Between Sobolev and Poincaré *

Rafał Latała and Krzysztof Oleszkiewicz
(Warsaw)

Abstract

Let $a \in [0,1]$ and $r \in [1,2]$ satisfy relation $r = 2/(2-a)$. Let $\mu(dx) = c_r^n \exp(-(|x_1|^r + |x_2|^r + \ldots + |x_n|^r))dx_1 dx_2 \ldots dx_n$ be a probability measure on the Euclidean space $(R^n, \|\cdot\|)$. We prove that there exists a universal constant C such that for any smooth real function f on R^n and any $p \in [1,2)$

$$E_\mu f^2 - (E_\mu |f|^p)^{2/p} \leq C(2-p)^a E_\mu \|\nabla f\|^2.$$

We prove also that if for some probabilistic measure μ on R^n the above inequality is satisfied for any $p \in [1,2)$ and any smooth f then for any $h : R^n \longrightarrow R$ such that $|h(x) - h(y)| \leq \|x - y\|$ there is $E_\mu |h| < \infty$ and

$$\mu(h - E_\mu h > \sqrt{C} \cdot t) \leq e^{-Kt^r}$$

for $t > 1$, where $K > 0$ is some universal constant.

Let us begin with few definitions.

Definition 1 *Let (Ω, μ) be a probability space and let f be a measurable, square integrable non-negative function on Ω. For $p \in [1,2)$ we define the p-variance of f by*

$$Var(p)_\mu(f) = \int_\Omega f(x)^2 \mu(dx) - (\int_\Omega f(x)^p \mu(dx))^{2/p} = E_\mu f^2 - (E_\mu f^p)^{2/p}.$$

Note that $Var(1)_\mu(f) = D^2_\mu(f) = Var_\mu(f)$ coincides with classical notion of variance, while

$$\lim_{p \to 2^-} \frac{Var(p)_\mu(f)}{2-p} = \frac{1}{2}(E_\mu f^2 \ln(f^2) - E_\mu f^2 \cdot \ln(E_\mu f^2)) = \frac{1}{2} Ent_\mu(f^2),$$

where Ent_μ denotes a classical entropy functional (see [L] for a nice introduction to the subject).

*Research partially supported by KBN Grant 2 P03A 043 15

Definition 2 *Let $\mathcal{E}$ be a non-negative functional on some class $\mathcal{C}$ of non-negative functions from $L^2(\Omega, \mu)$. We will say that $f \in \mathcal{C}$ satisfies*

- *the Poincaré inequality with constant C if $Var_\mu(f) \leq C \cdot \mathcal{E}(f)$,*
- *the logarithmic Sobolev inequality with constant C if $Ent_\mu(f^2) \leq C \cdot \mathcal{E}(f)$,*
- *the inequality $I_\mu(a)$ (for $0 \leq a \leq 1$) with constant C if $Var(p)_\mu(f) \leq C \cdot (2-p)^a \cdot \mathcal{E}(f)$ for all $p \in [1, 2)$.*

Lemma 1 *For a fixed $f \in \mathcal{C}$ and $p \in [1, 2)$ let*

$$\varphi(p) = \frac{Var(p)_\mu(f)}{1/p - 1/2}.$$

Then φ is a non-decreasing function.

Proof. Hölder's inequality yields that $\alpha(t) = t \ln(E_\mu f^{1/t})$ is a convex function for $t \in (1/2, 1]$. Hence also $\beta(t) = e^{2\alpha(t)} = (E_\mu f^{1/t})^{2t}$ is convex and therefore $\frac{\beta(t)-\beta(1/2)}{t-1/2}$ is non-decreasing on $(1/2, 1]$. Observation that

$$\varphi(p) = \frac{\beta(1/2) - \beta(1/p)}{1/p - 1/2}$$

completes the proof. $\square$

Corollary 1 *For $f \in \mathcal{C}$ the following implications hold true:*

- *f satisfies the Poincaré inequality with constant C if and only if f satisfies $I_\mu(0)$ with constant C,*
- *if f satisfies the logarithmic Sobolev inequality with constant C then f satisfies $I_\mu(1)$ with constant C,*
- *if f satisfies $I_\mu(1)$ with constant C then f satisfies the logarithmic Sobolev inequality with constant $2C$,*
- *if f satisfies $I_\mu(a)$ with constant C and $0 \leq \alpha \leq a \leq 1$ then f satisfies $I_\mu(\alpha)$ with constant C.*

Proof.

- To prove the first part of Corollary 1 it suffices to note that $p \longmapsto Var(p)_\mu(f)$ is a non-increasing function.

- The second part of Corollary 1 follows easily from the fact that

$$\lim_{p\to 2^-} \frac{Var(p)_\mu(f)}{2-p} = \frac{1}{2}\cdot Ent_\mu(f^2).$$

- To prove the third part of Corollary 1 use Lemma 1 and note that for $p \in [1,2)$ we have

$$\frac{Var(p)_\mu(f)}{2-p} = \frac{\varphi(p)}{2p} \leq \frac{\lim_{p\to 2^-}\varphi(p)}{2} = Ent_\mu(f^2).$$

- The last part of statement is trivial. □

Corollary 1 shows that inequalities $I_\mu(a)$ interpolate between Poincaré and logarithmic Sobolev inequalities. Note that $I_\mu(a)$ for $a<0$ would be equivalent to the Poincaré inequality and the only functions satisfying $I_\mu(a)$ for $a>1$ would be the constant functions (because in this case $I_\mu(a)$ would imply the logarithmic Sobolev inequality with constant 0). Therefore restriction to $a\in[0,1]$ is natural.

Definition 3 *Given probability space (Ω,μ), a class $\mathcal{C}\subseteq L^2_+(\Omega,\mu)$ and non-negative functional $\mathcal{E}$ on $\mathcal{C}$ we will say that a pair $(\mu,\mathcal{E})$ satisfies $I(a)$ (respectively the Poincaré or the logarithmic Sobolev) inequality if every $f\in\mathcal{C}$ satisfies $I_\mu(a)$ (resp. the Poincaré or the logarithmic Sobolev) inequality with constant C (for these particular μ and $\mathcal{E}$). For the sake of brevity we will assume that μ identifies probability space and $\mathcal{E}$ carries information about $\mathcal{C}$.*

An obvious modification of Corollary 1 for pairs $(\mu,\mathcal{E})$ follows. In some cases we can establish the precise relation between best possible constants in $I(1)$ and logarithmic Sobolev inequalities.

Let $m:(-a,a)\longrightarrow R$ be an even, strictly postive continuous density of some probability measure μ on $(-a,a)$, where $0<a\leq\infty$ and assume that $\int_{-a}^a x^2m(x)dx<\infty$. For $f\in C_0^\infty(-a,a)$ put

$$(Lf)(x) = xf'(x) - u(x)f''(x),$$

where $u(x)=\frac{\int_x^a tm(t)\ dt}{m(x)}\geq 0$. General theory (see [KLO] for detailed references and some related results) yields that L can be extended to a positive definite self-adjoint operator (denoted by the same symbol), defined on a dense subspace $Dom(L)$ of $L^2((-a,a),\mu)$, whose spectrum $\sigma(L)$ is contained in $\{0\}\cup[1,\infty)$. Moreover $P_t=e^{-tL}$ $(t\geq 0)$ is a Markov semigroup with invariant measure μ. Put $\mathcal{E}(f)=\|L^{1/2}f\|_2^2$ (we accept $\mathcal{E}(f)=+\infty$ for f which do not belong to $Dom(L^{1/2})$) and take $\mathcal{C}=L^2_+((-a,a),\mu)$.

Lemma 2 *Under the above assumptions the following equivalence holds true:*
$(\mu,\mathcal{E})$ satisifes the inequality $I(1)$ with constant C
if and only if
$(\mu,\mathcal{E})$ satisfies the logarithmic Sobolev inequality with constant $2C$.

