On the convex infimum convolution inequality with optimal cost function

Marta Strzelecka, Michał Strzelecki and Tomasz Tkocz

Institute of Mathematics, University of Warsaw,
Banacha 2, 02–097 Warsaw, Poland.
E-mail address: `martast@mimuw.edu.pl`, `michalst@mimuw.edu.pl`
URL: http://www.mimuw.edu.pl/~martast, http://www.mimuw.edu.pl/~michalst

Mathematics Department, Princeton University,
Fine Hall, Princeton, NJ 08544-1000, USA.
E-mail address: `ttkocz@math.princeton.edu`
URL: http://www.mimuw.edu.pl/~tkocz

Abstract. We show that every symmetric random variable with log-concave tails satisfies the convex infimum convolution inequality with an optimal cost function (up to scaling). As a result, we obtain nearly optimal comparison of weak and strong moments for symmetric random vectors with independent coordinates with log-concave tails.

1. Introduction

Functional inequalities such as the Poincaré, log-Sobolev, or Marton-Talagrand inequality to name a few, play a crucial role in studying concentration of measure, an important cornerstone of the local theory of Banach spaces. In this paper we focus on another example of such inequalities, the infimum convolution inequality, introduced by Maurey (1991).

Let X be a random vector with values in $\mathbb{R}^n$ and let $\varphi : \mathbb{R}^n \to [0, \infty]$ be a measurable function. We say that the pair (X, φ) satisfies the *infimum convolution inequality* (ICI for short) if for every bounded measurable function $f : \mathbb{R}^n \to \mathbb{R}$,

$$\mathbb{E}\, e^{f \square \varphi(X)} \, \mathbb{E}\, e^{-f(X)} \le 1, \tag{1.1}$$

where $f \square \varphi$ denotes the infimum convolution of f and φ defined as $f \square \varphi(x) = \inf\{f(y) + \varphi(x - y) : y \in \mathbb{R}^n\}$ for $x \in \mathbb{R}^n$. The function φ is called a *cost function* and f is called a *test function*. We also say that the pair (X, φ) satisfies the *convex infimum convolution inequality* if (1.1) holds for every convex function $f : \mathbb{R}^n \to \mathbb{R}$ bounded from below.

Maurey showed that Gaussian and exponential random variables satisfy the ICI with a quadratic and quadratic-linear cost function respectively. Thanks to the tensorisation property of the ICI, he recovered the Gaussian concentration inequality as well as the so-called Talagrand two-level concentration inequality for the exponential product measure. Moreover, Maurey proved that bounded random variables satisfy the convex ICI with a quadratic cost function (see Samson, 2007, Equation (3) for an improvement and consult Samson, 2003, Lemma 2.3 for results for non-symmetric Bernoulli measures; see also Shu and Strzelecki, 2017+, Lemma 3.2).

Later on, Maurey's idea was developed further by Latała and Wojtaszczyk (2008) who studied comprehensively the ICI. By testing with linear functions, they observed that the optimal cost function is given by the Legendre transform of the cumulant-generating function (here optimal means largest possible, up to a scaling constant, because the larger the cost function is, the better (1.1) gets). They introduced the notion of optimal infimum convolution inequalities, established them for log-concave product measures and uniform measures on ℓ_p-balls, and put forward important, challenging and far-reaching conjectures (see also Latała, 2017).

The recent works of Gozlan, Roberto, Samson and Tetali (2017) and Gozlan, Roberto, Samson, Shu and Tetali (2017+) enable to view the ICI from a different perspective. Gozlan et al. (2017) introduce weak transport-entropy inequalities and establish their dual formulations. The dual formulations are exactly the convex ICIs. Gozlan et al. (2017+) investigate extensively the weak transport cost inequalities on the real line, obtaining a characterisation for arbitrary cost functions which are convex and quadratic near zero, thus providing a tool for studying the convex ICI. Around the same time, the convex ICI for the quadratic-linear cost function was fully understood by Feldheim, Marsiglietti, Nayar and Wang (2018).

In this paper we go along Latała and Wojtaszczyk's line of research and study the optimal convex ICI. Using the aforementioned novel tools (Gozlan et al., 2017+), we show that product measures with symmetric marginals having log-concave tails satisfy the optimal convex ICI, which complements Latała and Wojtaszczyk's result about log-concave product measures. This has applications to concentration and moment comparison of any norm of such vectors in the spirit of celebrated Paouris' inequality (see Paouris, 2006 and Adamczak, Latała, Litvak, Oleszkiewicz, Pajor and Tomczak-Jaegermann, 2014) and addresses some questions posed lately by Latała and Strzelecka (2017+). We also offer an example showing that the assumption of log-concave tails cannot be weakened substantially.

2. Main results

For a random vector X in $\mathbb{R}^n$ we define

$$\Lambda_X^*(x) := \mathcal{L}\Lambda_X(x) := \sup_{y \in \mathbb{R}^n} \{\langle x, y \rangle - \ln \mathbb{E}\, e^{\langle y, X \rangle}\},$$

which is the Legendre transform of the cumulant-generating function

$$\Lambda_X(x) := \ln \mathbb{E} \, e^{\langle x, X \rangle}, \qquad x \in \mathbb{R}^n.$$

If X is symmetric and the pair (X, φ) satisfies the ICI, then $\varphi(x) \leq \Lambda_X^*(x)$ for every $x \in \mathbb{R}^n$ (see Latała and Wojtaszczyk, 2008, Remark 2.12). In other words, Λ_X^* is the optimal cost function φ for which the ICI can hold. Since this conclusion is obtained by testing (1.1) with linear functions, the same holds for the convex ICI. Following Latała and Wojtaszczyk (2008) we shall say that X satisfies (convex) IC(β) if the pair $(X, \Lambda_X^*(\cdot/\beta))$ satisfies the (convex) ICI.

