SJBS (M), (2016b), 34: 31-66 ISSN: 7615-1858

Verification of Dimension Free Concentration Interms of Series of Transportation Inequalities

Shawgy Hussein(1) and Mohammed Ali(2)

Abstract:

We show a generalization of a probability measure on R^d concentrates independently of the dimension like a Gaussian measure if and only if it verifies Talagrand's T_2 series of transportation-cost inequality. We give a verification of the proof of Otto and Villani result. The equivalent of Poincare' inequality and certain form of dimension free exponential concentration is shown.

Keywords: concentration-cost, logarithmic-Sobolev inequality, Gaussian measure, Poincare inequality, relative entropy, large deviation, empirical measure, Wasserstein metrics.

1.Introduction

In this paper we follow the theory and notation that appear in the work of Nathael Gozlan [39].A probability measure μ on $\mathbf{R}^d$ has the Gaussian dimension free concentration property if there are three non-negative constants a, b and r_0 such that for every integer n, the product measure μ^n verifies the following inequality:

(1) Mathematics Department, College of technology of mathematical sciences/ Alnelain University, Sudan.

(2) Bahri University, College of Science, department of Mathematics, Sudan..

$$\mu^n\left(A+(\varepsilon+r_0)B_2\right)\geq 1-be^{-a\varepsilon^2}\quad \forall\varepsilon\geq 0 \tag{1}$$

for a measurable subset A of $\left(\mathbf{R}^d\right)^n$ with $\mu^n\left(A\right)\geq\frac{1}{2}$, denoting by B_2 the Euclidean unit ball of $\left(\mathbf{R}^d\right)^n$ .If $p\in[1,2]$ the probability measure $d\,\mu_p\left(\sum_{j=1}^{k}x_j\right)=\left(\sum_{j=1}^{k}Z_j\right)_p^{-1}e^{-\left|\sum_{j=1}^{k}x_j\right|^p}d\left(\sum_{j=1}^{k}x_j\right)$ verifies a series of concentration inequalities similar to (1) with $(\varepsilon+r_0)^2$ replaced by $\min\left((\varepsilon+r_0)^p,(\varepsilon+r_0)^2\right)$. The logarithmic-Sobolev inequality is well known to imply (1); this is renowned Herbst argument (see Ledoux [25]). Among functional inequalities yielding concentration estimates let us mention: Poincare' inequalities [21,6], Logarithmic-Sobolev inequalities [6, 10, 18, 9]. Transportation-cost inequalities [30, 35, 4, 33, 7, 32, 5, 22], inf-convolution inequalities [31, 28], Latala-Oleskiewicz inequalities [3, 26, 8, 2] see[39] (see for instance [24], [1], or [36, 37]). This large variety of tools and points of view raise the following natural question: is one of these functional inequalities equivalent to say (1)?

Nathael Gozlan [39] show with a certain generality that Talagrand's transportation-cost inequalities are equivalent to dimension of free concentration of measure. He also show a new proof of a result of Otto and Villani, generalizations of other types of series of concentrations Poincare' inequality with exponential using deviations techniques methods. We present and state with a bit change on [39] series of transportation –cost inequalities first define the series of the optimal quadratic transportation-cost on $\mathrm{P}\left(\mathbf{R}^d\right)$, we define

$$\sum_{j=1}^{k}\mathrm{T}_2\left(v_j,\mu\right)=\inf_{\pi}\int\sum_{j=1}^{k}\left|x_j-y_j\right|_2^2 d\,\pi\left(x_j,y_j\right) \tag{2}$$

where π describes the set $\mathrm{P}\left(\sum_{j=1}^{k} \nu_j, \mu\right)$ of probability measures on $\mathbf{R}^d \times \mathbf{R}^d$ having $\sum_{j=1}^{k} \nu_j$ and μ for marginal distributions. We say that μ verifies the inequality $\mathrm{T}_2(\tilde{c})$, if

$$\sum_{j=1}^{k} \mathrm{T}_2(\nu_j, \mu) \le \tilde{c} \sum_{j=1}^{k} H(\nu_j | \mu), \quad \text{for every } \sum_{j=1}^{k} \nu_j \in \mathrm{P}(\mathbf{R}^d) \qquad (3)$$

where $\sum_{j=1}^{k} H(\nu_j | \mu),$ is the relative entropy of $\sum_{j=1}^{k} \nu_j$ with respect to μ defined by

$$\sum_{j=1}^{k} H(\nu_j | \mu), = \int \log\left(\frac{d\left(\sum_{j=1}^{k} \nu_j\right)}{d\mu}\right) d\left(\sum_{j=1}^{k} \nu_j\right),$$

if $\sum_{j=1}^{k} \nu_j$ is absolutely continuous with respect to μ and $+\infty$ otherwise. The idea of controlling an optimal series of transportation-cost inequalities by the relative entropy to obtain concentration first appeared in Marton's works [29, 30]. The inequality T_2 was then introduced by Talagrand in [35], where it was proved to be fulfilled by Gaussian probability measures in particular, if $\mu = \gamma$ is the standard Gaussian measure on $\mathbf{R}$, then the inequality (3) holds true with the sharp constant $\tilde{c} = 2$ (see[39]) .

Theorem1: Let μ be a probability measure on $\mathbf{R}^d$ and $a > 0$; the following propositions are equivalent:

(i) There are $r_0, b \ge 0$ such that for all n the probability μ^n verifies (1).

(ii) The probability measure μ verifies $\mathrm{T}_2(1/a)$.

We think that this new result confirms the relevance of the large deviations point of view for functional inequalities initiated by Leonard and the author in [19] and pursued in [20] by Guillin, Leonard. Wu and Yiao. Moreover Theorem 1: turns out to be a quite powerful tool. For example, the famous result by Otto and Vilani stating that the logarithmic Sobolev inequality (LSI) implies T_2 inequality (see [33]) is direct consequence of Theorem 1

2. Preliminaries

Let (χ_k, p) be a Polish space and the set of probability measures on χ is denoted by $\mathrm{P}(\chi_k)$. Let μ be a probability measure on χ_k and $(X_i)_i$ an i.i.d sequence of random variables with law μ defined on some probability space (Ω, P). The empirical measure L_n is defined for all integer n by

$$L_n = \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n}\left(\delta_{X_i}\right)_j,$$

where δ_x, stands for the Dirac mass at points x_j.

According to Varadarajan's Theorem [16], with probability 1 the sequence $(L_n)_n$ has a weak converges topology to μ in $\mathrm{P}(\chi_k)$ this means that there is a measurable subset N of Ω with $\mathrm{P}(N) = 0$ such that for all $\omega \notin N$,

$$\left|\int f \, dL_n(\omega) - \int f \, d\mu\right| < \varepsilon \text{ for all } n \in N$$

Continuous function f on χ_k.

Here, we consider the Wasserstein metrics. Let $p \geq 1$ and define

$$\mathrm{P}_p(\chi_k) = \left\{\sum_{j=1}^{k} v_j \in p(\chi_k) : \int \rho\left(\left(\sum_{j=1}^{k} x_j\right)_0, \left(\sum_{j=1}^{k} x_j\right)_1\right)^p d\left(\sum_{j=1}^{k} v_j\left(\sum_{j=1}^{k} x_j\right)\right) < \infty,\right\}$$

for some $\left(\sum_{j=1}^{k} x_j\right)_0 \in \chi_k$.

For all probability measures $\left(\sum_{j=1}^{k} v_j\right)_1, \left(\sum_{j=1}^{k} v_j\right)_2 \in \mathrm{P}_p(\chi_k)$, define

$$\sum_{j=1}^{k} \mathrm{T}_p\left((v_j)_1, (v_j)_2\right) = \inf_{\pi} \int \rho\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right)^p d\pi(x_j, y_j)$$

and

$$\sum_{j=1}^{k} W_p\left((v_j)_1, (v_j)_2\right) = \left(\sum_{j=1}^{k} \mathrm{T}_p\left(\left((v_j)_1, (v_j)_2\right)\right)\right)^{\frac{1}{p}}$$

where π describes the set $\mathrm{P}\left(\left(\sum_{j=1}^{k} v_j\right)_1, \left(\sum_{j=1}^{k} v_j\right)_2\right)$ of couplings of $\left(\sum_{j=1}^{k} v_j\right)_1$ and $\left(\sum_{j=1}^{k} v_j\right)_2$.

According to e.g [37], W_p is a metric on $\mathrm{P}_p(\chi_k)$ and for every sequence μ_n in $\mathrm{P}_p(\chi_k)$, $W_p(\mu_n, \mu) \to 0$, if and only if μ_n converges to μ, for the weak topology and $\int \rho\left(\left(\sum_{j=1}^{k} x_j\right)_0, \sum_{j=1}^{k} x_j\right)^p d\mu_n \to \int \rho\left(\left(\sum_{j=1}^{k} x_j\right)_0, \sum_{j=1}^{k} x_j\right)^p d\mu$, for some and any $\left(\sum_{j=1}^{k} x_j\right)_0 \in \chi_k$.

From these considerations, one can conclude that [39]if $\mu \in \mathrm{P}_p(\chi_k)$ then $W_p(L_n, \mu) \to 0$ with probability one, and in particular, $\mathrm{P}\left(W_p(L_n, \mu) \geq t\right) \to 0$ when $n \to +\infty$, for all $t > 0$. Moreover, supposing that $\mu \in \mathrm{P}_p(\chi_k)$ with $p > 1$, it is easy to check that the sequence $W_p(L_n, \mu)$ is bounded in $\mathrm{L}_p(\Omega, \mathrm{P})$, thus it is uniformly integrable and consequently $\mathrm{E}\left[W_p(L_n, \mu)\right] \to 0$. This is summarized in the following proposition [39]:

Proposition 2: If $\mu \in \mathrm{P}_p(\chi_k)$, then the sequence $W_p(L_n,\mu) \to 0$ almost surely (and thus -in probability) and if $p>1$, then the convergence is in $\mathrm{L}_1 : \mathrm{E}\left[W_p(L_n,\mu)\right] \to 0$.

