DIMENSION FREE WEAK CONCENTRATION OF MEASURE PHENOMENON *

S. G. Bobkov † and C. Houdré ‡

July 24, 1995

Abstract

For product probability measures μ^n, we obtain necessary and sufficient conditions (in terms of μ) for dimension free isoperimetric inequalities of the form $\mu^n(A + h[-1,1]^n) \geq R_h(\mu^n(A))$ to hold; for a function R such that $R(p) > p$, for all (some) $p \in (0,1)$, and for $h > 0$ large enough. Some questions related to the concentration of measure phenomenon are also discussed.

1 Introduction

Let μ be a probability measure on the real line $\mathbf{R}$, and let μ^n be the n–fold tensor product of μ with itself. Given a notion of enlargement $enl(A)$ for sets A, inequalities of isoperimetric type have the form

$$\mu^n(enl(A)) \geq R^{(n)}(\mu(A)).$$

Moreover, if $R = R^{(n)}$ is dimension free, such inequalities are often viewed as concentration inequalities. One question of interest which will be addressed here is, whether or not, there exists such a function R (of course, such that $R(p) > p$). Besides the measure, the answer essentially depends on the enlargement. Usually, it is built with the help of a metric, say ρ, by putting

$$enl(A) = A^h = \{x \in \mathbf{R}^n : \rho(x,a) \leq h \text{ for some } a \in A\},$$

*Key words: Concentration of measure, Isoperimetry

†Research supported in part by ISF NXZ000, NXZ300 and Grant No. 94–1.4–57 from the Grant Center of Fundamental Sciences in Sankt–Petersburg University

‡Research supported in part by an NSF Postdoctoral Fellowship. This author enjoyed the hospitality 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.

where $h > 0$ is a fixed number (for A compact, A^h is the closed h–neighbourhood of A with respect to ρ). To consider the weakest possible type of enlargement, we equip $\mathbf{R}^n$ with the supremum distance

$$\rho_\infty(x, y) = \sup_{1 \le i \le n} |x_i - y_i|$$

and consider the value

$$R_h^{(n)}(p) = \inf_{\mu^n(A) \ge p} \mu^n(A^h), \tag{1.1}$$

where the infimum is taken over all the Borel sets of measure $\mu^n(A) \ge p$, $p \in (0,1)$. Note here that if A is Borel measurable, so is A^h, hence (1.1) is well–defined. For other types of metrics this is not necessarily true, and therefore one rather takes the inner measure in (1.1). In his work on isoperimetry, M.Talagrand [4, Proposition 5.1] made the following observation: if $\inf_n R_h^{(n)}(1/2) > 1/2$, then μ has finite exponential moment, that is,

$$\int_{\mathbf{R}} \exp(\varepsilon |x|)\, d\mu(x) < +\infty,$$

for some $\varepsilon > 0$. In proving this result, he studied the behavior of $\mu^n(A^h)$ for the cubes A. As it turns out, also studying enlargements of cubes allows us to find necessary and sufficient conditions for the concentration inequality $\mu^n(A^h) \ge R(\mu(A))$, to hold for some R such that $R(p) > p$.

Definition 1.1 *A non–decreasing function $U^* : (0, +\infty) \longrightarrow [0, +\infty)$ is called a modulus, if for all positive a and b, $U^*(a + b) \le U^*(a) + U^*(b)$.*

In addition, if $U^*(0^+) = 0$, then U^* is a usual modulus of continuity (U^* is necessarily continuous). Clearly, if U^* is a modulus, then for some $a, b \ge 0$, $U^*(h) \le a + bh$, whenever $h > 0$.

Definition 1.2 *A function U defined on some interval $\Delta \subset \mathbf{R}$, (finite or not, closed or open, or semi–open) generates a finite modulus, if for all (equivalently, for some) $h > 0$,*

$$U^*(h) = \sup\{|U(x) - U(y)| : x, y \in \Delta, |x - y| \le h\} < +\infty.$$

Clearly U^*, above, is a modulus and U generates a finite modulus if and only if for some $a, b \ge 0$, $|U(x) - U(y)| \le a + b|x - y|$, whenever $x, y \in \Delta$.
Now define the function U_μ as follows:

$$U_\mu(x) = F_\mu^{-1}\left(\frac{1}{1 + \exp(-x)}\right), \quad x \in \mathbf{R},$$

where $F_\mu(x) = \mu((-\infty, x])$ is the distribution function of the measure μ, and where

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

is the minimal quantile of order p of μ. The map U_μ transforms the logistic distribution ν (given by $\nu((-\infty, x]) = 1/(1+\exp(-x)))$ into the probability measure μ.

Theorem 1.3 *Let $p \in (0,1)$. There exists $h > 0$ such that $\inf_n R_h^{(n)}(p) > p$, if and only if U_μ generates a finite modulus. In this case, setting $h^* = U_\mu^*(h)$, we have*

$$\inf_n R_{h^*}^{(n)}(p) \geq \frac{p}{p + (1-p)\exp(-h)} \tag{1.2}$$

with equality for $\mu = \nu$. In particular, the following alternative holds: either $\inf_n R_h^{(n)}(p)$ $= p$, for all $h > 0$, or $\inf_n R_h^{(n)}(p) \to 1$, as $h \to +\infty$.

Let ξ_n, $n \geq 1$ be a sequence of independent random variables, defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$, and with common distribution μ.