Proof. If $(\mu, \mathcal{E})$ satisifes the inequality $I(1)$ with constant C then by Corollary 1 it satisfies the logarithmic Sobolev inequality with constant $2C$. Now let us assume that $(\mu, \mathcal{E})$ satisfies the logarithmic Sobolev inequality with constant $2C$. Then for any $f \in L^2((-a,a),\mu)$ we have

$$Ent_\mu(f^2) = Ent_\mu(|f|^2) \leq 2C\mathcal{E}(|f|) \leq 2C\mathcal{E}(f)$$

(the last inequality is a well known property of Dirichlet forms of Markov semigroups - see for example Theorem 1. 3. 2 of [D]). Therefore classical hypercontractivity result [G] yields

$$\|P_{t(p)}f\|_2 \leq \|f\|_p,$$

where $t(p) = \frac{C}{2}\ln(\frac{1}{p-1})$ for $p \in [1,2)$; if $p=1$ then we put $t(p) = \infty$ and $P_\infty(f) = E_\mu f$. Hence

$$Efe^{-2t(p)L}f \leq (Ef^p)^{2/p}$$

or equivalently

$$Ef^2 - (Ef^p)^{2/p} \leq Ef(Id - e^{-2t(p)L})f$$

for any $f \in \mathcal{C}$. Now it suffices to prove that for any $\lambda \in \sigma(L)$ we have

$$1 - e^{-2t(p)\lambda} \leq (2-p)C\lambda,$$

i.e.

$$1 - (2-p)C\lambda \leq (p-1)^{C\lambda}.$$

For $\lambda = 0$ and $p \in (1,2)$ the inequality is trivial. It is known that if $(\mu, \mathcal{E})$ satisfies the logarithmic Sobolev inequality with constant $2C$ then (under the assumptions of Lemma 2) $C \geq 1$ - to see this consider the logarithmic Sobolev inequality for functions of the form $f(x) = |1 + \varepsilon x|$ with ε tending to zero (this is a special case of more general observation which says that, for functionals $\mathcal{E}$ satisfying certain natural conditions, if $(\mu, \mathcal{E})$ satisfies the logarithmic Sobolev inequality with constant $2C$ then it also satisfies the Poincaré inequality with constant C). We can restrict our considerations to the case $\lambda \geq 1$ since $\sigma(L) \setminus \{0\} \subseteq [1, \infty)$. Therefore $(p-1)^{C\lambda}$ is a convex function of p and to prove that

$$h(p) = (p-1)^{C\lambda} + (2-p)C\lambda - 1 \geq 0$$

for $p \in [1,2)$ it suffices to check that $h(2) = h'(2) = 0$ which is obvious. The case $p=1$ (omitted when $\lambda = 0$ because $(p-1)^{C\lambda}$ was not well defined) follows easily since the function $p \longmapsto (Ef^p)^{2/p}$ is continuous for $p \in [1,2]$. □

Corollary 2 *If μ is a $\mathcal{N}(0,1)$ Gaussian measure on real line, $\mathcal{E}(f) = E_\mu(f')^2$ and $\mathcal{C}$ is a class of non-negative smooth functions then $(\mu, \mathcal{E})$ satisfies $I(1)$ with constant 1.*

Proof. If μ is a $\mathcal{N}(0,1)$ Gaussian measure and operator L is defined as before then

$$E_\mu f L f = E_\mu (f')^2.$$

The assertion follows from Lemma 2 and well known fact ([G]) that Gaussian measures satisfy the logarithmic Sobolev inequality with constant 2. □

Remark 1 *Method used in Lemma 2 seems applicable also in more general situation (see [O] for possible directions of generalization). Let us mention just one interesting application. If* $\Omega = \{-1, 1\}$, $\mu(\{-1\}) = \mu(\{1\}) = 1/2$ *and* $\mathcal{E}(f) = (\frac{f(1)-f(-1)}{2})^2$ *then* $(\mu, \mathcal{E})$ *satisfies* $I(1)$ *with constant 1.*

Remark 2 *Let* μ *be a non-symmetric two-point distribution on* $\{-1, 1\}$, $\mu(\{1\}) = 1 - \mu(\{-1\}) = \alpha$ *with* $\alpha \in (0, 1/2) \cup (1/2, 1)$. *Then for any* $p \in [1, 2)$ *and any* $f : \{-1, 1\} \to R_+$ *the inequality*

$$E_\mu f^2 - (E_\mu f^p)^{2/p} \leq C_\alpha(p)(f(1) - f(-1))^2$$

holds with

$$C_\alpha(p) = \frac{\alpha^{1-2/p} - (1-\alpha)^{1-2/p}}{\alpha^{-2/p} - (1-\alpha)^{-2/p}}$$

and the constant cannot be improved.

Proof (sketch). To check the optimality of $C_\alpha(p)$ put $f(-1) = \alpha^{2/p}$ and $f(1) = (1-\alpha)^{2/p}$. To prove the inequality observe that for $p \in (1, 2)$, $\varphi(y) = ((1 + \sqrt{y})^p + (1 - \sqrt{y})^p)^{2/p}$ is a strictly convex function of $y \in (0, 1)$, since

$$\varphi^{'}(y) = [(1+\sqrt{y})^p + (1-\sqrt{y})^p]^{\frac{2}{p}-1} \frac{(1+\sqrt{y})^{p-1} - (1-\sqrt{y})^{p-1}}{\sqrt{y}}$$

$$= \big(2\sum_{k=0}^{\infty} \binom{p}{2k} y^k\big)^{\frac{2}{p}-1} 2 \sum_{k=0}^{\infty} \binom{p-1}{2k+1} y^k$$

is clearly increasing (note that $\binom{p}{2k}$ and $\binom{p-1}{2k+1}$ are positive for $k = 0, 1, \ldots$). Hence for each $y_0 \in (0, 1)$ and $p \in (1, 2)$ there exist unique real numbers A and B such that

$$\varphi(y^2) = ((1+y)^p + (1-y)^p)^{2/p} \geq A + By^2 \text{ for all } y \in (-1, 1)$$

with equality holding for $|y| = y_0$ only. By the homogenity we may assume that $f(-1) = (1-\alpha)^{-1/p}(1+y)$ and $f(1) = \alpha^{-1/p}(1-y)$. Putting $y_0 = \frac{(1-\alpha)^{1/p} - \alpha^{1/p}}{(1-\alpha)^{1/p} + \alpha^{1/p}}$, using the above inequality after some elementary, but a little involved computations one proves the assertion. □

Definition 4 *Let us denote by Φ the class of all continuous functions $\varphi : [0,\infty) \longrightarrow R$ having strictly positive second derivatve and such that $1/\varphi''$ is a concave function. Let us additionally include in Φ all functions φ of the form $\varphi(x) = ax + b$, where a and b are some real constants.*

Although it is not obvious, functions belonging to Φ form a convex cone. There are some interesting questions connected with the class Φ and its generalizations but we postpone them till the end of the note.

Lemma 3 *For any $\varphi \in \Phi$ and $t \in [0,1]$ the function $F_t : [0,\infty) \times [0,\infty) \longrightarrow R$ defined by*

$$F_t(x,y) = t\varphi(x) + (1-t)\varphi(y) - \varphi(tx + (1-t)y)$$

is non-negative and convex.

Proof. Non-negativity of F_t is an easy consequence of convexity of φ. Obviously F_t is continuous on $[0,\infty) \times [0,\infty)$ and twice differentiable on $(0,\infty) \times (0,\infty)$. Therefore it suffices to prove that $Hess\, F_t$ (second derivative matrix) is positive definite on $(0,\infty) \times (0,\infty)$. We skip the trivial case of φ being an affine function. Note that from the positivity of φ'' and the concavity of $1/\varphi''$ it follows that

$$\frac{1}{\varphi''(tx + (1-t)y)} \geq \frac{t}{\varphi''(x)} + \frac{1-t}{\varphi''(y)} \geq \frac{t}{\varphi''(x)}.$$

Therefore

$$\frac{\partial^2 F_t}{\partial x^2}(x,y) = t\varphi''(x) - t^2\varphi''(tx + (1-t)y) \geq 0.$$

