Theory of Stochastic Processes
Vol. 22 (38), no. 2, 2017, pp. 79–85

GEORGII I. ZELENOV

# ON DISTANCES BETWEEN DISTRIBUTIONS OF POLYNOMIALS

We estimate total variation distances between distributions of polynomials via $L^2$-norms.

## 1. Introduction

Let $f, g\colon \mathbb{R}^n \to \mathbb{R}$ be two polynomials of degree at most $d$ and let $\gamma_n$ be the standard Gaussian measure on $\mathbb{R}^n$. We study estimates of the type

$$\|\gamma_n \circ f^{-1} - \gamma_n \circ g^{-1}\|_{\mathrm{TV}} \le C(d, \|f\|, \|g\|)\|f - g\|_2^{1/d}, \tag{1.1}$$

where the number $C$ may depend on the degree $d$ of our polynomials and some (semi-)norm $\|\cdot\|$ on polynomials, but does not depend on the number of variables $n$. The latter requirement is crucial: it enables one to consider the case of measurable polynomials on an infinite-dimensional space with a Gaussian measure, in particular, the case of multiple stochastic integrals with respect to the Wiener process. About this circle of problems, we refer to the recent papers [2], [4], [5], [3], [9], [11], and [13].

Important results of this type are presented in [10, Theorem 1] and [9, equation (2) and Section 3.1], where the case of multiple stochastic integrals is considered, but the proofs can be easily applied to general polynomials; they use Fubini's theorem (or conditional measures) to derive the general case from the one-dimensional case. However, these proofs are based on the following nice theorem due to G.V. Martynova.

**Theorem 1.1.** *For each $d \in \mathbb{N}$, there is a number $c(d)$ such that for every pair of polynomials*

$$f = \sum_{k=0}^{d} a_k x^k, \quad g = \sum_{k=0}^{d} b_k x^k$$

*of degree $d$ on the real line one has*

$$\|\gamma_1 \circ f^{-1} - \gamma_1 \circ g^{-1}\|_{\mathrm{TV}} \le c(d)C(a_d)\max_k |a_k - b_k|, \tag{1.2}$$

*where $C(a_d)$ is a number depending only on $a_d$.*

The proof of this theorem, along with a number of other useful results, is contained in Martynova's PhD thesis [12] (see Theorem I.3), which unfortunately remained unpublished, but is available in some libraries in Saint Petersburg and Moscow. The aim of our paper is to provide an alternative estimate of this type and its multidimensional version (see Theorem 3.1 below). The power $1/d$ at the $L^2$-norm is the same as the one obtained by Martynova and Davydov [10] (see also [9]), and this power is best possible (a bound with an additional logarithmic factor was obtained in [3]), but we also obtain some new information about the constant in this estimate. The proof of our theorem is different from that of [12, Theorem I.3] and is based on the Malliavin calculus (see, e.g., [7] and [14]) and the Carbery–Wright inequality [8]. Moreover, in Theorem 3.5 we

2000 *Mathematics Subject Classification.* Primary 60E05, 60E015; Secondary 28C20, 60F99.

*Key words and phrases.* Distribution of a polynomial, total variation norm, Gaussian measure, logarithmically concave measure.

give certain generalizations of Theorems I.1 and I.2 from [12] to the case of log-concave measures.

## 2. Definitions and notation

Let $\gamma_n$ denote the standard Gaussian measure on $\mathbb{R}^n$, i.e., the measure on $\mathbb{R}^n$ with density

$$(2\pi)^{-n/2}\exp(-|x|^2/2)$$

with respect to Lebesgue measure.

The image of a measure $\mu$ on a measurable space under a measurable mapping $f$ with values in $\mathbb{R}^k$ is denoted by the symbol $\mu\circ f^{-1}$ and defined by the formula

$$\mu\circ f^{-1}(B)=\mu(f^{-1}(B))\quad \text{for Borel sets } B\subset\mathbb{R}^k.$$

We set $\|\varphi\|_\infty=\sup_x|\varphi(x)|$ for any bounded function $\varphi$ on any set.

