On measure concentration of vector valued maps

Michel Ledoux and Krzysztof Oleszkiewicz *

Abstract

In this work, we study concentration properties for vector valued maps. In particular, we describe inequalities which capture the exact dimensional behavior of Lipschitz maps with values in $\mathbb{R}^k$. To this task, we study in particular a domination principle for projections which might be of independent interest. We further compare our conclusions to earlier results by Pinelis in the Gaussian case, and discuss extensions to the infinite dimensional setting.

Keywords: concentration of measure, vector valued map, moment comparison, Gaussian measure.
2000 MSC: 60E15

Notation: In what follows, whenever we deal with $\mathbb{R}^k$, we endow it with the standard Euclidean structure with scalar product $\cdot$ and norm $\|\cdot\|$. By γ_n, we denote the standard $\mathcal{N}(0, \mathrm{Id}_n)$ Gaussian measure on $\mathbb{R}^n$ with density $d\gamma_n/dx = (2\pi)^{-n/2}e^{-\|x\|^2/2}$. Let $g, g_1, g_2 \dots$ be independent real $\mathcal{N}(0,1)$ random variables, so that $G_n = (g_1, \dots, g_n)$ is an $\mathbb{R}^n$-valued normal random vector with distribution γ_n. For $t \in \mathbb{R}$, let $T(t) = \gamma_1([t, \infty)) = \mathbb{P}(g \geq t)$. Obviously, $T(t) = 1 - \Phi(t)$, where Φ is the standard normal distribution function but using the function T will be more convenient in our computations. Let θ be a random vector uniformly distributed on the unit sphere $S^{k-1} \subseteq \mathbb{R}^k$, independent of $g, g_1, g_2 \dots$. For the sake of brevity, we denote throughout this work by C, C_1, $C_2 \dots$ different positive universal constants (i.e. numerical constants which do not depend on n, k or any other parameter). With little effort some more explicit numerical bounds can be deduced from the proofs.

1 Introduction

In the recent work [5], Gromov considers and analyses the question of isoperimetry of waists and measure concentration of maps. As a typical result, he shows that

*Research partially supported by the Polish KBN Grant 1 PO3A 012 29

whenever $f : \mathbb{R}^n \to \mathbb{R}^k$ is a continuous map, there exists $z \in \mathbb{R}^k$ such that for every $h > 0$,

$$\gamma_n\big(\big(f^{-1}(z)\big)_h\big) \geq \gamma_k\big(B(0,h)\big) \tag{1}$$

where $B(x,h)$ is the ball with center x and radius $h > 0$ in $\mathbb{R}^k$. When $k = 1$, this result follows from the Gaussian isoperimetric inequality with $z = m_f$ a median of f for γ_n. Similar conclusions hold for more general strictly log-concave measures and on the sphere [5].

Although this result is perhaps more of topological nature, it also has consequences to measure concentration. Namely, whenever $f : \mathbb{R}^n \to \mathbb{R}^k$ is 1-Lipschitz,

$$\big(f^{-1}(z)\big)_h \subset f^{-1}\big(B(z,h)\big).$$

In particular, inequality (1) provides an upper bound on the measure of the set $\{\|f - z\| \geq h\}$, namely

$$\gamma_n\big(\|f - z\| \geq h\big) \leq \gamma_k\big(x; \|x\| \geq h\big). \tag{2}$$

When $k = 1$, this amounts to the classical Gaussian control of the measure of the set $\{|f - m_f| \geq h\}$. In particular, (2) may be seen as part of the concentration of measure phenomenon. The aim of this note is actually to apply the general theory of measure concentration (for functions) to concentration of vector valued maps in the spirit of (2). We will deal with quantitative estimates up to numerical constants, as is usual for measure concentration. As in the scalar case, z will always be identified to a median or mean value of the Lipschitz function.

As a result, we first observe that whenever (X, d, μ) is a metric measure space with a Gaussian decay of the concentration function, then for any 1-Lipschitz function $f : X \to \mathbb{R}^k$ with mean zero,

$$\mu\big(\|f\| \geq r\big) \leq C_1\, \gamma_k\big(x; \|x\| \geq r/C_2\big)$$

for any $r \geq 0$ where $C_1, C_2 > 0$ are independent of k. The spirit of these concentration results is that they capture the exact dimensional behavior of Lipschitz maps with values in $\mathbb{R}^k$ (the various bounds are clearly sharp on linear maps). The approach relies on simple moment comparisons. We next try to reach sharper inequalities, in particular with $C_2 = 1$, and develop to this task a general domination principle to transfer concentration inequalities for (one-dimensional) projections to vector valued maps. We then compare our conclusions with earlier work by Pinelis [10] in the Gaussian case. We also discuss, following [10], comparison inequalities for maps with values in finite and infinite dimensional normed spaces based on an inequality put forward by Pisier [11], and describe general concentration results for maps on a Gaussian space. We conclude this investigation with several open questions and conjectures.

2 A general statement

We first recall some basic notions of measure concentration (cf. [8]). Let (X, d, μ) be a metric measure space in the sense of [4]. That is, (X, d) is a metric space and μ a probability measure on the Borel sets of X. The concentration function of (X, d, μ) is defined as

$$\alpha(r) = \alpha_{(X,d,\mu)}(r) = \sup\left\{1 - \mu(A_r); A \subset X, \mu(A) \geq \tfrac{1}{2}\right\}, \quad r > 0,$$

where $A_r = \{x \in X, d(x, A) < r\}$. The concentration function implies the following property of Lipschitz functions: whenever $f : X \to \mathbb{R}$ is 1-Lipschitz, and m_f is a median of f for μ, then, for every $r > 0$,

$$\mu\big(|f - m_f| \geq r\big) \leq 2\alpha(r).$$

Recall also that (X, d, μ) has Gaussian concentration whenever there are constants $\kappa \geq 1$ and $\sigma > 0$ such that

$$\alpha(r) \leq \kappa\, e^{-r^2/2\sigma^2}, \quad r > 0. \tag{3}$$

Typical examples that share Gaussian concentration are the standard Gaussian measures γ_n on $\mathbb{R}^n$ (with $\kappa = \sigma = 1$, independent of the dimension). While σ^2 may be interpreted as the observable diameter of (X, d, μ) (cf. [4], [8]), the constant κ is assumed for simplicity to be larger than or equal to 1.

A first general concentration result for vector valued maps is the following simple statement that relies on moment comparison.

Theorem 1. *Let (X, d, μ) be a metric measure space with Gaussian concentration (3). Then, for every 1-Lipschitz function $f : X \to \mathbb{R}^k$ with mean zero with respect to μ, and every $r \geq 0$,*

$$\mu\big(\|f\| \geq r\big) \leq C\,\kappa\,\gamma_k\big(x; \|x\| \geq r/C\sigma\big)$$

where $C > 0$ is numerical.

