Probab. Theory Relat. Fields 114, 245–277 (1999)

Probability Theory and Related Fields


# Discrete isoperimetric and Poincaré-type inequalities

**S. G. Bobkov[1,2,⋆], F. Götze[2]**

[1] Department of Mathematics, Syktyvkar University, 167001 Syktyvkar, Russia
[2] Department of Mathematics, Bielefeld University, P.O. Box 100131, D-33501, Bielefeld, Germany. e-mail: goetze@mathematik.uni-bielefeld.de

Received: 30 April 1997 / Revised version: 5 June 1998

**Abstract.** We study some discrete isoperimetric and Poincaré-type inequalities for product probability measures $\mu^n$ on the discrete cube $\{0, 1\}^n$ and on the lattice $\mathbf{Z}^n$. In particular we prove sharp lower estimates for the product measures of 'boundaries' of arbitrary sets in the discrete cube. More generally, we characterize those probability distributions $\mu$ on $\mathbf{Z}$ which satisfy these inequalities on $\mathbf{Z}^n$. The class of these distributions can be described by a certain class of monotone transforms of the two-sided exponential measure. A similar characterization of distributions on $\mathbf{R}$ which satisfy Poincaré inequalities on the class of convex functions is proved in terms of variances of suprema of linear processes.

*Mathematics Subject Classification (1991):* Primary 60E15; Secondary 26D15

## 1. Introduction

Let $(\Omega, \mathbf{P})$ be a product probability space, $\Omega = \Omega_1 \times \cdots \times \Omega_n$, $\mathbf{P} = \mu_1 \otimes \cdots \otimes \mu_n$. In this paper we derive some discrete isoperimetric inequalities of the form

$$\mathbf{P}^+(A) \geq c\, I(\mathbf{P}(A)), \qquad A \subseteq \Omega,\ c > 0\ , \tag{1.1}$$

---

⋆ Research supported in part by Alexander von Humboldt Foundation

*Keywords*: Concentration of measure – Isoperimetry – Poincaré inequalities

connecting suitably defined 'surface' measures $\mathbf{P}^+(A)$ and the 'volume' $\mathbf{P}(A)$ via some non-negative functions $I$ on [0, 1].

Various discrete and continuous isoperimetric inequalities may be written in the form (1.1). For instance for Gauss spaces $(\mathbf{R}^n, \gamma_n)$, i.e., Euclidean spaces equipped with the canonical Gaussian measure with density $(2\pi)^{-n/2}\exp(-|x|^2/2)$, $x \in \mathbf{R}^n$, the surface measure is defined by the so-called lower outer Minkowski $\gamma_n$-content, (or Gaussian perimeter)

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

where $A^h$ denotes the $h$-neighbourhood of $A$ in the Euclidean metric in $\mathbf{R}^n$. Let $\varphi$ denote the density of $\gamma_1$ and let $\Phi^{-1}$ denote the inverse of the standard Gaussian distribution function $\Phi(x) = \gamma_1((-\infty, x])$. The Gaussian isoperimetric inequality due to V.N. Sudakov, B.S. Tsirel'son [ST] and C. Borell [Bor] then asserts that, for any measurable set $A \subset \mathbf{R}^n$,

$$\gamma_n^+(A) \geq I(\gamma_n(A)) \tag{1.2}$$

with the Gaussian isoperimetric function $I(t) = \varphi(\Phi^{-1}(t))$, $t \in [0, 1]$. Here equality holds for an arbitrary half-space $A$. The most remarkable feature of this inequality is that the function $I$ is independent of the dimension $n$.

### *1.1. Product measures on* $\{0, 1\}^n$

In the case of the probability counting measure, say $\mathbf{P}$, on the discrete cube $\Omega = \{0, 1\}^n$ we are looking for suitable surface measures $\mathbf{P}^+$ which approximate the corresponding Gaussian perimeter (1.2) for growing $n$ and satisfy (1.1). To this end it may be helpful to review some measures used in related extremal set problems. Denote by $s_i(x)$ the 'neighbour' of $x \in \{0, 1\}^n$ obtained by changing the $i$-th coordinate: $[s_i(x)]_j = x_j$, for $j \neq i$, and $[s_i(x)]_i = 1 - x_i$. For any $A \subset \{0, 1\}^n$, the sets

$$\partial_+ A = \{x \in A : s_i(x) \notin A, \text{ for some } i \leq n\} \ ,$$

$$\partial_- A = \{x \notin A : s_i(x) \in A, \text{ for some } i \leq n\}$$

represent the inner-respectively the outer boundary of $A$ with respect to the Hamming distance in $\{0, 1\}^n$, or the inner- and the outer parts of the (two-sided) boundary $\partial A = \partial_+ A \cup \partial_- A$.

The classical discrete (vertex-)isoperimetric problem is to minimize $\mathbf{P}(\partial_- A)$ over all sets $A$ with fixed probability, say $\mathbf{P}(A) = t$. It has been solved by L.H. Harper [H2] (cf. also [WW], [FF]). In the case, when $t$ is of the form $2^{-n}\sum_{i\leq k} C_n^i$ , the extremal sets are the balls $A = \{x \in \{0, 1\}^n :$

$x_1 + \cdots + x_n \le k\}$ (which may also be regarded as half-spaces). The corresponding isoperimetric inequalities are of the type (1.1)

$$\mathbf{P}(\partial_- A) \ge \frac{1}{\sqrt{n}} I_n(\mathbf{P}(A)) \tag{1.3}$$

with functions $I_n$ closely related to the Gaussian isoperimetric function $I$. Note however, that these inequalities essentially depend on the dimension $n$ through the factor $1/\sqrt{n}$, and that the connection with (1.2) is not clear. It seems more promising to weight boundary points (or vertices) $x$ of subsets $A$ of $\{0, 1\}^n$ according to the number of outbound edges by introducing the functions

$$\bar{h}_A(x) = \mathrm{card}\{i \le n : (x \in A,\ s_i(x) \notin A) \text{ or } (x \notin A,\ s_i(x) \in A)\},$$

$$h_A = \bar{h}_A \mathbf{1}_A \ .$$

In particular, we may define boundaries and inner boundaries as well via $\partial A = \{\bar{h}_A > 0\}$, $\partial_+ A = \{h_A > 0\}$. For example, one might look at the socalled edge-isoperimetric problem corresponding to minimizing the functional

$$\int \bar{h}_A \, d\mathbf{P}$$

(which is, up to the constant $2^{-n-1}$, just the number of all boundary edges of $A$ ) over all sets $A$ with prescribed value $\mathbf{P}(A) = t$. For the discrete cube, the problem of (edge-) extremal sets was solved by L.H. Harper [H1] and J.H. Lindsey [Lin] (cf. R. Ahlswede and N. Cai [AC] for history of the problem and more general results). The edge-extremal sets are given by the first points in the lexicographic order of $\Omega$. In particular, when $t = 2^{-k}$, these are the subcubes $A = \{(0, \dots, 0)\} \times \{0, 1\}^{n-k}$. For such sets, $\int \bar{h}_A \, d\mathbf{P} = 2t \log_2(1/t)$ while for balls this quantity grows to infinity with $n$ like $\sqrt{n}$. Thus, similar to (1.3), this surface measure does not recover, even up to a multiplicative constant, the continuous isoperimetric inequality (1.2).

An appropriate normalization of $\int \bar{h}_A \, d\mathbf{P}$ by multiplication with $\mathbf{P}(\partial A)$ yields a closer analogue of (1.2). This is due to G.A. Margulis [M] who proved that $\mathbf{P}(\partial A) \int \bar{h}_A \, d\mathbf{P}$ can be bounded below in terms of $\mathbf{P}(A)$, only. More generally, for $p \in (0, 1)$, let $\mu_p^n$ denote the product measure on $\{0, 1\}^n$ of the Bernoulli measure assigning weights $q = 1 - p$ and $p$ to 0 and 1 respectively. Margulis investigated the (connectivity-) threshold behaviour of the function $p \to \mu_p^n(A)$ for monotone sets $A$ such that $h_A$ is large on $\partial_+ A$, by means of the inequality, ([M], Theorem 2.4),

$$\mu_p^n(\partial_+ A) \int h_A \, d\mu_p^n \ge c\,(p, \alpha) \ , \tag{1.4}$$

where $c\,(p,\alpha)$ depends on $p$ and $\alpha = \mu_p^n(A)$ but not on $n$. Using Cauchy-Schwarz' inequality, we have $\mu_p^n(\partial_+ A) \int h_A \geq \left(\int \sqrt{h_A}\right)^2$. Thus the lower bound $c\,(p,\alpha)$ can be derived from lower bounds (which do not involve $n$ directly) for $\int \sqrt{h_A}$. Using the latter as a new surface measure, M. Talagrand [T1] sharpened (1.4), proving that there exists a positive constant $c_p$ such that, for each subset $A$ of $\{0,1\}^n$,

$$\int \sqrt{h_A(x)}\, d\mu_p^n(x) \;\geq\; c_p\, J\,(\mu_p^n(A)\,(1-\mu_p^n(A)))\;, \tag{1.5}$$

where $J(t) = t\sqrt{\log(1/t)}$, $t \geq 0$. He studied as well the related Cheeger-type inequality

$$\int \sqrt{h_A(x)}\, d\mu_p^n(x) \;\geq\; c'_p\, \mu_p^n(A)\,(1-\mu_p^n(A)) \tag{1.6}$$

and applied it to get a quantitative version of Margulis' theorem on threshold behaviour. The 'surface measures' $\int \sqrt{h_A}\, d\mu_p^n$ and $\int \sqrt{\bar{h}_A}\, d\mu_p^n$ turn out to be the desired discrete analogues of the Gaussian perimeter $\gamma_n^+(A)$. Moreover, since $J(t(1-t))$ behaves like the Gaussian isoperimetric function $I(t)$, the inequality (1.5) itself may be viewed, up to a multiplicative absolute constant, as a discrete version of the Gaussian isoperimetric inequality (1.2). Using a different approach, we will study a correct order of constants in (1.5) and (1.6) as functions of parameter $p$ and prove in section 2 (Propositions 2.3–2.4):

**Theorem 1.1.** *The functions $c_p$ resp. $c'_p$ of $p$ in* (1.5) *and* (1.6) *can be chosen as*

$$c_p \approx 1/\sqrt{\log(1/p)}, \quad p \text{ small}, \quad \text{and} \quad c'_p = 1\;.$$

Exact values for $c_p$ are more involved and provided in Lemma 2.2 including more general functions $J$ in (1.5). The value $c'_p = 1$ implies in particular that in Margulis' lower bound we may choose $c(p,\alpha) = \alpha^2(1-\alpha)^2$, independent of $p$.

For the two sided boundary surface measures $\mathbf{P}^+(A) = \int \sqrt{\bar{h}_A}\, d\mu_p^n$, we prove in section 3 (Proposition 3.1):

**Theorem 1.2.** *For all* $p \in (0,1)$,

$$\int \sqrt{\bar{h}_A(x)}\, d\mu_p^n(x) \;\geq\; \frac{1}{\sqrt{2pq}}\, \varphi\big(\Phi^{-1}(\mu_p^n(A))\big), \quad A \subset \{0,1\}^n\;. \tag{1.7}$$

Thus, compared to (1.5) the two sided surface measure is bounded from below by a much more larger quantity, especially, when $p$ is small.

### 1.2. Functional inequalities on general product spaces

The bound (1.7) will be extended in section 3 (Theorem 3.2) to finite products of arbitrary probability spaces under an appropriate adaption of the notion of surface measure. As we believe, this illustrates a very general principle behind the Gaussian isoperimetric inequality. Our approach is based on the idea that the inequalities like (1.5)–(1.7) are particular cases of the following *functional* type inequalities on abstract product spaces $(\Omega, \mathbf{P})$

$$c\, I\left(\int f\, d\mathbf{P}\right) \le \int \sqrt{c^2 I(f)^2 + \mathrm{D}(f)^2}\, d\mathbf{P} \tag{1.8}$$

involving some measure of the size of a 'discrete gradient' $\mathrm{D}(f)$ of an arbitrary (measurable) function $f$ on $\Omega$ with values in [0, 1]. On indicator functions $f = \mathbf{1}_A$, (1.8) turns into an isoperimetric inequality (1.1) for the 'surface measure'

$$\mathbf{P}^+(A) = \int \mathrm{D}(\mathbf{1}_A)\, d\mathbf{P} \;\; . \tag{1.9}$$

For instance, in (1.7) we obtain the surface measure $\mathbf{P}^+(A) = \int \sqrt{\bar{h}_A}\, d\mu_p^n$ defined by (1.9) via the operator $\mathrm{D}f(x) = \sqrt{\sum_{i=1}^n |f(x) - f(s_i(x))|^2}$, i.e., the Euclidean length of the usual discrete gradient of $f$.

