Concentration of Distributions of the Weighted Sums with Bernoullian Coefficients*

S.G. Bobkov

School of Mathematics, University of Minnesota, 127 Vincent Hall, 206 Church St. S.E., Minneapolis, MN 55455, USA *bobkov@math.umn.edu*

Summary. For non-correlated random variables, we study a concentration property of the distributions of the weighted sums with Bernoullian coefficients. The obtained result is used to derive an "almost surely version" of the central limit theorem.

Let $X = (X_1, \dots, X_n)$ be a vector of n random variables with finite second moments such that, for all k, j,

$$\mathbf{E}\, X_k X_j = \delta_{kj} \tag{1}$$

where δ_{kj} is Kronecker's symbol. It is known that, for growing n, the distribution functions

$$F_\theta(x) = \mathbf{P}\left\{ \sum_{k=1}^{n} \theta_k X_k \le x \right\}, \quad x \in \mathbf{R},$$

of the weighted sums of (X_k), with coefficients $\theta = (\theta_1, \dots, \theta_n)$ satisfying $\theta_1^2 + \ldots + \theta_n^2 = 1$, form a family possessing a certain concentration property with respect to the uniform measure σ_{n-1} on the unit sphere S^{n-1}. Namely, most of F_θ's are close to the average distribution

$$F(x) = \int_{S^{n-1}} F_\theta(x)\, d\sigma_{n-1}(\theta)$$

in the sense that, for each $\delta > 0$, there is an integer n_δ such that if $n \ge n_\delta$ one can select a set of coefficients $\Theta \subset S^{n-1}$ of measure $\sigma_{n-1}(\Theta) \ge 1 - \delta$ such that $d(F_\theta, F) \le \delta$, for all $\theta \in \Theta$. This property was first observed by V.N. Sudakov [S] who stated it for the Kantorovich–Rubinshtein distance $d(F_\theta, F) = \int_{-\infty}^{+\infty} |F_\theta(x) - F(x)|\, dx$, with a proof essentially relying on the isoperimetric theorem on the sphere. A different approach to this result was suggested by H. von Weizsäcker [W] (cf. also [D-F]). V.N. Sudakov also considered "Gaussian coefficients" in which case, as shown in [W], there is a rather general infinite dimensional formulation. An important special situation where the random vector X is uniformly distributed over a centrally symmetric convex body in $\mathbf{R}^n$ was recently studied, for the uniform distance

* Supported by an NSF grant and EPSRC Visiting Fellowship.

V.D. Milman and G. Schechtman (Eds.): LNM 1807, pp. 27–36, 2003.

$\sup_x |F_\theta(x) - F(x)|$, by M. Antilla, K. Ball, and I. Perissinaki [A-B-P], see also [B] for refinements and extensions to log-concave distributions. One can find there quantitative versions of Sudakov's theorem, while in the general case, the following statement proven in [B] holds true: under (1), for all $\delta > 0$,

$$\sigma_{n-1}\big\{L(F_\theta, F) \geq \delta\big\} \leq 4n^{3/8}\, e^{-n\delta^4/8}. \tag{2}$$

Here $L(F_\theta, F)$ stands for the Lévy distance defined as the minimum over all $\delta \geq 0$ such that $F(x-\delta) - \delta \leq F_\theta(x) \leq F(x+\delta) + \delta$, for all $x \in \mathbf{R}$. As well as the Kantorovich–Rubinshtein distance d, the metric L is responsible for the weak convergence, and there is a simple relation $d(F_\theta, F) \leq 6L(F_\theta, F)^{1/2}$ (so one can give an appropriate estimate for d on the basis of (2)).