We are ready to present our first main result.

Theorem 2.1. *Let X be a symmetric random variable with log-concave tails, i.e. such that the function*

$$t \mapsto N(t) := -\ln \mathbb{P}(|X| \geq t), \quad t \geq 0,$$

is convex. Then there exists a universal constant $\beta \leq 1680e$ such that X satisfies convex IC(β).

The (convex) ICI tensorises and, consequently, the property (convex) IC tensorises: if independent random vectors X_i satisfy (convex) IC(β_i), $i = 1, \ldots, n$, then the vector $(X_1, \ldots, X_n)$ satisfies (convex) IC($\max \beta_i$) (see Maurey, 1991 and Latała and Wojtaszczyk, 2008). Therefore we have the following corollary.

Corollary 2.2. *Let X be a symmetric random vector with values in $\mathbb{R}^n$ and independent coordinates with log-concave tails. Then X satisfies convex IC(β) with a universal constant $\beta \leq 1680e$.*

Since, by the Prékopa–Leindler inequality, log-concave measures have log-concave tails, the class of distributions from Corollary 2.2 is wider than the class of symmetric log-concave product distributions considered by Latała and Wojtaszczyk (2008). Among others, it contains measures which do not have a connected support, e.g. a symmetric Bernoulli random variable.

In order to comment on the relevance of the assumptions of Theorem 2.1 and present applications to comparison of weak and strong moments, we need the following definition. Let X be a random vector with values in $\mathbb{R}^n$. We say that the moments of X *grow α-regularly* if for every $p \geq q \geq 2$ and every $\theta \in S^{n-1}$ we have

$$\|\langle X, \theta \rangle\|_p \leq \alpha \frac{p}{q} \|\langle X, \theta \rangle\|_q,$$

where $\|Y\|_p := (\mathbb{E} |Y|^p)^{1/p}$ is the p-th integral norm of a random variable Y. Clearly, if the moments of X grow α-regularly, then α has to be at least 1 (unless $X = 0$ a.s.).

Remark 2.3. If X is a symmetric random variable with log-concave tails, then its moments grow 1-regularly (this classical fact follows for instance from Guédon, Nayar and Tkocz, 2014, Proposition 5.5 and Latała and Wojtaszczyk, 2008, Proof of Proposition 3.8).

The assumption of log-concave tails in Theorem 2.1 cannot be replaced by a weaker one of α-regularity of moments: if X is a symmetric random variable defined by

$$\mathbb{P}(|X| > t) = 1_{[0,2)}(t) + \sum_{k=1}^{\infty} e^{-2^k} 1_{[2^k, 2^{k+1})}(t), \quad t \geq 0, \tag{2.1}$$

then the moments of X grow α-regularly (for some $\alpha < \infty$), but there does not exist $C > 0$ such that the pair $(X, x \mapsto \max\{(Cx)^2, C|x|\})$ satisfies the convex ICI. All the more, X cannot satisfy convex $IC(\beta)$ with any $\beta < \infty$ (see Section 5 for details). Thus it seems that the assumptions of Theorem 2.1 are not far from necessary conditions for the convex ICI to hold with an optimal cost function (random variables with moments growing regularly are akin to random variables with log-concave tails as the former can essentially be sandwiched between the latter, see Latała and Tkocz, 2015, Inequality (4.6)).

Our second main result is an application of Theorem 2.1 to moment comparison. Recall that for a random vector X its p-th weak moment associated with a norm $\|\cdot\|$ is the quantity defined as

$$\sigma_{\|\cdot\|, X}(p) := \sup_{\|t\|_* \leq 1} \|\langle t, X\rangle\|_p,$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$. The following version of Proposition 3.15 from Latała and Wojtaszczyk, 2008 holds (some non-trivial modifications of the proof are necessary in order to deal with the fact that the inequality (1.1) only holds for convex functions).

Theorem 2.4. *Let X be a symmetric random vector with values in $\mathbb{R}^n$ which moments grow α-regularly. Suppose moreover that X satisfies convex $IC(\beta)$. Then for every norm $\|\cdot\|$ on $\mathbb{R}^n$ and every $p \geq 2$ we have*

$$\left(\mathbb{E}\big|\|X\| - E\|X\|\big|^p\right)^{1/p} \leq C\alpha\beta\sigma_{\|\cdot\|, X}(p),$$

where C is a universal constant (one can take $C = 4\sqrt{2}e < 16$).

Immediately, in view of Corollary 2.2 and Remark 2.3, we obtain the following corollary in the spirit of the results of Paouris (2006), Adamczak et al. (2014), Latała and Strzelecka (2017+, 2016). Similar inequalities for Rademacher sums with the emphasis on exact values of constants have also been studied by Oleszkiewicz (2014, Theorem 2.1).

Corollary 2.5. *Let X be a symmetric random vector with values in $\mathbb{R}^n$ and with independent coordinates which have log-concave tails. Then for every norm $\|\cdot\|$ on $\mathbb{R}^n$ and every $p \geq 2$ we have*

$$\left(\mathbb{E}\|X\|^p\right)^{1/p} \leq \mathbb{E}\|X\| + D\sigma_{\|\cdot\|, X}(p), \tag{2.2}$$

where D is a universal constant (one can take $D = 6720\sqrt{2}e^2 < 70223$).

Note that each of the terms on the right-hand side of (2.2) is, up to a constant, dominated by the left-hand side of (2.2), so (2.2) yields the comparison of weak and strong moments of the norms of X.

