The Annals of Probability
1999, Vol. 27, No. 1, 416–431

THE SIZE OF SINGULAR COMPONENT AND SHIFT INEQUALITIES[1]

BY S. G. BOBKOV

Syktyvkar University

We study the question on how large the size of singular component of translates of product probability measures can be in terms of Fisher information. We then prove some shift inequalities.

1. Introduction. The well-known Kakutani theorem [4] asserts that any two product measures on $\mathbf{R}^\infty$ with equivalent marginal distributions are either equivalent or singular. In particular, this holds for a product measure μ^∞ and its translate μ_h^∞, defined by

$$\mu_h^\infty(A) = \mu^\infty(A + h), \qquad A \text{ is Borel in } \mathbf{R}^\infty,\ h \in \mathbf{R}^\infty.$$

In this important special situation, Feldman [3] and Shepp [8] found, respectively, simple sufficient and necessary conditions for equivalence of μ^∞ and μ_h^∞, provided $h \in l^2$ is arbitrary. Namely, the marginal distribution μ should have an absolutely continuous, a.e. positive density ρ with finite Fisher information

$$J(\mu) = \int_{-\infty}^{+\infty} \frac{\rho'(x)^2}{\rho(x)^2}\, d\mu(x),$$

where ρ' is a Radon–Nikodim derivative of ρ. The condition $J(\mu) < +\infty$ does not imply itself that ρ is a.e. positive, and, moreover, μ may have a compact support. Therefore, in this case μ^∞ and μ_h^∞ cannot be equivalent, at least for large values of $\|h\|_2 = (\sum_{i=1}^\infty h_i^2)^{1/2}$, and it is reasonable to ask about the size of a singular component of μ_h^∞ with respect to μ^∞, which we denote by $\mathrm{Sing}(\mu_h^\infty, \mu^\infty)$:

$$\mathrm{Sing}(\mu_h^\infty, \mu^\infty) = \sup\{\mu_h^\infty(A) : \mu^\infty(A) = 0\}.$$

An answer may easily be given if one uses elementary estimates

$$\mathrm{Sing}(\mu_h^\infty, \mu^\infty) \le \|\mu_h^\infty - \mu^\infty\|_{\mathrm{TV}} \le \|d_h \mu^\infty\|_{\mathrm{TV}} \le J(\mu)\|h\|_2, \tag{1.1}$$

where $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation norm, and $d_h \mu^\infty$ denotes the derivative of μ^∞ along the direction h (in the sense of Fomin–Skorokhod; cf. [1]). Consequently, for sufficiently small $\|h\|_2$, the measures μ^∞ and μ_h^∞ are still

Received April 1998.

[1]Supported in part by Russian Foundation for Fundamental Research Grant 96-01-00201 and Alexander von Humboldt Foundation.

AMS 1991 *subject classifications*. 60E15, 46G12, 60B11.

Key words and phrases. Translates of measures, shift inequalities.

not purely singular, and this observation may be applied, for example, to give a simple proof of Feldman's theorem (provided we may use Kakutani's theorem). Indeed, assume the density ρ of μ is a.e. positive and thus μ^∞ and μ_h^∞ have equivalent marginals. If $\|h\|_2 < 1/J(\mu)$, then $\operatorname{Sing}(\mu_h^\infty, \mu^\infty) < 1$, so μ_h^∞ and μ^∞ must be equivalent, by Kakutani's theorem. In the same way, since $J(\mu_{h_i}) = J(\mu)$, μ_{2h}^∞ and μ_h^∞ are equivalent, so are μ_{2h}^∞ and μ^∞. Repeating this process, the requirement $\|h\|_2 < 1/J(\mu)$ can be weakened to $\|h\|_2 < +\infty$.

While the two last inequalities in (1.1) cannot be essentially improved, the first one is rather rough, as the following statement shows.

THEOREM 1.1. *Assume μ has an absolutely continuous density with $\sigma^2 = J(\mu) < +\infty$ ($\sigma \geq 0$). Then, for all $h \in l^2$ such that $\sigma\|h\|_2 < \pi$,*

$$\operatorname{Sing}(\mu_h^\infty, \mu^\infty) \leq \sin^2\left(\frac{\sigma\|h\|_2}{2}\right). \tag{1.2}$$

Inequality (1.2) remains to hold for product measures μ on $\mathbf{R}^\infty$ with different marginals whose Fisher information is at most σ^2. As we will see, it is actually valid for many nonproduct measures on $\mathbf{R}^n$, satisfying certain correlation conditions on density. Note also that the condition $J(\mu) < +\infty$ cannot be dropped in (1.2). If, for example, μ is uniform distribution on $(0,1)$ and $h \in (0,1)^\infty$, then

$$\operatorname{Sing}(\mu_h^\infty, \mu^\infty) = 1 - \prod_{i=1}^{\infty}(1 - h_i).$$

Hence, it is possible to get $\operatorname{Sing}(\mu_h^\infty, \mu^\infty) = 1$ for every fixed value of $\|h\|_2$. Thus, the size of singular component in this case cannot be controlled in terms of $\|h\|_2$.

When the product measures μ^∞ and μ_h^∞ are equivalent, one could wonder how to measure or how to quantify their equivalence. One of the reasonable ways in this direction is to establish shift inequalities of the form

$$S_h(\mu^\infty(A)) \leq \mu_h^\infty(A) \leq R_h(\mu^\infty(A)), \tag{1.3}$$

for suitable functions R_h and S_h. For example, the optimal functions R_h and S_h in (1.3) for the canonical Gaussian measure $\mu = \gamma_1$ are given (cf. [6]) by

$$R_h(p) = \Phi(\Phi^{-1}(p) + \|h\|_2), \qquad S_h(p) = \Phi(\Phi^{-1}(p) - \|h\|_2),$$

where Φ is the distribution function of γ_1 and Φ^{-1} is the inverse of Φ. In this case, (1.3) becomes an equality for half-spaces which are orthogonal to vector h. In general, it is, however, rather difficult to compute the optimal function R_h. Nevertheless, such a shift inequality of Gaussian type holds true for a large family of probability measures μ.

