOF. If $\varepsilon > \frac{1}{12^3\beta_3^2}$, then $24\beta_3\sqrt{6\varepsilon} > \sqrt{2}$, and (3.2) is immediate. So, we may assume that $\varepsilon \le \frac{1}{12^3\beta_3^2}$. Put $T_1 = \frac{1}{6\beta_3}$, $T_n = nT_1$ $(n = 1, 2, \dots)$. By Proposition 3.2, for any integer k,

$$\operatorname{mes}\{t \in [T_k, T_{k+1}] : |v(t)|^2 \ge 1 - \varepsilon\} \le 2\sqrt{6\varepsilon}.$$

Given $T \ge T_1$, choose n such that $T_n \le T < T_{n+1}$. Summing these inequalities over $k = 0, 1, \dots, n$, we then get

$$\mathrm{mes}\{t \in [0, T] : |v(t)|^2 \ge 1 - \varepsilon\} \le 2(n+1)\sqrt{6\varepsilon}.$$

On the other hand, since $n \le \frac{T}{T_1}$, we have $T_1(n+1) \le T_1(\frac{T}{T_1} + 1) = T + T_1 \le 2T$. Hence

$$\mathrm{mes}\{t \in [0, T] : |v(t)|^2 \ge 1 - \varepsilon\} \le \frac{4T}{T_1}\sqrt{6\varepsilon} = 24\beta_3\sqrt{6\varepsilon}.$$

Since $|v(-t)| = |v(t)|$, the conclusion follows. □

Corollary 3.3 (with different numerical constants) is due to Esseen who actually considered multidimensional c.f.'s for isotropic, mean zero probability measures; cf. Theorem 2 in [7], page 94. Information about the diameter as in Proposition 3.2 is more precise and is needed in the proof of Theorem 1.4. Let us state Proposition 3.2 and Corollary 3.3 in a more flexible setting without the constraint on the variance.

PROPOSITION 3.4. *Let the r.v. ξ have variance $\sigma^2 = \mathrm{Var}(\xi)$ $(\sigma > 0)$ and finite moment $\beta_3 = \mathbb{E}|\xi|^3$, with c.f. $v(t)$. Given $0 < \varepsilon \le \sigma^6/(12^3\beta_3^2)$, for any interval $I \subset \mathbb{R}$ of length $|I| \le \sigma^2/(6\beta_3)$, we have*

$$\mathrm{diam}(U_\varepsilon \cap I) \le \frac{2}{\sigma}\sqrt{6\varepsilon}, \tag{3.3}$$

where $U_\varepsilon = U_\varepsilon(v)$. Moreover, if $0 < \varepsilon \le 1$ and $T \ge \frac{\sigma^2}{6\beta_3}$, then

$$\frac{1}{2T}\,\mathrm{mes}\{t \in [-T, T] : |v(t)|^2 \ge 1 - \varepsilon\} \le \frac{24\beta_3}{\sigma^3}\sqrt{6\varepsilon}. \tag{3.4}$$

PROOF. Indeed, the r.v. $\xi_\sigma = \xi/\sigma$ has variance 1, while its c.f. is given by $v_\sigma(t) = v(t/\sigma)$. Hence, $U_\varepsilon(v_\sigma) = \{t \in \mathbb{R} : |v_\sigma(t)|^2 > 1 - \varepsilon\} = \sigma U_\varepsilon$, and by (3.1),

$$\mathrm{diam}(\sigma U_\varepsilon \cap I) = \sigma\,\mathrm{diam}(U_\varepsilon \cap I/\sigma) \le 2\sqrt{6\varepsilon}, \qquad 0 < \varepsilon \le 1/(12^3\beta_3^2(\xi_\sigma)),$$

for any interval $I \subset \mathbb{R}$ of length $|I| \le \frac{1}{6\beta_3(\xi_\sigma)}$. Here, $\beta_3(\xi_\sigma) = \mathbb{E}|\xi_\sigma|^3 = \frac{1}{\sigma^3}\beta_3$, and replacing I/σ with I, we get (3.3) under the constraint $|I| \le \frac{\sigma^2}{6\beta_3}$. Also,

$$\begin{aligned}\frac{1}{2T}\,&\mathrm{mes}\{t \in [-T, T] : |v(t/\sigma)|^2 \ge 1 - \varepsilon\}\\ &= \frac{1}{2T/\sigma}\,\mathrm{mes}\{t \in [-T/\sigma, T/\sigma] : |v(t)|^2 \ge 1 - \varepsilon\}\end{aligned}$$

for any $T > 0$. Substituting $T' = T/\sigma$, we get, by (3.2),

$$\frac{1}{2T'}\,\mathrm{mes}\{t \in [-T', T'] : |v(t)|^2 \ge 1 - \varepsilon\} \le 24\beta_3(\xi_\sigma)\sqrt{6\varepsilon},$$

under the assumption $T \ge 1/(6\beta_3(X_\sigma))$, that is, $T' \ge \sigma^2/(6\beta_3)$. □