Functional inequalities like (1.8) were introduced in [B1] for the discrete cube and the uniform Bernoulli measure. Starting from dimension 1 with an arbitrary given function $I$, these inequalities can be extended by simple induction to the $n$-dimensional case for a wide class of operators D. This reduces the multidimensional isoperimetric problem of estimating of $\mathbf{P}^+(A)$ within all the sets with given (product-)measure $\mathbf{P}(A)$ to a one-dimensional problem about an optimal function $I$ or an optimal constant $c$ in (1.8). Of course, (1.1) is a weaker inequality than (1.8). However, in many cases of interest, (1.8) saves a lot of information about optimal constants in (1.1). This concerns in particular the inequalities (1.5)–(1.7). The induction argument for the general setting is isolated in section 2 (Lemma 2.1).

### 1.3. Product measures on $\mathbf{Z}^n$

For subsets $A$ of $\mathbf{Z}^n$, the boundary function $h_A$ is given by

$$h_A(x) = \mathrm{card}\{i \le n : x \in A,\ (x + e_i \notin A \text{ or } x - e_i \notin A)\} \;\; ,$$

where $(e_i)_{i \le n}$ denotes the canonical basis in $\mathbf{R}^n$. Using the functional form (1.8) with $I(t) = t(1-t)$, we obtain the following result (section 4, Theorem 4.1) which characterizes probability measures $\mu$ on $\mathbf{Z}$ satisfying a

Cheeger-type inequality on $\mathbf{Z}^n$. For the Minkowski perimeter with respect to continous probability measures $\mu^n$ on $\mathbf{R}^n$, an analogous problem of whether or not Cheeger-type inequalities extend from dimension 1 to dimension $n$ with a dimension free constant has been affirmatively settled in [BH1].

Let $F_\mu(x) = \mu((-\infty, x])$ denote the distribution function of $\mu$. Then we have

**Theorem 1.3.** *The following statements are equivalent*:
(1) *For some $c > 0$, the following Cheeger type inequality holds*:

$$\int_{\mathbf{Z}^n} \sqrt{h_A(x)}\, d\mu^n(x) \ \geq\ c\,\mu^n(A)\,(1-\mu^n(A)), \qquad \textit{for all } n \geq 1 \text{ and } A \subset \mathbf{Z}^n \ . \tag{1.10}$$

(2) *For some $c > 0$, the distribution function of $\mu$ satisfies*

$$\mu(\{x\}) \geq c\, F_\mu(x)\,(1 - F_\mu(x)), \qquad \textit{for all } x \in \mathbf{Z} \ . \tag{1.11}$$

(3) *There is a constant $C > 0$ such that a discrete Poincaré inequality on* $\mathbf{Z}$ *holds*:

$$\int_{\mathbf{Z}} f^2\, d\mu \ -\ \left(\int_{\mathbf{Z}} f\, d\mu\right)^2 \leq C \int_{\mathbf{Z}} \ (f(x+1)\ -\ f(x))^2\, d\mu(x), \qquad \textit{for all } f \text{ on } \mathbf{Z} \ . \tag{1.12}$$

(4) *The measure $\mu$ is a functional transform of the two-sided exponential distribution $\nu$, with density $\frac{1}{2}\exp(-|x|)$, $x \in \mathbf{R}$, under a nondecreasing map $U : \mathbf{R} \to \mathbf{Z}$ such that, for some $h > 0$, $\sup_{x\in\mathbf{R}} |U(x+h) - U(x)| \leq 1$.*

Note that (1.11) is just a particular case of (1.10) when $n = 1$ and $A = (-\infty, x] \cap \mathbf{Z}$. Two different characterizations of probability measures $\mu$ on $\mathbf{Z}$ satisfying the discrete Poincaré inequality (1.12) have recently been obtained by J.-H. Lou [Lou] in continuation of an earlier joint work with L.H.Y. Chen [CL].

The properties (1)–(4) in Theorem 1.3 are satisfied for example by the Poisson measures. This important case will be considered separately in section 3 on the basis of the functional form (1.8) for inequality (1.7) with the Gaussian isoperimetric function $I$. Using Poisson's limit theorem, we will obtain an isoperimetric inequality for products of Poisson measures which is very similar to the Gaussian isoperimetric inequality (1.2) (cf. Proposition 3.6).

### *1.4. Poincaré inequalities for probability measures on* $\mathbf{R}$

The discrete Poincaré inequality (1.12) makes sense for an arbitrary probability measure $\mu$ on $\mathbf{R}$, if we restrict it to the class of all monotone functions

$f$ on $\mathbf{R}$. More generally, one may wonder whether or not, for some couple $(C, a)$ of positive numbers, for all monotone $f$,

$$\int_{\mathbf{R}} f^2\, d\mu - \left(\int_{\mathbf{R}} f\, d\mu\right)^2 \le C \int_{\mathbf{R}} (f(x+a) - f(x))^2\, d\mu(x) \ .$$

Up to a constant, this inequality can be shown to be equivalent to the Poincaré inequality

$$\int_{\mathbf{R}} f^2\, d\mu - \left(\int_{\mathbf{R}} f\, d\mu\right)^2 \le C \int_{\mathbf{R}} |f'|^2\, d\mu \tag{1.13}$$

in the class of all convex functions $f$ on $\mathbf{R}$. Such 'convex' Poincaré inequalities are motivated by the study of maximum of random processes of the form

$$x(t) = a_0(t) + \sum_{n=1}^{\infty} a_n(t)\, \xi_n, \quad t \in T \ ,$$

where $(\xi_n)_{n\ge1}$ is a sequence of independent random variables with common distribution $\mu$, and where $a_n$ are arbitrary functions on a parametric set $T$. In analogy with Theorem 1.3 we will prove the following characterization (cf. Theorem 4.2):

**Theorem 1.4.** *For every probability measure $\mu$ on $\mathbf{R}$ with finite second moment, it is equivalent that:*
(1) *There exists a constant $C$ such that, for every bounded process $x(t)$ as above,*

$$\mathrm{Var}\,(\sup_t x(t)) \ \le \ C \sup_t \mathrm{Var}\,(x(t)) \ , \tag{1.14}$$

*where* Var *denotes the variance.*
(2) *There exists a constant $C$ such that* (1.13) *holds for all convex functions $f$ on $\mathbf{R}$.*
(3) *There exist constants $c > 0$ and $h > 0$ such that*

$$F_\mu(x) - F_\mu(x-h) \ge c\, F_\mu(x)\,(1 - F_\mu(x)), \quad \textit{for all } x \in \mathbf{R} \ .$$

(4) *The measure $\mu$ is a functional transform of $\nu$ under a nondecreasing map $U : \mathbf{R} \to \mathbf{R}$ such that $\sup_{x\in\mathbf{R}} |U(x+h) - U(x)| < +\infty$, for some (or, all) $h > 0$.*

For the Gaussian measure $\gamma_n$, the property (1.14) with optimal constant $C = 1$ is well-known as a consequence of the isoperimetric inequality (1.2). It is also a direct consequence of the Gaussian Poincaré inequality, i.e., (1.13) for $\gamma_n$.

Throughout the paper, we will often use probabilistic notations $\mathbf{E}_\mu f = \mathbf{E} f$, $\mathrm{Var}_\mu f = \mathrm{Var} f$ for the mathematical expectation and the variance of a function $f$ with respect to probability measure $\mu$.

## 2. Induction step. Sharp constants in Talagrand's inequalities (1.5) and (1.6)

First we separate the induction step which has been already shown to hold for some special cases in [B1] and [BaL]. For application to different forms of (1.8), we describe some general situation. Let $(\Omega, \mathbf{P})$ be a product probability space, $\Omega = \Omega_1 \times \cdots \times \Omega_n$, $\mathbf{P} = \mu_1 \otimes \cdots \otimes \mu_n$. Assume that, for each $i \leq n$, $\mathrm{D}_i$ denotes an operator acting on the set of all bounded measurable functions on $\Omega_i$. Set, for any function $f$ on $\Omega$,

$$\mathrm{D}(f) = \sqrt{\sum_{i=1}^{n} \mathrm{D}_i(f_i)^2}$$

where $f_i$ denotes the function of $x_i$ defined via $f_i(x_i) = f(x)$, $x \in \Omega$. The basic assumption on $\mathrm{D}_i$ is that

(a) for any bounded measurable function $f$ on $\Omega_i$, $\mathrm{D}_i(f)$ represents a non-negative measurable function on $\Omega_i$;

(b) $\mathrm{D}_i(f) = \mathrm{D}_i(f')$ $\mu_i$-almost surely, if $f = f'$ $\mu_i$-almost surely;

(c) $\mathrm{D}_i$ is 'convex' in the following sense: for any probability space $(Y, \lambda)$ and for any bounded measurable function $f$ on the product probability space $(\Omega_i \times Y, \mu_i \otimes \lambda)$, the function $y \to \mathrm{D}_i(f(x_i, y))$ is $\lambda$-measurable and, for $\mu_i$-almost all $x_i$,

$$\mathrm{D}_i(\mathbf{E}_\lambda f(x_i, y)) \ \leq \ \mathbf{E}_\lambda \mathrm{D}_i(f(x_i, y)) \ .$$

Under an appropriate continuity assumption, the above property will follow from inequality $\mathrm{D}_i(tu + (1-t)v) \ \leq \ t\mathrm{D}_i(u) + (1-t)\mathrm{D}_i(v)$, $0 \leq t \leq 1$.

**Lemma 2.1.** *Let $I$ be a non-negative, Borel measurable function on* $[0, 1]$. *For each $i \leq n$, assume that* $\mathrm{D}_i$ *satisfies the inequality*

$$I(\mathbf{E}_{\mu_i} f) \ \leq \ \mathbf{E}_{\mu_i} \sqrt{I(f)^2 + \mathrm{D}_i(f)^2} \ ,$$

*for every bounded measurable function $f$ on $\Omega_i$. Then* D *satisfies the same inequality*

$$I(\mathbf{E} f) \ \leq \ \mathbf{E} \sqrt{I(f)^2 + \mathrm{D}(f)^2} \ ,$$

*for any bounded measurable function $f$ on* $\Omega$ (*with mathematical expectations taken with respect to* $\mathbf{P}$ ).

The proof does not differ from that of Lemma 2 in [B1] and is omitted. The statement of Lemma 2.1 remains also true in a more general situation, when, for example, the above inequalities are considered for smaller classes of functions closed under the operation $x_i \to \mathbf{E}_\lambda f(x_i, y)$ (for $f$ as in $c$)).

Now, in order to obtain inequalities (1.5)–(1.6) for product measures on $\Omega = \{0, 1\}^n$, it suffices to establish 'one-dimensional' functional inequalities (1.8) for the space $(\{0, 1\}, \mu_p)$ with $I(t) = J(t(1-t))$ and $I(t) = t(1-t)$, respectively, and for $\mathrm{D}f = (f(1) - f(0))^+$, where we write $a^+ = \max(a, 0)$. Recall that $\mu_p$ assigns the value $p \in [0, 1]$ to 1 and $q = 1 - p$ to 0. Following M. Talagrand [T1], for any function $f$ on $\{0, 1\}^n$, define

$$Mf(x) = \sqrt{\sum_{i=1}^{n} ((f(x) - f(s_i(x))^+)^2} \ ,$$

so that $M\mathbf{1}_A = \sqrt{h_A}$. Such operators were used in [T1] in discrete Sobolev-type inequalities $c_p\, I(\mathrm{Var}_{\mu_p^n}(f)) \le \int Mf\, d\mu_p^n$ to establish (1.5)–(1.6) with $c_p' = \sqrt{2}\min(p, q)/\sqrt{pq}$, $c_p = c_p'/K$ where $K$ is universal. In addition, on the class of monotone sets $A \subset \{0, 1\}^n$, i.e., such that $\mathbf{1}_A$ is a non-decreasing function in each coordinate, the inequality (1.6) was proved with a better constant $c_p'' = \sqrt{2p/q}$. The advantage of (1.6) over (1.5) was that the constant $c_p''$ is explicit, and the form (1.6) is better adapted for applications to Margulis' theorem about threshold behaviour of functions $p \to \mu_p^n(A)$.