THEOREM 1.2. *Let μ be a probability measure on $\mathbf{R}$. There exists $c > 0$ such that, for all Borel sets $A \subset \mathbf{R}^\infty$ of measure $\mu^\infty(A) = p$ and every $h \in l^2$, an inequality*

$$\Phi(\Phi^{-1}(p) - c\|h\|_2) \le \mu_h^\infty(A) \le \Phi(\Phi^{-1}(p) + c\|h\|_2) \tag{1.4}$$

holds true, if and only if μ has an absolutely continuous density ρ with

$$\int_{-\infty}^{+\infty} \exp\left\{\varepsilon \frac{\rho'(x)^2}{\rho(x)^2}\right\} d\mu(x) \le 2 \quad \textit{for some } \varepsilon > 0. \tag{1.5}$$

We will also observe that the optimal constants c and ε in (1.4) and (1.5) are connected with the relation

$$\frac{1}{\sqrt{6\varepsilon}} \le c \le \frac{4}{\sqrt{\varepsilon}}. \tag{1.6}$$

The conditions of Theorem 1.2 are satisfied, for example, by the two-sided exponential measure ν with density $\rho(x) = \frac{1}{2}e^{-|x|}$, $x \in \mathbf{R}$. In this case, the functions $\xi_i(x) = (\rho'(x_i)/\rho(x_i)) = -\operatorname{sign}(x_i)$, $x \in \mathbf{R}^\infty$, form an i.i.d. Bernoulli sequence on probability space $(\mathbf{R}^\infty, \nu^\infty)$, and the problem of finding the optimal function R_h in (1.3) reduces in essence to the (still open) "average" isodiametral problem on the discrete cube $(\{-1,1\}^n, P_n)$. The latter consists of finding the minimum of $\int_A \int_A \|x - y\|_2^2 \, dP_n(x)\, dP_n(y)$ among all subsets A of $\{-1,1\}^n$ of a fixed measure $P_n(A)$.

Anyway, up to the constant c, for the exponential measure as well as for those satisfying the conditions (1.5) and having a finite second moment, the left- and the right-hand sides of (1.4) cannot be improved in terms of $\|h\|_2$. Indeed, if we apply (1.3) to half-spaces $A = \{x\colon x_1 + \cdots + x_n - an \le xb\sqrt{n}\}$ with a the mean and b the variance of μ, and to the "n-dimensional" vectors $h = (1/\sqrt{n})(t, \dots, t; 0, \dots)$ with $t = \|h\|_2 > 0$, we will obtain, by the central limit theorem, that $\Phi(x + \|h\|_2/b) \le R_h(\Phi(x))$. Thus, $\sup_{\{\|h\|_2 = t\}} R_h(p) \ge \Phi(\Phi^{-1}(p) + t/b)$. A similar argument applies to S_h.

The organization of the paper is the following. Theorem 1.1 is proved in Section 3, on the basis of a certain functional form for shift inequalities (1.3), which is discussed separately in Section 2. Theorem 1.2 is proved in Section 4. Here we consider also a question on how to bound $\|\mu_h^\infty - \mu^\infty\|_{\mathrm{TV}}$ in terms of $\|h\|_2$. In Section 5, we prove an analogue of Theorem 1.2 on the appropriate form for (1.3) under a weaker assumption on the density ρ.

2. A functional form of shift inequalities. Assume we are given a family of concave continuous functions $(R_t)_{t \ge 0}$ from $[0,1]$ onto itself with the following semigroup property: for all $t, s \ge 0$ and $p \in [0,1]$,

$$R_{t+s}(p) = R_t(R_s(p)). \tag{2.1}$$

Assume also that R_0 is the identity function, and the "generator" of the family

$$\lim_{t \to 0^+} \frac{R_t(p) - p}{t} = I(p), \qquad 0 < p < 1, \tag{2.2}$$

is well defined and represents a positive function on $(0,1)$. Clearly, the function I must be concave and thus continuous on $(0,1)$ as the limit of concave functions.

LEMMA 2.1. *Let μ be an absolutely continuous probability measure on $\mathbf{R}^n$. The following properties are equivalent*:

(a) *For all $h \in \mathbf{R}^n$ and every Borel set A in $\mathbf{R}^n$*,

$$\mu_h(A) \le R_{\|h\|_2}(\mu(A)). \tag{2.3}$$

(b) *For all $h \in \mathbf{R}^n$ and every Borel measurable function $f\colon \mathbf{R}^n \to [0,1]$*,

$$\int f(x-h)\,d\mu(x) \le R_{\|h\|_2}(\mathbf{E}_\mu f). \tag{2.4}$$

(c) *For every smooth function $f\colon \mathbf{R}^n \to [0,1]$ with compact support and such that* $0 < \mathbf{E}_\mu f < 1$,

$$\|\mathbf{E}_\mu \nabla f\|_2 \le I(\mathbf{E}_\mu f). \tag{2.5}$$

Here and elsewhere, $\mathbf{E}_\mu$ denotes the expectation (the integral) over μ, μ_h is the translate of μ on vector h, that is, $\mu_h(A) = \mu(A+h)$, for $A \subset \mathbf{R}^n$ and $\nabla f(x) = (\partial f(x)/\partial x_i)_{1 \le i \le n}$ is the gradient of f.

PROOF. On indicator functions $f = 1_A$, (2.4) turns into (2.3). Conversely, starting from (2.3), we get

$$\begin{aligned}
\mathbf{E}_\mu f(x-h) &= \int_0^1 \mu\{x\colon f(x-h) > t\}\,dt = \int_0^1 \mu_h\{x\colon f(x) > t\}\,dt \\
&\le \int_0^1 R_{\|h\|_2}(\mu\{x\colon f(x) > t\})\,dt \\
&\le R_{\|h\|_2}\left(\int_0^1 \mu\{x\colon f(x) > t\}\,dt\right) = R_{\|h\|_2}(\mathbf{E}_\mu f),
\end{aligned}$$

where we have also used Jensen's inequality. Thus, (2.3) and (2.4) are equivalent, and we need to show the equivalence of (2.4) and (2.5).

To derive (2.5) from (2.4), let $h \to 0$ in (2.4). According to the Taylor expansion and in view of (2.2), we get

$$-\langle \mathbf{E}_\mu \nabla f, h\rangle \le I(\mathbf{E}_\mu f)\|h\|_2 + o(\|h\|_2), \tag{2.6}$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbf{R}^n$. Applying this inequality to $h = \varepsilon \mathbf{E}_\mu \nabla f$ and letting $\varepsilon \to 0$, we arrive at (2.5).

To derive (2.4) from (2.5), we may restrict ourselves to smooth functions f with a compact support and such that $0 < \mathbf{E}_\mu f < 1$ (since μ is absolutely continuous). First note that if inequality (2.3) holds for vectors $h_1 = t_1 e$ and $h_2 = t_2 e$ with $t_1, t_2 \ge 0$, then it holds for $h = (t_1 + t_2)e$. Indeed, applying (2.1) and (2.3), we get

$$\mathbf{E}_\mu f(x-h) \le R_{h_1}(\mathbf{E}_\mu f(x-h_2)) \le R_{h_1}(R_{h_2}(\mathbf{E}_\mu f)) = R_h(\mathbf{E}_\mu f). \tag{2.7}$$

This argument shows that (2.4) holds for all h, if it holds for sufficiently small h. Below we precise this argument. Fix a vector $e \in \mathbf{R}^n$ of Euclidean length $|e| = 1$, take a number $c \in (0, 1)$, and consider the set

$$\Delta = \left\{t \geq 0: \mathbf{E}_\mu f(x - cte) \leq R_t(\mathbf{E}_\mu f)\right\}.$$

We will show that $\Delta = [0, +\infty)$ for any $c \in (0, 1)$. Then, letting $c \to 1^-$, we will get (2.4) for all h of the form $h = te$, and since e is arbitrary, we will be done. Clearly, it suffices to check that the following two properties are fulfilled:

(a) $t \in \Delta$, for all $t > 0$ small enough.
(b) if $t_1, t_2 \in \Delta$, then $t_1 + t_2 \in \Delta$.