Let us now turn to more general sets of the form

$$U = \{t \in \mathbb{R} : |v(t)|^2 > 1 - \varepsilon(t)\},$$

assuming that the function $\varepsilon(t)$ is even, positive, nonincreasing in $t > 0$, with $\varepsilon(0) \le \varepsilon_0 = \frac{\sigma^6}{12^3 \beta_3^2}$. In particular, $0 < \varepsilon(t) \le \varepsilon_0$ for all $t \in \mathbb{R}$. Put

$$T_1 = \sigma^2/(6\beta_3), \qquad T_n = nT_1, \qquad I_n = [T_n, T_{n+1}] \qquad (n \text{ integer}).$$

By Proposition 3.4 with $\varepsilon = \varepsilon_0$, we have $\operatorname{diam}(U_\varepsilon \cap I) \le \frac{2}{\sigma}\sqrt{6\varepsilon}$ for any interval I of length $|I| = T_1$. Since $U \subset U_\varepsilon$, a similar conclusion is true about U as well. Moreover, by the monotonicity of $\varepsilon(t)$, choosing $t_n = \inf\{U \cap I_n\}$ in case $U \cap I_n$ is nonempty, we have $U \subset U_{\varepsilon(t_n)}$ on I_n for any $n \ge 0$. As a result, $\operatorname{diam}(U \cap I_n) \le \frac{2}{\sigma}\sqrt{6\varepsilon(t_n)}$, and we get the following.

COROLLARY 3.5. *If the set U is unbounded, there exists an increasing sequence $t_n \ge T_n$ $(n \ge 1)$ with the following properties*:

(a) $|v(t_n)|^2 \ge 1 - \varepsilon(t_n)$.

(b) *For each $t \ge T_1$ with $|v(t)|^2 > 1 - \varepsilon(t)$, we have $t_n \le t \le t_n + \frac{2}{\sigma}\sqrt{6\varepsilon(t_n)}$ for some n.*

4. Products of characteristic functions. For the proof of Theorem 1.4, we need a general triangle-type inequality for c.f.'s.

LEMMA 4.1. *Let u be the c.f. of a r.v. ξ with variance b^2. For all $t, s \in \mathbb{R}$,*

$$1 - |u(t)|^2 \ge \frac{1}{2}\big(1 - |u(s)|^2\big) - b^2|t - s|^2.$$

PROOF. Since $\sin^2(y) \le 2\sin^2(x) + 2\sin^2(y - x)$ for all $x, y \in \mathbb{R}$, there is a general bound

$$2\sin^2(x) \ge \sin^2(y) - 2|x - y|^2.$$

Let $\eta = \xi - \xi'$, where ξ' is an independent copy of ξ, so that $1 - |u(t)|^2 = 2\mathbb{E}\sin^2(t\eta/2)$. Hence, the above inequality yields

$$1 - |u(t)|^2 \ge \frac{1}{2}\big(1 - |u(s)|^2\big) - \frac{|t - s|^2}{2}\mathbb{E}\eta^2. \qquad \square$$

PROOF OF THEOREM 1.4. We restrict ourselves to the values $\alpha_k \in (0, 1)$. The r.v. X in (1.1) has the product c.f.

$$f(t) = v(t)u_\alpha(t), \qquad u_\alpha(t) = v_1(\alpha_1 t)\cdots v_m(\alpha_m t), \qquad t \in \mathbb{R},$$

where v is the c.f. of ξ_0 and v_k's denote the c.f.'s of ξ_k. Note that the property $\beta_3 = \mathbb{E}|X|^3 < \infty$ is equivalent to $\beta_{3,k} = \mathbb{E}|\xi_k|^3 < \infty$ for each $k = 0, 1, \ldots, m$.

Without loss of generality, one may assume that the set $U = \{t \geq 0 : |v(t)|^2 > 1 - 2\varepsilon(t)\}$ is unbounded, and

$$\varepsilon(0) \leq \varepsilon_0 = \min_{0 \leq k \leq m} \frac{\sigma_k^6}{2 \cdot 12^3 \beta_{3,k}^2},$$

where $\sigma_k^2 = \mathrm{Var}(\xi_k)$. Indeed, since $\varepsilon(t) \to 0$ as $t \to \infty$, there exists $t_0 > 0$ such that $\varepsilon(t_0) \leq \varepsilon_0$. One may then extend $\varepsilon(t)$ from $[t_0, \infty)$ as a constant $\varepsilon(t) = \varepsilon(t_0)$ on $[0, t_0]$, and apply the assertion to the new function.

Put $T_1 = \sigma_0^2/(6\beta_{3,0})$ and take a sequence $t_n \geq nT_1$ as in Corollary 3.5 for the function $2\varepsilon(t)$, so that:

(a) $|v(t_n)|^2 \geq 1 - 2\varepsilon(t_n)$.

(b) For each $t \geq T_1$ with the property $|v(t)|^2 > 1 - 2\varepsilon(t)$, there exists n such that $t_n \leq t \leq t_n + \frac{2}{\sigma_0}\sqrt{12\,\varepsilon(t_n)}$.

