BOUNDS FOR STABLE MEASURES OF CONVEX SHELLS AND STABLE APPROXIMATIONS

V. Bentkus[1,2] A. Juozulynas[1,3,4] V. Paulauskas[3]

[2]University of Bielefeld, [3]University of Vilnius

22 March 1999

Abstract. The standard normal distribution Φ on $\mathbb{R}^d$ satisfies $\Phi\big((\partial C)^{\varepsilon}\big) \leqslant c_d\,\varepsilon$, for all $\varepsilon > 0$ and for all convex subsets $C \subset \mathbb{R}^d$, with a constant c_d which depends on the dimension d only. Here ∂C denotes the boundary of C, and $\big(\partial C\big)^{\varepsilon}$ stands for the ε-neighborhood of ∂C. Such bounds for the normal measure of convex shells are extensively used to estimate the accuracy of normal approximations.

We extend the inequality to all (non-degenerate) stable distributions on $\mathbb{R}^d$, with a constant which depends on the dimension, the characteristic exponent and the spectral measure of the distribution only. As a corollary we provide an explicit bound for the accuracy of stable approximations on the class of all convex subsets of $\mathbb{R}^d$.

1 . Introduction and formulation of results

Let $\mathbb{R}^d$ denote the standard real Euclidean space with the norm defined by $|x|^2 = x_1^2 + \cdots + x_d^2$ and the corresponding inner product $\langle x, x\rangle = |x|^2$. Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed (i.i.d.) $\mathbb{R}^d$-valued random vectors with distribution F. Denote by F_n the distribution of the sum

$$a_n^{-1}\sum_{i=1}^{n} X_i - b_n,$$

where $a_n > 0$ and $b_n \in \mathbb{R}^d$ are normalizing constants and centering vectors. It is well known that if F_n, as $n \to \infty$, converge weakly to a distribution, say G, it has to be a stable distribution with a characteristic exponent $0 < \alpha \leqslant 2$. The case $\alpha = 2$ corresponds to a Gaussian law.

1991 *Mathematics Subject Classification.* Primary 60E07; secondary 60F05.

Key words and phrases. Stable measure, ε-strip, convex set, convex shell, stable approximations, convergence rates.

[1]Research supported by the SFB 343 in Bielefeld.

[2]On leave from Vilnius Institute of Mathematics and Informatics.

[4]Research supported by the DELTA–Stiftung in Mannheim.

The characteristic function $\varphi(t) = \int\limits_{\mathbb{R}^d} \exp\{\, i\langle t, x\rangle\,\}\, G(dx)$ of a stable law G can be written as

$$\varphi(t) = \exp\Big\{ i\langle t, a\rangle - \int\limits_{S_{d-1}} \big|\langle t, y\rangle\big|^{\alpha}\, N(y,\alpha)\,\Gamma(dy)\Big\} \tag{1.1}$$

with

$$\begin{aligned} N(y,\alpha) &= 1 - i\ \mathrm{sign}\big(\langle t, y\rangle\big)\ \tan\frac{\pi\alpha}{2}, && \alpha \neq 1,\\ N(y,\alpha) &= 1 + i\ \frac{2}{\pi}\ \mathrm{sign}\big(\langle t, y\rangle\big)\ \log\big|\langle t, y\rangle\big|, && \alpha = 1, \end{aligned} \tag{1.2}$$

where $a \in \mathbb{R}^d$ and Γ denotes a finite non-negative σ-additive measure on the unit sphere $S_{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$. The measure Γ is called the spectral measure of a stable distribution. The triple (a, α, Γ) completely characterizes the stable distribution. Since all our results are independent of shifts of distributions, without loss of generality throughout we assume that $a = 0$. We write $G_{\alpha,\Gamma}$, φ_α, etc., in cases where we want to emphasize the dependence on the characteristic exponent α or on the spectral measure Γ. We denote the density of G with respect to the Lebesgue measure on $\mathbb{R}^d$ as g (if it exists). For more information about multivariate stable laws we refer to Samorodnitsky and Taqqu (1994).

A rather general formulation of the problem of convergence rates in the central limit theorem may be stated as follows (see, for example, Bhattacharya and Rao 1986, Paulauskas 1975, Sazonov 1968). Let $\mathcal{F}$ be a class of measurable functions $f\colon \mathbb{R}^d \to \mathbb{R}$. The goal is to estimate

$$\Delta_n\big(\mathcal{F}\big) := \sup_{f\in\mathcal{F}} \Big| \int\limits_{\mathbb{R}^d} f(x)\big(F_n - G\big)(dx)\Big|, \tag{1.3}$$

for example, as follows

$$\Delta_n\big(\mathcal{F}\big) \leqslant c_d\, c\big(\mathcal{F}, G\big)\, c\big(F, G\big)\, \delta_n \tag{1.4}$$

with some δ_n such that $\delta_n \to 0$, as $n \to \infty$. The constant $c\big(\mathcal{F}, G\big)$ depends on $\mathcal{F}$ and G, and $c\big(F, G\big)$ usually is a moment or pseudo-moment related to the distributions F and G.

Classes of indicator functions $\mathbf{I}\big(x; A\big)$ of subsets $A \subset \mathbb{R}^d$ are of special interest. We define $\mathbf{I}\big(x; A\big) = 1$ if $x \in A$, and $\mathbf{I}\big(x; A\big) = 0$ otherwise. A natural correspondence between classes $\mathcal{A}$ of sets $A \in \mathcal{A}$ and classes $\mathcal{F}$ of indicator functions is given by $A \leftrightarrow \mathbf{I}\big(\cdot\,; A\big)$. Hence, for such classes we can rewrite (1.3) as

$$\Delta_n\big(\mathcal{A}\big) := \sup_{A\in\mathcal{A}} \big|F_n\big(A\big) - G\big(A\big)\big|. \tag{1.5}$$

The constant $c\big(\mathcal{F}, G\big) = c\big(\mathcal{A}, G\big)$ in (1.4) usually depends on the quantities

$$\eta\big(\mathcal{A}, G, \varepsilon\big) := \sup_{A\in\mathcal{A}} G\big(\big(\partial A\big)^{\varepsilon}\big), \quad \varepsilon > 0,$$

$$\eta\big(\mathcal{A}, G\big) := \sup_{\varepsilon>0} \eta\big(\mathcal{A}, G, \varepsilon\big)/\varepsilon,$$

$$\chi(g, w_1, \dots, w_s) := \int_{\mathbb{R}^d} \big| g^{(s)}(x) w_1 \dots w_s \big| dx, \quad w_1, \dots, w_s \in \mathbb{R}^d, \tag{1.6}$$

$$\chi_s(g) := \sup \big\{ \chi(g, w_1, \dots, w_s) : \ |w_i| \leqslant 1, \text{ for all } i \big\}, \tag{1.7}$$

where $g^{(s)}(x)w_1 \dots w_s$ denotes the Fréchet derivatives. Using the directional derivatives

$$d_w g(x) := \lim_{t\to 0} \big(g(x+tw) - g(x)\big)/t,$$

we have

$$g^{(s)}(x) w_1 \dots w_s = d_{w_1} \dots d_{w_s} g(x).$$

The boundary of a set A we denote as ∂A, and $\big(\partial A\big)^{\varepsilon}$ is the ε-neighborhood of ∂A.

In the case of the standard normal distribution $G = \Phi$, the quantities χ in (1.6) and (1.7) are obviously finite. However, one needs a special proof in order to show that $\eta\big(\mathcal{A}_c, \Phi\big) < \infty$ for the class $\mathcal{A}_c$ of convex subsets of $\mathbb{R}^d$, see Bahr (1967), Bhattacharya and Rao (1986), Sazonov (1981). In the stable case $\alpha < 2$ the existence of $\chi_s(g)$ and $\eta\big(\mathcal{A}_c, G\big)$ was either imposed as a condition (see Paulauskas 1975, Bloznelis 1988), or special cases were considered such that it was possible to show the existence of η and χ. A list of the special cases consists of: i) the class $\mathcal{A}_r$ of rectangles (Banys 1971); ii) spherically symmetric distributions (see Bloznelis 1989, Paulauskas 1975, Mikhailova 1983); iii) the two dimensional case $d = 2$ (Paulauskas 1975); iv) stable random vectors with independent coordinates (Paulauskas 1975). The condition $\chi_s(g) < \infty$ is used to ensure the existence of some metrics related to stable distributions, see Ch. 14 in Rachev (1991).

The aim of the present paper is to show that all aforementioned quantities exist, for the class $\mathcal{A}_c$ of convex subsets and for any stable distribution which is non-degenerate in a subspace of $\mathbb{R}^d$. Furthermore, we obtain explicit bounds for these quantities.