The total variation distance $d_{\mathrm{TV}}(\mu,\nu)$ between two Borel measures $\mu,\nu$ on $\mathbb{R}^k$ is generated by the norm

$$\|\mu-\nu\|_{\mathrm{TV}}:=\sup\biggl\{\int\varphi\,d(\mu-\nu),\ \varphi\in C_b^\infty(\mathbb{R}^k),\ \|\varphi\|_\infty\le 1\biggr\},$$

where here and throughout we omit indicating the domain of integration in case of integrating over the whole space.

We need the following important inequality proved by Carbery and Wright [8]. There is an absolute constant $c$ such that, for every Gaussian measure (more generally, for every logarithmically concave measure) $\gamma$ on $\mathbb{R}^n$ and for every polynomial $f$ of degree $d$, one has

$$\gamma(|f|\le t)\biggl(\int|f|\,d\gamma\biggr)^{1/d}\le cdt^{1/d},\quad t\ge 0. \tag{2.1}$$

Some generalizations of this estimate to the case of measurable polynomials on infinite-dimensional locally convex spaces are discussed in [1] and [2].

**Lemma 2.1.** *For every $d\in\mathbb{N}$ there is a number $c(d)$ such that for each $n$ and each polynomial $f\colon\mathbb{R}^n\to\mathbb{R}$ of degree at most $d$ one has*

$$\int\frac{1}{f^2+\varepsilon}\,d\gamma_n\le\varepsilon^{-1+1/(2d)}c(d)\biggl(\int f^2\,d\gamma_n\biggr)^{-1/(2d)} \tag{2.2}$$

*Proof.* To simplify notation we set $\beta=1/2d$. Note that $f^2\ge 0$. Hence

$$\int(f^2+\varepsilon)^{-1}\,d\gamma_n=\int_0^{\varepsilon^{-1}}\gamma_n\bigl((f^2+\varepsilon)^{-1}\ge t\bigr)\,dt=\int_0^\infty(s+\varepsilon)^{-2}\gamma_n\bigl(f^2\le s\bigr)\,ds.$$

The Carbery–Wright inequality (2.1) yields

$$\begin{aligned}\int_0^\infty(s+\varepsilon)^{-2}\gamma_n\bigl(f^2\le s\bigr)\,ds&\le c(2d)\biggl(\int f^2\,d\gamma_n\biggr)^{-\beta}\int_0^\infty(s+\varepsilon)^{-2}s^\beta\,ds\\&=\varepsilon^{-1+\beta}2cd\biggl(\int f^2\,d\gamma_n\biggr)^{-\beta}\int_0^\infty(s+1)^{-2}s^\beta\,ds.\end{aligned}$$

Letting

$$c(d):=2cd\int_0^\infty(s+1)^{-2}s^{(1/2d)}\,ds$$

we arrive at the desired estimate. □

For a polynomial $f\colon \mathbb{R}^n \to \mathbb{R}$ of degree at most $d$ one has

$$\int |\nabla f|^2 \, d\gamma_n \le c(d) \int f^2 \, d\gamma_n, \tag{2.3}$$

where $c(d)$ does not depend on $n$. This fact follows from the equivalence of all Sobolev norms and all $L^p$-norms on the space of measurable polynomials of degree $d$ (see [6, Example 5.3.4] or [2]).

**Remark 2.2.** Consider the vector space $P_d(\mathbb{R})$ of all polynomials of degree at most $d$ in one variable. This is a $d+1$-dimensional vector space. Therefore, any two norms on $P_d(\mathbb{R})$ are equivalent. This enables us to replace the norm $\max_k |a_k - b_k|$ by the standard $L_2$-norm with respect to the measure $\gamma_1$.

## 3. Main results

Let us set

$$\|\nabla f\|_*^2 := \sup_{|e|=1} \int_{\mathbb{R}^n} |\partial_e f|^2 \, d\gamma_n. \tag{3.1}$$

Our first theorem is a multidimensional analog of [12, Chapter 3, Theorem I.3].

