The Annals of Probability
1997, Vol. 25, No. 1, 184–205

ISOPERIMETRIC CONSTANTS FOR PRODUCT PROBABILITY MEASURES

BY S. G. BOBKOV[1] AND C. HOUDRÉ[2]

Syktyvkar University and Georgia Institute of Technology

A dimension free lower bound is found for isoperimetric constants of product probability measures. From this, some analytic inequalities are derived.

1. Introduction. Let (X, d) be a metric space equipped with a separable Borel probability measure μ, and assume that μ is not a unit mass at a point. In the present paper we study the quantity

$$Is(\mu) = \inf \frac{\mu^+(A)}{\min(\mu(A), 1 - \mu(A))} \tag{1.1}$$

which was introduced by Cheeger [6] in a Riemannian geometry context. The infimum in (1.1) is taken over all Borel sets $A \subset X$ of measure $0 < \mu(A) < 1$ (such sets exist by the assumptions on μ), and μ^+ denotes the surface measure of A, that is,

$$\mu^+(A) = \liminf_{h \to 0^+} \frac{\mu(A^h) - \mu(A)}{h},$$

where $A^h = \{x \in X: d(x, a) < h \text{ for some } a \in A\}$ is the open h–neighborhood of A (for the metric d).

For any function $f: X \to \mathrm{R}$, we also define the modulus of its gradient

$$|\nabla f(x)| = \limsup_{d(x, y) \to 0^+} \frac{|f(x) - f(y)|}{d(x, y)},$$

to which we assign the value 0 whenever x is an isolated point in X; clearly, $|\nabla f|$ is always Borel measurable for f continuous. The space $X^n = X \times \cdots \times X$ is endowed with the metric d_n given by $d_n(x, y) = (\sum_{k=1}^n d^2(x_k, y_k))^{1/2}$ and with the probability measure μ^n which is the n-fold tensor product of μ with itself. Also, we assume that for any Lipschitz function f on (X^n, d_n), $|\nabla f|^2 =$

Received July 1995; revised April 1996.

[1]Research supported in part by ISF Grant NXZ000 and NXZ300 and by the Alexander von Humboldt-Stiftung. This author enjoyed the hospitality of the Faculty of Wiskunde and Informatica, Free University of Amsterdam and of the Faculty of Mathematics, Bielefeld University, while part of this research was carried out.

[2]Research supported in part by an NSF Postdoctoral Fellowship. This author enjoyed the hospitality of Le Cermics, ENPC, France, of the Steklov Mathematical Institute (Sankt Petersburg branch) and of the Department of Mathematics, University of Syktyvkar, Russia, while part of this research was carried out.

AMS 1991 *subject classifications.* 60E15, 28A35, 49Q20.

Key words and phrases. Isoperimetry, Poincaré Inequalities, Cheeger's inequality, Khintchine–Kahane inequality, Hölder's Inequality.

$\sum_{k=1}^{n} |\nabla_{x_k} f|^2$ almost everywhere (with respect to μ^n). On the Euclidean space $X = \mathrm{R}^n$ and via the Rademacher theorem, this standing assumption holds for any absolutely continuous probability measure.

With these notations, our main result can be stated as follows.

THEOREM 1.1. *For any triple* (X, d, μ) *as above,*

$$Is(\mu^n) \geq \frac{1}{2\sqrt{6}}\, Is(\mu), \tag{1.2}$$

for all $n = 1, 2, \ldots$. *Equivalently, and up to a universal constant, for any function* $f\colon X^n \to [0, 1]$ *which has finite Lipschitz constant on every ball in* (X^n, d_n),

$$K\, \mathrm{Var}(f) \leq \mathrm{E} \min\left(\frac{1}{Is(\mu)}\, |\nabla f|, \frac{1}{Is^2(\mu)}\, |\nabla f|^2\right). \tag{1.3}$$

Above, the expectation and the variance are taken with respect to μ^n, *and one can take* $K = 1/144$.

From (1.3), $KIs(\mu)\,\mathrm{Var}(f) \leq \mathrm{E}\,|\nabla f|$, and approximating the indicator function 1_A by Lipschitz functions f_k in such a way that $\liminf_k \mathrm{E}|\nabla f_k| \leq (\mu^n)^+(A)$ (see Lemma 3.5), gives

$$(\mu^n)^+(A) \geq KIs(\mu)\,\mathrm{Var}(1_A) \geq \tfrac{1}{2} KIs(\mu) \min(\mu^n(A), 1 - \mu^n(A)).$$

Therefore, (1.3) implies (1.2) with a worse but still universal constant.

One of the most interesting partial cases of Theorem 1.1 is when the measure μ is the double exponential distribution on the real line $X = \mathrm{R}$, $\nu(dx) = 2^{-1} \exp(-|x|)\, dx$. It is known (Talagrand [15]) that ν satisfies the isoperimetric inequality

$$\nu^+(A) \geq \min(\nu(A), 1 - \nu(A)), \tag{1.4}$$

with equality for the intervals $A = (-\infty, x]$, and thus, $Is(\nu) = 1$. It is then natural to ask whether (1.4) continue to hold for the product measure ν^n with a (multiplicative) constant independent of the dimension, that is, whether $\inf_n Is(\nu^n) > 0$. Equivalently, one can ask whether or not ν^n satisfies an L^1-Poincaré type inequality with a dimension free constant, that is, whether or not for all smooth functions f on R^n with $\mathrm{E}f = 0$,

$$K\, \mathrm{E}|f| \leq \mathrm{E}\,|\nabla f|. \tag{1.5}$$

Theorem 1.1 gives a positive answer to this question and in fact the following characterization holds.

THEOREM 1.2. *Let* μ *be a probability measure on the real line* R *which is not a unit mass at a point. The following properties are equivalent.*

(i) *The measure* μ^n *satisfies* (1.5) *for some positive constant independent of the dimension.*

(ii) *The measure* μ *satisfies* (1.5) *for some positive constant* ($n = 1$).

(iii) *There exists a function* U: R $\to$ R *with finite Lipschitz constant which transforms the double exponential measure* ν *into* μ.

(iv) $Is(\mu) > 0$.

In addition, (1.5) *holds with* $K = Is(\mu)/(2\sqrt{6})$, *and* $Is(\mu^n) \geq Is(\mu)/(2\sqrt{6})$, *for all* $n = 1, 2, \ldots$.

In the proof of the multidimensional inequality (1.5), Theorem 1.1 is applied to ν (at which point the assumption on the gradient is trivially verified); next, (1.5) is extended to all the Lipschitz images of ν. Hence, on the real line, our standing assumption on the gradient can be omitted in Theorem 1.1. However, this assumption is essential to perform the induction step in the metric space X where it is unlikely that it is possible to find a probability distribution with properties as specific as the ones of the double exponential measure ν on R.

Note that by (iii), μ necessarily has a finite exponential moment, and as easily seen from (iv), μ also has a nontrivial absolutely continuous component. Moreover, in terms of the distribution function, the property (iv) can be verified with the help of the following theorem.

THEOREM 1.3. *Let* $F(x) = \mu((-\infty, x])$ *be the distribution function of a probability measure* μ *on the real line* (*μ is not the unit mass at a point*) *and let* p *be the density of its absolutely continuous part. Then,*

$$Is(\mu) = \operatorname*{ess\,inf}_{a<x<b} \frac{p(x)}{\min(F(x), 1 - F(x))}, \tag{1.6}$$

where $a = \inf\{x\colon F(x) > 0\}$, $b = \sup\{x\colon F(x) < 1\}$. *Also,* $Is(\mu) = 1/\|U\|_{\mathrm{Lip}}$ *where* U *is the nondecreasing left-continuous function which transforms* ν *into* μ.

Clearly, such a nondecreasing function always exists, is unique and is also Lipschitz in order to satisfy $Is(\mu) > 0$. As for (1.6), it just tells us that in the case of the real line, it suffices to take, in the definition (1.1), the intervals $A = (-\infty, x]$, for all or even for almost all (with respect to the Lebesgue measure) x. For example, for the measures of the form

$$\mu_\alpha = \alpha\lambda + (1 - \alpha)\mu, \qquad \alpha \in (0, 1),$$

where λ is the uniform distribution on [0,1], and μ is an arbitrary probability measure on [0,1], we find from (1.6) that $Is(\mu_\alpha) \geq 2\alpha$, and also $Is(\mu_\alpha) = 2\alpha$ when μ is singular and nonatomic. On the other hand, the measure μ with density $p(x) = |x| \exp(-x^2)$ has Gaussian tails; however, $Is(\mu) = 0$ since the map $U(x) = \mathrm{sign}(x)\sqrt{|x|}$, which transforms ν into μ is not Lipschitz, or since $p(0) = 0$.