Proof. Under the Gaussian concentration hypothesis, whenever $\varphi : X \to \mathbb{R}$ is 1-Lipschitz with $\int \varphi d\mu = 0$, then

$$\mu\big(|\varphi| \geq r\big) \leq C_1\,\kappa\, e^{-r^2/2\sigma^2 C_1}, \quad r \geq 0,$$

for some universal $C_1 > 0$ (cf. [8], Proposition 1.8). Hence, for every $p \geq 1$,

$$\int |\varphi|^p d\mu = \int_0^\infty \mu\big(|\varphi| \geq r\big) d(r^p) \leq C_1\,\kappa \int_0^\infty e^{-r^2/2\sigma^2 C_1} d(r^p)$$

so that

$$\int |\varphi|^p d\mu \le 2\,\kappa\, p\, C_1^{\frac{p}{2}+1}\sigma^p M_{p-1}$$

where $M_q = \int_{\mathbb{R}} |x|^q d\gamma_1(x) = 2^{q/2}\pi^{-1/2}\Gamma(\frac{q+1}{2})$, $q \ge 0$.

Now, let $f : X \to \mathbb{R}^k$ be 1-Lipschitz with mean zero. Then, for every $y \in \mathbb{R}^k$, $y \cdot f : X \to \mathbb{R}$ is $\|y\|$-Lipschitz with mean zero. Hence, by the preceding,

$$\int |y\cdot f|^p d\mu \le 2\,\kappa\, p\, C_1^{\frac{p}{2}+1}\sigma^p M_{p-1}\|y\|^p.$$

Therefore, for any $p \ge 1$,

$$\begin{aligned}\int \|f\|^p d\mu &= M_p^{-1}\iint |y\cdot f(x)|^p d\mu(x) d\gamma_k(y)\\ &\le 2\,\kappa\, p\, C_1^{\frac{p}{2}+1}\sigma^p M_{p-1}M_p^{-1}\int \|y\|^p d\gamma_k(y).\end{aligned}$$

Easy calculation yields

$$\int \left\|\frac{f}{2\sigma\sqrt{C_1}}\right\|^p d\mu \le C_2\kappa \int \|y\|^p d\gamma_k(y)$$

where $C_2 > 1$ is some numerical constant. We are now left with the following lemma that we learned from Pinelis and which we formulate with probabilistic notation.

Lemma 1. *Let $U \ge 0$ be a random variable such that for any $p \ge 1$,*

$$\mathbb{E}(U^p) \le B\,\mathbb{E}\big(\|G_k\|^p\big)$$

where $B \ge 1$. Then, for any $r \ge 0$,

$$\mathbb{P}(U \ge r) \le CB\,\mathbb{P}\big(\|G_k\| \ge r/C\big)$$

for some numerical $C > 0$.

Proof. We may and do assume that $k \ge 2$. Let $a \in (0, 1/2)$ denote a universal constant, to be specified later. When $r \le \frac{1}{a}\sqrt{\frac{k}{2}}$, then

$$\mathbb{P}\big(\|G_k\| \ge ar\big) \ge \mathbb{P}\big(\|G_k\|^2 \ge \tfrac{k}{2}\big) \to 1$$

as $k \to \infty$ by the Law of Large Numbers. Hence the lemma holds in this case provided $C > 0$ is large enough.

Let now $r \geq \frac{1}{a}\sqrt{\frac{k}{2}}$. From the hypothesis, for any $p \geq 1$,

$$\mathbb{P}(U \geq r) \leq Br^{-2p}\,\mathbb{E}\big(\|G_k\|^{2p}\big) = B\Big(\frac{2}{r^2}\Big)^p \frac{\Gamma(p+\frac{k}{2})}{\Gamma(\frac{k}{2})}\,.$$

Choose then $p \geq 1$ such that $p+\frac{k}{2} = \frac{r^2}{2}$. It follows that, for some numerical constant $C_3 > 0$,

$$\mathbb{P}(U \geq r) \leq B\,\Gamma\big(\tfrac{k}{2}\big)^{-1}(C_3 r)^{k-1}e^{-r^2/2} \leq C_3 B\Gamma\big(\tfrac{k}{2}\big)^{-1}(C_3 r)^{k-2}e^{-r^2/4},$$

where we have used Stirling's formula. Now, integrating by parts (see the proof of Theorem 2 below), for every $k \geq 2$ and $r \geq 0$,

$$\mathbb{P}\big(\|G_k\| \geq ar\big) \geq \Gamma\big(\tfrac{k}{2}\big)^{-1}\Big(\frac{ar}{2}\Big)^{k-2}e^{-a^2r^2/2} \geq \Gamma\big(\tfrac{k}{2}\big)^{-1}\Big(\frac{ar}{2}\Big)^{k-2}e^{-r^2/8}.$$

Choose $a \in (0,1/2)$ small enough to have $\exp(\frac{1}{16a^2}) \geq \frac{2C_3}{a} \geq 1$. Then

$$e^{r^2/4}e^{-r^2/8} = e^{r^2/8} \geq \exp\Big(\frac{k}{16a^2}\Big) \geq \Big(\frac{2C_3}{a}\Big)^k \geq \Big(\frac{2C_3}{a}\Big)^{k-2}$$

and therefore $\mathbb{P}(U \geq r) \leq C_3 B\mathbb{P}\big(\|G_k\| \geq ar\big)$. It is then easily seen that the lemma holds for some well chosen C. $\square$

3 A domination principle

In this section, we develop a domination principle that will prove more precise than the general statement of the preceding section. Starting from a sharp Gaussian concentration inequality along linear functionals, the tail of vector valued maps in $\mathbb{R}^k$ will be controlled by the norm of the Gaussian vector in $\mathbb{R}^k$, with only a dimensional factor in front of the probability. We will need several lemmas. All of them are quite standard but we present their proofs for the sake of completeness.

Lemma 2. *For every $s > 0$, $T(s) \leq (2\pi)^{-1/2}s^{-1}e^{-s^2/2}$. Moreover,*

$$\lim_{s\to\infty} sT(s)e^{s^2/2} = (2\pi)^{-1/2}. \tag{4}$$

Proof. Indeed,

$$(2\pi)^{1/2}sT(s) = \int_s^\infty se^{-x^2/2}\,dx \leq \int_s^\infty xe^{-x^2/2}\,dx = e^{-s^2/2}.$$

The de l'Hospital rule easily yields that $\lim_{s\to\infty}\frac{T(s)}{s^{-1}e^{-s^2/2}} = (2\pi)^{-1/2}$. $\square$

Lemma 3. *There exists a constant $C_1 > 0$ such that for every $k \geq 2$ and all $\alpha \in (0,1)$,*

$$\mathbb{P}(\theta_1 \geq \alpha) \geq C_1 k^{-1/2}(1-\alpha^2)^{\frac{k-1}{2}}$$

and, for all $\alpha \in (k^{-1/2},\, 2^{-1/2})$,

$$\mathbb{P}(\theta_1 \geq \alpha) \geq C_1 k^{-1/2}\alpha^{-1}(1-\alpha^2)^{\frac{k-1}{2}}$$

where θ_1 denotes the first coordinate of an $\mathbb{R}^k$-valued random vector θ which is uniformly distributed on S^{k-1}.