**Theorem 3.1.** *For each $d \in \mathbb{N}$, there is a number $c(d)$ such that, for each $n$ and every pair of polynomials $f, g\colon \mathbb{R}^n \to \mathbb{R}$ of degree at most $d$, one has*

$$\|\gamma_n \circ f^{-1} - \gamma_n \circ g^{-1}\|_{\mathrm{TV}} \le c(d)\bigl(\|\nabla g\|_*^{-1/(d-1)} + 1\bigr)\|f - g\|_2^{1/d}.$$

**Remark 3.2.** The inequality in the theorem is of the type

$$\|\gamma_n \circ f^{-1} - \gamma_n \circ g^{-1}\|_{\mathrm{TV}} \le C\|f - g\|_2^{\Theta}.$$

In our case the constant $C$ does not depend on the number of variables $n$. Thus, the result can be generalized to the infinite-dimensional case in the spirit of the recent papers [2], [3], [4], [5], and [9] (see below). On measurable polynomials on an infinite-dimensional space with a Gaussian measure $\gamma$, see [6], [7], and [2]. For example, multiple stochastic integrals can be represented as measurable polynomials with respect to the Wiener measure, see [14, Section 1.1.2], [6, Section 2.11]. In fact, the original inequality of Martynova was obtained for the study of multiple stochastic integrals. We use the norm $\|\nabla f\|_*^2$ in place of the more standard norm of gradients to avoid dependence on $n$. It is also worth noting that in [3] a weaker estimate (with an extra logarithmic factor, which is not needed as we see) has been obtained, but the proof has not employed reduction to the one-dimensional case.

*Proof.* Let us fix a function $\varphi \in C_0^\infty(\mathbb{R})$ with $\|\varphi\|_\infty \le 1$ and a unit vector $e \in \mathbb{R}^n$. Consider the function

$$\Phi(t) = \int_{-\infty}^{t} \varphi(\tau) \, d\tau.$$

Note that

$$\partial_e(\Phi(f) - \Phi(g)) = \partial_e f \varphi(f) - \partial_e g \varphi(g) = \varphi(f)(\partial_e f - \partial_e g) + (\varphi(f) - \varphi(g))\partial_e g,$$

where $\partial_e$ denotes the partial derivative along the vector $e$. Thus, for each $\varepsilon > 0$ we have

$$\begin{aligned}
&\int [\varphi(f) - \varphi(g)]\, d\gamma_n \\
&= \int [\varphi(f) - \varphi(g)] \frac{(\partial_e g)^2}{(\partial_e g)^2 + \varepsilon}\, d\gamma_n + \varepsilon \int [\varphi(f) - \varphi(g)] \frac{1}{(\partial_e g)^2 + \varepsilon}\, d\gamma_n \\
&= \int \frac{\partial_e g \partial_e (\Phi(f) - \Phi(g))}{(\partial_e g)^2 + \varepsilon}\, d\gamma_n - \int \frac{\partial_e g \cdot \varphi(f)(\partial_e f - \partial_e g)}{(\partial_e g)^2 + \varepsilon}\, d\gamma_n \\
&+ \varepsilon \int [\varphi(f) - \varphi(g)]((\partial_e g)^2 + \varepsilon)^{-1}\, d\gamma_n.
\end{aligned} \tag{3.2}$$

Let us estimate each term separately. First, let us consider the last term. Applying (2.2) to the polynomial $\partial_e g$ we obtain the bound

$$\begin{aligned}
\varepsilon \int [\varphi(f) - \varphi(g)]((\partial_e g)^2 + \varepsilon)^{-1}\, d\gamma_n &\le 2\varepsilon \int ((\partial_e g)^2 + \varepsilon)^{-1}\, d\gamma_n \\
&\le 2dc_1 \|\partial_e g\|_2^{-1/(d-1)} \varepsilon^{1/(2d-2)} \int_0^\infty (s+1)^{-2} s^{1/(2d-2)}\, ds \\
&= c_1(d) \|\partial_e g\|_2^{-1/(d-1)} \varepsilon^{1/(2d-2)}.
\end{aligned}$$

The second term in (3.2) is easily estimated as follows (here we need (2.3)):

$$\begin{aligned}
-\int \frac{\partial_e g \cdot \varphi(f)(\partial_e f - \partial_e g)}{(\partial_e g)^2 + \varepsilon}\, d\gamma_n &\le \int \frac{|\partial_e g|}{(\partial_e g)^2 + \varepsilon} |\partial_e f - \partial_e g|\, d\gamma_n \\
&\le 2^{-1} \varepsilon^{-1/2} \int |\partial_e f - \partial_e g|\, d\gamma_n \le c_2(d) \varepsilon^{-1/2} \|f - g\|_2.
\end{aligned}$$