It is interesting here to compare (1.5) with the similar L^2-Poincaré type inequality

$$K\,\mathrm{E}|f|^2 \leq \mathrm{E}\,|\nabla f|^2, \tag{1.7}$$

where again $\mathrm{E}f = 0$ and $K > 0$. In contrast to (1.5), this inequality (as well as the related log-Sobolev inequality; see Gross [8]) is additive; that is, it can be extended to higher dimensions with the same constant K, and in this sense, (1.7) is better behaved than (1.5). The fact that (1.7) follows from the property $Is(\mu) > 0$ has been known since the paper of Cheeger, but apparently was first noted in the probability literature by Borovkov and Utev ([5], Theorem B) only in 1983. In addition, taking for μ the measure of density $|x|$, $|x| \le 1$, shows that the class of probability distributions satisfying (1.7) is strictly larger than the class of those satisfying (1.5). On the real line, assuming for simplicity that 0 is a median of μ, the condition $Ef = 0$ in (1.7) can be replaced by the condition $f(0) = 0$ (this only potentially changes the optimal constant K), and then (1.7) becomes a partial case of the broadly studied Hardy-type inequalities. An important result obtained (in a more general setting) by Artola, Talonti and Tomaselli (see [13]) asserts that a probability measure μ (with median 0) satisfies (1.7) if and only if

$$\sup_{x>0} (1 - F(x)) \int_0^x \frac{1}{p(t)}\, dt < +\infty, \qquad \sup_{x<0} F(x) \int_x^0 \frac{1}{p(t)}\, dt < +\infty,$$

where F and p are as in Theorem 1.3. When μ has a continuous positive density whose support is an interval, these conditions can be combined and rewritten in terms of the map U as follows:

$$\sup_{x>0} e^{-x} \int_0^x \big[U'(t)^2 + U'(-t)^2\big] e^t\, dt < +\infty.$$

In addition to Sobolev-type inequalities, (1.2) can also be linked to some concentration inequalities. Letting for simplicity $\mu = \nu$, (1.2) is equivalent (see [4], Theorem 2.1) to

$$\nu^n(A^h) \ge \nu\Big(\Big(-\infty, a + \frac{h}{2\sqrt{6}}\Big]\Big), \qquad h > 0, \tag{1.8}$$

where a is chosen such that $\nu^n(A) = \nu((-\infty, a])$ and where $A \subset \mathrm{R}^n$ is an arbitrary Borel set. In this setting, Talagrand [15] (see also Maurey [12]) proved that

$$\nu^n(A + \sqrt{h} B_2 + h B_1) \ge \nu\Big(\Big(-\infty, a + \frac{h}{K}\Big]\Big), \qquad h > 0, \tag{1.9}$$

where B_2 and B_1 are, respectively, the l^2 and l^1 unit balls in R^n and where K is a universal constant. Since $A^h = A + hB_2$, (1.9) is stronger than (1.8) for h large. However, for h small (which is important in obtaining sharp constants in Sobolev-type inequalities), (1.9) does not imply (1.8). It should nevertheless be noted here that (1.3) also involves a certain type of mixture of the L^1 and L^2 norms of the gradient.

A natural way to prove (1.2) is to establish its equivalent functional form (1.5) [with a dimension free constant $K(\mu)$]. In turn, a natural way of proving (1.5) is to use an induction procedure on the dimension. However, the space L^1 does not seem adequate to perform this induction. Instead, it is necessary to find a more suitable functional space and functional inequality which will allow an inductive proof. This functional inequality should enjoy the following two properties: (1) it should possess the additivity property; (2) it should become isoperimetric on indicator functions. For example, the inequality (1.7) as well as the log-Sobolev type inequality is additive, and one can derive from them appropriate concentration inequalities (see [7], [1], [10]). However, it is unlikely that it is possible to derive isoperimetric inequalities from these since they do not contain information when one approximates the indicator functions by Lipschitz functions. For our purposes, a more suitable inequality could be one of the form

$$I(\mathsf{E}f) \le \mathsf{E}\sqrt{I(f)^2 + |\nabla f|^2}, \tag{1.10}$$

which was introduced in [3] (see also [2]), where it was studied for the uniform distribution on the discrete cube. The inequality (1.10) clearly satisfies (1), and for dimension n it gives on indicator functions the isoperimetric inequality

$$(\mu^n)^+(A) \ge I(\mu^n(A)). \tag{1.11}$$

One can therefore wonder whether (1.10) is stronger than (1.11). In the Gaussian case on $X = \mathsf{R}$ with the Euclidean metric, and with the Gaussian isoperimetric function I, it turns out that the inequality (1.10) with $n = 1$ becomes (1.11) with $n = 2$. We extend this observation to the case $I(p) = Kp(1-p)$ and claim that the functional inequality (1.10) for dimension $n = 1$ (therefore, for all dimensions) is equivalent, up to a universal constant, to the isoperimetric inequality (1.11) for dimension $n = 2$; that is, the geometric information in (1.11) when $n \ge 2$ is contained "in the plane." On the other hand, for $n = 1$, (1.10) with I as above reduces in essence to the Poincaré-type inequality

$$\mathsf{E}N(f - \mathsf{E}f) \le \mathsf{E}N(K|\nabla f|), \tag{1.12}$$

with $N(x) = \sqrt{1+x^2} - 1$, $x \in \mathsf{R}$, which behaves like x^2 for $|x|$ small and like $|x|$ for $|x|$ large. In particular, the inequality (1.3) corresponds, up to a constant, to this choice of N. We will study (1.12) separately in the context of Cheeger-type inequalities. Then, the induction step will be performed for the inequality

$$\operatorname{Var} f \le \mathsf{E}N(K|\nabla f|),$$

which is related to (1.10). We are now ready for some preliminaries.

2. A generalization of Hölder's inequality.

Let (Ω, μ) be a measure space and let $N\colon \mathsf{R} \to \mathsf{R}$ be a differentiable convex function.

LEMMA 2.1. *Let f and g be measurable functions on Ω such that*

$$\int_\Omega N(g)\, d\mu \le \int_\Omega N(f)\, d\mu, \tag{2.1}$$

then (provided all the written integrals exist)

$$\int_\Omega N'(f) g \, d\mu \le \int_\Omega N'(f) f \, d\mu. \tag{2.2}$$

PROOF. It suffices to prove the result for f and g bounded and μ finite. First, by convexity,

$$\int_\Omega N((1-t)f + tg) \, d\mu \le (1-t) \int_\Omega N(f) \, d\mu + t \int_\Omega N(g) \, d\mu, \quad 0 \le t \le 1. \tag{2.3}$$

Now, (2.3) becomes equality at $t = 0$ (and $t = 1$) and the left-hand side of (2.3) is a convex function of t while the right-hand side is linear. Thus, at $t = 0$, the slope of the left-hand side of (2.3) is dominated by the slope of the right-hand side. Differentiating at $t = 0$ gives

$$\int_\Omega N'(f)(g - f) \, d\mu \le \int_\Omega N(g) \, d\mu - \int_\Omega N(f) \, d\mu.$$

The lemma follows.

The proof above is due to A. V. Zhubr, and very elegantly replaces the original one. Let now $\|\cdot\|_p$, $p > 1$, denote the L^p-norm with respect to μ, and let $q = p/(p-1)$. Applying Lemma 2.1, with $N(x) = |x|^p$, to $f = u^{1/(p-1)}$, $g = v$, where $u, v \ge 0$ are such that $\|u\|_q = 1$, $\|v\|_p = 1$, gives equality in (2.1), and (2.2) becomes

$$\int_\Omega uv \, d\mu \le 1 = \|u\|_q \|v\|_p.$$

3. An extension of Cheeger's inequality. We return to the setting of the introductory section. Let also N be a Young function; that is, $N\colon \mathrm{R} \to \mathrm{R}$ is even and nonnegative, with $N(0) = 0$ and $N(x) > 0$ for all $x \ne 0$. Moreover, assume that

$$C_N = \sup_{x > 0} \frac{x N'(x)}{N(x)} < +\infty, \tag{3.1}$$

where N' is a Radon–Nikodym derivative of N (clearly, C_N does not depend on the choice of N'). We also denote by $L_N(X, \mu)$ the Orlicz space of functions f such that

$$\|f\|_N = \inf\{\lambda > 0 \colon \mathrm{E} N(f/\lambda) \le 1\} < +\infty.$$

Finally, and for simplicity, we write $\|\nabla f\|_N = \| \, |\nabla f| \, \|_N$, while $m(f)$ denotes a median of f. It is worthwhile to note that, for f Lipschitz on every ball in X, the function $|\nabla f|$ is Borel measurable and finite. Indeed, the set X_0 of all isolated points in X is open, and $|\nabla f| = 0$ on X_0, while on its complement $X_1 = X \setminus X_0$,

$$|\nabla f(x)| = \lim_{n \to \infty} \sup_{d(x, y) < 1/n} \frac{|f(x) - f(y)|}{d(x, y)}$$

is the monotone limit of a sequence of lower semicontinuous functions on X_1. The finiteness follows from the Lipschitz property.

