Bernoulli **24**(1), 2018, 257–270
DOI: 10.3150/16-BEJ875

A note on the convex infimum convolution inequality

NAOMI FELDHEIM[*], ARNAUD MARSIGLIETTI[**], PIOTR NAYAR[†] and JING WANG[‡]

Institute for Mathematics and its Applications, University of Minnesota, 207 Church Street SE, 306 Lind Hall, Minneapolis, MN 55455, USA. E-mail: [*]trinomi@gmail.com; [**]arnaud.marsiglietti@ima.umn.edu; [†]nayar@mimuw.edu.pl; [‡]wangjing@illinois.edu

We characterize the symmetric real random variables which satisfy the one dimensional convex infimum convolution inequality of Maurey. We deduce Talagrand's two-level concentration for random vector $(X_1, \dots, X_n)$, where X_i's are independent real random variables whose tails satisfy certain exponential type decay condition.

Keywords: concentration of measure; convex sets; infimum convolution; Poincaré inequality; product measures

1. Introduction

In the past few decades, a lot of attention has been devoted to study the concentration of measure phenomenon, especially the concentration properties of $\mathbb{R}^n$-valued random vectors with independent coordinates. Through this note, we denote by $|x|_p$ the l_p norm on $\mathbb{R}^n$, namely $|x|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$, and let us take $B_p^n = \{x \in \mathbb{R}^n : |x|_p \le 1\}$. We say that an $\mathbb{R}^n$-valued random vector X satisfies concentration with a profile $\alpha_X(t)$ if for any set $A \subset \mathbb{R}^n$ with $\mathbb{P}(X \in A) \ge 1/2$ we have

$$\mathbb{P}\bigl(X \in A + tB_2^n\bigr) \ge 1 - \alpha_X(t) \qquad \text{for all } t \ge 0.$$

An equivalent statement is that for any 1-Lipschitz function f on $\mathbb{R}^n$ the random vector X satisfies the inequality

$$\mathbb{P}\bigl(f(X) > t + \operatorname{Med} f(X)\bigr) \le \alpha_X(t) \qquad \text{for all } t \ge 0,$$

where $\operatorname{Med} f(X)$ is a median of $f(X)$. The above framework includes classical theory of large deviations for sums of independent random variables. The concentration of measure has applications in functional analysis [15], theory of empirical processes (see [14], Chapter 7), random matrix theory and combinatorics (see [14], Chapter 8), and statistical mechanics (see [21]). It can be also investigated in the context of infinite dimensional diffusion generators (see [5]). For the study of concentration and isoperimetry under curvature bounds on Riemannian manifolds, see [18].

A usual way to reach concentration is via certain functional inequalities. For example, if an $\mathbb{R}^n$-valued random vector X satisfies Poincaré inequality with constant C, that is, for any $f \in$

$C^1(\mathbb{R}^n, \mathbb{R})$ we have $\mathrm{Var}(f(X)) \le C\mathbb{E}|\nabla f(X)|_2^2$, then it has exponential concentration with profile $\alpha_X(t) = 2\exp(-t/2\sqrt{C})$. The characterization of random vectors with independent entries satisfying the above inequality is well known, see [19], [17] and [14], Corollary 5.7. Similar characterization is valid for the so-called log-Sobolev inequality, which implies a stronger concentration -the Gaussian concentration phenomenon, namely with a profile $\alpha_X(t) = \exp(-t^2/2C)$, see [9] and [6].

The goal of this article is to investigate concentration properties of a wider class of random vectors, that is, vectors that may not even satisfy Poincaré inequality. For example, in the case of random vector whose law is purely atomic, one can easily construct a non-constant function f with $\mathbb{E}|\nabla f(X)|_2^2 = 0$. However, one can still hope to get concentration if one restricts set A to the class of convex sets. It turns out that to reach exponential concentration for convex sets, it suffices to prove that X satisfies the convex Poincaré inequality, which is stated as below.

Definition 1. *We say that a real random variable X satisfies the convex Poincaré inequality with a constant C_p if for every convex function $f : \mathbb{R} \to \mathbb{R}$ with f' bounded we have*

$$\mathrm{Var}\big(f(X)\big) \le C_p \mathbb{E} f'(X)^2. \tag{1}$$

Here we adopt the standard convention that for a locally Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ the gradient f' is defined by

$$f'(x) = \limsup_{h\to 0} \frac{f(x+h) - f(x)}{h}. \tag{2}$$

This definition applies in particular to convex f. If f is differentiable, (2) agrees with the usual derivative.

Now we consider a class of real random variables with exponentially decaying tails, which allows random variables with atomic distributions.

Definition 2. *Let $h > 0$ and $\lambda \in [0, 1)$. Let $\mathcal{M}(h, \lambda)$ be the class of symmetric real random variables such that for any $X \in \mathcal{M}(h, \lambda)$ it holds that $\mathbb{P}(X \ge x + h) \le \lambda \mathbb{P}(X \ge x)$ for $x \ge 0$. Moreover, let $\mathcal{M}^+(h, \lambda)$ be the class of $\mathbb{R}_+$-valued real random variables, satisfying the same condition.*

The convex Poincaré inequality has been investigated by Bobkov and Götze, see [8], Theorem 4.2. In particular, the authors proved (1) in the class $\mathcal{M}(h, \lambda)$ with a constant C_p depending only on h and λ. This leads to the exponential concentration for 1-Lipschitz convex functions f (as well as the exponential concentration for convex sets) via the standard Herbst argument (see, e.g., [14], Theorem 3.3). In the present article, we show that the class $\mathcal{M}(h, \lambda)$ admits even better concentration properties, that is, the so-called two level Talagrand concentration. We follow the approach developed by Maurey in [16]. Let us introduce the following definition.