A distribution G we call non-degenerate if $G\big(L\big) = 0$, for any linear subspace $L \subset \mathbb{R}^d$ such that $\dim L < d$. A stable non-degenerate distribution is absolutely continuous and, hence, has a density g. Write

$$\varkappa\big(\Gamma\big) := \inf_{|t|=1} \Big(\int_{S_{d-1}} \big|\langle t, y\rangle\big|^{\alpha} \Gamma(dy)\Big)^{1/\alpha}, \qquad \varkappa_0\big(\Gamma\big) := \Gamma\big(S_{d-1}\big). \tag{1.8}$$

Notice that $\varkappa(\Gamma) > 0$ if and only if the stable distribution $G = G_\Gamma$ is non-degenerate, and $\varkappa_0\big(\Gamma\big) < \infty$, for any stable distribution. Write

$$K_\alpha\big(\Gamma\big) = \varkappa_0^d\big(\Gamma\big)\, \varkappa^{-d\alpha-1}\big(\Gamma\big), \qquad \alpha \neq 1, \tag{1.9}$$

$$K_\alpha\big(\Gamma\big) = \varkappa_0^d\big(\Gamma\big)\, \varkappa^{-d-1}\big(\Gamma\big)\, \Big(1 + \big|\log \varkappa\big(\Gamma\big)\big|\Big)^d, \qquad \alpha = 1. \tag{1.10}$$

If the distribution G is symmetric, that is, the function $N(\,\cdot\,,\alpha)$ in the characteristic function (1.1) satisfies $N(\,\cdot\,,\alpha)\equiv 1$ or the measure Γ is symmetric, then we define $K_\alpha\bigl(\Gamma\bigr)$ by (1.9), for all $0<\alpha\leqslant 2$.

Theorem 1. *Let G be a non-degenerate stable distribution. Then $\chi_s(g)<\infty$, for $s=1,2,\dots$ Moreover, we have*

$$\chi_1(g)\leqslant c(\alpha,d)\,K_\alpha\bigl(\Gamma\bigr), \tag{1.11}$$

where $K_\alpha\bigl(\Gamma\bigr)$ is defined in (1.9)–(1.10), *and*

$$\chi_s(g)\leqslant\bigl(s^{1/\alpha}\chi_1(g)\bigr)^s, \quad \text{for } s=2,3,\dots \tag{1.12}$$

In particular, $g\in C^\infty\bigl(\mathbb{R}^d\bigr)$, for all α, and g is an analytic function, for $\alpha\geqslant 1$.

Theorem 1 refines a result of Bogachev (1986) who showed that $\chi_1(g)$ exists. This fact easily implies (1.12) and the analyticity for $\alpha\geqslant 1$. Also it is necessary to note that Bogachev asserts only the existence of $\chi_1(g)$ while we provide a constructive proof and an explicit bound. To have a constructive proof is necessary in applications such as simulations of stable random vectors by LePage series in the multidimensional case (to appear elsewhere; in papers Ledoux and Paulauskas 1996, Bentkus, Götze and Paulauskas 1996 the case $d=1$ is considered). The constant $c(\alpha,d)$ in (1.11) allows the following explicit bound

$$\begin{aligned} &c(\alpha,d)\leqslant 12\,(50/\alpha)^{2d}\,(2d)^{2d/\alpha}\,(1/\alpha)^{2/\alpha}\,\bigl(1+\bigl|\tan(\pi\alpha/2)\bigr|\bigr)^d, && \alpha\neq 1,\\ &c(1,d)\leqslant 18\,(8d)^{3d}, && \alpha=1. \end{aligned} \tag{1.13}$$

If the distribution G is symmetric, then

$$c(\alpha,d)\leqslant 12\,(50/\alpha)^{2d}\,(2d)^{2d/\alpha}\,(1/\alpha)^{2/\alpha}. \tag{1.14}$$

Notice that the bound (1.13), for $\alpha\neq 1$, is uniform in α from any compact subset of $(0,1)\cup(1,2]$, and it degenerates when $\alpha\downarrow 0$ or $\alpha\to 1$. It seems that the degeneration around 1 is an artifact of our methods. In the symmetric case the bound (1.14) is satisfactory since it degenerates only as $\alpha\downarrow 0$. Of course, the bounds (1.13)–(1.14) are not optimal. Writing them down we tried to reflect the uniformity in α and prefered simplicity of the form to accuracy.

Write

$$\zeta\bigl(\mathcal{A},G\bigr):=\sup_{A\in\mathcal{A}}\int_{\partial A} g(x)\,ds\,,$$

where ds denotes the surface area element on ∂A.

Theorem 2. *Let G be an arbitrary distribution on $\mathbb{R}^d$ such that its density g exists and is a continuous function. Then*

$$\eta\bigl(\mathcal{A}_c,G\bigr)=2\,\zeta\bigl(\mathcal{A}_c,G\bigr). \tag{1.15}$$

Let G be an arbitrary distribution on $\mathbb{R}^d$ such that $\chi_1(g)$ exists. Then

$$\eta\big(\mathcal{A}_c, G\big) \leqslant 4\, d^{3/2}\, \chi_1(g). \tag{1.16}$$

In particular, any stable non-degenerate G satisfies

$$\eta\big(\mathcal{A}_c, G\big) \leqslant 4\, d^{3/2}\, c(\alpha, d)\, K_\alpha\big(\Gamma\big)$$

with $c(\alpha, d)$ as in (1.13)–(1.14) *and $K_\alpha\big(\Gamma\big)$ as in* (1.9)–(1.10).

Using Theorem 2, the bounds for the accuracy of stable approximations in $\mathbb{R}^d$ obtained by Paulauskas (1975) and by Bloznelis (1988) extend to the whole class of non-degenerate stable distributions, see Theorem 3 below.

We hope that using (1.16) and applying the method used to prove Theorem 1, one can derive results similar to Theorem 2 for some classes of infinitely divisible distributions. As an initial step in this direction we provide an extension to the case of mixtures of stable distributions with the varying α, see Theorem 4 below.

In the special case of the standard normal distribution $G = \Phi$ simple calculations show that $\chi_1(g) \leqslant (2/\pi)^{1/2}$ and the estimate (1.16) yields

$$\eta\big(\mathcal{A}_c, \Phi\big) \leqslant (32/\pi)^{1/2} d^{3/2}, \tag{1.17}$$

which is worse than the best known estimate $\eta\big(\mathcal{A}_c, \Phi\big) \leqslant 8\, d^{1/4}$ (see Ball 1993). A precise dependence of $\eta\big(\mathcal{A}_c, \Phi\big)$ on dimension is not known and would be of interest in the context of estimates of normal approximations (see Bentkus 1986). The bound (1.17) depends on $d^{3/2}$ which is worse than $d^{1/2}$ in Bhattacharya and Rao (1986), where a proof adapted to the structure of Φ is provided. Our proof applies to arbitrary G such that $\chi_1(g) < \infty$ and seems to be simpler.

Recall that $X_1, X_2, \dots$ denote i.i.d. random vectors with common distribution F. Let henceforth F_n denote the distribution of the sum $n^{-1/\alpha} \sum\limits_{i=1}^{n} X_i$. Assume that

$$\int\limits_{\mathbb{R}^d} \big\langle t, x \big\rangle \big(F - G\big)(dx) = 0, \qquad \text{for all } t \in \mathbb{R}^d.$$

If $\alpha > 1$ then assume in addition that

$$\int\limits_{\mathbb{R}^d} \big\langle t, x \big\rangle \big\langle s, x \big\rangle \big(F - G\big)(dx) = 0, \qquad \text{for all } t, s \in \mathbb{R}^d.$$

Introduce the uniform distance

$$\rho = \rho\big(F, G\big) = \sup \big\{ \big| F\big(A\big) - G\big(A\big) \big| :\ A \in \mathcal{A}_c \big\}$$

between distributions F and G on the class $\mathcal{A}_c$ of convex sets.

The class $H^r(\mathbb{R}^d)$, $r > 0$, consists of the functions $f : \mathbb{R}^d \to \mathbb{R}$ such that f is m times Fréchet differentiable and the derivative $f^{(m)}$ satisfies

$$\sup_{|w_1|=\dots=|w_m|=1} \left| f^{(m)}(x) w_1 \dots w_m - f^{(m)}(y) w_1 \dots \ w_m \right| \leqslant |x-y|^\theta,$$

where a non-negative integer m and a positive θ satisfy $r = m + \theta$ and $0 < \theta \leqslant 1$. Introduce the metric

$$\zeta_r = \zeta_r(F, G) = \sup \Big\{ \Big| \int_{\mathbb{R}^d} f(x) (F-G)(dx) \Big| : f \in H^r(\mathbb{R}^d) \Big\}$$

and pseudo-moments

$$\nu_r = \nu_r(F, G) = \int_{\mathbb{R}^d} |x|^r \, \big| F - G \big|(dx),$$

where $\big| F - G \big|$ denotes the variation of the signed measure $F - G$. Note that $\zeta_r \leqslant \nu_r$.