Note also that the constant standing at $\mathbb{E}\|X\|$ is equal to 1. If we only assume that the coordinates of X are independent and their moments grow α-regularly, then (2.2) does not always hold (the example here is a vector with independent coordinates distributed like in (2.1); see Section 5 for details), although by Theorem 1.1 from Latała and Strzelecka, 2017+ it holds if we allow the constant at $\mathbb{E}\|X\|$ to be greater than 1 and to depend on α. Hence Corollary 2.5 and example (2.1) partially answer the following question raised by Latała and Strzelecka (2017+): "For which vectors does the comparison of weak and strong moments hold with constant 1 at the first strong moment?"

The organisation of the paper is the following. In Section 3 and 4 we present the proofs of Theorem 2.1 and Theorem 2.4 respectively. In Section 5 we discuss example (2.1) in details.

3. Proof of Theorem 2.1

Our approach is based on a characterisation – provided by Gozlan et al. (2017+) – of measures on the real line which satisfy a weak transport-entropy inequality. We emphasise that our optimal cost functions need not be quadratic near the origin, therefore we cannot apply their characterisation as is, but have to first fine-tune the cost functions a bit. We shall also need the following simple lemma.

Lemma 3.1. *If X is a symmetric random variable and $\mathbb{E}\,X^2 = \beta_1^{-2}$, then*

$$\Lambda_X^*(x/\beta_1) \leq x^2 \quad \text{for } |x| \leq 1.$$

Proof: Since X is symmetric, we have

$$\mathbb{E}\,e^{tX} = 1 + \sum_{k=1}^{\infty} \frac{\|X\|_{2k}^{2k} t^{2k}}{(2k)!} \geq 1 + \sum_{k=1}^{\infty} \frac{\|X\|_2^{2k} t^{2k}}{(2k)!} = 1 + \sum_{k=1}^{\infty} \frac{\beta_1^{-2k} t^{2k}}{(2k)!} = \cosh(\beta_1^{-1}|t|).$$

Moreover, $\mathcal{L}\big(\ln\cosh(\cdot)\big)(|u|) \leq |u|^2$ for $|u| \leq 1$ (see for example Latała and Wojtaszczyk, 2008, Proof of Proposition 3.3). Therefore

$$\Lambda_X^*(x/\beta_1) = \mathcal{L}(\Lambda_X(\beta_1\cdot))(x) \leq \mathcal{L}(\ln\cosh(\cdot))(x) \leq x^2 \qquad \text{for } |x| \leq 1. \qquad \square$$

Throughout the proof g^{-1} stands for the generalised inverse of a function g defined as

$$g^{-1}(y) := \inf\{x : g(x) \geq y\}.$$

Proof of Theorem 2.1: Note that $N(0) = 0$ and the function N is non-decreasing. First we tweak the assumptions and change the assertion to a more straightforward one.

Step 1 (first reduction). We claim that it suffices to prove the assertion for random variables for which the function N is strictly increasing on the set where it is finite (or, in other words, $N(t) = 0$ only for $t = 0$). Indeed, suppose we have done this and let now X be any random variable satisfying the assumptions of the theorem. Let X_ε be a symmetric random variable such that $\mathbb{P}(|X_\varepsilon| \geq t) = \exp(-N_\varepsilon(t))$, where $N_\varepsilon(t) = N(t) \vee \varepsilon t$. Consider versions of X and X_ε on the probability space $(0,1)$ (equipped with Lebesgue measure) constructed as the (generalised) inverses of their cumulative distribution functions. Then $|X_\varepsilon| \leq |X|$ a.s. (and also $X_\varepsilon \to X$ a.s. as $\varepsilon \to 0^+$). Hence $\Lambda_{X_\varepsilon} \leq \Lambda_X$ and therefore also $\Lambda_{X_\varepsilon}^* \geq \Lambda_X^*$.

The theorem applied to the random variable X_ε and the above inequality imply that the pair $(X_\varepsilon, \Lambda_X^*(\cdot/\beta))$ satisfies the convex ICI. Taking $\varepsilon \to 0^+$ we get the assertion for X (in the second integral we just use the fact that the test function f is bounded from below and thus e^{-f} is bounded from above; for the first integral it suffices to prove the convergence of integrals on any interval $[-M, M]$, and on such an interval we have $f\square\Lambda_X^*(x/\beta) \leq f(x) + \Lambda_X^*(0) = f(x)$, and thus $\exp(\max_{[-M,M]} f)$ is a good majorant).

Step 2 (second reduction). We claim that it suffices to prove the assertion for random variables such that $\Lambda_X < \infty$. Indeed, suppose we have done this and let X be any random variable satisfying the assumptions of the theorem. Let $N_\varepsilon(t) = N(t) \vee \varepsilon^2 t^2$ and let X_ε be a symmetric random variable such that $\mathbb{P}(|X_\varepsilon| \geq$

$t) = \exp(-N_\varepsilon(t))$. Then, similarly as in Step 1., $\Lambda_{X_\varepsilon} \leq \Lambda_Y < \infty$, where Y is symmetric and $\mathbb{P}(|Y| \geq t) = \exp(-\varepsilon^2 t^2)$. Thus we can apply the proposition to X_ε and we continue as in Step 1.

Step 3 (scaling). Due to the scaling properties of the Legendre transform, we can fix the value of $\mathbb{E}\,X^2 = \beta_1^{-2}$. We choose $\beta_1 := 2e$ (the case where $X \equiv 0$ is trivial), so that, by Markov's inequality, $e^{-N(1/2)} = \mathbb{P}(|X| \geq \frac{1}{2}) \leq 4\,\mathbb{E}\,X^2 = e^{-2}$ and equivalently

$$N(1/2) \geq 2. \tag{3.1}$$

Step 4 (reformulation). For $x \in \mathbb{R}$ let

$$\varphi(x) := \big(x^2 1_{\{|x|<1\}} + (2|x|-1)1_{\{|x|\geq 1\}}\big) \vee \Lambda_X^*(x/(2\beta_1)).$$

We claim that there exists a universal constant $\widetilde{b} \leq 1/420$, such that the pair $(X, \varphi(\widetilde{b}\,\cdot))$ satisfies the convex infimum convolution inequality. Of course the assertion follows immediately from that.