The aim of this note is to show that a property similar to (2) still holds with respect to very small pieces of the sphere. As a basic example, we consider coefficients of the special form $\theta = \frac{1}{\sqrt{n}}\,\varepsilon$ where $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ is an arbitrary sequence of signs ± 1. Thus, consider the weighted sums

$$S_\varepsilon = \frac{1}{\sqrt{n}} \sum_{k=1}^n \varepsilon_k X_k$$

together with their distribution functions $F_\varepsilon(x) = \mathbf{P}\{S_\varepsilon \leq x\}$ and the corresponding average distribution

$$F(x) = \int_{\{-1,1\}^n} F_\varepsilon(x)\, d\mu_n(\varepsilon) = \frac{1}{2^n} \sum_{\varepsilon_k = \pm 1} \mathbf{P}\left\{\frac{\varepsilon_1 X_1 + \ldots + \varepsilon_n X_n}{\sqrt{n}} \leq x\right\}. \tag{3}$$

Here and throughout, μ_n stands for the normalized counting measure on the discrete cube $\{-1,1\}^n$. We prove:

Theorem 1. *Under* (1)*, for all* $\delta > 0$*,*

$$\mu_n\big\{\varepsilon : L(F_\varepsilon, F) \geq \delta\big\} \leq Cn^{1/4}\, e^{-cn\delta^8}, \tag{4}$$

where C *and* c *are certain positive numerical constants.*

Note that the condition (1) is invariant under rotations, i.e., it is fulfilled for random vectors $U(X)$ with an arbitrary linear unitary operator U in $\mathbf{R}^n$. Being applied to such vectors, the inequality (4) will involve the average $F = F^U$ which of course depends on U. However, under mild integrability assumptions on the distribution of X, all these F^U (not just most of them) turn out to be close to the one appearing in Sudakov's theorem as the typical distribution for the uniformly distributed (on the sphere) or suitably squeezed Gaussian coefficients. In particular, one can give an analogue of (4) with a certain distribution F not depending on the choice of the basis in $\mathbf{R}^n$. On the other hand, some additional natural assumptions lead to the following version of the central limit theorem. We will denote by μ_∞ the canonical infinite product measure $\mu_1 \otimes \mu_1 \otimes \ldots$ on the product space $\{-1,1\}^\infty$.

Theorem 2. *Let* $\{X_{n,k}\}_{k=1}^n$ *be an array of random variables satisfying* (1) *for all* n *and such that in probability, as* $n \to \infty$,

a) $\frac{\max\{|X_{n,1}|,\dots,|X_{n,n}|\}}{\sqrt{n}} \to 0$,

b) $\frac{X_{1,1}^2+\dots+X_{n,n}^2}{n} \to 1$.

Then, for μ_∞*-almost all sequences* $\{\varepsilon_k\}_{k\geq 1}$ *of signs,*

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n \varepsilon_k X_{n,k} \to N(0,1), \quad \textit{as} \quad n \to \infty.$$

If we consider the sum $\frac{1}{\sqrt{n}} \sum_{k=1}^n \varepsilon_k X_{n,k}$ with ε_k regarded as independent Bernoullian random variables which are independent of all $X_{n,k}$, then the above statement will become much weaker and will express just the property that the average distribution F defined by (3) for the random vector $(X_{1,1},\dots,X_{n,n})$ is close to $N(0,1)$ (here is actually a step referring to the assumptions a) and b)). In addition to this property, we need to have a sufficiently good closeness (in spaces of finite dimension) of most of F_ε's to F and thus to the normal law.

Both the assumption a) and b) are important for the conclusion of Theorem 2. Under a), the property b) is necessary. To see that a) cannot be omitted, assume that the underlying probability space $(\Omega, \mathbf{P})$ is non-atomic and take a partition $A_{n,1},\dots,A_{n,n}$ of Ω consisting of the sets of $\mathbf{P}$-measure $1/n$. Then, the array $X_{n,k} = \sqrt{n}\, 1_{A_{n,k}}$, $1 \leq k \leq n$, satisfies (1), and

$$\frac{\max\{|X_{n,1}|,\dots,|X_{n,n}|\}}{\sqrt{n}} = 1, \qquad \frac{X_{1,1}^2+\dots+X_{n,n}^2}{n} = 1,$$

so, the property b) is fulfilled, while a) is not. On the other hand, for any sign sequence $(\varepsilon_1,\dots,\varepsilon_n)$, the random variable $\frac{1}{\sqrt{n}} \sum_{k=1}^n \varepsilon_k X_{n,k}$ takes only the two values ± 1, so it cannot be approximated by the standard normal distribution. Note, however, that Theorem 1 still holds in this degenerate case, with the μ_n-typical distribution F having two equal atoms at ± 1.