Theorem 3. *Assume that G is a stable non-degenerate distribution with the characteristic function given by* (1.1)–(1.2). *If $\alpha \neq 1$ then*

$$\Delta_n(\mathcal{A}_c) \leqslant c_{\alpha,d} \, n^{-1/\alpha} \Big(\rho + K_\alpha(\Gamma) \, \zeta_1 + K_\alpha^{\alpha+1}(\Gamma) \, \zeta_{1+\alpha} \Big). \tag{1.18}$$

For $\alpha = 1$ the bound (1.18) *holds for strictly stable G.*

The quantities $\Delta_n(\mathcal{A}_c)$ and $K_\alpha(\Gamma)$ are defined by (1.5) and (1.9)–(1.10), respectively. We shall derive (1.18) combining a bound proved by Bloznelis (1988) and Theorems 1 and 2. Recall that G is strictly stable if the distribution of the sum $n^{-1/\alpha} \sum_{i=1}^{n} Y_i$ equals G when $Y_1, \dots, Y_n$ are i.i.d. with the distribution G. As it is well known, for $\alpha \neq 1$, the strict stability means that the shift a from (1.1)–(1.2) equals zero. For $\alpha = 1$, it is equivalent to $a = 0$ and $\int_{S_{d-1}} \langle t, y \rangle \, \Gamma(dy) = 0$, for all $t \in \mathbb{R}^d$. The requirement of the strict stability in Theorem 3 seems to be superfluous and is inherited from the bound of Bloznelis. The constant in (1.18) satisfies

$$c_{\alpha,d} \leqslant c \, c(\alpha) \, (20)^{1/\alpha} \, d^{3/2} \, (1 + c(\alpha, d))^{\alpha+1}$$

with $c(\alpha, d)$ defined by (1.13)–(1.14), where c is an absolute constant and $c(\alpha) = 1$, for $\alpha \le 1$, $c(\alpha) = 1 - \alpha$, for $\alpha > 1$. Once again, the bound for $c_{\alpha,d}$ degenerates as $\alpha \downarrow 0$ or $\alpha \to 1$.

Let us conclude the introduction providing the aforementioned extension to mixtures of one-dimensional stable distributions with varying α. Consider a measurable function $\alpha : S_{d-1} \to [0, 2]$. Let G_m (respectively g_m) denote a distribution (respectively its density) which has the characteristic function defined by (1.1)–(1.2) with α and $N(y,\alpha)$ replaced by $\alpha(y)$ and $N(y,\alpha(y))$ respectively. Note that G_m can be interpreted as a mixture of one-dimensional stable distributions, say $G_{\alpha(L)}$, with the varying characteristic exponent $\alpha = \alpha\big(L\big)$ such that $G_{\alpha(L)}$ degenerates in a one-dimensional subspace $L \subset \mathbb{R}^d$, and any stable distribution $G = G_\alpha$ as a similar mixture with the constant α.

Let

$$\omega\big(\Gamma,\tau\big) = \inf_{|t|=1} \int_{S_{d-1}} \big|\big\langle t,y\big\rangle\big|^{\alpha(y)}\,\tau^{\alpha(y)}\,\Gamma(dy), \qquad \tau>0. \tag{1.19}$$

If α is a constant then $\omega\big(\Gamma,1\big) = \varkappa^\alpha\big(\Gamma\big)$ (cf. (1.8)). Let $\tau(\Gamma)$ denote a solution of the equation $\omega\big(\Gamma,\tau\big)=1$. Define the following counterpart of $K_\alpha\big(\Gamma\big)$ (cf. (1.9))

$$K(\Gamma) = \varkappa^d(\Gamma)\,\max\big\{\tau^{2d+1}(\Gamma),\ \tau^{\delta d+1}(\Gamma)\big\}. \tag{1.20}$$

Theorem 4. *Assume that $\omega(\Gamma,1)>0$ and that $2 \geq \alpha(y) \geqslant \delta > 0$, for some $\delta>0$. Then there exists a unique solution, say $\tau(\Gamma)$, of the equation $\omega\big(\Gamma,\tau\big)=1$ and*

$$\min\big\{\omega^{-1/\delta},\ \omega^{-1/2}\big\} \leq \tau(\Gamma) \leq \max\big\{\omega^{-1/\delta},\ \omega^{-1/2}\big\}, \qquad \omega := \omega(\Gamma,1). \tag{1.21}$$

Furthermore, assume that either G_m is a mixture of symmetric distributions (that is, $N(y,\alpha(y)) \equiv 1$), or that $|\alpha(y)-1|>\delta$, for all $y \in S_{d-1}$. Then there exists a constant $c(\delta,d)$ such that

$$\chi_1(g_m) \leqslant c(\delta,d)\,K\big(\Gamma\big), \tag{1.22}$$

$$\eta\big(\mathcal{A}_c, G_m\big) \leqslant c(\delta,d)\,K\big(\Gamma\big). \tag{1.23}$$

2. Proofs

We start this section with an auxiliary lemma. Fourier transforms we denote as

$$\widehat{f}(t) := \int_{\mathbb{R}^d} f(x)\exp\big\{i\big\langle x,t\big\rangle\big\}\,dx\,.$$

In particular, we have $\varphi = \widehat{g}$, see (1.1)–(1.2).

Lemma 3. *Let G be a stable non-degenerate distribution on $\mathbb{R}^d$. Then its density g is a function of the class $C^\infty\big(\mathbb{R}^d\big)$. Furthermore, the derivatives $g^{(s)}(x)w_1\dots w_s$ as functions of x are square integrable and vanish as $|x|\to\infty$, for any $w_1,\dots,w_s \in \mathbb{R}^d$.*

Proof. Without loss of generality we can assume that $|w_j| \leqslant 1$, for all j. Then

$$\big|\big\langle t,w_1\big\rangle\cdot\ldots\cdot\big\langle t,w_s\big\rangle\,\varphi(t)\big| \leqslant |t|^s\,\big|\varphi(t)\big| \leqslant |t|^s\,\exp\big\{-\varkappa^\alpha(\Gamma)\,|t|^\alpha\big\}. \tag{2.1}$$

The bound (2.1) implies that the functions $t \mapsto \langle t, w_1 \rangle \cdot \ldots \cdot \langle t, w_s \rangle \varphi(t) : \mathbb{R}^d \to \mathbb{C}^d$ are in $L_2(\mathbb{R}^d) \cap L_1(\mathbb{R}^d)$. Hence, using $(\widehat{d_w g})(t) = -i\langle t, w\rangle \widehat{g}(t)$, we easily derive all assertions of the lemma. $\square$

Unfortunately, Lemma 31 does not imply directly the integrability of the derivatives of g. A proof of this integrability is rather involved, see the proof of Theorem 1 below.

Proof of Theorem 1. The estimate (1.12) and the analyticity of g is contained as Theorem 1 in Bogachev (1986). For the sake of completeness, let us prove (1.12). Let $Y, Y_1, \ldots, Y_2$ be i.i.d. random vectors such that $\mathcal{L}(Y) = G$. Since G is α-stable, we have $\mathcal{L}(s^{-1/\alpha}(Y_1 + \cdots + Y_s)) = G$, for any $s \in \mathbb{N}$. Hence, the function g is an s-fold convolution, say g_s^{*s}, of the function

$$g_s(x) := s^{d/\alpha}\, g(s^{1/\alpha} x).$$

Using $\|f * g\|_1 \leqslant \|f\|_1 \|g\|_1$, where $\|f\|_1$ denotes the $L_1(\mathbb{R}^d)$-norm of f, we have

$$\begin{aligned}\chi(g, w_1, \ldots, w_s) &= \int_{\mathbb{R}^d} \left| d_{w_1} \ldots d_{w_s} \left(g_s^{*s}(x)\right) \right| dx = \int_{\mathbb{R}^d} \left| \left(d_{w_1} g_s * \cdots * d_{w_s} g_s\right)(x) \right| dx \\ &\leqslant \prod_{j=1}^{s} \|d_{w_j} g_s\|_1 \leqslant \left(s^{1/\alpha} \chi_1(g)\right)^s,\end{aligned}$$

whence (1.12) follows.