Since the values t_n grow linearly (at worst), using the monotonicity of $\varepsilon(t)$, the hypothesis (1.5) implies that

$$\sum_{n=1}^{\infty} \varepsilon(t_n)^{\frac{m}{2}} < \infty. \tag{4.1}$$

Now, fix a parameter $A > 2$. For t large enough, we have $A\varepsilon(t) < 1$, in which case the property $|u_\alpha(t)|^2 \geq 1 - A\varepsilon(t)$ implies $|v_k(\alpha t)|^2 \geq 1 - A\varepsilon(t)$ for each $k \leq m$. Let us apply the second part of Proposition 3.4 with $\varepsilon = A\varepsilon(t_n)$ and $T = t_n$, where n is large enough so that $t_n \geq \max_{k \leq m} \sigma_k^2/(6\beta_{3,k})$ and $A\varepsilon(t_n) < 1$. The inequality (3.4) then gives that the measure

$$P_n = \mathrm{mes}\{\alpha \in (0,1)^m : |u_\alpha(t_n)|^2 \geq 1 - A\varepsilon(t_n)\}$$

is bounded by

$$\begin{aligned}
&\mathrm{mes}\{\alpha \in (0,1)^m : |v_k(\alpha_k t_n)|^2 \geq 1 - A\varepsilon(t_n) \text{ for each } k \leq m\} \\
&\quad = \prod_{k=1}^{m} \mathrm{mes}\{\alpha_k \in (0,1) : |v_k(\alpha_k t_n)|^2 \geq 1 - A\varepsilon(t_n)\} \\
&\quad = \prod_{k=1}^{m} \frac{1}{2t_n} \mathrm{mes}\{t \in [-t_n, t_n] : |v_k(t)|^2 \geq 1 - A\varepsilon(t_n)\} \leq B^m \big(A\varepsilon(t_n)\big)^{\frac{m}{2}},
\end{aligned}$$

where $B = \max_k 24\sqrt{6}\beta_{3,k}/\sigma_k^3$. Hence, according to (4.1), $\sum_{n=1}^{\infty} P_n < \infty$. Applying the Borel–Cantelli lemma, it follows that, for almost all $\alpha \in (0,1)^m$, for all $n \geq n_\alpha$, we have

$$|u_\alpha(t_n)|^2 < 1 - A\varepsilon(t_n) \tag{4.2}$$

and then also

$$|f(t_n)|^2 = |v(t_n)|^2 |u_\alpha(t_n)|^2 < 1 - A\varepsilon(t_n) < 1 - 2\varepsilon(t_n).$$

Our next step is to extend this inequality to all t large enough, by replacing t_n with t on both sides. Given $t \geq T_1$, consider the two cases.

Case 1. If $|v(t)|^2 \leq 1 - 2\varepsilon(t)$, then also $|f(t)|^2 \leq 1 - 2\varepsilon(t)$.

Case 2. If $|v(t)|^2 > 1 - 2\varepsilon(t)$, we apply property b) and choose n such that

$$t_n \leq t, \qquad |t - t_n| \leq C\sqrt{\varepsilon(t_n)}, \qquad C = \frac{2}{\sigma_0}\sqrt{12}.$$

At this point, we apply Lemma 4.1 to the c.f. $u = u_\alpha$, which gives

$$1 - |u_\alpha(t)|^2 \geq \frac{1}{2}\left(1 - |u_\alpha(t_n)|^2\right) - b^2|t - t_n|^2, \qquad b^2 = \sum_{k=1}^m \sigma_k^2.$$

By (4.2), $1 - |u_\alpha(t_n)|^2 > A\varepsilon(t_n)$, while $|t - t_n| \leq C\sqrt{\varepsilon(t_n)}$. Hence

$$1 - |u_\alpha(t)|^2 \geq \frac{1}{2}\left(A - 2b^2C^2\right)\varepsilon(t_n).$$

Choosing A to be sufficiently large, the coefficient in front of $\varepsilon(t_n)$ can be made as large, as we wish. Since also $\varepsilon(t_n) \geq \varepsilon(t)$, we obtain that $|u_\alpha(t)|^2 \leq 1 - 2\varepsilon(t)$ for all t sufficiently large, and then again $|f(t)|^2 \leq 1 - 2\varepsilon(t)$. Finally, $1 - |f(t)| = \frac{1-|f(t)|^2}{1+|f(t)|} \geq \varepsilon(t)$, which is the required inequality. $\square$

5. Relationship with Diophantine inequalities. One may apply Theorem 1.4 with

$$\varepsilon(t) = \frac{1}{1 + t^{\frac{2}{m}}(\log(e+t))^{\frac{2}{m}+\delta}},$$

and then we get the following.