Note that φ is convex, increasing on $[0, \infty)$ (because $\Lambda_X^*(\cdot/(2\beta_1))$ is convex and symmetric and thus non-decreasing on $[0, \infty)$). Crucially, $\varphi(x) = x^2$ for $x \in [0, 1]$ (by Lemma 3.1), so the cost function φ is quadratic near zero. Moreover, by Lemma 3.1, $\varphi^{-1}(3) = 2$.

Let $U = F^{-1} \circ F_\nu$, where F, F_ν are the distribution functions of X and the symmetric exponential measure ν on $\mathbb{R}$, respectively. By Gozlan et al., 2017+, Theorem 1.1 we know that if there exists $b > 0$ such that for every $x, y \in \mathbb{R}$ we have

$$\big|U(x) - U(y)\big| \leq \frac{1}{b}\varphi^{-1}\big(1 + |x - y|\big), \tag{3.2}$$

then the pair $(X, \varphi(\widetilde{b}\,\cdot))$, where $\widetilde{b} = \frac{b}{210\varphi^{-1}(2+1^2)} = \frac{b}{420}$, satisfies the convex ICI. We will show that (3.2) holds with $b = 1$.

Step 5 (further reformulation). Let $a = \inf\{t > 0 : N(t) = \infty\}$. We have three possibilities (recall that N is left-continuous):

- $a = \infty$. Then N is continuous, increasing, and transforms $[0, \infty]$ onto $[0, \infty]$. Also, F is increasing and therefore F^{-1} is the usual inverse of F.
- $a < \infty$ and $N(a) < \infty$. Then X has an atom at a. Moreover, $N(a) = \lim_{t \to a^-} N(t)$.
- $a < \infty$ and $N(a) = \infty = \lim_{t \to a^-} N(t)$.

Of course, in the first case one can extend N by putting $N(a) = \infty$, so that all formulas below make sense.

Note that

$$F(t) = \begin{cases} \frac{1}{2}\exp(-N(|t|)) & \text{if } t < 0, \\ 1 - \frac{1}{2}\exp(-N_+(t)) & \text{if } t \geq 0, \end{cases}$$

where $N_+(t)$ denotes the right-sided limit of N at t (which is different from $N(t)$ only if $t = a$ and X has an atom at a). Hence, F is continuous on the interval $(-a, a)$, the image of $(-a, a)$ under F is the interval $\big(\frac{1}{2}\exp(-N(a)), 1 - \frac{1}{2}\exp(-N(a))\big)$, and we have $F(-a) = \frac{1}{2}\exp(-N(a))$ and $F(a) = 1$. Since the image of $\mathbb{R}$ under U is equal to the image of $(0, 1)$ under F^{-1}, we conclude that $U(\mathbb{R}) = (-a, a)$ if $N(a) = \infty$ and $U(\mathbb{R}) = [-a, a]$ if $N(a) < \infty$. Denote $A := U(\mathbb{R})$.

When $N(a) < \infty$, it suffices to check condition (3.2) for $x, y \in [-N(a), N(a)]$ (otherwise one can change x, y and decrease the right-hand side while not changing the value of the left-hand side of (3.2)). For $x \in [-N(a), N(a)]$ we can write

$U^{-1}(x) = N(|x|)\operatorname{sgn}(x)$ and $U^{-1}(x) \in \mathbb{R}$. When $N(a) = \infty$, U is a bijection (on its image), so we can obviously write again $U^{-1}(x) = N(|x|)\operatorname{sgn}(x)$ for any $x \in \mathbb{R}$.

Therefore, in order to verify (3.2) we need to check that

$$|x - y| \le \varphi^{-1}\big(1 + \big|N(|x|)\operatorname{sgn}(x) - N(|y|)\operatorname{sgn}(y)\big|\big) \qquad \text{for } x, y \in A. \qquad (3.3)$$

Since we consider the case when $\Lambda_X(t)$ is finite for every $t \in \mathbb{R}$, the Chernoff inequality applies, so for $t \ge \mathbb{E}\,X = 0$ we have

$$\frac{1}{2}e^{-N(t)} = \mathbb{P}(X \ge t) \le e^{-\Lambda_X^*(t)},$$

so

$$N(t) \ge \Lambda_X^*(t) - \ln 2. \qquad (3.4)$$

Note that $\varphi(|x - y|) < \infty$ for $x, y \in A$, since $\varphi(|x - y|) = \infty$ would imply $\Lambda_X^*(|x - y|/(2\beta_1)) = \infty$, and hence $\Lambda_X^*(|x - y|/2) = \infty$, and – by (3.4) – also $N(|x - y|/2) = \infty$, but for $x, y \in A$ we have $|x - y|/2 \in [0, a)$ when $N(a) = \infty$ or $|x - y|/2 \in [0, a]$ when $N(a) < \infty$ and in either case $N(|x - y|/2)$ is finite. Therefore for every $x, y \in A$ we have $\varphi(|x - y|) < \infty$. Since $\varphi^{-1}(\varphi(z)) = z$ for z such that $\varphi(z) < \infty$ (because φ is then continuous and increasing on $[0, z]$), the condition (3.3) is implied by

$$\varphi(|x - y|) \le 1 + \big|N(|x|)\operatorname{sgn} x - N(|y|)\operatorname{sgn} y\big| \qquad \text{for } x, y \in A. \qquad (3.5)$$

In the next step we check that this is indeed satisfied.

