V. Yu. Bentkus, A Lyapunov-type bound in R^d, *Teor. Veroyatnost. i Primenen.*, 2004, Volume 49, Issue 2, 400–410

DOI: https://doi.org/10.4213/tvp230

*BENTKUS V.**

A LYAPUNOV TYPE BOUND IN $\mathbf{R}^{D}$ [1)]

Пусть $X_1, \ldots, X_n$ — независимые случайные векторы со значениями в $\mathbf{R}^d$ такие, что $\mathbf{E}X_k = 0$ для любого k. Положим $S = X_1 + \cdots + X_n$. Будем предполагать, что ковариационный оператор суммы S — обозначим его C^2 — обратим. Пусть Z — центрированный гауссовский случайный вектор такой, что ковариации векторов S и Z равны. Обозначим $\mathscr{C}$ класс всех выпуклых подмножеств $\mathbf{R}^d$. Мы доказываем оценку типа Ляпунова для $\Delta = \sup_{A\in\mathscr{C}} |\mathbf{P}\{S \in A\} - \mathbf{P}\{Z \in A\}|$. А именно, $\Delta \leqslant cd^{1/4}\beta$ с $\beta = \beta_1 + \cdots + \beta_n$ и $\beta_k = \mathbf{E}|C^{-1}X_k|^3$, где c — абсолютная постоянная. Если случайные величины $X_1, \ldots, X_n$ независимы и одинаково распределены и X_k имеет единичную ковариацию, то полученная оценка преобразуется к виду $\Delta \leqslant cd^{1/4}\mathbf{E}\,|X_1|^3/\sqrt{n}$. Вопрос, может ли множитель $d^{1/4}$ быть убран или заменен на лучший (например, на 1), остается открытым.

Ключевые слова и фразы: многомерный случай, центральная предельная теорема, оценка Берри–Эссеена, оценка Ляпунова, зависимость от размерности, зависимые случайные величины, разнораспределенные случайные величины.

1. Introduction and results. We provide a Lyapunov type bound with hopefully optimal dependence on dimension. It extends to the nonidentically distributed case a Berry–Esseen bound for independent identically distributed summands established in [4], [3]. To prove the bound we use a new method recently introduced in [2], combining it with some technical elements from [4], [3].

Let $\mathbf{R}^d$ be a real Euclidean d-dimensional space of vectors $x = (x_1, \ldots, x_d)$ with the norm $|x|^2 = x_1^2 + \cdots + x_d^2$ and the scalar product $\langle x, x\rangle = |x|^2$. Let $X_1, \ldots, X_n$ be independent random vectors with common mean $\mathbf{E}X_j = 0$. Write $S = X_1 + \cdots + X_n$. Throughout we assume that S has a nondegenerated distribution in the sense that the covariance operator, say $C^2 = \operatorname{cov} S$, is invertible ($C$ stands for the positive root of C^2). Let Z be a Gaussian random vector such that $\mathbf{E}Z = 0$ and $\operatorname{cov} S$ and $\operatorname{cov} Z$ are equal, that is, $\langle C^2x, x\rangle = \mathbf{E}\langle S, x\rangle^2 = \mathbf{E}\langle Z, x\rangle^2$, for all $x \in \mathbf{R}^d$. Write

$$\beta = \beta_1 + \cdots + \beta_n, \qquad \beta_k = \mathbf{E}|C^{-1}X_k|^3,$$

and

$$\Delta(\mathscr{C}) = \sup_{A\in\mathscr{C}} |\mathbf{P}\{S \in A\} - \mathbf{P}\{Z \in A\}|,$$

where $\mathscr{C}$ stands for the class of all convex subsets of $\mathbf{R}^d$.

Our main result is the following bound.

Theorem 1.1. *There exists an absolute positive constant c such that*

$$\Delta(\mathscr{C}) \leqslant cd^{1/4}\beta. \tag{1.1}$$

If the random variables $X_1, \ldots, X_n$ are i.i.d. and X_k have identity covariance, then the bound (1.1) *specifies to $\Delta(\mathscr{C}) \leqslant cd^{1/4}\mathbf{E}\,|X_1|^3/\sqrt{n}$.*

---

* Vilnius Institute of Mathematics and Informatics, Akademijos 4, 232600, Vilnius, Lithuania; e-mail: bentkus@julius.ktl.mii.lt

1) Research supported by the DFG Forschergruppe FOR 399/1-1 «Spektrale Analysis, asymptotische Verteilungen und stochastische Dynamiken» at University of Bielefeld.

Theorem 1.1 is just a special case of Theorem 1.2 below, where we consider abstract classes of sets instead of the class $\mathscr{C}$.

The question of the dependence of the Berry–Esseen bound on dimension has drawn attention of many authors, let us mention here only [8] and [6], where the bounds depend on the dimension as d. We use a new method recently introduced in [2], combining it with basic technical features from [4] and [3]. Besides the basic representation (3.4), the proof involves induction in n and Taylor expansions. In the context of the Lindeberg method, direct induction proof is a commonly used tool to obtain convergence rates in the central limit theorem, see, for example, [12], [11], [13]. As in [4], [2], we apply a special smoothing with differentiable functions approximating indicator functions of convex sets. The smoothness properties of these functions rely on nice differentiability properties of Euclidean distances. We do not know how to prove the result using Fourier transforms as in [5]. Until now the Stein method led to the factor d, see [6].

Whether the dependence on d in (1.1) is optimal, remains an open question. One can show (see [8]) that $\Delta(\mathscr{C}) \leqslant c_d\beta$ implies $c_d \geqslant c_0$, where $c_0 > 0$ is an absolute positive constant. Hence the constant in $\Delta(\mathscr{C}) \leqslant c_d\beta$ satisfies $c_0 \leqslant c_d \leqslant cd^{1/4}$.

We need some definitions and conditions related to classes of subsets of $\mathbf{R}^d$ and the standard normal distribution Φ.

For a class, say $\mathscr{A}$, of subsets $A \subset \mathbf{R}^d$, consider the following conditions:

i) class $\mathscr{A}$ is invariant under affine symmetric transformations, that is, $DA + a \in \mathscr{A}$ if $a \in \mathbf{R}^d$ and $D\colon \mathbf{R}^d \to \mathbf{R}^d$ is a linear symmetric invertible operator;

ii) class $\mathscr{A}$ is invariant under taking ε-neighborhoods, for all $\varepsilon > 0$.

More precisely, A^ε, $A^{-\varepsilon} \in \mathscr{A}$ if $A \in \mathscr{A}$. Here

$$A^\varepsilon = \{x \in \mathbf{R}^d\colon \rho_A(x) \leqslant \varepsilon\} \quad \text{and} \quad A^{-\varepsilon} = \{x \in A\colon B_\varepsilon(x) \subset A\},$$

where $\rho_A(x) = \inf_{y\in A} |x - y|$ is the distance between $A \subset \mathbf{R}^d$ and $x \in \mathbf{R}^d$, and $B_\varepsilon(x) = \{y \in \mathbf{R}^d\colon |x - y| \leqslant \varepsilon\}$.

The class $\mathscr{C}$ of convex subsets of $\mathbf{R}^d$ satisfies the conditions i) and ii). The class of rectangles does not satisfy the condition ii) since an ε-neighborhood of a rectangle is not a rectangle.