It might be worthwhile also noting that in general it is not possible to state Theorem 2 for any prescribed coefficients, say, for $\varepsilon_k = 1$ – similarly to the case of independent variables, even if, for each n, $\{X_{n,k}\}$ are bounded, symmetrically distributed and pairwise independent. For example, start from a sequence of independent Bernoullian random variables $\xi_1,\dots,\xi_d$ (with $\mathbf{P}\{\xi_k = \pm 1\} = \frac{1}{2}$) and construct a double index sequence $X_{n,(k,j)} = \xi_k \xi_j$, $1 \leq k < j \leq d$. The collection $\{X_{n,(k,j)}\}$, of cardinality $n = d(d-1)/2$, satisfies the basic correlation condition (1), and since $|X_{n,(k,j)}| = 1$, both the assumption a) and b) are fulfilled. Nevertheless, in probability, as $d \to \infty$,

$$\frac{1}{\sqrt{n}} \sum_{1\leq k<j\leq d} X_{n,(k,j)} = \frac{1}{2\sqrt{n}} \left(\sum_{k=1}^d \xi_k\right)^2 - \frac{d}{2\sqrt{n}} \longrightarrow \frac{\zeta^2-1}{\sqrt{2}}$$

where $\zeta \in N(0,1)$.

We turn to the proof of Theorem 1. To this task, we first study the concentration property of the family $\{F_\varepsilon\}$ on the level of their characteristic functions

$$f_\varepsilon(t) = \mathbf{E}\, e^{itS_\varepsilon}, \quad t \in \mathbf{R}.$$

Concentration of $\{f_\varepsilon\}$ around its μ_n-mean

$$f(t) = \int f_\varepsilon(t)\, d\mu_n(\varepsilon) = \int_{-\infty}^{+\infty} e^{itx}\, dF(x)$$

can be then converted, with the help of standard facts from Fourier analysis, into the concentration of distributions in the form (4). This route somewhat different than that of [A-B-P] or [B] has apparently to be chosen in view of a specific form of concentration on the discrete cube.

With every complex-valued function f on $\{-1,1\}^n$, we connect the length of the discrete gradient $|\nabla f|$ defined by

$$|\nabla f(\varepsilon)|^2 = \sum_{k=1}^n \left| \frac{f(\varepsilon) - f(s_k(\varepsilon))}{2} \right|^2, \quad \varepsilon \in \{-1,1\}^n,$$

where $s_k(\varepsilon)$ is the neighbour of ε along kth coordinate, i.e., $(s_k(\varepsilon))_j = \varepsilon_j$ for $j \neq k$, and $(s_k(\varepsilon))_k = -\varepsilon_k$. Set $\|\nabla f\|_\infty = \max_\varepsilon |\nabla f(\varepsilon)|$.

Lemma 1. *For every f such that $\|\nabla f\|_\infty \leq \sigma$,*

$$\mu_n\left\{ \left| f - \int f\, d\mu_n \right| \geq h \right\} \leq 4e^{-h^2/(4\sigma^2)}, \quad h > 0.$$

This Gaussian bound is standard. It can be obtained using the so-called modified logarithmic Sobolev inequalities, see e.g. [B-G], [L]. In fact, for real-valued f, a sharper estimate holds true,

$$\mu_n\left\{ \left| f - \int f\, d\mu_n \right| \geq h \right\} \leq 2e^{-h^2/(2\sigma^2)},$$

while in general the latter can be applied separately to the real and the imaginary part of f to yield the inequality of Lemma 1.