On the other hand, Sanov's Theorem (see [17]) says that for all good sets $A, \mathrm{P}(L_n \in A)$, behaves like $e^{-nH(A|\mu)}$ when n is large, where $H(A|\mu)$ stands for the infimum of $H(.|\mu)$ on A. So, when A does not contain $\mu, H(A|\mu) > 0$ and this probability tends to 0 exponentially fast. With this in mind, one can expect that $\mathrm{P}\left(W_p(L_n,\mu) > t\right)$ behaves

like $e^{-nH(t)}$ **where** $H(t) = \inf\left\{\sum_{j=1}^{k} H(v_j|\mu) : \sum_{j=1}^{k} v_j \in \sum_{j=1}^{k} W_p(v_j,\mu) > t\right\}$**.**
The following result validates partially this heuristic, stating that $\mathrm{P}\left(W_p(L_n,\mu) > t\right)$ **tends to 0 not faster than** $e^{-nH(t)}$**.**

Theorem 3: If $\mu \in \mathrm{P}_p(\chi_k)$, **then for all t > 0,**

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathrm{P}\left(W_p(L_n,\mu) > t\right) \geq -\inf\left\{\sum_{j=1}^{k} H(v_j|\mu) : \sum_{j=1}^{k} v_j \in \mathrm{P}_p(\chi)\, s.t.\, \sum_{j=1}^{k} W_p(v_j,\mu) > t\right\}$$

Proof: Let $t \geq 0$ and define $A = \left\{\sum_{j=1}^{k} v_j \in \mathrm{P}_p(\chi_k)\, s.t.\, \sum_{j=1}^{k} W_p(v_j,\mu) > t\right\}$. Take $\sum_{j=1}^{n} v_j \in A$ such that $\sum_{j=1}^{k} H(v_j|\mu) < +\infty$. If $(Y_i)_i$ is an $i.i.d$ sequence of law $\sum_{j=1}^{n} v_j$, and $L_n^{\sum_{j=1}^{k} Y_j} = n^{-1} \sum_{j=1}^{k} \sum_{i=1}^{n} \left(\delta_{Y_i}\right)_j$, then L_n^Y converges to $\sum_{j=1}^{n} v_j$ almost surely for the W_p distance and so

$$\left(\sum_{j=1}^{n} v_j\right)^n (L_n \in A) = \mathrm{P}\left(W_p\left(L_n^Y, \mu\right) > t\right) \to \mathrm{P}\left(\sum_{j=1}^{k} W_p\left(v_j, \mu\right) > t\right) = 1,$$

when $n \longrightarrow +\infty$.

Applying [39] to A and $\sum_{j=1}^{k} v_j$ and taking the limit when $n \longrightarrow +\infty$, gives

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathrm{P}\left(W_p\left(L_n, \mu\right) > t\right) \geq -\sum_{j=1}^{k} H\left(v_j \,|\mu\right).$$

Optimizing over $\sum_{j=1}^{k} v_j$ gives the result.

For the sake of completeness, (see[39] and [19]).

3.The Gaussian Case

The product space χ_k^n will be equipped with the following metric:

$$\rho_2^n\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right) = \sum_{j=1}^{k}\left[\sum_{i=1}^{n} \rho\left(x_j^i, y_j^i\right)^2\right]^{\frac{1}{2}}$$

(here $\sum_{j=1}^{k} x_j = \left(\left(\sum_{j=1}^{k} x_j\right)^1, \left(\sum_{j=1}^{k} x_j\right)^2, \ldots, \left(\sum_{j=1}^{k} x_j\right)^n\right)$ with $\left(\sum_{j=1}^{k} x_j\right)^i \in \chi_k^n$ for all i, j).

In the general case, one says that a probability measure μ on (χ_k, ρ) verifies the dimension of the series of free Gaussian concentration property, if there are $r_0, a, b \geq 0$ such that for all n the probability μ^n verifies

$$\mu^n\left(A^{(\varepsilon+r_0)}\right) \geq 1 - be^{-a\varepsilon^2}, \quad \forall \varepsilon \geq 0 \tag{4}$$

for all measurable $A \subset \chi_k^n$ such that $\mu_n(A) \geq \frac{1}{2}$, where $A^{(\varepsilon+r_0)}$ denotes the $(\varepsilon+r_0)$-enlargement of A defined by

$$A^{(\varepsilon+r_0)} = \left\{ \sum_{j=1}^{k} x_j \in \chi_k^n \text{ such that there is } \sum_{j=1}^{k} \bar{x}_j \in A \text{ with } \rho_2^n \left(\sum_{j=1}^{k} x_j , \sum_{j=1}^{k} \bar{x}_j \right) \le (\varepsilon + r_0) \right\}$$

When $\chi_k = \mathbf{R}^d$ is equipped with its Euclidean metric one has $A^{(\varepsilon+r_0)} = A + (\varepsilon + r_0)B_2$ and one recovers the inequality (1).

The proof of the following well known result makes use of the so called Marton's argument.

Proposition 4 If μ verifies $\mathrm{T}_1(\tilde{c})$, then for a measurable subset A of χ_k, such that $\mu(A) \ge \frac{1}{2}$.

$$\mu\left(A^{(\varepsilon+r_0)}\right) \ge 1 - be^{-\tilde{c}^{-1}\varepsilon^2}, \quad \forall \varepsilon \ge 0$$

where $r_0 = \sqrt{\tilde{c}\log(2)}$.

Proof: Consider a subset A of χ_k, and define $d\mu_A = \mathrm{I}_A d\mu\left(\sum_{j=1}^{k} x_j\right) / \mu(A)$. Let $B = \chi_k \setminus A^{(\varepsilon+r_0)}$ and define μ_B accordingly. Since the distance between two points of A and B is always more than $(\varepsilon + r_0)$, one has $W_1(\mu_A, \mu_B) \ge (\varepsilon + r_0)$. The triangle inequality and the transportation-cost inequality $\mathrm{T}_1(\tilde{c})$ yield

$$\begin{aligned}(\varepsilon + r_0) &\le W_1(\mu_A, \mu_B) \le W_1(\mu_A, \mu) + W_1(\mu_B, \mu) \\ &\le \sqrt{\tilde{c}H(\mu_A|\mu)} + \sqrt{\tilde{c}H(\mu_B|\mu)} \\ &= \sqrt{\tilde{c}\log(1/\mu(A))} + \sqrt{\tilde{c}\log(1/\mu(B))}\end{aligned}$$

Rearranging terms gives the result.

Theorem 5: Let $\mu \in \mathrm{P}_2(\chi_k)$ and a > 0; the following Propositions are equivalent:

(i) There are $r_0, b \ge 0$ such that for all n the probability μ^n verifies (4),

(ii) The probability μ verifies $T_2(1/a)$.

Let us recall the definition of the series of the T_1 transportation-cost inequality. We say that a probability measure μ on χ_k verifies $T_2(1/a)$, if

$$\sum_{j=1}^{k} W_1(v_j,\mu) \le \sqrt{\tilde{c}\sum_{j=1}^{k} H(v_j|\mu)},\quad \text{for every } \sum_{j=1}^{k} v_j \in P(\chi_k)$$

Proof: Let us show that (ii) implies (i). The main point is that T_2 tensorizes ; this means that if μ verifies $T_2(1/a)$ then μ^n verifies $T_2(1/a)$ on the space χ_k^n equipped with ρ_2^n (see[19]) .Jensen's inequality implies that $W_1^{\,2} \le T_2$ and consequently μ^n verifies $T_1(1/a)$ (on χ_k^n equipped with ρ_2^n) for all n. Applying Proposition 4 to μ^n gives (1) with $r_0 = \sqrt{\log(2)/a},\quad b=1$ and a.

Let us show that (i) implies (ii). For every integer n, and $\sum_{j=1}^{k} x_j \in \chi_k^n$, define $L_n^{\left(\sum_{j=1}^{k} x_j\right)} = n^{-1}\sum_{j=1}^{k}\sum_{i=1}^{n}\delta_{\left(\sum_{j=1}^{k} x\right)^i}$. The map $\sum_{j=1}^{k} x_j \mapsto W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)},\mu\right)$ is $\frac{1}{\sqrt{n}} - Lipschitz$.

Indeed, if

$$\sum_{j=1}^{k} x_j = \left(\left(\sum_{j=1}^{k} x_j\right)^1, \left(\sum_{j=1}^{k} x_j\right)^2, \dots, \left(\sum_{j=1}^{k} x_j\right)^n\right)$$ and

$$\sum_{j=1}^{k} y_j = \left(\left(\sum_{j=1}^{k} y_j\right)^1, \left(\sum_{j=1}^{k} y_j\right)^2, \dots, \left(\sum_{j=1}^{k} y_j\right)^n\right)$$

are in χ_k^n, then thanks to the triangle inequality,

$$\left| W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)},\mu\right) - W_2\left(L_n^{\left(\sum_{j=1}^{k} y_j\right)},\mu\right)\right| \le W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, L_n^{\left(\sum_{j=1}^{k} y_j\right)}\right)$$

According to the convexity property of $\mathrm{T}_2(.,.)$ (see [37]), we have

$$\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^k x_j\right)}, L_n^{\left(\sum_{j=1}^k y_j\right)}\right) \le \frac{1}{n}\sum_{i=1}^n \mathrm{T}_2\left(\delta_{\left(\sum_{j=1}^k x_j\right)^i}, \delta_{\left(\sum_{j=1}^k y_j\right)^i}\right)$$

$$= \frac{1}{n}\sum_{i=1}^n \rho\left(\left(\sum_{j=1}^k x_j\right)^i, \left(\sum_{j=1}^k y_j\right)^i\right)^2 = \frac{1}{n}\rho_2^n\left(\sum_{j=1}^k x_j, \sum_{j=1}^k y_j\right)^2$$

which proves the claim.

Now, let $\left(\left(\sum_{j=1}^k X_j\right)_i\right)_i$ be an i.i.d sequence of law μ and let L_n be its empirical measure. Let m_n be the median of $W_2(L_n,\mu)$ and define $A = \left\{\sum_{j=1}^k x_j : W_2\left(L_n^{\left(\sum_{j=1}^k x_j\right)}, \mu\right) \le m_n\right\}$. Then $\mu^n(A) \ge 1/2$ and it is easy to show that $A^{(\varepsilon+r_0)} \subset \left\{\sum_{j=1}^k x_j : W_2\left(L_n^{\left(\sum_{j=1}^k x_j\right)}, \mu\right) \le m_n + \frac{(\varepsilon+r_0)}{\sqrt{n}}\right\}$.