With every function $I$ on $[0, 1]$, we connect a constant

$$d_p(I) \ = \sup_{1 \ge a > b \ge 0} \frac{((I(pa + qb) - qI(b))^+)^2 - p^2 I(a)^2}{p^2(a-b)^2} \ . \tag{2.1}$$

Applying Lemma 2.1 to $\mathrm{D}f = Mf$, we immediately obtain:

**Lemma 2.2.** *Let $I$ be a non-negative function on* $[0, 1]$ *such that* $I(0) = I(1) = 1$. *With respect to the measure $\mu_p^n$ on* $\{0, 1\}^n$, *the optimal constant $c$ in the inequality*

$$c\, I(\mathbf{E} f) \ \le \ \mathbf{E}\sqrt{c^2 I(f)^2 + M(f)^2} \ , \tag{2.2}$$

*for an arbitrary function $f$ on* $\{0, 1\}^n$, *is given by* $c = \min\left\{\frac{1}{\sqrt{d_p(I)}}, \frac{1}{\sqrt{d_q(I)}}\right\}$. *Thus,*

$$\int \sqrt{h_A(x)}\, d\mu_p^n(x) \ \ge \ c\, I(\mu_p^n(A)) \ , \tag{2.3}$$

*for any subset $A$ of* $\{0, 1\}^n$. *In addition, for the class of monotone functions $f$ (resp., for the class of monotone sets $A$)* (2.2)–(2.3) *hold with* $c = 1/\sqrt{d_p(I)}$.

Indeed, it suffices to consider the case $n = 1$. Setting $f(1) = a$, $f(0) = b$, $d = 1/c^2$ and assuming $a \geq b$, (2.2) may be written as

$$I(pa + qb) \leq p\sqrt{I(a)^2 + d(a-b)^2} + qI(b) \tag{2.4}$$

which is equivalent to $d \geq d_p(I)$. Thus, $c = 1/\sqrt{d_p(I)}$ is optimal in (2.2) for the class of monotone functions $f$. Repeating this argument for the case $a < b$ gives an equivalent condition $d \geq \max\{d_p(I), d_q(I)\}$ for (2.2) to hold with arbitrary $f$.

Clearly, the relations (2.2)–(2.3) with $c = \inf_p c_p(I)$ remain true for product measures on $\{0, 1\}^n$ with arbitrary marginal distributions.

In the particular case $I(t) = t(1-t)$, it is easy to compute $d_p(I)$: the expression under the sup in (2.1) is just $q((a-b)^2 + 2p(a - a^2))$, and its maximum is attained for $b = 0$, $a = 1/(2-q)$. This gives $d_p(I) = q/(2-q) \leq 1$ and leads to the following:

**Proposition 2.3.** *For any product measure* $\mathbf{P}$ *on* $\{0, 1\}^n$ *with possibly different marginal distributions, and for any subset* $A$ *of* $\{0, 1\}^n$,

$$\int \sqrt{h_A(x)}\, d\mathbf{P}(x) \ \geq\ \mathbf{P}(A)\,(1 - \mathbf{P}(A)) \ . \tag{2.5}$$

It might be worthwile to note that the optimal function $I$ in $\int \sqrt{h_A(x)}\, d\mathbf{P}(x) \geq I(\mathbf{P}(A))$ should satisfy $I(p) \leq p$. This follows from the particular case $n = 1$, $\mathbf{P} = \mu_p$, $A = \{1\}$. Hence, the right-hand side of (2.5) is of correct order when $\mathbf{P}(A)$ is close to 0. On the other hand, by Lemma 2.2, (2.5) can be sharpened for $\mathbf{P} = \mu_p^n$ and for monotone $A$: it holds with constant $\sqrt{(2-q)/q}$ in front of $\mathbf{P}(A)\,(1 - \mathbf{P}(A))$. As shown in [T1], the connection of a (2.5)-type inequality with Margulis' theorem is the following. Using the identity $\frac{d}{dp}\,\mu_p^n(A) = \frac{1}{p}\int h_A\, d\mu_p^n$, where $A$ is an arbitrary monotone set (Margulis–Russo's Lemma), one obtains from (2.5) that

$$\frac{d}{dp}\,\mu_p^n(A) \geq \frac{\sqrt{k}}{p}\,\mu_p^n(A)(1 - \mu_p^n(A)) \ ,$$

provided $h_A \geq k$ on $\partial_+ A$. Hence, for $\varepsilon \in (0, 1/2)$, if $\mu_{p_1}^n(A) = \varepsilon$, $\mu_{p_2}^n(A) = 1 - \varepsilon$,

$$2\log\frac{1-\varepsilon}{\varepsilon} = \int_{p_1}^{p_2} \frac{d}{dp}\,\log\frac{\mu_p^n(A)}{1 - \mu_p^n(A)}\,dp \geq \int_{p_1}^{p_2} \frac{\sqrt{k}}{p}\,dp = \sqrt{k}\log\frac{p_2}{p_1} \ .$$

Thus,

$$\log\frac{p_2}{p_1} \leq \frac{1}{\sqrt{k}}\log\frac{1-\varepsilon}{\varepsilon} \ . \tag{2.6}$$

Since $p_2 - p_1 \le \log \frac{p_2}{p_1}$, the function $\mu_p^n(A)$ jumps from $\varepsilon$ to $1-\varepsilon$ on an interval of the length at most $K(\varepsilon)/\sqrt{k}$ with $K(\varepsilon)$ of order $\log \frac{1-\varepsilon}{\varepsilon}$. The last is a quantitative version of Margulis' theorem due to M. Talagrand [T1] (cf. also [T2] for related results). And the relation (2.6) gives a little more information, especially when $p_1$ is small.

Let us return to the inequalities (2.3). For many functions $I$, it seems difficult to compute $d_p(I)$ exactly on the basis of (2.1) but its behaviour can be explored in some special cases. But even in this case, one can not guarantee that the constant $c = \min\{1/\sqrt{d_p(I)}, 1/\sqrt{d_q(I)}\}$ will be optimal (or, asymptotically optimal) in the inequality (2.3) which is weaker than its functional form (2.2). Nevertheless, for $I(t) = J(t(1-t))$, where $J(t) = t\sqrt{\log(1/t)}$ is as in (1.5), we have the following:

**Proposition 2.4.** *Let* $I(t) = J(t(1-t))$. $t \in [0,1]$. *For some positive* $K_0$ *and* $K_1$, *and for all* $p \in (0,1)$, *the optimal constant* $c$ *in* (2.3) *as well as the optimal constant* $c$ *in* (2.2) *satisfies*

$$\frac{K_0}{\sqrt{\log(1/(pq))}} \le c \le \frac{K_1}{\sqrt{\log(1/(pq))}} \ .$$

*Proof*. Fix $p \in (0,1)$ and $q = 1-p$. Since (2.2) is stronger than (2.3), and therefore the optimal constant $c$ in (2.2) is not grater than the optimal constant $c$ in (2.3), we need only to prove the upper estimate for $c$ in (2.3) and the lower estimate for $c$ in (2.2), for one dimension in the last case. To obtain the upper estimate, let us test (2.3) on the sets $A = \{S_n = 0\}$ where $S_n = x_1 + \cdots + x_n$. Since $h_A = n\mathbf{1}_{\partial_+ A}$, (2.3) becomes

$$\sqrt{n} q^n \ \ge \ c\, q^n (1-q^n) \sqrt{\log \frac{1}{q^n(1-q^n)}} \ .$$

Letting $n \to \infty$, we obtain $c \le 1/\sqrt{\log(1/q)}$. Testing (2.3) on the sets $A = \{S_n = n\}$, we arrive at $c \le 1/\sqrt{\log(1/p)}$. Both estimates give $c \le 1/\sqrt{\log(1/p')}$ where $p' = \min(p,q)$. But $p'$ is of order $pq$.

The proof of the lower bound is rather routine, and we will omit some technical details. We need to show that (2.4) holds, for all $1 \ge a > b \ge 0$, with some constant $d$ such that $K_0^2 d \le \log(1/(pq))$. Set

$$L(a,b) = I(pa+qb) - (pI(a) + qI(b)) \ .$$

The inequality (2.4) may be rewritten, in terms of $L$, as $\frac{L}{pI(a)} \le \sqrt{1+s^2} - 1$, where $s = \frac{\sqrt{d}\,(a-b)}{I(a)}$. Since $\frac{1}{3}\min(s,s^2) \le \sqrt{1+s^2} - 1$, (2.4) is a consequence of the

$$3\,\frac{L(a,b)}{p\,(a-b)} \;\leq\; \sqrt{d}, \qquad \text{and} \qquad 3\,\frac{L(a,b)I(a)}{p\,(a-b)^2} \;\leq\; d\ . \tag{2.7}$$

We use the following elementary properties of the function $I$: $\frac{1}{3} \leq -I''(x)$ $I(x) \leq 1$, for all $x \in (0.1)$, $I$ is thus concave on [0, 1], and for all $a \in (0, 1)$,

$$Q_p(a) \equiv \frac{I^2(pa) - p^2 I^2(a)}{(pa)^2} \leq 3\log\frac{1}{pq}\ \ , \tag{2.8}$$

$$R_p(a) \equiv \frac{I(pa) - pI(a)}{a} \leq 4pq\sqrt{\log\frac{1}{pq}} \tag{2.9}$$

(of course, the value of numerical constants in (2.8)–(2.9) is not essential). We start with the second inequality in (2.7). To prove it, first we observe that for all $1 > a > b \geq 0$,

$$\frac{L(a,b)}{(a-b)^2} \;\leq\; \frac{pq}{2I(a)} + 3\,\frac{L(a,0)}{a^2}\ \ . \tag{2.10}$$

Indeed, by Taylor's formula, for all $x \in [0, 1]$ and $c \in (0, 1)$,

$$I(x) = I(c) + I'(c)\,(x-c) + \int_0^1 I''(t\,c + (1-t)\,x\,)\,t\,dt\ \,(x-c)^2\ \ .$$

Writing there $c = pa + qb$ and $x = a$, $x = b$, and averaging with weights $p$ and $q$, we get a general identity:

$$\frac{L(a,b)}{(a-b)^2} = \;pq\left[q\int_0^1 -I''(t\,c + (1-t)\,a\,)\,t\,dt \right.$$
$$\left. +p\int_0^1 -I''(t\,c + (1-t)\,b\,)\,t\,dt\right]\ . \tag{2.11}$$

Since $-I'' \leq 1/I$,

$$\frac{L(a,b)}{(a-b)^2} \leq \;pq\left[\,q\int_0^1 \frac{1}{I(t\,c + (1-t)\,a\,)}\,t\,dt \right.$$
$$\left. +p\int_0^1 \frac{1}{I(t\,c + (1-t)\,b\,)}\,t\,dt\,\right]\ . \tag{2.12}$$

The function $1/I$ decreases in [0, 1/2] and increases in [1/2, 1], so, the maximum of $1/I$ on every interval $[\alpha, \beta]$ is attained at one of the endpoints of the interval. Therefore, $\frac{1}{I(x)} \leq \frac{1}{I(\alpha)} + \frac{1}{I(\beta)}$, for all $x \in [\alpha, \beta]$. If $t$ and $a$ are fixed, and $b$ varies in $[0, a]$, then $x = t\,c + (1-t)\,a$ varies in the interval $[\alpha, \beta] = [t\,pa + (1-t)\,a, a]$. Therefore,

$$\frac{1}{I(t\,c + (1-t)\,a)} \le \frac{1}{I(a)} + \frac{1}{I(t\,pa + (1-t)\,a)} \quad .$$

In the same way, $\frac{1}{I(t\,c+(1-t)\,b)} \le \frac{1}{I(a)} + \frac{1}{I(t\,pa)}$. Using these estimates in (2.12) yields

$$\frac{L(a,b)}{(a-b)^2} \le \frac{pq}{2I(a)} + pq\left[ q\int_0^1 \frac{1}{I(t\,pa + (1-t)\,a)}\,t\,dt \right.$$
$$\left. + p\int_0^1 \frac{1}{I(t\,pa)}\,t\,dt \right] \quad .$$

Applying the inequality $1/I \le -3I''$, we get

$$\frac{L(a,b)}{(a-b)^2} \le \frac{pq}{2I(a)} + 3\,pq\left[ q\int_0^1 -I''(t\,pa + (1-t)\,a)\,t\,dt \right.$$
$$\left. + p\int_0^1 -I''(t\,pa)\,t\,dt \right] \quad .$$

But the second term (without the constant 3) corresponds to the right hand side of (2.11) with $b = 0$ and is equal to $L(a,0)/a^2$. This proves (2.10). As a result,

$$3\,\frac{L(a,b)I(a)}{p(a-b)^2} \;\le\; \frac{3q}{2} + 9\,\frac{L(a,0)I(a)}{pa^2} \quad . \tag{2.13}$$

Since $I$ is concave, and $I(0) = 0$, we have $I(pa) \ge pI(a)$. Therefore,

$$L(a,0) = I(pa) - pI(a) = \frac{I^2(pa) - p^2I^2(a)}{I(pa) + pI(a)}$$
$$\le \frac{I^2(pa) - p^2I^2(a)}{2pI(a)} = \frac{(pa)^2 Q_p(a)}{2pI(a)} \quad .$$