THEOREM 3.1. *Let $Is(\mu) > 0$. Then, for all functions f which are Lipschitz on every ball in X and such that $m(f) = 0$,*

$$\|f\|_N \le \frac{C_N}{Is(\mu)} \|\nabla f\|_N, \tag{3.2}$$

$$\mathsf{E}N(f) \le \mathsf{E}N\left(\frac{C_N}{Is(\mu)} |\nabla f|\right). \tag{3.3}$$

LEMMA 3.2 (Co-area inequality). *Let f be a function on X with a finite Lipschitz constant, then*

$$\int_X |\nabla f(x)|\, d\mu(x) \ge \int_{-\infty}^{+\infty} \mu^+\{x \in X\colon f(x) > t\}\, dt. \tag{3.4}$$

REMARK 3.3. The integrand on the right-hand side of (3.4) is a measurable function on the real line. Indeed, let $A \subset X$ be Borel measurable, and let r take only rational values. Whenever $h > 0$, $\cup_{0<r<h} A^r = A^h$, hence for any $\varepsilon > 0$,

$$\inf_{0<h<\varepsilon} \frac{\mu(A^h) - \mu(A)}{h} = \inf_{0<r<\varepsilon} \frac{\mu(A^r) - \mu(A)}{r}.$$

Therefore,

$$\liminf_{r\to 0^+} \frac{\mu(A^r) - \mu(A)}{r} = \mu^+(A).$$

Hence, for any nonincreasing family of Borel sets A_t, $t \in \mathrm{R}$, the function $t \longrightarrow \mu^+(A_t)$ is Borel measurable.

REMARK 3.4. Equality in (3.4) requires some additional properties of μ, such as nonsingularity. In fact, let $X = \mathrm{R}$ with its usual metric, let μ be an arbitrary Borel probability measure on R and let μ_{ac} denote the absolutely continuous (with respect to the Lebesgue measure) part of μ. If $f(x) = x$, then $p(t) = \mu^+\{x \in X\colon f(x) > t\}$ is a Radon–Nikodym derivative (with respect to the Lebesgue measure) of μ_{ac}, and (3.4) becomes $1 \ge \mu_{\mathrm{ac}}(\mathrm{R})$. Therefore, and for $X = \mathrm{R}$, equality in (3.4) requires that $\mu = \mu_{\mathrm{ac}}$, that is, that μ is absolutely continuous. As well known, the usual co-area formula tells us that this property is also sufficient.

PROOF OF LEMMA 3.2 ([4]). First, let us assume that f is bounded. Then, without loss of generality, one may assume that $f \ge 0$, since the left- and the right-hand side of (3.4) remain unchanged if a constant is added to f. Since f is Lipschitz on X,

$$|f(x) - f(y)| \le c d(x, y), \tag{3.5}$$

for some $c > 0$ and all $x, y \in X$. Then, let

$$f_h(x) = \sup_{d(x,y)<h} f(y),$$

where $h > 0$, and let $A_t = \{x \in X: f(x) > t\}$. Then, for all $t \in \mathrm{R}$ and $h > 0$, the set $\{x \in X: f_h(x) > t\} = \{x \in X: f(x) > t\}^h = A_t^h$ is open as the open h-neighborhood of A_t. Therefore f_h is lower semicontinuous and in addition,

$$\int_X f_h \, d\mu = \int_0^{+\infty} \mu\{x \in X: f_h(x) > t\} \, dt = \int_0^{+\infty} \mu(A_t^h) \, dt.$$

Since $\int_X f \, d\mu = \int_0^{+\infty} \mu(A_t) \, dt$, we have

$$\int_X \frac{f_h - f}{h} \, d\mu = \int_0^{+\infty} \frac{\mu(A_t^h) - \mu(A_t)}{h} \, dt. \tag{3.6}$$

From (3.5), $f_h(x) - f(x) \le ch$, for all $x \in X$ and $h > 0$, hence the integrand on the left-hand side of (3.4) is bounded. Therefore, using (3.6), the Lebesgue dominated convergence theorem and Fatou's lemma and noting that

$$\limsup_{h \to 0^+} \frac{f_h(x) - f(x)}{h} = \limsup_{y \to x} \frac{f(y) - f(x)}{d(x, y)} \le |\nabla f(x)|,$$

we get

$$\begin{aligned}
\int_X |\nabla f| \, d\mu &\ge \int_X \limsup_{h \to 0^+} \frac{f_h - f}{h} \, d\mu \\
&\ge \limsup_{h \to 0^+} \int_X \frac{f_h - f}{h} \, d\mu \\
&\ge \liminf_{h \to 0^+} \int_X \frac{f_h - f}{h} \, d\mu \\
&= \liminf_{h \to 0^+} \int_0^{+\infty} \frac{\mu(A_t^h) - \mu(A_t)}{h} \, dt \\
&\ge \int_0^{+\infty} \liminf_{h \to 0^+} \frac{\mu(A_t^h) - \mu(A_t)}{h} \, dt \\
&= \int_0^{+\infty} \mu^+(A_t) \, dt.
\end{aligned}$$

Thus, (3.4) is established for f Lipschitz and bounded. Let now f be an arbitrary Lipschitz function. Let a_n be an increasing sequence of positive numbers such that $\lim_{n \to +\infty} a_n = +\infty$, and such that the sets $\{x \in X: |f(x)| = a_n\}$ have μ-measure 0, for all n. Let $A_n = \{x \in X: |f(x)| < a_n\}$, and let

$$f_n(x) = \begin{cases} f(x), & \text{if } |f(x)| < a_n, \\ a_n, & \text{if } f(x) \ge a_n, \\ -a_n, & \text{if } f(x) \le -a_n. \end{cases}$$

That is, $f_n(x) = \max\{-a_n, \min\{a_n, f(x)\}\}$, so f_n is also a Lipschitz function (of Lipschitz constant at most c) and thus one can apply (3.4) to f_n which then reads as

$$\int_{A_n} |\nabla f(x)|\, d\mu(x) \geq \int_{-a_n}^{a_n} \mu^+\{x \in X\colon f(x) > t\}\, dt.$$

Finally, apply Tonelli's monotone convergence theorem.

LEMMA 3.5. *For any Borel set $A \subset X$ with $0 < \mu(A) < 1$, there exists a sequence of Lipschitz functions f_n on X with values in $[0,1]$ such that $f_n \to 1_{\mathrm{clos}(A)}$ pointwise, as $n \to \infty$, and*

$$\limsup_{n\to\infty} \mathsf{E}\,|\nabla f_n| \leq \mu^+(A).$$

PROOF. Let $\mathrm{clos}(A)$ denote the closure of a set A. If $\mu(\mathrm{clos}(A)) > \mu(A)$, then by the very definition of μ^+, $\mu^+(A) = +\infty$, so there is nothing to prove. Next, let $\mu(\mathrm{clos}(A)) = \mu(A)$. Since for all $h < h'$, $\mathrm{clos}(A^h) \subset A^{h'}$, one has $\mu^+(A) = \liminf_{h\to 0^+}(\mu(\mathrm{clos}(A^h)) - \mu(A))/h$, hence there exists a sequence $h_n \to 0^+$ such that $(\mu(\mathrm{clos}(A^{h_n})) - \mu(A))/h_n \to \mu^+(A)$. Now take a sequence $c_n \in (0,1)$ such that $c_n \to 0$, and let

$$f_n(x) = \min\left\{1, \frac{d(A^{c_n h_n}, x)}{(1-c_n)h_n}\right\},$$

where $d(A,x) = \inf\{d(a,x)\colon a \in A\}$. This function has Lipschitz seminorm at most 1, hence $\|f_n\|_{\mathrm{Lip}} \leq 1/(1-c_n)h_n$, and therefore,

$$|\nabla f_n(x)| \leq 1/(1-c_n)h_n,$$

for all $x \in X$. Note also that for $x \notin A^{h_n}$, $d(A,x) \geq h_n$, hence by the triangle inequality, $d(A^{c_n h_n}, x) \geq (1-c_n)h_n$, and thus $f_n(x) = 1$. Therefore, $|\nabla f_n| = 0$ on the open set $X \setminus \mathrm{clos}(A^{h_n})$. In a similar way, $|\nabla f_n| = 0$ on the open set $A^{c_n h_n}$. Thus,

$$\mathsf{E}\,|\nabla f_n| \leq \frac{\mu(\mathrm{clos}(A^{h_n})) - \mu(A)}{(1-c_n)h_n} \to \mu^+(A).$$