Remark 1.4 When $\mu = \nu$, the right-hand side of (1.2) is simply $\mathbf{P}\{\xi_1 - m_p(\xi_1) \leq h\}$, where $m_p(\xi_1) = F_\nu^{-1}(p)$ is the quantile of order p of the random variable ξ_1. Let $f : \mathbf{R}^n \to \mathbf{R}$ be a Lipschitz function, of Lipschitz constant at most 1 (with respect to ρ_∞), and let $\eta = f(\xi_1, \cdots, \xi_n)$. Applying (1.2) to sets of the form $\{f \leq const.\}$, it easily follows that

$$\mathbf{P}\{\eta - m_p(\eta) \leq h\} \geq \mathbf{P}\{\xi_1 - m_p(\xi_1) \leq h\},$$

for all $p \in (0,1)$ and $h > 0$. This is equivalent to saying that there exists a Lipschitz function $L_f : \mathbf{R} \to \mathbf{R}$ such that the random variables η and $L_f(\xi_1)$ are identically distributed (for details, see [3]). Thus, the distribution of any Lipschitz function f with respect to ν^n is a Lipschitz image of ν.

Let us now give another equivalent wording for the concentration property $\inf_n R_h^{(n)}(p) > p$.

Theorem 1.5 *Let $p \in (0,1)$. There exists $h > 0$ such that $\inf_n R_h^{(n)}(p) > p$, if and only if the following two conditions hold:*

1a) $\sup_n \mathbf{Var} \max\{\xi_1, \cdots, \xi_n\} < +\infty$;

2a) $\sup_n \mathbf{Var} \min\{\xi_1, \cdots, \xi_n\} < +\infty$.

Equivalently, and more generally, if and only if for any fixed $\alpha \geq 1$,

1b) $\sup_n \mathbf{E}\,|\max\{\xi_1, \cdots, \xi_n\} - \mathbf{E}\max\{\xi_1, \cdots, \xi_n\}|^\alpha < +\infty$;

2b) $\sup_n \mathbf{E}\,|\min\{\xi_1, \cdots, \xi_n\} - \mathbf{E}\min\{\xi_1, \cdots, \xi_n\}|^\alpha < +\infty$.

Moreover, in 1b) and 2b) the second expectations can be replaced by the quantiles of order $p \in (0,1)$.

The properties 1a), 2a) are just 1b), 2b) for $\alpha = 2$. If the distribution μ is symmetric about a point, then 1a) and 2a) (and more generally, 1b) and 2b)) are equivalent, and thus they are separately characterized in terms of the notion of modulus. It is then natural to ask if there exist necessary and sufficient condition for 1b) to hold (in the general non–symmetric case). We discuss this question, when $\alpha = 1$, in Section 5 where it is proved that $\sup_n \mathbf{E}\,|\max\{\xi_1, \cdots, \xi_n\} - \mathbf{E}\max\{\xi_1, \cdots, \xi_n\}| < +\infty$ if and only if $\mathbf{E}\,|\xi_1| < +\infty$, and the function U_μ generates a finite modulus on the interval $[0, +\infty)$. This characterization is then applied in Section 6 to prove:

Theorem 1.6 *Let ξ be a random variable with values in $(0,1)$, and with distribution function F_ξ. Then,*

$$\mathbf{E}\,\xi^n = O\left(\frac{1}{n}\right), \text{ as } n \to \infty, \tag{1.3}$$

if and only if there exist $a, b \geq 0$ such that, for all $1/2 \leq p \leq q < 1$,

$$\int_p^q \frac{1}{1-t}\, dF_\xi(t) \;\leq\; a + b \log \frac{1-p}{1-q}. \tag{1.4}$$

The proofs of Theorem 1.3 and 1.5 are respectively given in Section 3 and 4, while Section 5 leads to the proof of Theorem 1.6 in Section 6. In Section 7, and in contrast to the alternative mentioned in Theorem 1.3, it is shown that $R_h^{(n)}(p) \to 1$, as $h \to +\infty$, when n is fixed (to be more precise, this property is proved in an abstract metric space). We start (Section 2) by giving some characterizations of the concentration of measure phenomenon for the sequence of maxima.

2 Concentration of maxima

Let ξ_n, $n \geq 1$, be a sequence of independent random variables with a common distribution function F. Let $M_n = \max\{\xi_1, \cdots, \xi_n\}$, and let $\mathcal{L}(\cdot)$ be the law (distribution) of a random variable. Below, we also use the standard notations $x^+ = \max(x, 0), x^- = \max(-x, 0)$.

Proposition 2.1 *The following are equivalent:*

1a) for some $p \in (0,1)$ and $h \in \mathbf{R}$, $\inf_n \mathbf{P}\{M_n - m_p(M_n) \leq h\} > p$;

1b) for some $p \in (0,1)$ and $h \in \mathbf{R}$, $\sup_n \mathbf{P}\{M_n - m_p(M_n) \leq h\} < p$;

2a) for any $p \in (0,1)$, there exists $h \in \mathbf{R}$ such that $\inf_n \mathbf{P}\{M_n - m_p(M_n) \leq h\} > p$;

2b) for any $p \in (0,1)$, there exists $h \in \mathbf{R}$ such that $\sup_n \mathbf{P}\{M_n - m_p(M_n) \leq h\} < p$;

3a) for all $p \in (0,1)$, $\inf_n \mathbf{P}\{M_n - m_p(M_n) \leq h\} \to 1$, as $h \to +\infty$;

3b) for all $p \in (0,1)$, $\sup_n \mathbf{P}\{M_n - m_p(M_n) \leq h\} \to 0$, as $h \to -\infty$;

4) for a sequence of real numbers a_n, $\sup_n \mathbf{P}\{|M_n - a_n| > h\} \to 0$, as $h \to +\infty$, that is, the sequence of laws $\mathcal{L}(M_n - a_n)$ is tight (i.e., it forms a precompact set in the space of probability measures on $\mathbf{R}$ equipped with the weak convergence topology);

5) for all (equivalently, for some) $p, q \in (0,1)$, $\sup_n |m_p(M_n) - m_q(M_n)| < +\infty$;

6) for all (equivalently, for some) $a \in \mathbf{R}$, the function $U(x) = F^{-1}(1/(1+\exp(-x)))$, $x \geq a$, generates a finite modulus;