In a similar way we prove that $\frac{\partial^2 F_t}{\partial y^2}(x,y) \geq 0$. Now it is enough to prove that $\det(Hess\, F_t) \geq 0$ i.e. that

$$\frac{\partial^2 F_t}{\partial x^2}(x,y) \cdot \frac{\partial^2 F_t}{\partial y^2}(x,y) \geq (\frac{\partial^2 F_t}{\partial x \partial y}(x,y))^2$$

which is equivalent to

$$(t\varphi''(x) - t^2\varphi''(tx + (1-t)y))((1-t)\varphi''(y) - (1-t)^2\varphi''(tx + (1-t)y))$$
$$\geq (-t(1-t)\varphi''(tx + (1-t)y))^2$$

or

$$\varphi''(x)\varphi''(y) \geq t\varphi''(y)\varphi''(tx + (1-t)y) + (1-t)\varphi''(x)\varphi''(tx + (1-t)y).$$

After dividing by $\varphi''(x)\varphi''(y)\varphi''(tx + (1-t)y)$ the last inequality follows from concavity of $1/\varphi''$ and the proof is complete. $\square$

Lemma 4 *For a non-negative real random variable Z defined on probability space (Ω, μ) and having finite first moment, and for $\varphi \in \Phi$ let*

$$\Psi_\varphi(Z) = E_\mu \varphi(Z) - \varphi(E_\mu Z).$$

Then for any non-negative real random variables X and Y defined on (Ω, μ) and having finite first moment, and for any $t \in [0, 1]$ the following inequality holds:

$$\Psi_\varphi(tX + (1-t)Y) \geq t\Psi_\varphi(X) + (1-t)\Psi_\varphi(Y);$$

in other words Ψ_φ is a convex functional on the convex cone of integrable non-negative real random variables defined on (Ω, μ).

Proof. Let us note that (under notation of Lemma 3)

$$\Psi_\varphi(tX + (1-t)Y) - t\Psi_\varphi(X) - (1-t)\Psi_\varphi(Y) =$$
$$(E_\mu \varphi(tX + (1-t)Y) - tE_\mu \varphi(X) - (1-t)E_\mu \varphi(Y)) -$$
$$(\varphi(tE_\mu X + (1-t)E_\mu Y) - t\varphi(E_\mu X) - (1-t)\varphi(E_\mu Y))$$
$$= E_\mu F_t(X, Y) - F_t(E_\mu X, E_\mu Y) = E_\mu F_t(X, Y) - F_t(E_\mu(X, Y)).$$

We are to prove that it is a non-negative expression and this follows easily from Jensen inequality. For the sake of clarity we present a detailed argument.

Let $x_0 = E_\mu X$ and $y_0 = E_\mu Y$. Lemma 3 yields that F_t is convex, so that there exist constants $a, b, c \in R$ such that

$$F_t(x, y) \geq ax + by + c$$

for any $x, y \in [0, \infty)$ and

$$F_t(x_0, y_0) = ax_0 + by_0 + c.$$

Therefore

$$E_\mu F_t(X, Y) \geq E_\mu(aX + bY + c) = ax_0 + by_0 + c = F_t(x_0, y_0) = F_t(E_\mu X, E_\mu Y)$$

and the proof is finished. □

Lemma 5 *Let (Ω_1, μ_1) and (Ω_2, μ_2) be probability spaces and let $(\Omega, \mu) = (\Omega_1 \times \Omega_2, \mu_1 \otimes \mu_2)$ be their product probability space. For any non-negative random variable Z defined on (Ω, μ) and having finite first moment and for any $\varphi \in \Phi$ the following inequality holds true:*

$$E_\mu \varphi(Z) - \varphi(E_\mu Z) \leq E_\mu([E_{\mu_1} \varphi(Z) - \varphi(E_{\mu_1} Z)] + [E_{\mu_2} \varphi(Z) - \varphi(E_{\mu_2} Z)]).$$

Proof. For $\omega_2 \in \Omega_2$ let $Z_{(\omega_2)}$ be a non-negative random variable defined on (Ω_1, μ_1) by the formula

$$Z_{[\omega_2]}(\omega_1) = Z(\omega_1, \omega_2).$$

By Lemma 4 used for the probability space (Ω_1, μ_1) and Jensen inequality used for the family of random variables $(Z_{[\omega_2]})_{\omega_2 \in \Omega_2}$ (this time we skip the detailed argument which the reader can easily repeat after the proof of Lemma 4) we get

$$E_{\mu_2}(E_{\mu_1}\varphi(Z) - \varphi(E_{\mu_1}Z)) \geq E_{\mu_1}\varphi(E_{\mu_2}Z) - \varphi(E_{\mu_1}(E_{\mu_2}Z))$$

which is equivalent to the assertion of Lemma 5. □

By an easy induction argument we obtain

Corollary 3 *Let $(\Omega_1, \mu_1), (\Omega_2, \mu_2), \ldots, (\Omega_n, \mu_n)$ be probability spaces and let $(\Omega, \mu) = (\Omega_1 \times \Omega_2 \times \ldots \times \Omega_n, \mu_1 \otimes \mu_2 \otimes \ldots \otimes \mu_n)$ be their product probability space. Let Z be any integrable non-negative real random variable defined on (Ω, μ). Then for any $\varphi \in \Phi$ the following inequality holds:*

$$E_\mu \varphi(Z) - \varphi(E_\mu Z) \leq \sum_{k=1}^{n} E_\mu(E_{\mu_k}\varphi(Z) - \varphi(E_{\mu_k}Z)).$$

Let us observe that the function φ defined by $\varphi(x) = x^{2/p}$ belongs to the class Φ if $p \in [1, 2]$. Therefore by applying Corollary 3 to the random variable $Z = f^p$, where $f \in L^2_+(\Omega, \mu)$, we obtain

Corollary 4 *Under the notation of Corollary 3 for any $f \in L^2_+(\Omega, \mu)$ we have*

$$E_\mu f^2 - (Ef^p)^{2/p} \leq \sum_{k=1}^{n} E_\mu(E_{\mu_k} f^2 - (E_{\mu_k} f^p)^{2/p}).$$

This sub-additivity property of functional $Var(p)_\mu$ immediately yields the following

Corollary 5 *Assume that pairs $(\mu_1, \mathcal{E}_1)$,$(\mu_2, \mathcal{E}_2)$,$\ldots$ $(\mu_n, \mathcal{E}_n)$ satisfy the inequality $I(a)$ with some constant C. Let $\mu = \mu_1 \otimes \mu_2 \otimes \ldots \otimes \mu_n$ and $\mathcal{E}(f) = E_\mu(\mathcal{E}_1(f_1) + \mathcal{E}_2(f_2) + \ldots + \mathcal{E}_n(f_n))$, where*

$$f_i(x) = f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)$$

for given $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$. Class $\mathcal{C}$ can be chosen in any way which assures that $f \in \mathcal{C}$ implies $f_i \in \mathcal{C}_i$, for example $\mathcal{C} = \mathcal{C}_1 \otimes \mathcal{C}_2 \otimes \ldots \otimes \mathcal{C}_n$. Then the pair $(\mu, \mathcal{E})$ also satisfies the inequality $I(a)$ with constant C.

The case we will concentrate on is $\mathcal{E}(f) = E_\mu \|\nabla f\|^2$.

Proposition 1 *Let* μ_1,μ_2,...μ_n *be probability measures on* R. *Let* $C > 0$ *and* $a \in [0, 1]$. *Assume that for any smooth function* $f : R \longrightarrow [0, \infty)$ *the inequality*

$$E_{\mu_i} f^2 - (E_{\mu_i} f^p)^{2/p} \leq C(2-p)^a E_{\mu_i} (f')^2$$

holds true for $p \in [1, 2)$ *and* $i = 1, 2, \ldots n$. *Then for* $\mu = \mu_1 \otimes \mu_2 \otimes \ldots \otimes \mu_n$ *the inequality*

$$E_\mu f^2 - (E_\mu f^p)^{2/p} \leq C(2-p)^a E_\mu \|\nabla f\|^2,$$

where $\|\cdot\|$ *denotes standard Euclidean norm, is satisfied for* $p \in [1, 2)$ *and any smooth function* $f : R^n \longrightarrow [0, \infty)$.

Proof. Use Corollary 5 and note that

$$E_\mu \|\nabla f\|^2 = E_\mu[(\frac{\partial f}{\partial x_1})^2 + \ldots + (\frac{\partial f}{\partial x_1})^2] = E_\mu[(f_1')^2 + \ldots + (f_n')^2]$$

$$= E_\mu[E_{\mu_1}(f_1')^2 + \ldots + E_{\mu_n}(f_n')^2]. \ \square$$

Now let us demonstrate that the inequality $I(a)$ for the $\mathcal{E}(f) = E_\mu \|\nabla f\|^2$ functional implies concentration of Lipschitz functions.