PROOF OF THEOREM 3.1. The isoperimetric constant $C = Is(\mu)$ is the optimal constant satisfying (3.2) and (3.3) when $N(x) = |x|$, that is, such that

$$C\mathsf{E}|f| \leq \mathsf{E}\,|\nabla f|, \tag{3.7}$$

for all integrable, Lipschitz functions f on X with $m(f) = 0$. Indeed, following an argument of Ledoux [11], and via the co-area inequality, we have from (3.7) that

$$\begin{aligned}\mathsf{E}\,|\nabla f| &\geq \int_{-\infty}^{+\infty} \mu^+(f > t)\, dt \\ &\geq Is(\mu)\int_{-\infty}^{0}(1 - \mu(f > t))\, dt + Is(\mu)\int_0^{+\infty} \mu(f > t)\, dt \\ &= Is(\mu)\mathsf{E}|f|.\end{aligned}$$

Therefore, $C \geq Is(\mu)$. In fact, a simple truncation argument (see [4], Section 4, for example) allows us to extend (3.7) to the (slightly) larger class of all integrable functions f, which are Lipschitz on every ball in X and such that $m(f) = 0$. To derive from (3.7) the converse inequality $C \leq Is(\mu)$, that is, the inequality $C \min(\mu(A), 1 - \mu(A)) \leq \mu^+(A)$, one can assume (as noted above) that $\mu(\mathrm{clos}(A)) = \mu(A)$, take $f_n \to 1_{\mathrm{clos}(A)}$ as in Lemma 3.5 and apply (3.7) to $g_n = f_n - m(f_n)$.

Now, let f be a function bounded, which is Lipschitz on every ball in X, with $m(f) = 0$ and such that $\|f\|_N = 1$, that is, such that $\mathrm{E}N(f) = 1$. Also, and without loss of generality, assume that N is differentiable on the whole real line, with, in particular, $N'(0) = 0$. Let $f_1 = \max(f, 0)$ and $f_2 = \max(-f, 0)$. Then, $m(f_1) = m(f_2) = 0$, and thus $m(N(f_1)) = m(N(f_2)) = 0$. Applying (3.7) to $N(f_1)$ and $N(f_2)$, respectively, gives

$$C\mathrm{E}N(f_1) \leq \mathrm{E}N'(f_1)|\nabla f_1| = \mathrm{E}N'(|f|)|\nabla f| 1_{(f>0)},$$
$$C\mathrm{E}N(f_2) \leq \mathrm{E}N'(f_2)|\nabla f_2| = \mathrm{E}N'(|f|)|\nabla f| 1_{(f<0)}.$$

Therefore,

$$C\mathrm{E}N(f) = C\mathrm{E}N(f_1) + C\mathrm{E}N(f_2) \leq \mathrm{E}N'(|f|)|\nabla f|.$$

Next, applying Lemma 2.1 to $|f|$ and $g = |\nabla f|/\|\nabla f\|_N$ gives

$$\begin{aligned} C\mathrm{E}N(f) &\leq \|\nabla f\|_N \mathrm{E}N'(|f|)g \\ &\leq \|\nabla f\|_N \mathrm{E}N'(|f|)|f| \\ &\leq C_N \|\nabla f\|_N \mathrm{E}N(f). \end{aligned}$$

Hence, $C \leq C_N \|\nabla f\|_N$, and since $\|f\|_N = 1$, (3.2) follows. To get (3.3), it is enough to apply (3.2) to the functions $N_\alpha(x) = N(x)/\alpha$, $\alpha > 0$. Indeed, if $\|f\|_{N_\alpha} \geq 1$, then $\||\nabla f|/\lambda\|_{N_\alpha} \geq 1$, $\lambda = C_N/Is(\mu)$. Equivalently, if $\mathrm{E}N(f) \geq \alpha$, then $\mathrm{E}N(|\nabla f|/\lambda) \geq \alpha$. Theorem 3.1 follows.

REMARK 3.6. The inequalities (3.2) and (3.3) are Poincaré-type inequalities. When, $N(x) = |x|^2$, and since $\|f - \mathrm{E}f\|_2 \leq \|f - m(f)\|_2$, (3.2) gives

$$C\|f - \mathrm{E}f\|_2 \leq \|\nabla f\|_2, \tag{3.8}$$

where $C \geq Is(\mu)/2$. Cheeger was the first to express the optimal constant C in (3.8) in terms of the isoperimetric constant (1.1) and so the inequality $C \geq Is(\mu)/2$ bears his name. Cheeger's inequality has thus been extended in the following way: the optimal constant in $C\|f - m(f)\|_N \leq \|\nabla f\|_N$ is such that

$$C \geq \frac{Is(\mu)}{C_N}. \tag{3.9}$$

For $N(x) = |x|^p$, the inequality (3.9) cannot be improved in terms of the isoperimetric constant. Indeed, taking $\mu = \nu$, (3.9) becomes equality as easily tested with the functions $\exp(\alpha x)$, $\alpha \to 1/p$.

4. Induction.

LEMMA 4.1. *Let $C > 0$ be such that*

$$\int_X \sqrt{1+f^2}\, d\mu \le \int_X \sqrt{1+C^2|\nabla f|^2}\, d\mu, \tag{4.1}$$

for all Lipschitz functions f on X with $m(f) = 0$. Then,

$$(\mu^n)^+(A) \ge \frac{2}{\sqrt{6}C}\mu^n(A)(1-\mu^n(A)), \tag{4.2}$$

for all Borel sets $A \subset X^n$.

PROOF. If for all $x \in X$, $0 \le f(x) \le a$, then $|f(x) - m(f)| \le a$ and

$$\sqrt{1+(f-m(f))^2} \ge 1 + K(a)(f-m(f))^2,$$

where $K(a) = (\sqrt{1+a^2}-1)/a^2$ is the optimal constant K satisfying $\sqrt{1+t^2} \ge 1 + Kt^2$, for all $|t| \le a$. Therefore,

$$\int_X \sqrt{1+(f-m(f))^2}\, d\mu \ge 1 + K(a)\int_X (f-m(f))^2\, d\mu$$
$$\ge 1 + K(a)\,\mathrm{Var}(f).$$

Thus, from (4.1),

$$1 + K(a)\,\mathrm{Var}(f) \le \int_X \sqrt{1+C^2|\nabla f|^2}\, d\mu, \tag{4.3}$$

for all Lipschitz functions f on X with $0 \le f \le a$. Now, we fix $a > 0$ and prove by induction that, for all Lipschitz functions $f\colon X^n \to [0, a]$,

$$1 + L\,\mathrm{Var}(f) \le \int_{X^n} \sqrt{1+C^2|\nabla f|^2}\, d\mu^n, \tag{4.4}$$

where L is an arbitrary positive number such that

$$L\left(\frac{1+La^2}{4}\right) \le K\left(\frac{a}{1+La^2/4}\right), \qquad L \le K(a). \tag{4.5}$$