Using (2.8), we arrive at $\frac{L(a,0)I(a)}{pa^2} \le \frac{3}{2}\log\frac{1}{pq}$, and by (2.13),

$$3\,\frac{L(a,b)I(a)}{p(a-b)^2} \;\le\; \frac{3q}{2} + \frac{27}{2}\log\frac{1}{pq} \;\le\; 15\log\frac{1}{pq} \quad .$$

Thus, the second inequality in (2.7) holds with $d = 15\log\frac{1}{pq}$. To establish the first one, consider, for $1 \ge a \ge b \ge 0$, the inequality

$$L(a,b) = I(pa + qb) - (pI(a) + qI(b)) \;\le\; C(p)\,(a-b) \quad . \tag{2.14}$$

Since $I$ is concave, for any $c \in (0,1)$, the left hand side of (2.14) is a convex function on the segment

$$\Delta(c) = \{(a, b) : pa + qb = c, 1 \geq a \geq b \geq 0\}$$

while the right hand side of (2.14) is a linear function. Therefore, (2.14) holds for all $(a, b) \in \Delta(c)$ if and only if it holds for the endpoints of the segment. Note that, for each of these endpoints with coordinates $(a, b)$, either $b = 0$, or $a = 1$. When $b = 0$, (2.14) becomes

$$I(pa) - pI(a) \leq C_0(p)\, a \tag{2.15}$$

with $C_0(p) = C(p)$, and for $a = 1$, it becomes, after a formal change of the variable $b$ to $a$, $I(qa) - qI(a) \leq C_0(q)\, a$, where again $C_0(q) = C(p)$. Thus, the optimal constants $C(p)$ in (2.14) are connected with the optimal constants $C_0(p)$ in (2.15) via

$$C(p) = \max(C_0(p), C_0(q)) \ .$$

By (2.9), $C_0(p) \leq 4pq\sqrt{\log \frac{1}{pq}}$, so, $C(p) \leq 4pq\sqrt{\log \frac{1}{pq}}$. Therefore, $3\,\frac{L(a,b)}{p(a-b)} \leq 12\sqrt{\log \frac{1}{pq}}$. As a result, the first inequality in (2.7) holds with $d = 144 \log \frac{1}{pq}$. Thus, Proposition 2.4 is proved with $K_0 = 1/12$. □

## 3. Gaussian-type isoperimetric inequality in abstract product spaces

As noted in [T1], the dependence of the right hand side of (1.5) on $\mu_p^n(A)$ given $p$ is of the right order. This is readily seen by checking (1.5) on the sets

$$A_n(a) = \left\{x \in \{0, 1\}^n : \frac{S_n - np}{\sqrt{npq}} \leq a\right\}$$

where $S_n = x_1 + \cdots + x_n$. And as we have seen, the optimal constant in (1.5) as a function of $p$ is of order $1/\sqrt{\log(1/(pq))}$. For the two-sided boundaries, the constants in (1.5) will change considerably:

**Proposition 3.1.** *Let $I(t) = \varphi(\Phi^{-1}(t))$, $0 \leq t \leq 1$. For any subset $A$ of $\{0, 1\}^n$,*

$$\int \sqrt{\bar{h}_A(x)}\, d\mu_p^n(x) \ \geq\ \frac{1}{\sqrt{2pq}}\, I(\mu_p^n(A)) \ . \tag{3.1}$$

Recall that

$$\bar{h}_A(x) = \mathrm{card}\{i \leq n : (x \in A,\ s_i(x) \notin A) \text{ or } (x \notin A,\ s_i(x) \in A)\} \ .$$

In particular, $\bar{h}_{A_n(a)} = (n-k)\,\mathbf{1}_{\{S_n=k\}} + (k+1)\,\mathbf{1}_{\{S_n=k+1\}}$ where $k = \big[\,np + a\sqrt{npq}\,\big]$ (here $[\,\cdot\,]$ denotes the integer part of a real number). By the de Moivré-Laplace' local limit theorem,

$$\int \sqrt{\bar{h}_{A_n(a)}}\, d\mu_p^n = \sqrt{n-k}\;\mu_p^n(S_n = k)$$
$$+\sqrt{k+1}\;\mu_p^n(S_n = k+1) \longrightarrow \frac{\sqrt{p}+\sqrt{q}}{\sqrt{pq}}\,\varphi(a)\ ,$$

as $n \to \infty$. On the other hand, by the central limit theorem,

$$\frac{1}{\sqrt{2pq}}\, I\big(\mu_p^n(A_n(a))\big) \ \longrightarrow\ \frac{1}{\sqrt{2pq}}\,\varphi(a)\ .$$

Therefore, since $1 \le \sqrt{p}+\sqrt{q} \le \sqrt{2}$, the dependence on the pair $(p, \mu_p^n(A))$ on the right-hand side of (3.1) is sharp. A remarkable feature of (3.1) however is that it can be extended to arbitrary product measures, and (3.1) follows from a more general and a simpler inequality. Let $(\Omega, \mathbf{P})$ be the product of a finite number of probability spaces $(\Omega_i, \mu_i)$, $i = 1, 2, \ldots, n$. For every (measurable) function $f$ on $\Omega$, we set

$$\mathrm{D}(f) = \sqrt{\sum_{i=1}^{n} \mathrm{Var}_{x_i}(f)}$$

where $\mathrm{Var}_{x_i}(f)$ denotes the variance of $f$ with respect to the $i$-th coordinate while the remaining of the variables are fixed. The function $\mathrm{D}(f)$ is related to the modulus of the gradient of a smooth function on Euclidean space. Thus, given a measurable set $A \subset \Omega$, one may consider the quantity

$$\mathbf{P}^+(A) = \mathbf{E}\,\mathrm{D}(1_A)$$

as a certain '$\mathbf{P}$-perimeter' of $A$ (the mathematical expectation is taken with respect to $\mathbf{P}$). We shall prove the following result:

**Theorem 3.2.** *For every measurable function $f$ on $\Omega$ with values in* $[0, 1]$,

$$I(\mathbf{E}\, f) \ \le\ \mathbf{E}\sqrt{I(f)^2 + 2\,\mathrm{D}(f)^2}\ . \tag{3.2}$$

*In particular*, *for any measurable subset $A$ of* $\Omega$,

$$\mathbf{P}^+(A) \ \ge\ \frac{1}{\sqrt{2}}\, I(\mathbf{P}(A))\ . \tag{3.3}$$

The inequality (3.3) reduces to (3.1) choosing $\Omega_i = \{0, 1\}$ with distribution $\mu_p$. We do not know an optimal function $I_0$ for which

$$\mathbf{P}^+(A) \ \geq \ I_0(\mathbf{P}(A)) \tag{3.4}$$

holds for all product probability spaces $(\Omega, \mathbf{P})$. However, let us note that if we apply (3.4) to $\Omega = \{0, 1\}^n$, $\mathbf{P} = \mu_p^n$ with $p = t^{1/n}$ and to $A = \{S_n = n\}$, and then let $n \to \infty$, we will arrive in the limit at $J(t) = t\sqrt{\log(1/t)} \geq I_0(t)$. Thus, by (3.3), and by symmetry of $I_0(t)$ around $1/2$,

$$\frac{1}{\sqrt{2}}\, I(t) \leq I_0(t) \leq \min\{J(t), J(1-t)\} \ , \tag{3.5}$$

for any $t \in (0, 1)$. Both sides of (3.5) represent equivalent functions as $t \to 0$ or $t \to 1$, so, the dependence of the right-hand side of (3.3) in $\mathbf{P}(A)$ is sharp, and the factor $1/\sqrt{2}$ in (3.3) and therefore the constant 2 in (3.2) can not be improved.

One may wonder of course whether (3.2) can be sharpened for some special measures, and in particular, when the constant 2 may be improved, e.g., replaced by 1 as the best possible case. Let us apply (3.2) with $n = 1$ to the functions $f = c + \varepsilon g$ with $g$ bounded such that $\mathbf{E}\, g = 0$. If $c \in (0, 1)$, Taylor's expansion around $\varepsilon = 0$ yields

$$\begin{aligned}\mathbf{E}\sqrt{I(f)^2 + \mathrm{Var}(f)} &= \mathbf{E}\sqrt{I(c+\varepsilon g)^2 + \varepsilon^2 \mathrm{Var}(g)} \\ &= I(c) + \frac{I'(c)}{6\, I(c)^2}\, \mathbf{E}\, g^3\, \varepsilon^3 + \mathrm{O}\left(\varepsilon^4\right) \ .\end{aligned}$$

Therefore, validity of the inequality $I(c) = I(\mathbf{E}\, f) \leq \mathbf{E}\sqrt{I(f)^2 + \mathrm{Var}(f)}$ requires that $\mathbf{E}\, g^3 = 0$. But if $\mu_1$ $(= \mathbf{P})$ is different from a Dirac measure in one point, the equality $\mathbf{E}\, g^3 = 0$ holds for all bounded $g$ with $\mathbf{E}\, g = 0$ if and only if $\mu_1$ has two support points with probability 1/2 each. Thus inequality (3.2) may hold in product spaces with the factor 2 replaced by 1 only if we restrict ourselves to the discrete cube $\{0, 1\}^n$ with the normalized uniform measure. As shown in [B1], this is indeed the case.

To prove (3.2), it suffices to consider the case $n = 1$: here we apply Lemma 2.1 with $\mathrm{D}_i(f) = \sqrt{2\mathrm{Var}(f)}$ (which is a functional rather than an operator) and with the Gaussian isoperimetric function

$$I(t) = \varphi(\Phi^{-1}(t)), \quad 0 \leq t \leq 1 \ .$$

In order to treat the case $n = 1$, we note several elementary properties of this function (the important property 3) below was for the first time noticed and used apparently in [AGK]):

(1) $I$ is concave on [0, 1] and is symmetric around the point 1/2;
(2) $I$ increases on [0, 1/2] and decreases on [1/2, 1]; $I(0) = I(1) = 1$;
(3) $I$ satisfies differential equation $I''I = -1$.

**Lemma 3.3.** *For every measurable function $f$ on $\Omega$ with values in* [0, 1],

$$I(\mathbf{E}\, f) - \mathbf{E}\, I(f) \ \le\ \frac{1}{I(\mathbf{E}\, f)}\ \mathrm{Var}(f)\ . \tag{3.6}$$

*Consequently*,

$$I(\mathbf{E}\, f) \ \le\ \mathbf{E}\sqrt{I(f)^2 + 2\,\mathrm{Var}(f)}\ . \tag{3.7}$$

*Proof*. (A two-point analogue of (3.6) for the function $J$ appears in [T1], Lemma 3.2). First let us see how (3.6) implies (3.7). We apply (3.6) and Jensen's inequality to get

$$\begin{aligned}(I(\mathbf{E}\, f))^2 - (\mathbf{E}\, I(f))^2 &= (I(\mathbf{E}\, f) - \mathbf{E}\, I(f))\,(I(\mathbf{E}\, f) + \mathbf{E}\, I(f)) \\ &\le (I(\mathbf{E}\, f) - \mathbf{E}\, I(f))\, 2I(\mathbf{E}\, f) \\ &\le 2\,\mathrm{Var}(f).\end{aligned}$$

Hence, $(I(\mathbf{E}\, f))^2 \ \le\ (\mathbf{E}\, I(f))^2 + 2\,\mathrm{Var}(f) \ \le\ \Big(\mathbf{E}\sqrt{I(f)^2 + 2\,\mathrm{Var}(f)}\Big)^2$ where we used the triangle inequality $\left(\int u\right)^2 + \left(\int v\right)^2 \le \left(\int \sqrt{u^2+v^2}\right)^2$ with $u = I(f)$ and $v = \sqrt{2\,\mathrm{Var}(f)}$.