Proof. Recall that the surface measure of the unit sphere $S^{k-1} \subset \mathbb{R}^k$ is given by the formula $\omega_{k-1} = 2\pi^{k/2}/\Gamma(k/2)$. Therefore

$$\begin{aligned}\mathbb{P}(\theta_1 \geq \alpha) &= \omega_{k-1}^{-1}\int_0^{\sqrt{1-\alpha^2}} \omega_{k-2}\, t^{k-2}(1-t^2)^{-1/2}\, dt \\ &= \frac{\Gamma(k/2)}{\Gamma(\frac{k-1}{2})\sqrt{\pi}}\int_0^{\sqrt{1-\alpha^2}} t^{k-2}(1-t^2)^{-1/2}\, dt.\end{aligned}$$

Obviously, for all $\alpha \in (0,1)$,

$$\int_0^{\sqrt{1-\alpha^2}} t^{k-2}(1-t^2)^{-1/2}\, dt \geq \int_0^{\sqrt{1-\alpha^2}} t^{k-2}\, dt = \frac{1}{k-1}\,(1-\alpha^2)^{\frac{k-1}{2}}.$$

We also have, for every $\alpha \in (k^{-1/2},\, 2^{-1/2})$,

$$\begin{aligned}\int_0^{\sqrt{1-\alpha^2}} t^{k-2}(1-t^2)^{-1/2}\, dt &\geq \frac{1}{\sqrt{2}\,\alpha}\int_{\sqrt{1-2\alpha^2}}^{\sqrt{1-\alpha^2}} t^{k-2}\, dt \\ &= \frac{1}{\sqrt{2}\,\alpha(k-1)}\Big((1-\alpha^2)^{\frac{k-1}{2}} - (1-2\alpha^2)^{\frac{k-1}{2}}\Big) \\ &\geq \frac{(1-e^{-1/4})(1-\alpha^2)^{\frac{k-1}{2}}}{\sqrt{2}\,\alpha(k-1)}\end{aligned}$$

since

$$(1-2\alpha^2)^{\frac{k-1}{2}}(1-\alpha^2)^{-\frac{k-1}{2}} \leq (1-\alpha^2)^{\frac{k-1}{2}} \leq \left(1-\frac{1}{k}\right)^{\frac{k-1}{2}} \leq e^{-\frac{k-1}{2k}} \leq e^{-1/4}.$$

To finish the proof observe that

$$\inf_{k\geq 2}\frac{\Gamma(k/2)}{\Gamma(\frac{k-1}{2})\sqrt{k}} > 0$$

by Stirling's formula. $\square$

Lemma 4. *Let ξ be an $\mathbb{R}^k$-valued random vector. Then for any $r > s > 0$,*

$$\mathbb{P}\big(\|\xi\| \geq r\big) \leq \sup_{v \in S^{k-1}} \frac{\mathbb{P}(|\xi \cdot v| \geq s)}{\mathbb{P}(|\theta_1| \geq s/r)}.$$

Proof. Without loss of generality we can assume that a random vector θ uniformly distributed on S^{k-1} is independent of ξ. By the rotation invariance of θ, for any $x \in \mathbb{R}^k$ and $s \geq 0$, $\mathbb{P}(|x \cdot \theta| \geq s) = \mathbb{P}(\|x\| \, |\theta_1| \geq s)$. Hence

$$\begin{aligned}
\sup_{v \in S^{k-1}} \mathbb{P}(|\xi \cdot v| \geq s) \geq \mathbb{P}(|\xi \cdot \theta| \geq s) & \\
&= \mathbb{E}_\xi \mathbb{P}_\theta(|\xi \cdot \theta| \geq s) \\
&= \mathbb{E}_\xi \mathbb{P}_\theta(\|\xi\| \, |\theta_1| \geq s) \\
&\geq \mathbb{P}(|\theta_1| \geq s/r, \|\xi\| \geq r) \\
&= \mathbb{P}(|\theta_1| \geq s/r) \, \mathbb{P}(\|\xi\| \geq r)
\end{aligned}$$

which is the conclusion. □

The next theorem describes the domination principle that allows us to deduce sharp concentration inequalities for vector valued maps from the corresponding bounds on one-dimensional projections with a good care in the constants depending upon the dimension.

Theorem 2. *Let $\kappa \geq 1/\sqrt{k}$. Assume that ξ is an $\mathbb{R}^k$-valued random vector such that for every $v \in S^{k-1}$ and $s \geq 0$, $\mathbb{P}(|\xi \cdot v| \geq s) \leq \kappa \, T(s)$. Then, for every $r \geq 0$,*

$$\mathbb{P}\big(\|\xi\| \geq r\big) \leq C\sqrt{k} \, \kappa \, \mathbb{P}\big(\|G_k\| \geq r\big)$$

where $C > 0$ is some numerical constant.

The result readily applies to probability measures μ on a metric space (X, d) and 1-Lipschitz mean zero maps $f : X \to \mathbb{R}^k$ such that, for any $v \in S^{k-1}$ and all $s \geq 0$,

$$\mu\big(|v \cdot f| \geq s\big) \leq \kappa \, T(s)$$

(if ζ has distribution μ, take $\xi = f(\zeta)$). We then have, for all $r \geq 0$,

$$\mu\big(\|f\| \geq r\big) \leq C \sqrt{k} \, \kappa \, \gamma_k\big(x; \|x\| \geq r\big).$$

The result applies in particular to the standard Gaussian measure γ_n on $X = \mathbb{R}^n$, although in this case the factor $\sqrt{k}$ is not necessary as we will see in the next section. As discussed in the remark below, it is however necessary in general.

Proof. For $k = 1$ the assertion is trivial, so assume $k \geq 2$. For $0 \leq r \leq \sqrt{k}$,

$$\mathbb{P}(\|G_k\| \geq r) \geq \inf_{j\geq 2} \mathbb{P}(\|G_j\| \geq \sqrt{j}) = \inf_{j\geq 2} \mathbb{P}\Big(\frac{g_1^2 + g_2^2 + \cdots + g_j^2 - j}{\sqrt{j}} \geq 0\Big)$$

and the last expression is a positive universal constant by the Central Limit Theorem (for another argument, giving more explicit estimate, see for example [7], Lemma 2). Hence it suffices to prove that for every $r > \sqrt{k}$,

$$\mathbb{P}(\|\xi\| \geq r) \leq C\sqrt{k}\,\kappa\,\mathbb{P}(\|G_k\| \geq r)$$

where $C > 0$ is some universal constant.

Assume $r > \sqrt{k}$ and put $s = (r^2 - (k-1))^{1/2}$ so that $r^2 - s^2 = k-1$. Observe that $\alpha = s/r \in (k^{-1/2}, 1)$. If $r \in (\sqrt{k}, \sqrt{2k-2})$, then we also have $\alpha < 2^{-1/2}$. Therefore Lemmas 4, 3 and 2 yield, for all $r \in (\sqrt{k}, \sqrt{2k-2})$,

$$\begin{aligned}
\mathbb{P}(\|\xi\| \geq r) &\leq \frac{\kappa T(s)}{\mathbb{P}(|\theta_1| \geq s/r)} \\
&\leq \frac{(2\pi)^{-1/2}\kappa s^{-1} e^{-s^2/2}}{2C_1 k^{-1/2}\alpha^{-1}(1-\alpha^2)^{\frac{k-1}{2}}} \\
&= \frac{\sqrt{k}\,\kappa\, r^{k-2} e^{-r^2/2} e^{(r^2-s^2)/2}}{C_1\sqrt{8\pi}\,(r^2-s^2)^{\frac{k-1}{2}}} \\
&= C_2\sqrt{k}\left(\frac{e}{k-1}\right)^{\frac{k-1}{2}} \kappa\, r^{k-2} e^{-r^2/2}
\end{aligned}$$