Step 6 (checking the condition). Let $x_0 = \inf\{x \ge 1 : 2x - 1 = \Lambda_X^*(\frac{x}{2\beta_1})\}$ (if $x_0 = \infty$ we simply do not have to consider Case 2 below). We consider three cases. We repeatedly use the fact that $uN(t) \ge N(ut)$ for $u \le 1$, $t \ge 0$, which follows by the convexity of N and the property $N(0) = 0$.

Case 1. $|x - y| \le 1$. Then $\varphi(|x - y|) = (x - y)^2 \le 1$, so (3.5) is trivially satisfied.

Case 2. $|x - y| \ge x_0$. Then $\varphi(|x - y|) = \Lambda_X^*(\frac{1}{2\beta_1}|x - y|) \le \Lambda_X^*(|x - y|/2)$. Inequality (3.4) implies that in order to prove (3.5) it suffices to show that if x, y are of the same sign, say $x, y \ge 0$, then $N(|x - y|/2) \le |N(x) - N(y)|$ and if x, y have different signs, we have $N((|x| + |y|)/2) \le N(|x|) + N(|y|)$.

By the convexity of N, for $s, t \ge 0$ we have

$$N((s + t)/2) \le \frac{1}{2}N(s) + \frac{1}{2}N(t) \le N(s) + N(t)$$

and

$$N(s/2) + N(t) \le N(s) + N(t) \le \frac{s}{s + t}N(s + t) + \frac{t}{s + t}N(s + t) = N(s + t).$$

This finishes the proof of (3.5) in Case 2.

Case 3. $1 \le |x - y| \le x_0$. Then $\varphi(|x - y|) = 2|x - y| - 1$. Consider two sub-cases:

(i) x, y have different signs. Without loss of generality we may assume $x \ge |y| \ge 0 \ge y$. Thus in order to obtain (3.5) it suffices to show that $N(x) \ge 2x + 2|y|$. Note that $1 \le x + |y| \le 2x$, so $x \ge \frac{1}{2}$. Thus

$$N(x) \ge N(1/2)2x \overset{(3.1)}{\ge} 4x \ge 2x + 2|y|,$$

which finishes the proof in case (i).

(ii) x, y have the same sign. Without loss of generality we may assume $x \geq y \geq 0$. Thus it suffices to show that $2(x - y) \leq N(x) - N(y)$. Note that due to the assumption of Case 3 we have $x \geq x - y \geq 1 \geq \frac{1}{2}$, so by the convexity of N we have

$$\frac{N(x) - N(y)}{x - y} \geq \frac{N(\frac{1}{2}) - N(0)}{\frac{1}{2} - 0} \overset{(3.1)}{\geq} 4 \geq 2$$

This ends the examination of case (ii) and the proof of the theorem. □

4. Comparison of weak and strong moments

The goal of this section is to establish the comparison of weak and strong moments with respect to any norm $\| \cdot \|$ for random vectors X with independent coordinates having log-concave tails (Corollary 2.5). In view of Theorem 2.1 and Remark 2.3, it is enough to show Theorem 2.4.

Our proof of Theorem 2.4 comprises three steps: first we exploit α-regularity of moments of X to control the size of its cumulant-generating function Λ_X, second we bound from below the infimum convolution of the optimal cost function with the convex test function being the norm $\| \cdot \|$ properly rescaled, and finally by the property convex $\mathrm{IC}(\beta)$ we obtain exponential tail bounds which integrated out give the desired moment inequality.

We start with two lemmas corresponding to the first two steps described above and then we put everything together.

Lemma 4.1. *Let $p \geq 2$ and suppose that the moments of a random vector X in $\mathbb{R}^n$ grow α-regularly. If for a vector $u \in \mathbb{R}^n$ we have $\|\langle u, X \rangle\|_p \leq 1$, then*

$$\Lambda_X((2e\alpha)^{-1}pu) \leq p.$$

Proof: Let k_0 be the smallest integer larger than p. If $\alpha e\|\langle u, X \rangle\|_p \leq 1/2$, then by α-regularity we have

$$\Lambda_X(pu) \leq \ln\Big(\sum_{k \geq 0} \frac{\mathbb{E}\,|\langle pu, X \rangle|^k}{k!}\Big) \leq \ln\Big(\sum_{0 \leq k \leq p} p^k \frac{\|\langle u, X \rangle\|_p^k}{k!} + \sum_{k > p}(\alpha k)^k \frac{\|\langle u, X \rangle\|_p^k}{k!}\Big)$$

$$\leq \ln\Big(\sum_{0 \leq k \leq p} \frac{p^k \|\langle u, X \rangle\|_p^k}{k!} + \sum_{k > p}\big(\alpha e\|\langle u, X \rangle\|_p\big)^k\Big)$$

$$\leq \ln\Big(\sum_{0 \leq k \leq p} \frac{p^k \|\langle u, X \rangle\|_p^k}{k!} + 2\big(\alpha e\|\langle u, X \rangle\|_p\big)^{k_0}\Big)$$

$$\leq \ln\Big(\sum_{0 \leq k \leq p} \frac{p^k \|\langle u, X \rangle\|_p^k}{k!} + \frac{\big(2\alpha e p\|\langle u, X \rangle\|_p\big)^{k_0}}{k_0!}\Big)$$

$$\leq \ln\Big(\sum_{0 \leq k \leq k_0} \frac{\big(2\alpha e p\|\langle u, X \rangle\|_p\big)^k}{k!}\Big) \leq 2\alpha e p\|\langle u, X \rangle\|_p \leq p.$$

Replace u with $(2e\alpha)^{-1}u$ to get the assertion. □

Lemma 4.2. *Let $\| \cdot \|$ be a norm on $\mathbb{R}^n$ and let X be a random vector with values in $\mathbb{R}^n$ and moments growing α-regularly. For $\beta > 0$, $p \geq 2$, and $x \in \mathbb{R}^n$ we have*

$$\big(\Lambda_X^*\left(\cdot/\beta\right) \square a\| \cdot \|\big)(x) \geq a\|x\| - p,$$

where $a = p(2e\alpha\beta\sigma_{\|\cdot\|,X}(p))^{-1}$.