Lemma 2. *Under* (1)*, for every $t \in \mathbf{R}$,*

$$\|\nabla f_\varepsilon(t)\|_\infty \leq \frac{|t| + t^2}{\sqrt{n}}.$$

Proof. Using the equality $f_\varepsilon(t) - f_{s_k(\varepsilon)}(t) = \mathbf{E}\, e^{itS_\varepsilon}(1 - e^{-2it\,\varepsilon_k X_k/\sqrt{n}})$, we may write

$$|\nabla f_\varepsilon(t)| = \sup\left|\mathbf{E}\, e^{itS_\varepsilon} \sum_{k=1}^n a_k\, \frac{1-e^{-2it\,\varepsilon_k X_k/\sqrt{n}}}{2}\right|$$
$$\leq \sup\, \mathbf{E}\left|\sum_{k=1}^n a_k\, \frac{1-e^{-2it\,\varepsilon_k X_k/\sqrt{n}}}{2}\right|,$$

where the supremum runs over all complex numbers $a_1, \dots, a_n$ such that $|a_1|^2 + \ldots + |a_n|^2 = 1$. Using the estimate $|e^{i\alpha} - 1 - i\alpha| \leq \frac{1}{2}\,\alpha^2$ ($\alpha \in \mathbf{R}$) and the assumption $\mathbf{E}X_k^2 = 1$, we can continue to get

$$|\nabla f_\varepsilon(t)| \leq \frac{|t|}{\sqrt{n}}\, \sup\, \mathbf{E}\left|\sum_{k=1}^n a_k\varepsilon_k X_k\right| + \frac{t^2}{n}\, \sup\, \mathbf{E}\sum_{k=1}^n |a_k| X_k^2$$
$$= \frac{|t|}{\sqrt{n}}\, \sup\, \mathbf{E}\left|\sum_{k=1}^n a_k\varepsilon_k X_k\right| + \frac{t^2}{\sqrt{n}}.$$

It remains to note that, by Schwarz' inequality and (1), $(\mathbf{E}\,|\sum_{k=1}^n a_k\varepsilon_k X_k\,|)^2 \leq \mathbf{E}\,|\sum_{k=1}^n a_k\varepsilon_k X_k\,|^2 = 1$.

We also need the following observation due to H. Bohman [Bo].

Lemma 3. *Given characteristic functions φ_1 and φ_2 of the distribution functions F_1 and F_2, respectively, if $|\varphi_1(t) - \varphi_2(t)| \leq \lambda|t|$, for all $t \in \mathbf{R}$, then, for all $x \in \mathbf{R}$ and $a > 0$,*

$$F_1(x-a) - \frac{2\lambda}{a} \leq F_2(x) \leq F_1(x+a) + \frac{2\lambda}{a}.$$

The particular case $a = \sqrt{2\lambda}$ gives an important relation

$$\frac{1}{2}\, L(F_1,F_2)^2 \leq \sup_{t>0}\left|\frac{\varphi_1(t)-\varphi_2(t)}{t}\right|. \tag{5}$$

Proof of Theorem 1. Fix a number $h > 0$. For $0 < t \leq \frac{2}{h}$, by Lemma 2, $\|\nabla f_\varepsilon(t)\|_\infty \leq \frac{t+t^2}{\sqrt{n}} \leq \frac{t}{\sqrt{n}}\,(1+\frac{2}{h})$, so that, by Lemma 1 applied to the function $\varepsilon \to f_\varepsilon(t)$, we get

$$\mu_n\left\{\varepsilon : \left|\frac{f_\varepsilon(t)-f(t)}{t}\right| \geq h\right\} \leq 4e^{-nh^4/4(h+2)^2}. \tag{6}$$