The property (b) has already been shown in (2.7). To prove (a), again by Taylor's expansion, by (2.6), and since $I(\mathbf{E}_\mu f) > 0$, we get, for $t > 0$ small enough,

$$\begin{aligned} \mathbf{E}_\mu f(x - cte) &= \mathbf{E}_\mu f - c\langle \mathbf{E}_\mu \nabla f, e\rangle t + o(t) \\ &\leq \mathbf{E}_\mu f + c\|\mathbf{E}_\mu \nabla f\|_2 t + o(t) \\ &\leq \mathbf{E}_\mu f + cI(\mathbf{E}_\mu f)t + o(t) < R_t(\mathbf{E}_\mu f), \end{aligned}$$

since $R_t(\mathbf{E}_\mu f) = \mathbf{E}_\mu f + I(\mathbf{E}_\mu f)t + o(t)$, as $t \to 0$. Lemma 3.1 is proved. □

REMARK 2.2. The same functional form can be established for converse shift inequalities. More precisely, let $(S_t)_{t \geq 0}$ form a family of convex continuous functions from $[0, 1]$ onto itself with the semigroup property (1.1) and the generating function $I(p) = \lim_{t \to 0^+}(p - S_t(p))/t$. Then, the inequality

$$\mu_h(A) \geq S_{\|h\|_2}(\mu(A))$$

is equivalent to the functional inequality

$$\int f(x - h)\, d\mu(x) \geq S_{\|h\|_2}(\mathbf{E}_\mu f), \qquad 0 \leq f \leq 1,$$

which in turn is equivalent to the property (c) in Lemma 2.1.

3. The size of singular component. Let μ be an absolutely continuous probability measure on $\mathbf{R}^n$ with a density ρ from the Sobolev space $W_1^1(\mathbf{R}^n)$. This may equivalently be expressed as the property that the function ρ (possibly modified on a set of zero Lebesgue measure) is absolutely continuous on almost all straight lines that are parallel to coordinate axes, and their Radon–Nikodym partial derivatives $\partial_{x_i}\rho(x) = (\partial\rho(x)/\partial x_i)$ are integrable on $\mathbf{R}^n$ (cf. [7], Section 1.1). Assume moreover that the functions $\xi_i(x) = (\partial_{x_i}\rho(x))/\rho(x)$ (which are well defined μ-a.e.) satisfy the following two conditions:

1. $\mathbf{E}_\mu \xi_i^2 \leq \sigma^2$, for all $i = 1, \ldots, n (\sigma \geq 0)$;
2. $\mathbf{E}_\mu \xi_i \xi_j = 0$, for all $1 \leq i < j \leq n$.

For any $t \in \mathbf{R}$, introduce the function R_t on $[0,1]$ as follows. Given $p \in [0,1]$, if $|\arcsin(2p-1)+t| \le \pi/2$, set

$$\begin{aligned} R_t(p) &= \frac{1+\sin(\arcsin(2p-1)+t)}{2} \\ &= \sin^2\left(\frac{t}{2}\right) + p\cos t + \sqrt{p(1-p)}\,\sin t. \end{aligned}$$

Otherwise, we set $R_t(p) = 1$ when $\arcsin(2p-1)+t \ge \pi/2$, and set $R_t(p) = 0$ when $\arcsin(2p-1)+t \le -\pi/2$.

THEOREM 3.1. *For all $h \in \mathbf{R}^n$ and for every Borel set A in $\mathbf{R}^n$,*

$$R_{-\sigma\|h\|_2}(\mu(A)) \le \mu_h(A) \le R_{\sigma\|h\|_2}(\mu(A)). \tag{3.1}$$

In particular,

$$\mathrm{Sing}(\mu_h, \mu) \le \sin^2\left(\frac{\sigma\|h\|_2}{2}\right).$$

PROOF. The functions R_t can also be defined as follows. Let F be the distribution function of the probability measure on $\mathbf{R}$ with density $\frac{1}{2}\cos x$, $|x| < \pi/2$, so that $F(x) = (1+\sin x)/2$ in the support interval. Let F^{-1}: $[0,1] \to [-\pi/2, \pi/2]$ denote its inverse. Then, for all $t \in \mathbf{R}$ and $p \in [0,1]$,

$$R_t(p) = F\big(F^{-1}(p) + t\big).$$

In particular, $R_t(0) = \sin^2(t/2)$, for $0 \le t \le \pi$. One can also observe that, for $0 < p \le 1$,

$$R_t(p) = \lambda\big(B(p)^t\big),$$

where λ is the normalized Lebesgue measure on the 2-sphere $S^2 \subset \mathbf{R}^3$ of unit radius, $B(p)$ is a ball (a cap) on the sphere of measure p and where $B(p)^t$ for $t \ge 0$ denotes the t-neighborhood of $B(p)$ with respect to the geodesic metric on S^2 (respectively, $|t|$-interior for $t \le 0$). Correspondingly, for $p = 0$ and $t \in [0, \pi]$, $R_t(0)$ expresses the area of balls on S^2 of radius t (in the sense of λ).

These definitions ensure that $(R_t)_{t \ge 0}$ (respectively, $(S_t) = (R_{-t})_{t\ge 0}$) forms a family of concave (respectively, convex) functions from $[0,1]$ onto itself with the semigroup property (2.1). According to (2.2), the "generator" of this family is given by $I(p) = \sqrt{p(1-p)}$. Therefore, by Lemma 2.1 and Remark 2.2, both inequalities in (3.1) are equivalent to

$$\|\mathbf{E}_\mu \nabla f\|_2 \le \sigma I(\mathbf{E}_\mu f) = \sigma\sqrt{\mathbf{E}_\mu f(1-\mathbf{E}_\mu f)}, \tag{3.2}$$

where f: $\mathbf{R}^n \to [0,1]$ is an arbitrary smooth function with a compact support.