Definition 3. *Define*

$$\varphi_0(x) = \begin{cases} \frac{1}{2}x^2, & |x| \le 1, \\ |x| - \frac{1}{2}, & |x| > 1. \end{cases}$$

We say that a real random variable X satisfies convex exponential property (τ) with constant C_τ if for any convex function $f : \mathbb{R} \to \mathbb{R}$ with $\inf f > -\infty$ *we have*

$$\big(\mathbb{E}e^{f \Box \varphi(X)}\big)\big(\mathbb{E}e^{-f(X)}\big) \le 1, \qquad \text{where } \varphi(x) = \varphi_0(x/C_\tau) \tag{3}$$

and $(f \Box \varphi)(x) = \inf_y\{f(y) + \varphi(x-y)\}$ *is the so-called infimum convolution.*

In [16], Maurey showed that every real random variable having values in the set of diameter 1 satisfies the inequality $\mathbb{E}e^{f \Box \varphi(X)}\mathbb{E}e^{-f(X)} \le 1$ for convex f, with $\varphi(x) = \frac{1}{4}|x|^2$ (see [16], Theorem 3). Our result, as stated below, extends both the above fact and the exponential concentration in the class $\mathcal{M}(h, \lambda)$ due to Bobkov and Götze.

Theorem 1.
The following conditions are equivalent

(a) $X \in \mathcal{M}(h, \lambda)$.
(b) *There is $C_p > 0$ such that X satisfies the convex Poincaré inequality with constant C_p.*
(c) *There exists $C_\tau > 0$ such that X satisfies the convex exponential property (τ) with constant C_τ.*

Moreover, (a) *implies* (c) *with the constant* $C_\tau = 17h/(1-\lambda)^2$, (c) *implies* (b) *with the constant* $C_p = \frac{1}{2}C_\tau^2$ *and* (b) *implies* (a) *with* $h = \sqrt{8C_p}$ *and* $\lambda = 1/2$.

This generalizes Maurey's theorem due to the fact that any symmetric $[-1, 1]$-valued real random variable belongs to $\mathcal{M}(1, 0)$.

It is well known that the convex property (τ) tensorizes, namely, if the independent real random variables $X_1, \dots, X_n$ satisfy convex property (τ) with cost functions $\varphi_1, \dots, \varphi_n$ respectively, then the vector $X = (X_1, \dots, X_n)$ has convex property (τ) with $\varphi(x) = \sum_{i=1}^n \varphi_i(x_i)$, see [16], Lemma 5. Therefore Theorem 1 implies the following corollary.

Corollary 1. *Let $X_1, \dots, X_n \in \mathcal{M}(h, \lambda)$ be independent and let us take $X = (X_1, \dots, X_n)$. Define the cost function $\varphi(x) = \sum_{i=1}^n \varphi_0(x_i/C_\tau)$, where $C_\tau = 17h/(1-\lambda)^2$. Then for any convex function f we have*

$$\big(\mathbb{E}e^{f \Box \varphi(X)}\big)\big(\mathbb{E}e^{-f(X)}\big) \le 1. \tag{4}$$

As a consequence, we deduce the two-level concentration for convex sets and convex functions in $\mathbb{R}^n$.

Corollary 2. *Let $X_1, \dots, X_n \in \mathcal{M}(h, \lambda)$ be independent. Let $C_\tau = 17h/(1-\lambda)^2$. Take $X = (X_1, \dots, X_n)$. Then for any convex set A with $\mathbb{P}(X \in A) > 0$ we have*

$$\mathbb{P}\big(X \in A + \sqrt{2t}C_\tau B_2^n + 2tC_\tau B_1^n\big) \ge 1 - \mathbb{P}(X \in A)^{-1}e^{-t}. \tag{5}$$

Corollary 3. *Let $X_1, \dots, X_n \in \mathcal{M}(h, \lambda)$ be independent. Let $C_\tau = 17h/(1-\lambda)^2$. Take $X = (X_1, \dots, X_n)$. Then for any convex function $f : \mathbb{R}^n \to \mathbb{R}$ with*

$$\big|f(x) - f(y)\big|_2 \le a|x-y|_2, \qquad \big|f(x) - f(y)\big|_1 \le b|x-y|_1, \qquad x, y \in \mathbb{R}^n, \tag{6}$$

we have

$$\mathbb{P}\big(f(X) > \mathrm{Med}\big(f(X)\big) + C_\tau t\big) \le 2\exp\left(-\frac{1}{8}\min\left\{\frac{t}{b}, \frac{t^2}{a^2}\right\}\right), \qquad t \ge 0, \tag{7}$$

and

$$\mathbb{P}\big(f(X) < \mathrm{Med}\big(f(X)\big) - C_\tau t\big) \le 2\exp\left(-\frac{1}{8}\min\left\{\frac{t}{b}, \frac{t^2}{a^2}\right\}\right), \qquad t \ge 0. \tag{8}$$

The rest of this article is organized as follows. The next section briefly describes other relevant results existing in the literature. In Section 3, we prove Theorem 1. In Section 4, we deduce Corollaries 2 and 3.

2. Discussion

Poincaré and log-Sobolev inequalities have formulations in the context of Markov processes via the notion of Dirichlet forms, see [3,4,14]. For applications to mixing time bounds, see [10]. Also of interest is Wang's generalized Beckner-type inequality, that interpolates the Poincaré and log-Sobolev inequalities, see [22]. This extends a result of Latała and Oleszkiewicz [12].

The property (τ) for general (not necessarily convex) function was introduced by Maurey in [16] and studied in more details by Latała and Wojtaszczyk (see [13]). An $\mathbb{R}^n$-valued random vector X is said to satisfy property (τ) with a nonnegative cost function φ if the inequality

$$\big(\mathbb{E}e^{f \Box \varphi(X)}\big)\big(\mathbb{E}e^{-f(X)}\big) \le 1 \tag{9}$$

holds for every bounded measurable function f on $\mathbb{R}^n$. Property (τ) implies concentration with respect to level-sets of φ; namely, for every measurable set A we have (see [16], Lemma 4)

$$\mathbb{P}\big(X \notin A + \{\varphi < t\}\big) \le \mathbb{P}(X \in A)^{-1}e^{-t}.$$

In [16], Maurey showed that the vector $Y = (Y_1, \dots, Y_n)$, where $Y_1, \dots, Y_n$ are independent symmetric exponential random variables, satisfies the infimum convolution inequality with the cost function $\varphi(x) = \sum_{i=1}^n \varphi_0(x_i)$, where $\varphi_0(t) = \min\{\frac{1}{36}t^2, \frac{2}{9}(|t|-2)\}$. This leads to a two-level concentration inequality (similar to (5), see [16], Corollary 1), which, as mentioned earlier, is stronger than the usual exponential concentration implied by Poincaré inequality. However, Bobkov, Gentil and Ledoux showed that the property (τ) with a cost function $\varphi_a(x) = \min\{|x|_2^2/a^2, |x|_2/a\}$ is in fact equivalent to Poincaré inequality for smooth $f : \mathbb{R}^n \to \mathbb{R}$ (see [7]).