Proof: For $f(x) = a\|x\|$ with positive a being arbitrary for now we bound the infimum convolution as follows:

$$\left(\Lambda_X^*(\cdot/\beta)\Box f\right)(x) = \inf_y \sup_z \left\{\beta^{-1}\langle y, z\rangle - \Lambda_X(z) + a\|x - y\|\right\}$$

$$= \inf_y \sup_u \left\{(2e\alpha\beta)^{-1}p\langle y, u\rangle - \Lambda_X((2e\alpha)^{-1}pu) + a\|x - y\|\right\}$$

$$\geq \inf_y \sup_{u:\|\langle u,X\rangle\|_p \leq 1} \left\{(2e\alpha\beta)^{-1}p\langle y, u\rangle - p + a\|x - y\|\right\},$$

where in the last inequality we have used Lemma 4.1. Choose $u = \sigma_{\|\cdot\|,X}(p)^{-1}v$ with $\|v\|_* \leq 1$ such that $\langle y, v\rangle = \|y\|$. Then clearly $\|\langle u, X\rangle\|_p \leq 1$ and thus

$$\Lambda_X^*(\cdot/\beta)\Box f(x) \geq \inf_y \left\{(2e\alpha\beta\sigma_{\|\cdot\|,X}(p))^{-1}p\|y\| - p + a\|x - y\|\right\}.$$

If we now set $a = p(2e\alpha\beta\sigma_{\|\cdot\|,X}(p))^{-1}$, then by the triangle inequality we obtain the desired lower bound

$$\left(\Lambda_X^*(\cdot/\beta)\Box a\|\cdot\|\right)(x) \geq a\|x\| - p. \qquad \Box$$

Proof of Theorem 2.4: Let $f(x) = a\|x\|$ with $a = p(2e\alpha\beta\sigma_{\|\cdot\|,X}(p))^{-1}$ as in Lemma 4.2. Testing the property convex IC(β) with f and applying Lemma 4.2 yields

$$\mathbb{E}\,e^{a\|X\|}\,\mathbb{E}\,e^{-a\|X\|} \leq e^p.$$

By Jensen's inequality we obtain that both $\mathbb{E}\,e^{a(\|X\|-\mathbb{E}\|X\|)}$ and $\mathbb{E}\,e^{a(-\|X\|+\mathbb{E}\|X\|)}$ are bounded above by e^p. Thus Markov's inequality implies the tail bound

$$\mathbb{P}\left(a\big|\|X\| - \mathbb{E}\|X\|\big| > t\right) \leq 2e^{-t}e^p \leq 2e^{-t/2}, \quad t \geq 2p.$$

Consequently,

$$a^p\,\mathbb{E}\big|\|X\| - \mathbb{E}\|X\|\big|^p = \int_0^\infty pt^{p-1}\mathbb{P}\left(a\big|\|X\| - \mathbb{E}\|X\|\big| > t\right)dt$$

$$\leq (2p)^p + 2\int_0^\infty pt^{p-1}e^{-t/2}dt = (2p)^p + 2\cdot 2^p p\Gamma(p)$$

$$\leq 2(2p)^p.$$

Plugging in the value of a gives the result (we can take $C = 4\sqrt{2}e < 16$). $\qquad \Box$

5. An example

Let X be a symmetric random variable defined by $\mathbb{P}(|X| > t) = T(t)$, where

$$T(t) := 1_{[0,2)}(t) + \sum_{k=1}^\infty e^{-2^k}1_{[2^k, 2^{k+1})}(t), \quad t \geq 0, \tag{5.1}$$

or, in other words, let $|X|$ have the distribution

$$(1 - e^{-2})\delta_2 + \sum_{k=2}^\infty \left(e^{-2^{k-1}} - e^{-2^k}\right)\delta_{2^k}.$$

Let us first show that the moments of X grow 3-regularly, but X does not satisfy IC(β) for any $\beta < \infty$ (we also prove a slightly stronger statement later).

Let Y be a symmetric exponential random variable. Then Y has log-concave tails, so the moments of Y grow 1-regularly (see Remark 2.3). Moreover, if X

and Y are constructed in the standard way by the inverses of their cumulative distribution functions on the probability space $(0, 1)$, then

$$|Y| \leq |X| \leq 2|Y| + 2.$$

Therefore, for $p \geq q \geq 2$,

$$\|X\|_p \leq 2\|Y\|_p + 2 \leq 2\frac{p}{q}\|Y\|_q + 2 \leq 3\frac{p}{q}\|X\|_q$$

(we used the fact that $|X| \geq 2$ in the last inequality). Thus the moments of X grow 3-regularly.

On the other hand, for every $h > 0$ there exists $t > 0$ such that

$$\mathbb{P}(|X| \geq t + h) = \mathbb{P}(|X| \geq t).$$

Therefore by Theorem 1 from Feldheim et al. (2018) there does not exist a constant C such that the pair $(X, \varphi(\cdot/C))$, where $\varphi(x) = \frac{1}{2}x^2 1_{\{|x|\leq 1\}} + (|x| - 1/2)1_{\{|x|>1\}}$, satisfies the convex infimum convolution inequality. But, by symmetry and the 3-regularity of moments of X,

$$\Lambda_X(s) \leq \ln\Big(1 + \sum_{k\geq 1}\frac{s^{2k}\,\mathbb{E}\,X^{2k}}{(2k)!}\Big) \leq \ln\Big(1 + \sum_{k\geq 1}\frac{s^{2k}(3k)^{2k}\big(\mathbb{E}\,X^2\big)^k}{(2k)!}\Big)$$

$$\leq \ln\Big(1 + \sum_{k\geq 1}s^{2k}(3e/2)^{2k}\big(\mathbb{E}\,X^2\big)^k\Big) = \ln\Big(1 + \sum_{k\geq 1}\big(9e^2 s^2\,\mathbb{E}\,X^2/4\big)^k\Big).$$