It remains to prove the bound (1.11) for $\chi_1(g)$. Let us describe the idea of the proof. It is based on integration by parts repeating this integration d times. In the case of the standard Gaussian distribution $G = \Phi$ one can proceed as follows. Write

$$\widehat{\mathbf{I}}(t;\, A) = (1 - \Delta)^d\, \widehat{\mathbf{I}}_*(t;\, A)$$

with

$$\mathbf{I}_*(x;\, A) := \mathbf{I}(x;\, A)/(1 + |x|^2)^d,$$

where $\widehat{\mathbf{I}}(t;\, A)$ denotes the Fourier transform of the function $x \mapsto \mathbf{I}(x;\, A)$ and Δ is the Laplace operator. The function $\mathbf{I}_*$ is integrable and $\left|\widehat{\mathbf{I}}_*(t)\right| \leqslant c_d$ with some constant c_d depending on the dimension. Integrating by parts we reduce (cf. (2.9)–(2.14)) the estimation of $\chi_1(g)$ to a proof that

$$\sup_{|w|=1} \int_{\mathbb{R}^d} \left|(1 - \Delta)^d \langle t, w\rangle \varphi(t)\right| dt \leqslant c_d\,,$$

which clearly holds since $\varphi(t) = \exp\left\{-|t|^2/2\right\}$. In the non-Gaussian case $\alpha < 2$ such simple arguments are not applicable since the characteristic function φ is not sufficiently smooth. For example, for $\alpha < 1$, it is differentiable at most once. This non-differentiability enforces us to use a complicated construction of a sequence of vector fields, say $b_1, \ldots, b_d$. Each of the fields b_j depends on $b_1, \ldots, b_{j-1}$ and on many other variables related to the construction. To the operator $(1-\Delta)^d$ corresponds a differential operator of the form $P(\partial) := (1 + i\, d_{b_1}) \ldots (1 + i\, d_{b_d})$ and to the function $\mathbf{I}_*$ a function of the type

$$\mathbf{I}\big(x;\, A\big) \prod_{i=1}^{d} \big(1 + \big|\big\langle x, b_i \big\rangle\big|\big)^{-1}. \tag{2.2}$$

We choose the fields b_j such that the function (2.2) is in $L_p\big(\mathbb{R}^d\big)$, for some $p > 1$. Furthermore, we have to construct the operator $P(\partial)$ such that the expression $P(\partial)\, \varphi(t)$ depends only on the first order derivatives of $\log \varphi(t)$.

Let us return to the proof of (1.11). Consider the normalized measure $\sigma = \Gamma / \varkappa^{\alpha}\big(\Gamma\big)$ and define the function

$$\psi(t) = \exp\Big\{ - \int_{S_{d-1}} H\big(\big\langle t, y \big\rangle\big)\, \sigma(dy) \Big\} \tag{2.3}$$

with

$$H(z) = |z|^{\alpha} \big(1 - i\ \mathrm{sign}(z)\ \tan(\pi\alpha/2)\big), \qquad \alpha \neq 1, \tag{2.4}$$

$$H(z) = |z| \Big(1 + (2\,i/\pi)\, \mathrm{sign}(z)\ \log\big|\, z / \varkappa\big(\Gamma\big) \big|\Big), \quad \alpha = 1. \tag{2.5}$$

Notice that $\varkappa(\sigma) = 1$ and $\varkappa_0(\sigma) = \varkappa_0\big(\Gamma\big)/\varkappa^{\alpha}\big(\Gamma\big)$.

Introduce the class $D \subset C^{\infty}\big(\mathbb{R}^d\big)$ of the functions v which satisfy

$$|v^{(s)}(t) w_1 \ldots w_s| \leqslant 1 + |t|, \ \text{ for all } s = 0, 1, \ldots \ \text{ and } \ |w_i| \leqslant 1.$$

Assuming that a set $A \subset \mathbb{R}^d$ is bounded, define

$$J_0\big(A\big) := \sup_{v \in D} \Big| \int_{\mathbb{R}^d} \overline{\widehat{\mathbf{I}}}(t;\, A)\, v(t)\, \psi(t)\, dt \Big|. \tag{2.6}$$

By $\widehat{\mathbf{I}}(t;\, A)$ we denote the Fourier transform of the function $x \mapsto \mathbf{I}(x;\, A)$, and $\overline{z}$ is the complex conjugate of z. Let us show that (1.11) is implied by the following bounds

$$J_0\big(A\big) \leqslant 2^{-1} (2\pi)^{-d}\, c(\alpha, d)\, \varkappa_0^d(\sigma), \qquad \alpha \neq 1, \tag{2.7}$$

$$J_0\big(A\big) \leqslant 2^{-1} (2\pi)^{-d}\, c(\alpha, d)\, \varkappa_0^d(\sigma)\, \Big(1 + \big|\log \varkappa\big(\Gamma\big)\big|\Big)^d, \quad \alpha = 1, \tag{2.8}$$

respectively, where the constant $c(\alpha, d)$ satisfies (1.13) (we shall prove (2.7) and (2.8) below, for any bounded measurable subset $A \subset \mathbb{R}^d$). The definition (1.7) yields

$$\chi_1(g) = \sup \big\{ \chi_1(g, w) : \ w \in \mathbb{R}^d, \ |w| = 1 \big\}. \tag{2.9}$$

It is clear that

$$\chi_1(g,w)=\int\limits_{\mathbb{R}^d}|d_wg(x)|\,dx=\sup_A\int\limits_A|d_wg(x)|\,dx\leqslant 2\sup_A\Big|\int\limits_A d_wg(x)\,dx\Big|=2\,J_* \tag{2.10}$$

with

$$J_*=\sup_A\Big|\int\limits_{\mathbb{R}^d}\mathbf{I}(x;\,A)\,d_wg(x)\,dx\,\Big|, \tag{2.11}$$

where $\sup\limits_A$ is taken over all bounded sets $A\subset\mathbb{R}^d$. Parseval's equality and the well known properties of the Fourier transforms imply

$$J_*=(2\pi)^d\sup_A\Big|\int\limits_{\mathbb{R}^d}\overline{\widehat{\mathbf{I}}}(t;\,A)\,\big\langle t,w\big\rangle\,\varphi(t)\,dt\,\Big|. \tag{2.12}$$

Changing variables in (2.12) as $t=u/\varkappa\big(\Gamma\big)$, we obtain

$$J_*=(2\pi)^d\varkappa^{-1}\big(\Gamma\big)\,J_{**} \tag{2.13}$$

with

$$J_{**}=\sup_A\Big|\int\limits_{\mathbb{R}^d}\overline{\widehat{\mathbf{I}}}(t;\,A)\,\big\langle t,w\big\rangle\,\psi(t)\,dt\,\Big|, \tag{2.14}$$

where ψ is defined by (2.3). Obviously, the function $t\mapsto\big\langle t,w\big\rangle$ belongs to the class D since $|w|=1$. Hence, $J_{**}\leqslant\sup\limits_A J_0\big(A\big)$. This inequality combined with (2.9)–(2.14) and $\varkappa_0(\sigma)=\varkappa_0\big(\Gamma\big)/\varkappa^\alpha\big(\Gamma\big)$ shows that instead of (1.11) it suffices to prove the estimates (2.7) and (2.8).

Let us prove (2.7) and (2.8). We shall integrate by parts d times. Let us start with a description of the first integration by parts. Choose any vector $b_1\in R^d$ such that $|b_1|=1$. Splitting the set $A=B_1\cup B_2$, where

$$B_1=\{u:\ \big\langle b_1,u\big\rangle\geqslant 0\}\cap A\quad\text{and}\quad B_2=\{u:\ \big\langle b_1,u\big\rangle<0\}\cap A,$$

we can write the following obvious identity

$$\mathbf{I}\big(x;\,A\big)=\mathbf{I}_1\big(x;\,B_1\big)\big(1+\big\langle b_1,x\big\rangle\big)+\mathbf{I}_1\big(x;\,B_2\big)\big(1-\big\langle b_1,x\big\rangle\big) \tag{2.15}$$

with

$$\mathbf{I}_1\big(x;\,C\big)=\frac{\mathbf{I}\big(x;\,C\big)}{1+\big|\big\langle b_1,x\big\rangle\big|},\qquad C=B_1,B_2.$$

Using the well known properties of the Fourier transform, the identity (2.15) yields

$$\widehat{\mathbf{I}}\big(t;\,A\big)=(1-id_{b_1})\,\widehat{\mathbf{I}}_1\big(t;\,B_1\big)+(1+id_{b_1})\,\widehat{\mathbf{I}}_1\big(t;\,B_2\big). \tag{2.16}$$

