Rates of normal approximation for typical weighted sums

Sergey Bobkov
University of Minnesota

(based on joint work with
G. P. Chistyakov and F. Götze)

Workshop “Concentration of Measure Phenomena”
Simons Institute for the Theory of Computing
Berkeley 19-23 October 2020 (online 22 October 2020)

Notations. Sudakov's theorem

Setting. $X = (X_1, \dots, X_n)$ isotropic random vector in $\mathbb{R}^n$

$$S_\theta = \theta_1 X_1 + \dots + \theta_n X_n = \langle X, \theta \rangle, \qquad \theta = (\theta_1, \dots, \theta_n) \in \mathbb{S}^{n-1}.$$

Isotropy: $\mathbb{E} S_\theta^2 = 1$ for all $\theta \in \mathbb{S}^{n-1}$. Equivalently: $\mathbb{E} X_i X_j = \delta_{ij}$.

Problem 1: Is it true that most of S_θ are nearly normal $N(0,1)$? (in the sense of the normalized Lebesgue measure $\mathfrak{s}_{n-1}$ on $\mathbb{S}^{n-1}$)
Problem 2: Is it true that most of S_θ are nearly equidistributed?

Theorem (Sudakov 1978). Yes for Problem 2, if n is large.
Also yes for Problem 1, if (and only if)

$$\frac{|X|^2}{n} = \frac{X_1^2 + \dots + X_n^2}{n} \approx 1.$$

Proof: Use of the spherical isoperimetric inequality (sufficient: the spherical concentration phenomenon).

Independent summands: Standard rate

For $x \in \mathbb{R}$, put

$$F_\theta(x) = \mathbb{P}\{S_\theta \le x\}, \qquad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy,$$

$$\rho(F_\theta, \Phi) = \sup_x |F_\theta(x) - \Phi(x)|.$$

Theorem (Berry-Esseen). If $X_1, \dots, X_n$ are independent, $\mathbb{E}X_k = 0$, $\mathbb{E}X_k^2 = 1$, $\mathbb{E}\,|X_k|^3 \le \beta_3$, then, for all $\theta \in \mathbb{S}^{n-1}$,

$$\rho(F_\theta, \Phi) \le c\beta_3 \sum_{k=1}^{n} |\theta_k|^3.$$

Note: $\sum_{k=1}^{n} |\theta_k|^3 \ge \frac{1}{\sqrt{n}}$ for all $\theta \in \mathbb{S}^{n-1}$. On the other hand, with high $\mathfrak{s}_{n-1}$-probability

$$\sum_{k=1}^{n} |\theta_k|^3 < \frac{2}{\sqrt{n}}.$$

Equal coefficients: If $\theta_k = \frac{1}{\sqrt{n}}$, then $\rho(F_\theta, \Phi) \le \frac{c\beta_3}{\sqrt{n}}$.

Bernoulli case: If $\mathbb{P}\{X_k = \pm 1\} = \frac{1}{2}$, then $\rho(F_\theta, \Phi) \sim \frac{1}{\sqrt{n}}$.

Independent summands: Improved rates

$X_1, \dots, X_n$ independent, $\mathbb{E} X_k = 0$, $\mathrm{Var}(X_k) = 1$.

Theorem (B.Klartag-S.Sodin 2011).

$$c\,\mathbb{E}_\theta\,\rho(F_\theta, \Phi) \le \frac{1}{n}\,\bar\beta_4, \qquad \bar\beta_4 = \frac{1}{n}\sum_{k=1}^n \mathbb{E} X_k^4.$$

Moreover, for any $r > 0$,

$$\mathfrak{s}_{n-1}\Big\{c\,\rho(F_\theta, \Phi) \ge \frac{1}{n}\,\bar\beta_4\, r\Big\} \,\le\, 2\, e^{-\sqrt{r}}.$$

Theorem (B. 2020). In the i.i.d. case with $\mathbb{E} X_1^3 = 0$,

$$c\,\mathbb{E}_\theta\,\rho(F_\theta, \Phi) \le \frac{1}{n^{3/2}}\,\beta_5, \qquad \beta_5 = \mathbb{E}\,|X_1|^5.$$