COROLLARY 5.1. *Given $\delta > 0$, for almost all $\alpha \in \mathbb{R}^m$ and $t \geq t_\alpha$,*

$$|f(t)| \leq 1 - t^{-\frac{2}{m}}(\log t)^{-\frac{2}{m}-\delta}. \tag{5.1}$$

Let us now describe the relationship between Diophantine inequalities and Theorem 1.4 specialized to the summands ξ_k with a symmetric Bernoulli distribution on $\{-1, 1\}$ [this will help us see, in particular, that the parameter δ may not be removed from (5.1)]. In this case, the c.f. of X is given by $f(t) = \cos(t)\cos(\alpha_1 t)\dots\cos(\alpha_m t)$.

The property that the system (1.6) has only finitely many rational solutions p_k/q may be written as

$$\max_{k \leq m}\|q\alpha_k\| \geq \sqrt{\varepsilon(q)} \qquad \text{for all } q \geq q_0 \text{ (i.e., large enough)}, \tag{5.2}$$

where $\|x\|$ denotes the closest distance from a real number x to integers. Assuming for simplicity that $\alpha_k \in (0,1)$, the above may be extended to all real $t \geq q_0$ as the the relation

$$\|t\|^2 + \|t\alpha_1\|^2 + \cdots + \|t\alpha_m\|^2 \geq \frac{1}{4}\varepsilon\bigl(q(t)\bigr), \tag{5.3}$$

where $q(t)$ denotes the closest integer to t [for definiteness, let $q(t) = q$ in case $t = q + 1/2$]. Indeed, write $t = q + \gamma$, $|\gamma| = \|t\|$, with $q = q(t)$. If $\|t\| \geq c\varepsilon(q)$, $c > 0$, then

$$M(t) \equiv \max\bigl\{\|t\|, \|t\alpha_1\|, \ldots, \|t\alpha_m\|\bigr\} \geq c\varepsilon(q).$$

Let now $\|t\| < c\varepsilon(q)$. By the assumption, $\|q\alpha_k\| \geq \varepsilon(q)$ for some $k \leq m$. Using the elementary inequalities $\|x\| \leq |x|$ and $|\|x\| - \|y\|| \leq \|x - y\|$ $(x, y \in \mathbb{R})$, we conclude that $t\alpha_k = q\alpha_k + \gamma\alpha_k$ satisfies

$$\begin{aligned}\|t\alpha_k\| &\geq \|q\alpha_k\| - \|\gamma\alpha_k\| \\ &\geq \|q\alpha_k\| - |\gamma\alpha_k| = \|q\alpha_k\| - \|t\|\alpha_k \geq (1 - c\alpha_k)\varepsilon(q).\end{aligned}$$

Here, $1 - c\alpha_k = c$ for $c = \frac{1}{1+\alpha_k}$, and then $\|t\alpha_k\| \geq c\varepsilon(q)$ in both cases. Hence

$$M(t) \geq \frac{1}{1+\alpha_k}\varepsilon(q) \geq \frac{1}{2}\varepsilon(q),$$

thus proving (5.3). Note that this inequality with integer values $t = q$ returns us to (5.2) with an additional factor $\frac{1}{2}$ on the right-hand side.

Using $|\cos(\pi x)| \leq \exp\{-\pi^2\|x\|^2/2\}$, from (5.3), we then obtain that

$$\bigl|f(\pi t)\bigr| \leq \exp\bigl\{-\pi^2\varepsilon\bigl(q(t)\bigr)/8\bigr\} \leq 1 - c\varepsilon\bigl(q(t)\bigr),$$

which is a slightly modified conclusion of Theorem 1.4. The argument may easily be reversed. Starting from $|f(t)| \leq 1 - \varepsilon(t)$, for the values $t = \pi q$ with integer q we then have

$$1 - \varepsilon(\pi q) \geq \bigl|f(\pi q)\bigr| = (1 - \delta_1)\cdots(1 - \delta_m) \geq 1 - (\delta_1 + \cdots + \delta_m),$$

where $\delta_k = 1 - |\cos(\pi q\alpha_k)|$. That is, $\varepsilon(\pi q) \leq \delta_1 + \cdots + \delta_m$. Using another inequality $1 - |\cos(\pi x)| \leq \frac{\pi^2}{2}\|x\|^2$, we have $\delta_k \leq \frac{\pi^2}{2}\|q\alpha_k\|^2$, and thus

$$\varepsilon(\pi q) \leq \frac{\pi^2}{2}\sum_{k=1}^m \|q\alpha_k\|^2 \leq \frac{m\pi^2}{2}\max_{k \leq m}\|q\alpha_k\|^2.$$

So, we return to (5.2) with the function $\varepsilon(\pi q)$ up to an m-dependent factor.

This shows that the Bernoulli case in Theorem 1.4 may be rephrased as the statement that, under the condition (1.5), the property (5.2) holds true for almost all $\alpha \in (0,1)^m$.