and, for all $r \geq \sqrt{2k-2}$,

$$\begin{aligned}
\mathbb{P}(\|\xi\| \geq r) &\leq \frac{\kappa T(s)}{\mathbb{P}(|\theta_1| \geq s/r)} \\
&\leq \frac{(2\pi)^{-1/2}\kappa s^{-1} e^{-s^2/2}}{2C_1 k^{-1/2}(1-\alpha^2)^{\frac{k-1}{2}}} \\
&= \frac{\sqrt{k}\,\kappa\,\alpha^{-1} r^{k-2} e^{-r^2/2} e^{(r^2-s^2)/2}}{C_1\sqrt{8\pi}\,(r^2-s^2)^{\frac{k-1}{2}}} \\
&\leq C_3\sqrt{k}\left(\frac{e}{k-1}\right)^{\frac{k-1}{2}} \kappa\, r^{k-2} e^{-r^2/2}
\end{aligned}$$

for some universal $C_2, C_3 > 0$. On the other hand

$$\begin{aligned}
\mathbb{P}(\|G_k\| \geq r) &= (2\pi)^{-k/2} \int_r^\infty \omega_{k-1} t^{k-1} e^{-t^2/2}\, dt \\
&\geq (2\pi)^{-k/2} \omega_{k-1} r^{k-2} \int_r^\infty t e^{-t^2/2}\, dt \\
&= (2\pi)^{-k/2} 2\pi^{k/2}\, \Gamma(k/2)^{-1} r^{k-2} e^{-r^2/2} \\
&= 2^{-\frac{k-2}{2}}\, \Gamma(k/2)^{-1} r^{k-2} e^{-r^2/2} \\
&\geq C_4 \Big(\frac{e}{k-1} \Big)^{\frac{k-1}{2}} r^{k-2} e^{-r^2/2}
\end{aligned}$$

for some universal $C_4 > 0$, by Stirling's formula. This ends the proof of the theorem. □

Remark 1. *In general the factor $\sqrt{k}$ in Theorem 2 is necessary.*

Proof. Fix $k \geq 2$. Choose $r > \sqrt{k}$ such that $p_k(r) = k(e/k)^{k/2} r^{k-2} e^{-r^2/2}$ satisfies $p_k(r) < 1$ and $p_k(r) \leq T(r/2)$. Some large enough r will do because of (4). We will prove that for any $s \in (0, r)$,

$$p_k(r)\, \mathbb{P}(\theta_1 \geq s/r) \leq C_5\, T(s), \tag{5}$$

where $C_5 > 0$ is numerical. Indeed, for $s \in (0, r/2]$ the inequality trivially follows from the fact that $T(s) \geq T(r/2)$ and from the way in which we chose r. If $s \in (r/2, r)$, then $\alpha = s/r \in (1/2, 1)$ so that

$$\begin{aligned}
\mathbb{P}(\theta_1 \geq s/r) &= \frac{\Gamma(k/2)}{\Gamma(\frac{k-1}{2})\sqrt{\pi}} \int_0^{\sqrt{1-\alpha^2}} t^{k-2} (1-t^2)^{-1/2}\, dt \\
&\leq C_6 \sqrt{k}\, \alpha^{-1} \int_0^{\sqrt{1-\alpha^2}} t^{k-2}\, dt \\
&\leq C_7\, k^{-1/2} (1-\alpha^2)^{\frac{k-1}{2}}
\end{aligned}$$

and therefore, by Lemma 2,

$$
\begin{aligned}
\frac{T(s)}{\mathbb{P}(\theta_1 \geq s/r)} &\geq C_8\, \frac{s^{-1}e^{-s^2/2}}{k^{-1/2}(1-\alpha^2)^{\frac{k-1}{2}}} \\
&\geq C_8 \sqrt{k}\, r^{k-2} e^{-r^2/2} \cdot \frac{e^{(r^2-s^2)/2}}{(r^2-s^2)^{\frac{k-1}{2}}} \\
&\geq C_8 \sqrt{k}\, r^{k-2} e^{-r^2/2} \cdot \inf_{u>0} u^{-\frac{k-1}{2}} e^{u/2} \\
&= C_8 \sqrt{k}\, r^{k-2} e^{-r^2/2} \Big(\frac{e}{k-1}\Big)^{\frac{k-1}{2}} \\
&\geq C_9\, p_k(r)
\end{aligned}
$$

where C_6, C_7, C_8, C_9 are some universal positive constants.

Let θ be, as before, uniformly distributed on S^{k-1} and let η be a random variable independent of θ with $\mathbb{P}(\eta = r) = p_k(r)$, $\mathbb{P}(\eta = 0) = 1 - p_k(r)$. Let $\xi = \eta\theta$. We have proved (5), which means that for $s > 0$ and all $v \in S^{k-1}$,

$$\mathbb{P}(|\xi \cdot v| \geq s) \leq 2\, C_5\, T(s).$$

On the other hand $\mathbb{P}(\|\xi\| \geq r) \geq p_k(r)$, whereas $\mathbb{P}(\|G_k\| \geq r) \leq C_{10}\, k^{-1/2} p_k(r)$ where $C_{10} > 0$ is numerical (to see it, modify the end of the proof of Theorem 2). Hence the factor $\sqrt{k}$ in Theorem 2 cannot be avoided in general. $\square$

4 Gaussian concentration results of Pinelis

In this section, we compare and discuss earlier results by Pinelis [10] based on moment comparison which provide improved constants in a Gaussian setting. Pinelis' investigation covers the case of Lipschitz maps with values in both Euclidean space $\mathbb{R}^k$ and arbitrary (finite or infinite dimensional) normed spaces.

A first optimal result in Euclidean space is the following statement from [10]. With respect to Theorem 2, it shows that the dimensional factor $\sqrt{k}$ is not necessary for Gaussian measures. Recall the standard Gaussian measure γ_n on $\mathbb{R}^n$.

Theorem 3. *Let $f : \mathbb{R}^n \to \mathbb{R}^k$ be a 1-Lipschitz function such that $\int f d\gamma_n = 0$. Then, for any convex function $\Psi : \mathbb{R} \to \mathbb{R}$,*

$$\int \Psi\big(\|f\|\big) d\gamma_n \leq \int \Psi\big(\|x\|\big) d\gamma_k(x).$$