To prove this induction step, take a Lipschitz function $f\colon X^{n+1} \to [0, a]$ and introduce the function

$$\alpha(y) = \int_{X^n} f(x, y)\, d\mu^n(x), \qquad y \in X.$$

Clearly, $\alpha\colon X \to [0, a]$ is Lipschitz and

$$|\nabla\alpha(y)| \le \int_{X^n} |\nabla_y f(x, y)|\, d\mu^n(x), \qquad y \in X, \tag{4.6}$$

where $|\nabla_y f|$ is the modulus of gradient with respect to the coordinate y. Now, by our standing assumption, $|\nabla f|^2 = |\nabla_x f|^2 + |\nabla_y f|^2$, for μ^{n+1}-almost all (x, y), and thus by Fubini's theorem,

$$\int_{X^n} \sqrt{1 + C^2|\nabla f|^2}\, d\mu^n(x) = \int_{X^n} \sqrt{(1 + C^2|\nabla_x f|^2) + C^2|\nabla_y f|^2}\, d\mu^n(x), \tag{4.7}$$

for μ-almost all y. The elementary inequality

$$\int \sqrt{u^2 + v^2} \geq \sqrt{\left(\int u\right)^2 + \left(\int v\right)^2} \tag{4.8}$$

applied in (4.7) to $u = \sqrt{1 + C^2|\nabla_x f|^2}$, $v = C|\nabla_y f|$ (keeping the coordinate y fixed for a while) gives

$$\begin{aligned}
&\int_{X^n} \sqrt{1 + C^2|\nabla f|^2}\, d\mu^n(x) \\
&\qquad \geq \sqrt{\left(\int_{X^n} \sqrt{1 + C^2|\nabla_x f|^2}\, d\mu^n(x)\right)^2 + C^2\left(\int_{X^n} |\nabla_y f|\, d\mu^n(x)\right)^2} \\
&\qquad \geq \sqrt{(1 + L\,\mathrm{Var}_x(f))^2 + C^2|\nabla\alpha(y)|^2},
\end{aligned}$$

where the last inequality follows from the induction hypothesis (4.4) as well as (4.6), and where $\mathrm{Var}_x(f)$ is the variance of f with respect to $x \in X^n$. Next we need to estimate this variance in terms of a. Trivially $\mathrm{Var}_x(f) \leq a^2$. However, to improve the final constant, we use following elementary lemma.

LEMMA 4.2. *Let ξ be a random variable such that $0 \leq \xi \leq a$, then $\mathrm{Var}(\xi) \leq a^2/4$.*

PROOF. Fix $c = \mathrm{E}\xi$ and let F be the distribution of ξ. Then $A(F) = \mathrm{Var}(\xi) = \int_0^a x^2\, dF(x) - c^2$ represents a continuous affine functional of F on the convex compact (for the topology of weak convergence) set $M(c)$ of all probability distributions on $[0, a]$ with mean c. Therefore, A attains its maximum on $M(c)$ at an extremal point of $M(c)$. But these extremal points have at most two atoms, that is, they are of the form $F = p\delta_x + (1 - p)\delta_y$, where $0 \leq p \leq 1$, $0 \leq x$, $y \leq a$. For such F, $A(F) = p(1 - p)(x - y)^2 \leq a^2/4$.

Thus, $\mathrm{Var}_x(f) \leq a^2/4$, and since $\sqrt{1 + t^2} - t$ is decreasing in $t > 0$, we get

$$\begin{aligned}
&\int_{X^n} \sqrt{1 + C^2|\nabla f|^2}\, d\mu^n(x) - (1 + L\,\mathrm{Var}_x f) \\
&\qquad \geq \sqrt{(1 + L\,\mathrm{Var}_x(f))^2 + C^2|\nabla\alpha(y)|^2} - (1 + L\,\mathrm{Var}_x f) \\
&\qquad \geq \sqrt{(1 + La^2/4)^2 + C^2|\nabla\alpha(y)|^2} - (1 + La^2/4) \\
&\qquad = (1 + La^2/4)\sqrt{1 + C^2|\nabla\alpha_1(y)|^2} - (1 + La^2/4),
\end{aligned} \tag{4.9}$$

where $\alpha_1 = \alpha/(1 + La^2/4)$. Integrating (4.9) over $y \in X$ and applying (4.3) to α_1 and $a_1 = a/(1 + La^2/4)$ gives

$$\begin{aligned}
&\int_{X^{n+1}} \sqrt{1 + C^2|\nabla f|^2}\, d\mu^{n+1}(x, y) - \int_X (1 + L\,\mathrm{Var}_x(f))\, d\mu(y) \\
&\quad \geq (1 + La^2/4)(1 + K(a_1)\,\mathrm{Var}(\alpha_1)) - (1 + La^2/4) \\
&\quad = (1 + La^2/4) K(a_1)\,\mathrm{Var}(\alpha_1) \\
&\quad = \frac{1}{1 + La^2/4} K(a_1)\,\mathrm{Var}(\alpha).
\end{aligned}$$

In other words,

$$\int_{X^{n+1}} \sqrt{1 + C^2|\nabla f|^2}\, d\mu^{n+1} \geq 1 + L \int_X \mathrm{Var}_x(f)\, d\mu(y) + \frac{K(a_1)\,\mathrm{Var}(\alpha)}{1 + La^2/4}. \tag{4.10}$$

Therefore, to finish the induction process via (4.10), it remains to show that

$$L \int_X \mathrm{Var}_x(f)\, d\mu(y) + \frac{1}{1 + La^2/4} K(a_1)\,\mathrm{Var}(\alpha) \geq L\,\mathrm{Var}(f). \tag{4.11}$$

Putting $\beta(y) = \int_{X^n} f^2(x, y)\, d\mu^n(x)$, we have

$$\begin{aligned}
\mathrm{Var}_x(f) &= \beta(y) - \alpha^2(y), \\
\mathrm{Var}(f) &= \int_X \beta(y)\, d\mu(y) - \left(\int_X \alpha(y)\, d\mu(y)\right)^2,
\end{aligned}$$

and (4.11) becomes

$$L\left(\int_X \beta - \int_X \alpha^2\right) + \frac{1}{1 + La^2/4} K(a_1)\,\mathrm{Var}(\alpha) \geq L\left(\int_X \beta - \left(\int_X \alpha\right)^2\right).$$

In turn, this is equivalent to

$$\frac{K(a_1)\,\mathrm{Var}(\alpha)}{1 + La^2/4} \geq L\,\mathrm{Var}(\alpha),$$

that is, to

$$L\left(\frac{1 + La^2}{4}\right) \leq K\left(\frac{a}{1 + La^2/4}\right),$$

but, by (4.5) this last inequality is true and under this condition (4.4) is proved. Now, from (4.4) using the inequality $\sqrt{1 + t^2} \leq 1 + t$, we obtain

$$L\,\mathrm{Var}(f) \leq C \int_{X^n} |\nabla f|\, d\mu^n,$$

for any Lipschitz function $f\colon X^n \to [0,a]$. That is, for every $f\colon X^n \to [0,1]$, and for all $a > 0$ and $L > 0$ satisfying (4.5),

$$La\,\mathrm{Var}(f) \le C \int_{X^n} |\nabla f|\, d\mu^n. \tag{4.12}$$

Applying (4.12) to a sequence of Lipschitz functions f_k converging pointwise to the indicator function 1_A so that to have $(\mu^n)^+(A) \ge \liminf \int_{X^n} |\nabla f_k|\, d\mu^n$, we get

$$(\mu^n)^+(A) \ge \frac{La}{C}\mu^n(A)(1-\mu^n(A)).$$

It just remains to show that

$$\sup_{L \text{ in (4.5)}} \sup_{a>0} La \ge \frac{2}{\sqrt{6}}.$$

Put $L = w/a$, so that $L(1+La^2/4) \to w^2/4$, that is, $a/(1+La^2/4) \to 4/w$, as $a \to +\infty$. Therefore, (4.5) is fulfilled for all a large enough if

$$\frac{w^2}{4} < K\left(\frac{4}{w}\right), \qquad w < 1, \tag{4.13}$$

since $K(a) \sim 1/a$ as $a \to +\infty$. However, the first inequality in (4.13) is equivalent to $\sqrt{1+(4t)^2} - 1 > 4$ $(t = 1/w)$. In turn, this is equivalent to $t^2 > 3/2$, that is, $w < \sqrt{2/3} = 2/\sqrt{6}$, and so the second inequality of (4.13) holds true.

Finally we get

$$\sup_{a>0} La \ge \sup_{w<2/\sqrt{6}} w = \frac{2}{\sqrt{6}}.$$

Lemma 4.1 is proved.