To prove (3.6), we may assume that $0 < f < 1$. Fix $c = \mathbf{E}\, f$. By Taylor's formula with integral remainder term, we have

$$I(f) = I(c) + I'(c)\,(f-c) + \int_0^1 I''(t\,c + (1-t)\,f\,)\,t\,dt\ \ (f-c)^2\ . \tag{3.8}$$

By properties (1) and (3), we obtain that

$$-I''(tc+(1-t)f)\,t = \frac{t}{I(tc + (1-t)f)} \le \frac{t}{tI(c) + (1-t)I(f)} \le \frac{1}{I(c)}\ .$$

Therefore, by (3.8),

$$I(c) - I(f) \le -I'(c)\,(f-c) + \frac{1}{I(c)}\,(f-c)^2\ . \tag{3.9}$$

It remains to take the expectations of the both sides in (3.9). This proves Lemma 3.3, and thus Theorem 3.2 is proved, as well. □

**Remark 3.4.** The functional inequality (3.2) can be essentially stronger than (3.3). For an illustration, let $\Omega_i = \mathbf{R}^d$ and let $\mu_i = \mu$ be a probability measure with mean 0 and with identity correlation operator. Given a smooth function $f$ on $\mathbf{R}^d$, consider the function $f_n$ on $\Omega = \mathbf{R}^{nd}$ of the form $f_n(x_1, \dots, x_n) = f((x_1 + \cdots + x_n)/\sqrt{n}\,)$, $x_1, \dots, x_n \in \mathbf{R}^d$. If $f$ has bounded first and second partial derivatives, then

$$\mathrm{D}(f_n)\,(x_1, \dots, x_n) \; = \; \left|\nabla f\left((x_1 + \cdots + x_n)/\sqrt{n}\,\right)\right| \left(1 + O\left(1/\sqrt{n}\,\right)\right), \quad \text{as } n \to \infty \; .$$

The application of the central limit theorem to the functions $f_n$ in (3.2) yields

$$I(\mathbf{E}_{\gamma_d} f) \; \le \; \mathbf{E}_{\gamma_d} \sqrt{I(f)^2 + 2\,|\nabla f|^2} \tag{3.10}$$

where the expectations are taken with respect to the Gaussian measure $\gamma_d$ on $\mathbf{R}^d$. Approximating the indicator function $\mathbf{1}_A$ by smooth functions yields in (3.10)

$$\gamma_d^+(A) \ge \frac{1}{\sqrt{2}}\, I(\gamma_d(A)) \; . \tag{3.11}$$

And, as noted, in the special case where $\mu$ is the standard Bernoulli measure on $\{-1, 1\}^d$, the factor 2 can be replaced by 1 in (3.2) and (3.10). On this way, one can recover the Gaussian isoperimetric inequality (1.2) but it is unlikely to reach (1.2) or even (3.11) via a direct application of the isoperimetric-type inequality (3.3): the quantity $\mathbf{P}^+(A)$ does not represent the perimeter of $A$ in the 'usual' geometric sense as in (1.2) and (3.11). Indeed, it could be natural to expect that similarly to the case of smooth functions $f_n$ defined above, for 'regular' sets $A \subset \mathbf{R}^d$ and normalized measure $\mathbf{P}$ on the cube $\left(\{-1, 1\}^d\right)^n$, the convergence

$$\lim_{n \to \infty} \mathbf{P}^+ \left\{(x_1, \dots, x_n) : x_i \in \{-1, 1\}^d, \; \frac{x_1 + \cdots + x_n}{\sqrt{n}} \in A\right\} = \gamma_d^+(A) \tag{3.12}$$

holds. Unexpectedly, (3.12) is false. For half-planes $A = \{(x_1, x_2) \in \mathbf{R}^2 : a_1 x_1 + a_2 x_2 \le a\}$ with $a_1^2 + a_2^2 = 1$, the limit in (3.12) exists but depends on $(a_1, a_2)$ while the right-hand side does not depend on $(a_1, a_2)$ and is equal to $\varphi(a)$.

**Remark 3.5.** Another unexpected observation is that Poisson probability measures satisfy inequalities of type (3.10) for moduli of discrete gradients (instead of usual gradients). D. Bakry and M. Ledoux proved in [BaL] that if a probability measure $\mu$ on the metric space $\Omega$ satisfies, for any ('smooth') function $f$ with values in [0, 1],

$$I(\mathbf{E}_\mu f) \ \le\ \mathbf{E}_\mu \sqrt{I(f)^2 + c^2\,|\nabla f|^2}\ ,$$

then, every function $f$ on $\Omega$ with Lipschitz constant equal to or less than $1/c$ has Gaussian tails and, more precisely, the distribution of $f$ represents a contraction of the canonical Gaussian measure $\gamma_1$. In contrast to this property, we shall derive the following corollary from Theorem 3.2.

Denote the Poisson distribution with parameter $\lambda > 0$ by $\Pi_\lambda$. This measure is concentrated on $\mathbf{Z}_+ = \{0, 1, 2, \ldots\}$ and, for $k \in \mathbf{Z}_+$, $\Pi_\lambda(\{k\}) = e^{-\lambda}\lambda^k/k!$. Note that the tails $\Pi_\lambda(\{i : i \ge k\})$ are 'heavier' than Gaussian tails and behave like $\Pi_\lambda(\{k\})$, as $k \to \infty$. Let $\Pi_\lambda^n$ denote the corresponding product measure on $\mathbf{Z}_+^n$. For any subset $A$ of $\mathbf{Z}_+^n$, define

$$\bar{h}_A(x) = \mathrm{card}\{i \le n : (x \in A,\ x + e_i \notin A) \ \text{or}\ (x \notin A,\ x + e_i \in A)\}\ ,$$

where $(e_i)_{1\le i\le n}$ denotes the canonical basis in $\mathbf{R}^n$. Put $\partial A = \{x \in \mathbf{R}^n : \bar{h}_A(x) > 0\}$. Define the difference operator $\Delta f(x) = f(x+1) - f(x)$, $x \in \mathbf{Z}_+$.

**Proposition 3.6.** *Let* $I(t) = \varphi(\Phi^{-1}(t))$, $0 \le t \le 1$. *For every function* $f$ *on* $\mathbf{Z}_+$ *with values in* [0, 1], *we have*

$$I\left(\int f\, d\Pi_\lambda\right) \le \int \sqrt{I(f)^2 + 2\lambda\,|\Delta f|^2}\, d\Pi_\lambda\ . \tag{3.13}$$

*Consequently, for any subset* $A$ *of* $\mathbf{Z}_+^n$,

$$\int \sqrt{\bar{h}_A}\, d\Pi_\lambda^n \ge \frac{1}{\sqrt{2\lambda}}\, I(\Pi_\lambda^n(A))\ , \tag{3.14}$$

$$\Pi_\lambda^n(\partial A) \ge \frac{1}{\sqrt{2\lambda n}}\, I(\Pi_\lambda^n(A))\ . \tag{3.15}$$

*Proof*. By Lemma 2.1, (3.13) extends to $n$-dimensional case for the operator $\mathrm{D}f(x) = \sqrt{\sum_{i=1}^n (f(x+e_i) - f(x))^2}$. For indicator functions, the $n$-dimensional variant of (3.13) implies (3.14). To prove (3.13), we apply (3.2) for the discrete cube $\Omega = \{0, 1\}^n$ with $p = \lambda/n$ to functions $f_n$ of the form $f_n(x) = f(x_1 + \cdots + x_n)$. Letting $n \to \infty$, it remains to apply Poisson's theorem. At last, (3.14) implies (3.15) since $\bar{h}_A \le n\,\mathbf{1}_{\partial A}$.

In particular, Poisson measures satisfy an inequality similar to (1.3) for the normalized Bernoulli measure. There is another interesting parallel with the discrete cube, connected with Talagrand's quantitative version of

Margulis' theorem. Let $A$ be a monotone subset of $\mathbf{Z}_+^n$, i.e., such that $x \in A \Rightarrow x + e_i \in A$, for all $i \leq n$. Assume that $\bar{h}_A \geq k \geq 1$ on $\partial A$. This is equivalent to saying that $x \notin A \Rightarrow x + e_i \in A$, for $k$ values of $i \leq n$, provided there is at least one such $i$. Using an obvious identity $\frac{d}{d\lambda}\,\Pi_\lambda^n(A) = \int \bar{h}_A\, d\Pi_\lambda^n$, one obtains from (3.14) that

$$\frac{d}{d\lambda}\,\Pi_\lambda^n(A) \geq \frac{\sqrt{k}}{\sqrt{2\lambda}}\, I(\Pi_\lambda^n(A)) \quad .$$

Hence, for $\varepsilon \in (0, 1/2)$, if $\Pi_{\lambda_1}^n(A) = \varepsilon$, $\Pi_{\lambda_2}^n(A) = 1 - \varepsilon$,

$$\begin{aligned}\Phi^{-1}(1-\varepsilon) - \Phi^{-1}(\varepsilon) &= \int_{\lambda_1}^{\lambda_2} \frac{d}{d\lambda}\,\Phi^{-1}(\Pi_\lambda^n(A))\, d\lambda \\ &\geq \int_{\lambda_1}^{\lambda_2} \frac{\sqrt{k}}{\sqrt{2\lambda}}\, d\lambda = \sqrt{k}\left(\sqrt{\lambda_2} - \sqrt{\lambda_1}\right) \quad .\end{aligned}$$

Thus, $\sqrt{\lambda_2} - \sqrt{\lambda_1} \leq \frac{2\,\Phi^{-1}(1-\varepsilon)}{\sqrt{k}}$, so, the function $\lambda \to \Pi_{\lambda^2}^n(A)$ jumps from $\varepsilon$ to $1-\varepsilon$ on an interval of the length at most $K(\varepsilon)/\sqrt{k}$. □

## 4. Inequalities on the lattice $\mathbf{Z}^n$. Discrete Poincaré inequalities on the real line

For any subset $A$ of $\mathbf{Z}^n$, we introduced the function

$$h_A(x) = \mathrm{card}\{i \leq n : x \in A,\ (x + e_i \notin A \text{ or } x - e_i \notin A)\} \ ,$$

where $(e_i)_{i \leq n}$ is the canonical basis in $\mathbf{R}^n$. The interior boundary $\partial_+ A = \{h_A > 0\}$ consists of those points $x$ in $A$ which can leave $A$ at least in one of the $n$ directions. For functions $f$ on $\mathbf{Z}^n$, the corresponding definition of $Mf$ should be

$$Mf(x) = \sqrt{\sum_{i=1}^n ((f(x) - f(x - e_i))^+)^2 + ((f(x) - f(x + e_i))^+)^2} \ ,$$

so that $h_A = M\,\mathbf{1}_A$. Define also the operator $Df(x) = \sqrt{\sum_{i=1}^n (f(x + e_i) - f(x))^2}$. In dimension one $Df(x) = |f(x+1) - f(x)|$ is the modulus of the difference operator $\Delta$.

For every probability measure $\mu$ on the real line $\mathbf{R}$ with distribution function $F_\mu(x) = \mu((-\infty, x])$, $x \in \mathbf{R}$, define the quantile function

$$F_\mu^{-1}(p) = \inf\{x : F_\mu(x) \geq p\}, \quad p \in (0, 1] \ ,$$

which takes values in $(-\infty, +\infty]$. As usual, we denote the corresponding product measure on $\mathbf{R}^n$ by $\mu^n$. As in section 1, we denote by $\nu$ the measure with density $\frac{1}{2}\exp(-|x|)$. Note that there always exists the (unique) left continuous non-decreasing map $U_\mu : \mathbf{R} \to \mathbf{R}$ which transports $\nu$ to $\mu$. Indeed, $U_\mu(x) = F_\mu^{-1}(F_\nu(x))$. If $\mu$ is concentrated on $\mathbf{Z}$, then $F_\mu$ is a step function, and $U_\mu$ takes values in $\mathbf{Z}$.

In this section, first we will describe the class of all probability measures on $\mathbf{Z}^n$ which satisfy (1.6)-type inequality with a dimension free constant.

**Theorem 4.1.** *Let a probability measure $\mu$ be concentrated on $\mathbf{Z}$. Assume that the support of $\mu$ represents an interval in $\mathbf{Z}$. The following properties* (a)–(g) *are equivalent*:

(a) *There exists $c > 0$ such that, for every $n \geq 1$ and every subset $A$ of $\mathbf{Z}^n$*,

$$\int \sqrt{h_A(x)}\, d\mu^n(x) \;\geq\; c\, \mu^n(A)\, (1 - \mu^n(A)) \; . \tag{4.1}$$

(b) *There exists $c > 0$ such that, for every $n \geq 1$ and every subset $A$ of $\mathbf{Z}^n$*,

$$\mu^n(\partial_+ A) \;\geq\; \frac{c}{\sqrt{n}}\, \mu^n(A)\, (1 - \mu^n(A)) \; . \tag{4.2}$$

(c) *There exists $c > 0$ such that, for every subset $A$ of $\mathbf{Z}$*,

$$\mu(\partial_+ A) \;\geq\; c\, \mu(A)\, (1 - \mu(A)) \; . \tag{4.3}$$

(d) *There exists $c > 0$ such that, for all $x \in \mathbf{Z}$*,

$$\mu(\{x\}) \;\geq\; c\, F_\mu(x)(1 - F_\mu(x)) \; . \tag{4.4}$$

(e) *There exists $c > 0$ such that, for every function $f$ on $\mathbf{Z}$*,

$$c\, \mathrm{Var}_\mu(f) \;\leq\; \mathbf{E}_\mu (Mf)^2 \; . \tag{4.5}$$

(f) *There exists $c > 0$ such that, for every function $f$ on $\mathbf{Z}$*,

$$c\, \mathrm{Var}_\mu(f) \;\leq\; \mathbf{E}_\mu (Df)^2 \; . \tag{4.6}$$

(g) $\sup_{x \in \mathbf{R}} [U_\mu(x+h) - U_\mu(x)] \leq 1$, *for some $h > 0$.*

Clearly, the assumption on the support of $\mu$ is necessary for all properties (a)–(g).