7) for all (equivalently, for some) $a \in \mathbf{R}$, the function $V(x) = F^{-1}(\exp(-\exp(-x)))$, $x \geq a$, generates a finite modulus. In addition, for any $p \in (0,1)$, and $h > 0$,

$$\sup_n \mathbf{P}\{M_n - m_p(M_n) > V_p^*(h)\} \leq \mathbf{P}\{Z - m_p(Z) > h\}, \tag{2.1}$$

where Z is a random variable with distribution function $\mathbf{P}\{Z \leq x\} = \exp(-\exp(-x))$, and where V_p^* is a modulus generated by V on the interval $[-\log\log(1/p), +\infty)$;

8) for any $p \in (0,1)$, there exists $\varepsilon > 0$ such that $\sup_n \mathbf{E}\exp\{\varepsilon(M_n - m_p(M_n))\} < +\infty$;

9) for all (equivalently, for some) $\alpha \geq 1$, and for all (equivalently, for some) $p \in (0,1)$, $\sup_n \mathbf{E}\left((M_n - m_p(M_n))^+\right)^\alpha < +\infty$.

Proof of Proposition 2.1.

1a) $\Longrightarrow$ **7)**. Let p and $h > 0$ be such that

$$q = \inf_n \mathbf{P}\{M_n - m_p(M_n) \leq h\} > p,$$

that is, such that $F(F^{-1}(p^{1/n}) + h) \geq q^{1/n}$. By the very definition of F^{-1}, p and h are such that $F^{-1}(p^{1/n}) + h \geq F^{-1}(q^{1/n})$, which is rewritten

$$F^{-1}(q^{1/n}) - F^{-1}(p^{1/n}) \leq h. \tag{2.2}$$

Putting $a = -\log\log(1/p)$, $b = -\log\log(1/q)$, (2.2) can also be rewritten as

$$V(b + \log(n)) - V(a + \log(n)) \leq h, \tag{2.3}$$

which holds for all $n \geq 1$. Clearly, V is a non–decreasing function on $\mathbf{R}$, and from (2.3), it follows that,

$$V_p^*(h) = \sup\{V(y) - V(x) : a \leq x \leq y,\, y - x \leq h\} < +\infty, \tag{2.4}$$

whenever $h > 0$. Now, fix any real number c such that $1 < c < \exp(b-a)$, and let $n_0 = n_0(a,b)$ be any positive integer such that $\log(n_0+1) - \log(n_0) \leq b - a$ and $(\exp(b-a) - c)n_0 \geq 1$. Let $a \leq x \leq y$. If $x < a + \log(n_0)$, then one has the estimate

$$\begin{aligned} V(y) - V(x) &= (V(y) - V(a + \log(n_0)) + (V(a + \log(n_0)) - V(x)) \\ &\leq (V(y) - V(a + \log(n_0))) + V(a + \log(n_0)). \end{aligned}$$

Hence,

$$\sup\{V(y)-V(x): a\le x\le y,\, y-x\le h\}\le$$
$$\sup\{V(y)-V(x): a+\log(n_0)\le x\le y,\, y-x\le h\}+V(a+\log(n_0)).$$

Note that for non–decreasing functions V, the property "V generates a finite modulus on $[a,+\infty)$" does not depend on a. Therefore, in order to prove (2.1), without loss of generality it is assumed that $a+\log(n_0)\le x\le y$. Now define a sequence n_k, $k\ge 1$, recursively, in the following way: let n_1 be the largest integer such that $a+\log(n_1)\le x$. If $k\ge 1$, let n_{k+1} be the largest integer such that $a+\log(n_{k+1})\le b+\log(n_k)$. By the definition of n_0, we then have $n_0\le n_k<n_{k+1}$, for all $k\ge 1$, since $a+\log(n_k+1)\le b+\log(n_k)$ for, $n_k\ge n_0$, provided $a+\log(n_0+1)\le b+\log(n_0)$.

Denote by K the smallest k such that $b+\log(n_k)\ge y$. Since, by construction, the intervals $[a+\log(n_k), b+\log(n_k)]$, $1\le k\le K$, cover the interval $[x,y]$, we get

$$V(y)-V(x)\le\sum_{k=1}^{K}V(b+\log(n_k))-V(a+\log(n_k))\le CK,$$

where $C=\sup_n V(b+\log(n))-V(a+\log(n))$. Our aim is now to find an estimate of K depending on $y-x$; we would then have an estimate for $V(y)-V(x)$ in terms of $y-x$. Denote by $[u]$ the integer part of a real u. Then, $n_{k+1}=\left[e^{b-a}n_k\right]$, hence

$$n_{k+1}\ge e^{b-a}n_k-1\ge cn_k,$$

since $n_k\ge n_0$ and $\exp(b-a)n_0\ge cn_0$. By induction, it is easy to see that $n_k\ge c^{k-1}n_1$, that is, $\log(n_k)\ge(k-1)\log(c)+\log(n_1)$. Thus, the inequality $b+\log(n_k)\ge y$ follows from $b+(k-1)\log(c)+\log(n_1)\ge y$, which holds if $k\ge 1+(y-b-\log(n_1))/\log(c)$. By the definition of n_1, we also have $\log(n_1)\le x-a$, so it suffices to take $k\ge 1+((y-x)-(b-a))/\log(c)$. Hence,

$$K\le 2+\frac{(y-x)-(b-a)}{\log(c)}\le 2+\frac{y-x}{\log(c)},$$

and (2.1) is proved for all $p\in(0,1)$.

To prove the second part of 7), fix $p\in(0,1)$, $h>0$, and set $q=\mathbf{P}\{Z-m_p(Z)\le h\}$, $a=-\log\log(1/p)$, $b=-\log\log(1/q)$. Then, as easily checked, $b-a=h$. As previously seen, inequalities of the form

$$\mathbf{P}\{M_n-m_p(M_n)\le V_p^*(h)\}\ge q \tag{2.5}$$

are equivalent to

$$V(b+\log(n))-V(a+\log(n))\le V_p^*(h)$$

(this is (2.3) with $V_p^*(h)$ instead of h). Since $b-a=h$, the above inequality holds true by the very definition of V_p^*. It just remains to note that (2.5) and (2.1) coincide.

