Concentration of measure and optimal transport

Nathaël Gozlan*

*LAMA Université Paris Est – Marne-la-Vallée

High Dimensional Probability
Banff, 2011

Outline

Talagrand's ineq. $(\mathbf{T}_2)$

$W_2 \leq \sqrt{H}$

Marton & Talagrand (90's)

Log-Sobolev ineq. (LSI)

$H \leq I$

Gross (70's)

Outline

Talagrand's ineq. (T_2)
$W_2 \leq \sqrt{H}$
Marton & Talagrand (90's)

$\Downarrow$

Log-Sobolev ineq. (LSI)
$H \leq I$
Gross (70's)

$\Downarrow$

Gaussian concentration
Milman, Talagrand, Ledoux ...

Outline

Talagrand's ineq. (T_2)
$W_2 \leq \sqrt{H}$
Marton & Talagrand (90's)

$\Longleftarrow$
Otto & Villani (00)

Log-Sobolev ineq. (LSI)
$H \leq I$
Gross (70's)

$\Downarrow$ $\Downarrow$

Gaussian concentration
Milman, Talagrand, Ledoux ...

Outline

Talagrand's ineq. ($\mathbf{T}_2$)
$W_2 \leq \sqrt{H}$
Marton & Talagrand (90's)

$\Longleftarrow$
Otto & Villani (00)

Log-Sobolev ineq. (LSI)
$H \leq I$
Gross (70's)

$\Downarrow$ $\Downarrow$

Gaussian concentration
Milman, Talagrand, Ledoux . . .

$\mathbf{T}_2 \neq$ **LSI** (Cattiaux and Guillin 2005)

Outline

Talagrand's ineq. (T_2) $\Longleftarrow$ **Log-Sobolev ineq. (LSI)**

$W_2 \leq \sqrt{H}$ — Otto & Villani (00) — $H \leq I$

Marton & Talagrand (90's) — Gross (70's)

$\Downarrow$ $\Downarrow$

Gaussian concentration

Milman, Talagrand, Ledoux . . .

$\mathbf{T}_2 \neq$ **LSI** (Cattiaux and Guillin 2005)

Main result of this talk : Talagrand's inequality is equivalent to a modified Log-Sobolev inequality.

Notation

- $(\mathcal{X}, d)$ is a polish metric space (i.e complete and separable)
- μ is a Borel probability measure on $\mathcal{X}$
- $\mathcal{P}(\mathcal{X})$ is the set of all Borel probability measures on $\mathcal{X}$

Talagrand's inequality

Quadratic optimal transport cost

For all $\nu, \mu \in \mathcal{P}_2(\mathcal{X})$

$$\mathcal{T}_2(\nu, \mu) = \inf \left\{ \mathbb{E}[d^2(X, Y)]; \mathcal{L}(X) = \nu \text{ and } \mathcal{L}(Y) = \mu \right\}.$$

The Wasserstein distance W_2 is defined by

$$W_2(\nu, \mu) = \sqrt{\mathcal{T}_2(\nu, \mu)}, \qquad \forall \nu, \mu \in \mathcal{P}_2(\mathcal{X}).$$

Relative entropy / Kullback-Leibler distance

For all $\nu, \mu \in \mathcal{P}(\mathcal{X})$,

$$H(\nu \mid \mu) = \int \log\left(\frac{d\nu}{d\mu}\right) d\nu, \quad \text{if } \nu \ll \mu, \quad \text{and } +\infty \quad \text{otherwise.}$$

Talagrand's inequality

Talagrand's inequality

The probability μ verifies $\mathbf{T}_2(C)$, if

$$\mathcal{T}_2(\nu,\mu) \leq CH(\nu|\mu), \quad \forall \nu \in \mathcal{P}_2(\mathcal{X}).$$

The idea of bounding a transport cost by a function of the relative entropy first appeared in a paper by Marton in 1986.

Talagrand (96) was the first to prove that the standard Gaussian measure on $\mathbf{R}$ satisfies $\mathbf{T}_2(2)$. The constant 2 is optimal.

Talagrand's inequality

Talagrand's inequality

The probability μ verifies $\mathbf{T}_2(C)$, if

$$\mathcal{T}_2(\nu, \mu) \leq CH(\nu|\mu), \quad \forall \nu \in \mathcal{P}_2(\mathcal{X}).$$

The idea of bounding a transport cost by a function of the relative entropy first appeared in a paper by Marton in 1986.

Talagrand (96) was the first to prove that the standard Gaussian measure on $\mathbf{R}$ satisfies $\mathbf{T}_2(2)$. The constant 2 is optimal.

Examples : More generally, if μ is a probability on $\mathbf{R}^k$ with a density of the form e^{-V} with V such that

$$\mathrm{Hess}\, V \geq C\mathrm{I}, \qquad \text{with } C > 0,$$

then μ verifies $\mathbf{T}_2(2/C)$.

Link with Gaussian concentration

Gaussian concentration

The probability μ verifies the Gaussian concentration property $\mathbf{CP}_2(a, t_o)$ for $a, t_o \geq 0$, if for all $A \subset \mathcal{X}$ such that $\mu(A) \geq 1/2$,

$$\mu(A^t) \geq 1 - e^{-a(t-t_o)^2}, \qquad \forall t \geq t_o,$$

where $A^t = \{x \in \mathcal{X}; d(x, A) \leq t\}$.

Link with Gaussian concentration

Gaussian concentration

The probability μ verifies the Gaussian concentration property $\mathbf{CP}_2(a, t_o)$ for $a, t_o \geq 0$, if for all $A \subset \mathcal{X}$ such that $\mu(A) \geq 1/2$,

$$\mu(A^t) \geq 1 - e^{-a(t-t_o)^2}, \qquad \forall t \geq t_o,$$

where $A^t = \{x \in \mathcal{X}; d(x, A) \leq t\}$.