Theorem 1 *Let* μ *be a probability measure on* R^n. *Assume that there exist constants* $C > 0$ *and* $a \in [0, 1]$ *such that the inequality*

$$E_\mu f^2 - (E_\mu f^p)^{2/p} \leq C(2-p)^a E_\mu \|\nabla f\|^2$$

is satisfied for any smooth function $f : R^n \longrightarrow [0, \infty)$ *and* $p \in [1, 2)$. *Let* $h : R^n \longrightarrow R$ *be a Lipschitz function with Lipschitz constant 1, i.e.* $|h(x) - h(y)| \leq \|x - y\|$ *for any* $x, y \in R^n$, *where* $\|\cdot\|$ *denotes a standard Euclidean norm. Then* $E_\mu |h| < \infty$ *and*

- *for any* $t \in [0, 1]$

$$\mu(h - E_\mu h \geq t\sqrt{C}) \leq e^{-Kt^2}$$

- *for any* $t \geq 1$

$$\mu(h - E_\mu h \geq t\sqrt{C}) \leq e^{-Kt^{\frac{2}{2-a}}}$$

where K *is some universal constant.*

Proof. Our proof will work for $K = 1/3$ but we do not know optimal constants (it is also interesting what the optimal K is for given value of parameter a). Note that it is essential part of the assumptions that we study the limit behaviour when $p \to 2$. For any fixed $p \in (1, 2)$ the inequality

$$E_\mu f^2 - (E_\mu f^p)^{2/p} \leq C(2-p)^a E_\mu \|\nabla f\|^2$$

is weaker than the Poincaré inequality with constant $C(2-p)^a$ and therefore it cannot imply anything stronger than the exponential concentration.

We will follow the aproach of [AS]. Assume first that h is bounded and smooth. Then $\|\nabla h\| \le 1$. Define $H(\lambda) = E_\mu e^{\lambda h}$ for $\lambda \ge 0$. Assumptions of Theorem 1 for $f = e^{\lambda h/2}$ give

$$H(\lambda) - H(\frac{p}{2}\lambda)^{2/p} \le \frac{C\lambda^2}{4}(2-p)^a E_\mu \|\nabla h\|^2 e^{\lambda h} \le \frac{C\lambda^2}{4}(2-p)^a H(\lambda).$$

Hence

$$H(\lambda) \le \frac{H(\frac{p}{2}\lambda)^{2/p}}{1 - \frac{C}{4}(2-p)^a\lambda^2}$$

for any $p \in [1,2)$ and $0 \le \lambda \le \frac{2}{\sqrt{C}}(2-p)^{-a/2}$. Applying the same inequality for $\frac{p}{2}\lambda$ and iterating, after m steps we get

$$H(\lambda) \le \frac{H((\frac{p}{2})^m\lambda)^{(2/p)^m}}{\prod_{k=0}^{m-1}(1 - \frac{C\lambda^2}{4}(2-p)^a \cdot (\frac{p}{2})^{2k})^{(2/p)^k}}.$$

Note that

$$1 - \frac{C\lambda^2}{4}(2-p)^a \cdot (\frac{p}{2})^{2k} \ge (1 - \frac{C\lambda^2}{4}(2-p)^a)^{(p/2)^{2k}}$$

since $(\frac{p}{2})^{2k} < 1$. Hence

$$H(\lambda) \le H((\frac{p}{2})^m\lambda)^{(2/p)^m}(1 - \frac{C\lambda^2}{4}(2-p)^a)^{-\sum_{k=0}^{m-1}(p/2)^k}.$$

As $\lim_{m\to\infty}(\frac{p}{2})^m = 0$ we get

$$\lim_{m\to\infty} H((\frac{p}{2})^m\lambda)^{(2/p)^m} = e^{\lambda E_\mu h}.$$

Therefore

$$E_\mu e^{\lambda(h - E_\mu h)} \le (1 - \frac{C\lambda^2}{4}(2-p)^a)^{-\frac{2}{2-p}}$$

and

$$\mu(h - E_\mu h \ge t\sqrt{C}) \le e^{-\lambda t\sqrt{C}} \cdot (1 - \frac{C\lambda^2}{4}(2-p)^a)^{-\frac{2}{2-p}}.$$

- Putting $p = 1$ and $\lambda = \frac{t}{\sqrt{C}}$ we get for any $t \in [0,2)$

  $$\mu(h - E_\mu h \ge t\sqrt{C}) \le e^{-t^2} \cdot (1 - \frac{t^2}{4})^{-2}.$$

  In particular for $t \in [0,1]$ we have $1 - \frac{t^2}{4} > e^{-t^2/3}$ and

  $$\mu(h - E_\mu h \ge t\sqrt{C}) \le e^{-t^2/3}.$$

- If $t \geq 1$, let us put $p = 2 - t^{-\frac{2}{2-a}}$ and $\lambda = t^{\frac{a}{2-a}}/\sqrt{C}$. Then we arrive at

$$\mu(h - E_\mu h \geq t\sqrt{C}) \leq e^{-t^{\frac{2}{2-a}}} \cdot (1 - \frac{1}{4})^{-2t^{\frac{2}{2-a}}} = (\frac{16}{9e})^{t^{\frac{2}{2-a}}}$$

which completes the proof (if h is bounded and smooth) since $\frac{16}{9e} \leq e^{-1/3}$.

Therefore by a standard approximation argument we prove the assertion for any bounded h which satisfies assumptions of Theorem 1. Finally for general h define its bounded truncations $(h_N)_{N=1}^{\infty}$ putting $h_N(x) = h(x)$ if $|x| \leq N$ and $h_N(x) = N \cdot sgn(x)$ if $|x| \geq N$. One can easily check that if h satisfies the assumptions of Theorem 1 then $|h_N|$ is also a Lipschitz function with a Lipschitz constant 1 and therefore using Theorem 1 for a bounded function $|h_N|$ we arrive at

$$\mu(|h_N| - E_\mu|h_N| \geq 4\sqrt{C}) \leq (\frac{16}{9e})^{4^{\frac{2}{2-a}}} \leq (\frac{16}{9e})^4 \leq \frac{1}{5}.$$

Similarly

$$\mu(|h_N| - E_\mu|h_N| \leq -4\sqrt{C}) = \mu(-|h_N| - E_\mu(-|h_N|) \geq 4\sqrt{C}) \leq \frac{1}{5}.$$

Hence

$$\mu(|\,|h_N| - E_\mu|h_N|\,| \geq 4\sqrt{C}) \leq \frac{2}{5}$$

and

$$\mu(|\,|h| - E_\mu|h_N|\,| \geq 4\sqrt{C}) \leq \frac{2}{5} + \mu(|h| > N).$$

Therefore

$$\mu(|E_\mu|h_N| - E_\mu|h_M|\,| \geq 8\sqrt{C}) \leq$$
$$\mu(|\,|h| - E_\mu|h_N|\,| \geq 4\sqrt{C}) + \mu(|\,|h| - E_\mu|h_M|\,| \geq 4\sqrt{C}) \leq$$
$$\frac{4}{5} + \mu(|h| > N) + \mu(|h| > M) \longrightarrow \frac{4}{5} < 1$$

as $\min(N, M) \longrightarrow \infty$, which means that the sequence $(E_\mu|h_N|)_{N=1}^{\infty}$ is bounded. As $|h_N|$ grows monotonically to $|h|$, by Lebesgue Lemma we get $E_\mu|h| < \infty$ and $E_\mu h_N \longrightarrow E_\mu h$ as $N \longrightarrow \infty$. Now an easy approximation argument completes the proof. $\square$

In order to prove that the order of concentration implied by Theorem 1 cannot be improved we will need the following

Theorem 2 *Let $a \in [0,1]$ and $r \in [1,2]$ satisfy $r = 2/(2-a)$. Put $c_r = \frac{1}{2\Gamma(1+1/r)} = \frac{r}{2\Gamma(1/r)}$. Then $\mu_r(dx) = c_r^n \exp(-(|x_1|^r + |x_2|^r + \ldots + |x_n|^r))dx_1 dx_2 \ldots dx_n$ is a probability measure on R^n and there exists a universal constant $C > 0$ (not depending on a or n) such that*

$$E_{\mu_r} f^2 - (E_{\mu_r} f^p)^{2/p} \leq C(2-p)^a E_{\mu_r} \|\nabla f\|^2$$

for any smooth non-negative function f on R^n and any $p \in [1, 2)$.