Moreover, for any $r > 0$,

$$\mathfrak{s}_{n-1}\Big\{c\,\rho(F_\theta, \Phi) \ge \frac{1}{n^{3/2}}\,\beta_5\, r\Big\} \,\le\, 2\, e^{-\sqrt{r}}.$$

Typical distributions

$X = (X_1, \dots, X_n)$ in $\mathbb{R}^n$, $\quad F_\theta = \mathcal{L}(\langle X, \theta \rangle)$, $\theta \in \mathbb{S}^{n-1}$

$$F = \mathbb{E}_\theta F_\theta = \mathcal{L}(|X|\theta_1) \qquad \text{(typical distribution)}$$

Sudakov's theorem: Most of F_θ are close to F (if n is large).
Note: $\sqrt{n}\theta_1 \approx N(0,1)$.

Theorem (B-C-G 2017).

$$\int_{-\infty}^{\infty} (1 + x^2)\, |F - \Phi|(dx) \le \frac{c}{n} \Big(1 + \mathrm{Var}(|X|)\Big).$$

In particular,

$$\rho(F, \Phi) \le \frac{c}{n} \Big(1 + \mathrm{Var}(|X|)\Big).$$

Simplification using $\mathrm{Var}(|X|) \le \frac{\mathrm{Var}(|X|^2)}{\mathbb{E}\,|X|^2}$. If $\mathbb{E}\,|X|^2 = n$,

$$\rho(F, \Phi) \le \frac{c}{n}\,(1 + \sigma_4^2), \qquad \sigma_4^2 = \frac{1}{n}\,\mathrm{Var}(|X|^2).$$

Standard rate

X in $\mathbb{R}^n$, $\quad F_\theta = \mathcal{L}(\langle X, \theta \rangle)$, $\quad F = \mathbb{E}_\theta F_\theta$.

Theorem (B-C-G 2018). If $\mathbb{E}\,|X|^2 = n$ (no isotropy), then

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \,\le\, c\,\big(M_3^3 + \sigma_4^{3/2}\big)\frac{1}{\sqrt{n}}$$

where

$$M_3^3 \;=\; \sup_{\theta \in \mathbb{S}^{n-1}} \mathbb{E}\,|S_\theta|^3.$$

Second order correlation condition

We say that a random vector $X = (X_1, \dots, X_n)$ in $\mathbb{R}^n$ satisfies a SOC with parameter Λ, if for all $a_{ij} \in \mathbb{R}$,

$$\mathrm{Var}\Big(\sum_{i,j=1}^{n} a_{ij} X_i X_j\Big) \le \Lambda \sum_{i,j=1}^{n} a_{ij}^2.$$

Optimal value $\Lambda = \Lambda(X)$ is the maximal eigenvalue of the covariance matrix for the random vector $(X_i X_j - \mathbb{E} X_i X_j)_{i,j=1}^n$ of dimension n^2.

Theorem (B-C-G 2019). If X is isotropic in $\mathbb{R}^n$ $(n \ge 2)$, and its distribution is symmetric around the origin, then

$$c\,\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \le \frac{\log n}{n}\,\Lambda.$$

Examples

1) $\Lambda \le 2\,\max_k \mathbb{E} X_k^4$, if X_k are independent.
2) $\Lambda \le 4/\lambda_1$, if X is isotropic and satisfies a Poincaré-type inequality

$$\lambda_1\, \mathrm{Var}(u(X)) \le \mathbb{E}\,|\nabla u(X)|^2$$

for all smooth u on $\mathbb{R}^n$.

Note: $\sigma_4^2 \le \Lambda$ (choose $a_{ij} = \delta_{ij}$).