In the case $t > \frac{2}{h}$, this inequality is immediate, since $|f_\varepsilon(t) - f(t)| \leq 2 < th$, for all ε. Thus, we have the estimate (6) for all t separately, but in order to apply Lemma 3, we need a similar bound holding true for the supremum over all $t > 0$. To this end, apply (6) to the points $t_r = rh^2$, $r = 1, 2, \dots, N = \left[\frac{2}{h}\right] + 1$, to get

$$\mu_n \left\{ \varepsilon : \max_{1 \leq r \leq N} \left| \frac{f_\varepsilon(t_r) - f(t_r)}{t_r} \right| \geq h \right\} \leq 4N e^{-nh^4/4(h+2)^2}. \qquad (7)$$

Since $\mathbf{E}S_\varepsilon = 0$, $\mathbf{E}S_\varepsilon^2 = 1$, we have $|f_\varepsilon'(t)| \leq 1$, $f_\varepsilon'(0) = 0$, $|f_\varepsilon''(t)| \leq 1$, and similarly for f. Therefore, $|f_\varepsilon(t) - f(t)| \leq t^2 \leq th$, for all ε, as soon as $0 \leq t \leq h$. In case $h \leq t \leq \frac{2}{h}$, since $t_N \geq \frac{2}{h}$, one can pick an index $r = 1, \dots, N-1$ such that $t_r < t \leq t_{r+1}$. Assuming that $|\frac{f_\varepsilon(t_r) - f(t_r)}{t_r}| < h$, and recalling that $t_{r+1} - t_r = h^2$, we may write

$$\begin{aligned} |f_\varepsilon(t) - f(t)| &\leq |f_\varepsilon(t) - f_\varepsilon(t_r)| + |f_\varepsilon(t_r) - f(t_r)| + |f(t_r) - f(t)| \\ &< 2|t - t_r| + t_r h \,\leq\, 2h^2 + t_r h \,<\, 3th. \end{aligned}$$

Consequently, (7) implies

$$\begin{aligned} \mu_n \left\{ \sup_{t>0} \left| \frac{f_\varepsilon(t) - f(t)}{t} \right| \geq 3h \right\} &\leq 4N e^{-nh^4/4(h+2)^2} \\ &\leq 4\left(\frac{2}{h} + 1\right) e^{-nh^4/4(h+2)^2}. \end{aligned}$$

Therefore, by (5),

$$\mu_n \left\{ \frac{1}{2}\, L(F_\varepsilon, F)^2 \geq 3h \right\} \leq 4 \left(\frac{2}{h} + 1 \right) e^{-nh^4/4(h+2)^2}.$$

Replacing $6h$ with δ^2 and noticing that only $0 < \delta \leq 1$ should be taken into consideration, one easily arrives at the estimate $\mu_n\{L(F_\varepsilon, F) \geq \delta\} \leq \frac{C}{\delta^2}\, e^{-cn\delta^8}$ with some positive numerical constants C and c. On the other hand, in the latter inequality, we may restrict ourselves to values $\delta > c_1 n^{-1/8}$ which make the bound $\frac{C}{\delta^2}\, e^{-cn\delta^8}$ less than 1, and then we arrive at the desired inequality (4).

Theorem 1 has been proved, and we may state its immediate consequence:

Corollary 1. *Under* (1), *for at least* 2^{n-1} *sequences* $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ *of signs,* $L(F_\varepsilon, F) \leq C(\frac{\log n}{n})^{1/8}$*, where* C *is a universal constant.*

Let us now turn to the second task: approximation of the μ_n-typical F by more canonical distributions. Namely, denote by G the distribution function of the random variable $\zeta\, \frac{|X|}{\sqrt{n}}$ where ζ is a standard normal random variable independent of the Euclidean norm $|X| = (X_1^2 + \dots + X_n^2)^{1/2}$. Clearly, G represents a mixture of a family of centered Gaussian measures on the line and has characteristic function

$$g(t) = \mathbf{E}e^{-t^2|X|^2/(2n)}, \quad t \in \mathbf{R}, \qquad (8)$$

while F has characteristic function

$$f(t) = \mathbf{E} \prod_{k=1}^{n} \cos\left(\frac{tX_k}{\sqrt{n}}\right). \tag{9}$$

In order to bound the Lévy distance $L(F,G)$, the following general elementary observation, not using the condition (1), can be applied.