The first part of 7 $\Longrightarrow$ 1a). Let $a=-\log\log(1/p)$. Then, as shown in the previous steps, 1a) holds if and only if (2.4) holds for some $b>a$, some $h>0$, and all $n\geq 1$. But (2.4) holds for all $b>a$ and $h>0$ since $V(b+\log(n))-V(a+\log(n))\leq V_p^*(b-a)$.

1a) $\Longleftrightarrow$ 2a) is clear recalling the above description and noting that, for a non–decreasing function V, the property "V generates a finite modulus on $[a,+\infty)$" does not depend on a.

6) $\Longleftrightarrow$ 7). Note that $U(x)=V(T(x))$, where $T(x)=-\log(\exp(e^{-x})-1)$. Then, T is an increasing bijection from $\mathbf{R}$ to $\mathbf{R}$, and has a finite Lipschitz constant on every interval $[a,+\infty)$, and similarly, for its inverse T^{-1}. Therefore, U generates a finite modulus on $[a,+\infty)$ if and only if V generates a finite modulus on $[a,+\infty)$.

1a), 2a), 3a) $\Longleftrightarrow$ 1b), 2b), 3b). Simply note (recalling (2.2)) that, for all $0<p<q<1$ and all $h>0$.

$$\mathbf{P}\{M_n-m_p(M_n)\leq h\}\geq q \iff \mathbf{P}\{M_n-m_q(M_n)\leq -h\}\leq p.$$

7) $\Longrightarrow$ 3a) $\Longrightarrow$ 2a) $\Longrightarrow$ 1a). Indeed, (2.1) implies 3a) which is stronger than 2a) which is stronger than 1a).

1a), 2a), 3a) $\Longleftrightarrow$ 5). Note that $m_p(M_n)=F^{-1}(p^{1/n})$. Then, as previously shown (see (2.2)), for all $0<p<q<1$, and $h>0$, $m_q(M_n)-m_p(M_n)\leq h$ is equivalent to $\mathbf{P}\{M_n-m_p(M_n)\leq h\}\geq q$.

3a), 3b) $\Longrightarrow$ 4). Take $a_n=m_p(M_n)$.

4) $\Longrightarrow$ 1a). By assumption, there exists h_0 such that $\mathbf{P}\{|M_n-a_n|>h_0\}<1/2$, for all $n\geq 1$. Hence, for $p=1/2$, we have $|m_p(M_n)-a_n|\leq h_0$. Therefore,

$$\sup_n \mathbf{P}\{M_n-m_p(M_n)>h\}\leq \sup_n \mathbf{P}\{M_n-a_n>h-h_0\}\to 0,\ \ as\ h\to+\infty.$$

7) $\Longrightarrow$ 8) Clear by (2.1).

8) $\Longrightarrow$ 9) Immediate.

9) $\Longrightarrow$ 3a) Use Chebyshev's inequality.

Proposition 2.1 is proved.

3 Proof of Theorem 1.3

Necessity. Assume that there exist $h > 0$, and $0 < p < q < 1$, such that $\inf_n R_h^{(n)}(p) \geq q$. Thus, for all integers $n \geq 1$ and for all Borel sets $A \subset \mathbf{R}^n$ with $\mu^n(A) \geq p$, we have

$$\mu^n(A^h) \geq q. \tag{3.1}$$

Let ξ_n, $n \geq 1$, be a sequence of independent random variables with a common distribution function F_μ, and let $M_n = \max\{\xi_1, \cdots, \xi_n\}$. First, apply the inequality (3.1) to the cubes $A_n(p) = (-\infty, F_\mu^{-1}(p^{1/n})]^n$. Since $\mu^n(A_n(p)) = F_\mu^n(F_\mu^{-1}(p^{1/n})) \geq p$, we get

$$\mathbf{P}\{M_n - m_p(M_n) \leq h\} = F_\mu^n(F_\mu^{-1}(p^{1/n}) + h) = \mu^n(A_n(p)^h) \geq q,$$

that is, the property 1a) from Proposition 2.1 is fulfilled. Therefore, so is the property 6): the function U_μ generates a finite modulus on the interval $[0, +\infty)$. Now, apply (3.1) to the cubes $B_n(p) = [-F_\mu^{-1}(1 - p^{1/n}), +\infty)^n$. In a similar way, we observe that the property 1a) from Proposition 2.1 is satisfied for the random variables $-\xi_n$, $n \geq 1$, whose distribution function is given by $G(x) = 1 - F_\mu((-x)^-)$. Since $G^{-1}(p) = -F_\mu^{-1}(1-p)$, we obtain the property 6) for the function

$$W(x) = G^{-1}(1/(1 + \exp(-x))) = -F_\mu^{-1}(1 - 1/(1 + \exp(-x))) = -U_\mu(-x).$$

The above identity implies that W generates a finite modulus on the interval $[0, +\infty)$ if and only if U_μ generates a finite modulus on the interval $(-\infty, 0]$. Thus, U_μ generates a finite modulus on the intervals $[0, +\infty)$ and $(-\infty, 0]$, and therefore it generates a finite modulus on the whole real line.