Non-symmetric case. Large deviations

Theorem (B-C-G 2020). Let X be isotropic in $\mathbb{R}^n$, $n \geq 2$, with mean zero and a positive Poincaré constant λ_1. Then

$$c\,\mathbb{E}_\theta\,\rho(F_\theta, \Phi) \;\leq\; \frac{\log n}{n}\,\lambda_1^{-1}.$$

Moreover, for all $r > 0$,

$$\mathfrak{s}_{n-1}\Big\{c\,\rho(F_\theta, \Phi) \geq \frac{\log n}{n}\,\lambda_1^{-1}\,r\Big\} \;\leq\; 2\,e^{-\sqrt{r}}.$$

Note: Recall that, for the typical distribution $F = \mathbb{E}_\theta F_\theta$,

$$\rho(F, \Phi) \leq \frac{c}{n}\,(1 + \sigma_4^2) \leq \frac{c'}{n}\,\Lambda \leq \frac{c''}{n}\,\lambda^{-1},$$

where $\sigma_4^2 = \frac{1}{n}\,\mathrm{Var}(|X|^2)$. Hence

$$\rho(F_\theta, \Phi) = \rho(F_\theta, F) + O\Big(\frac{1}{n}\,\lambda^{-1}\Big).$$

Relationship with thin shell/KLS

The next assertions are equivalent up to constants c, β (different in different places) for the entire class of isotropic random vectors X in $\mathbb{R}^n$ with symmetric log-concave distributions:

(i) $\sup_X \lambda_1^{-1}(X) \le c\,(\log n)^\beta$

(ii) $\sup_X \mathrm{Var}(|X|) \le c\,(\log n)^\beta$

(ii') $\sup_X \mathrm{Var}(|X|^2) \le cn\,(\log n)^\beta$

(iii) $\sup_X \mathbb{E}_\theta\, \rho(F_\theta, \Phi) \le \frac{c}{n}\,(\log n)^\beta$.

$(\mathrm{i}) \Rightarrow (\mathrm{ii})$: Apply Poincaré-type inequality to $u(x) = |x|$.
$(\mathrm{i}) \Rightarrow (\mathrm{ii}')$: Apply Poincaré-type inequality to $u(x) = |x|^2$.
$(\mathrm{ii}) \Rightarrow (\mathrm{i})$: R. Eldan (2013), stochastic localization.
$(\mathrm{iii}) \Rightarrow (\mathrm{ii})$: In view of a general relation

$$c\,\mathrm{Var}(|X|) \ \le \ n\,(\log n)^4\, \mathbb{E}_\theta\, \rho(F_\theta, \Phi) + 1.$$

H.Jiang, Y.T.Lee and S.S.Vempala (2020): Formulation of (i) as a CLT for $\langle X, Y\rangle$ where Y is an independent copy of X.

Berry-Esseen-type bounds

Lemma 1. Given distribution functions U and V with characteristic functions $\hat{U}$ and $\hat{V}$, for all $T > 0$,

$$c\,\rho(U,V) \;\le\; \int_0^T \frac{|\hat{U}(t) - \hat{V}(t)|}{t}\,dt + \frac{1}{T}\int_0^T |\hat{V}(t)|\,dt.$$

Lemma 2. If X is a random vector in $\mathbb{R}^n$ such that $\mathbb{E}\,|X|^2 = n$, then, for all $T \ge T_0 \ge 1$ and $\theta \in \mathbb{S}^{n-1}$,

$$\begin{aligned} c\,\rho(F_\theta, \Phi) \;\le\; & \int_0^{T_0} \frac{|f_\theta(t) - f(t)|}{t}\,dt \\ & + \int_{T_0}^{T} \frac{|f_\theta(t)|}{t}\,dt + \frac{\Lambda}{n}\Big(1 + \log\frac{T}{T_0}\Big) + \frac{1}{T} + e^{-T_0^2/4}. \end{aligned}$$

Good choice: $T \sim n$, $T_0 \sim \sqrt{\log n}$.