Lemma 4. *Assume* $\mathbf{E}|X|^2 \leq n$. *For all* $\alpha > 0$ *and* $|t| \leq \frac{1}{2\alpha}$,

$$|f(t) - g(t)| \leq \frac{1}{9}\,\alpha^2 t^4 + 2\mathbf{P}\left\{\frac{\max\{|X_1|,\dots,|X_n|\}}{\sqrt{n}} > \alpha\right\}.$$

Proof. By Taylor's expansion, in the interval $|s| \leq \frac{1}{2}$, we have $\cos(s) = e^{-\frac{s^2}{2} - u(s)}$ with u satisfying $0 \leq u(s) \leq \frac{s^4}{9}$. Therefore, provided that $|\frac{X_k}{\sqrt{n}}| \leq \alpha$, for all $k \leq n$, and $\alpha|t| \leq \frac{1}{2}$,

$$\prod_{k=1}^{n} \cos\left(\frac{tX_k}{\sqrt{n}}\right) = \exp\left\{-\frac{t^2|X|^2}{2n} - \sum_{k=1}^{n} u\left(\frac{tX_k}{\sqrt{n}}\right)\right\}$$

with $0 \leq \sum_{k=1}^n u(\frac{tX_k}{\sqrt{n}}) \leq \frac{1}{9}\, \max_k |\frac{tX_k}{\sqrt{n}}|^2 \sum_{k=1}^n |\frac{tX_k}{\sqrt{n}}|^2 \leq \frac{\alpha^2 t^4}{9}\, \frac{|X|^2}{n}$. So,

$$e^{-\frac{t^2|X|^2}{2n}} \geq \prod_{k=1}^{n} \cos\left(\frac{tX_k}{\sqrt{n}}\right) \geq e^{-\frac{t^2|X|^2}{2n} - \frac{\alpha^2 t^4}{9}\,\frac{|X|^2}{n}}.$$

Taking the expectations and using $|\prod_{k=1}^n \cos(\frac{tX_k}{\sqrt{n}}) - e^{-t^2|X|^2/(2n)}| \leq 2$ for the complementary event $\frac{\max\{|X_1|,\dots,|X_n|\}}{\sqrt{n}} > \alpha$, we thus get

$$|f(t) - g(t)| \leq 2\mathbf{P}\left\{\frac{\max\{|X_1|,\dots,|X_n|\}}{\sqrt{n}} > \alpha\right\} + \mathbf{E}\, e^{-\frac{t^2|X|^2}{2n}}\left(1 - e^{-\frac{\alpha^2 t^4}{9}\,\frac{|X|^2}{n}}\right).$$

The last term is bounded by $\mathbf{E}(1 - e^{-\frac{\alpha^2 t^4}{9}\,\frac{|X|^2}{n}}) \leq 1 - e^{-\frac{\alpha^2 t^4}{9}\,\frac{\mathbf{E}|X|^2}{n}} \leq \frac{\alpha^2 t^4}{9}$ where we applied Jensen's inequality together with the assumption $\mathbf{E}|X|^2 \leq n$. Lemma 4 follows.

Via the inequality of Lemma 4, with mild integrability assumptions on the distribution of X, one can study a rate of closeness of F and thus of F_ε to the distribution function G. One can start, for instance, with the moment assumption

$$\mathbf{E}|X_k|^4 \leq \beta, \quad 1 \leq k \leq n, \tag{10}$$

implying $\mathbf{P}\{\frac{\max\{|X_1|,\dots,|X_n|\}}{\sqrt{n}} > \alpha\} \leq \frac{\beta}{\alpha^4 n}$, so that, by Lemma 4,

$$|f(t) - g(t)| \leq \frac{1}{9}\,\alpha^2 t^4 + \frac{2\beta}{\alpha^4 n}, \quad \text{as soon as} \quad |t| \leq \frac{1}{2\alpha}.$$