Let $h(z) := H'(z)$ denote the derivative of the function $H(z)$ defined by (2.4) and (2.5), that is,

$$h(z) = \alpha\,|z|^{\alpha-1}\,\mathrm{sign}(z)\big(1 - i\,\mathrm{sign}(z)\,\tan(\pi\alpha/2)\big), \qquad \alpha \neq 1, \tag{2.17}$$

$$h(z) = \mathrm{sign}(z) + (2\,i/\pi) + (2\,i/\pi)\,\log\big|\,z/\varkappa\big(\Gamma\big)\big|, \qquad \alpha = 1. \tag{2.18}$$

Using (2.6) and (2.16), we obtain

$$J_0\big(A\big) \leqslant 2 \max_{B=B_1,\,B_2}\ \sup_{v\in D}\ \big(Q_1 + Q_2\big)\,, \tag{2.19}$$

where

$$Q_1 = \Big|\int_{\mathbb{R}^d} \overline{\widehat{\mathbf{I}}}_1\big(t;\,B\big)\,v(t)\,\psi(t)\,dt\Big|, \qquad Q_2 = \Big|\int_{\mathbb{R}^d} d_{b_1}\overline{\widehat{\mathbf{I}}}_1\big(t;\,B\big)\,v(t)\,\psi(t)\,dt\Big|.$$

Introducing the Dirac measure δ_b on $\mathbb{R}^d$ such that $\delta_b\big(C\big) = \mathbf{I}\big(b \in C\big)$, for $C \subset \mathbb{R}^d$, we can write

$$Q_1 = \int_{S_{d-1}} \Big|\int_{\mathbb{R}^d} \overline{\widehat{\mathbf{I}}}_1\big(t;\,B\big)\,v(t)\,\psi(t)\,\langle y_1, b_1\rangle\,dt\,\Big|\,\delta_{b_1}(dy_1) \tag{2.20}$$

since $\big\langle b_1, b_1\big\rangle = 1$. Integrating by parts, using $|b_1| = 1$ and $d_{b_1}v \in D$ together with

$$d_{b_1}\psi(t) = \int_{S_{d-1}} \psi(t)\,h\big(\big\langle t, y_1\big\rangle\big)\big\langle y_1, b_1\big\rangle\,\sigma(dy_1)\,,$$

we have

$$Q_2 \leqslant \int_{S_{d-1}} \Big|\int_{\mathbb{R}^d} \overline{\widehat{\mathbf{I}}}_1\big(t;\,B\big)\,v(t)\,\psi(t)\,h\big(\big\langle t, y_1\big\rangle\big)\big\langle y_1, b_1\big\rangle\,dt\,\Big|\,\sigma(dy_1) + \sup_{v\in D}\ Q_1\,. \tag{2.21}$$

In order to rewrite the bounds (2.19)–(2.21) in a more compact form and to proceed with integration by parts, let us introduce additional notation. We shall denote by $\Omega \subset \Omega_j$ subsets of the set $\Omega_j = \{1, 2, \ldots, j\}$. Collecting (2.19)–(2.21) we obtain

$$J_0\big(A\big) \leqslant 6\,J_1\big(A\big), \tag{2.22}$$

where

$$J_1\big(A\big) = \max_{\Omega\subset\Omega_1}\ \sup_{v\in D}\ \sup_{B\subset A}\ \int_{S_{d-1}} \big|J_1^*\big(B\big)\big| \prod_{k\in\Omega}\sigma(dy_k) \prod_{l\in\Omega_1\setminus\Omega}\delta_{b_l}(dy_l)$$

with

$$J_1^*\big(B\big) = \int_{\mathbb{R}^d} \overline{\widehat{\mathbf{I}}}_1\big(t;\,B\big)\,v(t)\,\psi(t)\,\langle y_1, b_1\rangle \prod_{k\in\Omega} h\big(\big\langle t, y_k\big\rangle\big)\,dt.$$

The inequality (2.22) provides the result of the first integration by parts. Let $2 \leqslant j \leqslant d$ and $y_1, \ldots, y_d \in S_{d-1}$. In order to describe further integrations by parts, consider vector valued measurable functions

$$b_1,\ b_2 = b_2(y_1), \ldots,\ b_d = b_d(y_1, \ldots, y_{d-1}) \tag{2.23}$$

such that $b_j \in S_{d-1}$, and

$$\big\langle b_j, y_l \big\rangle = 0, \tag{2.24}$$

for all $1 \leqslant j \leqslant d$ and $1 \leqslant l \leqslant j-1$. It is clear that such functions b_j exist. Denote

$$J_j\big(A\big) = \max_{\Omega\subset\Omega_j} \sup_{v\in D} \sup_{B\subset A} \int\limits_{S_{d-1}} \dots \int\limits_{S_{d-1}} \big|J_j^*\big(B\big)\big| \prod_{k\in\Omega} \sigma(dy_k) \prod_{l\in\Omega_j\setminus\Omega} \delta_{b_l}(dy_l), \tag{2.25}$$

where $\Omega_j = \{1, \dots, j\}$, and

$$J_j^*\big(B\big) = \int\limits_{\mathbb{R}^d} \bar{\bar{\mathbf{I}}}_j\big(t;\, B\big)\, v(t)\, \psi(t) \prod_{l=1}^{j} \big\langle y_l, b_l\big\rangle \prod_{k\in\Omega} h\big(\big\langle t, y_k\big\rangle\big)\, dt$$

with

$$\mathbf{I}_j\big(x;\, B\big) = \mathbf{I}\big(x;\, B\big) \prod_{i=1}^{j} \big(1 + \big|\big\langle x, b_i\big\rangle\big|\big)^{-1}.$$

The integrals $J_j\big(A\big)$ defined by (2.25) satisfy

$$J_{j-1}\big(A\big) \leqslant 6\, J_j\big(A\big), \ \text{ for all } j = 1, \dots, d\,. \tag{2.26}$$

To see that (2.26) holds indeed, notice that our choice of b_j as in (2.23) and (2.24) guarantees that

$$d_{b_j} \prod_{k\in\Omega} h\big(\big\langle t, y_k\big\rangle\big) = 0, \quad \text{for } \Omega \subset \Omega_{j-1}\,.$$

Hence, in order to prove (2.26) we can estimate $J_{j-1}\big(A\big)$ proceeding similarly as in (2.15)–(2.21), which led to the bound (2.22) for $J_0\big(A\big)$.

In particular, the bound (2.26) yields

$$J_0\big(A\big) \leqslant 6^d\, J_d\big(A\big), \tag{2.27}$$

where $J_d\big(A\big)$ is defined by (2.25) with $j = d$ and

$$J_d^*\big(B\big) = \prod_{l=1}^{d} \big\langle y_l, b_l\big\rangle \int\limits_{\mathbb{R}^d} \bar{\bar{\mathbf{I}}}_d\big(t;\, B\big)\, v(t)\, \psi(t) \prod_{k\in\Omega} h\big(\big\langle t, y_k\big\rangle\big)\, dt. \tag{2.28}$$

We shall prove that

$$\big|\, J_d^*\big(B\big)\big| \leqslant c_*(\alpha, d), \ \alpha \neq 1, \ \text{ and } \ \big|\, J_d^*\big(B\big)\big| \leqslant c_*(1, d)\Big(1 + \big|\log \varkappa\big(\Gamma\big)\big|\Big)^d, \ \alpha = 1. \tag{2.29}$$

The constant $c_*(\alpha, d)$ is specified below, see (2.44) and the text below, where estimates (1.13)–(1.14) for $c(\alpha, d)$ are proved. Using the definition (2.25) of $J_d\big(A\big)$ and integrating the bounds (2.29) with respect to the measure σ on S_{d-1}, we obtain

$$J_d\big(A\big) \leqslant c_*(\alpha, d) \max_{1\leqslant i\leqslant d} \varkappa_0^i(\sigma) \leqslant c_*(\alpha, d)\, \varkappa_0^d(\sigma), \quad \alpha \neq 1, \tag{2.30}$$

which combined with the inequality (2.27) proves (2.7). While proving (2.30) we used $1 = \varkappa(\sigma) \leqslant \varkappa_0^\alpha(\sigma)$. Similarly, integrating the second inequality in (2.29), we derive (2.8).

To conclude the proof of the theorem we have to verify (2.29). Consider the matrix $\mathbb{E} = \big(\langle y_i, b_j\rangle\big)_{i,j=1,\dots,d}$. By our choice (see (2.23) and (2.24)) of the vectors b_j, all entries above the diagonal of the matrix $\mathbb{E}$ are equal to zero. Therefore

$$\det \mathbb{E} = \prod_{l=1}^{d} \big\langle y_l, b_l \big\rangle \tag{2.31}$$

and it is clear that

$$\big|\det \mathbb{E}\big| \leqslant 1 \tag{2.32}$$

since $\big|\big\langle y_l, b_l\big\rangle\big| \leqslant 1$, for $|b_l| = |y_l| = 1$.