Proof. Proposition 1 shows that it is enough to prove Theorem 2 in the case $n = 1$. Therefore the assertion easily follows from the two following propositions. $\square$

Proposition 2 *Let $a \in [0,1]$ and $r \in [1,2]$ satisfy $r = 2/(2-a)$. Put $c_r = \frac{1}{2\Gamma(1+1/r)}$, so that $\mu_r(dx) = c_r \exp(-|x_1|^r)dx$ is a probability measure on R. Let $\lambda(dx) = \frac{1}{2}e^{-|x|}$ be a symmetric exponential probability measure on R. Under these assumptions the following implications hold true:*

- *If $C > 0$ is a constant such that for any smooth function $f : R \longrightarrow [0,\infty)$ and any $p \in [1,2)$ there is*

$$E_{\mu_r} f^2 - (E_{\mu_r} f^p)^{2/p} \leq C(2-p)^a E_{\mu_r}(f')^2$$

  then for any smooth function $g : R \longrightarrow [0,\infty)$ and any $p \in [1,2)$ there is

$$\int_R g(x)^2\lambda(dx) - (\int_R g(x)^p\lambda(dx))^{2/p} \leq 600C(2-p)^a \int_R \max(1,|x|^a)g'(x)^2\lambda(dx).$$

- *Conversely, if $C > 0$ is such a constant that for any smooth function $g : R \longrightarrow [0,\infty)$ and any $p \in [1,2)$ there is*

$$\int_R g(x)^2\lambda(dx) - (\int_R g(x)^p\lambda(dx))^{2/p} \leq C(2-p)^a \int_R \max(1,|x|^a)g'(x)^2\lambda(dx)$$

  then for any smooth function $f : R \longrightarrow [0,\infty)$ and any $p \in [1,2)$ there is

$$E_{\mu_r} f^2 - (E_{\mu_r} f^p)^{2/p} \leq 50C(2-p)^a E_{\mu_r}(f')^2.$$

Proposition 3 *There exists a universal constant C such that for any $a \in [0,1]$, any $p \in [1,2)$ and any smooth function $g : R \longrightarrow [0,\infty)$ the following inequality holds*

$$\int_R g(x)^2\lambda(dx) - (\int_R g(x)^p\lambda(dx))^{2/p} \leq C(2-p)^a \int_R \max(1,|x|^a)g'(x)^2\lambda(dx).$$

We will start with proof of Proposition 2. The proof of Proposition 3 will be postponed to the end of the paper.

Proof of Proposition 2. Let us define the function $z_r : R \longrightarrow R$ by

$$\frac{1}{2}\int_{z_r(x)}^{\infty} e^{-|t|}dt = c_r \int_x^{\infty} e^{-|t|^r}dt,$$

where $c_r = \frac{r}{2\Gamma(1/r)} = \frac{1}{2\Gamma(1+1/r)}$. It is easy to see that z_r is a homeomorphism of R onto itself and

$$z_r'(x) = 2c_r e^{|z_r(x)|-|x|^r}.$$

Therefore z_r is a C^1-diffeomorphism of R onto itself. Binding f and g by relation $f(x) = g(z_r(x))$ and using standard change of variables formula we reduce the proof of Proposition 2 to the following lemma. $\square$

Lemma 6 *Under notation introduced above*

$$\frac{1}{50}\max(1,|x|^a) \le (z_r'(z_r^{-1}(x)))^2 \le 600\max(1,|x|^a)$$

for any $x \in R$.

Proof. First let us note that $1/3 \le c_r \le e/2$. Indeed,

$$\Gamma(1/r) = \int_0^\infty x^{\frac{1}{r}-1}e^{-x}dx \le \int_0^1 x^{\frac{1}{r}-1}dx + \int_1^\infty e^{-x}dx = r + 1/e.$$

Hence $c_r \ge \frac{r}{2r+2/e} \ge 1/3$. On the other hand

$$\Gamma(1/r) = \int_0^\infty x^{\frac{1}{r}-1}e^{-x}dx \ge \frac{1}{e}\int_0^1 x^{\frac{1}{r}-1}dx = r/e.$$

Therefore $c_r \le e/2$. Let us also notice that by obvious symmetry we can consider only the case $x > 0$. Now let us estimate from below $z_r^{-1}(1)$. We have

$$\frac{e}{2}z_r^{-1}(1) \ge c_r z_r^{-1}(1) \ge c_r \int_0^{z_r^{-1}(1)} e^{-t^r}dt = \frac{1}{2}\int_0^1 e^{-t}dt = \frac{1}{2}(1-1/e)$$

and therefore $z_r^{-1}(1) \ge \frac{e-1}{e^2} \ge 1/5$. Note that by definition of $z_r(x)$ for $x > 0$ we have

$$\frac{1}{2}e^{-z_r(x)} = c_r\int_x^\infty e^{-t^r}dt \le c_r\int_x^\infty \frac{rt^{r-1}}{rx^{r-1}}e^{-t^r}dt = \frac{c_r e^{-x^r}}{rx^{r-1}}$$

and therefore

$$z_r'(x) = 2c_r e^{z_r(x)-x^r} \ge rx^{r-1}.$$

Hence also $z_r(x) \ge x^r$ and $z_r^{-1}(x) \le x^{1/r}$ for all positive x. If $x \ge 1/5$ then

$$\int_x^\infty e^{-t^r}dt \ge \int_x^{6x} e^{-t^r}dt \ge \frac{1}{r(6x)^{r-1}}\int_x^{6x} rt^{r-1}e^{-t^r}dt =$$

$$6^{1-r}\frac{e^{-x^r}-e^{-6^r x^r}}{rx^{r-1}} \ge \frac{1}{12}\frac{e^{-x^r}}{rx^{r-1}},$$

since $6^r x^r \ge x^r + 1$ for $x \ge 1/5$ and $r \in [1,2]$. Therefore for $x \ge z_r^{-1}(1) \ge 1/5$ we have

$$z_r'(x) \le 12rx^{r-1} \le 24x^{r-1}$$

and

$$z_r(x) \le z_r(z_r^{-1}(1)) + 12\int_{z_r^{-1}(1)}^x rt^{r-1}dt = 1 + 12(x^r - [z_r^{-1}(1)]^r) \le 1 + 12x^r \le 37x^r.$$

Hence $z_r^{-1}(x) \geq (x/37)^{1/r}$ for $x \geq z_r^{-1}(1)$. If $x \geq 1$ then $z_r^{-1}(x) \geq 1/5$ and therefore

$$z_r'(z_r^{-1}(x)) \leq 24[z_r^{-1}(x)]^{r-1} \leq 24x^{\frac{r-1}{r}} = 24x^{a/2}.$$

Also if $x \geq 1$ then $z_r^{-1}(x) \geq z_r^{-1}(1)$ and

$$z_r'(z_r^{-1}(x)) \geq r[z_r^{-1}(x)]^{r-1} \geq (x/37)^{\frac{r-1}{r}} \geq 37^{\frac{1}{r}-1}x^{a/2} \geq \frac{1}{7}x^{a/2}.$$

This proves Lemma 6 for $|x| \geq 1$. For any $x \geq 0$ we have

$$z_r'(z_r^{-1}(x)) = 2c_r e^{x - z_r^{-1}(x)^r} \geq 2c_r \geq 2/3.$$

We used the previously proved fact that $z_r^{-1}(x) \leq x^{1/r}$. Now it remains only to establish upper estimate on $z_r'(z_r^{-1}(x))$ for $x \in [0,1]$. Note that if $x \leq z_r^{-1}(1)$ then

$$c_r \int_x^\infty e^{-t^r} dt = \frac{1}{2}\int_{z_r(x)}^\infty e^{-t} dt \geq \frac{1}{2}\int_1^\infty e^{-t} dt = \frac{1}{2e}$$

and therefore

$$z_r'(x) = \frac{2c_r e^{-x^r}}{2c_r \int_x^\infty e^{-t^r} dt} \leq \frac{c_r}{c_r \int_x^\infty e^{-t^r} dt} \leq 2ec_r \leq e^2 \leq 8.$$

Hence $z_r'(z_r^{-1}(x)) \leq 8$ for any $|x| \leq 1$ and the proof is finished. □

Lemma 7 *For $s \in (1,2]$ and $x, y \geq 0$ put*

$$\rho_s(x,y) = (\frac{x^s + y^s}{2} - (\frac{x+y}{2})^s)^{1/2}.$$

Then ρ_s is a metric on $[0,\infty)$.