It is worth mentioning that a much stronger condition than being in $\mathcal{M}(h, \lambda)$, namely $\mathbb{P}(X \ge x + C/x) \le \lambda\mathbb{P}(X \ge x)$, $x \ge m$, where m is a fixed positive number, has been considered. It implies log-Sobolev inequalities for log-convex functions and the Gaussian concentration for convex sets. We refer to the nice study [1] for the details.

Very recently Adamczak and Strzelecki established related results in the context of modified log-Sobolev inequalities, see [2]. For simplicity, we state their result in the case of symmetric real random variables. For $\lambda \in [0,1)$, $\beta \in [0,1]$ and $h, m > 0$ the authors defined the class of random variables $\mathcal{M}^{\beta}_{\mathrm{AS}}(h,\lambda,m)$ satisfying the condition $\mathbb{P}(X \geq x + h/x^{\beta}) \leq \lambda \mathbb{P}(X \geq x)$ for $x \geq m$. Note that $\mathcal{M}^{0}_{\mathrm{AS}}(h,\lambda,0) = M(h,\lambda)$. They proved that any vector $X = (X_1, \dots, X_n)$, where $X_1, \dots, X_n \in \mathcal{M}^{\beta}_{\mathrm{AS}}(h,\lambda,m)$ are independent, satisfy the inequality

$$\mathrm{Ent}\big(e^{f(X)}\big) \leq C_{\mathrm{AS}} \mathbb{E}\Big(e^{f(X)} \big|\nabla f(X)\big|_2^2 \vee |\nabla f|_{\frac{\beta+1}{\beta}}^{\frac{\beta+1}{\beta}}\Big) \tag{10}$$

for any smooth convex function $f : \mathbb{R}^n \to \mathbb{R}$. Here $\mathrm{Ent}(f) = \mathbb{E}(f \ln f) - (\mathbb{E}f)\ln(\mathbb{E}f)$. As a consequence, for any convex set A in $\mathbb{R}^n$ with $\mathbb{P}(X \in A) \geq 1/2$ we have

$$\mathbb{P}\big(X \in A + t^{\frac{1}{2}} B_2^n + t^{\frac{1}{1+\beta}} C_\tau B_{1+\beta}^n\big) \geq 1 - e^{-C'_{\mathrm{AS}} t}, \qquad t \geq 0.$$

Here the constants $C_{\mathrm{AS}}, C'_{\mathrm{AS}}$ depend only on the parameters β, m, h and λ. They also established inequality similar to (7), namely for a convex function f with

$$\big|f(x) - f(y)\big|_2 \leq a|x-y|_2, \qquad \big|f(x) - f(y)\big|_{1+\beta} \leq b|x-y|_{1+\beta}, \qquad x, y \in \mathbb{R}^n,$$

one gets

$$\mathbb{P}\big(f(X) > \mathrm{Med}\big(f(X)\big) + 2t\big) \leq 2\exp\left(-\frac{3}{16}\min\left\{\frac{t^{1+\beta}}{b^{1+\beta} C_{\mathrm{AS}}^{\beta}}, \frac{t^2}{a^2 C_{\mathrm{AS}}}\right\}\right), \qquad t \geq 0.$$

However, the authors in [2] were not able to get (8). In fact one can show that for $\beta = 0$ our Theorem 1 is stronger than (10). In particular, the inequality (10) is equivalent to $\mathbb{E}e^{f \square \varphi(X)} \leq e^{\mathbb{E}f(X)}$, see [2], which easily follows from (4).

While writing this note, we were aware of a work in progress by Gozlan, Roberto, Samson, Shu and Tetali (private communication), which shows an equivalence between the convex property (τ) on the real line and certain mass transportation inequalities (see also [11]).

3. Proof of Theorem 1

We need the following lemma, which is essentially included in Theorem 4.2 of [8]. For reader's convenience, we provide a straightforward proof of this fact.

Lemma 1. *Let $X \in \mathcal{M}^{+}(h,\lambda)$ and let $g : \mathbb{R} \to [0,\infty)$ be non-decreasing with $g(0) = 0$. Then*

$$\mathbb{E}g(X)^2 \leq \frac{2(1+\lambda)}{(1-\lambda)^2} \cdot \mathbb{E}\big(g(X) - g(X-h)\big)^2.$$

Proof. We first prove that $\lambda\mathbb{E}g(X) \geq \mathbb{E}g(X-h)$ for any non-decreasing $g : \mathbb{R} \to [0, \infty)$ such that $g(0) = 0$. Both sides of this inequality are linear in g. Therefore, it is enough to consider only functions of the form $g(x) = \mathbf{1}_{[a,\infty)}(x)$ for $a \geq 0$, since g can be expressed as a mixture of these functions. For $g(x) = \mathbf{1}_{[a,\infty)}(x)$ the above inequality reduces to $\lambda\mathbb{P}(X \geq a) \geq \mathbb{P}(X \geq a+h)$, which is clearly true due to our assumption on X.