We now estimate the first term in (3.2). Integrating by parts we obtain

$$\begin{aligned}
&\int \frac{\partial_e g \partial_e (\Phi(f) - \Phi(g))}{(\partial_e g)^2 + \varepsilon}\, d\gamma_n \\
&\quad = -\int (\Phi(f) - \Phi(g)) \left[ \frac{\partial_e^2 g - \langle x, e\rangle \partial_e g}{(\partial_e g)^2 + \varepsilon} - 2\frac{(\partial_e g)^2 \partial_e^2 g}{((\partial_e g)^2 + \varepsilon)^2} \right] d\gamma_n \\
&\quad \le 3 \int |f - g| \frac{|\partial_e^2 g|}{(\partial_e g)^2 + \varepsilon}\, d\gamma_n + 2^{-1} \varepsilon^{-1/2} \int |f(x) - g(x)|\, |\langle x, e\rangle|\, \gamma_n(dx).
\end{aligned}$$

Let us show that this sum is bounded by $c_3(d)\varepsilon^{-1/2}\|f-g\|_2$. To this end, we first consider the one-dimensional case $n = 1$. In this case $x$ and $e$ belong to $\mathbb{R}$. Therefore, $e = \pm 1$ and $\partial_e g = \pm g'(x)$, $\partial_e^2 g = g''(x)$, $|\langle x, e\rangle| = |x|$. Thus,

$$\begin{aligned}
\int |f - g| \frac{|\partial_e^2 g|}{(\partial_e g)^2 + \varepsilon}\, d\gamma_1 &= \int |f - g| \frac{|g''|}{(g')^2 + \varepsilon}\, d\gamma_1 \\
&= \int |f(x) - g(x)| \frac{|g''(x)|}{(g'(x))^2 + \varepsilon} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx.
\end{aligned}$$

Observe that there is a number $C(d)$ depending only on $d$ such that

$$\frac{|f(x) - g(x)|}{\|f - g\|_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \le C(d).$$

This follows from Remark 2.2, because

$$\|f\| = \sup_{x \in \mathbb{R}} (f(x) e^{-\frac{x^2}{2}})$$

is a norm on the space of polynomials of degree at most $d$. Thus,

$$\int |f-g|\frac{|\partial_e^2 g|}{(\partial_e g)^2+\varepsilon}\,d\gamma_1 \le C(d)\,\|f-g\|_2\int \frac{|g''(x)|}{(g'(x))^2+\varepsilon}\,dx.$$

The function $g''\colon \mathbb{R}\to\mathbb{R}$ is a polynomial of degree at most $d-2$ and it has $k\le d-2$ zeros $a_1<\ldots<a_k$. Let us consider the intervals $(a_i,a_{i+1})$ with $i=0,\ldots,k$ and

$$a_0=-\infty; \qquad a_{k+1}=+\infty.$$

On each of the $k+1\le d-1$ intervals $(a_i,a_{i+1})$ the polynomial $g''$ is of constant sign. Therefore,

$$\begin{aligned}\int_{a_i}^{a_{i+1}}\frac{|g''|}{(g')^2+\varepsilon}\,dx &= \operatorname{sign}\bigl(g''(x)\bigr)|_{x\in(a_i,a_{i+1})}\cdot\int_{a_i}^{a_{i+1}}\frac{g''}{(g')^2+\varepsilon}\,dx\\ &=\left|\int_{a_i}^{a_{i+1}}\frac{g''}{(g')^2+\varepsilon}\,dx\right|=\left|\int_{a_i}^{a_{i+1}}\frac{1}{(g')^2+\varepsilon}\,d(g'(x))\right|\le\frac{\pi}{\sqrt{\varepsilon}}.\end{aligned}$$

Taking into account the obtained estimates and the bound $k+1\le d-1$ we have

$$\begin{aligned}\int |f-g|\frac{|g''|}{(g')^2+\varepsilon}\,d\gamma_1 &\le C(d)\|f-g\|_2\int_{\mathbb{R}}\frac{|g''|}{(g')^2+\varepsilon}\,dx\\ &= C(d)\|f-g\|_2\sum_{i=0}^{k}\int_{a_i}^{a_{i+1}}\frac{|g''|}{(g')^2+\varepsilon}\,dx \le C(d)(d-1)\|f-g\|_2\frac{\pi}{\sqrt{\varepsilon}}.\end{aligned}$$