Now, let us note that $\mathbf{E}_\mu \xi_i = 0$, for all $i = 1, \dots, n$. Indeed, let for definiteness $i = n$. The set $C = \{y = (x_1, \dots, x_{n-1}) \in \mathbf{R}^{n-1}$: $\int \rho(y, x_n)\,dx_n < +\infty\}$ has a full Lebesgue measure. The integral $\int |\partial_{x_n}\rho(x)|\,dx = \mathbf{E}_\mu|\xi_n|$ is finite; hence, for almost all points $y \in C$, the function $x_n \to \rho(y, x_n)$ has bounded total variation on $\mathbf{R}$. In particular, there exist finite nonnegative limits of $\rho(y, x_n)$, as $x_n \to -\infty$ and $x_n \to +\infty$. These limits must be zero, since otherwise $\int \rho(y, x_n)\, dx_n = +\infty$ that contradicts the property $y \in C$. Therefore, for almost all $y \in C$,

$$\int_{-\infty}^{+\infty} \partial_{x_n} \rho(y, x_n)\, dx_n = \lim_{a \to -\infty} \lim_{b \to +\infty} [\rho(y, b) - \rho(y, a)] = 0.$$

The next integration over $y \in C$ yields $\mathbf{E}_\mu \xi_n = 0$.

It remains to do the last step to prove Theorem 3.1. Integrating by parts, we observe that

$$\mathbf{E}_\mu \partial_{x_i} f = -\int_{\mathbf{R}^n} f(x)\, \partial_{x_i} \rho(x)\, dx = -\mathbf{E}_\mu f \xi_i. \tag{3.3}$$

(note that all the integrals are well defined). By assumption (1), the functions ξ_i belong to the space $L^2(\mu)$, and their norms satisfy $\sigma_i = \|\xi_i\|_{L^2(\mu)} \le \sigma$. By assumption (2), these functions are orthogonal: $\langle \xi_i, \xi_j \rangle_{L^2(\mu)} = 0$, for $i < j$, where the scalar product is now taken in $L^2(\mu)$. As noted above, $\mathbf{E}_\mu \xi_i = 0$. Set $\psi_i = \xi_i / \sigma_i$. Then, from (3.3) and by Parseval's inequality,

$$\begin{aligned} \|\mathbf{E}_\mu \nabla f\|_2^2 &= \sum_{i=1}^n \sigma_i^2 \langle f - \mathbf{E}_\mu f, \psi_i \rangle_{L^2(\mu)}^2 \\ &\le \sigma^2 \sum_{i=1}^n \langle f - \mathbf{E}_\mu f, \psi_i \rangle_{L^2(\mu)}^2 \le \sigma^2 \|f - \mathbf{E}_\mu f\|_{L^2(\mu)}^2, \end{aligned}$$

that is,

$$\|\mathbf{E}_\mu \nabla f\|_2^2 \le \sigma^2 \operatorname{Var}_\mu(f) = \sigma^2 \Big[\mathbf{E}_\mu f^2 - (\mathbf{E}_\mu f)^2\Big]. \tag{3.4}$$

Inequalities like (3.4) were earlier studied by many authors in the context of getting lower estimates for variance. For example, the fact that (3.4) with $\sigma = 1$ holds for $\mu = \gamma_n$ was proved by Cacoullos ([2], Proposition 3.7). Now, if $0 \le f \le 1$, we have $\mathbf{E}_\mu f^2 \le \mathbf{E}_\mu f$, so (3.4) implies (3.2). Theorem 3.1 is thus proved. □

Now assume that μ is a product measure on $\mathbf{R}^n$, $\mu = \mu_1 \otimes \cdots \otimes \mu_n$, with density $\rho(x) = \rho_1(x_1) \cdots \rho_n(x_n)$, where ρ_i are absolutely continuous on $\mathbf{R}$. Then, the functions

$$\xi_i(x) = \frac{\partial_{x_i} \rho(x)}{\rho(x)} = \frac{\rho_i'(x_i)}{\rho_i(x_i)}$$

represent independent random variables on probability space $(\mathbf{R}^n, \mu)$. Condition (1) becomes

$$J(\mu_i) = \mathbf{E}_\mu \xi_i^2 \le \sigma^2,$$

which implies in particular that $\mathbf{E}_\mu \xi_i = 0$. By independency, condition (2) of Theorem 3.1 is therefore fulfilled, so the inequality (3.1) holds with $\sigma^2 = \max_i J(\mu_i)$.

In particular, this applies to product probability measures μ^n on $\mathbf{R}^n$ with equal marginals: inequality (3.1) holds with $\sigma^2 = J(\mu)$. Since this inequality is dimension free, it extends to $(\mathbf{R}^\infty, \mu^\infty)$. Note that the transition to the infinite-dimensional inequality can more easily be performed on the basis of the functional inequality (2.4). Thus, we arrive at the statement of Theorem 1.1.

Another broad class of measures when the conditions of Theorem 3.1 are satisfied is the class of measures with densities ρ-symmetric in the following sense: for all $(x_1, \dots, x_n) \in \mathbf{R}^n$ and for any sequence $\varepsilon_1, \dots, \varepsilon_n \in \{-1, 1\}$, $\rho(\varepsilon_1 x_1, \dots, \varepsilon_n x_n) = \rho(x_1, \dots, x_n)$. For example, one can consider densities of the form $\rho(x) = \rho(|x|)$, that is, the measures which are invariant under rotations. Of course, this class contains all product measures μ with symmetric (around 0) marginal distributions μ_i. It should be pointed out that, for nonproduct measures μ with marginals μ_i, in general $\mathbf{E}_\mu |(\partial_{x_i} \rho)/\rho|^2 \neq J(\mu_i)$ (this is easily seen for Gaussian nonproduct measures).

4. Shift inequalities of Gaussian type. In this section, we prove Theorem 1.2. Let us introduce the standard notations: for $x \in [-\infty, +\infty]$, set

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2), \qquad \Phi(x) = \int_\infty^x \varphi(t)\, dt.$$

Let $\Phi^{-1}\colon [0,1] \to [-\infty, +\infty]$ denote the inverse of Φ.

Given a probability measure μ on $\mathbf{R}$, the inequality (1.4) requires that all translates of μ on $\mathbf{R}$ are equivalent to μ. This already implies that μ is equivalent to Lebesgue measure, that is, it has an a.e. positive density ρ (cf. [8], Lemma 5). Moreover, according to Shepp's necessary condition ([8], Theorem 1), ρ can be chosen to be absolutely continuous with finite Fisher information $J(\mu)$. This will be further assumed. Note that this property is weaker than (1.5).