The above inequality is equivalent to

$$(1-\lambda)\mathbb{E}g(X) \leq \mathbb{E}\big(g(X) - g(X-h)\big). \tag{11}$$

Now, let us use (11) with g^2 instead of g. Then,

$$\begin{aligned}
\mathbb{E}g(X)^2 &\leq \frac{1}{1-\lambda}\mathbb{E}\big(g(X)^2 - g(X-h)^2\big) \\
&= \frac{1}{1-\lambda}\mathbb{E}\big(g(X) - g(X-h)\big)\big(g(X) + g(X-h)\big) \\
&\leq \frac{1}{1-\lambda}\Big(\mathbb{E}\big(g(X) - g(X-h)\big)^2\Big)^{1/2}\Big(\mathbb{E}\big(g(X) + g(X-h)\big)^2\Big)^{1/2}.
\end{aligned}$$

Moreover, again using (11) for g^2, we get

$$\mathbb{E}\big(g(X) + g(X-h)\big)^2 \leq 2\mathbb{E}\big(g(X)^2 + g(X-h)^2\big) \leq 2(1+\lambda)\mathbb{E}g(X)^2.$$

We arrive at

$$\big(\mathbb{E}g(X)^2\big)^{1/2} \leq \frac{\sqrt{2(1+\lambda)}}{1-\lambda}\Big(\mathbb{E}\big(g(X) - g(X-h)\big)^2\Big)^{1/2}.$$

Our assertion follows. □

In the rest of this note, we take $f : \mathbb{R} \to \mathbb{R}$ to be convex. Let x_0 be a point where f attains its minimal value. Note that this point may not be unique. However, one can check that what follows does not depend on the choice of x_0. Moreover, if f is increasing (decreasing) we adopt the notation $x_0 = -\infty$ ($x_0 = \infty$). Let us define a discrete version of gradient of f,

$$(Df)(x) = \begin{cases} f(x) - f(x-h), & x > x_0 + h, \\ f(x) - f(x_0), & x_0 - h \leq x \leq x_0 + h, \\ f(x) - f(x+h), & x < x_0 - h. \end{cases}$$

Lemma 2. *Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function with $f(0) = 0$ and let $X \in \mathcal{M}(h, \lambda)$. Then*

$$\mathbb{E}\big(e^{f(X)/2} - e^{-f(X)/2}\big)^2 \leq \frac{8}{(1-\lambda)^2}\mathbb{E}\Big(e^{f(X)}\big(Df(X)\big)^2\Big).$$

Proof. *Step 1.* We first assume that f is non-negative and non-decreasing. It follows that $f(x) = 0$ for $x \leq 0$. Correspondingly, the function $g = e^{f/2} - e^{-f/2}$ is non-negative, non-decreasing and

$g(0)=0$. Note that $|X|\in\mathcal{M}^+(h,\lambda)$. From Lemma 1, we get

$$\mathbb{E}\big(e^{f(X)/2}-e^{-f(X)/2}\big)^2=\frac{1}{2}\mathbb{E}g\big(|X|\big)^2\le\frac{1+\lambda}{(1-\lambda)^2}\mathbb{E}\big(g\big(|X|\big)-g\big(|X|-h\big)\big)^2$$
$$=\frac{2(1+\lambda)}{(1-\lambda)^2}\mathbb{E}\big(g(X)-g(X-h)\big)^2.$$

Observe that

$$\begin{aligned}g(x)-g(x-h)&=e^{\frac{f(x)}{2}}-e^{-\frac{f(x)}{2}}-e^{\frac{f(x-h)}{2}}+e^{-\frac{f(x-h)}{2}}\\&=\big(e^{\frac{f(x)}{2}}-e^{\frac{f(x-h)}{2}}\big)\big(1+e^{-\frac{f(x)}{2}-\frac{f(x-h)}{2}}\big)\\&\le 2\big(e^{\frac{f(x)}{2}}-e^{\frac{f(x-h)}{2}}\big)\le e^{\frac{f(x)}{2}}\big(f(x)-f(x-h)\big),\end{aligned}$$

where the last inequality follows from the mean value theorem. Since $\lambda\le 1$, we arrive at

$$\mathbb{E}\big(e^{f(X)/2}-e^{-f(X)/2}\big)^2\le\frac{4}{(1-\lambda)^2}\mathbb{E}e^{f(X)}\big(Df(X)\big)^2.$$

Step 2. Now let f be non-decreasing but not necessarily non-negative. From convexity of f and the fact that $f(0)=0$, we get $|f(-x)|\le f(x)$ for $x\ge 0$. This implies the inequality $|e^{f(-x)}-e^{-f(-x)}|\le|e^{f(x)}-e^{-f(x)}|$, $x\ge 0$. From the symmetry of X, one gets

$$\mathbb{E}\big(e^{f(X)/2}-e^{-f(X)/2}\big)^2\le\mathbb{E}\big(e^{f(|X|)/2}-e^{-f(|X|)/2}\big)^2.$$

Let $\tilde{f}=f\mathbf{1}_{[0,\infty)}$. From Step 1, one gets

$$\mathbb{E}\big(e^{f(|X|)/2}-e^{-f(|X|)/2}\big)^2=2\mathbb{E}\big(e^{\tilde{f}(X)/2}-e^{-\tilde{f}(X)/2}\big)^2\le\frac{8}{(1-\lambda)^2}\mathbb{E}e^{\tilde{f}(X)}\big(D\tilde{f}(X)\big)^2$$
$$\le\frac{8}{(1-\lambda)^2}\mathbb{E}e^{f(X)}\big(Df(X)\big)^2.$$

Step 3. The conclusion of Step 2 is also true in the case of non-increasing functions with $f(0)=0$. This is due to the invariance of our assertion under the symmetry $x\to -x$, which is an easy consequence of the symmetry of X and the fact that for $F(x)=f(-x)$ we have $(DF)(x)=(Df)(-x)$.