In particular, for any $r \geq 0$,

$$\gamma_n\big(\|f\| \geq r\big) \leq e\, \gamma_k\big(x; \|x\| \geq r\big).$$

For the reader's convenience we extract from the Pinelis' paper a direct argument showing that the convex domination implies the tail inequality with factor e (Pinelis traces this argument back to Kemperman and cites the book by Shorack and Wellner [12], pages 797-799). It is well known and quite easy to prove that the random variable $\|G_k\|$ has logarithmically concave tails, i.e. $\gamma_k\big(x; \|x\| \geq t\big) = e^{-w(t)}$ for some convex, increasing function $w : [0, \infty) \to [0, \infty)$. Given $r > 0$ one can find an affine function $t \mapsto a + bt$, with $a \in \mathbb{R}$ and $b > 0$, supporting the function w at point $t = r$, so that $\mathbb{P}(\|G_k\| \geq r) = e^{-a-br}$ and $\mathbb{P}(\|G_k\| \geq t) \leq e^{-a-bt}$ for $t \geq 0$. In particular, by setting $t = 0$ we deduce that $a \leq 0$. Let $c = r - 1/b$. If $c \leq 0$ then $br \leq 1$, so that also $a + br \leq 1$ and therefore

$$e\,\gamma_k\big(x; \|x\| \geq r\big) = e\mathbb{P}(\|G_k\| \geq r) = e^{1-a-br} \geq 1 \geq \gamma_n\big(\|f\| \geq r\big).$$

If $c > 0$ then consider a nondecreasing, convex function $\Psi(t) = (t - c)_+$ and observe that

$$\gamma_n(\|f\| \geq r) = b(r - c)_+ \gamma_n(\|f\| \geq r) = b\,\Psi(r)\gamma_n(\|f\| \geq r).$$

Therefore,

$$\begin{aligned}\gamma_n(\|f\| \geq r) &\leq b \int \Psi\big(\|f\|\big)d\gamma_n \\ &\leq b \int \Psi\big(\|x\|\big)d\gamma_k(x) = b\,\mathbb{E}(\|G_k\| - c)_+.\end{aligned}$$

Now,

$$\begin{aligned}b\,\mathbb{E}(\|G_k\| - c)_+ &= b \int_0^\infty \mathbb{P}\Big((\|G_k\| - c)_+ \geq t\Big)\,dt \\ &= b \int_c^\infty \mathbb{P}(\|G_k\| \geq t)\,dt \\ &\leq b \int_c^\infty e^{-a-bt}\,dt = e^{-a-bc}\end{aligned}$$

and the conclusion follows since $e^{1-a-br} = e\,\gamma_k(x; \|x\| \geq r)$.

Let $d\mu = e^{-V}dx$ on $\mathbb{R}^n$ with $V'' \geq c\,\mathrm{Id}$, $c > 0$. By a theorem of Caffarelli [3], the Brenier map [2] $S : \mathbb{R}^n \to \mathbb{R}^n$ that transports γ_n to μ is Lipschitz with norm $c^{-1/2}$. Theorem 3 thus readily extends to this family of log-concave measures. In particular, if $f : \mathbb{R}^n \to \mathbb{R}$ is 1-Lipschitz and $\int f d\mu = 0$, for any $p \geq 1$,

$$\int \|f\|^{2p} d\mu \leq c^{-p} \int \|x\|^{2p} d\gamma_k(x).$$

It is worthwhile mentioning that a slight improvement of this moment comparison may be obtained by an alternate semigroup proof which we briefly discuss now. For a probability measure μ on $\mathbb{R}^n$, denote by λ_1 its Poincaré constant defined as the largest λ such that for all smooth enough functions $f : \mathbb{R}^n \to \mathbb{R}$ with $\int f d\mu = 0$,

$$\lambda \int f^2 d\mu \leq \int \|\nabla f\|^2 d\mu.$$

Proposition 1. *Let $d\mu = e^{-V}dx$ on $\mathbb{R}^n$ with $V'' \geq c\,\mathrm{Id}$, $c > 0$. Then, for any 1-Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^k$ such that $\int f d\mu = 0$ and any integer $p \geq 1$,*

$$\int \|f\|^{2p} d\mu \leq p! \prod_{i=0}^{p-1} \frac{1}{c\,i + \lambda_1} \int \|x\|^{2p} d\gamma_k(x).$$

It is classical (cf. [8]) that under the assumptions of the theorem, $\lambda_1 \geq c$ (with equality in the Gaussian case). In particular thus,

$$\int \|f\|^{2p} d\mu \leq c^{-p} \int \|x\|^{2p} d\gamma_k(x).$$

Proposition 1 provides a somewhat sharper result than the conjunction of Caffarelli's theorem together with the Gaussian case in Theorem 3 since the inequality $\lambda_1 \geq c$ can be strict.

Proof. Let $(P_t)_{t\geq 0}$ be semigroup generated by the second order differential operator $\Delta - \nabla \cdot \nabla V$. Since $V'' \geq c\,\mathrm{Id}$, it is known (cf. e.g. [8]) that for all smooth enough functions $\varphi : \mathbb{R}^n \to \mathbb{R}$ and all $t \geq 0$,

$$\|\nabla P_t \varphi\|^2 \leq e^{-2ct} P_t\big(\|\nabla\varphi\|^2\big).$$

In particular, if φ is 1-Lipschitz, $\|\nabla P_t\varphi\|^2 \leq e^{-2ct}$.

Given now $\varphi : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz smooth and such that $\int \varphi d\mu = 0$, write, for every $t \geq 0$,

$$\begin{aligned}\int (P_t\varphi)^{2p} d\mu &= -\int_t^\infty \frac{d}{ds}\bigg(\iint (P_s\varphi)^{2p} d\mu\bigg) ds \\ &\leq 2p(2p-1)\int_t^\infty e^{-2cs}\bigg(\iint (P_s\varphi)^{2p-2} d\mu\bigg) ds.\end{aligned}$$

Iterating,

$$\begin{aligned}\int \varphi^{2p} d\mu \leq 2p(2p-1)(2p-2)\cdots 3 \int_0^\infty e^{-2ct_1} \cdots \\ \cdots \int_{t_{p-2}}^\infty e^{-2ct_{p-1}} \int (P_{t_{p-1}}\varphi)^2 d\mu dt_1 \cdots dt_{p-1}.\end{aligned}$$

Now, the Poincaré inequality provides an exponential decay in $L^2(\mu)$ along the semigroup P_t in the form (cf. e.g. [8])

$$\int (P_{t_{p-1}}\varphi)^2 d\mu \le e^{-2\lambda_1 t_{p-1}} \int \varphi^2 d\mu \le \frac{1}{\lambda_1}\, e^{-2\lambda_1 t_{p-1}} .$$

Therefore,

$$\int \varphi^{2p} d\mu \le \frac{(2p)!}{2^p} \prod_{i=0}^{p-1} \frac{1}{c\, i + \lambda_1} .$$

This is the result in the one-dimensional case.

Let now $f = (\varphi_1, \ldots, \varphi_k) : \mathbb{R}^n \to \mathbb{R}^k$. Write

$$\int \|f\|^{2p} d\mu = M_{2p}^{-1} \int\!\!\int \bigg| \sum_{i=1}^{k} y_i \varphi_i(x) \bigg|^{2p} d\mu(x) d\gamma_k(y)$$

where we recall that $M_{2p} = \int_{\mathbb{R}} x^{2p} d\gamma_1$. For every fixed $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$, the map $x \mapsto \sum_{i=1}^k y_i \varphi_i(x)$ is Lipschitz with Lipschitz coefficient less than or equal to $\|y\|$. The conclusion then follows from the preceding since $M_{2p} = \frac{(2p)!}{2^p p!}$. $\square$

We next turn to Lipschitz functions on Gaussian spaces with values in arbitrary vector spaces, and point out several extensions and generalizations. As developed in [10], comparison results are obtained here from a Poincaré type inequality put forward by Pisier [11]. In the following, F denotes a normed vector space.