In the other case, $\sum_{q=1}^{\infty} \varepsilon(q)^{\frac{m}{2}} = \infty$, the "divergence" (more difficult) part of Khinchine's theorem asserts that (5.2) holds true for almost no α (cf. [6, 10, 11]). In particular, given $c > 0$, for almost all α, the inequality

$$\max_{k \le m} \|q\alpha_k\| < c(q \log q)^{-\frac{1}{m}}$$

has infinitely many integer solutions $q > 1$. Equivalently, the inequality $|f(t)| > 1 - t^{-\frac{2}{m}} (\log t)^{-\frac{2}{m}}$ is fulfilled for infinitely many integer multiples of 2π. Therefore, the parameter δ may not be removed from (5.1).

Let us mention that, in connection with Diophantine inequalities, the properties of the c.f.'s such as $\frac{1}{1-|f(t)|} = O(t^r)$ (under the name "weak Cramér") were recently considered in [1]; see also [3].

6. Berry–Esseen inequality for Edgeworth corrections. Let us now consider the i.i.d. r.v.'s $X, X_1, X_2, \dots$ with mean zero, standard deviation $\sigma > 0$ and c.f. $f(t)$. The closeness of the distribution functions F_n of the normalized sums Z_n to the Edgeworth correction Φ_s of a given integer order $s \ge 3$ in terms of the Kolmogorov distance

$$\Delta_n^{(s)} = \sup_x \bigl|F_n(x) - \Phi_s(x)\bigr|, \qquad n = 1, 2, \dots,$$

essentially depends on the behavior of $f(t)$ on large intervals of the real line. This may be seen from the following statement.

LEMMA 6.1. *Assume that $\beta_{s+1} = \mathbb{E}\,|X|^{s+1}$ is finite. Then, for all $T \ge t_0 = (\sigma^2/\beta_{s+1})^{\frac{1}{s-1}}$, with some constant $c_s > 0$ depending on s only, we have*

$$c_s \Delta_n^{(s)} \le \frac{\beta_{s+1}}{\sigma^{s+1}} n^{-\frac{s-1}{2}} + \frac{1}{T\sigma\sqrt{n}} + \int_{t_0}^{T} \frac{|f(t)|^n}{t}\,dt. \tag{6.1}$$

The derivation of similar inequalities can be found in [8, 9], with final formulations which often assume the Cramér condition. For completeness, we include a standard argument based on the general Berry–Esseen bound.

PROOF. Let F be a distribution function, and G be a differentiable function of bounded variation such that $G(-\infty) = 0$, $G(\infty) = 1$, $\sup_x |G'(x)| \le D$. The Berry–Esseen bound asserts that, for all $T > 0$,

$$c \sup_x \bigl|F(x) - G(x)\bigr| \le \int_0^T \frac{|f(t) - g(t)|}{t}\,dt + \frac{D}{T} \tag{6.2}$$

with some absolute constant $c > 0$, where

$$f(t) = \int_{-\infty}^{\infty} e^{itx}\,dF(x), \qquad g(t) = \int_{-\infty}^{\infty} e^{itx}\,dG(x)$$

are corresponding Fourier–Stieltjes transforms (cf. [3, 7–9]).

One may apply (6.2) with F_n in place of F and $G = \Phi_s$. The Fourier–Stieltjes transform of F_n is the c.f. of Z_n given by $f_n(t) = f(\frac{t}{\sigma\sqrt{n}})^n$, while, according to (6.1), the Fourier–Stieltjes transform of Φ_s is

$$g_s(t) = e^{-t^2/2} \sum \frac{1}{k_1! \dots k_{s-2}!} \left(\frac{\gamma_3}{3!}\right)^{k_1} \cdots \left(\frac{\gamma_s}{s!}\right)^{k_{s-2}} n^{-j/2} \sigma^{-k} (it)^k.$$

Here, the summation is running over all nonnegative integers $k_1, \dots, k_{s-2}$ such that $k_1 + 2k_2 + \cdots + (s-2)k_{s-2} \le s-2$, with $k = 3k_1 + \cdots + sk_{s-2}$ and $j = k_1 + 2k_2 + \cdots + (s-2)k_{s-2}$. Note that Φ_s has density (i.e., derivative) described by a similar expression

$$\varphi_s(x) = \varphi(x) \sum \frac{1}{k_1! \dots k_{s-2}!} \left(\frac{\gamma_3}{3!}\right)^{k_1} \cdots \left(\frac{\gamma_s}{s!}\right)^{k_{s-2}} n^{-j/2} \sigma^{-k} H_k(x).$$

Thus, by (6.2), for any $T_1 > 0$,

$$c\Delta_n^{(s)} \le \int_0^{T_1} \frac{|f_n(t) - g_s(t)|}{t}\, dt + \frac{1}{T_1} \sup_x |\varphi_s(x)|. \tag{6.3}$$