Step 4. Let us now eliminate of the assumption of monotonicity of f. Suppose that f is not monotone. Then f has a (not necessarily unique) minimum attained at some point $x_0\in\mathbb{R}$. Due to the remark of Step 3 we can assume that $x_0\le 0$. Since $f(0)=0$, we have $f(x_0)\le 0$. Take $y_0=\inf\{y\in\mathbb{R}: f(y)=0\}$. Clearly, $y_0\le x_0$. We define

$$f_1(x)=\begin{cases}f(x), & x\ge x_0,\\ f(x_0), & x<x_0,\end{cases}\qquad f_2(x)=\begin{cases}0, & x\ge y_0,\\ f(x), & x<y_0.\end{cases}$$

Note that f_1 is non-decreasing and f_2 is non-increasing. Moreover, $f_1(0) = f_2(0) = 0$. Therefore, from the previous steps applied for f_1 and f_2 we get

$$\begin{aligned}&\left(e^{f(x_0)/2} - e^{-f(x_0)/2}\right)^2 \mathbb{P}(X \le x_0) + \mathbb{E}\left(e^{f(X)/2} - e^{-f(X)/2}\right)^2 \mathbf{1}_{\{X \ge x_0\}} \\ &\quad \le \frac{8}{(1-\lambda)^2} \mathbb{E}\left(e^{f(X)} \left(Df(X)\right)^2 \mathbf{1}_{\{X \ge x_0\}}\right)\end{aligned} \tag{12}$$

and

$$\mathbb{E}\left(\left(e^{f(X)/2} - e^{-f(X)/2}\right)^2 \mathbf{1}_{\{X \le y_0\}}\right) \le \frac{8}{(1-\lambda)^2} \mathbb{E}\left(e^{f(X)} \left(Df(X)\right)^2 \mathbf{1}_{\{X \le y_0\}}\right). \tag{13}$$

Moreover, since $|f(x)| \le |f(x_0)|$ on $[y_0, x_0]$, we have

$$\begin{aligned}\mathbb{E}\left(\left(e^{f(X)/2} - e^{-f(X)/2}\right)^2 \mathbf{1}_{\{y_0 \le X \le x_0\}}\right) &\le \left(e^{f(x_0)/2} - e^{-f(x_0)/2}\right)^2 \mathbb{P}\left(X \in [y_0, x_0]\right) \\ &\le \left(e^{f(x_0)/2} - e^{-f(x_0)/2}\right)^2 \mathbb{P}(X \le x_0).\end{aligned} \tag{14}$$

Combining (12), (13) and (14), we arrive at

$$\mathbb{E}\left(e^{f(X)/2} - e^{-f(X)/2}\right)^2 \le \frac{8}{(1-\lambda)^2} \mathbb{E}\left(e^{f(X)} \left(Df(X)\right)^2\right).$$

□

The following lemma provides an estimate on the infimum convolution.

Lemma 3. *Let $C_1, h > 0$. Define $\varphi_1(x) = \frac{1}{C_1}\varphi_0(x/h)$. Assume that a convex function f satisfies $|f'| \le 1/(C_1 h)$. Then*

$$(f \square \varphi_1)(x) \le f(x) - \frac{C_1}{2}\left|(Df)(x)\right|^2.$$

Proof. Let us consider the case when $x \ge x_0 + h$. We take $\theta \in [0, 1]$ and write $y = \theta(x-h) + (1-\theta)x$. Note that $x - y = h\theta$. By the convexity of f, we have

$$\begin{aligned}(f \square \varphi_1)(x) &\le f(y) + \varphi_1(x-y) \le \theta f(x-h) + (1-\theta) f(x) + \varphi_1(h\theta) \\ &= \theta f(x-h) + (1-\theta) f(x) + \frac{1}{2C_1}\theta^2.\end{aligned}$$

Let us now take $\theta = C_1(f(x) - f(x-h))$. Note that $0 \le f' \le 1/C_1 h$ yields $\theta \in [0, 1]$. We get

$$(f \square \varphi_1)(x) \le f(x) - \theta\left(f(x) - f(x-h)\right) + \frac{1}{2C_1}\theta^2 = f(x) - \frac{C_1}{2}\left(f(x) - f(x-h)\right)^2.$$

The case $x \le x_0 - h$ follows by similar computation (one has to take $y = \theta(x+h) + (1-\theta)x$). Also, in the case $x \in [x_0 - h, x_0 + h]$ it is enough to take $y = \theta x_0 + (1-\theta)x$ and use the fact that $|x - y| = |\theta(x - x_0)| \le h\theta$. □

Proof of Theorem 1. We begin by showing that (a) implies (c). We do this in three steps.

Step 1. We first show that it is enough to consider only the case when f satisfies $|f'| \le 1/C_\tau$. To this end, for any convex $f : \mathbb{R} \to \mathbb{R}$ bounded from below, we consider $g(x) = \sup_{y\in\mathbb{R}}\{(f\Box\varphi)(y) - \varphi(x-y)\}$. Thus, $g(x) = \sup_{y\in\mathbb{R}} \inf_{z\in\mathbb{R}}\{f(z) + \varphi(y-z) - \varphi(x-y)\}$. By taking $z = x$, we get $g(x) \le f(x)$ for all $x \in \mathbb{R}$. Moreover, since f is bounded from below, it is easy to see that g is also bounded from below. Since f is convex, one can check that the function $f\Box\varphi$ is also convex (classical property of infimum convolution). Therefore writing $g(x) = \sup_{u\in\mathbb{R}}\{(f\Box\varphi)(x-u) - \varphi(u)\}$, the function g turns out to be convex as a supremum of convex functions. Moreover, since $x \mapsto \varphi(x-y)$ is $(1/C_\tau)$-Lipschitz for every y, the function g is also $(1/C_\tau)$-Lipschitz. Finally, we check that $g\Box\varphi = f\Box\varphi$. Indeed, the inequality $g\Box\varphi \le f\Box, \varphi$ follows from $g \le f$. The other direction is obtained by writing that

$$(g\Box\varphi)(x) = \inf_{y\in\mathbb{R}} \sup_{z\in\mathbb{R}} \inf_{w\in\mathbb{R}} \big\{ f(w) + \varphi(z-w) - \varphi(y-z) + \varphi(x-y) \big\}$$

and by taking $z = x$. Using $g\Box\varphi = f\Box\varphi$, $g \le f$ and the fact that g is $(1/C_\tau)$-Lipschitz, we arrive at

$$\big(\mathbb{E}e^{f\Box\varphi(X)}\big)\big(\mathbb{E}e^{-f(X)}\big) \le \big(\mathbb{E}e^{g\Box\varphi(X)}\big)\big(\mathbb{E}e^{-g(X)}\big) \le 1.$$

Step 2. The inequality (3) stays invariant when we add a constant to the function f. Thus, we may assume that $f(0) = 0$. Note that from the elementary inequality $4ab \le (a+b)^2$ we have

$$4\big(\mathbb{E}e^{f\Box\varphi(X)}\big)\big(\mathbb{E}e^{-f(X)}\big) \le \big(\mathbb{E}\big(e^{f\Box\varphi(X)} + e^{-f(X)}\big)\big)^2.$$