Theorem 4. *For every convex function $\Psi : F \to \mathbb{R}$ and every (smooth, sufficiently integrable) function $f : \mathbb{R}^n \to F$ with $\int f d\gamma_n = 0$,*

$$\int \Psi(f) d\gamma_n \le \int\!\!\int \Psi\Big(\frac{\pi}{2}\, y \cdot \nabla f(x)\Big) d\gamma_n(x) d\gamma_n(y).$$

The example of $F = \ell^1$ shows that the factor $\frac{\pi}{2}$ in this inequality cannot be improved (cf. [11]). It might be worthwhile briefly recalling the simple proof of Theorem 4. Let G be a random vector with distribution γ_n and G' an independent copy of G. For any $\theta \in \mathbb{R}$, set $G_\theta = G \sin\theta + G' \cos\theta$ and $G'_\theta = G\cos\theta - G'\sin\theta$. Then, for a smooth enough function $f : \mathbb{R}^n \to F$ such that $\int f d\gamma_n = 0$,

$$f(G) - f(G') = \int_0^{\pi/2} \frac{d}{d\theta} f(G_\theta) d\theta = \int_0^{\pi/2} G'_\theta \cdot \nabla f(G_\theta) d\theta.$$

Apply then Ψ and take expectation. On the one hand, by Jensen's inequality (in G'), $\mathbb{E}(\Psi(f(G) - f(G')) \geq \mathbb{E}\Psi(f(G))$ since f has mean zero, and on the other, by Jensen's inequality again but in $d\theta$,

$$\mathbb{E}\Psi\big(f(G)\big) \leq \int_0^{\pi/2} \mathbb{E}\bigg(\Psi\Big(\frac{\pi}{2}\, G'_\theta \cdot \nabla f(G_\theta)\Big)\bigg) \frac{d\theta}{\pi/2}\,.$$

The conclusion follows since for each θ, the couple (G_θ, G'_θ) has the same distribution as (G, G').

Although the extension below is not strictly necessary for the purposes of measure concentration, it might be worthwhile mentioning that Caffarelli's contraction theorem extends Theorem 4 to all strictly log-concave measures on $\mathbb{R}^n$. We leave the details to the reader.

Corollary 1. *Let $d\mu = e^{-V}dx$ on $\mathbb{R}^n$ with $V'' \geq c\,\mathrm{Id}$, $c > 0$. Then, for every convex function $\Psi : F \to R$ and every (smooth, sufficiently integrable) vector valued function $f : \mathbb{R}^n \to F$ with $\int f d\mu = 0$,*

$$\int \Psi(f) d\mu \leq \int\!\!\int \Psi\Big(\frac{\pi}{2\sqrt{c}}\, y \cdot \nabla f(x)\Big) d\mu(x) d\gamma_n(y).$$

Theorem 4 allows us to derive concentration inequalities for functions on Gaussian spaces with values in arbitrary vector spaces that are Lipschitz in an appropriate sense.

The first result concerns maps $f : \mathbb{R}^n \to F$ that are Lipschitz in the usual sense. If $\Psi(x) = \psi(\|x\|)$, $x \in F$, where $\psi : \mathbb{R}_+ \to \mathbb{R}$ is convex and non-decreasing, for any 1-Lipschitz map $f : \mathbb{R}^n \to F$ (with respect to the norm on F) with $\int f d\gamma_n = 0$,

$$\int \psi\big(\|f\|\big) d\gamma_n \leq \int \psi\Big(\frac{\pi}{2}\, \|y\|\Big) d\gamma_n(y).$$

By the comparison theorems of [10] (see the comment following Theorem 3), it follows that

$$\gamma_n\big(\|f\| \geq r\big) \leq e\, \gamma_n\big(x; \|x\| \geq 2r/\pi\big)$$

for every $r \geq 0$.

Let now ν be a centered Gaussian measure on a real separable Banach space F. A map $f : \mathbb{R}^n \to F$ is then said to be 1-Lipschitz with respect to ν if for every $\xi \in F'$, $\langle \xi, f\rangle : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz with coefficient $\big(\int \langle \xi, x\rangle^2 d\nu(x)\big)^{1/2}$. Of course, the choice of $\nu = \gamma_k$ on $F = \mathbb{R}^k$ leads to the usual definition of 1-Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^k$. With the help of Theorem 4, we may thus extend the concentration of maps to Lipschitz functions with respect to a given Gaussian measure ν.

Corollary 2. *Let $d\mu = e^{-V}dx$ on $\mathbb{R}^n$ with $V'' \geq c\,\mathrm{Id}$, $c > 0$. Let furthermore ν be a centered Gaussian measure on a Banach space F. Then, for any function $f : \mathbb{R}^n \to F$ 1-Lipschitz with respect to ν and such that $\int f d\mu = 0$,*

$$\mu\big(\|f\| \geq r\big) \leq K\,\nu\big(x; \|x\| \geq \sqrt{c}\,r/K\big)$$

for every $r \geq 0$, where K is some positive universal constant.

Proof. By standard smoothing arguments (convoluting f with $\mathcal{C}_0^\infty$ approximation of δ_0) we can assume that f is smooth. By Caffarelli's result, it is enough to deal with the Gaussian case $\mu = \gamma_n$ (alternatively, use Corollary 1). By definition of 1-Lipschitz with respect to ν, for any fixed x, and any $\xi \in F'$,

$$\int \langle \xi, y \cdot \nabla f(x)\rangle^2 d\gamma_n(y) \leq \int \langle \xi, y\rangle^2 d\nu(y).$$

This covariance domination implies that ν is a convolution of $(\nabla f(x))_*\gamma_n$ (the image of γ_n under linear transportation by $\nabla f(x)$) with some other centered Gaussian measure. Therefore, by Jensen's inequality, for every convex function $\Psi : F \to \mathbb{R}$ and any $x \in \mathbb{R}^n$,

$$\int \Psi(y \cdot \nabla f(x)) d\gamma_n(y) \leq \int \Psi(y) d\nu(y).$$

Now, by Theorem 4

$$\int \Psi(2f/\pi) d\gamma_n \leq \iint \Psi(y \cdot \nabla f(x)) d\gamma_n(y) d\gamma_n(x) \leq \int \Psi(y) d\nu(y).$$

The comment following Theorem 3 does not apply here since the norm on F may differ from the Euclidean norm induced by ν. We need another argument. Let G be an F-valued Gaussian random vector with distribution ν. Denote by M the median of $\|G\|$ and let $\sigma = \sup_{\xi \in F' : \|\xi\|=1} \big(\mathbb{E}\langle\, \xi, G\,\rangle^2\big)^{1/2}$. The Gaussian isoperimetry implies that for $g \sim \mathcal{N}(0,1)$ there is $\mathbb{E}\exp\big((\|G\| - M)^2/(4\sigma^2)\big) \leq \mathbb{E}e^{g^2/4} = \sqrt{2}$. Let $\Psi(y) = \exp\big(\|y\|^2/(8\sigma^2)\big)$. Since $\|G\|^2 \leq 2M^2 + 2(\|G\| - M)^2$ we have

$$\int \Psi(2f/\pi) d\gamma_n \leq \mathbb{E}\Psi(G) \leq e^{M^2/(4\sigma^2)}\,\mathbb{E}\exp\big((\|G\| - M)^2/(4\sigma^2)\big) \leq \sqrt{2}\,e^{M^2/(4\sigma^2)}.$$