Sufficiency. It is known that, for the measure ν with $\nu((-\infty, x]) = 1/(1 + \exp(-x))$,

$$\nu^n(A^h) \geq \frac{p}{p + (1-p)\exp(-h)}, \tag{3.2}$$

whenever $\nu^n(A) \geq p$, with equality at the standard half-spaces $A = \{x : x_1 \leq \mathit{const}\}$ (different proofs of (3.2) can respectively be found in [1] and in [2, Corollary 15.3]). Introduce the function $i(x_1, \cdots, x_n) = (U_\mu(x_1), \cdots, U_\mu(x_n))$ which transforms ν^n into μ^n. Let $h > 0$, $h^* = U_\mu^*(h)$. Observe the following inclusion: for any set $A \subset \mathbf{R}^n$,

$$\left(i^{-1}(A)\right)^h \subset i^{-1}\left(A^{h^*}\right). \tag{3.3}$$

Indeed, if $x \in (i^{-1}(A))^h$, then for some $y \in i^{-1}(A)$ we have $\rho_\infty(x, y) \leq h$, that is, $|x_k - y_k| \leq h$, for all $1 \leq k \leq n$. Since $i(y) = (U_\mu(y_1), \cdots, U_\mu(y_n)) \in A$ and since $|U_\mu(x_k) - U_\mu(y_k)| \leq h^*$, we get $\rho_\infty(i(x), i(y)) \leq h^*$, and therefore $i(x) \in A^{h^*}$, hence,

$x \in i^{-1}\left(A^{h^*}\right)$. Now, combine (3.2) and (3.3) to prove (1.2). Let $\nu^n(A) = \mu^n(i^{-1}(A)) \geq p$. Then,

$$\nu^n\left(A^{h^*}\right) = \mu^n\left(i^{-1}(A^{h^*})\right) \geq \mu^n\left(\left(i^{-1}(A)\right)^h\right) \geq \frac{p}{p+(1-p)\exp(-h)}.$$

Theorem 1.1 is proved.

4 Proof of Theorem 1.5

Sufficiency. Assume that 1b) is true. Then, for the sequence $a_n = \mathbf{E}M_n$, or $a_n = m_p(M_n)$, Chebyshev's inequality implies that $\sup_n \mathbf{P}\{|M_n - a_n| > h\} \to 0$, as $h \to +\infty$. Hence, the property 4) of Proposition 2.1 holds true. Therefore, the function U_μ generates a finite modulus on the interval $[0, +\infty)$. The assumption 2b) is just 1b) for the sequence $(-\xi_n)$, $n \geq 1$, hence, again by Proposition 2.1, the function U_μ generates a finite modulus on the interval $(-\infty, 0]$. As a result, U_μ generates a finite modulus on the whole real line. It now remains to make use of Theorem 1.3 to obtain the sufficiency part of the proof.

Necessity. As before, let $M_n = \max\{\xi_1, \cdots, \xi_n\}$. Let ζ_n, $n \geq 1$, be a sequence of independent random variables with common (logistic) distribution ν, and let $Z_n = \max\{\zeta_1, \cdots, \zeta_n\}$. Since U_μ transforms ν into μ, M_n and $U_\mu(Z_n)$ are identically distributed. Therefore,

$$\mathbf{E}\,|M_n - M_n'|^\alpha = \mathbf{E}\,|U_\mu(Z_n) - U_\mu(Z_n')|^\alpha, \tag{4.1}$$

where (M_n', Z_n') is an independent copy of (M_n, Z_n). By Theorem 1.3, there exist constants $a, b \geq 0$ such that $|U_\mu(x) - U_\mu(y)| \leq a + b|x-y|$, whenever $x, y \in \mathbf{R}$. thus, for all $\alpha \geq 1$, the left–hand side of (4.1) is bounded when $n \to \infty$, if the same is true of Z_n instead of M_n. That is, we reduced our tentative proof to the case $\mu = \nu$. So, one may assume that $\xi_n = \zeta_n$, and that $M_n = Z_n$, for all n. By Remark 1.4 (in the particular case where $f(x) = \max_{1 \leq k \leq n} x_k$), there exist Lipschitz functions $L_n : \mathbf{R} \to \mathbf{R}$, such that the random variables M_n and $L_n(\xi_1)$ are identically distributed (of course, in this particular case, this is easily verified directly). Therefore, for all $n \geq 1$,

$$\mathbf{E}\,|M_n - M_n'|^\alpha = \mathbf{E}\,|L_n(\xi_1) - L_n(\xi_1')|^\alpha \leq \mathbf{E}\,|\xi_1 - \xi_1'|^\alpha = C_\alpha,$$

where ξ_1' is an independent copy of ξ_1. Now by Hölder's inequality,

$$\mathbf{E}\,|M_n - \mathbf{E}\,M_n|^\alpha = \mathbf{E}\,|\mathbf{E}'\,(M_n(\omega) - M_n'(\omega'))|^\alpha \leq \mathbf{E}\,|M_n - M_n'|^\alpha \leq C_\alpha,$$

where $\mathbf{E}'$ is taken with respect to the random variable M_n'. This proves 1b). The property 2b) is proved in a similar way, taking into account that ν is symmetric about 0.

In order to prove the last statement on the quantiles, one can apply (1.2) to the cubes $\{x : x_i \le const$ for all $i \le n\}$. This gives

$$\mathbf{P}\{M_n - m_p(M_n) > h^*\} \le \mathbf{P}\{\zeta_1 - m_p(\zeta_1) > h\},$$

$$\mathbf{P}\{M_n - m_p(M_n) < -h^*\} \le \mathbf{P}\{\zeta_1 - m_p(\zeta_1) < -h\},$$

for all $p \in (0,1)$, $h > 0$. Since $h^* \le a + bh$, these inequalities immediately imply that

$$\sup_n \mathbf{E}\,|M_n - m_p(M_n)|^\alpha < +\infty.$$

Theorem 1.5 follows.