The discrete Poincaré-type inequality (4.6) has recently been studied by J.-H.Lou [Lou]. He proves that $\mu$ satisfies (4.6) if and only if, for some $c > 0$, the linear difference equation

$$\Delta f(x-1)\, p(x-1) - (\Delta f(x) + c f(x))\, p(x) = 0,$$

$$\text{where} \quad p(x) = \mu(\{x\}), \; x \in \mathbf{Z} \; , \tag{4.7}$$

has a strictly monotone solution $f$ such that $\int_{\mathbf{Z}} |f|\, d\mu < +\infty$. Another equivalent condition he found indicates that, for some point $x_0 \in \mathbf{Z}$ in the support of $\mu$,

$$\sup_{x \geq x_0} (1 - F_\mu(x)) \sum_{k=x_0}^{x} \frac{1}{p(k)} < +\infty, \quad \sup_{x \leq x_0 - 1} F_\mu(x) \sum_{k=x}^{x_0-1} \frac{1}{p(k)} < +\infty \ . \tag{4.8}$$

J.-H. Lou considers (4.8) as a discrete analogue of the so-called Muckenhoupt's condition for a probability measure $\mu$ on $\mathbf{R}$ satisfying the 'usual' Poincaré-type inequality (1.13) in the class of all smooth functions $f$:

$$\sup_{x > x_0} (1 - F_\mu(x)) \int_{x_0}^{x} \frac{1}{p(t)}\, dt < +\infty, \quad \sup_{x < x_0} F_\mu(x) \int_{x}^{x_0} \frac{1}{p(t)}\, dt < +\infty \ . \tag{4.9}$$

Here $p$ is the density of the absolutely continuous component of $\mu$, and $x_0$ is an arbitrary point such that $0 < F_\mu(x_0) < 1$. The description (4.9) is a particular case of a result due to M. Artola, G. Talenti and G. Tomaselli about Hardy-type inequalities (cf. [Mu]). Apparently it remained unknown for a long time in Probability Theory, since certain attempts have been made to find a characterization of probability measures satisfying (1.13), especially after the rediscovery of the Gaussian Poincaré inequality by H. Chernoff [C] (cf. e.g. [BU], [K], [CL]).

As for the first condition (4.7), it is similar to another characterization for (1.13) due L.H.Y. Chen and J.-H. Lou [CL], formulated in terms of properties of an appropriate Sturm–Liouville equation. Our approach to (4.6) is based on some special properties of the two-sided exponential distribution $\nu$ and leads to different conditions $d$) and $g$). It is however easy to see why (4.4) is equivalent to a formally stronger property (4.8). By (4.4), for all $k \geq x_0$, $1 - F_\mu(k) \leq \alpha(1 - F_\mu(k-1))$ where $\alpha = \frac{1}{1 + cF_\mu(x_0)}$. It follows by induction that $\frac{1-F_\mu(x)}{1-F_\mu(k)} \leq \alpha^{x-k}$ whenever $x_0 \leq k \leq x$. Hence, once more applying (4.4), we get:

$$\begin{aligned}\sum_{k=x_0}^{x} \frac{1 - F_\mu(x)}{p(k)} &\leq \sum_{k=x_0}^{x} \frac{1 - F_\mu(x)}{cF_\mu(x_0)(1 - F_\mu(k))} \\ &\leq \frac{1}{cF_\mu(x_0)} \sum_{k=x_0}^{x} \alpha^{x-k} \leq \frac{1 + cF_\mu(x_0)}{(cF_\mu(x_0))^2} \ .\end{aligned}$$

This proves the first statement in (4.8), and a similar argument yields the second one.

As announced, we will also prove the following analogous statement for probability measures $\mu$ on the real line:

**Theorem 4.2.** *The following properties are equivalent:*
(a) *There exist $h > 0$ and $c > 0$ such that, for every non-decreasing function $f$ on* $\mathbf{R}$,

$$c\,\mathrm{Var}_\mu(f) \;\leq\; \int |f(x+h) - f(x)|^2 \, d\mu(x) \;\; . \tag{4.10}$$

(b) *There exists $c > 0$ such that, for every differentiable convex function $f$ on* $\mathbf{R}$,

$$c\,\mathrm{Var}_\mu(f) \;\leq\; \mathbf{E}_\mu |f'|^2 \;\; . \tag{4.11}$$

(c) *There exists $c > 0$ such that, for all $n \geq 1$ and every convex function $f$ on $\mathbf{R}^n$ with Lipschitz constant at most* 1,

$$c\,\mathrm{Var}_{\mu^n}(f) \leq 1 \;\; . \tag{4.12}$$

(d) *There exist constants $c > 0$ and $h > 0$ such that, for any $x \in \mathbf{R}$*,

$$F_\mu(x) - F_\mu(x-h) \;\geq\; c\, F_\mu(x)(1 - F_\mu(x)) \;\; . \tag{4.13}$$

(e) *For some (or, equivalently, for all) $h > 0$*,

$$\sup_{x \in \mathbf{R}} \, [U_\mu(x+h) - U_\mu(x)] \;<\; +\infty \;\; . \tag{4.14}$$

The 'convex' Poincaré inequality (4.11) is of course a weaker property than the 'usual' Poincaré-type inequality when no assumption on $f$ (except absolute continuity or, equivalently, smoothness) is made. It is known, for example, that any probability measure $\mu$ on $\mathbf{R}$ with a compact support satisfies (b) (cf. [Led], [B2]). A simple argument of getting (4.11) for such measures is the following: for any $f$ convex and differentiable, we have $(f(x) - f(y))^2 \leq (f'(x)^2 + f'(y)^2)\,(x-y)^2$. Hence, if $\mu([a,b]) = 1$,

$$\begin{aligned} \mathrm{Var}_\mu(f) \;&=\frac{1}{2}\int_a^b \int_a^b (f(x) - f(y))^2 \, d\mu(x)\, d\mu(y) \\ &\leq\; (b-a)^2 \int_a^b f'(x)^2 \, d\mu(x) \;\; . \end{aligned}$$

On the other hand, the characterization (4.9) of measures $\mu$ satisfying (4.11) for all smooth $f$ implies in particular that $\mu$ should have a non-trivial absolutely continuous component which is positive on the interval support of $\mu$. Compactness of the support is not sufficient: the measure $\mu$ with density $|x|$ on the interval $[-1, 1]$ provides a counter-example. To compare the 'usual' Pioncaré-type inequality with the property (4.14), one can also express the condition (4.9) in terms of the map $U_\mu$. When $\mu$ has a density which is positive a.e. on the interval support, (4.9) may be written as follows:

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

Let us return to Theorem 4.1 and discuss how to connect isoperimetric-type inequalities (4.1)–(4.3) with discrete Poincaré-type inequalities (4.5). Poincaré-type inequalities are well-known to be of additive type, that is, they can be extended to higher dimensions without any loss in constants. However, the multidimensional variant of (4.5) is unlikely to be appropriate to yield (4.1). To perform the induction step, it is better to work instead with the inequality

$$I(\mathbf{E}_\mu f) \ \le \ \mathbf{E}_\mu \sqrt{I(f)^2 + d\,(Mf)^2} \tag{4.15}$$

where $I(t) = t(1-t)$ and $f$ takes values in [0, 1], as in the proof of Proposition 2.3. According to Lemma 2.1, this inequality is additive as well, and the application of the multidimensional variant of (4.15) to indicator functions yields (4.1) (while the multidimensional variant of (4.5) yields on indicator functions just an estimate for $\int h_A \, d\mu^n$).

Now we are left with the question whether (4.5) and (4.15) are equivalent in one dimension. If $\mu$ is continuous and $Mf = |f'|$, then (4.5) and (4.15) are not equivalent ((4.15) is stronger). However, for the discrete-type gradient, the situation changes considerably. The crucial observation is that $Mf \le \sqrt{2}$ holds whenever $0 \le f \le 1$. Indeed, first let us apply (4.15) to functions $\frac{1}{2} + \varepsilon f$ with $f$ bounded and such that $\mathbf{E}_\mu f = 0$. We obtain, as $\varepsilon \to 0$,

$$\mathrm{Var}_\mu(f) \le 2d \, \mathbf{E}_\mu (Mf)^2 \ .$$

Hence, (4.15) implies (4.5) with $c = 1/(2d)$. To prove the converse, note that, for all $u \in [0, 1/4]$ and $v \in [0, \sqrt{2}]$, $\sqrt{u^2 + dv^2} - u \ge Cv^2$ with $d = \frac{C}{2} + 2C^2$. For $u = I(f)$, $v = Mf$, and for $C = 1/c$, we thus obtain

$$\begin{aligned} &\mathbf{E}_\mu \left[ \sqrt{I(f)^2 + d\,(Mf)^2} - I(f) \right] \\ &\quad \ge C \, \mathbf{E}_\mu (Mf)^2 \ge \mathrm{Var}_\mu(f) = I(\mathbf{E}_\mu f) - \mathbf{E}_\mu I(f) \ . \end{aligned}$$

Hence, (4.5) also implies (4.15) with the constant $d = \frac{1}{2c} + \frac{2}{c^2}$.

Note also that $h_A \le n \mathbf{1}_{\partial_+ A}$ so that (4.2) is a direct consequence of (4.1). In turn, (4.3) is one-dimensional variant of (4.2), and (4.4) is a particular case of (4.3) when $A = (-\infty, x] \cap \mathbf{Z}$. Thus, via (4.15) and Lemma 2.1, we have reduced Theorem 3.1 to one dimension. With the above remarks, in order to prove the theorem, it remains to show the implications (e) & (f) $\Rightarrow$ (d) $\Rightarrow$ (g) $\Rightarrow$ (e) & (f), only. We start with some preparations.

**Lemma 4.3.** *If* $x \le z \le y$, *then*

$$\frac{F_\nu(y) - F_\nu(x)}{F_\nu(z)\,(1 - F_\nu(z))} \;\le\; 2\,\left(e^{y-x} - 1\right) \quad .$$

*Proof* . Clearly, it suffices to consider the cases $z = x$ and $z = y$, only. These cases are equivalent since $\nu$ is symmetric around 0, so let $z = x$. Recall that $F_\nu(u) = \frac{1}{2}\exp(u)$, for $u \le 0$, and $F_\nu(u) = 1 - \frac{1}{2}\exp(-u)$, for $u \ge 0$. Set $h = y - x$. If $0 \le x \le y$, then

$$\frac{F_\nu(y) - F_\nu(x)}{F_\nu(x)\,(1 - F_\nu(x))} = \frac{1 - e^{-h}}{F_\nu(x)} \le 2\,(1 - e^{-h}) \le 2\,(e^h - 1) \quad .$$

If $x \le y \le 0$, then

$$\frac{F_\nu(y) - F_\nu(x)}{F_\nu(x)\,(1 - F_\nu(x))} = \frac{e^h - 1}{1 - F_\nu(x)} \le 2\,(e^h - 1) \quad .$$

Finally, if $x \le 0 \le y$, then

$$\begin{aligned} \frac{F_\nu(y) - F_\nu(x)}{F_\nu(x)\,(1 - F_\nu(x))} &= \frac{2e^{-x} - e^{-2x-h} - 1}{1 - F_\nu(x)} \\ &\le 2\,(2e^{-x} - e^{-2x-h} - 1) \le 2\,(e^h - 1) \quad . \end{aligned}$$

□

**Lemma 4.4.** *Properties* (e) *and* (f) *imply* (d).

*Proof* . We obtain (4.4) applying (4.5) or (4.6) to the indicator function $f = \mathbf{1}_A$ of the set $A = (-\infty, x] \cap \mathbf{Z}$, $x \in \mathbf{Z}$ in which case we have for both operators $Mf = Df = \mathbf{1}_{\{x\}}$.

Now we show that the property (d) implies the property (g). For the purpose of Theorem 4.2, we actually need a more general statement (the particular case $\delta = 1$ below corresponds to the implication (d) $\Rightarrow$ (g) in Theorem 4.1). □

**Lemma 4.5.** *Let* $\mu$ *be a probability measure on the real line* $\mathbf{R}$. *Let* $c > 0$ *and* $\delta > 0$ *be such that*, *for all* $a \in \mathbf{R}$,

$$F_\mu(a) - F_\mu(a - \delta) \ge c\, F_\mu(a)\,(1 - F_\mu(a)) \quad . \tag{4.16}$$

*Then*, $\sup_{x \in \mathbf{R}}\, [U_\mu(x + h) - U_\mu(x)] \le \delta$, *with* $h = \log(1 + c/2)$.