Proof. Since $k_t(a,b) = e^{-(a+b)t}$ is obviously positive definite integral kernel and $K(a,b) = s(s-1)(a+b)^{s-2} = \frac{s(s-1)}{\Gamma(2-s)}\int_0^\infty t^{1-s}k_t(a,b)\,dt$ we get, by Schwartz inequality (applied to a scalar product defined by the kernel $K(a,b)$), that for any $y \geq x \geq 0$ and $z \geq t \geq 0$ the following inequality is true:

$$\int_{x/2}^{y/2}\int_{t/2}^{z/2} K(a,b)\,da\,db$$
$$\leq (\int_{x/2}^{y/2}\int_{x/2}^{y/2} K(a,b)\,da\,db)^{1/2}(\int_{t/2}^{z/2}\int_{t/2}^{z/2} K(a,b)\,da\,db)^{1/2}.$$

Now, as

$$K(a,b) = \frac{\partial^2}{\partial a\,\partial b}(a+b)^s,$$

we get by integration by parts

$$(\frac{y+z}{2})^s + (\frac{x+t}{2})^s - (\frac{x+z}{2})^s - (\frac{y+t}{2})^s \leq$$

$$(x^s + y^s - 2(\frac{x+y}{2})^s)^{1/2}(z^s + t^s - 2(\frac{z+t}{2})^s)^{1/2}$$

Putting $t = y$ we arrive at

$$(\frac{x+y}{2})^s + (\frac{y+z}{2})^s - (\frac{x+z}{2})^s - y^s \leq 2\rho_s(x,y)\rho_s(y,z)$$

which is equivalent to

$$\rho_s(x,z)^2 - \rho_s(x,y)^2 - \rho_s(y,z)^2 \leq 2\rho_s(x,y)\rho_s(y,z).$$

Hence $\rho_s(x,z) \leq \rho_s(x,y) + \rho_s(y,z)$. For $x \leq y \leq z$ we have also easily $\rho_s(x,z) \geq \rho_s(x,y)$ and $\rho_s(x,z) \geq \rho_s(y,z)$, so that $\rho_s(x,y) \leq \rho_s(x,z) + \rho_s(z,y)$ and $\rho_s(y,z) \leq \rho_s(y,x) + \rho_s(x,z)$. This completes the proof of triangle inequality for $s < 2$. Other metric properties of ρ_s as well as the case $s = 2$ are trivial. □

Remark 3 *In a similar way one can prove that* $\rho_s(x,y) = |\frac{x^s+y^s}{2} - (\frac{x+y}{2})^s|^{1/2}$ *is a metric on* $(0,\infty)$ *for* $s \in (-\infty, 0) \cup (0,1)$. *It was pointed out to the authors by B. Maurey that Lemma 7 follows also from isometrical immersion of* $([0,\infty), \rho_s)$ *into* $L^2([0,\infty), \kappa_s^{-1} t^{-s-1} dt)$, *where* $x \in [0,\infty)$ *is sent to the function* $e^{-xt} - 1$ *and* $\kappa_s = 2^{s+1} \int_0^\infty (e^{-u} - 1 + u) u^{-s-1} du$.

Lemma 8 *Let* $s \in [1,2]$, $t \in [0,1]$ *and* c, d, x *be nonnegative numbers. The following inequality holds*

$$(1-t)c^s + td^s - ((1-t)c + td)^s \leq$$

$$K[(1-t)c^s + td^s + x^s - ((1-t)c + tx)^s - (td + (1-t)x)^s]. \qquad (1)$$

under anyone of the following additional assumptions

- x *lies outside the open interval* (c,d) *and* $K = 1$
- $t = \frac{1}{2}$ *and* $K = 2$
- $t \leq \frac{1}{2}$, $c \geq d$ *and* $K = 12$

Proof. Let us remind that

$$F_t(x,y) = tx^s + (1-t)y^s - (tx + (1-t)y)^s$$

is a convex function on $[0,\infty) \times [0,\infty)$. Note that the inequality of Lemma 8 is equivalent to

$$F_t(d,c) \leq K[F_t(d,x) + F_t(x,c)].$$

- As
$$\frac{\partial}{\partial x}[F_t(d,x)+F_t(x,c)]$$
$$= s[(1-t)(x^{s-1}-(td+(1-t)x)^{s-1})+t(x^{s-1}-(tx+(1-t)c)^{s-1})],$$
we see that the right-hand side of the inequality as a function of x is increasing on $(\max(c,d),\infty)$ and decreasing on $[0,\min(c,d))$. For $x = \max(c,d)$ and $x=\min(c,d)$ the inequality is trivially satisfied with $K=1$. This completes the case of x which does not lie between c and d.

- The second part of Lemma 8 follows easily by Lemma 7, as
$$F_{1/2}(d,c)=\rho_s(d,c)^2\le(\rho_s(d,x)+\rho_s(x,c))^2\le$$
$$2[\rho_s(d,x)^2+\rho_s(x,c)^2]=2[F_{1/2}(d,x)+F_{1/2}(x,c)].$$

- To prove the last part of the statement we will use convexity of F_t. Since $F_t(d,x)+F_t(x,c)\ge F_t(\frac{d+x}{2},\frac{x+c}{2})$, it suffices to prove that $F_t(d,c)\le 12F_t(\frac{d+x}{2},\frac{x+c}{2})$. Thanks to the first part of Lemma 8 we can restrict our considerations to the case $x\in[d,c]$. Note that
$$\frac{\partial}{\partial x}F_t(\frac{d+x}{2},\frac{x+c}{2})$$
$$=\frac{s}{2}[t(\frac{d+x}{2})^{s-1}+(1-t)(\frac{x+c}{2})^{s-1}-(t(\frac{d+x}{2})+(1-t)(\frac{x+c}{2}))^{s-1}]\le 0,$$
since the function $\varphi(u)=u^{s-1}$ is concave. Therefore it is enough to prove that
$$F_t(d,c)\le 12F_t(\frac{d+c}{2},c).$$
Using the homogenity of the above formula we can reduce our task to proving that
$$F_t(1-u,1)\le 12F_t(1-u/2,1)$$
for any $u\in[0,1]$ and $t\in[0,1/2]$.

  Using the Taylor expansion we get
$$F_t(1-u,1)=t(1-u)^s+1-t-(1-tu)^s=$$
$$s(s-1)u^2t(1-t)\cdot[\frac{1}{2}+\sum_{k=1}^{\infty}\frac{u^k}{(k+1)(k+2)}\sum_{m=0}^{k}t^m\cdot\prod_{l=1}^{k}(1-\frac{s-1}{l})].$$
Therefore
$$F_t(1-u/2,1)\ge\frac{1}{2}s(s-1)(u/2)^2t(1-t)$$

and

$$F_t(1-u,1) \le s(s-1)u^2 t(1-t)\cdot\Big[\frac{1}{2}+2\sum_{k=1}^{\infty}\frac{1}{(k+1)(k+2)}\Big]$$

$$= \frac{3}{2}s(s-1)u^2t(1-t)$$

because $\sum_{m=0}^{\infty} t^m \le 2$. Hence

$$F_t(1-u,1) \le 12F_t(1-u/2,1)$$

which completes the proof. $\Box$

Lemma 9 *Let $a \in [0,1]$, $0 \le x_1 < x_2$ and g be a smooth function on $[x_1, x_2]$ such that $g(x_1) = y_1, g(x_2) = y_2$. Then*

$$\int_{x_1}^{x_2} \max(1, x^a) g'(x)^2 d\lambda(x) \ge \frac{(y_2-y_1)^2}{4(e^{x_2}-e^{x_1})}\max(1, x_2^a). \tag{2}$$