5 Concentration of maxima in L^1–norm

Again, let ξ_n, $n \ge 1$, be a sequence of independent random variables with a common distribution function F (which is assumed right–continuous). The associated measure will be also denoted by F, and again $M_n = \max\{\xi_1, \cdots, \xi_n\}$.
The function

$$F^{-1}(p) = \inf\{x \in \mathbf{R} : F(x) \ge p\},$$

which is the minimal quantile of order p of F, is well-defined on $(0,1)$, non–decreasing and left–continuous. It therefore generates a non–negative Lebesgue–Stieltjes measure on $(0,1)$ which we also denote by F^{-1}. In particular, $F^{-1}([a,b)) = F^{-1}(b) - F^{-1}(a)$, for $0 < a \le b < 1$. For example, if $F = \delta_x$ (as measures) is a unit mass at a point x, then $F^{-1}(p) = x$, for all $p \in (0,1)$, hence as a measure, $F^{-1} = 0$. If F is a Bernoulli measure, $F = \frac{1}{2}\delta_x + \frac{1}{2}\delta_y$, $x < y$, and then $F^{-1} = (y-x)\delta_{1/2}$.

We return to Theorem 1.5 to express the property 1b), for $\alpha = 1$, in terms of F^{-1}.

Theorem 5.1 *The following are equivalent:*
1) $\sup_n \mathbf{E}\,|M_n - \mathbf{E}\,M_n| < +\infty$;
2) $\int_0^1 p^n(1-p)\,dF^{-1}(p) = O\left(\frac{1}{n}\right)$, *as* $n \to \infty$;
3) $\int_0^1 p(1-p)\,dF^{-1}(p) < +\infty$, *and there exist* $a, b \ge 0$ *such that, for all* $1/2 \le p \le q < 1$,

$$F^{-1}(q) - F^{-1}(p) \ \le\ a + b\log\frac{1-p}{1-q}. \tag{5.1}$$

The condition (5.1) in 3) expresses the fact that the function $U(x) = F^{-1}(1/(1+\exp(-x))$ generates a finite modulus on the interval $[0,+\infty)$. This implies that $\mathbf{E}\,\xi_1^+ < +\infty$, and moreover $\mathbf{E}\,\exp(\varepsilon\xi_1) < +\infty$, for some $\varepsilon > 0$. But this says nothing about the behavior of F at $-\infty$. This behavior is regulated by the first condition in 3) which is equivalent to $\mathbf{E}\,|\xi_1| < +\infty$.

Before proving this theorem, we prove an elementary lemma. Below, the right integral in (5.3) is taken over the semi–open interval $[p^{1/n}, 1)$ (the measure F^{-1} can have an atom at the point $p^{1/n}$); while the right integral in (5.4) is taken over the open interval $(0, p^{1/n})$.

Lemma 5.2 *Let M_n' be an independent copy of M_n, $p \in (0,1)$. Then,*

$$\mathbf{E}\,|M_n - M_n'| = 2\int_0^1 t^n(1-t^n)\,dF^{-1}(t). \tag{5.2}$$

$$\mathbf{E}\,(M_n - m_p(M_n))^+ = \int_{p^{1/n}}^1 (1-t^n)\,dF^{-1}(t). \tag{5.3}$$

$$\mathbf{E}\,(M_n - m_p(M_n))^- = \int_0^{p^{1/n}} t^n\,dF^{-1}(t). \tag{5.4}$$

Proof. We use the following identity

$$\int_0^1 h(F^{-1}(t))\,dt = \int_{-\infty}^{+\infty} h(x)\,dF(x), \tag{5.5}$$

where h is an arbitrary non–negative Borel measurable function on the real line. To prove (5.5), it suffices to verify it on the indicator functions $h = \chi_{(x,y]}$, $x < y$. The right–hand side of (5.5) becomes $F(y) - F(x)$, while the left–hand side is the length of the interval of all points $t \in (0,1)$ such that $x < F^{-1}(t) \le y$. But, by the very definition of F^{-1}, these inequalities are equivalent to $F(x) < t \le F(y)$. This proves (5.5). Now by (5.5),

$$\begin{aligned}\mathbf{E}\,|M_n - M_n'| &= \int_{\mathbf{R}}\int_{\mathbf{R}} |x-y|\,dF(x)^n dF(y)^n = 2\iint_{x<y} (y-x)\,dF(x)^n dF(y)^n \\ &= 2\iint_{0<t<s<1} (F^{-1}(s) - F^{-1}(t))\,dt^n ds^n.\end{aligned}$$

Since $F^{-1}(t)$ is left–continuous, an integration by parts over t while $s \in (0,1)$ is fixed, and taking into account that $tF^{-1}(t) \to 0$, as $t \to 0^+$, gives

$$\int_0^s (F^{-1}(s) - F^{-1}(t))\,dt^n = \int_0^s t^n\,dF^{-1}(t).$$

Hence, by Fubini's Theorem,

$$\iint_{0<t<s<1} (F^{-1}(s) - F^{-1}(t))\,dt^n ds^n = \iint_{0<t<s<1} t^n\,dF^{-1}(t) ds^n = \int_0^1 t^n(1-t^n)\,dF^{-1}(t).$$

This, in turn, proves (5.2). Analogously, since $m_p(M_n) = F^{-1}(p^{1/n})$,

$$\begin{aligned}\mathbf{E}\,(M_n - m_p(M_n))^+ &= \int_{\mathbf{R}} (x - F^{-1}(p^{1/n}))^+ \, dF(x)^n = \int_{p^{1/n}}^1 (F^{-1}(t) - F^{-1}(p^{1/n}))\, dt^n \\ &= \int_{[p^{1/n},1)} (1 - t^n)\, dF^{-1}(t).\end{aligned}$$

$$\begin{aligned}\mathbf{E}\,(M_n - m_p(M_n))^- &= \int_{\mathbf{R}} (x - F^{-1}(p^{1/n}))^- \, dF(x)^n = \int_0^{p^{1/n}} (F^{-1}(p^{1/n}) - F^{-1}(t))\, dt^n \\ &= \int_0^{p^{1/n}} t^n \, dF^{-1}(t),\end{aligned}$$

and the lemma is proved.