Minimizing the right-hand side over all $\alpha > 0$, we obtain that

$$|f(t) - g(t)| \leq \frac{\beta^{1/3}|t|^{16/3}}{3n^{1/3}}, \quad \text{provided that} \ \ |t| \leq \frac{n^{1/4}}{24\beta^{1/2}}.$$

Now apply Zolotarev's estimate, [Z], [P], to get

$$\begin{aligned} L(F,G) &\leq \frac{1}{\pi} \int_0^T \left| \frac{f(t) - g(t)}{t} \right| dt + 2e\, \frac{\log T}{T} \quad (T > 1.3) \\ &\leq \frac{\beta^{1/3}\, T^{16/3}}{16\pi\, n^{1/3}} + 2e\, \frac{\log T}{T}, \qquad \text{if} \ \ 1.3 < T \leq \frac{n^{1/4}}{24\beta^{1/2}}. \end{aligned}$$

Taking $T = \frac{n^{1/19}}{\beta^{1/19}}$ and using $\beta \geq 1$, we will arrive at the estimate of the form

$$L(F,G) \leq C\, \frac{\beta^{1/19} + \log n}{n^{1/19}}, \quad n \geq C\beta^{37/15},$$

up to some numerical constant C. Higher moments or exponential integrability assumption improve this rate of convergence, but it seems, with the above argument, the rate of Corollary 1 cannot be reached.

On the other hand, the closeness of G to the normal distribution function Φ requires some additional information concerning the rate of convergence of $\frac{X_1^2 + \ldots + X_n^2}{n}$ to 1. For example, the property $\mathrm{Var}(|X|^2) \leq O(n)$ guarantees a rate of the form $L(G,\Phi) = O(n^{-c})$ with a certain power $c > 0$. Thus, together with the moment assumption (10), one arrives at the bound $L(F,\Phi) = O(n^{-c})$.

Finally, let us note that G is determined via the distribution of the Euclidean norm $|X|$, so it is stable under the choice of the basis in $\mathbf{R}^n$. The condition (10) is stated for the canonical basis in $\mathbf{R}^n$, and the appropriate basis free assumption may read as

$$\sup_{\theta \in S^{n-1}} \mathbf{E}\, |\langle \theta, X \rangle|^p \leq \beta_p, \quad p > 2. \tag{11}$$

Then, at the expense of the rate of closeness, one may formulate an analogue of Theorem 1 for the distribution G in the place of F and with respect to an arbitrary basis in $\mathbf{R}^n$. The inequality (11) includes many interesting classes of distributions such as log-concave probability measures satisfying (1), for example.

Proof of Theorem 2. Denote by f_n and g_n the characteristic functions defined for the random vectors $(X_{n,1}, \ldots, X_{n,n})$ according to formulas (9) and (8), respectively. Also, according to (3), denote by $F^{(n)}$ the corresponding average distribution functions.

In view of the assumption a), one can select a sequence $\alpha_n \downarrow 0$ such that $\mathbf{P}\{\frac{\max\{|X_{1,1}|, \ldots, |X_{n,n}|\}}{\sqrt{n}} > \alpha_n\} \to 0$, as $n \to \infty$. Then, by Lemma 4, for all $t \in \mathbf{R}$, $|f_n(t) - g_n(t)| \to 0$, as $n \to \infty$. On the other hand, the condition b) readily implies $g_n(t) \to e^{-t^2/2}$, so $f_n(t) \to e^{-t^2/2}$. Thus, $L(F^{(n)}, \Phi) \to 0$.