Thus, it is enough to show that

$$\mathbb{E}\big(e^{f\Box\varphi(X)} + e^{-f(X)}\big) \le 2.$$

Step 3. Take $C_1 = 17/(1-\lambda)^2$, $C_\tau = C_1 h$ and $\varphi(x) = \varphi_0(x/C_\tau)$. Assume that $|f'| \le 1/C_\tau$. By the convexity of φ_0 we get $\varphi(x) \le \frac{1}{C_1}\varphi_0(x/h)$, since $C_1 > 1$. Thus, by Lemma 3 we get $f\Box\varphi \le f(x) - \frac{1}{2}C_1|(Df)(x)|^2$. By the mean value theorem $|(Df)(x)|/h \le 1/C_\tau$. Therefore, $\frac{1}{2}C_1|(Df)(x)|^2 \le \frac{1}{2}C_1(\frac{h}{C_\tau})^2 = 1/2C_1$. Let $\alpha(C_1) = 2C_1(1 - \exp(-\frac{1}{2C_1}))$. The convexity of the exponential function yields $e^{-s} \le 1 - \alpha(C_1)s$, $s \in [0, 1/2C_1]$. Therefore,

$$\begin{aligned}\mathbb{E}\big(e^{f\Box\varphi(X)} + e^{-f(X)}\big) &\le \mathbb{E}\big(e^{f(X) - \frac{1}{2}C_1|Df(X)|^2} + e^{-f(X)}\big)\\ &\le \mathbb{E}\bigg(e^{f(X)}\Big(1 - \frac{1}{2}C_1\alpha(C_1)\big|Df(X)\big|^2\Big) + e^{-f(X)}\bigg).\end{aligned}$$

Therefore, since $e^f + e^{-f} - 2 = (e^{f/2} - e^{-f/2})^2$, we are to prove that

$$\mathbb{E}\big(e^{f(X)/2} - e^{-f(X)/2}\big)^2 \le \frac{C_1}{2}\alpha(C_1)\mathbb{E}\big(e^{f(X)}\big|Df(X)\big|^2\big).$$

From Lemma 2, this inequality is true whenever $\frac{1}{2}C_1\alpha(C_1) \geq \frac{8}{(1-\lambda)^2}$. It suffices to observe that

$$\frac{1}{2}C_1\alpha(C_1) = C_1^2\big(1 - e^{-\frac{1}{2C_1}}\big) \geq C_1^2\left(1 - \frac{1}{1+\frac{1}{2C_1}}\right) = \frac{C_1}{2+\frac{1}{C_1}} \geq \frac{C_1}{2+\frac{1}{8}} = \frac{8}{(1-\lambda)^2}.$$

We now sketch the proof of the fact that (c) implies (b), which is well known, see Corollary 3 in [16]. Due to the standard approximation argument one can assume that f is a convex C^2 smooth function with bounded first and second derivative (note that in the definition of the convex Poincaré inequality we assumed that f' is bounded). Consider the function $f_\varepsilon = \varepsilon f$. The infimum of $\psi_x(y) = \varphi(y) + \varepsilon f(x-y)$ is attained at the point y satisfying the equation $\psi_x'(y) = \varphi'(y) - \varepsilon f'(x-y) = 0$. Note that φ' is strictly increasing on the interval $[-C_\tau, C_\tau]$. If ε is sufficiently small, it follows that the above equation has a unique solution y_x and that $y_x \in [-C_\tau, C_\tau]$. Thus, $y_x = C_\tau^2\varepsilon f'(x-y_x)$. This implies $y_x = \varepsilon C_\tau^2 f'(x) + o(\varepsilon)$, where the $o(\varepsilon)$ dependence is uniform in x. We get

$$\begin{aligned} f\Box\varphi(x) &= \varphi(y_x) + \varepsilon f(x-y_x) = \frac{1}{2C_\tau^2}y_x^2 + \varepsilon f\big(x - \varepsilon C_\tau^2 f'(x)\big) + o\big(\varepsilon^2\big) \\ &= \frac{1}{2}\varepsilon^2 C_\tau^2 f'(x)^2 + \varepsilon f(x) - \varepsilon^2 C_\tau^2 f'(x)^2 + o\big(\varepsilon^2\big) \\ &= \varepsilon f(x) - \frac{1}{2}\varepsilon^2 C_\tau^2 f'(x)^2 + o\big(\varepsilon^2\big). \end{aligned}$$

Therefore, from the infimum convolution inequality we get

$$\big(\mathbb{E}e^{\varepsilon f(X) - \frac{1}{2}\varepsilon^2 C_\tau^2 f'(X)^2 + o(\varepsilon^2)}\big)\big(\mathbb{E}e^{-\varepsilon f(X)}\big) \leq 1.$$

Testing (3) with $f(x) = |x|/C_\tau$ one gets that

$$(f\Box\varphi)(x) \geq \inf_y\left(\frac{|y|}{C_\tau} + \frac{|x-y|}{C_\tau} - \frac{1}{2}\right) = \frac{|x|}{C_\tau} - \frac{1}{2}$$

and thus $\mathbb{E}e^{|X|/C_\tau} < \infty$. Also, there exists a constant $c > 0$ such that $|f(x)| \leq c(1+|x|)$, $x \in \mathbb{R}$. As a consequence, one can consider the Taylor expansion of the above quantities in $\varepsilon = 0$. This gives

$$\mathbb{E}\left(1 + \varepsilon f(X) - \frac{1}{2}\varepsilon^2 C_\tau^2 f'(X)^2 + \frac{1}{2}\varepsilon^2 f(X)^2 + o\big(\varepsilon^2\big)\right)\mathbb{E}\left(1 - \varepsilon f(X) + \frac{1}{2}\varepsilon^2 f(X)^2 + o\big(\varepsilon^2\big)\right) \leq 1.$$

Comparing the terms in front of ε^2 leads to

$$\mathbb{E}f(X)^2 - \big(\mathbb{E}f(X)\big)^2 \leq \frac{1}{2}C_\tau^2\mathbb{E}f'(X)^2.$$

This is exactly the Poincaré inequality with constant $\frac{1}{2}C_\tau^2$.