5. Proof of Theorem 1.1. For the Young function $N(x) = \sqrt{1+x^2} - 1$, we have $C_N = 2$. Now, combine Theorem 3.1 and Lemma 4.1. By (3.3), the inequality (4.1) holds with $C = 2/Is(\mu)$, hence from (4.2),

$$(\mu^n)^+(A) \ge \frac{Is(\mu)}{\sqrt{6}}\mu^n(A)(1-\mu^n(A)) \tag{5.1}$$

$$\ge \frac{Is(\mu)}{2\sqrt{6}} \min(\mu^n(A), 1-\mu^n(A)). \tag{5.2}$$

Therefore, (1.2) follows. Next, applying once more (3.3) to (X^n, d_n, μ^n) we have

$$\mathrm{E}N(f - m(f)) \le \mathrm{E}N\left(\frac{4\sqrt{6}}{Is(\mu)}|\nabla f|\right), \tag{5.3}$$

for any Lipschitz on every ball function f on X^n. If $0 \le f \le 1$, then $|f - m(f)| \le 1$ and so $N(f-m(f)) \ge (f-m(f))^2/3$, therefore (5.3) gives

$$\frac{1}{3}\mathrm{Var}(f) \le \mathrm{E}\sqrt{1 + \left(\frac{4\sqrt{6}}{Is(\mu)}\right)^2 |\nabla f|^2} - 1. \tag{5.4}$$

Now, note that for all $x \in \mathbb{R}$, $\sqrt{1+4x^2} - 1 \le 2\min(|x|, x^2)$. Hence, the right-hand side of (5.4) is estimated by

$$2\mathrm{E}\min\left(\frac{2\sqrt{6}}{Is(\mu)}\,|\nabla f|,\ \left(\frac{2\sqrt{6}}{Is(\mu)}\right)^2 |\nabla f|^2\right)$$
$$\le 48\,\mathrm{E}\min\left(\frac{1}{Is(\mu)}\,|\nabla f|,\ \frac{1}{Is^2(\mu)}\,|\nabla f|^2\right).$$

REMARK 5.1. Of course, (5.1) is a bit better than (5.2). In its functional form, (5.1) reads as (1.5) and this is why we claimed in Theorem 1.2 that (1.5) holds with $K = Is(\mu)/(2\sqrt{6})$ which is the same as the estimate for $Is(\mu^n)$ in (1.2). Most likely, the multiplicative constant $C = 1/(2\sqrt{6})$ is not optimal. Anyhow, in order to satisfy (1.2) for all measures μ, it has to be less than 1. For individual measures, the optimal constant C_μ in (1.2) depends on μ and satisfies $C_\mu \le 1$. When μ is Gaussian, we have $C_\mu = 1$, as seen from the isoperimetric inequality in Gaussian space. We do not know if there exist other probability distributions with this property.

REMARK 5.2. It is clear from the proof of Theorem 1.1 that for triples (X_i, d_i, μ_i), $i = 1, \dots, m$, satisfying the hypotheses of Theorem 1.1, (1.2) takes the form

$$Is(\mu_1 \otimes \cdots \otimes \mu_m) \ge \frac{1}{2\sqrt{6}} \min_{1 \le i \le m} Is(\mu_i).$$

The converse inequality,

$$Is(\mu_1 \otimes \cdots \otimes \mu_m) \le \min_{1 \le i \le m} Is(\mu_i),$$

is trivial. So, for product measures μ^n, our results give Cheeger and Buser-type inequalities (see [11]) independent of the dimension.

REMARK 5.3. It is clear that the above upper bound holds more generally for a probability measure μ on X^n and its marginals μ_i, $i = 1, 2, \dots, n$. Such is not the case of the lower bound. In $\mathbb{R}^n$, let C be a closed curve of length ℓ and let μ be the uniform distribution on C. Then it is easily seen that $Is(\mu) = 4/\ell$. Now note that the marginals could be arcsin distributions on $[-1, 1]$ while ℓ could be arbitrarily large. Indeed, given integers $a_1, \dots, a_n$, let $x(t) = (\sin a_1 t, \dots, \sin a_n t)$, $-\pi/2 \le t \le \pi/2$, and let μ be the distribution of the function $x(t)$ with respect to the uniform distribution on $[-\pi/2, \pi/2]$. Then, ℓ is of order $\sqrt{a_1^2 + \cdots + a_n^2}$ which can arbitrarily large even for $n \ge 2$ fixed.

6. Proof of Theorems 1.2 and 1.3.

PROOF OF THEOREM 1.3. Let

$$K(\mu) = \operatorname*{ess\,inf}_{a<x<b} \frac{p(x)}{\min(F(x), 1 - F(x))},$$

where, as in Section 1, $a = \inf\{x\colon F(x) > 0\}$, $b = \sup\{x\colon F(x) < 1\}$, and p is a density of the absolutely continuous (with respect to the Lebesgue measure) part of μ. Clearly, $K(\mu)$ does not depend on the choice of p, and $K(\mu) \geq Is(\mu)$. Indeed, taking $A(x) = (-\infty, x]$ with $x \in (a, b)$, we obtain

$$\mu^+(A(x)) = \liminf_{\varepsilon \to 0^+} \frac{F(x+\varepsilon) - F(x)}{\varepsilon} \equiv p_+(x);$$

therefore,

$$Is(\mu) \leq \frac{\mu^+(A(x))}{\min(\mu(A(x)), 1 - \mu(A(x)))} = \frac{p_+(x)}{\min(F(x), 1 - F(x))}.$$

It remains to take the ess inf over all $x \in (a, b)$, noting that p_+ is a version of the density of the absolutely continuous part of μ. In particular, we obtain $K(\mu) = Is(\mu)$, if $K(\mu) = 0$. To establish (1.6), we thus need to prove the converse inequality $K(\mu) \leq Is(\mu)$ assuming that $K(\mu) > 0$.

So, let us assume that $K(\mu) > 0$. Since for almost all $x \in (a, b)$,

$$K(\mu) \min(F(x), 1 - F(x)) \leq p(x),$$

we have in particular that $p(x) > 0$ almost everywhere on (a, b) and therefore F is strictly increasing on (a, b). Hence, the minimal quantile function

$$F^{-1}(p) = \min\{x\colon F(x) \geq p\}$$

is nondecreasing, continuous on $(0, 1]$ and takes values in $(a, b]$. We extend this function to $[0, 1]$ by putting $F^{-1}(0) = a$.

Let F_ν be the distribution function of the measure ν with density $p_\nu(x) = 2^{-1} \exp(-|x|)$, and let $F_\nu^{-1}\colon [0, 1] \to [-\infty, +\infty]$ be the inverse of F_ν. As noted just after Theorem 1.3, and without any assumption on μ, there exists only one nondecreasing left-continuous function $U\colon \mathrm{R} \to \mathrm{R}$ which transforms ν into μ. Indeed, such a function should satisfy the equality $\nu(\{t\colon U(t) \leq x\}) = F(x)$; that is for all x, the set $\{t\colon U(t) \leq x\}$ is the interval $(-\infty,\ F_\nu^{-1}(F(x))]$. Hence for all t and x real, the two inequalities $U(t) \leq x$ and $t \leq V(x)$, where

$$V(x) = F_\nu^{-1}(F(x)),$$

are equivalent. Therefore, at the point $t = V(x)$, the only way to define U is to put $U(t) = x$. For the other points, then necessarily $U = x$ on the intervals $(V(x^-), V(x))$, $U = a$ on $(-\infty, V(a)]$ and $U = b$ on $[V(b), +\infty)$. Thus, U is unique and can be expressed as

$$U(x) = F^{-1}(F_\nu(x)).$$

In the next step, we show that

$$\frac{1}{\|U\|_{\mathrm{Lip}}} \geq K(\mu). \tag{6.1}$$

By Lebesgue theorem, F is almost everywhere differentiable, with $F'(x) = p(x)$ almost everywhere on (a, b). Therefore, differentiating the function V and noting that $p_\nu(F_\nu^{-1}(p)) = \min(p, 1-p)$, we get

$$V'(x) \geq \frac{p(x)}{\min(F(x), 1 - F(x))} \geq K(\mu),$$

for almost all $x \in (a, b)$. Therefore, for all $a \leq x < y \leq b$,

$$V(y) - V(x) \geq \int_x^y V'(t)\, dt \geq K(\mu)(y - x).$$

Let $\mathrm{Im}(F) = \{F(x) > 0\colon x \in \mathbf{R}\}$. For all $p \in \mathrm{Im}(F)$, one always has $F(F^{-1}(p)) = p$, hence $V(U(x)) = x$, whenever $F_\nu(x) \in \mathrm{Im}(F)$. Noting that U takes values in $[a, b]$, we thus see that for all $x < y$,

$$y - x = V(U(y)) - V(U(x)) \geq K(\mu)(U(y) - U(x)), \tag{6.2}$$

provided that $F_\nu(x)$ and $F_\nu(y) \in \mathrm{Im}(F)$. However, (6.2) remains true for all $x < y$, since the function F^{-1} is continuous and constant on the intervals of the form $[F(z^-), F(z)]$. We have thus proved (6.1).