If $\det \mathbb{E} = 0$ then $J_d^*\big(B\big) = 0$ (cf. (2.28) and (2.31)) and (2.29) is obviously fulfilled. Hence, without loss of generality we may assume in the proof of (2.29) that $\det \mathbb{E} \neq 0$. Let $\|f\|_p$ stand for the $L_p\big(\mathbb{R}^d\big)$ norm of a function f. Using Hölder's inequality with $1/p + 1/q = 1$ such that $1 < p \leqslant 2,\ q \geqslant 2$, the relation (2.28) yields

$$\big|\, J_d^*\big(B\big)\,\big| \leqslant \big|\det \mathbb{E}\big|\, \|\widehat{\mathbf{I}}_d\big(\,\cdot\,; B\big)\|_q\, \|\vartheta\|_p\,, \tag{2.33}$$

where we denote for brevity

$$\vartheta(t) = v(t)\,\psi(t) \prod_{k\in\Omega} h\big(\big\langle t, y_k\big\rangle\big)\,.$$

To estimate $\|\widehat{\mathbf{I}}_d\big(\,\cdot\,; B\big)\|_q$ we shall use the fact that the Fourier transform is a bounded operator from $L_p\big(\mathbb{R}^d\big)$ to $L_q\big(\mathbb{R}^d\big)$, $1 < p \leqslant 2$. More precisely, we have $\|\widehat{f}\|_q \leqslant \|f\|_p$ (see the classical Riesz–Thorin interpolation theorem in Triebel 1978), whence

$$\|\widehat{\mathbf{I}}_d\big(\,\cdot\,; B\big)\|_q \leqslant \|\mathbf{I}_d\big(\,\cdot\,; B\big)\|_p\,.$$

Changing variables $t = u$ with $u = (u_1, \dots, u_d)$ such that $u_i = \big\langle t, b_i\big\rangle$ and introducing the matrix

$$\mathbb{B}{:}= (b_{i,j})_{i,j=1,\dots,d}, \qquad b_i = (b_{i,1}, \dots, b_{i,d})\,, \tag{2.34}$$

we obtain

$$\|\widehat{\mathbf{I}}_d\big(\,\cdot\,; B\big)\|_q \leqslant \|\mathbf{I}_d\big(\,\cdot\,; B\big)\|_p \leqslant \big|\det \mathbb{B}\big|^{-1/p}\, c_0 \tag{2.35}$$

with

$$c_0 = \Big(\int_{\mathbb{R}} (1+|u|)^{-p} du\Big)^{d/p} = \big(2/(p-1)\big)^{d/p}. \tag{2.36}$$

Let us estimate $\|\vartheta\|_p$. The function v belongs to the class D and $\varkappa(\sigma) = 1$. Therefore $\big|\psi(t)\big| \leqslant \exp\big\{-|t|^\alpha\big\}$ and we get

$$\big|v(t)\big|\,\big|\psi(t)\big| \leqslant (1+|t|)\,\big|\psi(t)\,\big| \leqslant c_1 \exp\big\{-|t|^\alpha/2\big\}, \tag{2.37}$$

where $c_1 = 1 + \alpha^{-1/\alpha}$. The estimate (2.37) and the obvious inequality

$$d\,|t|^\alpha = \sum_{i=1}^{d} |t|^\alpha \geqslant \sum_{i=1}^{d} \bigl|\bigl\langle t, y_i\bigr\rangle\bigr|^\alpha, \quad \text{for } |y_i| = 1, \tag{2.38}$$

yield

$$\|\vartheta\|_p^p \leqslant c_1^p \int\limits_{\mathbb{R}^d} \prod_{j=1}^{d} \exp\bigl\{-\tfrac{p}{2d}|\langle t, y_j\rangle|^\alpha\bigr\} \prod_{k\in\Omega} \bigl|h\bigl(\langle t, y_k\rangle\bigr)\bigr|^p dt. \tag{2.39}$$

Changing in (2.39) variables $t = u$ with $u = (u_1, \dots, u_d)$ such that $u_i = \bigl\langle t, y_i\bigr\rangle$ and introducing the matrix

$$\mathbb{A} := (y_{j,i})_{i,j=1,\dots,d}, \qquad y_i = (y_{i,1}, \dots, y_{i,d}), \tag{2.40}$$

we get (notice that $0 \leqslant |\Omega| \leqslant d$ and use $a^\beta b^\gamma \leqslant a + b$, for nonnegative β and γ such that $\beta + \gamma = 1$)

$$\|\vartheta\|_p \leqslant \bigl|\det \mathbb{A}\bigr|^{-1/p} c_1\, c_2^{|\Omega|/p}\, c_3^{(d-|\Omega|)/p} \leqslant \bigl|\det \mathbb{A}\bigr|^{-1/p} c_1 \bigl(c_2^{d/p} + c_3^{d/p}\bigr), \tag{2.41}$$

where

$$c_2 = \int\limits_{\mathbb{R}} \exp\bigl\{-\tfrac{p}{2d}|u|^\alpha\bigr\}\,|h(u)|^p\,du, \qquad c_3 = \int\limits_{\mathbb{R}} \exp\bigl\{-\tfrac{p}{2d}|u|^\alpha\bigr\}\,du. \tag{2.42}$$

Multiplying the matrices $\mathbb{A}$ and $\mathbb{B}$ we see that $\mathbb{A}\mathbb{B} = \mathbb{E}$. In particular, both changes (2.34) and (2.40) of variables are well defined since $\det \mathbb{E} = \det \mathbb{A} \det \mathbb{B}$ and we assume that $\det \mathbb{E} \neq 0$. The relation $\det \mathbb{A} \det \mathbb{B} = \det \mathbb{E}$ combined with the inequalities (2.33), (2.35), (2.41) yields

$$\bigl|J_d^*\bigl(B\bigr)\bigr| \leqslant c_0\, c_1 \bigl(c_2^{d/p} + c_3^{d/p}\bigr) \bigl|\det \mathbb{E}\bigr|^{1-1/p} \leqslant c_0\, c_1 \bigl(c_2^{d/p} + c_3^{d/p}\bigr), \tag{2.43}$$

since (2.32) implies $\bigl|\det \mathbb{E}\bigr|^{1-1/p} \leqslant 1$, for $p \geqslant 1$. The inequality (2.29) follows from (2.43) with

$$c_*(\alpha, d) = c_0\, c_1 \bigl(c_2^{d/p} + c_3^{d/p}\bigr), \qquad \alpha \neq 1, \tag{2.44}$$

$$c_*(1, d) = \sup_{\varkappa(\Gamma)} \Bigl(\bigl(1 + \bigl|\log \varkappa\bigl(\Gamma\bigr)\bigr|\bigr)^{-d} c_0\, c_1 \bigl(c_2^{d/p} + c_3^{d/p}\bigr)\Bigr), \quad \alpha = 1,$$

and with $c_*(\alpha, d)$ as in (2.44) in the symmetric case, for all α. The bound (2.29) yields (2.7) and (2.8) with

$$c(\alpha, d) \leqslant 2\,(12\pi)^d\, c_*(\alpha, d).$$

In order to prove bounds (1.13) and (1.14) for $c(\alpha, d)$ we have to estimate $c_*(\alpha, d)$. To bound $c_*(\alpha, d)$ it suffices to estimate constants c_2 and c_3 (they are defined in (2.42)) since c_0 is given by explicit formula (2.36) and $c_1 = 1 + 1/\alpha$. This estimation is in essence elementary, however somewhat cumbersome. Therefore, let us describe only some hints used to this end. Recall that $1 < p \leqslant 2$. The constant c_3 is obviously finite and can be simply estimated. Using (2.42) and the definition (2.17)–(2.18) of the function h, we have

$$\begin{aligned} c_2 &= \int_{\mathbb{R}} \exp\big\{-\tfrac{p}{2d}|u|^\alpha\big\}\, |h(u)|^p\, du \\ &\leqslant \alpha^p \big(1 + |\tan(\pi\alpha/2)|\big)^p \int_{\mathbb{R}} \exp\big\{-\tfrac{p}{2d}|u|^\alpha\big\}\, |u|^{p(\alpha-1)}\, du, \end{aligned} \tag{2.45}$$

for $\alpha \neq 1$, and

$$\begin{aligned} c_2 \leqslant 4^p \big(1 + \big|\log \varkappa\big(\Gamma\big)\big|\big)^p &\int_{\mathbb{R}} \exp\big\{-\tfrac{p|u|}{2d}\big\}\, du \\ &+ 2^p \int_{\mathbb{R}} \exp\big\{-\tfrac{p|u|}{2d}\big\}\, |\log|u||^p\, du, \end{aligned} \tag{2.46}$$

for $\alpha = 1$. The integrals in (2.45) and (2.46) exist if $p(\alpha - 1) > -1$. Therefore, we can choose

$$p = 2, \ \text{ for } \alpha \geqslant 1, \qquad p = \frac{3}{2}, \text{ for } \frac{1}{2} \leqslant \alpha < 1, \quad p = \frac{1}{2}\Big(1 + \frac{1}{1-\alpha}\Big), \text{ for } 0 < \alpha < \frac{1}{2}. \tag{2.47}$$