We consider the following condition on a class $\mathscr{A}$ of subsets of $\mathbf{R}^d$ and the standard normal distribution Φ with the density

$$\eta(x) = (2\pi)^{-d/2} \exp\left\{ -\frac{|x|^2}{2} \right\}, \qquad x \in \mathbf{R}^d \tag{1.2}$$

there exist constants, say $a_d = a_d(\mathscr{A})$, depending only on d and $\mathscr{A}$ such that

$$\Phi\{A^\varepsilon \setminus A\} \leqslant a_d\varepsilon, \quad \Phi\{A \setminus A^{-\varepsilon}\} \leqslant a_d\varepsilon, \quad \text{for all} \quad A \in \mathscr{A} \text{ and } \varepsilon > 0. \tag{1.3}$$

We shall refer to $a_d(\mathscr{A})$ as to the isoperimetric constant of $\mathscr{A}$. For the class $\mathscr{B}$ of balls we have $\sup_{d\geqslant 1} a_d(\mathscr{B}) < \infty$. Combining an upper bound from [1] and a lower bound from [9], [10], we have

$$c \leqslant d^{-1/4} \sup_{A\in\mathscr{C}} \int_{\partial A} \eta(x)\, ds \leqslant 4, \qquad cd^{1/4} \leqslant a_d(\mathscr{C}) \leqslant 4d^{1/4}, \tag{1.4}$$

where c is an absolute constant and ds is the surface measure on the boundary ∂A of A. A proof of Nazarov's result is provided in [2]. We shall use the upper bound $a_d(\mathscr{C}) \leqslant 4d^{1/4}$. Indirectly (1.4) supports a conjecture that the dependence on d as in Theorem 1.1 is optimal or near to optimal since in all cases which we can analyze the isoperimetric constant controls the constant in the Berry–Esseen bound up to an absolute factor.

Theorem 1.1 is an obvious particular case of the following result.

Theorem 1.2. *Let a class $\mathscr{A}$ of convex sets satisfy the conditions* i) *and* ii). *Furthermore, assume that $\mathscr{A}$ and the standard normal distribution satisfy the condition* (1.3). *Then there exists an absolute constant $M > 0$ such that*

$$\Delta(\mathscr{A}) \leqslant Mb_d\beta, \qquad b_d = \max\{1, a_d(\mathscr{A})\}.$$

The conditions i)–ii) on the class $\mathscr{A}$ can be relaxed using a little more refined techniques. In the i.i.d. case one can relax the requirement i) on $\mathscr{A}$ assuming that $\mathscr{A}$ is invariant under rescaling by scalars and shifting, see [3].

Applications of Theorem 1.2 to other classes of sets are restricted due to our poor knowledge about isoperimetric constants $a_d(\mathscr{A})$. It is easy to see that $a_d(\mathscr{H}) = 1/\sqrt{2\pi}$ in the case of the class $\mathscr{H}$ of all affine half-spaces of $\mathbf{R}^d$. However, already for a minimal class, say $\mathscr{E}$, satisfying i) and ii) and containing all ellipsoids, the behavior of $a_d(\mathscr{E})$ is not known. A similar question is open as well in the case of a minimal class, say $\mathscr{H}_m$, satisfying i) and ii) and containing all sets of the form $A_1 \cap \cdots \cap A_m$ with $A_k \in \mathscr{H}$, for all k.

2. Auxiliary lemmas. The following smoothing lemma is contained in [3].

Lemma 2.1. *Let $\varepsilon > 0$ be a positive number. Let S and Z be arbitrary random vectors. Write*

$$\gamma(A) = \mathbf{P}\{S \in A\} - \mathbf{P}\{Z \in A\}, \qquad \gamma(\varphi) = \mathbf{E}\varphi(S) - \mathbf{E}\varphi(Z).$$

Let a class $\mathscr{A}$ of sets satisfy ii). *Assume that we have a family, say $\{\varphi_{\varepsilon,A}\colon A \in \mathscr{A}\}$, of functions $\varphi_{\varepsilon,A}\colon \mathbf{R}^d \to \mathbf{R}$ such that $0 \leqslant \varphi_{\varepsilon,A} \leqslant 1$, $\varphi_{\varepsilon,A}(x) = 1$ for $x \in A$, $\varphi_{\varepsilon,A}(x) = 0$ for $x \notin A^\varepsilon$. Then*

$$\sup_{A\in\mathscr{A}} |\gamma(A)| \leqslant \sup_{A\in\mathscr{A}} |\gamma(\varphi_{\varepsilon,A})| + \max\Big\{ \sup_{A\in\mathscr{A}} \mathbf{P}\{Z \in A^\varepsilon \setminus A\},\ \sup_{A\in\mathscr{A}} \mathbf{P}\{Z \in A \setminus A^{-\varepsilon}\}\Big\}.$$

For a sufficiently smooth function $f\colon \mathbf{R}^d \to \mathbf{R}$, we shall apply the Taylor formula

$$f(x+h) = f(x) + f'(x)\,h + \cdots + \frac{1}{s!}\,f^{(s)}(x)\,h^s + \frac{1}{s!}\,\mathbf{E}(1-\tau)^s f^{(s+1)}(x+\tau h)\,h^{s+1}, \qquad (2.1)$$

where τ is a random variable uniformly distributed on the interval $[0,1]$ and independent of all other random variables. In (2.1) we write $f^{(s)}(x)\,h^s \equiv f^{(s)}(x)\,h\cdots h$ for the s-th Fréchet derivative of f in the direction h. The following lemma is taken from [4], see [14] for an earlier version.

Lemma 2.2. *Let a set $A \subset \mathbf{R}^d$ be convex. Then for any $\varepsilon > 0$ there exists a function φ (which depends only on ε and A) such that $\varphi(x) = 1$, for $x \in A$, $\varphi(x) = 0$, for $x \in \mathbf{R}^d \setminus A^\varepsilon$, $0 \leqslant \varphi \leqslant 1$, and*

$$|\varphi'(x)\,h| \leqslant \frac{2}{\varepsilon}|h|, \quad |\varphi'(x)\,h - \varphi'(y)\,h| \leqslant \frac{8|x-y|}{\varepsilon^2}\,|h|, \qquad h \in \mathbf{R}^d. \qquad (2.2)$$

Furthermore, we can choose φ to get the form $\varphi(x) = \psi(\rho(x)/\varepsilon)$, where $\rho(x)$ is the distance between x and A, and $\psi\colon \mathbf{R} \to \mathbf{R}$ is a continuously differentiable nonnegative nonincreasing function such that $\int_{\mathbf{R}} |\psi'(t)|\,dt = 1$.