Hence, using the inequality $\ln(1 + x) \leq x$ and summing the geometric series, we arrive at

$$\Lambda_X(s) \leq \frac{9e^2\,\mathbb{E}\,X^2}{4 - 9e^2 s^2\,\mathbb{E}\,X^2} \cdot s^2,$$

provided that $|s| \leq \varepsilon$ for some small enough ε. Thus for some $A, \varepsilon > 0$ we have $\Lambda_X(s) \leq As^2$ for $|s| \leq \varepsilon$. Having chosen ε and possibly increasing A, we can also guarantee that $2A\varepsilon^2 \geq 1$. Hence

$$\Lambda_X^*(t) \geq \sup_{|s|\leq\varepsilon}\big\{st - As^2\big\} = \tfrac{1}{4A}t^2 1_{\{|t|\leq 2A\varepsilon\}} + (\varepsilon|t| - A\varepsilon^2)1_{\{|t|>2A\varepsilon\}}$$

$$= 2A\varepsilon^2\varphi\big(t/(2A\varepsilon)\big) \geq \varphi\big(t/(2A\varepsilon)\big).$$

We conclude that X cannot satisfy IC(β) for any β.

Remark 5.1. Let us also sketch an alternative approach. Take $a, c > 0$, $b \in \mathbb{R}$, and denote $\varphi(x) = \min\{x^2, |x|\}$, $f(x) = f_{a,b}(x) = a(x - b)_+$ for $x \in \mathbb{R}$. One can check that

$$\big(f\Box\varphi(c\cdot)\big)(x) = \begin{cases} 0 & \text{if } x \leq b, \\ c^2(x - b)^2 & \text{if } b < x \leq b + 1/c, \\ c(x - b) & \text{if } x > b + 1/c, \end{cases}$$

if $a > 2c$. It is rather elementary but cumbersome to show that for any $c > 0$ there exist $a > 0$ and $b \in \mathbb{R}$ such that (1.1) is violated by the test function f. We omit the details.

In fact, the above example shows that even a slightly stronger statement is true: for vectors with independent coordinates with α-regular growth of moments the comparison of weak and strong moments of norms does not hold with the constant 1 at the first strong moment. More precisely, let $X_1, X_2, \ldots$ be independent random

variables with distribution given by (5.1). We claim that there does *not* exist any $K < \infty$ such that

$$\big(\mathbb{E}\max_{i\leq n}|X_i|^p\big)^{1/p} \leq \mathbb{E}\max_{i\leq n}|X_i| + K\sup_{\|t\|_1\leq 1}\Big(\mathbb{E}|\sum_{i=1}^n t_iX_i|^p\Big)^{1/p} \qquad (5.2)$$

holds for every $p \geq 2$ and $n \in \mathbb{N}$ (note that we chose the ℓ^∞-norm as our norm). We shall estimate the three expressions appearing in (5.2).

We have

$$\sup_{\|t\|_1\leq 1}\Big(\mathbb{E}|\sum_{i=1}^n t_iX_i|^p\Big)^{1/p} \leq \sup_{\|t\|_1\leq 1}\sum_{i=1}^n|t_i|\|X_i\|_p = \|X_1\|_p \qquad (5.3)$$

(this inequality is in fact an equality). Since the moments of X_1 grow 3-regularly, the last term in (5.2) is bounded by $\widetilde{K}p$ for some $\widetilde{K} < \infty$.

To estimate the remaining two terms we need the following standard fact.

Lemma 5.2. *For independent events $A_1, \ldots, A_n$,*

$$(1-e^{-1})\Big(1\wedge\sum_{i=1}^n\mathbb{P}(A_i)\Big) \leq \mathbb{P}\Big(\bigcup_{i=1}^n A_i\Big) \leq 1\wedge\sum_{i=1}^n\mathbb{P}(A_i).$$

In particular, for i.i.d. non-negative random variables $Y_1, \ldots, Y_n$,

$$(1-e^{-1})\int_0^\infty\Big[1\wedge n\mathbb{P}(Y_1 > t)\Big]dt \leq \mathbb{E}\max_{i\leq n}Y_i \leq \int_0^\infty\Big[1\wedge n\mathbb{P}(Y_1 > t)\Big]dt.$$

Proof: The upper bound is just the union bound. The lower bound follows from de Morgan's laws combined with independence, which imply that

$$\mathbb{P}\Big(\bigcup_{i=1}^n A_i\Big) = 1 - \prod_{i=1}^n(1 - \mathbb{P}(A_i)),$$

and from the inequalities $1 - x \leq e^{-x}$ and $1 - e^{-y} \geq (1-e^{-1})y$ for $x \in \mathbb{R}$, $y \in [0,1]$. $\qquad\square$

Fix $m \geq 2$ and let $e^{2^{m-1}} \leq n < e^{2^m}$. Then

$$1 \wedge nT(t) = \begin{cases} 1 & \text{if } 0 < t < 2^m, \\ nT(t) & \text{if } t \geq 2^m. \end{cases}$$

By the above lemma,

$$\mathbb{E}\max_{i\leq n}|X_i| \leq \int_0^{2^m}dt + n\int_{2^m}^\infty T(t)dt = 2^m + n\sum_{j=m}^\infty e^{-2^j}(2^{j+1} - 2^j)$$

$$= 2^m + n\sum_{j=m}^\infty e^{-2^j}2^j \leq 2^m + ne^{-2^m}2^m\sum_{j=0}^\infty(2e^{-2^m})^j = 2^m + \frac{ne^{-2^m}2^m}{1 - 2e^{-2^m}}.$$