If $r < \pi M$ then obviously $\gamma_n(\|f\| \geq r) \leq 2\mathbb{P}(\|G\| \geq r/\pi)$. If $r \geq \pi M$ then, by Chebyshev's inequality,

$$\gamma_n(\|f\| \geq r) \leq \sqrt{2}\,e^{M^2/(4\sigma^2)} e^{-r^2/(2\pi^2\sigma^2)} \leq \sqrt{2}\,e^{-r^2/(4\pi^2\sigma^2)} \leq \sqrt{2}\,A^{-1} \cdot T\Big(\frac{r}{2\pi\sigma}\Big),$$

where $A = \inf_{s \geq 0} T(s) e^{s^2}$ is a positive universal constant (see Lemma 2). Choose $\xi \in F'$ such that $\|\xi\| = 1$ and $\big(\mathbb{E}\langle \xi, G \rangle^2\big)^{1/2} \geq \sigma/2$. Then

$$\nu\Big(x; \|x\| \geq \frac{r}{4\pi}\Big) = \mathbb{P}\Big(\|G\| \geq \frac{r}{4\pi}\Big) \geq \mathbb{P}\Big(\langle \xi, G \rangle \geq \frac{r}{4\pi}\Big) \geq$$

$$\mathbb{P}\Big(\frac{\sigma}{2} g \geq \frac{r}{4\pi}\Big) = T\Big(\frac{r}{2\pi\sigma}\Big) \geq \frac{A}{\sqrt{2}} \cdot \gamma_n\Big(\|f\| \geq r\Big)$$

and the proof is finished by setting $K = \max(\sqrt{2}/A, 4\pi)$. □

The couple $(\mathbb{R}^n, \gamma_n)$ may be replaced in the above statements by an abstract Wiener space. Lipschitz has then to be understood in the directions of the reproducing kernel Hilbert space.

The preceding results have counterparts on the discrete cube $\{0,1\}^n$. It has been shown by Pisier [11] that for every $f : \{0,1\}^n \to F$ with mean zero with respect to the uniform measure μ on the cube, and every $p \geq 1$,

$$\int \|f\|^p d\mu \leq C^p \int\!\!\int \Big\| \sum_{i=1}^n y_i D_i f(x) \Big\|^p d\mu(x) d\mu(y) \tag{6}$$

where $D_i f(x) = \frac{1}{2}\,[f(x) - f(s_i(x))]$ and $s_i(x)$ is obtained from $x \in \{0,1\}^n$ by changing the i-th coordinate. In general, the constant $C = 2e \log n$ and may not be improved for arbitrary spaces F. It is however independent of n in the case of $F = \mathbb{R}^k$ with its classical Euclidean structure (see [13]).

By the comparison between Rademacher and Gaussian averages, we may increase the right-hand side of (6) replacing $d\mu(y)$ by $d\gamma_n(y)$ (at the expense of a multiplicative factor). Now, the same reasoning as for Theorem 1 may be applied. If $f : \{0,1\}^n \to \mathbb{R}^k$ is such that $\int f d\mu = 0$ and, for every $\xi \in \mathbb{R}^k$,

$$\sum_{i=1}^n (\xi \cdot D_i f(x))^2 \leq \|\xi\|^2$$

uniformly in x, then

$$\int \|f\|^p d\mu \leq C^p \int \|y\|^p d\gamma_k(y).$$

Together with Lemma 1, we conclude that

$$\mu\big(\|f\| \geq r\big) \leq C\gamma_k\big(x; \|x\| \geq r/C\big)$$

for every $r \geq 0$.

5 Concluding comments and questions

In what follows $\sup_f$ denotes the supremum over all 1-Lipschitz functions $f : \mathbb{R}^n \to \mathbb{R}^k$. In view of Gromov's result [5] described in Introduction it is natural to ask what is the optimal rate of concentration of f around some **value** of f - namely, what is the asymptotics of $\sup_n \sup_f \inf_{x\in\mathbb{R}^n} \|f(x) - \mathbb{E}f(G_n)\|$ as $k \to \infty$ and, for fixed k, what is the asymptotics (as $t \to \infty$) of $\sup_n \sup_f \inf_{x\in\mathbb{R}^n} \mathbb{P}(\|f(x) - f(G_n)\| \geq t)$?

Dealing with concentration properties of $(X \times X, \mu \otimes \mu)$ rather than (X, μ) (see e.g. Barthe's isoperimetric inequality for $S^{n-1} \times S^{n-1}$ [1], Proposition 11) can lead to the concentration results of slightly different form: instead of estimating from above $\mathbb{P}(\|f(G_n) - \mathbb{E}f(G_n)\| \geq t)$ one can bound $\mathbb{P}(\|f(G_n) - f(G_n')\| \geq t)$ where G_n' is an independent copy of G_n. Another possible direction of research is related to the following definition.

Definition 1. *Let F be a separable real Banach space and let X and Y be F-valued random vectors. We will say that X is weakly dominated by Y if for every bounded linear functional $\varphi \in F'$ and all $t > 0$,*

$$\mathbb{P}(|\langle\varphi, X\rangle| \geq t) \leq \mathbb{P}(|\langle\varphi, Y\rangle| \geq t).$$

It is of interest under what additional assumptions about distributions of X and Y does weak domination imply $\mathbb{E}\|X\| \leq C\,\mathbb{E}\|Y\|$, or even

$$\mathbb{P}(\|X\| \geq t) \leq C\,\mathbb{P}(\|Y\| \geq t/C) \quad \text{for all } t > 0.$$

Note that the latter inequality easily implies $\mathbb{E}\|X\| \leq C^2\mathbb{E}\|Y\|$.

It is not very difficult to see that this is always so if both X and Y are centered Gaussian vectors (see [9] or [6], Chapter 5.5 - we have used a similar approach in the proof of Corollary 2). Some results of the present paper, especially Theorem 2, refer to the case when F is equal to $\mathbb{R}^k$ equipped with the standard Euclidean structure. Recently Kwapień and Latała (private communication) obtained several interesting results concerning the case when we make some additional assumptions about Y only. Also, Latała proved that the following natural conjecture would be a corollary to the so-called Bernoulli Conjecture of Talagrand (which is still open, see [14]):

Conjecture 1. *Let $r_1, r_2, \ldots$ be i.i.d. sequence of symmetric ± 1 random variables. There exists a universal constant $C > 0$ such that for any separable real Banach space F and every choice of vectors $v_1, w_1, v_2, w_2, \ldots, v_n, w_n \in F$ such that $X = \sum_{j=1}^n r_j v_j$ is weakly dominated by $Y = \sum_{j=1}^n r_j w_j$, there is also*

$$\mathbb{P}(\|X\| \geq t) \leq C\,\mathbb{P}(\|Y\| \geq t/C) \quad \textit{for all } t > 0.$$

Below we will show an example of $\mathbb{R}^k$-valued random vectors X and Y, both rotation invariant with respect to the standard Euclidean structure, indicating that even under such additional assumptions the weak domination cannot in general imply that $\mathbb{P}(\|X\| \geq t) \leq C\,\mathbb{P}(\|Y\| \geq t/C)$ for all $t > 0$.