Deviations on $\mathbb{S}^{n-1}$ at standard rate

Let u be a smooth function on $\mathbb{S}^{n-1}$.
Poincaré inequality:

$$\mathrm{Var}_\theta(u) \le \frac{1}{n-1} \int |\nabla u|^2 \, d\mathfrak{s}_{n-1}.$$

Logarithmic Sobolev inequality (C.E. Mueller-F.B. Weissler 1982):

$$\mathrm{Ent}_\theta(u^2) \le \frac{2}{n-1} \int |\nabla u|^2 \, d\mathfrak{s}_{n-1},$$

where

$$\mathrm{Ent}(\xi) = \mathbb{E}\xi \log \xi - \mathbb{E}\xi \log \mathbb{E}\xi \qquad (\xi \ge 0).$$

Concentration of measure (V.D. Milman 1970s): If $\|u\|_{\mathrm{Lip}} \le 1$, then, for all $r \ge 0$,

$$\mathfrak{s}_{n-1}\Big\{|u - \mathbb{E}_\theta u| \ge r\Big\} \ \le \ e^{-r^2/2(n-1)}.$$

Informally

$$|u - \mathbb{E}_\theta u| \ \prec \ c\,\frac{|Z|}{\sqrt{n}}, \qquad Z \sim N(0,1).$$

Deviations on $\mathbb{S}^{n-1}$ at improved rate

Let u be defined and C^2-smooth in some neighborhood of $\mathbb{S}^{n-1}$.

Lemma 3 (B-C-G 2017). If u is orthogonal to all linear functions in $L^2(\mathfrak{s}_{n-1})$, then, for any $a \in \mathbb{R}$,

$$\mathrm{Var}_\theta(u) \le \frac{5}{(n-1)^2} \int \|\nabla^2 u - aI_n\|_{\mathrm{HS}}^2 \, d\mathfrak{s}_{n-1}.$$

Moreover, if $\|\nabla^2 u - aI_n\| \le 1$ (operator norm) on $\mathbb{S}^{n-1}$ and this integral is bounded by b, then

$$\int \exp\Big\{\frac{n-1}{2(1+4b)}\,|u|\Big\}\, d\mathfrak{s}_{n-1} \;\le\; 2.$$

Informally

$$|u - \mathbb{E}_\theta u| \;\prec\; c_b \Big(\frac{|Z|}{\sqrt{n}}\Big)^2, \qquad Z \sim N(0,1).$$

Even functions: No linear component, if $u(-\theta) = u(\theta)$.
Example: $u(\theta) = f_\theta(t) = \mathbb{E}\, e^{it\langle X,\theta\rangle}$, assuming that X has a symmetric distribution on $\mathbb{R}^n$.

Deviations on $\mathbb{S}^{n-1}$ at improved rate (cont.)

Generalization: Given a function u in the (complex) Hilbert space $L^2 = L^2(\mathbb{R}^n, \mathfrak{s}_{n-1})$, consider its orthogonal projection

$$l = \mathrm{Proj}_H u$$

onto the linear space H in L^2 generated by linear functions on $\mathbb{R}^n$ (l is a linear part of u).

Lemma 4. Let u be C^2-smooth in some neighborhood of $\mathbb{S}^{n-1}$ and have $\mathfrak{s}_{n-1}$-mean zero. For any $a \in \mathbb{C}$,

$$\|u\|_{L^2}^2 \le \frac{c}{n^2} \int \|\nabla^2 u - aI_n\|_{\mathrm{HS}}^2 \, d\mathfrak{s}_{n-1} + \|l\|_{L^2}^2.$$

Moreover, if $\|\nabla^2 u - aI_n\| \le 1$ on $\mathbb{S}^{n-1}$, then

$$\|u\|_{\psi_1} \le \frac{c}{n} + \frac{c}{n} \int \|\nabla^2 u - aI_n\|_{\mathrm{HS}}^2 \, d\mathfrak{s}_{n-1} + 6\,\|l\|_{L^2}.$$

Orlicz ψ_1-norm:

$$\|u\|_{\psi_1} = \inf\left\{\lambda > 0 : \mathbb{E}_\theta\, e^{|u|/\lambda} \le 2\right\}.$$

Note: $u - l$ has zero linear part and the same Hessian as u.