Proof of Theorem 5.1. The conditions 1), 2) and 3) require $\mathbf{E}\,|\xi_1| < +\infty$, so this is assumed throughout. In particular, $t(1-t)F^{-1}(t) \to 0$, as $t \to 0^+$ and as $t \to 1^-$.
1) $\Longleftrightarrow$ **2)**. Clearly, for all $t \in (0,1)$,

$$nt^{2n-1}(1-t) \le t^n(1-t^n) \le nt^n(1-t). \tag{5.6}$$

Therefore, by (5.2), the property 2) is equivalent to $A = \sup_n \mathbf{E}\,|M_n - M_n'| < +\infty$. Indeed, by the right inequality in (5.6), $2A \le B = \sup_n \left[n \int_0^1 t^n (1-t)\, dF^{-1}(t) \right]$. Now, by the left inequality in (5.6), for $k = 2n-1$ odd,

$$k \int_0^1 t^k (1-t)\, dF^{-1}(t) \le 2n \int_0^1 t^{2n-1}(1-t)\, dF^{-1}(t) \le 2 \int_0^1 t^n (1-t^n)\, dF^{-1}(t) \le A,$$

while for $k = 2n$ even,

$$k \int_0^1 t^k (1-t)\, dF^{-1}(t) \le k \int_0^1 t^{k-1}(1-t)\, dF^{-1}(t) = 2n \int_0^1 t^{2n-1}(1-t)\, dF^{-1}(t) \le A,$$

as above. Therefore, $B \le A$. Then, it remains to note that

$$\mathbf{E}\,|M_n - \mathbf{E}\, M_n| \le \mathbf{E}\,|M_n - M_n'| \le 2\mathbf{E}\,|M_n - \mathbf{E}\, M_n|.$$

1) $\Longrightarrow$ **3)** This is clear from Proposition 2.1, since the property 9) holds with $\alpha = 1$.

3) $\Longrightarrow$ **1)**. By Theorem 1.5, this result is true for F symmetric about a point, and we have reduced the general case to the symmetric one. Let $p = 1/2$. Let F_0 be the distribution function which (as a measure) is symmetric about the median $m = m_p(\xi_1)$ and such that $F_0(x) = F(x)$, for all $x \ge m$. Therefore, $F_0^{-1}(t) = F^{-1}(t)$, for all $t \in [1/2, 1)$. By assumption, the function $U_0(x) = F_0^{-1}(1/(1+\exp(-x))$ generates a finite modulus on the interval $[0, +\infty)$, hence, since $U_0(-x) = -U_0(x)$, for all x real, it generates a finite modulus on the whole real line. Hence, by Theorem 1.5, the property 1) of Theorem 5.1 holds true with respect to F_0. Thus,

$$\sup_n \mathbf{E}\,|Q_n - Q_n'| < +\infty,$$

where $Q_n = \max\{\zeta_1, \cdots, \zeta_n\}$, where $(\zeta_n)_{n\geq 1}$ are independent random variables with common distribution function F_0, and where Q_n' is an independent copy of Q_n. As well-known, for $p = 1/2$,

$$\mathbf{E}\,|Q_n - m_p(Q_n)| \leq \mathbf{E}\,|Q_n - \mathbf{E}\,Q_n|,$$

hence, $\sup_n \mathbf{E}\,|Q_n - m_p(Q_n)| < +\infty$. Thus, (5.3) and (5.4) imply that for some constant C, and for all n,

$$\mathbf{E}\,(Q_n - m_p(Q_n))^+ = \int_{p^{1/n}}^1 (1 - t^n)\, dF_0^{-1}(t) \leq C, \tag{5.7}$$

$$\mathbf{E}\,(Q_n - m_p(Q_n))^- = \int_0^{p^{1/n}} t^n\, dF_0^{-1}(t) \leq C. \tag{5.8}$$

Since $p^{1/n} \geq 1/2$, we have $F_0^{-1}(t) = F^{-1}(t)$, for all $t \in [p^{1/n}, 1)$, and therefore (5.7) is satisfied for M_n and F instead of Q_n and F_0. For similar reasons, the sequence

$$\int_{1/2}^{p^{1/n}} t^n\, dF_0^{-1}(t) = \int_{1/2}^{p^{1/n}} t^n\, dF^{-1}(t)$$

is bounded. But

$$\int_0^{1/2} t^n\, dF^{-1}(t) \leq \int_0^{1/2} t\, dF^{-1}(t) = \mathbf{E}\,\xi_1^- < +\infty,$$

hence, (5.8) is also satisfied for M_n and F instead of Q_n and F_0 (maybe with a different constant C'). Again, by (5.3) and (5.4), we have that the sequence

$$\mathbf{E}\,|M_n - m_p(M_n)| = \mathbf{E}\,(M_n - m_p(M_n))^+ + \mathbf{E}\,(M_n - m_p(M_n))^-$$

is bounded. As a result,

$$\mathbf{E}\,|M_n - M_n'| \leq \mathbf{E}\,|M_n - m_p(M_n)| + \mathbf{E}\,|M_n' - m_p(M_n)| = 2\mathbf{E}\,|M_n - m_p(M_n)|$$

is also bounded, and so is the sequence $\mathbf{E}\,|M_n - \mathbf{E}\,M_n|$. Theorem 5.1 follows.

6 Proof of Theorem 1.6

We start by proving another elementary lemma.

Lemma 6.1 *For any non–negative Lebesgue–Stieltjes measure λ on $(0,1)$, there exists a distribution function F such that $F^{-1} = \lambda$ (as measures).*

Proof. Let $H(p) = \lambda([1/2, p))$, for $1/2 \le p < 1$, and let $H(p) = -\lambda([p, 1/2))$, for $0 \le p \le 1/2$. Clearly, H is non–decreasing and left–continuous. Now, it suffices to find a distribution function F with $F^{-1}(p) = H(p)$, for all $p \in (0,1)$. But, one can choose

$$F(x) = \sup\{p \in (0,1) : H(p) \le x\},$$

(with the convention $\sup \emptyset = 0$). The sup above can be replaced by max, since H is left–continuous. Therefore, $F(x) \ge p \Longleftrightarrow H(q) \le x$, for some $q \ge p$, hence

$$F^{-1}(p) = \min\{x \in \mathbf{R} : F(x) \ge p\} = \min\{x \in \mathbf{R} : H(q) \le x, \text{ for some } q \ge p\} = H(p).$$

Lemma 6.1 follows.