Lemma 2.3. *Let $p\colon \mathbf{R}^d \to \mathbf{R}$ be an infinitely many times differentiable function decaying together with its derivatives more rapidly than powers $|x|^{-m}$, as $|x| \to \infty$. Assume that a function $f\colon \mathbf{R}^d \to \mathbf{R}$ satisfies the Lipschitz condition $|f(x) - f(y)| \leqslant a|x-y|$. Then*

$$J = \left|\int_{\mathbf{R}^d} f(y)\,p'(y)\,h\,dy\right| \leqslant a|h| \int_{\mathbf{R}^d} \mathbb{I}\{y \in \operatorname{supp} f\}\,|p(y)|\,dy, \qquad h \in \mathbf{R}^d, \qquad (2.3)$$

where $\operatorname{supp} f = \overline{\{x\colon f(x) = 0\}}$ *is the support of f.*

P r o o f. Let $\delta > 0$. Using $p'(y)\,h = \lim_{t\downarrow 0}(p(y+th) - p(y))/t$, changing the variables and applying the Lipschitz condition, we have

$$\begin{aligned} J &\leqslant \limsup_{t\downarrow 0} \left|\int_{\mathbf{R}^d} t^{-1} f(y)(p(y+th) - p(y))\,dy\right| \\ &= \limsup_{t\downarrow 0} \left|\int_{\mathbf{R}^d} \mathbb{I}\{y \in (\operatorname{supp} f)^\delta\}\, t^{-1}(f(y-th) - f(y))\,p(y)\,dy\right| \\ &\leqslant |a|\,|h| \int_{\mathbf{R}^d} \mathbb{I}\{y \in (\operatorname{supp} f)^\delta\}\,|p(y)|\,dy. \end{aligned} \qquad (2.4)$$

Passing in (2.4) to the limit as $\delta \downarrow 0$, we conclude the proof of (2.3).

3. Proof of Theorem 1.2. By $\tau, \tau_1, \tau_2, \dots$ we denote independent copies of a random variable τ uniformly distributed in the interval $[0,1]$. By $N, N_1, N_2, \dots$ we denote independent copies of a standard normal random vector N. We assume everywhere that all random variables and vectors are independent in the aggregate if the contrary is not clear from the context. Throughout $A \ll B$ means that there exists an absolute positive constant, say c, such that $A \leqslant cB$.

Assuming that $C^2 = \operatorname{cov} S = \mathbb{I}$ is the identity operator, we shall prove that

$$\Delta \stackrel{\text{def}}{=} \Delta(\mathscr{A}) \leqslant M b_d \beta \tag{3.1}$$

with $\beta = \beta_1 + \dots + \beta_n$, $\beta_k = \mathbf{E}\,|X_k|^3$ holds for all n provided that an absolute constant $M \geqslant 4$ is sufficiently large (recall that $b_d = \max\{1, a_d\}$). The assumption $C = \mathbb{I}$ does not restrict the generality since the class $\mathscr{A}$ is invariant with respect to symmetric linear transformations and therefore we can rescale random vectors S and Z replacing them by $C^{-1}S$ and $C^{-1}Z$, respectively.

We use induction in n.

Let us prove (3.1) for small n such that $n \leqslant d^3M^2$. Of course, $n = 1$ satisfies $n \leqslant d^3M^2$, since $M \geqslant 1$ and $d \geqslant 1$. Using $C = \mathbb{I}$ and applying Hölder's inequality twice, we have

$$d = \mathbf{E}|S|^2 = \sum_{k=1}^{n} \mathbf{E}\,|X_k|^2 \leqslant n^{1/3} \Bigl(\sum_{k=1}^{n} (\mathbf{E}\,|X_k|^2)^{3/2} \Bigr)^{2/3} \leqslant n^{1/3}\beta^{2/3},$$

Hence, $1 \leqslant \beta d^{-3/2} n^{1/2}$. Therefore the trivial bound $\Delta \leqslant 1$ and $b_d \geqslant 1$ yield (3.1). Indeed, $\Delta \leqslant \beta d^{-3/2} n^{1/2} \leqslant M\beta \leqslant M b_d \beta$, provided that $d^{-3/2}n^{1/2}/M \leqslant 1$. But $d^{-3/2}n^{1/2}/M \leqslant 1$ is equivalent to the assumption $n \leqslant d^3M^2$.

Write $Z \stackrel{\mathscr{D}}{=} Y_1 + \dots + Y_n$, where $Y_1, \dots, Y_n$ are independent centered Gaussian vectors such that $\operatorname{cov} X_k = \operatorname{cov} Y_k$, for all k (recall, that we assume that Y's are independent of X's as well). We use the following notation:

$$U_k = S - X_k = \sum_{m \neq k,\ 1 \leqslant m \leqslant n} X_m, \qquad V_k = Z - Y_k = \sum_{m \neq k,\ 1 \leqslant m \leqslant n} Y_m. \tag{3.2}$$

First we exclude trivial cases where there is an X_k with too large covariance. More precisely, let us show that in the proof of (3.1) we can assume that the covariances, say

$$P_k^2 \stackrel{\text{def}}{=} \operatorname{cov} U_k = \operatorname{cov} V_k, \qquad Q_k \stackrel{\text{def}}{=} P_k^{-1},$$

satisfy $\|Q_k\| \leqslant 2$, where $\|A\|$ stands for the maximal eigenvalue of A. Indeed, otherwise there is k such that $\|Q_k\| > 2$. Let us show that $\|Q_k\| > 2$ trivially implies (3.1). If $\|Q_k\| > 2$, then P_k^2 has an eigenvalue, say σ^2, such that $\sigma^2 < \frac{1}{4}$. Since $P_k^2 = I - \operatorname{cov} X_k$, it follows that $\operatorname{cov} X_k$ has the eigenvalue $1 - \sigma^2 > \frac{3}{4}$. Applying Hölder's inequality, we have

$$1 < \frac{4(1-\sigma^2)}{3} \leqslant 2\mathbf{E}\,|X_k|^2 \leqslant 2(\mathbf{E}\,|X_k|^3)^{2/3} \leqslant 2\beta^{2/3},$$

and it follows that $1 < 4\beta$. Therefore $\Delta \leqslant 4\beta$ since $\Delta \leqslant 1$. The bound $\Delta \leqslant 4\beta$ clearly implies (3.1) since $M \geqslant 4$ and $b_d \geqslant 1$.

Now assume that (3.1) holds for $1, \dots, n-1$. We shall prove that then (3.1) holds for n as well. As it is noted above, in the proof of (3.1) we can assume that $n \geqslant d^3M^2$ and $\|Q_k\| \leqslant 2$. Let $0 < \varepsilon < 1$ be a number to be chosen later. We apply the smoothing inequality (see Lemma 2.1) with $\varphi = \varphi_{\varepsilon, A}$ from Lemma 2.2. Using (1.3) to estimate the isoperimetric constant, we get

$$\Delta \leqslant b_d \varepsilon + \sup_{A \in \mathscr{A}} |\Delta(\varphi)|, \qquad \Delta(\varphi) = \mathbf{E}\varphi(Z) - \mathbf{E}\varphi(S). \tag{3.3}$$

To estimate $\Delta(\varphi)$, we use the following representation (see [2]). Let α be a real valued random variable uniformly distributed in the interval $[0, \pi/2]$ and independent of all other random variables. Write

$$p = \cos\alpha, \quad q = \sin\alpha, \quad \overline{X}_k = pX_k + qY_k, \quad \overline{X}'_k = -qX_k + pY_k,$$

and $\overline{U}_k = \overline{X}_1+\dots+\overline{X}_{k-1}+\overline{X}_{k+1}+\dots+\overline{X}_n = pU_k+qV_k$. Let φ: $\mathbf{R}^d \to \mathbf{C}$ be a sufficiently smooth complex-valued function. Then

$$\Delta(\varphi) = \mathbf{E}\varphi(Z) - \mathbf{E}\varphi(S) = \frac{\pi}{2}\sum_{k=1}^{n}\Delta_k(\varphi) \tag{3.4}$$

with $\Delta_k(\varphi) = \mathbf{E}\varphi'(\overline{U}_k + \overline{X}_k)\,\overline{X}'_k$. For given α, the vectors $\overline{X}_k$ and $\overline{X}'_k$ satisfy

$$\mathbf{E}L(\overline{X}_k) = \mathbf{E}L(\overline{X}'_k) = 0, \qquad \mathbf{E}B(\overline{X}_k, \overline{X}'_k) = 0, \tag{3.5}$$

provided that L: $\mathbf{R}^d \to \mathbf{C}$ is a linear function, and B: $\mathbf{R}^d \times \mathbf{R}^d \to \mathbf{C}$ is a bilinear form.