Applying (4) to A gives

$$\mathrm{P}\left(W_2(L_n,\mu) > m_n + (\varepsilon+r_0)\sqrt{n}\right) \le b\ \exp\left(-a\varepsilon^2\right), \quad \forall \varepsilon \ge 0$$

Equivalently, as soon as $\sqrt{n}\left(\sum_{j-1}^k u_j - m_n\right) \ge r_0$, we have

$$\mathrm{P}\left(W_2(L_n,\mu) > \sum_{j-1}^k u_j\right) \le b\ \exp\left(-a\left(\sqrt{n}\left(\sum_{j-1}^k u_j - m_n\right) - r_0\right)^2\right).$$

Now, since $W_2(L_n,\mu)$ converges to 0 in probability (see Proposition 2), the sequence $m_n \to 0$ when $n \to +\infty$. Consequently,

$$\limsup_{n\to+\infty}\frac{1}{n}\log \mathrm{P}\left(W_2\left(L_n,\mu\right)>\sum_{j=1}^{k}u_j\right)\le -a\left(\sum_{j-1}^{k}u_j\right)^2,\quad \forall\sum_{j-1}^{k}u_j>0$$

The final step is given by Large deviations. According to Theorem 4,

$$\liminf_{n\to+\infty}\frac{1}{n}\log \mathrm{P}\left(W_2\left(L_n,\mu\right)>\sum_{j-1}^{k}u_j\right)\ge -\inf\left\{\sum_{j=1}^{k}H\left(v_j\,|\mu\right):\sum_{j=1}^{k}v_j\in \mathrm{P}_2\left(\chi_k\right)s.t.\sum_{j=1}^{k}W_2\left(v_j,\mu\right)>\sum_{j=1}^{k}u_j\right\}$$

This together with the preceding inequality yields

$$\inf\left\{\sum_{j=1}^{k}H\left(v_j\,|\mu\right):\sum_{j=1}^{k}v_j\in \mathrm{P}_2\left(\chi_k\right)s.t.\sum_{j=1}^{k}W_2\left(v_j,\mu\right)>\sum_{j=1}^{k}u_j\right\}\ge a\left(\sum_{j=1}^{k}u_j\right)^2$$

or in other words,

$$a\sum_{j=1}^{k}W_2\left(v_j,\mu\right)^2\le\sum_{j=1}^{k}H\left(v_j\,|\mu\right)$$

and this achieves the proof.

According to Jensen's inequality, the inequality $\mathrm{T}_1\left(\tilde{c}\right)$ is weaker than $\mathrm{T}_2\left(\tilde{c}\right)$; it was completely characterized in terms of square exponential integrability in [14].

Let us recall that a probability measure μ on χ_k verifies the Logarithmic-Sobolev inequality with constant $\tilde{c}>0$ ($\mathrm{LSI}\left(\tilde{c}\right)$ for short) if

$$H\left(\mu\middle|f^2\right)\le\tilde{c}\int\left|\nabla f\right|^2 d\mu$$

for all locally Lipschitz f, where the entropy functional is defined by

$$H\left(\mu\middle|f^2\right)=\int f\,\log f\,d\mu-\int f\,d\mu\log\left(\int f\,d\mu\right),\quad f\ge 0,$$

and the length of the gradient is defined by

$$|\nabla f|\left(\sum_{j=1}^{k} x_j\right) = \limsup_{y\to x} \frac{\left| f\left(\sum_{j=1}^{k} x_j\right) - f\left(\sum_{j=1}^{k} y_j\right)\right|}{\rho\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right)} \quad (5)$$

(when x_j are isolated points, we put $|\nabla f|\left(\sum_{j=1}^{k} x_j\right) = 0$.

In[32], Otto and Villani proved that if a probability measure μ on a Riemannian manifold M, satisfies the inequality $\mathrm{LSI}(\tilde{c})$ then it also satisfies the inequality $\mathrm{T}_2(\tilde{c})$. Their proof was rather involved and uses partial differential equations, optimal transportation results, and fine observations relating relative entropy and Fisher information[5]. It makes use of the dual formulation of transportation-cost inequalities [4], [5] relies on hypercontractivity properties of the Hamilton-Jacobi semigroup. Otto and Villani's result was successfully generalized by Wang on paths spaces in [38]. Lott and Villani showed that implication $\mathrm{LSI} \Rightarrow \mathrm{T}_2$ remains true on a length space provided the measure μ, satisfies a doubling condition and a local Poincare inequality (see [27]). (For the converse see [32]).

In [13], Cattiaux and Guillin give an example of a probability measure verifying T_2 and not LSI.

With Theorem 4 in hand, one could think that the implication $\mathrm{LSI} \Rightarrow \mathrm{T}_2$ is now completely straightforward. Namely, it is well known that the Logarithmic-Sobolcv inequality implies dimension free Gaussian concentration; since this latter is equivalent to Talagrand's T_2 inequality it should be clear that the Logarithmic-Sobolev inequality implies T_2. If μ, verifies the $\mathrm{LSI}(\tilde{c})$ inequality, then according to the additive property of the Logarithmic-Sobolcv inequality, one can conclude that the product measure μ^n verifies

$$H\left(\mu^n \left|f^2\right.\right) \le \tilde{c}\int \sum_{i=1}^{n} \left|\nabla_i f\right|^2 \left(\sum_{j=1}^{k} x_j\right) d\mu^n \left(\sum_{j=1}^{k} x_j\right) \quad (6)$$

where the length of the 'partial derivative' $\left|\nabla_i f\right|$ is defined according to (5). The problem is that, $\sum_{i=1}^{n} \left|\nabla_i f\right|^2 \left(\sum_{j=1}^{k} x_j\right)$ and $\left|\nabla f\right|^2 \left(\sum_{j=1}^{k} x_j\right)$ (computed with respect to ρ_2^n) may be different. The tensorized Logarithmic-Sobolev inequality will yield a series of concentration inequalities for functions such that $\sum_i \left|\nabla_i f\right|^2 \left(\sum_{j=1}^{k} x_j\right) \le \frac{1}{\mu^n} -$ almost everywhere and this class of functions may not contain I-Lipschitz functions for the ρ_2^n metric. This difficulty can be shown in the following theorems[39].

Theorem 6: Let μ, be a probability measure on χ_k and suppose that for all integer n the function F_n defined on χ_k^n by $F_n\left(\sum_{j=1}^{k} x_j\right) = W_2\left(L_n^{\sum_{j=1}^{k} x_j}, \mu\right)$ verifies $\sum_{i=1}^{n} \left|\nabla_i F_n\right|^2 \left(\sum_{j=1}^{k} x_j\right) \le \frac{1}{n}$ for μ^n almost every $\sum_{j=1}^{k} x_j \in \chi_k^n$ (7)

If μ, verifies the inequality $\mathrm{LSI}(\tilde{c})$, then μ, verifies the inequality $\mathrm{T}_2(\tilde{c})$. Suppose that $\chi_k = \mathbf{R}^d$ or a Riemannian manifold M, then according to Rademacher's Theorem, F_n is almost everywhere differentiable on $\left(\mathbf{R}^d\right)^n$ (resp. M^n) with respect to the Lebesgue measure. It is thus easy to show that condition (7) is fulfilled when μ, is absolutely continuous with respect to Lebesgue measure. This permits us to recover Otto and Villani's result as stated in [32].

Proof: As we said above the product measure μ^n verifies the inequality (6). Apply this inequality to $f = e^{\frac{s}{2}F_n}$, with $s \in \mathbf{R}^+$. It is easy to show that $\left|\nabla_i e^{\frac{s}{2}F}\right| = \frac{s}{2} e^{\frac{s}{2}F_n} \left|\nabla_i F_n\right|$, thus, using condition (7), we see that the right hand side of (6) is less than $\tilde{c}\,\frac{s^2}{4n}\int e^{sF_n}\,d\mu^n$.

Letting $Z(s) = \int e^{sF_n} d\mu^n$, we get the differential inequality:

$$\frac{Z'(s)}{sZ(s)} - \frac{\log Z(s)}{s^2} \le \frac{\tilde{c}}{4n}$$

Integrating this yields:

$$Z(s) = \int e^{sF_n} d\mu^n \le e^{s\int F_n d\mu^n + \frac{\tilde{c}s^2}{4n}}, \quad \forall s \in \mathbf{R}^+$$

This implies that

$$P\left(W_2(L_n, \mu) \ge t + E\left[W_2(L_n, \mu)\right]\right) \le e^{-nt^2/\tilde{c}}$$

According to Proposition 2 $E\left[W_2(L_n, \mu)\right] \to 0$. Arguing exactly as in proof of Theorem 4, one concludes that the inequality $T_2(\tilde{c})$ holds.

We show (see [39,11]) in the following Theorem that the implication $\mathrm{LSI} \Rightarrow T_2$ is true with a relaxed constant:

Theorem 7: Let μ be a probability measure on χ_k such that

$$\mu\left\{\sum_{j=1}^{k} x_j \in \chi_k : \rho^2\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} u_j\right) - \rho^2\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} v_j\right) = Q\right\} = 0,$$
$$\forall Q \in \mathbf{R}, \forall \sum_{j=1}^{k} u_j \ne \sum_{j=1}^{k} v_j \in \chi_k \tag{8}$$

If μ verifies the inequality $\mathrm{LSI}(\tilde{c})$ then μ satisfies $T(2\tilde{c})$.

For μ on $\mathbf{R}^d$, the condition (8) amounts to say that μ does not charge hyperplanes. We think that working better it would be possible to obtain the right constant $\tilde{c}$ instead of $2\tilde{c}$.