Some general properties of g_s and its closeness to f_n can be stated in terms of the Lyapunov coefficients $L_p = \frac{\beta_p}{\sigma^p} n^{-\frac{p-2}{2}}$, where $\beta_p = \mathbb{E}\,|X|^p$. We refer to the following (cf., e.g., [4]):

$$|f_n(t) - g_s(t)| \le C_s L_{s+1} \min(1, |t|^{s+1})\, e^{-t^2/8} \qquad \text{for } |t| \le 1/L_3 \tag{6.4}$$

with some constant C_s depending on s only. Moreover, if $L_{s+1} \le 1$, then

$$|g_s(t)| \le C_s L_{s+1}\, e^{-t^2/8} \qquad \text{for } |t| L_{s+1}^{\frac{1}{3(s-1)}} \ge 1/8, \tag{6.5}$$

and $\sup_x |\varphi_s(x)| \le C_s$. In addition, without any condition on L_{s+1},

$$\int_{-\infty}^{\infty} |\varphi_s(x) - \varphi(x)|\, dx \le s\sqrt{(3(s-2))!}\, \max\{L_{s+1}^{\frac{1}{s-1}}, L_{s+1}^{\frac{s-2}{s-1}}\}.$$

The latter implies a rough upper bound

$$\Delta_n^{(s)} \le C_s \max\{1, L_{s+1}\}, \tag{6.6}$$

which may be used in the (noninteresting) case where L_{s+1} is large.

Put

$$T_0 = L_{s+1}^{-\frac{1}{s-1}} = \sigma^{\frac{s+1}{s-1}} \beta_{s+1}^{-\frac{1}{s-1}} \sqrt{n} = t_0\, \sigma \sqrt{n}.$$

Since the function $p \to L_p^{\frac{1}{p-2}}$ is nondecreasing in $p > 2$, we have $L_3 \le L_{s+1}^{\frac{1}{s-1}}$. Therefore, the bound (6.4) holds true on the smaller interval $|t| \le T_0$. Thus, in case $L_{s+1} \le 1$, that is, if $T_0 \ge 1$, (6.3) yields

$$c_s \Delta_n^{(s)} \le L_{s+1} + \frac{1}{T_1} + \int_{T_0}^{T_1} \frac{|f_n(t) - g_s(t)|}{t}\, dt, \qquad T_1 \ge T_0, \tag{6.7}$$

with some constant $c_s > 0$ depending on s only. Moreover, (6.5) gives

$$\int_{T_0}^{T_1} \frac{|g_s(t)|}{t}\,dt \le C_s \int_{T_0}^{\infty} e^{-t^2/8}\,dt < C_s\sqrt{2\pi}\,e^{-T_0^2/8} < \frac{C_s'}{T_0^{s-1}}.$$

As a result, (6.7) is simplified to

$$c_s\Delta_n^{(s)} \le L_{s+1} + \frac{1}{T_1} + \int_{T_0}^{T_1} \frac{|f(t/\sigma)|^n}{t}\,dt,$$

which continues to hold in case $L_{s+1} \ge 1$ as well, due to (6.6). Finally, putting $T_1 = T\sigma\sqrt{n}$ and changing the variable, we arrive at (6.1). □

7. Theorem 1.2 and its extension. Keeping the assumptions of the previous section, Lemma 6.1 may be used to obtain a variety of bounds on $\Delta_n^{(s)}$ depending on the asymptotic behavior of $f(t)$ at infinity. First, let us describe one general situation, still assuming that X has a finite absolute moment $\beta_{s+1} = \mathbb{E}|X|^{s+1}$ of an integer order with $s \ge 3$.

PROPOSITION 7.1. *Suppose that, for some $p > 0$ and $q \in \mathbb{R}$,*

$$\frac{1}{1-|f(t)|} = O\big(t^p(\log t)^q\big) \qquad \text{as } t \to \infty. \tag{7.1}$$

Then

$$\Delta_n^{(s)} = O\big(n^{-\frac{1}{2}-\frac{1}{p}}(\log n)^{\frac{q+1}{p}} + n^{-\frac{s-1}{2}}\big). \tag{7.2}$$

PROOF. Suppose that $q \ne 0$. By the assumption, $|f(t)| < 1$ for all t large enough and hence for all $t > 0$ (since otherwise X_1 would have a lattice distribution). Moreover, for all $T \ge t_0 = (\frac{\sigma^2}{\beta_{s+1}})^{\frac{1}{s-1}}$, we have

$$M(T) = \max_{t_0 \le t \le T} |f(t)| \le 1 - \frac{a}{T^p \log^q(2+T)}$$

with some constant $a > 0$ which does not depend on T. Using $1-u \le e^{-u}$, we then get

$$|f(t)|^n \le M(T)^n \le \exp\Big\{-\frac{na}{T^p\log^q(2+T)}\Big\},$$

so that

$$\int_{t_0}^{T} \frac{|f(t)|^n}{t}\,dt \le \exp\Big\{-\frac{na}{T^p\log^q(2+T)}\Big\}\log(T/t_0).$$

Thus, by (6.1),

$$c_s\Delta_n^{(s)} \le \frac{\beta_{s+1}}{\sigma^{s+1}}n^{-\frac{s-1}{2}} + \frac{1}{T\sigma\sqrt{n}} + \exp\Big\{-\frac{na}{T^p\log^q(2+T)}\Big\}\log(T/t_0). \tag{7.3}$$