Let $0 < \gamma < \pi/2$. Below we prove that

$$|\Delta_k(\varphi)| \ll \varepsilon^{-2}(M\beta + b_d\varepsilon)\,\beta_k \sin\gamma + \frac{M\beta\beta_k}{\sin\gamma} + \beta_k \tag{3.6}$$

with $\beta_k = \mathbf{E}\,|X_k|^3$. Summing (3.6) with respect to $k = 1, \dots, n$ and using (3.4), we obtain

$$|\Delta(\varphi)| \ll \varepsilon^{-2}(M\beta + b_d\varepsilon)\,\beta \sin\gamma + \frac{M\beta^2}{\sin\gamma} + \beta. \tag{3.7}$$

Combining (3.7) and (3.3) and using $b_d \geqslant 1$, we derive

$$\Delta \ll b_d\Big(\varepsilon + \varepsilon^{-2}(M\beta + \varepsilon)\,\beta\sin\gamma + \frac{M\beta^2}{\sin\gamma} + \beta\Big). \tag{3.8}$$

We choose $0 < \gamma < \pi/2$ such that $\sin\gamma = \varepsilon$. The choice is possible since $0 < \varepsilon < 1$. Then (3.8) yields

$$\Delta \ll b_d(\varepsilon + \beta + M\beta^2\varepsilon^{-1}). \tag{3.9}$$

We choose $\varepsilon = \beta\sqrt{M}$. We have to ensure that $\varepsilon < 1$. But if $\varepsilon = \beta\sqrt{M} \geqslant 1$, then using the trivial bound $\Delta \leqslant 1$, we obtain $\Delta \leqslant \beta\sqrt{M} \leqslant b_d M\beta$, since $b_d \geqslant 1$ and $M \geqslant 1$, and there is nothing to prove. Thus, setting $\varepsilon = \beta\sqrt{M}$ the bound (3.9) yields

$$\Delta \ll b_d(\beta\sqrt{M} + \beta) \ll b_d\beta\sqrt{M}, \tag{3.10}$$

since $M \geqslant 1$. We can rewrite (3.10) as $\Delta \leqslant c b_d\beta\sqrt{M}$ with some absolute constant c. This inequality yields the desired bound $\Delta \leqslant b_d M\beta$ provided that

$$c\sqrt{M} \leqslant M \Leftrightarrow c^2 \leqslant M.$$

Taking $M = c^2$ completes the proof of (3.1).

It remains to obtain the bound (3.6) for $\Delta_k(\varphi)$. While estimating $\Delta_k(\varphi)$, we omit the index k, writing everywhere

$$X, \overline{X}, Y, U, \overline{U}, V, P, Q \qquad \text{instead of} \quad X_k, \overline{X}_k, Y_k, U_k, \overline{U}_k, V_k, P_k, Q_k$$

with a sole exception $\beta_k = \mathbf{E}\,|X|^3$.

Introduce the indicator functions $\mathbb{I}_1$ and $\mathbb{I}_2$ such that

$$\mathbb{I}_1 = \mathbb{I}\{0 \leqslant \alpha \leqslant \gamma\}, \quad \mathbb{I}_2 = \mathbb{I}\Big\{\gamma \leqslant \alpha \leqslant \frac{\pi}{2}\Big\}, \qquad \mathbb{I}_1 + \mathbb{I}_2 \equiv 1.$$

The number $0 < \gamma < \pi/2$ is specified above. Let us note that the value of γ is inessential for the further proof. Split $\Delta_k(\varphi) = I_1 + I_2$ with

$$I_1 = \mathbf{E}\,\mathbb{I}_1\varphi'(\overline{U} + \overline{X})\,\overline{X}', \qquad I_2 = \mathbf{E}\,\mathbb{I}_2\varphi'(\overline{U} + \overline{X})\,\overline{X}'. \tag{3.11}$$

Write

$$I_3 = \mathbf{E}\,\mathbb{I}_2\varphi'(PN + \overline{X})\,\overline{X}'. \tag{3.12}$$

It is clear that $|\Delta_k(\varphi)| \leqslant |I_1| + |I_2 - I_3| + |I_3|$. Below we prove that

$$|I_1| \ll \varepsilon^{-2}(M\beta + b_d\varepsilon)\,\beta_k \sin\gamma, \tag{3.13}$$

$$|I_2 - I_3| \ll \frac{M\beta\beta_k}{\sin\gamma}, \tag{3.14}$$

$$|I_3| \ll \beta_k. \tag{3.15}$$

Bounds (3.13)–(3.15) yield (3.6). Hence, to conclude the proof of the theorem, it remains to establish (3.13)–(3.15). We start with the proof of (3.15) since in this case the proof is relatively simple due to the presence of the large Gaussian component N. The density of N is a smooth function with bounded derivatives, and the proof reduces to an expansion in a Taylor series. The idea of proof of (3.13) and (3.14) is to use the smoothness properties of φ (see (2.2)) and of the density of the Gaussian component qV in $\overline{U} = pU + qV$, as well as the induction assumption together with bounds for Gaussian measures of $A^\varepsilon \setminus A$. Altogether we need 3 derivatives. In essence, th(function φ can provide 2 derivatives. The density of the Gaussian component qV is a function of the class C^∞ and its s-th derivative is bounded by q^{-s}. Choosing γ above we just compared the size of derivatives of φ and of the Gaussian density. On the interval $\alpha \in (0,\gamma)$ the derivatives of φ are smaller than derivatives of the Gaussian density since $q^{-s} = (\sin\alpha)^{-s} \sim \alpha^{-s} \to \infty$ as $\alpha \downarrow 0$. Therefore for $\alpha \in (0,\gamma)$ (that is, while estimating $|I_1|$), we use the maximal available smoothness of φ, which brings 2 derivatives. The lacking derivative provides the Gaussian component. On the interval $\alpha \in (\gamma, 1)$ (that is, while estimating $|I_2 - I_3|$), we use only smoothness properties of the Gaussian component. We note as well, that in the proof it is essential to carefully follow the behavior of factors depending on p and q.