Proof. By the Schwartz inequality

$$|y_2 - y_1| \le \int_{x_1}^{x_2} |g'(x)|dx$$

$$\le (\int_{x_1}^{x_2} \max(1,x^a)g'(x)^2 d\lambda(x))^{1/2}(2\int_{x_1}^{x_2}\min(1,x^{-a})e^x dx)^{1/2}.$$

Therefore to show (2) it is enough to prove that

$$f_1(x_2) = \int_{x_1}^{x_2}\min(1,x^{-a})e^x dx \le 2\min(1,x_2^{-a})(e^{x_2}-e^{x_1}) = f_2(x_2).$$

For $x_2 \le 2$ this is obvious because for $0 < x < x_2 \le 2$ we have $\min(1, x^{-a}) \le 1 \le 2\min(1, x_2^{-a})$, and for $x \ge 2$ we have

$$f_2'(x) = 2x^{-a}(e^x - ax^{-1}(e^x - e^{x_1})) \ge x^{-a}e^x = f_1'(x).\Box$$

Lemma 10 *Let $0 \le y_1 < y_2$, $0 \le x_1 < x_2$ and g is defined on $(-\infty, x_2)$ by the formula*

$$g(x) = \begin{cases} y_1 & \text{for } x \le x_1 \\ y_1 + (e^x - e^{x_1})\frac{y_2-y_1}{e^{x_2}-e^{x_1}} & \text{for } x \in (x_1, x_2] \end{cases}.$$

Then

$$\int_{-\infty}^{x_2} g'(x)^2 d\lambda(x) = \frac{(y_2-y_1)^2}{2(e^{x_2}-e^{x_1})}. \tag{3}$$

and for all $p \ge 1$

$$\int_{-\infty}^{x_2} g(x)^p d\lambda(x) \le \lambda(-\infty, x_2)[(1-\frac{x_2}{2}e^{-x_2})y_1^p + \frac{x_2}{2}e^{-x_2}y_2^p]. \tag{4}$$

Proof. Equation (3) follows by direct calculations. It is easy to see that $g(x)$ is maximal (for fixed values of x_2, y_1 and y_2) when $x_1 = 0$, so to prove (4) we may and will assume that this is the case. To easy the notation we will denote x_2 by x. First we will consider $p = 1$. After some standard calculations (4) is equivalent in this case to

$$\frac{e^x(x-1+e^{-x})}{(2e^x-1)(e^x-1)} \le \frac{1}{2}xe^{-x} \text{ for all } x > 0,$$

that is

$$2 + 3x \le xe^{-x} + 2e^x \text{ for all } x > 0,$$

which immeditely follows from well known estimates $e^{-x} \ge 1 - x$ and $e^x \ge 1 + x + x^2/2$.

Now, for arbitrary $p \ge 1$ notice that $g(x) = (1-\theta(x))y_1 + \theta(x)y_2$ with $0 \le \theta(x) \le 1$. Therefore we have by the convexity of x^p

$$\int_{-\infty}^{x_2} g(x)^p d\lambda(x) \le \int_{-\infty}^{x_2} ((1-\theta(x))y_1^p + \theta(x)y_2^p) d\lambda(x) \le$$

$$\lambda(-\infty, x_2)[(1 - \frac{x_2}{2}e^{-x_2})y_1^p + \frac{x_2}{2}e^{-x_2}y_2^p],$$

where the last inequality follows by the previously established case $p = 1$. □

Lemma 11 *Suppose that* $s \in (1, 2]$, $t \in (0, 1)$, $u = \frac{s}{4(s-1)}e^{-s/2(s-1)}$ *and positive numbers* $a, b, c, d, \tilde{a}, \tilde{c}, x$ *satisfy the following conditions*

$$c < x < d, c^s \le a, d^s \le b, \tilde{c}^s \le \tilde{a}, \tilde{c} \le (1-u)c + ux.$$

Then

$$(1-t)a + tb - ((1-t)c + td)^s \le$$
$$8[(1-t)\tilde{a} + tb - ((1-t)\tilde{c} + td)^s + (1-t)a + tx^s - ((1-t)c + tx)^s]. \quad (5)$$

Proof. Without loss of generality we may assume that $a = c^s, b = d^s, \tilde{a} = \tilde{c}^s$. Since the function $y \to (1-t)y^s - ((1-t)y + td)^s$ is nonincreasing on $[0, d]$, it is enough to show that

$$(1-t)c^s + td^s - ((1-t)c + td)^s \le$$
$$3[(1-t)((1-u)c + ud)^s + td^s - ((1-t)(1-u)c + (t + (1-t)u)d)^s].$$

By the homogenity we may and will assume that $d = 1$. We are then to show that

$$f((1-c)) \le 8f((1-u)(1-c)), \quad (6)$$

where

$$f(x) = (1-t)(1-x)^s + t - (1-(1-t)x)^s = \sum_{i=2}^{\infty} (-1)^i \binom{s}{i} (1-t)(1-(1-t)^{i-1})x^k.$$

We use the following simple observation: if a_i, b_i are two summable sequences of positive numbers such that for any $i > j$, $a_i/a_j \geq b_i/b_j$ then for any nondecreasing nonnegative sequence c_i

$$\frac{\sum a_i c_i}{\sum a_i} \geq \frac{\sum b_i c_i}{\sum b_i}.$$

We apply the above to the sequences $a_i = (-1)^i \binom{s}{i}(1-t)(1-(1-t)^{i-1})x^i$, $b_i = (i-1)(-1)^i\binom{s}{i}$ and $c_i = (1-u)^i$, $i = 2, 3, \ldots$ and notice that

$$h(y) := \sum_{i=2}^{\infty} b_i y^i = 1 - (1-y)^{s-1}(1+(s-1)y) \text{ for } y \in [0,1]$$

Therefore we get

$$f((1-u)x) \geq \frac{h(1-u)}{h(1)} = \big(1 - u^{s-1}(1+(s-1)(1-u))\big) f(x)$$

Inequality (6) follows if we notice that

$$u^{s-1}(1+(s-1)(1-u)) \leq su^{s-1} = \frac{s^s}{4^{s-1}} e^{-s/2} (\frac{1}{s-1})^{s-1} \leq 1 e^{-1/2} e^{1/e} \leq \frac{7}{8} \square$$

Proposition 4 *Suppose that for all $p \in [1,2)$ and all nonnegative smooth functions g we have*

$$\int_R g^2 d\lambda - (\int_R g^p d\lambda)^{2/p} \leq K_1 (2-p)^i \int_R (g'(x))^2 \max(1, |x|^i) d\lambda(x) \text{ for } i = 0, 1, \quad (7)$$

where K_1 is a universal constant. Then for all p and g as above we have

$$\int_R g^2 d\lambda - (\int_R g^p d\lambda)^{2/p} \leq$$

$$K_2 (2-p)^a \int_R (g'(x))^2 \max(1, |x|^a) d\lambda(x) \text{ for } a \in (0,1), \quad (8)$$

where $K_2 \leq 32K_1$ is some universal constant.

Proof. An easy approximation argument shows that (7) holds for any continuous function g, continuously differentiable everywhere except possibly finitely many points.

First we assume that g is constant on R^- or R^+, without loss of generality say it is R^-, and we show that (8) holds with $K_2 = 16K_1$. Let us fix $p \in [1,2)$ and define

$$x_p = (2-p)^{-1}, y = g(x_p), t = \lambda(x_p, \infty), s = \frac{2}{p},$$

$$a = \frac{1}{1-t}\int_{-\infty}^{x_p} g^2 d\lambda, b = \frac{1}{t}\int_{x_p}^{\infty} g^2 d\lambda$$

$$c = \frac{1}{1-t}\int_{-\infty}^{x_p} g^p d\lambda \text{ and } d = \frac{1}{t}\int_{x_p}^{\infty} g^p d\lambda.$$

Notice that by Hölder's inequality we have

$$a \geq c^s \text{ and } b \geq d^s. \tag{9}$$

We will consider two cases

Case 1. y^p lies outside (c, d) or $c > d$.