Therefore, for the first term in (3.2) we have

$$\int \frac{g'(\Phi(f)-\Phi(g))'}{(g')^2+\varepsilon}\,d\gamma_1\le \widetilde{C}(d)\|f-g\|_2\frac{\pi}{\sqrt{\varepsilon}}+\frac{1}{2\sqrt{\varepsilon}}\int |f(x)-g(x)|\,|x|\,\gamma_1(dx).$$

By the Cauchy–Bunyakovskii inequality the integral on the right is estimated by $\|f-g\|_2$. Hence in the one-dimensional case we have

$$\int \frac{g'(\Phi(f)-\Phi(g))'}{(g')^2+\varepsilon}\,d\gamma_1\le C'(d)\|f-g\|_2\frac{1}{\sqrt{\varepsilon}}$$

with some new number $C'(d)$.

We now proceed to the case $n>1$. The space $\mathbb{R}^n$ is decomposed into the sum $\mathbb{R}^n=\langle e\rangle\oplus\langle e\rangle^{\perp}$. The corresponding representation of the measure is $\gamma_n=\gamma_{n-1}\otimes\gamma_1$, where $\gamma_1$ is the standard Gaussian measure on $\langle e\rangle\simeq\mathbb{R}$ and $\gamma_{n-1}$ is the standard Gaussian measure on $\langle e\rangle^{\perp}\simeq\mathbb{R}^{n-1}$. By Fubini's theorem we have

$$\int_{\mathbb{R}^n}\frac{\partial_e g\,\partial_e(\Phi(f)-\Phi(g))}{(\partial_e g)^2+\varepsilon}\,d\gamma_n=\int_{\langle e\rangle^{\perp}}\int_{\mathbb{R}}\frac{g_x'(t)(\Phi(f_x(t))-\Phi(g_x(t)))'}{(g_x'(t))^2+\varepsilon}\,\gamma_1(dt)\gamma_{n-1}(dx),$$

where we use the notation $f_x(t):=f(x+te)$, $g_x(t):=g(x+te)$.

For any fixed $x\in\langle e\rangle^{\perp}$, the functions $f_x(t)$ and $g_x(t)$ are polynomials of variable $t$ and have degrees at most $d$. Therefore, we can apply the estimate proved in the one-dimensional case and obtain

$$\int\frac{g_x'(\Phi(f_x)-\Phi(g_x))'}{(g_x')^2+\varepsilon}\,d\gamma_1\le C'(d)\|f_x-g_x\|_2\frac{1}{\sqrt{\varepsilon}},$$

where $f_x=f_x(t)$ and $g_x=g_x(t)$ are considered as functions of one variable $t$. Now we integrate in $x\in\langle e\rangle^{\perp}$ with respect to $\gamma_{n-1}$ and find that

$$\int_{\langle e\rangle^{\perp}}\int_{\mathbb{R}}\frac{g_x'(\Phi(f_x)-\Phi(g_x))'}{(g_x')^2+\varepsilon}\,\gamma_1(dt)\,\gamma_{n-1}(dx)\le C'(d)\frac{1}{\sqrt{\varepsilon}}\int_{\langle e\rangle^{\perp}}\|f_x-g_x\|_2\,\gamma_{n-1}(dx).$$

By the Cauchy–Bunyakovskii inequality, $\|F\|_{L^1(P)} \le \|F\|_{L^2(P)}$ for any probability measure $P$. Applying this inequality to $F(x) = \|f_x - g_x\|_2$ we get

$$\int_{\mathbb{R}^n} \frac{g_x'(\Phi(f_x)-\Phi(g_x))'}{(g_x')^2+\varepsilon}\,\gamma_n(dx\,dt) = \int_{\langle e\rangle^\perp}\int_{\mathbb{R}} \frac{g_x'(\Phi(f_x)-\Phi(g_x))'}{(g_x')^2+\varepsilon}\,\gamma_1(dt)\,\gamma_{n-1}(dx)$$
$$\le c_3(d)\frac{1}{\sqrt{\varepsilon}}\biggl(\int_{\langle e\rangle^\perp}\|f_x-g_x\|_2^2\,\gamma_{n-1}(dx)\biggr)^{1/2} = c_3(d)\frac{1}{\sqrt{\varepsilon}}\|f-g\|_2.$$

The third term is estimated.