P r o o f o f (3.15). We have

$$I_3 = \mathbf{E}\,\mathbb{I}_2 \varphi'(PN + \overline{X})\,\overline{X}' = \mathbf{E}\,\mathbb{I}_2 \int_{\mathbf{R}^d} \varphi'(Pu + \overline{X})\,\overline{X}'\eta(u)\,du,$$

where η is the standard normal density (see (1.2)). We apply the following integration by parts formula

$$\int_{\mathbf{R}^d} f'(u)\,h\,g(u)\,du = -\int_{\mathbf{R}^d} f(u)\,g'(u)\,h\,du \tag{3.16}$$

with $f(u) = \varphi(Pu + \overline{X})$, $h = Q\overline{X}'$, $g(u) = \eta(u)$. We obtain

$$I_3 = -\mathbf{E}\,\mathbb{I}_2 \int_{\mathbf{R}^d} \varphi(Pu + \overline{X})\,\eta'(u)\,Q\overline{X}'\,du.$$

Changing the variable $u = x - Q\overline{X}$ and using the Taylor series (2.1) with

$$s = 1, \quad f(x) = \eta'(x)\,Q\overline{X}', \quad h = -Q\overline{X},$$

we derive

$$I_3 = I_4 + I_5 + I_6$$

with

$$I_4 = -\mathbf{E}\,\mathbb{I}_2 \int_{\mathbf{R}^d} \varphi(Px)\,\eta'(x)\,Q\overline{X}'\,dx, \tag{3.17}$$

$$I_5 = \mathbf{E}\,\mathbb{I}_2 \int_{\mathbf{R}^d} \varphi(Px)\,\eta''(x)\,Q\overline{X}'Q\,\overline{X}\,dx, \tag{3.18}$$

$$I_6 = -\mathbf{E}\,\mathbb{I}_2(1-\tau) \int_{\mathbf{R}^d} \varphi(x)\,\eta'''(x - \tau Q\overline{X})\,Q\overline{X}'(Q\overline{X})^2\,dx. \tag{3.19}$$

We note that the expression under the sign of integral in (3.17) (respectively, in (3.18)) is a linear function of $\overline{X}'$ (respectively, a bilinear function of variables $\overline{X}'$ and $\overline{X}$). Hence, (3.5) ensures that $I_4 = I_5 = 0$. Therefore, in order to complete the proof of (3.15) it remains to show that $|I_6| \ll \beta_k$. Estimating $|\mathbb{I}_2| \leqslant 1$ and $|\varphi| \leqslant 1$, and changing the variable, we have

$$|I_6| \ll \mathbf{E} \int_{\mathbf{R}^d} |\eta'''(u)\,Q\overline{X}'(Q\overline{X})^2|\,du. \tag{3.20}$$

It is easy to check that

$$\eta'''(u)\,w^2 g = \eta(u)\Big(2\langle w,g\rangle\,\langle u,w\rangle + \langle w,w\rangle\langle u,g\rangle - \langle u,g\rangle\,\langle u,w\rangle^2\Big). \tag{3.21}$$

Thus (3.20) implies

$$|I_6| \ll \mathbf{E}|\langle w,g\rangle\,\langle N,w\rangle| + \mathbf{E}|\langle w,w\rangle\,\langle N,g\rangle| + \mathbf{E}|\langle N,g\rangle|\,\langle N,w\rangle^2 \tag{3.22}$$

with $g = Q\overline{X}'$ and $w = Q\overline{X}$. The random variable $\langle N, x\rangle$ is a normal random variable with mean zero and variance $|x|^2$. In particular,

$$\mathbf{E}\langle N,x\rangle^2 = |x|^2, \quad \mathbf{E}|\langle N,x\rangle| \ll |x|, \quad \mathbf{E}|\langle N,x\rangle|^4 \ll |x|^4.$$

Using these inequalities, the bound (3.22) implies $|I_6| \ll \mathbf{E}|w|^2|g|$. Estimating

$$\|Q\| \ll 1, \quad |g| \leqslant \|Q\|\,|\overline{X}'| \ll |X| + |Y|, \quad |w| \ll |\overline{X}| \leqslant |X| + |Y|, \tag{3.23}$$

we derive $|I_6| \ll \mathbf{E}\,|X|^3 + \mathbf{E}\,|Y|^3$. It is well known that $\mathbf{E}\,|Y|^3 \ll \mathbf{E}\,|X|^3$ (see [7]). Hence $|I_6| \ll \mathbf{E}\,|X|^3 = \beta_k$, which completes the proof of (3.15).

Proof of (3.13). Using $\overline{X}' = -qX + pY$, we can split

$$I_1 = J_1 + J_2, \quad J_1 = -\mathbf{E}\,\mathbb{I}_1 q\varphi'(\overline{U} + \overline{X})\,X, \quad J_2 = \mathbf{E}\,\mathbb{I}_1 p\varphi'(\overline{U} + \overline{X})\,Y. \tag{3.24}$$

Write $\vartheta = 1/\sqrt{2}$, $T = pU + \vartheta qPN$ and

$$J = \sqrt{2}\mathbf{E}\,\mathbb{I}_1 p \int_{\mathbf{R}^d} \varphi'(T + \vartheta qPu)\,X\,\eta'(u)\,QX\,du.$$

In view of (3.24), to prove (3.13) it suffices to show that

$$J_1 = J + J_5, \qquad |J_5| \ll \varepsilon^{-2}(M\beta + b_d\varepsilon)\,\beta_k \sin\gamma, \tag{3.25}$$

and

$$J_2 = -J + J_{12}, \qquad |J_{12}| \ll \varepsilon^{-2}(M\beta + b_d\varepsilon)\,\beta_k \sin\gamma. \tag{3.26}$$

Let us prove (3.25). Note that QV is a standard normal vector. Using the relation $QV \overset{\mathcal{D}}{=} \vartheta N + \vartheta N_1$, we can write

$$\overline{U} + \overline{X} \overset{\mathcal{D}}{=} pU + \vartheta qPN + \vartheta qPN_1 + pX + qY, \qquad \vartheta = \frac{1}{\sqrt{2}}.$$

Recalling that $T = pU + \vartheta qPN$, using the standard normal density η, and changing the variable $y = u - (\vartheta qP)^{-1}(pX + qY)$, we obtain

$$\begin{aligned} J_1 &= -\mathbf{E}\,\mathbb{I}_1 q\varphi'(\overline{U} + \overline{X})\,X = -\mathbf{E}\,\mathbb{I}_1 q \int_{\mathbf{R}^d} \varphi'(T + \vartheta qPy + pX + qY)\,X\,\eta(y)\,dy \\ &= -\mathbf{E}\,\mathbb{I}_1 q \int_{\mathbf{R}^d} \varphi'(T + \vartheta qPu)\,X\,\eta(u + v + w)\,du \end{aligned} \tag{3.27}$$

with $v = -(\vartheta qP)^{-1}qY$ and $w = -(\vartheta qP)^{-1}pX$.

For sufficiently smooth functions $f\colon \mathbf{R}^d \to \mathbf{R}$, we use the following simple Taylor type expansion:

$$f(u + v + w) = f(u + v) + \mathbf{E}f'(u)\,w + \mathbf{E}f''(u + \tau_1 v + \tau_1\tau w)\,w(v + \tau w). \tag{3.28}$$

To see that (3.28) holds, expand first $f(s + w) = f(s) + \mathbf{E}f'(s + \tau w)\,w$ with $s = u + v$, and then again in powers of $v + \tau w$.