Since inequality (1.4) is dimension free, it can equivalently be written in terms of product measures μ^n on $\mathbf{R}^n$ as

$$\Phi\big(\Phi^{-1}(p) - c\|h\|_2\big) \le \mu_h^n(A) \le \Phi\big(\Phi^{-1}(p) + c\|h\|_2\big), \tag{4.1}$$

where $c > 0$, $A \subset \mathbf{R}^n$ is a Borel set of measure $\mu^n(A) = p$ and where $h \in \mathbf{R}^n$. The family of functions $R_t(p) = \Phi(\Phi^{-1}(p) + t)$, $t \in \mathbf{R}$, forms a group with generator

$$I(p) = \varphi\big(\Phi^{-1}(p)\big), \qquad 0 \le p \le 1,$$

which is a concave function on $[0,1]$ with $I(0) = I(1) = 0$. Therefore, according to Lemma 2.1, shift inequalities in (4.1) have a functional form

$$\|\mathbf{E}\nabla f\|_2 \le cI(\mathbf{E}f), \tag{4.2}$$

where $f\colon \mathbf{R}^n \to [0,1]$ is an arbitrary smooth function with a compact support, and the expectations are with respect to measure $\mathbf{P} = \mu^n$. Integration by parts as in (3.3) shows that, in terms of functions $\xi_i(x) = \rho'(x_i)/\rho(x_i)$, $x \in \mathbf{R}^n$, (4.2) is exactly the property

$$\left(\sum_{i=1} |\mathbf{E} f \xi_i|^2\right)^{1/2} \le cI(\mathbf{E} f), \qquad 0 \le f \le 1. \tag{4.3}$$

Here, the smoothness condition can easily be weakened to just measurability. To further simplify (4.3), let us note that its left-hand side is a convex functional in f while the right-hand side is concave, by concavity of I. Therefore, (4.3) holds for all f with $0 \le f \le 1$ if and only if it holds for indicator functions $f = 1_A$ of Borel sets $A \subset \mathbf{R}^n$:

$$\left(\sum_{i=1}^{n} \left|\int_A \xi_i \, d\mathbf{P}\right|^2\right)^{1/2} \le cI(\mathbf{P}(A)). \tag{4.4}$$

One may also rewrite this inequality as

$$\sup_{\alpha} \sup_{\mathbf{P}(A)=p} \int_A \sum_{i=1} \alpha_i \xi_i \, d\mathbf{P} \le cI(p), \qquad 0 < p < 1, \tag{4.5}$$

where the first supremum is taken over all collections $\alpha = (\alpha_1, \dots, \alpha_n)$ of real numbers such that $\sum_{i=1} \alpha_i^2 = 1$, and the second one is over all Borel sets A in $\mathbf{R}^n$ with $\mathbf{P}(A) = p$.

Recall that ξ_i, $1 \le i \le n$, are independent random variables with $\mathbf{E}\xi_i = 0$ and $\mathbf{E}\xi_i^2 = J(\mu) < +\infty$. If these r.v.'s have continuous distribution, then the second infimum in (4.5) is clearly attained at the set $A(p) = \{x \in \mathbf{R}^n$: $\xi = \sum_{i=1}^n \alpha_i \xi_i > c\}$ where the constant $c = m_p(\xi)$ is quantile of ρ of random variable ξ. For example, in the case of the canonical Gaussian measure $\mathbf{P} = \gamma_n$, we see that $\xi_i(x) = -x_i$ have distribution γ_1, so does ξ, and

$$\int_{A(p)} \xi \, d\mathbf{P} = \int_{\Phi^{-1}(p)}^{+\infty} x \, d\gamma_1(x) = I(p).$$

Thus, inequality (4.5) holds with $c = 1$, and we recover the Gaussian shift inequality

$$\Phi\big(\Phi^{-1}(p) - \|h\|_2\big) \le \gamma_n(A + h) \le \Phi\big(\Phi^{-1}(p) + \|h\|_2\big), \qquad \gamma_n(A) = p.$$

In the general, non-Gaussian case, it is probably rather difficult to find an exact value of the left-hand sides in (4.5). However, one can easily prove that the condition (1.5),

$$\mathbf{E}\exp\big(\varepsilon \xi_1^2\big) \le 2, \qquad \varepsilon > 0, \tag{4.6}$$

makes that supremum behave like $I(p)$. To complete the proof of Theorem 1.2, we need an elementary lemma.

LEMMA 4.1. *Let* $\xi_1, \dots, \xi_n$ *be i.i.d. random variables on a probability space* $(\Omega, \mathscr{F}, \mathbf{P})$ *with* $\mathbf{E}\xi_1 = 0$ *and such that, for some* $c > 0$,

$$\left| \int_A \xi_1 \, d\mathbf{P} \right| \le cI(\mathbf{P}(A)), \qquad A \in \mathscr{F}. \tag{4.7}$$

Then (4.6) *holds with* $\varepsilon = 1/(4c^2)$. *Conversely*, (4.6) *implies* (4.7) *and moreover* (4.4) *with* $c = 2/\sqrt{\varepsilon}$.

First we prove a calculus lemma.

LEMMA 4.2.

$$p\sqrt{\frac{1}{2}\log\frac{1}{p}} \le I(p) \le p\sqrt{2\log\frac{1}{p}}\,, \tag{4.8}$$

where the left inequality holds for all $p \in [0, \frac{1}{2}]$, *and the right inequality holds for all* $p \in [0, 1]$.

PROOF. The right inequality is simple [note that $I(p)$ behaves like $p\sqrt{2\log(1/p)}$, as $p \to 0$]. To prove the left one, consider a function $u(p) = I(p) - dp\sqrt{\log(1/p)}$ with some $d > 0$. We have $u(0) = 0$, and, by definition of I, $u(\frac{1}{2}) \ge 0$ if and only if $d^2 \le 2/(\pi \log 2)$. In particular, $d = 1/\sqrt{2}$ satisfies this condition. Next, using the property $I'' = -1/I$ and applying the right inequality of the lemma, we have for $p \in (0, 1)$,

$$\begin{aligned} u''(p) &= d\left[\frac{1}{2p\sqrt{\log(1/p)}} + \frac{1}{4p\log(1/p)\sqrt{\log(1/p)}}\right] - \frac{1}{I(p)} \\ &\le d\left[\frac{1}{2p\sqrt{\log(1/p)}} + \frac{1}{4p\log(1/p)\sqrt{\log(1/p)}}\right] - \frac{1}{p\sqrt{2\log(1/p)}} \\ &= \frac{1}{p\sqrt{\log(1/p)}}\left[d\left(\frac{1}{2} + \frac{1}{4\log(1/p)}\right) - \frac{1}{\sqrt{2}}\right] \le 0, \end{aligned}$$

for $d = 1/\sqrt{2}$ and $0 < p \le \frac{1}{2}$ in the last inequality. Thus, u is concave on $[0, \frac{1}{2}]$; hence, it is nonnegative on that interval for this value of d.