Concentration of characteristic functions (standard rate)

Lemma 5. Given an isotropic random vector X in $\mathbb{R}^n$, for all $t \in \mathbb{R}$,

$$\mathbb{E}_\theta \, |f_\theta(t) - f(t)|^2 \le \frac{t^2}{n-1}.$$

Moreover,

$$\| f_\theta(t) - f(t) \|_{\psi_2} \le \frac{c|t|}{\sqrt{n}}.$$

Proof. Define the smooth functions

$$u_t(\theta) = f_\theta(t) = \mathbb{E}\, e^{it\langle X,\theta\rangle}, \quad \theta \in \mathbb{R}^n.$$

Gradients:

$$\langle \nabla u_t(\theta), w\rangle = it\, \mathbb{E}\, \langle X, w\rangle\, e^{it\langle X,\theta\rangle}, \quad w \in \mathbb{C}^n.$$

By the isotropy, writing $w = w_0 + iw_1$, $w_0, w_1 \in \mathbb{R}^n$, we have

$$\begin{aligned} |\, \langle \nabla u_t(\theta), w\rangle \,|^2 &\le \mathbb{E}\, |\, \langle X, w\rangle \,|^2 \\ &= \mathbb{E}\, \langle X, w_0\rangle^2 + \mathbb{E}\, \langle X, w_1\rangle^2 = |w_0|^2 + |w_1|^2 = |w|^2. \end{aligned}$$

This gives a uniform bound $|\nabla u_t(\theta)| \le |t|$. Then apply the spherical Poincaré inequality and the Gaussian concentration on the sphere.

Concentration of characteristic functions (improved rate)

Lemma 6. Given an isotropic random vector X in $\mathbb{R}^n$, in the interval $|t| \le An^{1/5}$,

$$c\,\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2 \;\le\; \|l_t\|_{L^2}^2 + \frac{\Lambda t^4}{n^2}$$

where $f(t) = \mathbb{E}_\theta f_\theta(t)$ and l_t is a linear part of $f_\theta(t)$ in $L^2(\mathfrak{s}_{n-1})$ with constant $c = c(A) > 0$. Moreover, for $|t| \le An^{1/6}$,

$$c\,\|f_\theta(t) - f(t)\|_{\psi_1} \;\le\; \|l_t\|_{L^2} + \frac{\Lambda t^2}{n}.$$

Proof (with a worse upper bound in t and assuming that $l_t = 0$). In the isotropic case, the SOC is described as

$$\mathbb{E}\left|\sum_{j,k=1}^n z_{jk}\,(X_j X_k - \delta_{jk})\right|^2 \;\le\; \Lambda$$

with arbitrary $z_{jk} \in \mathbb{C}$ such that $\sum_{j,k=1}^n |z_{jk}|^2 = 1$.
To employ an improved spherical concentration (Lemma 3), we need to estimate the operator norm $\|\nabla^2 u_t - aI_n\|$ and the Hilbert-Schmidt norm $\|\nabla^2 u_t - aI_n\|_{\mathrm{HS}}$. For any fixed $t \in \mathbb{R}$,

$$\left[\nabla^2 u_t(\theta)\right]_{jk} = \frac{\partial^2}{\partial\theta_j \partial\theta_k}\, f_\theta(t) = -t^2\,\mathbb{E}\,X_j X_k\, e^{it\langle X,\theta\rangle}.$$

That is (in matrix form), for any $w \in \mathbb{C}^n$,

$$\big\langle \nabla^2 u_t(\theta)\, w, w \big\rangle = -t^2\, \mathbb{E}\, |\, \langle X, w\rangle\, |^2\, e^{it\langle X,\theta\rangle}.$$

Choosing $a = -t^2 f(t)$, by the isotropy assumption, if $|w| = 1$,

$$\big|\, \big\langle (\nabla^2 u_t(\theta) - aI_n)\, w, w \big\rangle \,\big| \ \le\ t^2\, \mathbb{E}\, |\, \langle X, w\rangle\, |^2 + |a|\, |w|^2 \ \le\ 2t^2.$$