*Proof* . Define the function $U_\mu^{-1} : \mathbf{R} \to [-\infty, +\infty]$, using the identity

$$\{t : U_\mu(t) \le a\} = (-\infty, U_\mu^{-1}(a)], \quad a \in \mathbf{R} \quad .$$

Since $U_\mu$ is left continuous, we may define the function $U_\mu^{-1}$ via $U_\mu^{-1}(a) = F_\nu^{-1}(F_\mu(a))$. From this definition, $F_\mu(a) = F_\nu(U_\mu^{-1}(a))$, for all $a \in \mathbf{R}$. Hence, for $x = U_\mu^{-1}(a-\delta)$, $z = y = U_\mu^{-1}(a)$, (4.16) takes the form

$$F_\nu(y) - F_\nu(x) \geq c\, F_\nu(z)\,(1 - F_\nu(z)) \ .$$

By Lemma 4.3, $c \leq 2\left(e^{y-x} - 1\right)$, provided $x$ and $y$ are finite. Anyway, $y \geq x + \log\left(1 + \frac{c}{2}\right)$, that is,

$$U_\mu^{-1}(a) \geq U_\mu^{-1}(a-\delta) + \log(1 + c/2) \ , \tag{4.17}$$

for all $a \in \mathbf{R}$. Since, by definition of $U_\mu^{-1}$,

$$(-\infty, x] \subset \{t : U_\mu(t) \leq U_\mu(x)\} = (-\infty, U_\mu^{-1}(U_\mu(x))] \ ,$$

we get $U_\mu^{-1}(U_\mu(x)) \geq x$, for all $x \in \mathbf{R}$. Therefore, putting in (4.17) $a = U_\mu(x) + \delta$ and recalling that $h = \log(1 + \frac{c}{2})$, we obtain

$$U_\mu^{-1}(U_\mu(x) + \delta) \geq U_\mu^{-1}(U_\mu(x)) + h \geq x + h \ .$$

The definition of $U_\mu^{-1}$ implies $U_\mu(U_\mu^{-1}(a)) \leq a$, for all $a \in \mathbf{R}$, where, if necessary, we understand the values $U_\mu(-\infty)$ and $U_\mu(+\infty)$ in the usual (limit) sense. Thus

$$U_\mu(x+h) \leq U_\mu(U_\mu^{-1}(U_\mu(x) + \delta)) \leq U_\mu(x) + \delta \ ,$$

and therefore, $\sup_{x \in \mathbf{R}} [U_\mu(x+h) - U_\mu(x)] \leq \delta$, which proves Lemma 4.3. □

For the part (g) $\Rightarrow$ (e) & (d), we need to study some properties of the measure $\nu$.

**Lemma 4.6.** *For any non-negative measurable function $f$ on $\mathbf{R}$, and for all $h \in \mathbf{R}$,*

$$e^{-|h|}\, \mathbf{E}_\nu f(x) \ \leq \ \mathbf{E}_\nu f(x+h) \ \leq \ e^{|h|}\, \mathbf{E}_\nu f(x) \ ,$$

$$e^{-2|h|}\, \mathrm{Var}_\nu f(x) \leq \mathrm{Var}_\nu f(x+h) \leq e^{2|h|}\, \mathrm{Var}_\nu f(x) \ .$$

*Proof.* The density $p_\nu$ of $\nu$ satisfies $p_\nu(x+h)/p_\nu(x) \leq e^{|h|}$. □

The second inequality of Lemma 4.6 will only be used to prove Lemma 4.8 below. So, while referring to Lemma 4.6 in other places, we mean the first inequality. The following two lemmas will be used in the proof of Lemma 4.9 in order to establish (4.5)–(4.6).

**Lemma 4.7.** *For any $a \in \mathbf{R}$ and $h > 0$, there exists $c = c(a,h) > 0$ such that, for every non-decreasing function $f$ on $\mathbf{R}$ with $f(a) = 0$,*

$$c\,\mathbf{E}_\nu f(x)^2 \;\le\; \mathbf{E}_\nu \;\Big[(f(x) - f(x-h))^2 \mathbf{1}_{\{x>a\}} + (f(x+h) - f(x))^2 \mathbf{1}_{\{x \le a-h\}}\Big] \;.$$

*Proof*. By Lemma 4.6, the above inequality holds with $c(a,h) = e^{-2|a|} c(0,h)$, so we may only consider the case $a = 0$. First we show, that for any non-decreasing function $f$ such that $f \equiv 0$ on $(-\infty, 0]$,

$$\mathbf{E}_\nu f(x+h) \ge (2 - e^{-h})\,\mathbf{E}_\nu f(x) \;. \tag{4.18}$$

Indeed, this inequality is linear in $f$, so it suffices to check it on indicator functions $f = \mathbf{1}_{(a,+\infty)}$, $a \ge 0$ (since every non-decreasing, left continuous function $f$ vanishing on $(-\infty, 0]$ can be represented as a mixture of such indicators). For $f = \mathbf{1}_{(a,+\infty)}$, (4.18) takes the form

$$1 - F_\nu(a-h) \ge (2 - e^{-h})\,(1 - F_\nu(a)), \quad a, h \ge 0 \;.$$

When $a \ge h \ge 0$, this is simply $e^h + e^{-h} \ge 2$. When $0 \le a \le h$, it becomes

$$1 - \frac{1}{2}\,e^{a-h} \ge \frac{1}{2}\,(2 - e^{-h})\,e^{-a}$$

which is also evident (with equality for $a = 0$). Let us now rewrite (4.18) in the following way:

$$(1 - e^{-h})\,\mathbf{E}_\nu f(x) \;\le\; \mathbf{E}_\nu (f(x+h) - f(x)) \tag{4.19}$$

When $f \equiv 0$ on $(-\infty, 0]$, one can also apply (4.19) to $f^2$ so that, by Cauchy-Schwarz inequality and by Lemma 4.6,

$$\begin{aligned}
(1 - e^{-h})\,\|f\|_2^2 = \;& (1 - e^{-h})\,\mathbf{E}_\nu f(x)^2 \\
&\le \mathbf{E}_\nu (f(x+h) - f(x))\,(f(x+h) + f(x)) \\
&\le \|f(x+h) - f(x)\|_2\,\|f(x+h) + f(x)\|_2 \\
&\le \|f(x+h) - f(x)\|_2 \left(e^{h/2}\|f\|_2 + \|f\|_2\right) \;.
\end{aligned}$$

Hence, $\left(\frac{1-e^{-h}}{1+e^{h/2}}\right)^2 \mathbf{E}_\nu f(x)^2 \le \mathbf{E}_\nu (f(x+h) - f(x))^2$. Estimating the right-hand side of this inequality by Lemma 4.6, and noting that $f(x) - f(x-h) = 0$ for $x \le 0$, we obtain

$$c(h)\,\mathbf{E}_\nu f(x)^2 \ \le\ \mathbf{E}_\nu (f(x) - f(x-h))^2 \mathbf{1}_{\{x>0\}} \tag{4.20}$$

where $c(h) = e^{-h}\big(\frac{1-e^{-h}}{1+e^{h/2}}\big)^2$. By symmetry, for any non-decreasing function $f$ such that $f \equiv 0$ on $[0,+\infty)$, we get as well

$$c(h)\,\mathbf{E}_\nu f(x)^2 \ \le\ \mathbf{E}_\nu (f(x+h) - f(x))^2 \mathbf{1}_{\{x<0\}}\ . \tag{4.21}$$

Consider now the general case. Let $f_0 = f\,\mathbf{1}_{(-\infty,0]}$ and $f_1 = f\,\mathbf{1}_{[0,+\infty)}$. Applying (4.20) to $f_1$ and (4.21) to $f_0$, adding these inequalities and noting that, for all $x \in \mathbf{R}$,

$$f_1(x+h) - f_1(x) \le f(x+h) - f(x), \quad \text{and}$$

$$f_0(x) - f_0(x-h) \le f(x) - f(x-h)\ ,$$

we obtain

$$\begin{aligned} c(h)\,\mathbf{E}_\nu f(x)^2 \le{}& \mathbf{E}_\nu (f(x) - f(x-h))^2 \mathbf{1}_{\{x>0\}} \\ &+ \mathbf{E}_\nu (f(x+h) - f(x))^2 \mathbf{1}_{\{x<0\}}\ . \end{aligned}$$

But, again by Lemma 4.6,

$$\begin{aligned} \mathbf{E}_\nu (f(x+h) - f(x))^2 \mathbf{1}_{\{-h<x<0\}} &\le e^h\,\mathbf{E}_\nu (f(x) - f(x-h))^2 \mathbf{1}_{\{0<x<h\}} \\ &\le e^h\,\mathbf{E}_\nu (f(x) - f(x-h))^2 \mathbf{1}_{\{x>0\}}\ , \end{aligned}$$

so,

$$\begin{aligned} c(h)\,\mathbf{E}_\nu f(x)^2 \ \le{}& (1+e^h)\,\mathbf{E}_\nu (f(x) - f(x-h))^2 \mathbf{1}_{\{x>0\}} \\ &+ \mathbf{E}_\nu (f(x+h) - f(x))^2 \mathbf{1}_{\{x\le -h\}}\ . \end{aligned}$$

This yields the desired result with $c(0,h) = c(h)/(1+e^h)$. □

**Lemma 4.8.** *For any real $a > b$, there exists a constant $c = c(a,b) > 0$ such that, for every non-decreasing function $f$,*

$$c\,\mathrm{Var}_\nu(f(x)) \ \le\ \mathbf{E}_\nu (f(x+a) - f(x+b))^2\ .$$

*Proof.* Due to Lemma 4.6, it suffices to consider the case $a = h > 0$, $b = -h$. In addition, one may assume that $f(0) = 0$. Then, apply Lemma 4.7 with $a = 0$ and note that $f(x) - f(x-h) \le f(x+h) - f(x-h)$, $f(x+h) - f(x) \le f(x+h) - f(x-h)$. □

The last step in the proof of Theorem 4.1 is the following:

**Lemma 4.9.** *Property* (g) *implies properties* (e) *and* (f).

*Proof* . Let $h > 0$ be such that $U_\mu(x + h) \le U_\mu(x) + 1$, for all $x \in \mathbf{R}$. Let $f$ be a non-decreasing function on $\mathbf{Z}$. Applying Lemma 4.8 with $a = h$ and $b = 0$ to the function $f(U_\mu)$, we obtain

$$\begin{aligned} c\,\mathrm{Var}_\mu(f(y)) = c\,\mathrm{Var}_\nu(f(U_\mu(x))) &\le \mathbf{E}_\nu(f(U_\mu(x + h)) - f(U_\mu(x)))^2 \\ &\le \mathbf{E}_\nu(f(U_\mu(x) + 1) - f(U_\mu(x)))^2 \\ &= \mathbf{E}_\mu(f(y + 1) - f(y))^2 \ . \end{aligned}$$

That is, $c\,\mathrm{Var}_\mu(f) \le \mathbf{E}_\mu(Df)^2$, and (4.6) is thus proved for the class of non-decreasing functions $f$. To prove this inequality in general, we may assume that $\mu$ is not a Dirac measure. Let for definiteness $p = \mu(\{0\}) > 0$ and $p_1 = \mu(\{1\}) > 0$. Then, for any function $f \in L^2(\mathbf{Z}, \mu)$ with $f(0) = 0$, we get by the Cauchy-Schwarz inequality,

$$(\mathbf{E}_\mu f)^2 = (\mathbf{E}_\mu f\,\mathbf{1}_{\{x\neq 0\}})^2 \le (1 - p)\,\mathbf{E}_\mu f^2 \ .$$