Let us return to the proof of (3.25). We expand the standard normal density η using (3.28). Then (3.27) yields

$$J_1 = J_3 + J_4 + J_5$$

with

$$\begin{aligned} J_3 &= -\mathbf{E}\,\mathbb{I}_1 q \int_{\mathbf{R}^d} \varphi'(T + \vartheta qPu)\,X\,\eta(u)\,du, \\ J_4 &= -\mathbf{E}\,\mathbb{I}_1 q \int_{\mathbf{R}^d} \varphi'(T + \vartheta qPu)\,X\,\eta'(u)\,w\,du, \\ J_5 &= -\mathbf{E}\,\mathbb{I}_1 q \int_{\mathbf{R}^d} \varphi'(T + \vartheta qPu)\,X\,\eta''(u + \tau_1 v + \tau_1\tau w)\,w(v + \tau w)\,du. \end{aligned} \tag{3.29}$$

We note that $J_3 = 0$. Indeed, $T = pU + \vartheta qPN$ is independent of X and the expression under the sign of integral in (3.29) is a linear function of X. Therefore $\mathbf{E}X = 0$ implies that $J_3 = 0$. Furthermore, recalling that $w = -(\vartheta qP)^{-1}pX$, we note that $J_4 = J$. To conclude the proof of (3.25), it remains to obtain the bound for $|J_5|$. This bound follows provided that we check that

$$|J_5| \ll J_6 \varepsilon^{-2} \beta_k \sin\gamma, \tag{3.30}$$

$$J_6 \ll M\beta + b_d \varepsilon \tag{3.31}$$

with

$$J_6 = \sup_* \mathbf{P}\{T \in A^\varepsilon \setminus A\}, \qquad T = pU + \vartheta q P N, \quad \vartheta = \frac{1}{\sqrt{2}},$$

where $\sup_*$ is taken over all $A \in \mathscr{A}$ and all nonrandom $0 < p,\, q < 1$ such that $p^2 + q^2 = 1$.

Let us prove (3.30). We apply Lemma 2.3. Recall that the function $\varphi' = \varphi'_{\varepsilon,A}$ satisfies the Lipschitz condition

$$|\varphi'(y+z)\,u - \varphi'(y)\,u| \leqslant 8\varepsilon^{-2}|z|\,|u|$$

(see (2.2)) and vanishes outside the set $A^\varepsilon \setminus A$. Therefore the function

$$f(u) = \varphi'(pU + \vartheta q P N + \vartheta q P u)\, X$$

is a Lipschitz function with constant $8\varepsilon^{-2}|X|\vartheta q\|P\| \ll q\varepsilon^{-2}|X|$ since $\|P\| \leqslant 1$ (to prove $\|P\| \leqslant 1$ note that $P^2 + \operatorname{cov} X = \mathbb{I}$ and that both operators P^2 and $\operatorname{cov} X$ are nonnegative). Therefore for estimation of J_5 we can use Lemma 2.3 with $h = v + \tau w$. We derive

$$|J_5| \ll \varepsilon^{-2}\mathbf{E}\,\mathbb{I}_1 q^2 |h|\,|X| \int_{\mathbf{R}^d} \mathbb{I}\{T + \vartheta q P u \in A^\varepsilon \setminus A\}\, |\eta'(u + \tau_1 h)\, w|\, du. \tag{3.32}$$

Recalling that $v = -(\vartheta q P)^{-1} q Y$ and $w = -(\vartheta q P)^{-1} p X$, and that $\|P^{-1}\| = \|Q\| \ll 1$, we have

$$|h| = |v + \tau w| \leqslant |v| + |w| \ll |Y| + pq^{-1}|X| \ll q^{-1}(|Y| + |X|)$$

since $0 \leqslant p,\, q \leqslant 1$. The random vectors U and N in $T = pU + \vartheta q P N$ are independent of all other random vectors and variables under the sign of expectation in (3.32). Let $\mathbf{P}_*$ stand for the conditional probability given all other vectors and variables except U and N. Then (3.32) implies

$$|J_5| \ll \varepsilon^{-2}\mathbf{E}\,\mathbb{I}_1 q(|X|^2 + |X|\,|Y|) \int_{\mathbf{R}^d} R|\eta'(u + \tau_1 h)\, w|\, du$$

with

$$\begin{aligned} R &= \mathbf{P}_*\{T + \vartheta q P u \in A^\varepsilon \setminus A\} \leqslant \sup_{z \in \mathbf{R}^d} \mathbf{P}_*\{pU + \vartheta q P N + z \in A^\varepsilon \setminus A\} \\ &= \sup_{z \in \mathbf{R}^d} \mathbf{P}_*\{pU + \vartheta q P N \in (A - z)^\varepsilon \setminus (A - z)\} \leqslant J_6. \end{aligned}$$

Collecting the bounds, we get

$$|J_5| \ll J_6 \varepsilon^{-2}\mathbf{E}\,\mathbb{I}_1 q\,(|X||Y| + |X|^2)\, D \tag{3.33}$$

with

$$D = \int_{\mathbf{R}^d} |\eta'(u + \tau_1 v + \tau_1 \tau w)\, w|\, du = \int_{\mathbf{R}^d} |\eta'(u)\, w|\, du = \int_{\mathbf{R}^d} |\langle u, w\rangle|\, \eta(u)\, du = \mathbf{E}|\langle N, w\rangle|.$$

The random variable $\langle N, w\rangle$ is a Gaussian random variable with mean zero and variance $|w|^2$. Hence $\mathbf{E}|\langle N, w\rangle| \ll |w| \ll pq^{-1}|X|$. Using in addition

$$\mathbf{E}\,|X|^2|Y| \leqslant \mathbf{E}\,|X|^2(\mathbf{E}\,|Y|^2)^{1/2} = (\mathbf{E}\,|X|^2)^{3/2} \leqslant \mathbf{E}\,|X|^3 = \beta_k,$$

the bound (3.33) yields

$$|J_5| \ll \varepsilon^{-2}\beta_k J_6 \mathbf{E}\,\mathbb{I}_1 p.$$

It is easy to check that $\mathbf{E}\,\mathbb{I}_1 p = (2/\pi)\int_0^\gamma \cos\alpha\, d\alpha \leqslant \sin\gamma$, and (3.30) follows.