Proof: We will use a sort of symmetrization argument. First observe that the probability measure $\mu^n \times \mu^n$ verifies the following Logarithmic-Sobolcv inequality:

$$H\left(\mu^n \times \mu^n \left| f^2\right.\right) \le \tilde{c}\sum_{i=1}^{n}\left|\nabla_{i,1} f\right|^2\left(\sum_{j=1}^{k}x_j, \sum_{j=1}^{k}y_j\right)$$

$$+\left|\nabla_{i,2} f\right|^2\left(\sum_{j=1}^{k}x_j, \sum_{j=1}^{k}y_j\right) d\,\mu^n\left(\sum_{j=1}^{k}x_j\right) d\,\mu^n\left(\sum_{j=1}^{k}y_j\right)$$

for all $f:\chi_{nk}^n \times \chi_k^n \to \mathbf{R} : \left(\sum_{j=1}^{k}x_j, \sum_{j=1}^{k}y_j\right) \mapsto f\left(\sum_{j=1}^{k}x_j, \sum_{j=1}^{k}y_j\right)$, where $\left|\nabla_{i,1} f\right|$ (resp. $\left|\nabla_{i,2} f\right|$) denotes the length of the gradient with respect to the $\left(\sum_{j=1}^{k}x_j\right)^i$-coordinate (resp. the $\left(\sum_{j=1}^{k}y_j\right)^i$-coordinate).

Define $G_n\left(\sum_{j=1}^{k}x_j, \sum_{j=1}^{k}y_j\right) = W_2\left(L_n^{\left(\sum_{j=1}^{k}x_j\right)}, L_n^{\left(\sum_{j=1}^{k}y_j\right)}\right)$ for all $\sum_{j=1}^{k}x_j, \sum_{j=1}^{k}y_j \in \chi_k^n$. We want to apply the tensorized Logarithmic--Sobolev inequality to the function G_n. To do so we need to compute the length of its partial derivatives. Let us explain how to compute $L = \left|\nabla_{1,1} G_n\right|\left(\sum_{j=1}^{k}a_j, \sum_{j=1}^{k}b_j\right)$, for instance. For every $\sum_{j=1}^{k}z_j \in \chi_k$, let $\sum_{j=1}^{k}z_j a_j = \left(\sum_{j=1}^{k}z_j, \left(\sum_{j=1}^{k}a_j\right)^2, \ldots, \left(\sum_{j=1}^{k}a_j\right)^n\right)$; obviously,

$$L = \limsup_{z \to a^1} \frac{\left| W_2\left(L_n^{\left(\sum_{j=1}^k z_j a_j\right)}, L_n^{\left(\sum_{j=1}^k a_j\right)} \right) - W_2\left(L_n^{\left(\sum_{j=1}^k a_j\right)}, L_n^{\left(\sum_{j=1}^k b_j\right)} \right) \right|}{\rho\left(\left(\sum_{j=1}^k a_j\right), \left(\sum_{j=1}^k a_j\right)^1 \right)}$$

$$= \frac{1}{2W\left(L_n^{\left(\sum_{j=1}^k a_j\right)}, L_n^{\left(\sum_{j=1}^k b_j\right)} \right)} \limsup_{z \to a^1} \frac{\left| \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^k z_j a_j\right)}, L_n^{\left(\sum_{j=1}^k b_j\right)} \right) - \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^k a_j\right)}, L_n^{\left(\sum_{j=1}^k b_j\right)} \right) \right|}{\rho\left(\left(\sum_{j=1}^k z_j\right), \left(\sum_{j=1}^k a_j\right)^1 \right)}$$

According to the condition (8), the probability measure μ is diffuse; so the probability of points $\sum_{j=1}^k x_j \in \chi_k^n$ having distinct coordinates is one. So, one can suppose without restriction that the coordinates $of \sum_{j=1}^k a_j$ (resp. $\sum_{j=1}^k b_j$) are all different. If $\sum_{j=1}^k z_j$ is sufficiently close to $\left(\sum_{j=1}^k a_j\right)^1$, the coordinates of $\sum_{j=1}^k z_j a_j$ arc all distinct too. According to [36], the optimal transport of $L_n^{\left(\sum_{j=1}^k a_j\right)}$ on $L_n^{\left(\sum_{j=1}^k b_j\right)}$ is given by a permutation, this means that there is at least one permutation a of $\{1,\ldots,n\}$ such that

$$\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^k a_j\right)}, L_n^{\left(\sum_{j=1}^k b_j\right)} \right) = n^{-1} \sum_{i=1}^n \rho\left(\left(\sum_{j=1}^k a_j\right)^i, \left(\sum_{j=1}^k b_j\right)^{\sigma(i)} \right)^2$$

Let us denote by S the set of series of these permutations and define accordingly the set $S_{\left(\sum_{j=1}^{k} z_j\right)}$ of permutations realizing the optimal transport of $L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}$ on $L_n^{\left(\sum_{j=1}^{k} b_j\right)}$.

Without loss of generality, one can suppose that S is a singleton. Indeed, let σ and $\tilde{\sigma}$ be two distinct permutations and consider

$$H_{\sigma,\tilde{\sigma}} = \left\{ \left(\sum_{j=1}^{k} x_j\right) \in \chi_k{}^n : \sum_{i=1}^{n} \rho\left(\left(\sum_{j=1}^{k} x_j\right)^i, \left(\sum_{j=1}^{k} b_j\right)^{\sigma(i)}\right)^2 = \sum_{i=1}^{n} \rho\left(\left(\sum_{j=1}^{k} x_j\right)^i, \left(\sum_{j=1}^{k} b_j\right)^{\sigma(i)}\right)^2 \right\}$$

Applying Fubini's Theorem together with the condition (8), one gets easily that $\mu^n\left(H_{\sigma,\tilde{\sigma}}\right) = 0$. This readily proves the claim. In the sequel we will set $S = \left\{\sigma^*\right\}$.

Now we claim that if $\left(\sum_{j=1}^{k} z_j\right)$ is sufficiently close to $\left(\sum_{j=1}^{k} a_j\right)^1$, then $S_{\left(\sum_{j=1}^{k} z_j\right)} = \left\{\sigma^*\right\}$. Indeed, let

$$\varepsilon_0 = \min_{\sigma \neq \tilde{\sigma}} \left\{ n^{-1} \sum_{i=1}^{n} \rho\left(\left(\sum_{j=1}^{k} a_j\right)^i, \left(\sum_{j=1}^{k} b_j\right)^{\sigma(i)}\right)^2 - \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, L_n^{\left(\sum_{j=1}^{k} b_j\right)}\right) \right\} > 0$$

then there is a neighborhood V of $\left(\sum_{j=1}^{k} a_j\right)^1$ such that for all $\left(\sum_{j=1}^{k} z_j\right) \in V$, we have

$$\left| \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}, L_n^{\left(\sum_{j=1}^{k} b_j\right)} \right) - \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, L_n^{\left(\sum_{j=1}^{k} b_j\right)} \right) \right| \le \varepsilon_0 / 3$$

and for all permutation σ ,

$$\left| n^{-1} \sum_{i=1}^{n} \rho\left(\left(\sum_{j=1}^{k} z_j a_j \right)^i, \left(\sum_{j=1}^{k} b_j \right)^{\sigma(i)} \right)^2 - n^{-1} \sum_{i=1}^{n} \rho\left(\left(\sum_{j=1}^{k} a_j \right)^i, \left(\sum_{j=1}^{k} b_j \right)^{\tilde{\sigma}(i)} \right)^2 \right| \le \varepsilon_0 / 3$$

Now, if $\sum_{j=1}^{k} z_j \in V$ and $\sigma \in S_{\sum_{j=1}^{k} z_j}$, we have

$$\left| n^{-1} \sum_{i=1}^{n} \rho\left(\left(\sum_{j=1}^{k} a_j \right)^i, \left(\sum_{j=1}^{k} b_j \right)^{\sigma(i)} \right)^2 \le n^{-1} \sum_{i=1}^{n} \rho\left(\left(\sum_{j=1}^{k} z_j a_j \right)^i, \left(\sum_{j=1}^{k} b_j \right)^{\sigma(i)} \right)^2 \right| + \varepsilon_0 / 3$$

$$= \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}, L_n^{\left(\sum_{j=1}^{k} b_j\right)} \right) + \varepsilon_0 / 3 \le \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, L_n^{\left(\sum_{j=1}^{k} b_j\right)} \right) + 2\varepsilon_0 / 3$$

By the definition of the number ε_0, one concludes that $\sigma = \sigma^*$,

which proves the claim. Now, if $\sum_{j=1}^{k} z_j \in V$, then

$$\frac{\left| \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}, L_n^{\left(\sum_{j=1}^{k} a_j\right)} \right) - \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, L_n^{\left(\sum_{j=1}^{k} b_j\right)} \right) \right|}{\rho\left(\sum_{j=1}^{k} z_j, \left(\sum_{j=1}^{k} a_j \right)^1 \right)}$$

$$= \frac{\left| \rho\left(\sum_{j=1}^{k} z_j, \left(\sum_{j=1}^{k} b_j \right)^{\sigma^*(1)} \right)^2 - \rho\left(\left(\sum_{j=1}^{k} a_j \right)^1, \left(\sum_{j=1}^{k} b_j \right)^{\sigma^*(1)} \right)^2 \right|}{n\rho\left(\sum_{j=1}^{k} z_j, \left(\sum_{j=1}^{k} a_j \right)^1 \right)} \le \frac{1}{n}\left(\rho\left(\sum_{j=1}^{k} z_j, \left(\sum_{j=1}^{k} b_j \right)^{\sigma^*(1)} \right) + \rho\left(\left(\sum_{j=1}^{k} a_j \right)^1, \left(\sum_{j=1}^{k} b_j \right)^{\sigma^*(1)} \right) \right)$$

So letting $\sum_{j=1}^{k} z_j \to \left(\sum_{j=1}^{k} a_j\right)^1$ yields $L \le \dfrac{\rho\left(\left(\sum_{j=1}^{k} a_j\right)^1, \left(\sum_{j=1}^{k} b_j\right)^{\sigma^{*(1)}}\right)}{nW_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, L_n^{\left(\sum_{j=1}^{k} b_j\right)}\right)}$.