We show that (b) implies (a). Suppose that a symmetric real random variable X satisfies the convex Poincaré inequality with a constant C_p. Consider the function $f_u(x) = \max\{x-u, 0\}$, $u \geq 0$. We have $\mathbb{E}|f_u'(X)|^2 = \mathbb{P}(X \geq u)$. Let Y be an independent copy of X. Since $f_u(y) = 0$ for $y \leq 0$ and $\mathbb{P}(Y \leq 0) \geq 1/2$, one gets

$$\begin{aligned}\mathrm{Var}\big(f_u(X)\big) &= \frac{1}{2}\mathbb{E}\big(f_u(X) - f_u(Y)\big)^2 \geq \frac{1}{2}\mathbb{E}\big(\big(f_u(X) - f_u(Y)\big)^2 \mathbf{1}_{\{Y \leq 0\}}\big) \\ &\geq \frac{1}{4}\mathbb{E}\bigg(\big(f_u(X)\big)^2 \geq \frac{1}{4}\mathbb{E}\big(f_u(X)\big)^2 \mathbf{1}_{\{X \geq u + \sqrt{8C_p}\}}\bigg) \geq 2C_p \mathbb{P}(X \geq u + \sqrt{8C_p}).\end{aligned}$$

These two observations, together with Poincaré inequality, yield that $X \in \mathcal{M}(\sqrt{8C_p}, 1/2)$. $\square$

4. Concentration properties

We show that the convex property (τ) implies concentration for convex sets.

Proposition 1. *Suppose that an $\mathbb{R}^n$-valued random vector X satisfies the property (τ) with a non-negative cost function φ, restricted to the family of convex functions. Let $B_\varphi(t) = \{x \in \mathbb{R}^n : \varphi(x) \leq t\}$. Then for any convex set A we have*

$$\mathbb{P}\big(X \in A + B_\varphi(t)\big) \geq 1 - \mathbb{P}(X \in A)^{-1} e^{-t}.$$

The proof of this proposition is similar to the proof of Proposition 2.4 in [13]. We recall the argument.

Proof of Proposition 1. Let $f = 0$ on A and $f = \infty$ outside of A. Note that f is convex (to avoid working with functions having values $+\infty$ one can consider a family of convex functions $f_n = n\,\mathrm{dist}(A, x)$ and take $n \to \infty$). Suppose that $(f \Box \varphi)(x) \leq t$. Then there exists $y \in \mathbb{R}^n$ such that $f(y) + \varphi(x-y) \leq t$. Thus, $y \in A$ and $x - y \in B_\varphi(t)$. Therefore $x \in A + B_\varphi(t)$. It follows that $x \notin A + B_\varphi(t)$ implies $(f \Box \varphi)(x) > t$. Applying the infimum convolution inequality, we get

$$e^t\big(1 - \mathbb{P}\big(X \in A + B_\varphi(t)\big)\big) \cdot \mathbb{P}(X \in A) \leq \big(\mathbb{E}e^{f \Box \varphi(X)}\big)\big(\mathbb{E}e^{-f(X)}\big) \leq 1.$$

Our assertion follows. $\square$

We are ready to derive the two-level concentration for convex sets.

Proof of Corollary 2. The argument is similar to [16], Corollary 1. Due to Corollary 1, X satisfies property (τ) with the cost function $\varphi(x) = \sum_{i=1}^n \varphi_0(x_i/C_\tau)$. Suppose that $\varphi(x) \leq t$. Define $y, z \in \mathbb{R}^n$ in the following way. Take $y_i = x_i$ if $|x_i| \leq C_\tau$ and $y_i = 0$ otherwise. Take $z_i = x_i$ if $|x_i| > C_\tau$ and $z_i = 0$ otherwise. Then $x = y + z$. Moreover,

$$\sum_{i=1}^n \varphi(y_i/C_\tau) + \sum_{i=1}^n \varphi(z_i/C_\tau) = \sum_{i=1}^n \varphi(x_i/C_\tau) \leq t.$$

In particular $|y|_2^2 \le 2C_\tau^2 t$ and $t \ge \sum_{i=1}^n \varphi_0(z_i/C_\tau) \ge \frac{1}{2}|z|_1/C_\tau$, since $|z_i|/C_\tau - \frac{1}{2} \ge \frac{1}{2}|z_i|/C_\tau$ for $|z_i| \ge 1/C_\tau$. This gives $x \in \sqrt{2t}C_\tau B_2^n + 2tC_\tau B_1^n$. Our assertion follows from Proposition 1. $\square$

Finally, we prove concentration for convex Lipschitz functions.

Proof of Corollary 3. The proof of (7) is similar to the proof of Proposition 4.18 in [14]. Let us define a convex set $A = \{f \le \mathrm{Med}\, f(X)\}$ and observe that $\mathbb{P}(X \in A) \ge 1/2$. Moreover,

$$A + C_\tau\big(\sqrt{2t}B_2^n + 2tB_1^n\big) \subset \big\{f \le \mathrm{Med}\, f(X) + C_\tau(a\sqrt{2t} + 2bt)\big\}.$$

Applying Corollary 2, we get

$$\mathbb{P}\big(f(X) > \mathrm{Med}\, f(X) + C_\tau(a\sqrt{2t} + 2bt)\big) \le 2e^{-t}, \qquad \text{for all } t \ge 0,$$

where a, b are Lipschitz constants given in (6). Take $s = C_\tau(a\sqrt{2t}+2bt)$ and $r = s/C_\tau$. Suppose that $\frac{r}{b} \le \frac{r^2}{a^2}$. Then

$$a\sqrt{2t} + 2tb = r \ge \frac{1}{2}\sqrt{a^2 r/b} + \frac{1}{2}r = a\sqrt{2\frac{r}{8b}} + 2b\frac{r}{4b} \ge a\sqrt{2\frac{r}{8b}} + 2b\frac{r}{8b}.$$