PROOF OF LEMMA 4.1. If we apply (4.7) to the sets $A = \{\xi_1 > x\}$ with $x \ge 0$ and make use of the right estimate in (4.8), we get

$$x\mathbf{P}\{\xi_1 > x\} \le cI(\mathbf{P}(A)) \le c\mathbf{P}\{\xi_1 > x\}\sqrt{2\log(1/\mathbf{P}\{\xi_1 > x\})}\,,$$

so that $\mathbf{P}\{\xi_1 > x\} \le \exp\{-x^2/2c^2\}$. With a similar bound for $\mathbf{P}\{\xi_1 < -x\}$, we obtain that

$$\mathbf{P}\{|\xi_1| > x\} \le 2\exp\left\{-\frac{x^2}{2c^2}\right\}, \qquad x \ge 0.$$

Therefore, for $\varepsilon < 1/2c^2$,

$$\mathbf{E}\exp\big(\varepsilon\xi_1^2\big) = 2\varepsilon\int_0^{+\infty} \mathbf{P}\{\xi_1 > x\}\, x\exp(\varepsilon x^2)\, dx + 1$$

$$\leq 1 + 4\varepsilon\int_0^{+\infty} x\exp\left(\varepsilon x^2 - \frac{x^2}{(2c^2)}\right) dx = \frac{1 + 2\varepsilon c^2}{1 - 2\varepsilon c^2}.$$

Hence, $\mathbf{E}\exp(\varepsilon\xi_1^2) \leq 2$ as soon as $\varepsilon \leq 1/6c^2$.

Conversely, assume (4.6) holds true. Using also the fact that $\mathbf{E}\xi_1 = 0$, it can easily be shown that, for all $t \in \mathbf{R}$,

$$\mathbf{E}\exp(t\xi_1) \leq \exp\left\{\frac{2t^2}{\varepsilon}\right\}. \tag{4.9}$$

Indeed, the function $v(t) = \log\mathbf{E}\exp(t\xi_1)$ is convex with $v(0) = v'(0) = 0$ and $v''(t) \leq \mathbf{E}\xi_1^2\exp(t\xi_1)$. Now, $t\xi_1 \leq (1/2\varepsilon)t^2 + (\varepsilon/2)\xi_1^2$, so that

$$\mathbf{E}\xi_1^2\exp(t\xi_1) \leq \exp\left(\frac{t^2}{2\varepsilon}\right)\mathbf{E}\xi_1^2\exp\big(\varepsilon\xi_1^2/2\big)$$

$$\leq \frac{2}{\varepsilon}\exp\left(\frac{t^2}{2\varepsilon}\right)\mathbf{E}\exp\big(\varepsilon\xi_1^2 - 1\big) \leq \frac{4}{e\varepsilon}\exp\left(\frac{t^2}{2\varepsilon}\right),$$

where we have applied the inequality $xe^x \leq e^{2x-1}$ with $x = \varepsilon\xi_1^2/2$. For $|t| \leq \sqrt{2\varepsilon}$, we thus get $v''(t) \leq 4/\varepsilon$, so that $v(t) \leq (2t^2/\varepsilon)$ according to Taylor's expansion. This implies (4.9) for such (small) values of t. Once more, since $t\xi_1 \leq (1/4\varepsilon)t^2 + \varepsilon\xi_1^2$, we get $\mathbf{E}\xi_1^2\exp(t\xi_1) \leq 2\exp(t^2/(4\varepsilon))$. This gives (4.9) for $|t| \geq \sqrt{2\varepsilon}$.

Now, let $\xi = \sum_{i=1}^n \alpha_i\xi_i$ with $\sum_{i=1}^n \alpha_i^2 = 1$. Since ξ_i are independent and identically distributed, we obtain from (4.9) that, for all $t \in \mathbf{R}$,

$$\mathbf{E}\exp(t\xi) \leq \exp\left\{\frac{2t^2}{\varepsilon}\right\}. \tag{4.10}$$

Introducing the functional $\mathbf{Ent}(g) = \mathbf{E}g\log g - \mathbf{E}g\log\mathbf{E}g = \sup\{\mathbf{E}g\eta\colon \mathbf{E}e^\eta \leq 1\}$ $(g \geq 0)$, we see that (4.10) is equivalent to saying that, for all $g \geq 0$,

$$\mathbf{E}\left(t\xi - \frac{2t^2}{\varepsilon}\right)g \leq \mathbf{Ent}(g).$$

Optimizing this inequality over all $t \in \mathbf{R}$, we come to the equivalent inequality

$$|\mathbf{E}\xi g| \leq \sqrt{\frac{8}{\varepsilon}\mathbf{E}g\,\mathbf{Ent}(g)}\,.$$

Applying this to indicator function $g = 1_A$ of a set of measure $\mathbf{P}(A) = p \in [0, \frac{1}{2}]$ and further using the left inequality of Lemma 4.1, we get

$$\left|\int_A \xi\, d\mathbf{P}\right| \leq p\sqrt{\frac{8}{\varepsilon}\log\frac{1}{p}} \leq \frac{4}{\sqrt{\varepsilon}}I(p).$$

As a result, we arrive at (4.5) and therefore at (4.4) with $c = 4/\sqrt{\varepsilon}$. If $p \geq \frac{1}{2}$, then we can apply (4.5) to $\Omega \setminus A$. Since $\mathbf{E}\xi_i = 0$, this inequality remains to hold for all p. This proves Lemma 4.1 and Theorem 1.2, together with inequalities in (1.6).

We can now derive from Theorem 1.2 the following corollary.

COROLLARY 4.3. *Let μ be a probability measure on $\mathbf{R}$ with an absolutely continuous density ρ such that*

$$\int_{-\infty}^{+\infty} \exp\left\{ \varepsilon \frac{\rho'(x)^2}{\rho(x)^2} \right\} d\mu(x) \leq 2,$$

where $\varepsilon > 0$. Then, for all $h \in l^2$,

$$\| \mu_h^\infty - \mu^\infty \|_{\mathrm{TV}} \leq 2\left(2\Phi\left(\frac{2\|h\|_2}{\sqrt{\varepsilon}} \right) - 1 \right). \tag{4.11}$$

PROOF. First let us note that, for every $t \geq 0$ and every unimodal, symmetric around 0, probability distribution F on $\mathbf{R}$, the function $u_t(p) = F(F^{-1}(p) + t) - p$, $0 < p < 1$, is maximized at the point $p = F(-t/2)$. Hence,

$$\sup_p u_t(p) = F(t/2) - F(-t/2) = 2F(t/2) - 1.$$

In particular, if $F = \Phi$, $\sup_p u_t(p) = 2\Phi(t/2) - 1$. From (1.4), for every set $A \subset \mathbf{R}^\infty$ of measure $\mu^\infty(A) = p$,

$$| \mu_h^\infty(A) - \mu^\infty(A) | \leq R_{c\|h\|_2}(p) - p = u_{c\|h\|_2}(p) \leq 2\Phi(c\|h\|_2/2) - 1.$$

This proves (4.11), since $\| \mu_h^\infty - \mu^\infty \|_{\mathrm{TV}} = 2\sup_A | \mu_h^\infty(A) - \mu^\infty(A)|$, and since $c \leq 4/\sqrt{\varepsilon}$ according to (1.6).