Doing the same for the other partial derivatives yields:

$$\sum_{i=1}^{n} \left|\nabla_{i,1} G_n\right|^2 \left(\sum_{j=1}^{k} a_j, \sum_{j=1}^{k} b_j\right) \le \frac{\sum_{i=1}^{n} \rho\left(\left(\sum_{j=1}^{k} a_j\right)^i, \left(\sum_{j=1}^{k} b_j\right)^{\sigma^{*(i)}}\right)^2}{n^2 \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, L_n^{\left(\sum_{j=1}^{k} b_j\right)}\right)} = \frac{1}{n}$$

Finally,

$$\sum_{i=1}^{n} \left|\nabla_{i,1} G_n\right|^2 \left(\sum_{j=1}^{k} a_j, \sum_{j=1}^{k} b_j\right) + \left|\nabla_{i,2} G_n\right|^2 \left(, \sum_{j=1}^{k} a_j, \sum_{j=1}^{k} b_j\right) \le \frac{2}{n}$$

for $\mu^n \times \mu^n$ almost every $\left(\sum_{j=1}^{k} a_j, \sum_{j=1}^{k} b_j\right) \in \chi_k^n \times \chi_k^n$.

Now reasoning as in the proof of Theorem 7, one concludes that

$$\mathrm{P}\left(W_2\left(L_n^{\sum_{j=1}^{k} X_j}, L_n^{\sum_{j=1}^{k} Y_j}\right) > t + \mathrm{E}\left[W_2\left(L_n^{\sum_{j=1}^{k} X_j}, L_n^{\sum_{j=1}^{k} Y_j}\right)\right]\right) \le e^{-nt^{2/(2\tilde{c})}}$$

On the other hand, an easy adaptation of Proposition 2 yields

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathrm{P}\left(W_2\left(L_n^{\left(\sum_{j=1}^{k} X_j\right)}, L_n^{\left(\sum_{j=1}^{k} Y_j\right)}\right) > t + \mathrm{E}\left[W_2\left(L_n^{\left(\sum_{j=1}^{k} X_j\right)}, L_n^{\left(\sum_{j=1}^{k} Y_j\right)}\right)\right]\right)$$

$$-\inf\left\{\left(\sum_{j=1}^{k} H\left(\left(v_j\right)_1 | \mu\right) + \sum_{j=1}^{k} H\left(\left(v_j\right)_2 | \mu\right)\right) : \left(\sum_{j=1}^{k} v_j\right)_1, \left(\sum_{j=1}^{k} v_j\right)_2 \in \mathrm{P}_2\left(\chi_k\right) s.t. \sum_{j=1}^{k} W_2\left(\left(v_j\right)_1, \left(v_j\right)_2\right) > t\right\}$$

From this follows as before that

$$\sum_{j=1}^{k} \mathrm{T}_2\left(\left(v_j\right)_1 | \mu\right) \le 2\tilde{c}\left(\sum_{j=1}^{k} H\left(\left(v_j\right)_1 | \mu\right) + \sum_{j=1}^{k} H\left(\left(v_j\right)_2 | \mu\right)\right)$$

holds for all probability measures $\left(\sum_{j=1}^{k} v_j\right)_1, \left(\sum_{j=1}^{k} v_j\right)_2$ belonging to $\mathrm{P}_2(\chi_k)$. Taking $\left(\sum_{j=1}^{k} v_j\right)_2 = \mu$ gives the inequality $\mathrm{T}_2(2\tilde{c})$.

(See [39]) to recover and extend a result of Lott and Villani using Hamilton-Jacobi. Following [27], one says that a probability measure μ on χ_k verifies the inequality $\mathrm{LSI}^+(\tilde{c})$ if

$$H\left(\mu \middle| f^2\right) \le \tilde{c}\int \left|\nabla^- f\right|^2 d\mu$$

holds true for all locally Lipschitz f, where the subgradient norm $\left|\nabla^- f\right|$ is defined by

$$\left|\nabla^- f\right|\left(\sum_{j=1}^{k} x_j\right) = \limsup_{y \to x} \frac{\left[f\left(\left(\sum_{j=1}^{k} y_j\right)\right) - f\left(\left(\sum_{j=1}^{k} x_j\right)\right)\right]_+}{\rho\left(\left(\sum_{j=1}^{k} x_j\right), \left(\sum_{j=1}^{k} y_j\right)\right)},$$

with $\left|\sum_{j=1}^{k} a_j\right| = \max\left(\sum_{j=1}^{k} a_j, 0\right)$. Since $\left|\nabla^- f\right| \le \left|\nabla f\right|$, the inequality $\mathrm{LSI}^+(\tilde{c})$ is stronger than LSI ; more precisely, $\mathrm{LSI}^+(\tilde{c}) \Rightarrow \mathrm{LSI}(\tilde{c})$.

Theorem 8: If μ verifies the inequality $\mathrm{LSI}^+(\tilde{c})$, then μ verifies $\mathrm{T}_2(\tilde{c})$.

Proof: The inequality LSI^+ tensorizes, so μ^n verifies[39]

$$H\left(\mu^n \middle| f^2\right) \le c\int \sum_{i=1}^{n} \left|\nabla_i^- f\right|^2 d\mu^n$$

Take $f = e^{\frac{s}{2}F_n}, \quad s \in \mathbf{R}^+$ with $F_n\left(\sum_{j=1}^{k} x_j\right) = W_2\left(L_n^{\left(\sum_{j=1}^{k} x_j\right)}, \mu\right)$. Once again, it is easy to check that $\left|\nabla_i^- e^{\frac{s}{2}F_n}\right| = \frac{s}{2} e^{\frac{s}{2}F_n} \left|\nabla_i^- F_n\right|$ (note that the function $\sum_{j=1}^{k} x_j \mapsto e^{s\left(\sum_{j=1}^{k} x_j\right)}$ is non decreasing). Reasoning as in the proof of Theorem 7 it is enough to show that $\sum_i \left|\nabla_i^- F_n\right|^2 \left(\sum_{j=1}^{k} x_j\right) \le 1/n$ for μ^{n_-} almost all $\sum_{j=1}^{k} x_j \in \chi_k^n$. Let us show how to compute $\left|\nabla_i^- F_n\right|$. Let $\sum_{j=1}^{k} z_j \in \sum_{j=1}^{k} X_j$, $\sum_{j=1}^{k} a_j = \left(\left(\sum_{j=1}^{k} a_j\right)^1, \ldots, \left(\sum_{j=1}^{k} a_j\right)^n\right) \in \chi_k^n$ and set - $\sum_{j=1}^{k} z_j a_j = \left(\sum_{j=1}^{k} z_j, \left(\sum_{j=1}^{k} a_j\right)^2, \ldots, \left(\sum_{j=1}^{k} a_j\right)^n\right)$

$$\left|\nabla_i^- F_n\right|\left(\sum_{j=1}^{k} a_j\right) = \frac{1}{2F_n\left(\left(\sum_{j=1}^{k} a_j\right)\right)} \limsup_{z \to a^1} \frac{\left[\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)}, \mu\right) - \mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, \mu\right)\right]_+}{\rho\left(\sum_{j=1}^{k} a_j, \left(\sum_{j=1}^{k} a_j\right)^1\right)}$$

Let $\pi \in \mathrm{P}\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)}, \mu\right)$ be an optimal coupling; it is not difficult to see that we can write

$$\pi\left(d\left(\sum_{j=1}^{k}x_j\right),d\left(\sum_{j=1}^{k}y_j\right)\right)=p\left(\left(\sum_{j=1}^{k}x_j\right),d\left(\sum_{j=1}^{k}y_j\right)\right)L_n^{\left(\sum_{j=1}^{k}a_j\right)}\left(d\left(\sum_{j=1}^{k}x_j\right)\right)$$

,where $p\left(\left(\sum_{j=1}^{k}a_j\right)^i,d\sum_{j=1}^{k}y_j\right)=\left(\sum_{j=1}^{k}v_j\right)_i\left(d\left(\sum_{j=1}^{k}y_j\right)\right)$ with $\left(\sum_{j=1}^{k}v_j\right)_1,\ldots,\left(\sum_{j=1}^{k}v_j\right)_n$ probability measures on χ_k such that $n^{-1}\left(\left(\sum_{j=1}^{k}v_j\right)_1+\ldots+\left(\sum_{j=1}^{k}v_j\right)_n\right)=\mu$.

Let $\tilde{p}$ be defined as p with $\left(\sum_{j=1}^{k}z_j\right)$ in place of $\left(\sum_{j=1}^{k}v_j\right)^1$; then $\tilde{\pi}=\tilde{p}\left(\sum_{j=1}^{k}x_j,d\left(\sum_{j=1}^{k}y_j\right)\right)L_n^{\left(\sum_{j=1}^{k}z_ja_j\right)}\left(d\left(\sum_{j=1}^{k}y_j\right)\right)$ belongs to $P\left(L_n^{\left(\sum_{j=1}^{k}z_ja_j\right)},\mu\right)$

We have

$$\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k}z_ja_j\right)},\mu\right)-\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k}a_j\right)},\mu\right)\le\int\rho\left(\sum_{j=1}^{k}x_j,\sum_{j=1}^{k}y_j\right)^2d\tilde{\pi}\left(\sum_{j=1}^{k}x_j,\sum_{j=1}^{k}y_j\right)-\int\rho\left(\sum_{j=1}^{k}x_j,\sum_{j=1}^{k}y_j\right)^2d\pi\left(\sum_{j=1}^{k}x_j,\sum_{j=1}^{k}y_j\right)$$

$$=\frac{1}{n}\sum_{i=1}^{n}\int\rho\left(\left(\sum_{j=1}^{k}z_ja_j\right)^i,\sum_{j=1}^{k}y_j\right)^2d\left(\sum_{j=1}^{k}v_j\right)_i\left(\sum_{j=1}^{k}y_j\right)-\frac{1}{n}\sum_{i=1}^{n}\int\rho\left(\left(\sum_{j=1}^{k}a_j\right)^i,\sum_{j=1}^{k}y_j\right)^2d\left(\sum_{j=1}^{k}v_j\right)_i\left(\sum_{j=1}^{k}y_j\right)$$

$$=\frac{1}{n}\int\rho\left(\sum_{j=1}^{k}z_j,\sum_{j=1}^{k}y_j\right)^2-\rho\left(\left(\sum_{j=1}^{k}a_j\right)^1,\sum_{j=1}^{k}y_j\right)^2d\left(\sum_{j=1}^{k}v_j\right)_1\left(\sum_{j=1}^{k}y_j\right)$$

$$\le\frac{1}{n}\rho\left(\sum_{j=1}^{k}z_j,\left(\sum_{j=1}^{k}a_j\right)^1\right)\int\rho\left(\sum_{j=1}^{k}z_j,\sum_{j=1}^{k}y_j\right)+\rho\left(\left(\sum_{j=1}^{k}a_j\right)^1,\sum_{j=1}^{k}y_j\right)d\left(\sum_{j=1}^{k}v_j\right)_1\left(\sum_{j=1}^{k}y_j\right)$$