We now apply (3.7) to the measure ν: since $C = Is(\nu) = 1$, we have

$$\mathrm{E}|f(\xi)| \leq \mathrm{E}|f'(\xi)|, \tag{6.3}$$

where the random variable ξ has distribution ν, and where f is an arbitrary locally Lipschitz function on the real line such that $\mathrm{E}|f(\xi)| < +\infty$, and $m(f(\xi)) = 0$. Put $\eta = U(\xi)$ and apply (6.3) to $f = g(U)$: since $|f'(\xi)| \leq \|U\|_{\mathrm{Lip}} |g'(U)|$, we obtain for the random variable $\eta = U(\xi)$ that

$$\frac{1}{\|U\|_{\mathrm{Lip}}}\, \mathrm{E}|g(\eta)| \leq \mathrm{E}|g'(\eta)|. \tag{6.4}$$

Hence by (6.1),

$$K(\mu)\mathrm{E}|g(\eta)| \leq \mathrm{E}|g'(\eta)|, \tag{6.5}$$

which holds for any locally Lipschitz function g such that $\mathrm{E}|g(\eta)| < +\infty$, and $m(g(\eta)) = 0$. Since η has distribution μ, again by (3.7), the optimal constant in (6.5) in front of $\mathrm{E}|g(\eta)|$ is $Is(\mu)$, and thus we conclude that $Is(\mu) \geq K(\mu)$. As a result

$$K(\mu) = Is(\mu). \tag{6.6}$$

With a similar reasoning, since, again by (3.7), the optimal constant in (6.4) in front of $\mathrm{E}|g(\eta)|$ is $Is(\mu)$, we get

$$\frac{1}{\|U\|_{\mathrm{Lip}}} \leq Is(\mu). \tag{6.7}$$

Comparing (6.7) with (6.1) via (6.6), one obtains

$$\frac{1}{\|U\|_{\mathrm{Lip}}} = Is(\mu). \tag{6.8}$$

To complete the proof, it remains to establish (6.8) when $K(\mu) = Is(\mu) = 0$, that is, we need to show that in this case $\|U\|_{\mathrm{Lip}} = +\infty$. Otherwise, $\|U\|_{\mathrm{Lip}} < +\infty$, and one can mimic the reasoning from (6.3) to (6.4), and obtain according to (6.7) that $Is(\mu) > 0$. Theorem 1.3 is proved.

REMARK 6.1. Repeating the step from (6.3) to (6.4), it is easy to see that (6.7) holds for any function U with $\|U\|_{\mathrm{Lip}} < +\infty$, which transforms ν into μ, without the nondecreasing assumption on U. As a result, we conclude that the property (iii) in Theorem 1.2 implies $Is(\mu) > 0$.

PROOF OF THEOREM 1.2. As one can note, we have just proved the equivalence of (ii), (iii) and (iv). In addition, (i) implies (ii) is trivial. To obtain (i), assume that $Is(\mu) > 0$. Let U: R $\to$ R be the nondecreasing function which transforms ν into μ, so $\|U\|_{\mathrm{Lip}} < +\infty$. Let $\xi = (\xi_1, \ldots, \xi_n)$ be a random vector with distribution ν^n, so that $\eta = (U(\xi_1), \ldots, U(\xi_n))$ has law μ^n. By Theorem 1.1 (remembering that $Is(\nu) = 1$), we have $Is(\nu^n) \geq 1/(2\sqrt{6})$; therefore, as shown in Section 3, the measure ν^n satisfies the inequality

$$\frac{1}{2\sqrt{6}} \mathrm{E}\,|f(\xi)| \leq \mathrm{E}|\nabla f(\xi)|, \tag{6.9}$$

where f is an arbitrary locally Lipschitz function on R^n such that $\mathrm{E}|f(\xi)| < +\infty$ and $m(f(\xi)) = 0$. Hence, applying (6.9) to $f = g(U)$ with g locally Lipschitz and noting [recall (6.8)] that

$$|\nabla g \circ U| \leq \|U\|_{\mathrm{Lip}}\, |(\nabla g)(U)| = \frac{1}{Is(\mu)}\, |(\nabla g)(U)|,$$

we obtain

$$\frac{Is(\mu)}{2\sqrt{6}}\ \mathrm{E}\,|g(\eta)| \leq \mathrm{E}\,|\nabla g(\eta)|, \tag{6.10}$$

for any locally Lipschitz function g such that $\mathrm{E}|g(\eta)| < +\infty$ and $m(g(\eta)) = 0$. Approximating the indicator function 1_A of a Borel set $A \subset \mathrm{R}^n$ by a sequence of Lipschitz functions g_k as in Lemma 3.5, it follows from (6.10) that

$$(\mu^n)^+(A) \ \geq\ \frac{Is(\mu)}{2\sqrt{6}} \min(\mu^n(A), 1 - \mu^n(A)),$$

hence $2\sqrt{6} Is(\mu^n) \geq Is(\mu)$. Now use (5.1) for the measure ν and repeat the above argument: for any locally Lipschitz function f on R^n such that $\mathrm{E}|f(\xi)| < +\infty$ we have

$$\frac{1}{2\sqrt{6}} \mathrm{E}\,|f(\xi) - \mathrm{E}\, f(\xi)| \leq \mathrm{E}|\nabla f(\xi)|. \tag{6.11}$$

Indeed, (6.11) is fulfilled for all the locally Lipschitz functions if and only if it is fulfilled for indicator functions (in an asymptotic sense; see [4]). But for indicator functions, (6.11) becomes

$$(\nu^n)^+(A) \geq \frac{1}{\sqrt{6}}\,\nu^n(A)(1-\nu^n(A)),$$

which is a particular case of (5.1) since $Is(\nu) = 1$. Applying (6.11) to the functions $f = g(U)$, we get as above that

$$\frac{Is(\mu)}{2\sqrt{6}}\,\mathsf{E}\,|g(\eta) - \mathsf{E}\,g(\eta)| \leq \mathsf{E}\,|\nabla g(\eta)|. \tag{6.12}$$

Now recall that (6.12) is just (1.5) with $K = Is(\mu)/(2\sqrt{6})$. Theorem 1.2 is thus proved.

7. Poincaré type inequalities. Here, Theorem 1.1 is used to obtain the statements of Theorem 3.1 in the n-dimensional space (X^n, d_n, μ^n) under the more natural assumption $\mathsf{E}f = 0$. Again, let N satisfy the same hypothesis as before, let $\|\cdot\|_N$ denote the norm in the Orlicz space $L_N(X^n, \mu^n)$ and let E be the expectation with respect to μ^n.

THEOREM 7.1. *For any Lipschitz on every ball function f on X^n with $\mathsf{E}f = 0$,*

$$\|f\|_N \leq \frac{4\sqrt{6}C_N}{Is(\mu)}\|\nabla f\|_N. \tag{7.1}$$

In particular,

$$\mathsf{E}N(f) \leq \mathsf{E}N\left(\frac{4\sqrt{6}C_N}{Is(\mu)}|\nabla f|\right). \tag{7.2}$$

PROOF. On $X^n \times X^n$, let $g(x, y) = f(x) - f(y)$, $x, y \in X$. Since $m(g) = 0$ with respect to μ^{2n}, applying Theorem 3.1 and (1.2) gives

$$\|g\|_N \leq \frac{2\sqrt{6}C_N}{Is(\mu)}\,\|\nabla g\|_N,$$

where now, $\|\cdot\|_N$ denotes the norm in $L_N(X^{2n}, d_{2n}, \mu^{2n})$. Since $|\nabla g(x, y)| = \sqrt{|\nabla f(x)|^2 + |\nabla f(y)|^2} \leq |\nabla f(x)| + |\nabla f(y)|$, we get $\|g\|_N \leq 2\|f\|_N$. Thus,

$$\|g\|_N \leq \left\|\frac{4\sqrt{6}C_N}{Is(\mu)}\,|\nabla f|\right\|_N. \tag{7.3}$$

Applying (7.3) to the functions $N_\alpha(t) = N(t)/\alpha$, $t \in \mathbb{R}$, $\alpha > 0$, one easily obtains

$$\int_{X^n}\int_{X^n} N(g(x, y))\,d\mu^n(x)\,d\mu^n(y) \leq \int_{X^n} N\left(\frac{4\sqrt{6}C_N}{Is(\mu)}\,|\nabla f|\right)d\mu^n. \tag{7.4}$$

However, by the convexity of N, $\int_{X^n}\int_{X^n} N(f(x)-f(y)) \geq \int_{X^n} N(f(x)-\int_{X^n} f)$. This gives (7.1). In turn, applying (7.1) to the functions N_α gives (7.2) and the theorem is proved.