Now, given an infinite sequence $\varepsilon \in \{-1,1\}^\infty$, denote by $T_n(\varepsilon)$ its projection $(\varepsilon_1,\dots,\varepsilon_n)$. It remains to show that $L(F_{T_n(\varepsilon)}, F^{(n)}) \to 0$, for μ_∞-almost all ε. Fix any small number $p>0$, and take a sequence $\delta_n \to 0^+$ such that

$$\sum_{n=1}^{\infty} Cn^{1/4}\, e^{-cn\delta_n^8} \leq p,$$

where C and c are numerical constants from Theorem 1 (δ_n may depend on p). Then the application of (4) yields

$$\begin{aligned} &\mu_\infty\big\{\varepsilon : L(F_{T_n(\varepsilon)}, F^{(n)}) > \delta_n, \ \text{for some } n \geq 1\big\} \\ &\leq \sum_{n=1}^{\infty} \mu_\infty\big\{L(F_{T_n(\varepsilon)}, F^{(n)}) > \delta_n\big\} \\ &= \sum_{n=1}^{\infty} \mu_n\big\{\varepsilon = (\varepsilon_1,\dots,\varepsilon_n) \colon L(F_\varepsilon, F^{(n)}) > \delta_n\big\} \leq p. \end{aligned}$$

Therefore, $L(F_{T_n(\varepsilon)}, F^{(n)}) \leq \delta_n$, for all $n \geq 1$ and for all ε except for a set of μ_∞-measure at most p. That is,

$$\mu_\infty\left\{\varepsilon : \sup_{n\geq 1}\Big(L(F_{T_n(\varepsilon)}, F^{(n)}) - \delta_n\Big) \leq 0\right\} \geq 1-p. \tag{12}$$

But since $\delta_n \to 0$,

$$\begin{aligned} \sup_{n\geq 1}\Big(L(F_{T_n(\varepsilon)}, F^{(n)}) - \delta_n\Big) &\geq \limsup_{n\to\infty}\Big(L(F_{T_n(\varepsilon)}, F^{(n)}) - \delta_n\Big) \\ &= \limsup_{n\to\infty} L(F_{T_n(\varepsilon)}, F^{(n)}). \end{aligned}$$

Consequently, (12) implies $\mu_\infty\big\{\limsup_{n\to\infty} L(F_{T_n(\varepsilon)}, F^{(n)}) = 0\big\} \geq 1-p$. The probability on the left does not depend on p, and letting $p \to 0$ finishes the proof.

Acknowledgement. We would like to thank B. Zegarlinski and A. Grigoryan for the invitation to visit to the Imperial College, London, where this work was partly done.

References

[A-B-P] Antilla, M., Ball, K., Perissinaki, I. (1998): The central limit problem for convex bodies. Preprint

[B] Bobkov, S.G. On concentration of distributions of random weighted sums. Ann. Probab., to appear

[B-G] Bobkov, S.G., Götze, F. (1999): Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. J. Funct. Anal., **163**, No. 1, 1–28

[Bo] Bohman, H. (1961): Approximate Fourier analysis of distribution functions. Arkiv Mat., **4**, No. 2-3, 99–157

[D-F] Diaconis, P., Freedman, D. (1984): Asymptotics of graphical projection pursuit. Ann. Stat., **12**, No. 3, 793–815

[Le] Ledoux, M. (1999): Concentration of measure and logarithmic Sobolev inequalities. Séminaire de Probabilités, XXXIII, Lecture Notes in Math., **1709**, 120–216

[P] Petrov, V.V. (1987): Limit theorems for sums of independent random variables. Moscow, Nauka (in Russian)

[S] Sudakov, V.N. (1978): Typical distributions of linear functionals in finite-dimensional spaces of higher dimension. Soviet Math. Dokl., **19**, No. 6, 1578–1582. Translated from: Dokl. Akad. Nauk SSSR, **243**, No. 6

[W] von Weizsäcker, H. (1997): Sudakov's typical marginals, random linear functionals and a conditional central limit theorem. Probab. Theory Rel. Fields, **107**, 313–324

[Z] Zolotarev, B.M. (1971): An estimate for the difference between distributions in Lévy metric. Proceedings of Steklov's Math. Institute, **112**, 224–231 (in Russian)