By the monotonicity of $x \mapsto a\sqrt{2x} + 2xb$, $x \ge 0$ it follows that $\frac{1}{8}\min\{\frac{r}{b}, \frac{r^2}{a^2}\} = \frac{r}{8b} \le t$. On the other hand, if $\frac{r}{b} \ge \frac{r^2}{a^2}$, then

$$a\sqrt{2t} + 2tb = r \ge \frac{1}{2}r + \frac{br^2}{2a^2} = a\sqrt{2\frac{r^2}{8a^2}} + 2b\frac{r^2}{4a^2} \ge a\sqrt{2\frac{r^2}{8a^2}} + 2b\frac{r^2}{8a^2}.$$

Therefore, $\frac{1}{8}\min\{\frac{r}{b}, \frac{r^2}{a^2}\} = \frac{r^2}{8a^2} \le t$. Thus,

$$\mathbb{P}\big(f(X) > \mathrm{Med}\, f(X) + rC_\tau\big) \le 2e^{-t} \le 2\exp\left(-\frac{1}{8}\min\left\{\frac{r}{b}, \frac{r^2}{a^2}\right\}\right), \qquad t \ge 0.$$

For the proof of (8), we follow [20]. Define a convex set $B = \{f < \mathrm{Med}\, f(X) - C_\tau(a\sqrt{2t} + 2bt)\}$ with $t \ge 0$. It follows that

$$B + C_\tau\big(\sqrt{2t}B_2^n + 2tB_1^n\big) \subset \big\{f < \mathrm{Med}\, f(X)\big\}$$

and thus Corollary 2 yields

$$\frac{1}{2} \ge \mathbb{P}\big(X \in B + C_\tau\big(\sqrt{2t}B_2^n + 2tB_1^n\big)\big) \ge 1 - \mathbb{P}(X \in B)^{-1}e^{-t}.$$

Therefore $\mathbb{P}(X \in B) \le 2e^{-t}$. To finish the proof we proceed as above. $\square$

Acknowledgements

We would like to thank Radosław Adamczak for his useful comments regarding concentration estimates for Lipschitz functions.

This research was supported in part by the Institute for Mathematics and its Applications with funds provided by the National Science Foundation. The research of the third named author was supported in part by NCN grant DEC-2012/05/B/ST1/00412.

References

[1] Adamczak, R. (2005). Logarithmic Sobolev inequalities and concentration of measure for convex functions and polynomial chaoses. *Bull. Pol. Acad. Sci. Math.* **53** 221–238. MR2163396

[2] Adamczak, R. and Strzelecki, M. (2015). Modified log-Sobolev inequalities for convex functions on the real line. Sufficient conditions. *Studia Math.* **230** 59–93. MR3456588

[3] Bakry, D. and Émery, M. (1984). Hypercontractivité de semi-groupes de diffusion. *C. R. Acad. Sci. Paris Sér. I Math.* **299** 775–778. MR0772092

[4] Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **348**. Cham: Springer. MR3155209

[5] Bakry, D. and Ledoux, M. (1996). Lévy–Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator. *Invent. Math.* **123** 259–281. MR1374200

[6] Barthe, F. and Roberto, C. (2003). Sobolev inequalities for probability measures on the real line. *Studia Math.* **159** 481–497. Dedicated to Professor Aleksander Pełczyński on the occasion of his 70th birthday (Polish). MR2052235

[7] Bobkov, S.G., Gentil, I. and Ledoux, M. (2001). Hypercontractivity of Hamilton–Jacobi equations. *J. Math. Pures Appl.* (9) **80** 669–696. MR1846020

[8] Bobkov, S.G. and Götze, F. (1999). Discrete isoperimetric and Poincaré-type inequalities. *Probab. Theory Related Fields* **114** 245–277. MR1701522

[9] Bobkov, S.G. and Götze, F. (1999). Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** 1–28. MR1682772

[10] Diaconis, P. and Saloff-Coste, L. (1996). Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** 695–750. MR1410112

[11] Gozlan, N., Roberto, C., Samson, P.-M. and Tetali, P. (2014). Kantorovich duality for general transport costs and applications. Preprint. Available at arXiv:1412.7480.

[12] Latała, R. and Oleszkiewicz, K. (2000). Between Sobolev and Poincaré. In *Geometric Aspects of Functional Analysis. Lecture Notes in Math.* **1745** 147–168. Berlin: Springer. MR1796718

[13] Latała, R. and Wojtaszczyk, J.O. (2008). On the infimum convolution inequality. *Studia Math.* **189** 147–187. MR2449135

[14] Ledoux, M. (2001). *The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs* **89**. Providence, RI: Amer. Math. Soc. MR1849347

[15] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces. Isoperimetry and Processes. Ergebnisse der Mathematik und Ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)] **23**. Berlin: Springer. MR1102015

[16] Maurey, B. (1991). Some deviation inequalities. *Geom. Funct. Anal.* **1** 188–197. MR1097258

[17] Miclo, L. (2008). Quand est-ce que des bornes de Hardy permettent de calculer une constante de Poincaré exacte sur la droite? *Ann. Fac. Sci. Toulouse Math.* (6) **17** 121–192. MR2464097

[18] Milman, E. (2009). On the role of convexity in isoperimetry, spectral gap and concentration. *Invent. Math.* **177** 1–43. MR2507637

[19] Muckenhoupt, B. (1972). Hardy's inequality with weights. *Studia Math.* **44** 31–38. Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity. MR0311856

[20] Talagrand, M. (1988). An isoperimetric theorem on the cube and the Kintchine–Kahane inequalities. *Proc. Amer. Math. Soc.* **104** 905–909. MR0964871

[21] Talagrand, M. (2011). *Mean Field Models for Spin Glasses. Volume I. Basic Examples. Ergebnisse der Mathematik und Ihrer Grenzgebiete.* 3 *Folge. A Series of Modern Surveys in Mathematics* [*Results in Mathematics and Related Areas.* 3*rd Series. A Series of Modern Surveys in Mathematics*] **54**. Berlin: Springer. MR2731561

[22] Wang, F.-Y. (2005). A generalization of Poincaré and log-Sobolev inequalities. *Potential Anal.* **22** 1–15. MR2127729

Received October 2015 and revised April 2016