Hence

$$\|\nabla^2 u_t(\theta) - aI_n\| \le 2t^2.$$

In addition, putting $a(\theta) = -t^2 f_\theta(t)$, we have

$$\begin{aligned}
\big\|\nabla^2 u_t(\theta) - a(\theta) I_n\big\|_{\mathrm{HS}}^2 \ &=\ \sum_{j,k=1}^n \big|\nabla^2 u_t(\theta)_{jk} - a(\theta)\, \delta_{jk}\big|^2 \\
&=\ \sup \bigg| \sum_{j,k=1}^n z_{jk} \left(\nabla^2 u_t(\theta)_{jk} - a(\theta)\, \delta_{jk}\right)\bigg|^2 \\
&\le\ t^4 \sup \mathbb{E} \bigg| \sum_{j,k=1}^n z_{jk} \left(X_j X_k - \delta_{jk}\right)\bigg|^2,
\end{aligned}$$

as long as $\sum_{j,k=1}^n |z_{jk}|^2 = 1$. Thus, for all θ,

$$\big\|\nabla^2 u_t(\theta) - a(\theta) I_n\big\|_{\mathrm{HS}}^2 \ \le\ \Lambda t^4.$$

On the other hand, by Lemma 5,

$$\mathbb{E}_\theta \left\|(a(\theta) - a)I_n\right\|_{\mathrm{HS}}^2 \ = \ nt^4\, \mathbb{E}_\theta \,|f_\theta(t) - f(t)|^2 \ \le \ 2t^6.$$

The two bounds give

$$\mathbb{E}_\theta \left\|\nabla^2 u_t(\theta) - aI_n\right\|_{\mathrm{HS}}^2 \le \ 2\Lambda t^4 + 4t^6.$$

Finally, apply Lemma 3 with

$$u(\theta) = \frac{1}{2t^2}\, u_t(\theta) = \frac{1}{2t^2}\,(f_\theta(t) - f(t))$$

for which

$$\|\nabla^2 u(\theta) - aI_n\| \le 1$$

and

$$b \ \equiv \ \mathbb{E}_\theta \left\|\nabla^2 u_t(\theta) - aI_n\right\|_{\mathrm{HS}}^2 \ \le \ \frac{1}{2}\,\Lambda + t^2.$$

Second assertion of Lemma 3:

$$\mathbb{E}_\theta \exp\left\{\frac{n-1}{2(1+4b)}\,|u(\theta)|\right\} \ \le \ 2,$$

or equivalently

$$\|u\|_{\psi_1} \ \le \ \frac{2(1+4b)}{n-1} \ \le \ \frac{4}{n} + \frac{16}{n}\left(\frac{1}{2}\,\Lambda + t^2\right).$$

This yields

$$\|f_\theta(t) - f(t)\|_{\psi_1} \ \le \ \frac{ct^2}{n}\,(\Lambda + t^2).$$

L^2-norm of linear parts

Let u be a C^2-smooth function on S^{n-1} with mean zero, and

$$l = \mathrm{Proj}_H(u)$$

be its projection onto the space H of all linear functions in $L^2(\mathfrak{s}_{n-1})$. One may choose for the orthonormal basis

$$l_k(\theta) = \sqrt{n}\,\theta_k, \quad k = 1, \dots, n, \quad \theta = (\theta_1, \dots, \theta_n) \in \mathbb{S}^{n-1}.$$

Hence

$$l(\theta) \,=\, \sum_{k=1}^{n} \langle u, l_k \rangle_{L^2} \, l_k(\theta) \,=\, \langle v, \theta \rangle$$

with

$$v = n \int \theta u(\theta)\, d\mathfrak{s}_{n-1}(\theta)$$

which implies

$$\|l\|_{L^2}^2 \,=\, \frac{1}{n}\,|v|^2 \,=\, n\; \mathbb{E}_{\theta}\, \mathbb{E}_{\theta'} \,\langle \theta, \theta' \rangle\, u(\theta) \bar{u}(\theta').$$

Characteristic function of linear functions

Linear functionals $l(\theta) = \langle \theta, v \rangle$ with $|v| = 1$ viewed as random variables on $(\mathbb{S}^{n-1}, \mathfrak{s}_{n-1})$ have equal distributions with density

$$c_n \left(1 - x^2\right)_+^{\frac{n-3}{2}}, \quad x \in \mathbb{R}, \qquad c_n = \frac{\Gamma(\frac{n}{2})}{\sqrt{\pi}\,\Gamma(\frac{n-1}{2})}.$$