Then, in particular, we change the variables $cu^\alpha = y$ and apply the following inequality $y^\beta \exp(-y) \leqslant c(\beta) \exp(-y/2)$. The symmetric case is less complicated since in this case $h(z) = \alpha\,|z|^{\alpha-1}$ which is simpler than h defined by (2.17)–(2.18). $\square$

Proof of Theorem 2. We shall use a reduction to polyhedrons with finite number of faces as in Bhattacharya and Rao (1986). A convex set P we call a polyhedron if there exist distinct unit vectors $u_1, \dots, u_m \in S_{d-1}$ and $d_1, \dots, d_m \in \mathbb{R}$ such that

$$P = \big\{x \in \mathbb{R}^d : \big\langle u_j, x\big\rangle \leqslant d_j,\ 1 \leqslant j \leqslant m\big\}. \tag{2.48}$$

Let $\mathcal{P}$ denote the class of compact sets of the form (2.48) with non-empty interior. An inspection of the proof of Theorem 3.1 in Bhattacharya and Rao (1986) shows that

$$\frac{1}{2\varepsilon} \int_{(\partial A)^\varepsilon} g(x)\, dx \leqslant \zeta\big(\mathcal{P}, g\big), \quad \varepsilon > 0, \tag{2.49}$$

for $A \in \mathcal{A}_c$ and continuous g.

Let us prove (1.15). Taking in (2.49) the limit as $\varepsilon \to 0$ yields $\zeta\big(\mathcal{A}_c, g\big) \leqslant \zeta\big(\mathcal{P}, g\big)$, which together with the obvious reverse inequality $\zeta\big(\mathcal{A}_c, g\big) \geqslant \zeta\big(\mathcal{P}, g\big)$ implies the relation $\zeta\big(\mathcal{A}_c, g\big) = \zeta\big(\mathcal{P}, g\big)$. Dividing by 2ε the inequality

$$\int_{(\partial A)^\varepsilon} g(x)\, dx \leqslant \varepsilon\, \eta\big(\mathcal{A}_c, g\big), \qquad \text{for } A \in \mathcal{A}_c,$$

and passing to the limit as $\varepsilon \to 0$ yields $\zeta\big(\mathcal{A}_c, g\big) \leqslant \eta\big(\mathcal{A}_c, g\big)/2$. The inequality (2.49) means that $\eta\big(\mathcal{A}_c, g\big) \leqslant 2\,\zeta\big(\mathcal{P}, g\big) = 2\,\zeta\big(\mathcal{A}_c, g\big)$, and (1.15) follows.

Let us prove (1.16). Due to (1.15) and $\zeta\big(\mathcal{A}_c, g\big) = \zeta\big(\mathcal{P}, g\big)$ it suffices to verify that

$$\int\limits_{\partial P} g(x)\, ds \leqslant 2\, d^{3/2}\, \chi_1(g), \qquad \text{for } P \in \mathcal{P}. \tag{2.50}$$

Let $n(x)$ denote the unit outer normal vector of ∂P at point $x \in \partial P$. The normal is defined for almost all $x \in \partial P$ with respect to the surface measure ds on ∂P. Introducing the standard orthonormal vectors $e_1, \dots, e_d$ in $\mathbb{R}^d$ and writing $e_{-i} = -e_i$, $e_0 = 0$, it is clear that

$$\partial P \subset \bigcup_{i=-d}^{d} Q_i, \quad \text{where} \quad Q_i = \big\{ x \in \partial P : \big\langle n(x), e_i \big\rangle \geqslant d^{-1/2} \big\}, \tag{2.51}$$

since any vector $n \in S_{d-1}$ has at least one coordinate, say n_j, such that $|n_j| \geqslant d^{-1/2}$. Using the representation (2.51), we have

$$\int\limits_{\partial P} g(x)\, ds \leqslant \sum_{i=-d}^{d} \int\limits_{Q_i} g(x)\, ds. \tag{2.52}$$

The inequality $\big\langle n(x), e_i \big\rangle \geqslant d^{-1/2}$ yields

$$\int\limits_{Q_i} g(x) dx \leqslant d^{1/2} \int\limits_{Q_i} \big\langle n(x), e_i \big\rangle\, g(x)\, ds. \tag{2.53}$$

According to the construction (2.51) of the surface Q_i, it is the graph of a piecewise linear function, say $x_i = f(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d)$. Introducing the subgraph $V_i = Q_i + (-\infty, 0]\, e_i$ of the graph of f and applying Stokes' Theorem or just using Fubini's Theorem and integrating with respect to the i-th coordinate of x, we have

$$\begin{aligned} \int\limits_{Q_i} \big\langle n(x), e_i \big\rangle\, g(x)\, ds &= \int\limits_{\partial V_i} \big\langle n(x), e_i \big\rangle\, g(x)\, ds = \int\limits_{V_i} d_{e_i} g(x)\, dx \\ &\leqslant \int\limits_{\mathbb{R}^d} |d_{e_i} g(x)|\, dx \leqslant \chi_1(g). \end{aligned} \tag{2.54}$$

Combining (2.52)–(2.54), we obtain (2.50). $\square$

Proof of Theorem 3. Write $a = \eta\big(\mathcal{A}_c, G\big)$ and $B = \chi_1(g)$. Theorem 1 in a paper of Bloznelis (1988) says that

$$\Delta_n\big(\mathcal{A}_c\big) \leqslant c c(\alpha)\, (20)^{1/\alpha}\, n^{-1/\alpha} \big(\rho + a\,\zeta_1 + R_\alpha\, \zeta_{1+\alpha}\big), \tag{2.55}$$

where c is an absolute constant,

$$R_\alpha = a + (a+1)\, \big(B + B^2\big), \qquad \alpha \leqslant 1,$$

$$R_\alpha = \big(a + (a+1)\big(B + B^2\big)\big)\, B, \qquad 1 < \alpha \leqslant 2,$$

and $c(\alpha) = 1$, for $\alpha \le 1$, $c(\alpha) = |1-\alpha|$, for $\alpha > 1$. Let a random variable Y have the distribution G. Notice that $\Delta_n\big(\mathcal{A}_c\big)$ does not change if we replace $X_1,\dots,X_n$ by $\tau X_1,\dots,\tau X_n$ and the distribution G by the distribution of τY respectively, for any fixed $\tau > 0$. Similarly, the metric ρ remains invariant and to $a = \eta\big(\mathcal{A}_c, G\big)$, $B = \chi_1(g)$ and $\zeta_{1+\alpha}$ correspond a/τ, B/τ and $\tau^{1+\alpha}\zeta_{1+\alpha}$ respectively. Hence, the bound (2.55) yields an estimate for $\Delta_n\big(\mathcal{A}_c\big)$ as (2.55) but with R_α replaced by $\tau^{\alpha+1}R_\alpha(\tau)$, where

$$R_\alpha(\tau) = a/\tau + (a/\tau+1)\big(B/\tau + B^2/\tau^2\big), \qquad \alpha \leqslant 1,$$

$$R_\alpha(\tau) = \big(a/\tau + (a/\tau+1)\big(B/\tau + B^2/\tau^2\big)\big)\,B/\tau, \qquad 1 < \alpha \leqslant 2.$$

Choosing $\tau = B$, we see that (2.55) holds with R_α replaced by

$$B^{\alpha+1}\,R_\alpha\big(B\big) \leqslant 3\,(a+B)\,B^\alpha.$$

By Theorems 1 and 2 we can estimate

$$a \leqslant c\,d^{3/2}\,c(\alpha,d)\,K_\alpha\big(\Gamma\big), \qquad B \leqslant c(\alpha,d)\,K_\alpha\big(\Gamma\big),$$

and Theorem 3 follows. $\square$

Proof of Theorem 4. To prove (1.21) it suffices to use $\tau > 0$ and the definition (1.19) of $\omega(\Gamma,\tau)$. The bound (1.22) together with the estimate (1.16) in Theorem 2 yields (1.23). Therefore we have to prove (1.22) only. In order to prove (1.22) we can proceed similarly to the proof of Theorem 1. Let us point out necessary changes. First of all, we have to replace everywhere α by $\alpha(y)$.