Now we can combine our estimates for all the three terms in (3.2) and get

$$\int[\varphi(f)-\varphi(g)]\,d\gamma_n \le C(d)\|f-g\|_2\frac{1}{\sqrt{\varepsilon}} + c_1(d)\|\partial_e g\|_2^{-1/(d-1)}\varepsilon^{1/(2d-2)}$$

with some new constants. Taking $\varepsilon = \|f-g\|_2^{(2d-2)/d}$ we arrive at the estimate

$$\int[\varphi(f)-\varphi(g)]\,d\gamma_n \le \Bigl(C(d) + c_1(d)\|\partial_e g\|_2^{-1/(d-1)}\Bigr)\|f-g\|_2^{1/d}.$$

Finally, we take the supremum over all smooth functions $\varphi$ with $\|\varphi\|_\infty \le 1$ and all unit vectors $e$ and complete the proof. □

There is a straightforward infinite-dimensional extension of our estimate. For simplicity we formulate it for the countable power $\gamma$ of the standard Gaussian measure on the real line defined on the space $\mathbb{R}^\infty$ of all real sequences. The Cameron–Martin space of this measure is the usual Hilbert space $H = l^2$; this space is used to define a norm on gradients. Measurable polynomials of degree $d$ are defined as $L^2$-limits of sequences of finite-dimensional polynomials of degree $d$. There is an analog of (3.1) defined by the same formula with $\gamma$ in place of $\gamma_n$ and $|e| = |e|_{l^2}$.

**Corollary 3.3.** *For all measurable polynomials $f$ and $g$ of degree $d$ on $\mathbb{R}^\infty$ we have*

$$\|\gamma\circ f^{-1} - \gamma\circ g^{-1}\|_{\mathrm{TV}} \le c(d)\bigl(\|\nabla g\|_*^{-1/(d-1)} + 1\bigr)\|f-g\|_2^{1/d}.$$

To formulate our second result, we recall the definition of a log-concave measure: a probability measure $\mu$ on $\mathbb{R}^n$ is called log-concave if for all $\lambda \in (0,1)$ and all Borel sets $A$ and $B$ one has

$$\mu(\lambda A + (1-\lambda)B) \ge \mu(A)^\lambda \mu(1-A)^{1-\lambda}.$$

An equivalent description is this: the measure $\mu$ is concentrated on an affine subspace and is given there by a density of the form $e^{-V}$ with a convex function $V$ (with values in $(-\infty,+\infty]$). Examples of log-concave measures are Gaussian measures and uniform distributions in convex sets. For all such measures the Carbery–Wright inequality (2.1) holds true. Thus, one can prove the following analog of (2.2).

**Lemma 3.4.** *For each $d \in \mathbb{N}$, there is a number $c(d)$ such that, whenever $\mu$ is a log-concave measure on the real line and $\nu$ is the measure given by a bounded density $\varrho$ with respect to $\mu$, for every polynomial $f$ of degree at most $d$ on the real line we have*

$$\int \frac{1}{f^2+\varepsilon}\,d\nu \le \varepsilon^{-1+1/(2d)}c(d)\|\varrho\|_\infty\biggl(\int f^2\,d\mu\biggr)^{-1/(2d)}. \tag{3.3}$$

This lemma enables us to prove the following result.

**Theorem 3.5.** *Let $\mu$ be an absolutely continuous log-concave measure on the real line. Suppose that $\nu$ is a probability measure given by a bounded density $\varrho$ with respect to $\mu$ such that $\varrho$ is a function of bounded variation on the whole real line. Then there is a* *number $c(d,\varrho,\mu)$, depending on $d$, $\varrho$ and $\mu$, such that for every pair of polynomials $f,g$ of degree at most $d$ one has*

$$\|\nu\circ f^{-1}-\nu\circ g^{-1}\|_{\mathrm{TV}}\le c(d,\varrho,\mu)\bigl(\|g'\|_{L^2(\mu)}^{-1/(d-1)}+1\bigr)\|f-g\|_{L^2(\nu)}^{1/d}.$$

The main steps in proof of this theorem are similar to the ones in the proof of Theorem 3.1, so we do not repeat them. The stated conditions on the measure $\nu$ come from Lemma 3.3 and the need to integrate by parts, for which it is important that the generalized derivative of $\nu$ is also a bounded measure. Note also that in this theorem we have $n=1$ and thus can use Remark 2.2 in place of inequality (2.3).