REMARK 7.2. If $N(x) = |x|^p$, $p \geq 1$, then $C_N = p$, and (7.1) becomes

$$\|f - \mathsf{E}f\|_p \leq \frac{4\sqrt{6}}{Is(\mu)} p \|\nabla f\|_p, \tag{7.5}$$

where the constant is of sharp order in p. This can be tested for the measure ν on the function $f(x) = \exp(\alpha x)$, $x \in \mathsf{R}$, letting $\alpha \to 1/p$. When μ is Gaussian on $X = \mathsf{R}$, (7.3) with a better constant (of order $\sqrt{p}$) can be found in Pisier [14].

8. Khintchine–Kahane type inequalities. Let $\xi_1, \ldots, \xi_n$, be i.i.d. random variables on the real line R with $\mathsf{E}\xi_1 = 0$, $\xi_1 \neq 0$, a.s. and with a law μ such that $Is(\mu) > 0$. As already noted (Theorem 1.2), this last condition implies that μ has a finite exponential moment.

Let N be a Young function such that

$$K_N = \|\Lambda\|_N < +\infty,$$

where Λ is a random variable which has a double exponential distribution.

THEOREM 8.1. *There exists a finite positive constant $C = C(N, \mu)$ such that for any Banach space $(B, \|\cdot\|_B)$ and vectors $v_1, \ldots, v_n \in B$,*

$$\|S\|_N \leq C\|S\|_1, \tag{8.1}$$

where $S = \|\xi_1 v_1 + \cdots + \xi_n v_n\|_B$.

In (8.1), one can take

$$C = 2 + \frac{8\sqrt{3}K_N}{Is(\mu)\mathsf{E}|\xi_1|}.$$

When $N(x) = |x|^p$ and μ is Gaussian, (8.1) is well known (see e.g., [14], page 179). Under the above general conditions, (8.1) might also be known for the Lebesgue norms and with possibly a different constant. Indeed, it can easily be obtained from Talagrand's inequality (1.9), since in the crucial inequality (8.3) below, only the large values of h are important.

PROOF. The inequality (1.2),

$$(\mu^n)^+(A) \geq \frac{Is(\mu)}{2\sqrt{6}} \min(\mu^n(A), 1 - \mu^n(A)),$$

$A \subset \mathsf{R}^n$, can easily be integrated (see [4]) to give

$$\mu^n(A^h) \geq R_{Kh}(\mu^n(A)), \qquad h > 0, \ K = \frac{Is(\mu)}{2\sqrt{6}}, \tag{8.2}$$

where the function R_h is defined by $R_h(p) = \nu((-\infty, a + h])$, $p = \nu(-\infty, a]$, $a \in \mathsf{R}$. In particular, when $\mu^n(A) \geq 1/2$, (8.2) gives

$$1 - \mu^n(A^{h/K}) \leq \tfrac{1}{2}e^{-h} = \mathsf{P}(\Lambda > h). \tag{8.3}$$

Therefore, for functions $f\colon \mathsf{R}^n \to \mathsf{R}$ with $\|f\|_{\mathsf{Lip}} < +\infty$, applying (8.3) to the sets $A = \{f < t\}$, we have

$$\mu^n(|f - m(f)| > K^{-1}\|f\|_{\mathsf{Lip}} h) \le \mathsf{P}(|\Lambda| > h).$$

By the very definition of the Orlicz norm, this gives

$$\|f - m(f)\|_N \le K^{-1}\|f\|_{\mathsf{Lip}}\|\Lambda\|_N = K^{-1}K_N\|f\|_{\mathsf{Lip}}. \tag{8.4}$$

Now, let $f(x) = \sup \sum_{k=1}^n \langle w, v_k\rangle x_k$, where the sup is taken over the unit ball of B^* the dual of B. The function f is such that

$$\|f\|_{\mathsf{Lip}}^2 \le \sigma^2 = \sup_{\|w\|_{B^*}} \sum_{k=1}^n \langle w, v_k\rangle^2,$$

and moreover it has (with respect to μ^n) the same distribution as S. Hence, (8.4) can be rewritten as

$$\|S - m(S)\|_N \le K^{-1}K_N\sigma. \tag{8.5}$$

To estimate σ via $\mathsf{E}S$, we replace our original argument by a folklore one provided to us by the referee. Let η_i be i.i.d. symmetric Bernoulli random variables independent of the ξ_i's. By the triangle inequality and symmetry,

$$2\mathsf{E}S \ge \mathsf{E}\left\|\sum_i \eta_i(\xi_i - \xi_i')v_i\right\|_B \ge \mathsf{E}\left\|\sum_i \eta_i\xi_i v_i\right\|_B = \mathsf{E}\left\|\sum_i \eta_i|\xi_i|v_i\right\|_B,$$

where $\{\xi_i'\}$ is an independent copy of the sequence $\{\xi_i\}$ and is also independent of the sequence $\{\eta_i\}$. By Jensen's inequality along the ξ_i's,

$$\mathsf{E}\left\|\sum_i \eta_i|\xi_i|v_i\right\|_B \ge \mathsf{E}|\xi_1| E\left\|\sum_i \eta_i v_i\right\|_B.$$

Then, using the usual Khintchine–Kahane inequality with the optimal constant (see [9]), we get

$$\sigma \le \frac{2\sqrt{2}\|S\|_1}{\mathsf{E}|\xi_1|}. \tag{8.6}$$

Finally, the elementary inequality $m(S) \le 2\mathsf{E}S$, as well as (8.5) and (8.6) give

$$\|S\|_N \le m(S) + K^{-1}K_N\sigma \le \left(2 + \frac{2\sqrt{2}K^{-1}K_N}{\mathsf{E}|\xi_1|}\right)\|S\|_1.$$

Acknowledgment. Many thanks to an anonymous referee for his careful reading of the paper and for a suggestion which led to the present, improved version of Theorem 1.2.

REFERENCES

[1] AIDA, S., MASUDA, T. and SHIGEKAWA, I. (1994). Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.* 126 83–101.

[2] BAKRY, D. and LEDOUX, M. (1996). Lévy–Gromov isoperimetric inequality for an infinite dimensional diffusion generator. *Invent. Math.* 123 259–286.
[3] BOBKOV, S. G. (1997). An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space. *Ann. Probab.* 25 206–214.
[4] BOBKOV, S. G. and HOUDRÉ, C. (1995). Some connections between isoperimetric and Sobolev-type inequalities and isoperimetry. *Memoirs AMS.* To appear.
[5] BOROVKOV, A. A. and UTEV, S. A. (1983). On an inequality and a related characterization of the normal distribution. *Probab. Theory Appl.* 28 219–228.
[6] CHEEGER, J. (1970). A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in Analysis: Symposium in Honor of S. Bochner* 195–199. Princeton Univ. Press.
[7] GROMOV, M. and MILMAN, V. (1983). A topological application of the isoperimetric inequality. *Amer. J. Math.* 105 843–854.
[8] GROSS, L. (1993). Logarithmic Sobolev inequalities and contractivity properties of semigroups Dirichlet forms. *Lecture Notes in Math.* 1563 54–88. Springer, Berlin.
[9] LATALA, R. and OLESZKIEWICZ, K. (1994). On the best constant in the Khinchin–Kahane inequality. *Studia Math.* 109 101–104.
[10] LEDOUX, M. (1994a). Isoperimetry and Gaussian analysis. *Ecole d'Été de Probabilités de Saint-Flour. Lecture Notes in Math.* Springer, Berlin. To appear.
[11] LEDOUX, M. (1994b). A simple analytic proof of an inequality by P. Buser. *Proc. Amer. Math. Soc.* 121 951–959.
[12] MAUREY, B. (1991). Some deviation inequalities. *Geom. and Funct. Anal.* 1 188–197.
[13] MUCKENHOUPT, B. (1972). Hardy's inequality with weights. *Studia Math.* 44 31–38.
[14] PISIER, G. (1986). Probabilistic methods in the geometry of Banach spaces. *Probability and Analysis. Lecture Notes in Math.* 1206 167–241. Springer, Berlin.
[15] TALAGRAND, M. (1991). A new isoperimetric inequality and the concentration of measure phenomenon. *Israel Seminar (GAFA). Lecture Notes in Math.* 1469 94–124. Springer, Berlin.

DEPARTMENT OF MATHEMATICS
SYKTYVKAR UNIVERSITY
167001 SYKTYVKAR
RUSSIA

CENTER FOR APPLIED PROBABILITY
SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30322
E-MAIL: houdre@math.gatech.edu