Set $\theta = \theta(m,n) = ne^{-2^m} \in [e^{-2^{m-1}}, 1)$. Then

$$\mathbb{E}\max_{i\leq n}|X_i| \leq 2^m\Big(1 + \frac{\theta}{1 - 2e^{-2^m}}\Big). \qquad (5.4)$$

Similarly,

$$\mathbb{E} \max_{i \leq n} |X_i|^p \geq (1 - e^{-1}) \int_0^\infty 1 \wedge nT(t^{1/p})dt$$

$$= (1 - e^{-1}) \Big[\int_0^{2^{mp}} dt + n \int_{2^{mp}}^\infty T(t^{1/p})dt \Big]$$

$$= (1 - e^{-1}) \Big[2^{mp} + n \sum_{j=m}^\infty e^{-2^j} \big(2^{(j+1)p} - 2^{jp} \big) \Big].$$

Hence

$$\mathbb{E} \max_{i \leq n} |X_i|^p > (1 - e^{-1})ne^{-2^m} \big(2^{(m+1)p} - 2^{mp} \big) = (1 - e^{-1})\theta 2^{mp}(2^p - 1). \quad (5.5)$$

Putting (5.3), (5.4), and (5.5) together, we see that (5.2) would imply

$$(1 - e^{-1})^{1/p}\theta^{1/p}2^m(2^p - 1)^{1/p} \leq 2^m \Big(1 + \frac{\theta}{1 - 2e^{-2^m}} \Big) + \widetilde{K}p$$

for every $p \geq 2$, $m \geq 2$, and $\theta \in [e^{-2^{m-1}}, 1)$ of the form ne^{-2^m}, $n \in \mathbb{N}$. Take $p = 1/\theta$ and $\theta \sim 1/m$ to get

$$(1 - e^{-1})^\theta \theta^\theta (2^{1/\theta} - 1)^\theta \leq 1 + \frac{\theta}{1 - 2e^{-2^m}} + \frac{\widetilde{K}}{2^m \theta}.$$

Since $\theta \to 0$ and $2^m\theta \to \infty$ as $m \to \infty$ this inequality yields $2 \leq 1$, which is a contradiction. Hence inequality (5.2) cannot hold for all $p \geq 2$ and $n \in \mathbb{N}$.

Acknowledgments

References

R. Adamczak, R. Latała, A. E. Litvak, K. Oleszkiewicz, A. Pajor and N. Tomczak-Jaegermann. A short proof of Paouris' inequality. *Canad. Math. Bull.* **57** (1), 3–8 (2014). MR3150710.

N. Feldheim, A. Marsiglietti, P. Nayar and J. Wang. A note on the convex infimum convolution inequality. *Bernoulli* **24** (1), 257–270 (2018). MR3706756.

N. Gozlan, C. Roberto, P.-M. Samson, Y. Shu and P. Tetali. Characterization of a class of weak transport-entropy inequalities on the line (2017+). To appear in *Ann. Inst. Henri Poincaré Probab. Stat.*

N. Gozlan, C. Roberto, P.-M. Samson and P. Tetali. Kantorovich duality for general transport costs and applications. *J. Funct. Anal.* **273** (11), 3327–3405 (2017). MR3706606.

O. Guédon, P. Nayar and T. Tkocz. Concentration inequalities and geometry of convex bodies. In *Analytical and probabilistic methods in the geometry of convex bodies*, volume 2 of *IMPAN Lect. Notes*, pages 9–86. Polish Acad. Sci. Inst. Math., Warsaw (2014). MR3329056.

R. Latała. On some problems concerning log-concave random vectors. In *Convexity and Concentration*, volume 161 of *IMA Vol. Math. Appl.*, pages 525–539. Springer, New York (2017). DOI: 10.1007/978-1-4939-7005-6_16.

R. Latała and M. Strzelecka. Weak and strong moments of ℓ_r-norms of log-concave vectors. *Proc. Amer. Math. Soc.* **144** (8), 3597–3608 (2016). MR3503729.

R. Latała and M. Strzelecka. Comparison of weak and strong moments for vectors with independent coordinates (2017+). To appear in *Mathematika*.

R. Latała and T. Tkocz. A note on suprema of canonical processes based on random variables with regular moments. *Electron. J. Probab.* **20**, no. 36, 17 (2015). MR3335827.

R. Latała and J. O. Wojtaszczyk. On the infimum convolution inequality. *Studia Math.* **189** (2), 147–187 (2008). MR2449135.

B. Maurey. Some deviation inequalities. *Geom. Funct. Anal.* **1** (2), 188–197 (1991). MR1097258.

K. Oleszkiewicz. Precise moment and tail bounds for Rademacher sums in terms of weak parameters. *Israel J. Math.* **203** (1), 429–443 (2014). MR3273447.

G. Paouris. Concentration of mass on convex bodies. *Geom. Funct. Anal.* **16** (5), 1021–1049 (2006). MR2276533.

P.-M. Samson. Concentration inequalities for convex functions on product spaces. In *Stochastic inequalities and applications*, volume 56 of *Progr. Probab.*, pages 33–52. Birkhäuser, Basel (2003). MR2073425.

P.-M. Samson. Infimum-convolution description of concentration properties of product probability measures, with applications. *Ann. Inst. H. Poincaré Probab. Statist.* **43** (3), 321–338 (2007). MR2319700.

Y. Shu and M. Strzelecki. A characterization of a class of convex log-Sobolev inequalities on the real line (2017+). To appear in *Ann. Inst. Henri Poincaré Probab. Stat.*