Characteristic function (multiple of Bessel function of the first kind with index $\nu = \frac{n}{2} - 1$):

$$J_n(t) \,=\, \mathbb{E}_\theta\, e^{it\theta_1} \,=\, c_n \int_{-1}^{1} e^{itx} \, (1 - x^2)^{\frac{n-3}{2}} \, dx.$$

Lemma 7. For all $t \in \mathbb{R}$,

$$\Big|J_n(t\sqrt{n}) - \Big(1 - \frac{t^4}{4n}\Big)\, e^{-t^2/2}\Big| \,\le\, \frac{c}{n^2}\, \min\{1, t^4\},$$

$$\Big|\big(J_n(t\sqrt{n})\big)' + t\,\Big(1 + \frac{4t^2 - t^4}{4n}\Big)\, e^{-t^2/2}\Big| \,\le\, \frac{c}{n^2}\, \min\{1, |t|^3\}.$$

Linear part of characteristic functions

Let X be an isotropic random vector in $\mathbb{R}^n$, and let $l_t(\theta)$ be the linear part of the characteristic function

$$f_\theta(t) = \mathbb{E}\, e^{it\langle X,\theta\rangle}$$

as a function of θ on the sphere. Squared $L^2(\mathfrak{s}_{n-1})$-norm:

$$I(t) \;=\; \|l_t\|_{L^2}^2 \;=\; n\,\mathbb{E}_\theta\,\mathbb{E}_{\theta'}\,\langle \theta, \theta' \rangle\, f_\theta(t)\bar{f}_{\theta'}(t).$$

Let Y be an independent copy of X. Lemma 7 implies:

Lemma 8. For any $t \in \mathbb{R}$,

$$I(t) = \frac{t^2}{n}\,\mathbb{E}\,\langle X, Y\rangle\,\Big(1 - \frac{(U^2+V^2)\,t^4 - 8R^2t^2}{4n}\Big)\,e^{-R^2t^2} + O\Big(\frac{t^2}{n^{5/2}}\Big),$$

where

$$R^2 = \frac{|X|^2+|Y|^2}{2n}, \quad U = \frac{|X|^2}{n}, \quad V = \frac{|Y|^2}{n}.$$

Putting $T_0 = 4\sqrt{\log n}$, we have

$$\int_0^{T_0} \frac{I(t)}{t^2}\,dt \;\le\; \frac{c}{n}\,\mathbb{E}\,\frac{\langle X, Y\rangle}{R} + O\Big(\frac{\Lambda^2}{n^2}\Big).$$

Back to (the part of) Berry-Esseen

Lemma 9. If X has mean zero and Poincaré constant $\lambda_1 > 0$, then

$$\mathbb{E}\,\frac{\langle X, Y\rangle}{R} \le \frac{c}{\lambda_1^2 n}.$$

Now, by Lemma 6, for $|t| \le An^{1/6}$,

$$c\,\|f_\theta(t) - f(t)\|_{\psi_1} \le \sqrt{I(t)} + \frac{\Lambda t^2}{n},$$

so that

$$\begin{aligned} c\left\|\int_0^{T_0} |f_\theta(t) - f(t)|\,\frac{dt}{t}\right\|_{\psi_1} &\le c\int_0^{T_0} \|f_\theta(t) - f(t)\|_{\psi_1}\,\frac{dt}{t} \\ &\le \frac{\Lambda T_0^2}{2n} + \int_0^{T_0} \frac{\sqrt{I(t)}}{t}\,dt. \end{aligned}$$

Here, by Lemmas 8-9,

$$\int_0^{T_0} \frac{I(t)}{t^2}\,dt \;=\; O\Big(\frac{1}{\lambda_1^2 n^2}\Big).$$

It follows that

$$\begin{aligned}\int_0^{T_0} \frac{\sqrt{I(t)}}{t}\,dt \;&\le\; \sqrt{T_0}\left(\int_0^{T_0} \frac{I(t)}{t^2}\,dt\right)^{1/2} \\ &\le\; \frac{c'}{\lambda_1 n}\,(\log n)^{1/4}.\end{aligned}$$

Recall that $\Lambda \le \frac{4}{\lambda_1}$.