Let us now come back to the proof of Theorem 1.6. By Lemma 6.1, it is possible to find F such that the Radon–Nikodym derivative

$$\frac{dF_\xi(p)}{dF^{-1}(p)} = 1 - p, \quad 0 < p < 1,$$

and then, the equivalence between (1.3) and (1.4) becomes the equivalence between the properties 2) and 3) in Theorem 5.1. Thus, Theorem 1.6 immediately follows from Theorem 5.1.

7 Appendix

Let (X, ρ, μ) be a metric space equipped with a Borel probability measure μ. As in Section 1, define the open h–neighbourhood of a set $A \subset X$ as

$$A^h = \{x \in \mathbf{R}^n : \rho(x, a) < h \text{ for some } a \in A\}, \ h > 0,$$

and the associated ("integral" isoperimetric) function

$$R_h(p) = \inf_{\mu(A) \ge p} \mu(A^h), \quad 0 < p < 1,$$

where the infimum is taken over all the Borel sets of measure $\mu(A) \ge p$.

Theorem 7.1 $R_h(p) \to 1$, *as* $h \to +\infty$, *whenever* $p \in (0,1)$.

Proof. First note that, given a random variable ξ, there always exists a concave increasing function $U : [0, +\infty) \to [0, +\infty)$, with $U(0) = 0$, $U(+\infty) = +\infty$, such that

$$\mathbf{E}\,(U(\xi))^2 < +\infty.$$

Now, take $\xi = \rho$ (as a random variable on the probability space $(X \times X, \mu \otimes \mu)$) and take such a U. Since U is a modulus of continuity, the function $d(x, y) = U(\rho(x, y))$ is also a metric, and the enlargement $A^{U(h)}$ (with respect to the metric d) coincides with the enlargement A^h (with respect to the metric ρ). Hence, one can assume in Theorem 7.1 (replacing ρ by d if needed) that

$$\sigma^2 = \frac{1}{2} \int_X \int_X \rho^2(x, y)\, d\mu(x) d\mu(y) < +\infty.$$

In particular, on the probability space (X, μ),

$$\mathbf{Var}(f) = \frac{1}{2} \int_X \int_X |f(x) - f(y)|^2\, d\mu(x) d\mu(y) \leq \sigma^2 < +\infty,$$

for any Lipschitz function f on X (with Lipschitz constant ≤ 1). The functions

$$f_A(x) = dist(x, A) = \inf\{\rho(a, x) : a \in A\}$$

are Lipschitz, of Lipschitz constant at most 1. Therefore, $\mathbf{Var}(f_A) \leq \sigma^2$, and by definition $\mu\{x \in X : f_A(x) = 0\} \geq p$, if $\mu(A) \geq p$. But, it is easy to verify that for any non–negative random variable ξ with $\mathbf{Var}(\xi) \leq \sigma^2$, and $\mu\{\xi = 0\} \geq p$,

$$\mathbf{E}\,\xi \leq c(p) = \sigma \sqrt{\frac{p}{1-p}}.$$

Hence, by Chebyshev's inequality, for all $h > 0$,

$$\mu\{f_A \geq c(p) + h\} \leq \frac{\sigma^2}{h^2}. \tag{7.1}$$

Since the right–hand side of (7.1) does not depend on A, and since $A^h = \{x \in X : f_A(x) < h\}$, we have, for $h > c(p)$,

$$1 - R_h(p) \leq \frac{\sigma^2}{(h - c(p))^2} \to 0, \quad as\ h \to +\infty.$$

This proves Theorem 7.1.

The statement of Theorem 7.1 remains true if the metric is replaced by a pseudo–metric ρ such that $\rho(x, y) < +\infty$, for almost all (x, y) with respect to measure $\mu \otimes \mu$. In Section

1, the concentration property $\inf_n R_h^{(n)}(p) > p$ could be expressed as $R_h(p) > p$, for the space $X = \mathbf{R}^\infty$ equipped with the pseudo–metric

$$\rho_\infty(x,y) = \sup_{i \geq 1} |x_i - y_i|,$$

and the product measure μ^∞. In this case, $\rho(x,y) = +\infty$, for almost all (x,y) with respect to μ^∞, whenever the measure μ has non–compact support, on the real line.

References

1. Bobkov, S.G. Extremal properties of half-spaces for log-concave distributions. Center for stochastic processes, Dept. of Statistics, Univ. of North Carolina at Chapel Hill. Tech. Report No 396 (1993).

2. Bobkov, S.G., Houdré, Ch. Some connections between Sobolev-type inequalities and isoperimetry. School of Mathematics, GaTech, Tech. Report No 008 (1995).

3. Bobkov, S.G., Houdré, Ch. A characterization of Gaussian measures via the isoperimetric property of half-spaces. School of Mathematics, Gatech, Tech. Report No 014 (1995). To appear in *Problems in the Theory of Probability distributions. Zap. Nauchn. Semin. Sankt–Petersburg's branch of the Steklov's Math. Inst.* (In Russian).

4. Talagrand, M. A new isoperimetric inequality and the concentration of measure phenomenon. *Israel Seminar (GAFA), Springer Verlag, Lect. Notes in Math.* 1469 (1991), 94–124.

Sergei. G. Bobkov
Department of Mathematics
Syktyvkar University
167001 Syktyvkar
Russia

Christian Houdré
Center for Applied Probability
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30322, USA