**Remark 3.6.** Theorem 3.5 is a generalization of Theorems I.1 and I.2 from Chapter 3 in Martynova's dissertation [12]. The latter theorems deal with the case where $\mu$ is the uniform distribution on some interval $[a,b]\subset\mathbb{R}$ and impose the condition that the density $\varrho$ is Lipschitz on $[a,b]$. This condition is stronger than having bounded variation.

**Acknowledgement.** This research is supported by the Russian Science Foundation (Grant 17-11-01058 at Lomonosov Moscow State University).

## References

1. L. M. Arutyunyan and E. D. Kosov, *Estimates for integral norms of polynomials on spaces with convex measures*, Matem. Sb. **206** (2015), no. 8, 3–22 (in Russian). English transl.: Sbornik Math. **206** (2015), no. 8, 1030–1048.
2. V. I. Bogachev, *Distributions of polynomials on multidimensional and infinite-dimensional spaces with measures*, Uspehi Mat. Nauk **71** (2016), no. 4, 107–154 (in Russian). English transl.: Russian Math. Surveys **71** (2016), no. 4, 703–749.
3. V. I. Bogachev, E. D. Kosov and G. I. Zelenov, *Fractional smoothness of distributions of polynomials and a fractional analog of the Hardy–Landau–Littlewood inequality*, Trans. Amer. Math. Soc. **370** (2018), no. 6, 4401–4432.
4. V. I. Bogachev and G. I. Zelenov, *On convergence in variation of weakly convergent multidimensional distributions*, Dokl. Akad. Nauk **461** (2015), no. 1, 14–17 (in Russian). English transl.: Dokl. Math. **91** (2015), no. 2, 138–141.
5. V. I. Bogachev, G. I. Zelenov and E. D. Kosov, *Membership of distributions of polynomials in the Nikolskii–Besov class*, Dokl. Akad. Nauk **469** (2016), no. 6, 651–655 (in Russian). English transl.: Dokl. Math. **94** (2016), no. 2, 453–457.
6. V. I. Bogachev, *Gaussian measures*, Amer. Math. Soc., Providence, Rhode Island, 1998.
7. V. I. Bogachev, *Differentiable measures and the Malliavin calculus*, Amer. Math. Soc., Providence, Rhode Island, 2010.
8. A. Carbery and J. Wright, *Distributional and $L_q$ norm inequalities for polynomials over convex bodies in $\mathbb{R}^n$*, Math. Research Lett. **8** (2001), no. 3, 233–248.
9. Y. A. Davydov, *On distance in total variation between image measures*, Statistics & Probability Letters **129** (2017), 393–400.
10. Y. A. Davydov and G. V. Martynova, *Limit behavior of multiple stochastic integral*, Statistics and Control of Random Processes, Nauka, Preila, Moscow, 1987, pp. 55–57 (in Russian).
11. E. D. Kosov, *Fractional smoothness of images of logarithmically concave measures under polynomials*, J. Math. Anal. Appl. **462** (2018), no. 1, 390–406.
12. G. V. Martynova, *Limit theorems for the functionals of random procceses*, Candidate of science dissertation (PhD thesis), Leningrad, 1987 (in Russian). Russian State Library (Moscow), N 1-1-5/LGU.
13. I. Nourdin and G. Poly, *Convergence in total variation on Wiener chaos*, Stochastic Process. Appl. **123** (2013), no. 2, 651–674.
14. D. Nualart, *The Malliavin calculus and related topics*, 2nd ed., Springer-Verlag, Berlin, 2006.

Faculty of Mechanics and Mathematics, Moscow State University, Moscow, 119991 Russia
*E-mail address*: zelenovyur@gmail.com