Since the function $\sum_{j=1}^{k}x_j\mapsto\left[\sum_{j=1}^{k}x_j\right]_+$ is non decreasing, we have

$$\frac{\left(\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} z_j a_j\right)},\mu\right)-\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)},\mu\right)\right)_+}{\rho\left(\sum_{j=1}^{k} z_j,\left(\sum_{j=1}^{k} a_j\right)^1\right)} \le \frac{1}{n}\int \rho\left(\sum_{j=1}^{k} z_j,\sum_{j=1}^{k} y_j\right)+\rho\left(\left(\sum_{j=1}^{k} a_j\right)^1,\sum_{j=1}^{k} y_j\right) d\left(\sum_{j=1}^{k} v_j\right)_1\left(\sum_{j=1}^{k} y_j\right)$$

Letting $\sum_{j=1}^{k} z_j \to \left(\sum_{j=1}^{k} a_j\right)^1$ yields

$$\left|\nabla_i^- F_n\right|\left(\sum_{j=1}^{k} a_j\right)^2 \le \frac{\int \rho\left(\left(\sum_{j=1}^{k} a_j\right)^1,\sum_{j=1}^{k} a_j\right)^2 d\left(\sum_{j=1}^{k} v_j\right)_1\left(\sum_{j=1}^{k} y_j\right)}{n^2\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)},\mu\right)}$$

. Doing the same computations for the other derivatives (with the same optimal coupling π), we get

$$\left|\nabla_i^- F_n\right|\left(\sum_{j=1}^{k} a_j\right)^2 \le \frac{\int \rho\left(\left(\sum_{j=1}^{k} a_j\right)^i,\sum_{j=1}^{k} y_j\right)^2 d\left(\sum_{j=1}^{k} v_j\right)_i\left(\sum_{j=1}^{k} y_j\right)}{n^2\mathrm{T}_2\left(L_n^{\left(\sum_{j=1}^{k} a_j\right)},\mu\right)}$$

.

Summing these inequalities gives $\sum_i \left|\nabla_i^- F_n\right|^2\left(\sum_{j=1}^{k} a_j\right) \le 1/n$ for all $\sum_{j=1}^{k} a_j \in \chi_k^n$, which achieves the proof.

4. Non Gaussian Concentration

The following theorem can be established with exactly the same proof as Theorem 1(see [39]).

Theorem 9: Let μ be a probability measure on χ_k^n, $p \geq 2$ and $a > 0$. The following propositions are equivalent:

(i)There are $r_0, b \geq 0$ such that for every n the probability measure μ^n verifies for A subset of χ_k^n with $\mu^n(A) \geq \frac{1}{2}$,

$$\mu^n\left(A^{(\varepsilon+r_0)}\right) \geq 1 - be^{-a\varepsilon^p}, \quad \forall \varepsilon \geq 0 \qquad (9)$$

where the enlargement $A^{(\varepsilon+r_0)}$ is performed with respect to the metric ρ_p^n on χ_k^n defined by

$$\rho_p^n\left(\sum_{j=1}^k x_j, \sum_{j=1}^k y_j\right)\left[\sum_{i=1}^n \rho\left(\left(\sum_{j=1}^k x_j\right)^i, \left(\sum_{j=1}^k y_j\right)^i\right)^p\right]^{1/p}, \quad \forall \sum_{j=1}^k x_j, \sum_{j=1}^k y_j \in \chi_k^n$$

(ii)The probability measure μ verifies the following series of transportation cost inequality:

$$\sum_{j=1}^k \mathrm{T}_p\left(v_j, \mu\right) \leq \left(\sum_{j=1}^k a_j\right)^{-1} \sum_{j=1}^k H\left(v_j \middle| \mu\right), \quad \forall \sum_{j=1}^k v_j \in \mathrm{P}_p\left(\chi_k\right)$$

To find the series of the transportation-cost inequality equivalent to Talagrand's two level concentration inequalities as an example which are well adapted to concentration rates between exponential and Gaussian.

Let us say that a probability measure μ on $\mathbf{R}^d$ satisfies a two level dimension free concentration inequality of order $p \in [1,2]$ if there are two non-negative constants a and b such that for every n the inequality

$$\mu^n\left(A + \sqrt{(\varepsilon + r_0)}B_2 + \sqrt[p]{(\varepsilon + r_0)}B_p\right) \geq 1 - be^{-a(\varepsilon+r_0)}, \quad \forall \varepsilon \geq -r_0 \qquad (10)$$

holds for all measurable subset A of $\left(\mathbf{R}^d\right)^n$ such that $\mu^n(A) \geq \frac{1}{n}$, where B_2 and B_p are the standard unit balls of $\left(\mathbf{R}^d\right)^n$. Inequalities of this form appear in [35], where it is proved that the measure $d\mu_p\left(\sum_{j=1}^{k} x_j\right) = \left(\sum_{j=1}^{k} Z_j\right)_p^{-1} e^{-\left|\sum_{j=1}^{k} x_j\right|^p},\ p \geq 1$ verifies such a bound.

The series of transportation-cost adapted to this kind of concentration is defined for all probability measures $\left(\sum_{j=1}^{k} v_j\right)_1, \left(\sum_{j=1}^{k} v_j\right)_2$ on $\left(\mathbf{R}^{d^n}\right)$ by

$$\sum_{j=1}^{k} \mathrm{T}_{2,p}\left(v_j, \mu\right) = \inf_{\pi \in \mathrm{P}\left(\left(\sum_{j=1}^{k} v_j\right)_1, \left(\sum_{j=1}^{k} v_j\right)_2\right)} \int \sum_{i=1}^{n} \sum_{j=1}^{d} \alpha_p\left(x_j^i - y_j^i\right) d\pi\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right)$$

where $\alpha_p\left(\sum_{j=1}^{k} u_j\right) = \min\left(\left|\sum_{j=1}^{k} u_j\right|^2, \left|\sum_{j=1}^{k} u_j\right|^p\right)$ (here $\sum_{j=1}^{k} x_j = \left(\left(\sum_{j=1}^{k} x_j\right)^1, \ldots, \left(\sum_{j=1}^{k} x_j\right)^n\right)$ with $\left(\sum_{j=1}^{k} x_j\right)^i \in \mathbf{R}^d$ for all i).

The following lemma collects different facts that are needed in the proof see[22,34].

Lemma 10: For all $\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j \geq 0,\ \alpha_p\left(\sum_{j=1}^{k} x_j + \sum_{j=1}^{k} y_j\right) \leq 2\alpha_p\left(\sum_{j=1}^{k} x_j\right) + 2\alpha_p\left(\sum_{j=1}^{k} y_j\right)$.

For all integer $n > 1$ and all probability measures $\left(\sum_{j=1}^{n} v_j\right)_1, \left(\sum_{j=1}^{n} v_j\right)_2$ and $\left(\sum_{j=1}^{n} v_j\right)_3$ on $\left(\mathbf{R}^d\right)^n$,

$$\sum_{j=1}^{n} \mathrm{T}_{2,p}\left(\left(v_j\right)_1, \left(v_j\right)_3\right) \leq 2\sum_{j=1}^{n} \mathrm{T}_{2,p}\left(\left(v_j\right)_1, \left(v_j\right)_2\right) + 2\sum_{j=1}^{n} \mathrm{T}_{2,p}\left(\left(v_j\right)_2, \left(v_j\right)_3\right).$$

For all integer $n > 1$ and all $\varepsilon \geq -r_0$, define

$$B_{2,p}\left(\varepsilon+r_0\right)=\left\{\sum_{j=1}^{k}x_j\in\left(\mathbf{R}^d\right)^n:\sum_{i=1}^{n}\sum_{j=1}^{d}\alpha_p\left(x_j^i\right)\leq\left(\varepsilon+r_0\right)\right\}.$$

Then for all $p\in[1,2]$,

$$\frac{1}{12}\left(\sqrt{(\varepsilon+r_0)}B_2+\sqrt[p]{(\varepsilon+r_0)}B_p\right)\subset\sqrt{(\varepsilon+r_0)}B_{2,p}\left(\varepsilon+r_0\right)\subset\sqrt{(\varepsilon+r_0)}B_2+\sqrt[p]{(\varepsilon+r_0)}B_p$$

Theorem 11: Let μ be a probability measure on $\mathbf{R}^d$ and $p\in[1,2]$. The following propositions are equivalent:

(i)The two level concentration (10) holds for some non-negative $\sum_{j=1}^{k}a_j,\sum_{j=1}^{k}b_j$ independent of n

(ii)The probability measure μ verifies the series of transportation-cost inequality

$$\sum_{j=1}^{k}\mathrm{T}_{2,p}\left(v_j,\mu\right)\leq\tilde{c}\sum_{j=1}^{k}H\left(v_j\,|\mu\right),\quad\forall\sum_{j=1}^{n}v_j\in\mathrm{P}\left(\mathbf{R}^d\right)$$

for some constant $\tilde{c}$.