The assumptions $\delta < \alpha(y) \le 2$ and the positivity of $\omega(\Gamma,1)$ guaranty that the function $\tau \mapsto \omega(\Gamma,\tau)$ is continuous strictly increasing and such that $\lim_{\tau\downarrow 0}\omega(\Gamma,\tau) = 0$ and $\lim_{\tau\to\infty}\omega(\Gamma,\tau) = \infty$. Hence, there exists a unique $0 < \tau_0 < \infty$ such that $\omega(\Gamma,\tau_0) = 1$.

Introduce the functions

$$\beta(u) = \min\big\{|u|^\delta,\ u^2\big\}, \qquad \gamma(u) = \max\big\{|u|^{\delta-1},\ |u|\big\}, \quad u \in \mathbb{R}. \tag{2.56}$$

A repetition of the proof which led to the reduction to integrals $J_0(A)$ produces now the inequality (cf. the proof of (2.7)–(2.8))

$$\chi_1(g_m) \le c_d\,\tau(\Gamma)\,\sup_A J_0(A). \tag{2.57}$$

Deriving (2.57), we use the change $t = \tau(\Gamma)\,u$ instead of the change of variables $t = u/\varkappa(\Gamma)$ used to prove (2.13). Therefore the measure σ is replaced by the measure $\sigma(dy) = \tau^{\alpha(y)}(\Gamma)\,\Gamma(dy)$. The function ψ now is defined by (2.3) with $\alpha(y)$ instead of α, since either we exclude the case $\alpha(y) = 1$ or assume that G is symmetric. We can word by word repeat the proof which led to (2.27). The only difference is that in the factors $h\big(\langle t, y_k\rangle\big)$ it is necessary to replace α by $\alpha(y_k)$. Below we shall prove that instead of (2.29) the following inequality holds:

$$\big|\, J_d^*\big(B\big)\,\big| \leqslant c(\delta, d). \tag{2.58}$$

This inequality implies (1.22) since using (2.57), the definition (2.25) of $J_d\big(A\big)$ and integrating the bound (2.58) with respect to the measure σ on S_{d-1}, we obtain

$$J_0(A) \le 6^d J_d(A) \leqslant c(\delta, d) \max_{1\leqslant i\leqslant d} \sigma^i(S_{d-1}) \leqslant c(\delta, d)\, \varkappa^d(\Gamma)\, \max\big\{\, \tau^{2d}(\Gamma),\ \tau^{\delta d}(\Gamma)\,\big\}$$

since

$$1 = \omega\big(\Gamma, \tau(\Gamma)\big) \le \sigma(S_{d-1}) \le\ \varkappa(\Gamma)\, \max\big\{\, \tau^2(\Gamma),\ \tau^\delta(\Gamma)\,\big\}.$$

It remains to verify (2.58). Using the definitions of $\omega\big(\Gamma, \tau\big)$ and β (see (1.20) and (2.56)), we have

$$\begin{aligned} \int_{S_{d-1}} \big|\langle t, y\rangle\big|^{\alpha(y)}\, \sigma(dy) &= \int_{S_{d-1}} \big|\langle t, y\rangle\big|^{\alpha(y)}\, \tau^{\alpha(y)}(\Gamma)\, \Gamma(dy) \\ &\ge \beta\big(|t|\big)\, \omega\big(\Gamma, \tau(\Gamma)\big) = \beta\big(|t|\big). \end{aligned} \tag{2.59}$$

Hence

$$\big|\, \psi(t)\,\big| = \exp\Big\{ -\int_{S_{d-1}} \big|\langle t, y\rangle\big|^{\alpha(y)}\, \sigma(dy)\Big\} \leqslant \exp\big\{ -\beta\big(|t|\big)\,\big\}. \tag{2.60}$$

Using $\big|1 - \alpha(y_k)\big| \geqslant \delta$ we have $\tan \dfrac{\pi\alpha}{2} \leqslant c(\delta)$. This inequality combined with the explicit formula (2.17) implies

$$\big|\, h\big(\langle t, y_k\rangle\big)\,\big| \leqslant c(\delta)\, \big|\langle t, y_k\rangle\big|^{\alpha(y_k)-1} \leqslant c(\delta)\, \gamma\big(\langle t, y_k\rangle\big). \tag{2.61}$$

The bound (2.61) holds as well for symmetric G without the assumption $\big|1 - \alpha(y_k)\big| \geqslant \delta$. The inequality (2.38) is replaced by

$$d\, \beta(t) \leqslant \sum_{i=1}^{d} \beta\big(\langle t, y_i\rangle\big). \tag{2.62}$$

The inequalities (2.60)–(2.62) allow to repeat the remaining part of the proof. Note that the constant c_1 is replaced by some constant $c(\delta)$. Using (2.60) and (2.61), the constants c_2 and c_3 now are

$$c_2 = \int_{\mathbb{R}} \exp\big\{ -\tfrac{p}{2d}\, \beta(u)\big\}\, \gamma(u)\, du\,, \qquad c_3 = \int_{\mathbb{R}} \exp\big\{ -\tfrac{p}{2d}\, \beta(u)\big\}\, du.$$

In order to show that $c_2, c_3 \le c(\delta, d)$ it suffices to use the definition (2.56) of the functions β and γ and to choose p as in (2.47) replacing α by δ. $\square$

References

Ball, K., *The reverse isoperimetric problem for Gaussian measure*, Discrete and Computational Geometry **10** (1993), 411–420.

Banys, I., *An estimate of the rate of convergence in the multidimensional integral limit theorem in the case of convergence to a stable symmetric law*, Litovsk. Mat. Sb. **11** (1971), 497–509.

Bentkus V., *Dependence of the Berry–Esseen estimate on the dimension*, Lithuanian Math. J. **26** (1986), no. 2, 110–114.

Bentkus, V.; Götze, F.; Paulauskas, V., *Bounds for the accuracy of Poissonian approximations of stable laws*, Stoch. Processes Appl. **65** (1996), no. 1, 55–68.

Bhattacharya, R.N.; Rao, R.R., *Normal approximation and asymptotic expansions*, Krieger, Malabar, Florida, 1986.

Bloznelis, M., *Rate of convergence to a stable law in the space* $\mathbb{R}^d$, Lithuanian Math. J. **28** (1988), no. 1, 21–29.

Bloznelis, M., *Nonuniform estimate of the rate of convergence to a stable law in the multidimensional central limit theorem*, Lithuanian Math. J. **29** (1989), no. 2, 97–109.

Bogachev, V.I., *Some results on differentiable measures*, Math. USSR Sb. **55** (1986), no. 2, 335–349.

Ledoux, M.; Paulauskas, V., *A rate of convergence in the Poissonian representation of stable distributions*, Lithuanian Math. J. **36** (1996), no. 4, 388–399.

Mikhailova, I. Yu., *The rate of approximation by a multidimensional stable law*, Vestnik Moskov. Univ. Ser. XV Vychisl. Mat. Kibernet. **3** (1983), 60–62.

Paulauskas, V., *The rate of convergence in the multidimensional limit theorem in the case of a stable limit law*, Litovsk. Mat. Sb. **15** (1975), no. 1, 207–228.

Rachev, S.T., *Probability metrics and the stability of stochastic models*, Wiley, New York, 1991.

Samorodnitsky, G.; Taqqu, M.S., *Stable non-Gaussian random processes* (*Stochastic Models with Infinite Variance*), Chapman and Hall, New York, 1994.

Sazonov, V., *The rate of convergence in the multidimensional central limit theorem*, Teor. Verojatnost. i Primenen. **13** (1968), 191–194.

Sazonov, V.V., *Normal Approximation—Some recent Advances*, Lect. Notes Math., vol. 879, Springer, Berlin, 1981.

Triebel, H., *Interpolation theory, function spaces, differential Operators*, North-Holland, Amsterdam, 1978.

Vidmantas Bentkus
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld 1
Germany
E-mail address: bentkus@mathematik.uni-bielefeld.de

Almantas Juozulynas
Department of Mathematics
University of Vilnius
Naugarduko 24
2006 Vilnius
Lithuania
E-mail address: almas@ieva.maf.vu.lt

Vygantas Paulauskas
Department of Mathematics
University of Vilnius
Naugarduko 24
2006 Vilnius
Lithuania
E-mail address: vpaul@ieva.maf.vu.lt