Hence, $\mathrm{Var}_\mu(f) \ge p\,\mathbf{E}_\mu f^2$. For every non-decreasing $f$ with $f(0) = 0$, we arrive at

$$cp\,\mathbf{E}_\mu f(y)^2 \ \le \ \mathbf{E}_\mu(Df(y))^2 \ . \tag{4.22}$$

We will also need an analogous inequality for the operator $M$. Let $a \in \mathbf{R}$ be the maximal solution to the equation $U_\mu(a) = 0$ (such an $a$ exists since the function $U_\mu$ is left continuous and $\mu([1, +\infty)) > 0$). Then, for all $x > a$, we have $y = U_\mu(x) \ge 1$, while, for $x \le a - h$, we have $y = U_\mu(x) \le 0$. Applying Lemma 4.7 with these values of $a$ and $h$ to the function $f(U_\mu)$ and noting that $f(U_\mu(a)) = f(0) = 0$, we obtain in the same way as above that

$$c\,\mathbf{E}_\mu f(y)^2 \ \le \ \mathbf{E}_\mu(Df(y - 1))^2\mathbf{1}_{\{y\ge 1\}} + \mathbf{E}_\mu(Df(y))^2\mathbf{1}_{\{y\le 0\}} \ , \tag{4.23}$$

where $c = c(a, h)$ is from Lemma 4.7. The second expectation on the right contains the term $(Df(0))^2 p$ which can be estimated from above by $\frac{p}{p_1}\mathbf{E}_\mu(Df(y - 1))^2\mathbf{1}_{\{y\ge 1\}}$. Thus, we get from (4.23)

$$c_1\,\mathbf{E}_\mu f(y)^2 \ \le \ \mathbf{E}_\mu(Df(y - 1))^2\mathbf{1}_{\{y\ge 1\}} + \mathbf{E}_\mu(Df(y))^2\mathbf{1}_{\{y\le -1\}} \ , \tag{4.24}$$

with the constant $c_1 = c/(1 + \frac{p}{p_1})$. Now we extend (4.22) and (4.24) to an arbitrary function $f$ with $f(0) = 0$. Define the function $g$ as follows. Let $g(0) = 0$,

$$g(y) = \sum_{i=1}^{y} |f(i) - f(i-1)|, \quad \text{for } y \geq 1,$$

$$g(y) = -\sum_{i=y}^{-1} |f(i) - f(i+1)|, \quad \text{for } y \leq -1 \ .$$

Then, for all $y \in \mathbf{Z}$, $|g(y)| \geq |f(y)|$ and $g(y+1) - g(y) = |f(y+1) - f(y)|$. Since $g$ is non-decreasing and $g(0) = 0$, we get, by (4.22):

$$cp\, \mathbf{E}_\mu f(y)^2 \leq cp\, \mathbf{E}_\mu g(y)^2 \leq \ \mathbf{E}_\mu (Dg(y))^2 = \mathbf{E}_\mu (Df(y))^2 \ .$$

By similar arguments, (4.24) holds for $f$, too. Thus, $cp\ \mathrm{Var}_\mu(f) \leq \mathbf{E}_\mu (Df)^2$, and, since the condition $f(0) = 0$ can be assumed in (4.6), (4.6) is proved in general. Now we prove (4.5) using (4.24). First let us observe that, for all $y \in \mathbf{Z}$,

$$(Df(y))^2 \leq (Mf(y))^2 + (Mf(y+1))^2 \ .$$

Hence, by (4.24),

$$\begin{aligned} c_1\, \mathrm{Var}_\mu(f) &\leq c_1\, \mathbf{E}_\mu f(y)^2 \\ &\leq \mathbf{E}_\mu \left[ (Mf(y))^2 \mathbf{1}_{\{y \geq 1\}} + (Mf(y-1))^2 \mathbf{1}_{\{y \geq 1\}} \right] \\ &\quad + \mathbf{E}_\mu \left[ (Mf(y))^2 \mathbf{1}_{\{y \leq -1\}} + (Mf(y+1))^2 \mathbf{1}_{\{y \leq -1\}} \right] \ . \end{aligned}$$

Therefore, it suffices to show that the inequalities

$$\mathbf{E}_\mu (Mf(y-1))^2 \mathbf{1}_{\{y \geq 1\}} \leq \alpha\, \mathbf{E}_\mu (Mf(y))^2 \mathbf{1}_{\{y \geq 0\}} \ , \tag{4.25}$$

$$\mathbf{E}_\mu (Mf(y+1))^2 \mathbf{1}_{\{y \leq -1\}} \leq \beta\, \mathbf{E}_\mu (Mf(y))^2 \mathbf{1}_{\{y \leq 0\}} \tag{4.26}$$

hold with some constants $\alpha$ and $\beta$ not depending on $f$. In fact, (4.25)–(4.26) are valid for an arbitrary non-negative function $V$ and thus for $V = (Mf)^2$ in particular. Indeed, since we have proved (4.6), we can use the inequality (4.4) which is its consequence. From (4.4), for any $y \geq 0$ integer,

$$\mu(\{y\}) \geq c\, F_\mu(y)\,(1 - F_\mu(y)) \geq c\, F_\mu(0)\, \mu(\{y+1\}) \ .$$

Hence, $\mathbf{E}_\mu V(y-1)\, \mathbf{1}_{\{y \geq 1\}} = \sum_{y=0}^{\infty} V(y) \mu(\{y+1\}) \leq \frac{1}{c\, F_\mu(0)}\, \mathbf{E}_\mu V(y)\, \mathbf{1}_{\{y \geq 0\}}$. This proves (4.25). Analogously, for $y \leq -1$,

$$\mu(\{y\}) \geq c\, F_\mu(y)\,(1 - F_\mu(y)) \geq c\, \mu(\{y-1\})\,(1 - F_\mu(-1)) \ .$$

Hence,

$$\begin{aligned}\mathbf{E}_\mu V(y+1)\,\mathbf{1}_{\{y\le -2\}} &= \sum_{y=-\infty}^{-1} V(y)\mu(\{y-1\}) \\ &\le \frac{1}{c\,(1-F_\mu(-1))}\,\mathbf{E}_\mu V(y)\,\mathbf{1}_{\{y\le -1\}} \ .\end{aligned}$$

In addition,

$$\mathbf{E}_\mu V(y+1)\,\mathbf{1}_{\{y=-1\}} \ = \ V(0)\,\mu(\{-1\}) \ \le \ \frac{1}{p}\,\mathbf{E}_\mu V(y)\,\mathbf{1}_{\{y=0\}} \ .$$

Since $F_\mu(0) \ge p > 0$ and $1 - F_\mu(-1) \ge p > 0$, we get (4.26) summing the above inequalities. Thus Lemma 4.9 and Theorem 4.1 are proved. □

**Proof of Theorem 4.2.** We go according to (a) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (a), and then (e) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (e).

(a) $\Rightarrow$ (d): (4.10) becomes (4.13) on indicator functions $f = \mathbf{1}_{(x,+\infty)}$.

(d) $\Rightarrow$ (e): by Lemma 4.5, (4.14) holds with $h = \log(1 + c/2)$.

(e) $\Rightarrow$ (a). Since (4.10) holds for $\nu$ (apply Lemma 4.8 to $a = h$ and $b = 0$), we get (4.10) for $\mu$ with $h^* = \sup_{x\in\mathbf{R}}\,[U_\mu(x+h) - U_\mu(x)]$ which replaces $h$.

(e) $\Rightarrow$ (b). Again apply Lemma 4.8, but now to $a = 0$ and $b = -h$: there exists a constant $C(h)$ such that, for any non-decreasing function $g$,

$$\mathrm{Var}_\nu(g(x)) \ \le \ C(h)\,\mathbf{E}_\nu(g(x) - g(x-h))^2 \ . \tag{4.27}$$

Assuming that $f$ is non-decreasing, apply now (4.27) to $g = f(U_\mu)$. We then obtain

$$\mathrm{Var}_\mu(f(y)) \ \le \ C(h)\,\mathbf{E}_\mu(f(y) - f(y-h^*))^2$$

with $h^*$ defined in the previous step. Since $f$ is convex and non-decreasing,

$$0 \le f(y) - f(y-h^*) \le h^* f'(y) \ .$$

Hence, (4.11) holds with constant $c = 1/(C(h)\,h^{*2})$ for all convex and non-decreasing functions $f$. Similarly, it holds for all convex and non-increasing functions $f$. Consider now the third case: $f$ is non-increasing on an interval $(-\infty, a]$ and is non-decreasing on $[a, +\infty)$. One may assume that $f(a) = 0$. As already proved, (4.27) holds for functions $f_0 = f\,\mathbf{1}_{[a,+\infty)}$ and $f_1 = f\,\mathbf{1}_{(-\infty,a]}$ since these functions are convex and monotone. Writing

$a_0 = \mathbf{E}_\mu f_0$, $a_1 = \mathbf{E}_\mu f_1$, and noting that $f = f_0 + f_1$, we obtain

$$\begin{aligned}\sqrt{\mathrm{Var}_\mu(f)} = \|f - \mathbf{E}_\mu f\|_2 &\leq \|f_0 - a_0\|_2 + \|f_1 - a_1\|_2 \\ &\leq \sqrt{C(h)}\, h^* \,\|f_0'\|_2 + \sqrt{C(h)}\, h^* \,\|f_1'\|_2 \\ &\leq \sqrt{2C(h)}\, h^* \,\|f'\|_2 \ ,\end{aligned}$$

where $\|\cdot\|_2$ denotes the norm in the space $L^2(\mu)$. In the last inequality we applied $\sqrt{u_0} + \sqrt{u_1} \leq \sqrt{2(u_0 + u_1)}$ to $u_0 = \mathbf{E}_\mu(f_0')^2$ and $u_1 = \mathbf{E}_\mu(f_1')^2$. Thus, (4.11) is proved in the general case with $C = 2\,C(h)h^{*2}$.

(b) $\Rightarrow$ (c). Apply the multidimensional variant of (4.11) to $f$.

(c) $\Rightarrow$ (e). One can apply (4.12) to the functions $f_n(x) = \max(x_1, \ldots, x_n)$ and $f_n(x) = -\min(x_1, \ldots, x_n)$, for arbitrary $n$. As shown in [BH2], the variances of such functions are bounded if and only if the property (e) is satisfied.

## References


[AC] Ahlswede, R., Cai, N.: General edge-isoperimetric inequalities. I. Information-theoretical methods. Europian J. Combin. **18**, No. 4, 355–372 (1997)

[AGK] Ahlswede, R., Gács, P., Körner, J.: Bounds on conditional probabilities with applications in multi-user communication. Z. Wahrscheinlichkeitstheorie verw. Gebiete **34**, 157–177 (1976)

[BaL] Bakry, D., Ledoux, M.: Lévy-Gromov isoperimetric inequality for an infinite dimensional diffusion generator. Invent. Math. **123**, 259–281 (1996)

[B1] Bobkov, S.G.: An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space. Abstracts of the 4th World Congress of the Bernoulli Society. Vienna, Austria, August 26–31, 1996, p. 130. Ann. Probab. **25**, No. 1, 206–214 (1997)

[B2] Bobkov, S.G.: Some extremal properties of Bernoulli distribution. Probab. Theory Appl. **41**, No. 4, 877–884 (1996)

[BH1] Bobkov, S.G., Houdré, C.: Isoperimetric constants for product probability measures. Ann. Probab. **25**, No. 1, 184–205 (1997)

[BH2] Bobkov, S.G., Houdré, C.: Dimension free weak type of the concentration of measure. Preprint. To appear in: Bernoulli (1995)

[Bor] Borell, C.: The Brunn-Minkowski inequality in Gauss Space. Invent. Math. **30**, Fasc. 2, 207–216 (1975)

[BU] Borovkov, A.A., Utev, S.A.: On an inequality and a related characterization of the normal distribution. Theory Probab. Appl. **28**, 209–218 (1983)

[CL] Chen, L.H.Y., Lou, J.-H.: A characterization of probability measures which admit Poincaré inequalities. Preprint (1991).

[C] Chernoff, H.: A note on an inequality involving the normal distribution. Ann. Probab. **9**, 533–535 (1981)

[FF] Frankl, P., Füredi, Z.: A short proof for a theorem of Harper about Hamming spheres. Discrete Math. **34**, 311–313 (1981)

[H1] Harper, L.H.: Optimal assignment of numbers to vertices. J. Soc. Ind. Appl. Math. **12**, No. 1, 131–135 (1964)

[H2] Harper, L.H.: Optimal numbering and isoperimetric problems on graphs. J. Comb. Theor. **1**, 385–393 (1966)

[K] Klaassen, C.A.J.: On an inequality of Chernoff. Ann. Probab. **13**, 966–974 (1985)

[Led] Ledoux, M.: On Talagrand's deviation inequalities for product measures. ESAIM Probability & Statistics **1**, 63–87 (1996)

[Lin] Lindsey, J.H.: Assignment of numbers to vertices. Amer. Math Monthly. **71**, 508–516 (1964)

[Lou] Lou, J.-H.: On discrete analog of Poincaré-type inequalities and density representation. Preprint

[M] Margulis, G.A.: Probabilistic characteristics of graphs with large connectivity. Problemy peredachi informatsii. **10,** No. 2, 101–108 (Russian) (1974). English translation: Problems of information transmission. **10**, 174–179 (1975)

[Mu] Muckenhoupt, B.: Hardy's inequality with weights. Studia Math. XLIV, 31–38 (1972)

[ST] Sudakov, V.N., Tsirel'son, B.S.: Extremal properties of half-spaces for spherically invariant measures. Zap. Nauchn. Semin. Leningr. Otd. Mat. Inst. Steklova. **41**, 14–24 (1974). English translation: J. Soviet. Math. **9**, 9–18 (1978)

[T] Talagrand, M.M.: Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem. Geom. and Funct. Anal. **3**, No. 3, 295–314 (1993)

[WW] Wang, D.L., Wang, P.: Extremal configurations on a discrete torus and a generalization of the generalized Macaulay theorem. Siam J. Appl. Math. **33**, No. 1, 55–59 (1977)