More precisely, if (10) holds for some constants $\sum_{j=1}^{k}a_j,\sum_{j=1}^{k}b_j$, then the series of the series of the transportation-cost inequality holds with the constant $\tilde{c}=\dfrac{\tilde{c}}{\sum_{j=1}^{k}a_j}$ (for $j=1,\tilde{c}=288$ see[39]). Conversely, if the transportation-cost inequality holds for some constant $\tilde{c}$, then (10) is true for $\sum_{j=1}^{k}b_j=2$ and

$$\sum_{j=1}^{k} a_j = 1/(2\tilde{c})$$

Proof: Let us recall the proof of (ii) implies (i). According to the tensorization property, for all n and all probability measure v on $\left(\mathbf{R}^d\right)^n$,

$$\mathrm{T}_{2,p}\left(\sum_{j=1}^{n} v_j, \mu^n\right) \le \tilde{c} H\left(\sum_{j=1}^{n} v_j \,\middle|\, \mu^n\right)$$

holds. Take A and B in $\left(\mathbf{R}^d\right)^n$ and define $d\mu_A^n = \mathrm{I}_A d\mu / \mu^n(A)$ and $d\mu_B^n = \mathrm{I}_B d\mu / \mu^n(B)$. According to point (ii) of Lemma10 and the transportation-cost inequality satisfied by μ^n, one has

$$\begin{aligned}\mathrm{T}_{2,p}\left(\mu_A^n, \mu_B^n\right) &\le 2\mathrm{T}_{2,p}\left(\mu_A^n, \mu^n\right) + 2\mathrm{T}_{2,p}\left(\mu_B^n, \mu^n\right)\\ &\le 2\tilde{c}H\left(\mu_A^n \,\middle|\, \mu^n\right) + 2\tilde{c}H\left(\mu_B^n \,\middle|\, \mu^n\right)\\ &= -2\tilde{c}\log\left(\mu^n(A)\,\mu^n(B)\right)\end{aligned}$$

Define

$$\tilde{c}_{2,p}(A,B) = \inf\left\{(\varepsilon + r_0) \ge 0 \; s.t.\; \left(A + B_{2,p}(\varepsilon + r_0)\right) \cap B \ne \phi\right\}$$

then $\mathrm{T}_{2,p}\left(\mu_A^n, \mu_B^n\right) \ge \tilde{c}_{2,p}(A,B)$ and so

$$\mu^n(A)\,\mu^n(B) \le e^{-\tilde{c}_{2,p}(A,B)/2\tilde{C}}.$$

Now, if $\mu^n(A) \ge \frac{1}{2}$ and $B = \left(\mathbf{R}^d\right)^n \setminus \left(A + B_{2.p}(\varepsilon + r_0)\right)$, one has $c_{2,p}(A,B) = (\varepsilon + r_0)$ and so $\mu^n\left(A + B_{2,p}(\varepsilon + r_0)\right) \ge 1 - 2e^{-(\varepsilon + r_0)/2\tilde{c}}$. Using point (iii) of Lemma 10 gives

$$\mu^n\left(A + \sqrt{(\varepsilon + r_0)}B_2 + \sqrt[p]{(\varepsilon + r_0)}B_p\right) \ge 1 - 2e^{-(\varepsilon + r_0)/2\tilde{c}}.$$

Now let us prove the converse. Let $(X_i)_i$ be an i.i.d sequence of law μ and let L_n be its empirical measure. Consider

$A=\left\{\sum_{j=1}^{k}x_j\in\left(\mathbf{R}^d\right)^n\ s.t.\ \mathrm{T}_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}x_j\right)},\mu\right)\le m_n\right\}$ where m_n denotes the median of $T_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}x_j\right)},\mu\right)$. According to point (iii) of Lemma10 $A+\sqrt{(\varepsilon+r_0)}B_2+\sqrt[p]{(\varepsilon+r_0)}B_p\subset A+12B_{2,p}\left(\varepsilon+r_0\right)$. Let $\sum_{j=1}^{k}x_j\in A+12B_{2,p}\left(\varepsilon+r_0\right)$; there is some $\sum_{j=1}^{k}\bar{x}_j\in A+12B_{2,p}\left(\varepsilon+r_0\right)$ there is some $\sum_{j=1}^{k}\bar{x}_j\in A$ such that

$$\sum_{i=1}^{n}\sum_{j=1}^{d}\alpha_p\left(\frac{x_j^i-\bar{x}_j^i}{12}\right)\le\left(\varepsilon+r_0\right)$$

(here $\sum_{j=1}^{k}x_j=\left(\left(\sum_{j=1}^{k}x_j\right)^1,\left(\sum_{j=1}^{k}x_j\right)^2,\ldots,\left(\sum_{j=1}^{k}x_j\right)^n\right)$ with $\left(\sum_{j=1}^{k}x_j\right)^i\in\mathbf{R}^d$). Since $\dfrac{\alpha_p\left(\sum_{j=1}^{k}x_j\ /12\right)\ge\alpha_p\left(\sum_{j=1}^{k}x_j\right)}{\tilde{c}_1}$, we get $\mathrm{T}_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}x_j\right)},L_n^{\left(\sum_{j=1}^{k}\bar{x}_j\right)}\right)\le\dfrac{\tilde{c}_1\left(\varepsilon+r_0\right)}{n}$.

(for $j=1,\tilde{c}_3=144$ (see [39]). According to point (ii) of Lemma10

$$\mathrm{T}_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}x_j\right)},\mu\right)\le 2\mathrm{T}_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}x_j\right)},L_n^{\left(\sum_{j=1}^{k}\bar{x}_j\right)}\right)+2\mathrm{T}_{2,p}\left(L_n^{\left(\sum_{j=1}^{k}\bar{x}_j\right)},\mu\right)\le 2m_n+\frac{\tilde{c}\left(\varepsilon+r_0\right)}{n}.$$

Consequently, the following holds for all n :

$$\mathrm{P}\left(\mathrm{T}_{2,p}\left(L_n,\mu\right)\ge 2m_n+\tilde{c}\left(\varepsilon+r_0\right)/n\right)\le\sum_{j=1}^{k}b_je^{-\sum_{j=1}^{k}a_j\left(\varepsilon+r_0\right)},\quad\forall\varepsilon\ge -r_0$$

Reasoning as in the proof of Theorem 1one concludes that

$$\sum_{j=1}^{n} \mathrm{T}_{2,p}\left(\nu_j,\mu\right) \leq \frac{\tilde{c}}{\sum_{j=1}^{k} a_j} \sum_{j=1}^{n} H\left(\nu_j \,|\mu\right),$$

for every $\sum_{j=1}^{n} \nu_j \in \mathrm{P}\left(\mathbf{R}^d\right)$

Remark 12: (i) If $\mu^n(A) = \frac{1}{2}$ we have $\frac{1}{2}\mu^n(B) \leq e^{-\tilde{c}_{2,p}(A,B)/2\tilde{c}}$,approximately, we have for $\tilde{c}_{2,p}(A,B) = 0$ that $\mu^n(B) \leq 2$

(iii) We can deduce that $2e^{-\tilde{c}_{2,p}(A,B)/2\tilde{c}} \geq 1$ and $\tilde{c}_{2,p}(A,B) \leq 2\log 2$ .Hence $\tilde{c} \geq \frac{(\varepsilon + r_0)}{2\log 2}$

5. Poincare' inequality and series of the exponential concentration

We consider more carefully the case $p = 1$ of the preceding one. Let us recall that a probability measure μ on $\mathbf{R}^d$ satisfies the Poincare inequality with constant $\tilde{c} > 0$ if

$$Var_\mu(f) \leq \tilde{c} \int |\nabla f|_2^2 d\mu \qquad (11)$$

for all smooth f .

The following theorem (see [39]) proves the equivalence between Poincare inequality, dimension of the series of free exponential concentration and the corresponding series of transportation-cost inequality.

Theorem 13: Let μ be a probability measure on $\mathbf{R}^d$. The following propositions are equivalent:

(i) The probability measure μ verifies Poincare inequality with a constant $\tilde{c}$.

(ii) The probability measure μ verifies for some constants $\sum_{j=1}^{n} a_j, \sum_{j=1}^{n} b_j > 0$

$$\mu^n\left(A + D_{2,1}(\varepsilon + r_0)\right) \geq 1 - \sum_{j=1}^{n} b_j e^{-\left(\sum_{j=1}^{n} a_j\right)(\varepsilon + r_0)}, \quad \forall \varepsilon \geq -r_0$$

for a subset A of $\left(\mathbf{R}^d\right)^n$ such that $\mu^n(A) \geq 1/2$, where the set $D_{2,1}(\varepsilon + r_0)$ is defined by

$$D_{2,1}(\varepsilon + r_0) = \left\{ \sum_{j=1}^{k} x_j \in \left(\mathbf{R}^d\right)^n \ s.t. \sum_{i=1}^{n} \alpha_1 \left(\left| \left(\sum_{j=1}^{k} x_j \right)^i \right|_2 \right) \leq (\varepsilon + r_0) \right\}$$

(iii) The probability measure μ verifies the following series of transportation-cost inequality for some constant $c_2 > 0$.

$$\mathrm{T}_{SG}\left(\sum_{j=1}^{n} v_j, \mu\right) = \inf \int \alpha_1 \left(\left| \sum_{j=1}^{k} x_j - \sum_{j=1}^{k} y_j \right|_2 \right) d\pi\left(\sum_{j=1}^{k} x_j, \sum_{j=1}^{k} y_j\right) \leq \tilde{c}_2 H\left(\sum_{j=1}^{n} v_j \,|\, \mu\right),$$

for every

$$\sum_{j=1}^{n} v_j \in \mathrm{P}\left(\mathbf{R}^d\right)$$

More precisely:

(i) implies (ii) with $\sum_{j=1}^{n} a_j = K \max\left(\tilde{c}_1, \sqrt{\tilde{c}_1}\right)^{-1}, K$ being a universal constant.

(ii) implies (iii) with $\tilde{c}_2 = \dfrac{2}{\sum_{j=1}^{n} a_j}$.

(iii) implies (i) with $\tilde{c}_1 = \dfrac{\tilde{c}_2}{2}$.

The equivalence between (i) and (iii) was first obtained by Bobkov, Gentil and Ledoux in [5] with the Hamilton-Jacobi approach.

Proof: According to [6] and[39],(i) implies (ii) with $\sum_{j=1}^{n} a_j = 1$ and a depending only on c_1; one can take $\sum_{j=1}^{n} a_j = K \max\left(\tilde{c}_1\sqrt{\tilde{c}_1}\right)^{-1}, K$, where K is a universal constant.