To conclude the estimation of $|J_5|$ we have to prove (3.31). Write $W = pV + \vartheta q P N$. We have $J_6 \leqslant J_7 + J_8$ with

$$J_7 = \sup_* |\mathbf{P}\{T \in A^\varepsilon \setminus A\} - \mathbf{P}\{W \in A^\varepsilon \setminus A\}|, \qquad J_8 = \sup_* \mathbf{P}\{W \in A^\varepsilon \setminus A\}.$$

To prove (3.31) it suffices to check that $J_7 \ll M\beta$ and $J_8 \ll b_d\varepsilon$. Recall that now p and q are nonrandom. Let $\mathbf{P}^*$ stand for the conditional probability given N. Using that $\mathscr{A}$ is invariant under shifting and rescaling, we have

$$\begin{aligned} J_7 &= \sup_* \Big|\mathbf{E}(\mathbf{P}^*\{T \in A^\varepsilon \setminus A\} - \mathbf{P}^*\{W \in A^\varepsilon \setminus A\})\Big| \\ &\leqslant \mathbf{E}\sup_* \Big|\mathbf{P}^*\{pU + \vartheta qPN \in A^\varepsilon \setminus A\} - \mathbf{P}^*\{pV + \vartheta qPN \in A^\varepsilon \setminus A\}\Big| \\ &\leqslant \sup_* \Big|\mathbf{P}\{pU \in A^\varepsilon \setminus A\} - \mathbf{P}\{pV \in A^\varepsilon \setminus A\}\Big| \\ &\leqslant 2\sup_* \Big|\mathbf{P}\{pU \in A\} - \mathbf{P}\{pV \in A\}\Big| = 2J_9 \end{aligned}$$

with $J_9 \stackrel{\text{def}}{=} \sup_{A\in\mathscr{A}} |\mathbf{P}\{U \in A\} - \mathbf{P}\{V \in A\}|$. Recall (see (3.2)) that U (respectively, V) is a sum of $n-1$ independent random (respectively, Gaussian) vectors. Using $\operatorname{cov} U = \operatorname{cov} Y = P$, the induction assumption and $P^{-1} = Q$, $\|Q\| \ll 1$, we have

$$J_9 \leqslant M \sum_{m\neq k,\, 1\leqslant m\leqslant n} \mathbf{E}|QX_m|^3 \ll M \sum_{1\leqslant m\leqslant n} \mathbf{E}\,|X_m|^3 = M\beta, \tag{3.34}$$

which proves the desired bound $J_7 \ll M\beta$.

Let us estimate J_8. Since now p and q are nonrandom, we have

$$W = pV + \vartheta qPN \stackrel{\mathscr{D}}{=} \vartheta pPN_1 + \vartheta pPN_2 + \vartheta qPN \stackrel{\mathscr{D}}{=} \vartheta pPN_1 + \vartheta PN.$$

Hence, using the independency of N and N_1, estimating $\|(\vartheta P)^{-1}\| \leqslant c$ with some absolute constant c, we get

$$\begin{aligned} J_8 &= \sup_* \mathbf{P}\{\vartheta pPN_1 + \vartheta PN \in A^\varepsilon \setminus A\} \leqslant \sup_{A\in\mathscr{A}} \mathbf{P}\{\vartheta PN \in A^\varepsilon \setminus A\} \\ &\leqslant \sup_{B\in\mathscr{A}} \mathbf{P}\{N \in B^{c\varepsilon} \setminus B\} \ll b_d\varepsilon, \end{aligned}$$

which proves (3.31) and concludes the estimation of J_1.

Let us prove (3.26). Just exchanging the roles of X and Y, the proof is similar to that one in the case of (3.25). Instead of (3.27) we get

$$J_2 = \mathbf{E}\,\mathbb{I}_1 p \int_{\mathbf{R}^d} \varphi'(T + \vartheta qPu)\, Y\, \eta(u+v+w)\, du$$

with $w = -(\vartheta qP)^{-1}qY$ and $v = -(\vartheta qP)^{-1}pX$ (note that now v and w differ from those ones in (3.27)). Using the Taylor type expansion (3.28) we get $J_2 = J_{10} + J_{11} + J_{12}$. The same arguments as proving $J_3 = 0$ show that $J_{10} = 0$. Furthermore,

$$J_{11} = -\mathbf{E}\,\mathbb{I}_1 p \int_{\mathbf{R}^d} \varphi'(T + \vartheta qPu)\, Y\, \eta'(u)((\vartheta qP)^{-1}qY)\, du, \tag{3.35}$$

and $J_{11} = -J$. Indeed, $T = pU + \vartheta qPN$ is independent of X and Y and the expression under the sign of integral in (3.35) is a bilinear function of Y. Therefore, using $\operatorname{cov} X = \operatorname{cov} Y$ we can replace Y by X, which yields $J_{11} = -J$. Estimation of J_{12} is similar to that of J_5. A slight difference is that now we have to use the estimate $\mathbf{E}\,|Y|^3 + \mathbf{E}\,|Y|^2\mathbf{E}\,|X| \ll \beta_k$, which follows from Hölder's inequality and relations $\mathbf{E}\,|Y|^3 \ll \mathbf{E}\,|X|^3$ (see [7]) and $\mathbf{E}\,|Y|^2 = \mathbf{E}\,|X|^2$.

P r o o f o f (3.14). We have to estimate $|I_2 - I_3|$ with I_2 and I_3 given by (3.11) and (3.12) respectively. For given α, we have $PN \stackrel{\mathscr{D}}{=} pPN + qV$ since $p^2 + q^2 = 1$. Recall that $\overline{U} = pU + qV$. Therefore

$$I_2 = \mathbf{E}\,\mathbb{I}_2\varphi'(pU + qV + \overline{X})\,\overline{X}', \qquad I_3 = \mathbf{E}\,\mathbb{I}_2\varphi'(pPN + qV + \overline{X})\,\overline{X}'.$$

Using the standard normal density and applying integration by parts formula (3.16), we can write (cf. the proof of (3.15), where a similar representation for I_3 is obtained)

$$\begin{aligned} I_2 &= -\mathbf{E}\,\mathbb{I}_2 \int_{\mathbf{R}^d} \varphi(pU + qPu + \overline{X})\, \eta'(u)\, Q\overline{X}'\, du, \\ I_3 &= -\mathbf{E}\,\mathbb{I}_2 \int_{\mathbf{R}^d} \varphi(pPN + qPu + \overline{X})\, \eta'(u)\, Q\overline{X}'\, du. \end{aligned} \tag{3.36}$$

Note that we can replace PN by V in (3.36) since $PN \stackrel{\mathcal{D}}{=} V$. Changing the variables $u = x - q^{-1}Q\overline{X}$ and using $\overline{X}' = -qX + pY$, we derive

$$I_2 - I_3 = K_1 - K_2$$

with

$$\begin{aligned} K_1 &= -\mathbf{E}\,\mathbb{I}_2 \int_{\mathbf{R}^d} \delta\eta'\Big(u - \frac{Q\overline{X}}{q}\Big)\, QX\, du, \\ K_2 &= -\mathbf{E}\,\mathbb{I}_2 pq^{-1} \int_{\mathbf{R}^d} \delta\eta'\Big(u - \frac{Q\overline{X}}{q}\Big)\, QY\, du, \\ \delta &= \mathbf{E}_p\varphi(pU + qPu) - \mathbf{E}_p\varphi(pV + qPu), \end{aligned}$$

where $\mathbf{E}_p$ stands for the conditional expectation given p and q. Write $-Q\overline{X}/q = v + w$, $w = -QpX/q$, $v = -QY$ and $-Q\overline{X}/q = v_0 + w_0$, $v_0 = -QpX/q$, $w_0 = -QY$. We expand the standard normal density using the Taylor type expansion (3.28), replacing u and w by u_0 and w_0 in the case of K_2. As in the proof of (3.13) terms which are linear (respectively, bilinear) in X and Y vanish (respectively, cancel). We get $|K_1 - K_2| \ll K_3 + K_4$ with

$$\begin{aligned} K_3 &= \mathbf{E}\,\mathbb{I}_2 \int_{\mathbf{R}^d} |\delta| \,\Big|\eta'''(u + \tau_1 v + \tau_1\tau w)\, QXw(v + \tau w)\Big|\, du, \\ K_4 &= \mathbf{E}\,\mathbb{I}_2 pq^{-1} \int_{\mathbf{R}^d} |\delta| \,\Big|\eta'''(u + \tau_1 v_0 + \tau_1\tau w_0)\, QYw_0(v_0 + \tau w_0)\Big|\, du. \end{aligned}$$