Let us take $T = T_n = (bn)^{1/p}(\log n)^{-r}$ with $r = \frac{q+1}{p}$ and $b > 0$. Then

$$\frac{na}{T_n^p \log^p(2+T_n)} \geq \frac{ap^q}{b} \log n + O(\log\log n) \geq (m+1)\log n,$$

where the last inequality holds true with $b = ap^q/s$ for all n large enough. In this case, the last term in (7.3) is estimated from above by $O(n^{-(s-2)})$. In case $q = 0$ with choice $r = 1/p$, we clearly arrive at the same conclusion. Therefore, (7.3) yields

$$\Delta_n^{(s)} = O\left(n^{-\frac{s-1}{2}} + \frac{1}{T_n\sqrt{n}}\right) = O\big(n^{-\frac{s-1}{2}} + n^{-\frac{1}{p}-\frac{1}{2}}(\log n)^r\big). \qquad \square$$

Combining Proposition 7.1 with Corollary 5.1, we arrive at the following more general variant of Theorem 1.2. Suppose that X admits the representation (1.1) for independent r.v.'s ξ_k having nondegenerate distributions with mean zero. To emphasize the dependence on the coefficients, we write

$$\Delta_n^{(s)}(\alpha) = \sup_x \big|\mathbb{P}\{Z_n(\alpha) \leq x\} - \Phi_s(x)\big|$$

where $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$ and

$$Z_n(\alpha) = \frac{X_1 + \dots + X_n}{\sigma\sqrt{n}}, \qquad \sigma^2 = \mathrm{Var}(X) = \mathrm{Var}(\xi_0) + \sum_{k=1}^n \alpha_k^2 \,\mathrm{Var}(\xi_k).$$

THEOREM 7.2. *If* $\mathbb{E}\,|\xi_k|^{s+1} < \infty$ $(0 \leq k \leq m)$ *for some* $s = 3, \dots, m+1$, *then for almost all* $\alpha \in \mathbb{R}^m$, *as* $n \to \infty$ *we have*

$$\Delta_n^{(s)}(\alpha) = O\big(n^{-\frac{s-1}{2}}\big). \tag{7.4}$$

Moreover, in case $s = m+2$, *the relation* (1.3) *holds true with any* $\delta > 0$.

Indeed, the hypothesis (7.1) is fulfilled with $p = \frac{2}{m}$ and $q = \frac{2}{m} + \delta$ with arbitrary $\delta > 0$ (Corollary 5.1). In this case, relation (7.2) reduces to

$$\Delta_n^{(s)} = O\big(n^{-\frac{m+1}{2}}(\log n)^{\frac{q+1}{p}} + n^{-\frac{s-1}{2}}\big).$$

If $s < m+2$, the latter leads to (7.4), and we get (1.3) in case $s = m+2$, where o may replace O by choosing a smaller value of δ.

8. Theorem 1.3 and its extension. Similarly to Theorem 7.2, let us now derive a more general variant of Theorem 1.3.

THEOREM 8.1. *If* $\mathbb{E}\,|\xi_k|^{s+1} < \infty$ $(0 \leq k \leq m)$ *for some* $s = 3, \dots, m+1$, *then as* $n \to \infty$

$$\int_{(-1,1)^m} \Delta_n^{(s)}(\alpha)\, d\alpha = O\big(n^{-\frac{s-1}{2}}\big). \tag{8.1}$$

Moreover, in case $s = m+2$, *the relation* (1.4) *holds true.*

PROOF. The c.f. of X in (1.1) is given by $f(t) = v_0(t)v_1(\alpha_1 t)\cdots v_m(\alpha_m t)$, where v_k denotes the c.f. of ξ_k. One may appeal to Lemma 6.1 once more in order to estimate the quantity $\Delta_n^{(s)}(\alpha)$ integrally over the cube $(-1,1)^m$. To simplify the Berry–Esseen inequality (6.1), note that $\sigma^2 = \mathrm{Var}(X) \geq \sigma_0^2 = \mathrm{Var}(\xi_0)$ while, by Jensen's inequality,

$$\beta_{s+1} = \mathbb{E}\,|X|^{s+1} \leq \beta \equiv (m+1)^{s+1} \sum_{k=0}^{m} \mathbb{E}\,|\xi_k|^{s+1}.$$

Therefore, (6.1) yields, for any $T \geq t_1 = (\sigma_0^2/\beta)^{1/(s-1)}$,

$$c \int_{(0,1)^m} \Delta_n^{(s)}(\alpha)\, d\alpha \leq n^{-\frac{s-1}{2}} + \frac{1}{T\sqrt{n}} + \int_{(0,1)^m} \int_{t_0}^{T} \frac{|f(t)|^n}{t}\, dt\, d\alpha, \tag{8.2}$$

where c is a positive constant which does not depend on n. Changing the order of integration, the last double integral may be written as

$$J_n(T) = \int_{t_0}^{T} \frac{|v_0(t)|^n}{t} \prod_{k=1}^{m} \psi_{k,n}(t)\, dt, \qquad \psi_{k,n}(t) = \frac{1}{t} \int_{-t}^{t} |v_k(s)|^n\, ds.$$