As we mentioned before, if μ has a finite variance b, then, by the central limit theorem,

$$\sup_{\|h\|_2 = t} \sup_{\mu^\infty(A) = p} \mu_h^\infty(A) \geq \Phi(\Phi^{-1}(p) + t/b).$$

Substracting p from both sides and maximizing over all $p \in (0,1)$ as above, we get $\sup_{\|h\|_2 = t} \sup_A [\mu_h^\infty(A) - \mu^\infty(A)] \geq 2\Phi(t/2b) - 1$; that is,

$$\sup_{\|h\|_2 = t} \| \mu_h^\infty - \mu^\infty \|_{\mathrm{TV}} \geq 2(2\Phi(t/2b) - 1). \tag{4.12}$$

This shows that (4.11) is sharp up to a constant in front of $\|h\|_2$, provided we want to control the total variation norm in terms of the Euclidean length of h. Actually, for all log-concave probability distributions μ, an estimate like (4.11) can be reversed in the sense that, after a modification of the right-hand side of (4.12), the supremum in (4.12) can be replaced with the infimum

$$\inf_{\|h\|_2 = t} \| \mu_h^\infty - \mu^\infty \|_{\mathrm{TV}} \geq 2(1 - \exp(-ct)). \tag{4.13}$$

This is a consequence of a recent result due to Krugova [5], who showed that, given a Radon log-concave probability measure μ on a locally convex space E, for every vector $h \in E$, along which μ is differentiable,

$$\|\mu_h - \mu\|_{\mathrm{TV}} \geq 2\big(1 - \exp(-\|d_h\,\mu\|_{\mathrm{TV}}/2)\big). \tag{4.14}$$

Recall that μ is differentiable in direction $h \in E$ (in the sense of Fomin; cf. [1]) if, for every Borel set A in E, the limit

$$\lim_{\varepsilon \to 0} \frac{\mu(A + \varepsilon h) - \mu(A)}{\varepsilon} = d_h\,\mu(A)$$

exists and defines a finite signed Radon measure on E. In the case of the product measure μ^∞ on $E = \mathbf{R}^\infty$ with finite Fisher information $J(\mu)$, the derivative $d_h\,\mu$ is well defined for all $h \in l^2$ and absolutely continuous with respect to μ^∞ with density $\sum_i h_i\,\xi_i$, where $\xi_i(x) = \rho'(x_i)/\rho(x_i)$. By Marcinkiewicz–Zygmund's inequality, for some universal constant $K > 0$, we have

$$\begin{aligned}\|d_h\,\mu\|_{\mathrm{TV}} &= \int_{R^\infty} \left| \sum_{i=1}^{\infty} h_i\,\xi_i \right| d\mu^\infty \geq K \int_{\mathbf{R}^\infty} \sqrt{\sum_{i=1}^{\infty} h_i^2 \xi_i^2}\; d\mu^\infty \\ &\geq K \sqrt{\sum_{i=1}^{\infty} h_i^2 (\mathbf{E}|\xi_i|)^2} = K\mathbf{E}|\xi_1|\,\|h\|_2,\end{aligned}$$

where the expectations are with respect to μ. It then follows from (4.14) that (4.13) holds for every log-concave measure μ with $c = K\mathbf{E}|\xi_1|/2$. For example, for the two-sided exponential distribution $\mu = \nu$, we have

$$2\big(1 - \exp(-c_1\|h\|_2)\big) \leq \|\nu_h^\infty - \nu^\infty\|_{\mathrm{TV}} \leq 2\big(2\Phi(c_2\|h\|_2) - 1\big),$$

where both sides are sharp up to some numerical constants c_1 and c_2 [the left inequality is sharp for the vectors $h = (t, 0, 0, \ldots)$].

5. Shift inequalities of exponential type. By a similar argument, we prove in this section the following analogue of Theorem 1.2. Set

$$H(p) = p \log \frac{1}{p} + (1 - p)\log \frac{1}{1 - p}, \qquad 0 \leq p \leq 1$$

and define the distribution function F on the real line via its inverse

$$F^{-1}(p) = \int_{1/2}^{p} \frac{dt}{H(t)}, \qquad 0 < p < 1.$$

THEOREM 5.1. *Let μ be a probability measure on $\mathbf{R}$. There exists $c > 0$ such that, for all Borel sets $A \subset \mathbf{R}^\infty$ of measure $\mu^\infty(A) = p$ and every $h \in l^2$, an inequality*

$$F\big(F^{-1}(p) - c\|h\|_2\big) \leq \mu_h^\infty(A) \leq F\big(F^{-1}(p) + c\|h\|_2\big) \tag{5.1}$$

holds true, if and only if μ *has an absolutely continuous density* ρ *with*

$$\int_{-\infty}^{+\infty} \exp\left\{\varepsilon\left|\frac{\rho'(x)}{\rho(x)}\right|\right\} d\mu(x) \le 2 \quad \textit{for some } \varepsilon > 0. \tag{5.2}$$

As we will see, the optimal constants c and ε in (5.1) and (5.2) are connected with the relation

$$\frac{1}{6\varepsilon} \le c \le \frac{4}{\varepsilon}. \tag{5.3}$$

The function H is concave, symmetric around the point $p = 1/2$, with $H(0) = H(1) = 0$, so F is a distribution function of a symmetric log-concave probability measure on $\mathbf{R}$. Since $t\log(1/t) \ge (1-t)\log(1/(1-t))$, for $t \in [0, \frac{1}{2}]$, it follows from the definition of F that

$$2^{-\exp(-x/2)} \le F(x) \le 2^{-\exp(-x)}, \qquad x \le 0. \tag{5.4}$$

Hence, F may be viewed as a symmetrized double exponential distribution.

The family of functions $R_t(p) = F(F^{-1}(p) + t)$, $t \in \mathbf{R}$, forms a group with generator $H(p)$. Therefore, according to Lemma 2.1 with $I = H$, and as in the pass from (4.1) to (4.4), the shift inequalities in (5.1) are equivalent to the property

$$\left(\sum_{i=1}^{n}\left|\int_A \xi_i \, d\mathbf{P}\right|^2\right)^{1/2} \le cH(\mathbf{P}(A)), \tag{5.5}$$

where $\mathbf{P} = \mu^n$ (n is an arbitrary positive integer), $\xi_i(x) = \rho'(x_i)/\rho(x_i)$, $x \in \mathbf{R}^n$ and where A is an arbitrary Borel set in $\mathbf{R}^n$. To complete the proof of Theorem 5.1, we need a statement similar to Lemma 4.1.

LEMMA 5.2. *Let* $\xi_1, \dots, \xi_n$ *be i.i.d. random variables on a probability space* $(\Omega, \mathscr{F}, \mathbf{P})$ *with* $\mathbf{E}\xi_1 = 0$. *If inequality* (5.5) *holds with* $n = 1$ *for some* $c > 0$, *then*

$$\mathbf{E}\exp(\varepsilon|\xi_1|) \le 2 \tag{5.6}$$

with $\varepsilon = 1/(6c)$. *Conversely*, (5.6) *implies* (5.5) *for all* n *with* $c = 4/\varepsilon$.