Recall that T is a continuous and strictly decreasing function. Fix $C > 1$. Choose $x_C > 0$ so great that $2CT(x_C) \leq 1/4$. Then choose $\beta \in (0, \frac{1}{2C})$ so small that

$$\frac{2C\beta}{1-\beta} \leq \inf_{x \in [0, x_C]} \frac{T(2Cx)}{T(x)}.$$

The way in which we chose x_C implies that, for all $x \geq x_C$

$$2C\beta T(x) - (1-\beta)T(2Cx) \leq \beta/4. \tag{7}$$

Now we will choose $b \in (0,1)$ so little that for all $x > 0$,

$$2C\beta T(x) \leq (1-\beta)T(2Cx) + \beta T(bx). \tag{8}$$

From the way in which we chose β we deduce that (8) is satisfied whenever $x \in [0, x_C]$ and $b > 0$. Hence it suffices to choose the proper b for $x \geq x_C$. One can easily check that (4) implies $T^{-1}(s)/\sqrt{2\ln(1/s)} \longrightarrow 1$ as $s \longrightarrow 0^+$ and therefore

$$\lim_{x \to \infty} T^{-1}\Big(2CT(x) - (1-\beta)T(2Cx)/\beta\Big)/x = 1,$$

so that there exists $y > x_C$ such that for every $x \geq y$,

$$2C\beta T(x) \leq (1-\beta)T(2Cx) + \beta T(x/2).$$

On the other hand by (7), we have for every $x \in [x_C, y]$,

$$2C\beta T(x) \leq (1-\beta)T(2Cx) + \beta/4 \leq (1-\beta)T(2Cx) + \beta T\Big(T^{-1}(1/4)x/y\Big).$$

Therefore $b = \min\Big(1/2, T^{-1}(1/4)/y\Big)$ satisfies our requirements. Recall that $\mathcal{L}(G_k) = \mathcal{N}(0, Id_k)$ and let $\|\cdot\|$ denote the standard Euclidean norm on $\mathbb{R}^k$, as usually. Consider random vectors (Gaussian mixtures) X and Y with distributions given by $\mathcal{L}(X) = (1-2C\beta)\delta_0 + 2C\beta\mathcal{L}(G_k)$ and $\mathcal{L}(Y) = (1-\beta)\mathcal{L}\Big((2C)^{-1}G_k\Big) + \beta\mathcal{L}(G_k/b)$. The inequality (8) means that X is weakly dominated by Y. By the Law of Large Numbers $\lim_{k\to\infty} \mathbb{P}(\|G_k\| \geq w\sqrt{k})$ is equal to 0 if $w > 1$ and it is equal to 1 if $w \in (0,1)$, so that

$$\mathbb{P}(\|X\| \geq 0.9\sqrt{k}) = 2C\beta\,\mathbb{P}(\|G_k\| \geq 0.9\sqrt{k}) \stackrel{k\to\infty}{\longrightarrow} 2C\beta,$$

whereas

$$C\,\mathbb{P}(\|Y\| \geq 0.9\sqrt{k}/C) \leq C(1-\beta)\,\mathbb{P}(\|G_k\| \geq 1.8\sqrt{k}) + C\beta \overset{k\to\infty}{\longrightarrow} C\beta.$$

Hence, in general, the weak domination cannot yield the inequality

$$\mathbb{P}(\|X\| \geq t) \leq C\,\mathbb{P}(\|Y\| \geq t/C) \quad \text{for all } t > 0.$$

for any universal C. However, one can quite easily prove such inequality with C depending on k.

On the other hand, note that for this example (and for any pair of rotation invariant $\mathbb{R}^k$-valued X and Y such that X is weakly dominated by Y), for every $p > 0$,

$$\mathbb{E}\|X\|^p \leq \mathbb{E}\|Y\|^p$$

for **any** norm $\|\cdot\|$ on $\mathbb{R}^k$ (not necessarily Euclidean).

Indeed, because of the rotation invariance we have $\mathbb{E}\|X\|^p = \mathbb{E}\|X\|_\circ^p$ and $\mathbb{E}\|Y\|^p = \mathbb{E}\|Y\|_\circ^p$, where $\|v\|_\circ = \Big(\int_{O(k)} \|U(v)\|^p \, d\sigma_H(U)\Big)^{1/p}$ (the integral is taken with respect to the normalized Haar measure σ_H) for $v \in \mathbb{R}^k$. The norm $\|\cdot\|_\circ$ is rotation invariant and our assertion follows from the fact that $\|\cdot\|_\circ$ must be proportional to another rotation invariant norm $\|v\|_{\circ\circ} := (\mathbb{E}|\theta\cdot v|^p)^{1/p}$. Obviously, $\mathbb{E}|\theta\cdot X|^p \leq \mathbb{E}|\theta\cdot Y|^p$.

Acknowledgements We thank I. Pinelis for his interest in this work and his help in the proof of Lemma 1. We also thank the anonymous referee for useful comments. Part of the work was done when the second named author visited Institut de Mathématiques at Université Paul Sabatier in Toulouse. It is a pleasure to acknowledge its kind hospitality.

References

[1] F. Barthe, *Extremal properties of central half-spaces for product measures*, J. Funct. Anal. **182** (2001), 81-107

[2] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), 375–417

[3] L. Caffarelli, *Monotonicity properties of optimal transportation and the FKG and related inequalities*, Comm. Math. Phys. **214** (2000), 547–563

[4] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Birkhäuser, 1998

[5] M. Gromov, *Isoperimetry of waists and concentration of maps*, Geom. Funct. Anal. **13** (2003), 178–215

[6] S. Kwapień and W. A. Woyczyński, *Random series and stochastic integrals: single and multiple*, Probability and its Applications, Birkhäuser, 1992

[7] R. Latała and K. Oleszkiewicz, *Small ball probability estimates in terms of width*, Studia Math. **169** (2005), 305-314

[8] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs **89** (2001), American Mathematical Society

[9] M. Ledoux and M. Talagrand, *Probability in Banach spaces. Isoperimetry and processes* Ergebnisse der Mathematik und ihrer Grenzgebiete (3), **23** (1991), Springer-Verlag

[10] I. Pinelis, *Optimal tail comparison based on comparison of moments*, in *High Dimensional Probability, Oberwolfach (Germany) 1996*, Progress in Probability **43** (1998), 297–314, Birkhäuser

[11] G. Pisier, *Probabilistic methods in the geometry of Banach spaces* in Probability and Analysis, Varenna (Italy) 1985, Lecture Notes in Math. **1206** (1986), 167–241, Springer

[12] G. R. Shorack and J. A. Wellner, *Empirical Processes with Applications to Statistics,* Wiley, New York, 1986

[13] M. Talagrand, *Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem*, Geom. Funct. Anal. 3 (1993), 295-314

[14] M. Talagrand, *Chaining method ???* recent

Institut de Mathématiques, Université Paul-Sabatier,
31062 Toulouse, France
e-mail: ledoux@math.ups-tlse.fr

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