We first apply inequality (7) for $i = 1$ and a function $gI_{(-\infty,x_p)} + yI_{[x_p,\infty)}$ to get

$$(1-t)a + ty^2 - ((1-t)c + ty^p)^s \leq K_1(2-p)\int_0^{x_p} (g'(x))^2 \max(1, |x|) d\lambda(x) \leq$$

$$K_1(2-p)^a \int_0^{x_p} (g'(x))^2 \max(1, |x|^a) d\lambda(x).$$

In a similar way using the case of $i = 0$ for the function $yI_{(-\infty,x_p)} + gI_{[x_p,\infty)}$ we get

$$tb + (1-t)y^2 - (td + (1-t)y^p)^s \leq K_1 \int_{x_p}^{\infty} (g'(x))^2 d\lambda(x) \leq$$

$$K_1(2-p)^a \int_{x_p}^{\infty} (g'(x))^2 \max(1, |x|^a) d\lambda(x).$$

Notice also that

$$\int_R g^2 d\lambda - (\int_R g^p d\lambda)^{2/p} = (1-t)a + tb - ((1-t)c + td)^s \leq$$

$$12\big[(1-t)a + ty^2 - ((1-t)c + ty^p)^s + tb + (1-t)y^2 - (td + (1-t)y^p)^s\big] \leq$$

$$12K_1(2-p)^a \int_R (g'(x))^2 \max(1, |x|^a) d\lambda(x).$$

The middle inequality follows by Lemma 8 with $x = y^p$ together with estimates (9).

Case 2. $c < y^p < d$, we can then find $0 < x_0 < x_p$ such that $g(x_0) = c^{1/p}$. Define new function f by the formula

$$f(x) = \begin{cases} g(x) & \text{for } x > x_p \\ c^{1/p} + \frac{y - c^{1/p}}{e^{x_p} - e^{x_0}}(e^x - e^{x_0}) & \text{for } x \in [x_0, x_p] \\ c^{1/p} & \text{for } x < x_0. \end{cases}$$

Let

$$\tilde{a} = \frac{1}{1-t}\int_{-\infty}^{x_p} f^2 d\lambda \text{ and } \tilde{c} = \frac{1}{1-t}\int_{-\infty}^{x_p} f^p d\lambda.$$

By Lemma 9 and 10 we have

$$\int_R f'(x)^2 d\lambda(x) \le 2(2-p)^a \int_R \max(1,|x|^a) g'(x)^2 d\lambda(x).$$

Therefore by (7) with $i = 0$, used for the function f, we have

$$(1-t)\tilde{a} + tb - ((1-t)\tilde{c}) + td)^s \le 2K_1(2-p)^a \int \max(1,|x|^a) g'(x)^2 d\lambda(x).$$

We conclude as in the previous case using Lemmas 10 and 11 instead of Lemma 8.

Finally suppose that g is arbitrary. A similar argument as in case 1 (but now with $x_p = 0$ and $t = 1/2$) together with the already proved case of g constant on R_- or R_+ proves the assertion in this case. □

Proof of Proposition 3. We need only to prove that assumptions of Proposition 4 are satisfied. But in view of Proposition 2 they are equivalent to the Poincaré inequality for symmetric exponential probability measure ($i = 0$) and the logarithmic Sobolev inequality for the centered $\mathcal{N}(0, \sqrt{2}/2)$ Gaussian measure ($i = 1$) which are well known to hold with some universal constants. This completes the proof. □

In the end of the paper we would like to come back to the class Φ introduced in Definition 4. It is easy to check that if Lemma 5 holds for some function $\varphi \in C^2((0,\infty)) \cap C([0,\infty))$ for any $(\Omega_1,\mu_1),(\Omega_2,\mu_2)$ and any Z then $\varphi \in \Phi$. Indeed, it is even true if we restrict our consideration to (Ω_1,μ_1) and (Ω_2,μ_2) being two-point probability spaces whose atoms have 1/2 measures. This gives a natural characterization of the class Φ.

One can try to generalize the definition of Φ. Let U be an open, convex subset of R^d. We will say that a continuous function $f : U \longrightarrow R$ belongs to the class $C_n(U)$ if for any probability spaces $(\Omega_1,\mu_1),\ldots,(\Omega_n,\mu_n)$ and any integrable random variable Z with values in U, defined on $(\Omega,\mu) = (\Omega_1 \times \ldots \times \Omega_n, \mu_1 \otimes \ldots \otimes \mu_n)$ the following inequality is satisfied:

$$\sum_{K\subseteq\{1,2,\ldots,n\}} (-1)^{|K|} E_{K^c} f(E_K Z) \ge 0,$$

where E_K denotes expectation with respect to μ_k for all $k \in K$. One can easily see that $C_1(U)$ is just a set of all convex functions on U, while $C_2((0,\infty))$ is closely related to the class Φ. In fact $f \in C_2((0,\infty))$ if and only if it is an affine function or it has a continuous strictly positive second derivative such that $1/f''$ is a concave function. One can prove that always $C_{n+1}(U) \subseteq C_n(U)$ and therefore it is natural to define $C_\infty(U)$ as an intersection of all $C_n(U)$. Then it appears that $f \in C_\infty(U)$ if and only if f is given by the formula $f(x) = Q(x,x) + x^*(x) + y$, where Q is a non-negative definite symmetric quadratic form, x^* is a linear functional on R^d and y is a constant. The above inclusions do not need to be strict. For example it is easy to see that $C_2(R)$=$C_\infty(R)$. It would be interesting to know some nice characterization of $C_2(U)$ for general U and $C_n((0,\infty))$ for $n > 2$. It is not clear what applications of C_n for $n > 2$ could be found but it is easy to see that this class has some tensorization property. By now, we do not know even the answer to the following question: For which $p \in [1,2]$ does $f(x) = x^p$ belong to $C_n((0,\infty))$? We can only give some estimates.

These problems will be discussed in a separate paper.

Remark 4 *Recently some new results were announced to the authors by F. Barthe (private communication) - he proved (using Theorem 2 above) that if a log-concave probability measure μ on the Euclidean space $(R^n, \|\cdot\|)$ satisfies inequality $\mu(\{x \in R^n\,; ||x|| > t\}) \le ce^{-(t/c)^r}$ for some constants $c > 0, r \in [1,2]$ and any $t > 0$ then it satisfies also inequality*

$$E_\mu f^2 - (E_\mu f^p)^{2/p} \le C(c,n,r)(2-p)^a E_\mu \|\nabla f\|^2$$

for any non-negative smooth function f on R^n and $p \in [1,2)$, where $C(c,n,r)$ is some positive constant depending on c, n and r only and $a = 2 - 2/r$.

Acknowledgements. The article was inspired by the questions of Prof. Stanisław Kwapień and Prof. Gideon Schechtman. This work was done while the first named author was visiting Southeast Applied Analysis Center at School of Mathematics, Georgia Institute of Technology and was partially supported by NSF Grant DMS 96-32032. The research of the second named author was performed at the Weizmann Institute of Science in Rehovot, Israel and Equipe d'Analyse, Université Paris VI.

References

[AS] S. Aida, D. Stroock *Moment estimates derived from Poincaré and logarithmic Sobolev inequalities,* Math. Res. Lett. 1 (1994), 75-86

[D] E. B. Davies *Heat Kernels and spectral Theory,* Cambridge Tracts in Mathematics, 92, Cambridge University Press, 1989

[G] L. Gross *Logarithmic Sobolev inequalities,* Amer. J. Math. 97 (1975), 1061-1083

[KLO] S. Kwapień, R. Latała, K. Oleszkiewicz *Comparison of Moments of Sums of Independent Random Variables and Differential Inequalities,* J. Funct. Anal. 136 (1996), 258-268

[L] M. Ledoux *Concentration of measure and logarithmic Sobolev inequalities. Séminaire de Probabilités XXXIII,* Lecture Notes in Math., to appear, Springer

[O] K. Oleszkiewicz *Comparison of moments via Poincaré-type inequality,* in Contemporary Mathematics, 234, *Advances in Stochastic Inequalities,* AMS, Providence, 1999, 135-148

Institute of Mathematics
Warsaw University
Banacha 2
02-097 Warszawa
Poland
E-mail: rlatala@mimuw.edu.pl, koles@mimuw.edu.pl