According to (a slightly different version of) Theorem 13 with $p=1$, (ii) implies (iii) (with $\tilde{c}_2 = 2/\sum_{j=1}^{n} a_j$). It remains to prove that (iii) implies (i). This last point is classical; let us simply sketch the proof. The series of transportation-cost inequalities are equivalent to the following property([4],[19]) for all bounded f on $\mathbf{R}^d$,

$$\int e^{Qf}\, d\mu \le e^{\int f\, d\mu},$$

where $Qf\left(\sum_{j=1}^{n} x_j\right) = \inf_{y\in\mathbf{R}^d}\left\{ f\left(\sum_{j=1}^{n} y_j\right) + \tilde{c}_2^{-1}\alpha_1\left(\left|\sum_{j=1}^{n} x_j - \sum_{j=1}^{n} y_j\right|_2\right)\right\}$. Let f be a smooth function and apply the preceding inequality to tf. When $t \longrightarrow 0$, it can be shown that

$$Q(tf)\left(\sum_{j=1}^{n} x_j\right) - tf\left(\sum_{j=1}^{n} x_j\right) = -\frac{\tilde{c}_2 t^2}{4}|\nabla f|_2^2\left(\sum_{j=1}^{n} x_j\right) + 0\left(t^2\right)$$

so $\int e^{Q(tf)} d\mu = 1 + t\int f\, d\mu + \frac{t^2}{2}\int f^2 d\mu \frac{\tilde{c}_2 t^2}{4}\int|\nabla f|_2^2 d\mu + 0\left(t^2\right)$. On the other hand,

$$e^t\int f\, d\mu = 1 + t\int f\, d\mu + \frac{t^2}{2}\int\left(f\, d\mu\right)^2$$

We conclude, that

$$Var(f) \le \frac{\tilde{c}_2}{2}\int|\nabla f|^2\, d\mu$$

which achieves the proof.

The series of the transportation-cost inequalities are closely related to the so called (τ) property introduced by Maurey in [31,39,23,13].

Other sufficient conditions were obtained by Bobkov and Ledoux in [7] with an approach based on the Prekopa-Lcindler inequality, or in [12] by Cordero-Erausquin, Gangbo and Houdre with an optimal transportation method .

$$\log \mu^n \left(L_n \in A\right) \geq \log\left(\sum_{j=1}^{n} v_j\right)^n (B) - \frac{\int_B \log h(x) d\left(\sum_{j=1}^{n} v_j\right)^n}{\left(\sum_{j=1}^{n} v_j\right)^n (B)}$$

Since $\sum_{j=1}^{k} H\left(\left(v_j\right)^n \left|\mu^n\right.\right) = \int \log h\left(\sum_{j=1}^{k} x_j\right) d\left(\sum_{j=1}^{k} v_j\right)^n,$ We conclude that

$$\log \mu^n \left(L_n \in A\right) \geq \log\left(\sum_{j=1}^{n} v_j\right)^n (B)$$

$$-\frac{\sum_{j=1}^{k} H\left(\left(v_j\right)^n \left|\mu^n\right.\right)}{\left(\sum_{j=1}^{n} v_j\right)^n (B)} + \frac{\int_{B^c} \log h\left(\sum_{j=1}^{k} x_j\right) h\left(\sum_{j=1}^{k} x_j\right) d\mu^n}{\left(\sum_{j=1}^{n} v_j\right)^n (B)} \tag{13}$$

But for all $x > 0, \quad x \log x \geq -1/e$, so

$$\frac{\int_{B^c} \log h\left(\sum_{j=1}^{k} x_j\right) h\left(\sum_{j=1}^{k} x_j\right) d\mu^n}{\sum_{j=1}^{n} v_j^{\,n} (B)} \geq \frac{\mu^n (B)}{e\left(\sum_{j=1}^{n} v_j\right)^n (B)} \geq \frac{1}{e\left(\sum_{j=1}^{n} v_j\right)^n (B)} \tag{14}$$

Putting (14) into (13) and using

$$\sum_{j=1}^{k} H\left(\left(v_j\right)^n \left|\mu^n\right.\right) = nH\left(\sum_{j=1}^{n} v_j \left|\mu\right.\right) and \left(\sum_{j=1}^{n} v_j\right)^n (B) = \left(\sum_{j=1}^{n} v_j\right)^n \left(L_n \in A\right)$$

gives the desired inequality.

References:

[1] C. AmJ, S', Blachere, D. Chafa, P.Fougeres, 1. Gentil, F. Malrieu, C. Roberto,, and G; Scheffer. Sur les inegalites de Sabalev logarithmiquescs, volume 10 of Panoramas el Synheses [Panoramas and Syntheses]. Societe Mathematique de France, Paris, 2000.

[2] F. Barthe, P. Cattiaux, and C. Roberto. Interpolated inequalities between exponential and Gauss. ian, Orlicz hypercontractivity and isoperimetry. Rev. Mat. Iberoamericana, 22(3):993-1067, 2006.

[3] W. Beckner. A generalized Poincare inequality for Gaussian measures. Proc. Amer. Math. Sac., 105:397-400, 1989.

[4] S. G. Bobkov and F. Gotze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. J. Funct. Anal., 163(1):1-28, 1999.

[5] S. G. Bobkov, 1. Gentil, and M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. J. Math. Pures Appl. (9), 80(7):669~696, 2001.

[6] S. G. Bobkov an M. Ledoux: Poincare's inequalities and Talagrand's concentration phenomenon for the exponential distnbution. Probab. Theory Related Fields, 107(3):383-400, 1007.

[7] S. G. Bobkov and M. Ledoux. From Brunn-Minkowski to Brascamp-Lieb and to logarithmic. Sobolev inequalities. Geam. Funct. Anal., 10(5):1028-152, 2000.

[8] F. Barthe and C. Roberto. Sobolev inequalities for probability measures on the real line. Studio. Math, 159(3):481-497, 2003.

[9] F. Barthe and C. Roberto. Modified Logarithmic-Sobolev inequalities. Preprint, 2006.

[10] S. G. Bobkov and B. Zegarlinski. Entropy bounds and isoperimetry. Mem. Amer. Math. Soc. 176(829):x+60,2005.

[11] J. A. Cuesta-Albertos and A. Tuero-Diaz. A characterization for the solution of the Monge Kantorovich mass transference problem. Statist. Probab. Lett., 16(2):147-152, 1993.

[12] D. Cordero-Erausquin, W. Gangbo, and C. Houdre. Inequalities for generalized entropy and optimal transportation. In Recent advances in the theary and applicatians af mass transpart, volume .-3.53 of Cantemp. Math., pages 73-94. Amer. Math. Soc., Providence, RI, 2004.

[13] P. Cattiaux and A. Guillin. Talagrand's like quadratic transportation cost inequalities. J. Math. Pures Appl.,86(9):341-361, 2006.

[14] H. Djellout, A. Cuillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. Ann. Probab., 32(3B):2702-2732, 2004.

[15] J. D. Deuschel and D. W. Stroock. Large deviatians, volume 137 of Pure and Applied Mathematics. Academic Press Inc., Boston, MA, 1989.

[16] R. M. Dudley. Real analysis and probability. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1989.

[17] A. Dembo and O. Zeitouni. Large deviatians techniques and applicatians. Secand editian. Applications of Mathematics 38. Springer Verlag, 1998.

[18] 1. Gentil, A. Guillin, and L. Miclo. Modified logarithmic Sobolev inequalities and transportation Winequalities. Probab. Theary Related Fields, 133(3):409-436, 2005.

[19] N. Gozlan ane] C. Leonard. A large deviation approach to some transportation cost inequalities. Probab. Theory Related Fields, 139(1-2):235-283, 2007.

[20] A. Guillin, C. Leonard, L. M. Wu, and N. Yiao. Transportation-information inequalities for Markov processes. Preprint. Availlable on the ArXiv, 2007.

[21] M. Gromov and V.D. Milmall. A topological applicatioll of the isoperimetric inequality. Am.cr. J., Math., 10".843-8,)4, 1983.

[22] N. Gozlan. Characterization of Talagrand's like transportation-cost inequality on the real line. . J Funct. Anal., 50(2):400-425, 2007.

[23] N. Gozlan Poincare' inequalities and dimensions free concentration of measure. preprint, 2008.

[24] M. Ledoux on Talgarands, deviation inequality for product measure. ESAIM Probab. Statsit 1: 63-87.

[25] M. Ledoux. The concentration of measure phenomenon, volume 89, of Mathematical Surveys and '1Monographs. American Mathematical Society, Providence, RI, 2001.

[26] R. Latala and K. Oleszkiewicz. Between Sobolev and Poincare. In geometric aspects of Functional Analy. Lect. Notes Math., 1745:147-168, 2000.

[27] J. Lott and C. Villani. Hamilton-Jacobi semigroup on length spaces and applications. J. Math. Pures Appl. (9), 88(3):219-229, 2007.

[28] R. Lata!a and J. O. WOjtaszczyk. On the infimum convolution inequality. Pre print. Availlable on the ArXiv, 2008.

[29] K. Marton. A simple proof of the blowing-up lemma. IEEE Trans. Inform. Theory, 32(3):445-446, 1986.

[30] K. larton. Bounding $\bar{d}$ -distance by inforational divergence: a method to prove measure concentratlon. Ann. Probab., 24(2).857-866, 1996.

[31] B. Maurey. Some deviation inequalities. Geom. Funct. Anal., 1(2):188-197, 1991.

[32] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev incquality. 1. Funct. Anal., 173(2):361-400, 2000.

[33] P. M. Samson. Concentration of measure inequalities for Markov chains and Φ -mixing processes. Ann. Probab., 28(1):416-461, 2000.

[34] M. Talagrand. The supremum of some canonical processes. Amer. J. Math., 116(2):283-325, 1994.

[35] M. Talgmod. Traonsportation cost for, Gaussian and other product, measure. Geom. Functional analysis 6(3):587-600, 1996.

[36] C. Villani. Topics in Optimal Transportation. Graduate Studies in Mathematics 58. American. 1 latheruatical Society, Providence RI, 2003.

[37] C. Villani. Optimal transport, old and new. In preparation, 2008.
[38] F. Y Wang. Probability distance inequalities on Riemannian manifolds and path spaces. Anal., 206(1):167-190,2004.
[39] N. A. Gozlan, A characterization of dimension free concentration in terms of transportation cost inequalities, preprint(2008).