PROOF. First assume (5.5) holds with $n = 1$,

$$\left|\int_A \xi_1 \, d\mathbf{P}\right| \le cH(\mathbf{P}(A)), \qquad A \in \mathscr{F}. \tag{5.7}$$

For the sets $A(x) = \{\xi_1 > x\}$, $x \ge 0$, we have $\int_{A(x)} \xi_1 \, d\mathbf{P} \ge x\mathbf{P}(A(x))$. Recalling that $H(p) \le 2p\log(1/p)$ for $p \le \frac{1}{2}$ and assuming $\mathbf{P}(A(x)) \le \frac{1}{2}$, we get from (5.7) that $\mathbf{P}(A(x)) \le \exp\{-x/(2c)\}$. With a similar argument concerning $-\xi_1$, one can conclude that

$$\mathbf{P}\{|\xi_1| > x\} \le \min\{1, 2\exp(-x/(2c))\}, \qquad x \ge 0.$$

Therefore, for $\varepsilon < 1/(2c)$,

$$\begin{aligned}\mathbf{E}\exp(\varepsilon|\xi_1|) &= 1 + \varepsilon\int_0^{+\infty}\exp(\varepsilon x)P\{|\xi_1| > x\}\,dx\\ &\le 1 + \varepsilon\int_0^{+\infty}\exp(\varepsilon x)\min\left\{1, 2\exp\left(-\frac{x}{2c}\right)\right\}dx\\ &= \frac{2^{2\varepsilon c}}{1-2\varepsilon c}.\end{aligned}$$

It is now easy to see that $\mathbf{E}\exp(\varepsilon|\xi_1|) \le 2$ for $\varepsilon c \le 1/6$.

To prove the converse, we need the following lemma.

LEMMA 5.3. *Let η be a random variable.*

(a) *If $\mathbf{E}e^{|\eta|} \le 2$, $\mathbf{E}\eta = 0$ and $|t| \le \frac{1}{2}$, then $\mathbf{E}\exp(t\eta) \le \exp(5t^2/2)$.*
(b) *If $\mathbf{E}e^{\eta} \le 2$ and $\mathbf{P}(A) \le \frac{1}{2}$, then $\int_A \eta\,d\mathbf{P} \le 2\mathbf{P}(A)\log(1/\mathbf{P}(A))$.*

PROOF. (a) As in the proof of Lemma 4.2, on the interval $|t| \le \frac{1}{2}$, consider the function $v(t) = \log\mathbf{E}\exp(t\eta)$. It is convex with $v(0) = v'(0) = 0$ and $v''(t) \le \mathbf{E}\eta^2\exp(t\eta) \le \mathbf{E}\eta^2\exp(|\eta|/2)$. Applying the inequality $\eta^2\exp(|\eta|/2) \le \frac{5}{2}\exp(|\eta|)$, we get $v''(t) \le 5$. This implies $v(t) \le 5t^2/2$.

To prove (b), recall that the inequality $\mathbf{E}\exp(\eta) \le 2$ is equivalent to the property $\mathbf{E}(\eta - \log 2)g \le \mathbf{Ent}(g)$, for all (bounded) $g \ge 0$. For indicator function $g = 1_A$, this gives $\int_A \eta\,d\mathbf{P} \le \mathbf{P}(A)\log 2 + \mathbf{P}(A)\log(1/\mathbf{P}(A))$. Since $\log 2 \le \log(1/\mathbf{P}(A))$, we obtain the result.

Now we can derive (5.5) from (5.6). Since $\mathbf{E}\xi_i = 0$ and since the function H is symmetric around the point $1/2$, we may assume that $\mathbf{P}(A) \le 1/2$. By Lemma 5.3(a) with $\eta = \varepsilon\xi_i$, we have $\mathbf{E}\exp(t\varepsilon\xi_i) \le \exp(5t^2/2)$, for all $|t| \le \frac{1}{2}$. Therefore, the same inequality holds for r.v.'s $\xi = a_1\xi_1 + \cdots + a_n\xi_n$, where $a_1^2 + \cdots + a_n^2 = 1$. In particular, for $t = 1/2$, $\mathbf{E}\exp(\varepsilon\xi/2) \le \exp(5/8) \le 2$. By Lemma 5.3(b) with $\eta = \varepsilon\xi/2$,

$$\frac{\varepsilon}{2}\int_A \sum_{i=1}^n a_i\xi_i\,d\mathbf{P} \le 2\mathbf{P}(A)\log\frac{1}{\mathbf{P}(A)}.$$

Optimizing over the coefficients (a_i) and since $p\log(1/p) \le H(p)$, we obtain (5.5) with $c = 4/\varepsilon$. Lemma 5.2 and Theorem 5.1 are proved. □

Using (5.2)–(5.4), with the proof very much similar to the one of Corollary 4.3, we get the following.

COROLLARY 5.4. *Let μ be a probability measure on $\mathbf{R}$ with an absolutely continuous density ρ satisfying* (5.2). *Then, for all $h \in l^2$,*

$$\|\mu_h^\infty - \mu^\infty\|_{\mathrm{TV}} \le 2\left(1 - 2^{1-\exp\{\|h\|_2/\varepsilon\}}\right). \tag{5.8}$$

Acknowledgment. I am grateful to W. Li for stimulating discussions. The paper was partly prepared during the author's stay at the School of Mathematics, Georgia Institute of Technology, Atlanta.

REFERENCES

[1] BOGACHEV, V. I. and SMOLYANOV, O. G. (1990). Analytic properties of infinite-dimensional distributions. *Russian Math. Surveys* **45** 3–83.

[2] CACOULLOS, T. (1982). On upper and lower bounds for the variance of a function of a random variable. *Ann. Probab.* **10** 799–809.

[3] FELDMAN, J. (1961). Examples of non-Gaussian quasi-invariant distributions on Hilbert space. *Trans. Amer. Math. Soc.* **99** 342–349.

[4] KAKUTANI, S. (1948). On equivalence of infinite product measures. *Ann. Math.* **49** 214–224.

[5] KRUGOVA, E. P. (1997). On translates of convex measures. *Sbornik Math.* **188** 227–236.

[6] KUELBS, J. and LI, W. V. (1998). Some shift inequalities for Gaussian measures. *Progr. Probab.* **43** 233–243.

[7] MAZ'JA, V. G. (1985). *Sobolev Spaces*. Springer, Berlin.

[8] SHEPP, L. A. (1965). Distinguishing a sequence of random variables from a translate of itself. *Ann. Math. Statist.* **36** 1107–1112.

DEPARTMENT OF MATHEMATICS
SYKTYVKAR UNIVERSITY
167001 SYKTYVKAR
RUSSIA
E-MAIL: bobkov@ssu.edu.komi.ru