Here, $\psi_{k,n}$ are connected with concentration functions for the distributions of sums of independent copies of ξ_k. Recall that, for the c.f. u of any r.v. ξ, for any $t > 0$, we have, up to an absolute constant $c > 0$,

$$\frac{1}{t} \int_{-t}^{t} |u(s)|^2\, ds \leq c\,Q(\xi, 1/t) \qquad \text{where } Q(\xi, h) = \sup_x \mathbb{P}\{x \leq \xi \leq x + h\}$$

(cf. [9], page 60, Lemma 7). Therefore, for any integer $N \geq 1$,

$$\frac{1}{t} \int_{-t}^{t} |u(s)|^{2N}\, ds \leq c\,Q(S_N, 1/t)$$

for the sum $S_N = (\eta_1 - \eta_1') + \cdots + (\eta_N - \eta_N')$, where η_j, η_j' are independent copies of ξ. On the other hand, if the distribution of ξ is nondegenerate, then

$$Q(S_N, h) \leq c\, \frac{h+1}{\sqrt{N}}, \qquad h \geq 0,$$

where the constant does not depend on h and n ([9], page 76, Theorem 11). These results ensure that $\psi_{k,n}(t) \leq \frac{c}{\sqrt{n}}$ for all $t \geq t_1$ with $n \geq 2$. Therefore,

$$J_n(T) \leq \frac{c}{n^{m/2}} \int_{t_0}^{T} \frac{|v_0(t)|^n}{t}\, dt.$$

Now, putting $I_n(t) = \int_0^t |v(s)|^n\, ds$ and using again the bound $\psi_{0,n}(t) = \frac{1}{t} I_n(t) \leq \frac{c}{\sqrt{n}}$, we have, integrating by parts,

$$\int_{t_0}^{T} \frac{|v(t)|^n}{t}\, dt = \int_{t_0}^{T} \frac{1}{t}\, dI_n(t) \leq \psi_{0,n}(T) + \int_{t_0}^{T} \frac{\psi_{0,n}(t)}{t}\, dt \leq \frac{c}{\sqrt{n}} + \frac{c}{\sqrt{n}} \log(T/t_0).$$

As a result, $J_n(T) \le cn^{-(m+1)/2} \log(Te/t_0)$, and (8.2) leads to

$$c \int_{(-1,1)^m} \Delta_n^{(s)}(\alpha)\, d\alpha \le n^{-\frac{m+1}{2}} + \frac{1}{T\sqrt{n}} + n^{-\frac{m+1}{2}} \log T,$$

where $c > 0$ does not depend on n. Choosing $T = T_n \sim n^{m/2}$, the latter bound yields (8.1) in the case $s \le m+1$ and (1.4) for $s = m+2$. $\square$

Acknowledgment. We would like to thank the referee for a careful reading of the manuscript and useful remarks.

REFERENCES

[1] ANGST, J. and POLY, G. (2017). A weak Cramér condition and application to Edgeworth expansions. *Electron. J. Probab.* **22** Paper No. 59, 24. MR3683368

[2] BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York. MR0436272

[3] BOBKOV, S. G. (2016). Closeness of probability distributions in terms of Fourier–Stieltjes transforms. *Uspekhi Mat. Nauk* **71** 37–98. MR3588939

[4] BOBKOV, S. G. (2017). Asymptotic expansions for products of characteristic functions under moment assumptions of non-integer orders. In *Convexity and Concentration* (E. Carlen, M. Madiman and E. Werner, eds.). *IMA Vol. Math. Appl.* **161** 297–357. MR3837276

[5] BOBKOV, S. G. (2017). Central limit theorem and Diophantine approximations. *J. Theoret. Probab.* **31** 2390–2411. MR3866618

[6] CASSELS, J. W. S. (1957). *An Introduction to Diophantine Approximation. Cambridge Tracts in Mathematics and Mathematical Physics* **45**. Cambridge Univ. Press, New York. MR0087708

[7] ESSEEN, C.-G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. *Acta Math.* **77** 1–125. MR0014626

[8] PETROV, V. V. (1975). *Sums of Independent Random Variables. Ergebnisse der Mathematik und ihrer Grenzgebiete* **82**. Springer, New York. Translated from the Russian by A. A. Brown. MR0388499

[9] PETROV, V. V. (1987). *Limit Theorems for Sums of Independent Random Variables* (Russian). Nauka, Moscow. MR0896036

[10] SCHMIDT, W. M. (1980). *Diophantine Approximation. Lecture Notes in Math.* **785**. Springer, Berlin. MR0568710

[11] SCHMIDT, W. M. (1991). *Diophantine Approximations and Diophantine Equations. Lecture Notes in Math.* **1467**. Springer, Berlin. MR1176315

SCHOOL OF MATHEMATICS
UNIVERSITY OF MINNESOTA
127 VINCENT HALL, 206 CHURCH S.E.
MINNEAPOLIS, MINNESOTA 55455
USA
E-MAIL: bobkov@math.umn.edu