Below we prove that $|\delta| \ll M\beta$. Using this bound and changing the variables, we get

$$K_3 \ll M\beta\,\mathbf{E}\,\mathbb{I}_2 \int_{\mathbf{R}^d} \Big|\eta'''(u)\, QXw(v + \tau w)\Big|\, du, \tag{3.37}$$

$$K_4 \ll M\beta\,\mathbf{E}\,\mathbb{I}_2 pq^{-1} \int_{\mathbf{R}^d} \Big|\eta'''(u)\, QYw_0(v_0 + \tau w_0)\Big|\, du. \tag{3.38}$$

Applying the identity (3.21) to the derivatives under the sign of integrals in (3.37) and (3.38), and estimating similarly to (3.22)–(3.23), we arrive at

$$K_3 \ll M\beta\,\mathbf{E}\,\mathbb{I}_2 pq^{-2}\mathbf{E}\,|X|^2(|X| + |Y|), \tag{3.39}$$

$$K_4 \ll M\beta\,\mathbf{E}\,\mathbb{I}_2 pq^{-2}\mathbf{E}\,|Y|^2(|X| + |Y|). \tag{3.40}$$

It is clear that

$$\mathbf{E}\,\mathbb{I}_2 pq^{-2} \leqslant \int_\gamma^{\pi/2} \frac{\cos\alpha}{(\sin\alpha)^2}\, d\alpha \ll \frac{1}{\sin\gamma}.$$

Earlier (see the proof of (3.13) or (3.15)) we estimated moments of combinations of $|X|$ and $|Y|$ from above by β_k. Hence, (3.39) and (3.40) lead to $K_3 + K_4 \ll M\beta\beta_k/\sin\gamma$, which completes the proof of (3.14) since $|I_2 - I_3| \ll K_3 + K_4$.

Proof of $|\delta| \ll M\beta$. Recall (see Lemma 2.2) that $\varphi(x) = \psi(\rho(x)/\varepsilon)$, where $\rho(x)$ is the distance between x and A, and $\psi\colon \mathbf{R} \to \mathbf{R}$ is a continuously differentiable nonnegative nonincreasing function such that $\int_{\mathbf{R}} |\psi'(t)|\, dt = 1$. Therefore, integrating by parts, using that neighborhoods of convex sets are again convex sets, and that the class $\mathscr{A}$ of convex sets is invariant with respect to linear symmetric transformations, we have

$$\begin{aligned} \delta &\leqslant \sup_{z \in \mathbf{R}^d} \Big|\mathbf{E}_p\psi\Big(\frac{\rho(pU + z)}{\varepsilon}\Big) - \mathbf{E}_p\psi\Big(\frac{\rho(pU + z)}{\varepsilon}\Big)\Big| \\ &= \sup_{z \in \mathbf{R}^d} \Big|\int_{\mathbf{R}} \psi(t)\, d\Big(\mathbf{P}_p\Big\{\frac{\rho(pU + z)}{\varepsilon} \leqslant t\Big\} - \mathbf{P}_p\Big\{\frac{\rho(pV + z)}{\varepsilon} \leqslant t\Big\}\Big)\Big| \\ &\leqslant \sup_{z \in \mathbf{R}^d} \int_{\mathbf{R}} |\psi'(t)|\, |\mathbf{P}_p\{\rho(pU + z) \leqslant t\varepsilon\} - \mathbf{P}_p\{\rho(pV + z) \leqslant t\varepsilon\}|\, dt \\ &\leqslant \sup_{A \in \mathscr{A}} |\mathbf{P}_p\{pU + z \in A\} - \mathbf{P}_p\{pV + z \in A\}| \\ &= \sup_{A \in \mathscr{A}} |\mathbf{P}\{U \in A\} - \mathbf{P}\{V \in A\}| \leqslant M\beta. \end{aligned} \tag{3.41}$$

The last inequality in (3.41) is established earlier as (3.34).

REFERENCES

1. *Ball K.* The reverse isoperimetric problem for Gaussian measure. — Discrete Comput. Geom., 1993, v. 10, №4, p. 411–420.
2. *Bentkus V.* On normal approximations, approximations of semigroups of operators, and approximations by accompanying laws. — Lithuanian Math. J. (to appear); www.mathematik.uni-bielefeld.de/fgweb/
3. *Bentkus V.* On the dependence of the Berry–Esseen bound on dimension. — J. Statist. Planning and Inference, 2003, v. 113, №2, p. 385–402.
4. *Бенткус В. Ю.* О зависимости оценки Берри–Эссеена от размерности. — Литов. матем. сб., 1986, т. 26, №2, с. 205–210.
5. *Bhattacharya R. N.*, *Ranga Rao R.* Normal Approximation and Asymptotic Expansions. Melbourne, FL: Krieger, 1986, 291 p.
6. *Götze F.* On the rate of convergence in the multivariate CLT. — Ann. Probab., 1991, v. 19, №2, p. 724–739.
7. *Нагаев С. В.* О вероятностных и моментных неравенствах для зависимых случайных величин. — Теория вероятн. и ее примен., 2000, т. 45, в. 1, с. 194–202.
8. *Nagaev S. V.* An estimate of the remainder term in the multidimensional central limit theorem, — Lecture Notes in Math., 1976, v. 550, p. 419–438.
9. *Nazarov F.* On the asymptotic behavior of the maximal perimeter of convex set in $\mathbf{R}^n$ with respect to the Gaussian measure. — A manuscript, 2001, p. 1–5.
10. *Nazarov F.* On the maximal perimeter of a convex set in $\mathbf{R}^n$ with respect to the Gaussian measure II (anisotropic case). — A manuscript, 2001, p. 1–11.
11. *Paulauskas V.*, *Račkauskas A.* Approximation Theory in the Central Limit Theorem. Exact Results in Banach Spaces. Dordrecht: Kluwer, 1989, 156 p.
12. *Sazonov V. V.* Normal Approximations — Some Recent Advances. Berlin–New York: Springer, 1981, 105 p.
13. *Senatov V. V.* Normal Approximations: New Results, Methods and Problems. Utrecht: VSP, 1998, 363 p.
14. *Сенатов В. В.* Одна оценка метрики Леви–Прохорова. — Теория вероятн. и ее примен., 1984, т. 29, в. 1, с. 108–113.
15. *Сенатов В. В.* Несколько оценок скорости сходимости в многомерной предельной теореме. — Докл. АН СССР, 1980, т. 254, №4, с. 809–812.

Поступила в редакцию
18.I.2004