Example.
The standard Gaussian measure on $\mathbf{R}^n$ verifies $\mathbf{CP}_2(1/2, 0)$ for all n.

Link with Gaussian concentration

Gaussian concentration

The probability μ verifies the Gaussian concentration property $\mathbf{CP}_2(a, t_o)$ for $a, t_o \geq 0$, if for all $A \subset \mathcal{X}$ such that $\mu(A) \geq 1/2$,

$$\mu(A^t) \geq 1 - e^{-a(t-t_o)^2}, \qquad \forall t \geq t_o,$$

where $A^t = \{x \in \mathcal{X}; d(x, A) \leq t\}$.

Example.
The standard Gaussian measure on $\mathbf{R}^n$ verifies $\mathbf{CP}_2(1/2, 0)$ for all n.

Proposition

The probability μ verifies the Gaussian concentration property $\mathbf{CP}_2(a, t_o)$ if and only if for all 1-Lipschitz function $f : \mathcal{X} \to \mathbf{R}$, it holds

$$\mu(f > m + t) \leq e^{-a(t-t_o)^2}, \qquad \forall t \geq t_o,$$

where m is a median of f.

Marton's argument

Theorem

If μ verifies $\mathbf{T}_2(C)$, then μ verifies $\mathbf{CP}_2\left(\frac{1}{C}, t_o\right)$, with $t_o = \sqrt{C\log(2)}$

Proof.
Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$ and define $B = \mathcal{X} \setminus A^t$, $t > 0$.

Marton's argument

Theorem

If μ verifies $\mathbf{T}_2(C)$, then μ verifies $\mathbf{CP}_2\left(\frac{1}{C}, t_o\right)$, with $t_o = \sqrt{C\log(2)}$

Proof.
Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$ and define $B = \mathcal{X} \setminus A^t$, $t > 0$.

Marton's argument

Theorem

If μ verifies $\mathbf{T}_2(C)$, then μ verifies $\mathbf{CP}_2\left(\frac{1}{C}, t_o\right)$, with $t_o = \sqrt{C\log(2)}$

Proof.
Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$ and define $B = \mathcal{X} \setminus A^t$, $t > 0$.
Set $d\mu_A = \frac{\mathbb{1}_A}{\mu(A)}\, d\mu$ and $d\mu_B = \frac{\mathbb{1}_B}{\mu(B)}\, d\mu$.

Marton's argument

Theorem

If μ verifies $\mathbf{T}_2(C)$, then μ verifies $\mathbf{CP}_2\left(\frac{1}{C}, t_o\right)$, with $t_o = \sqrt{C\log(2)}$

Proof.
Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$ and define $B = \mathcal{X} \setminus A^t$, $t > 0$.
Set $d\mu_A = \frac{\mathbb{1}_A}{\mu(A)}\, d\mu$ and $d\mu_B = \frac{\mathbb{1}_B}{\mu(B)}\, d\mu$.

$$W_2(\mu_A, \mu_B) \leq W_2(\mu_A, \mu) + W_2(\mu_B, \mu)$$

Marton's argument

Theorem

If μ verifies $\mathbf{T}_2(C)$, then μ verifies $\mathbf{CP}_2\left(\frac{1}{C}, t_o\right)$, with $t_o = \sqrt{C\log(2)}$

Proof.

Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$ and define $B = \mathcal{X} \setminus A^t$, $t > 0$.
Set $d\mu_A = \frac{\mathbb{1}_A}{\mu(A)}\, d\mu$ and $d\mu_B = \frac{\mathbb{1}_B}{\mu(B)}\, d\mu$.

$$\begin{aligned} W_2(\mu_A, \mu_B) &\leq W_2(\mu_A, \mu) + W_2(\mu_B, \mu) \\ &\leq \sqrt{C\,\mathrm{H}(\mu_A \mid \mu)} + \sqrt{C\,\mathrm{H}(\mu_B \mid \mu)} \end{aligned}$$

Marton's argument

Theorem

If μ verifies $\mathbf{T}_2(C)$, then μ verifies $\mathbf{CP}_2\left(\frac{1}{C}, t_o\right)$, with $t_o = \sqrt{C\log(2)}$

Proof.
Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$ and define $B = \mathcal{X} \setminus A^t$, $t > 0$.
Set $d\mu_A = \frac{\mathbb{1}_A}{\mu(A)}\, d\mu$ and $d\mu_B = \frac{\mathbb{1}_B}{\mu(B)}\, d\mu$.

$$\begin{aligned} W_2(\mu_A, \mu_B) &\leq W_2(\mu_A, \mu) + W_2(\mu_B, \mu) \\ &\leq \sqrt{C\,\mathrm{H}(\mu_A \mid \mu)} + \sqrt{C\,\mathrm{H}(\mu_B \mid \mu)} \\ &\leq \sqrt{-C\log(\mu(A))} + \sqrt{-C\log(\mu(B))} \end{aligned}$$

Marton's argument

Theorem

If μ verifies $\mathbf{T}_2(C)$, then μ verifies $\mathbf{CP}_2\left(\frac{1}{C}, t_o\right)$, with $t_o = \sqrt{C\log(2)}$

Proof.
Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$ and define $B = \mathcal{X} \setminus A^t$, $t > 0$.
Set $d\mu_A = \frac{\mathbb{1}_A}{\mu(A)}\, d\mu$ and $d\mu_B = \frac{\mathbb{1}_B}{\mu(B)}\, d\mu$.

$$\begin{aligned} W_2(\mu_A, \mu_B) &\leq W_2(\mu_A, \mu) + W_2(\mu_B, \mu) \\ &\leq \sqrt{C\,\mathrm{H}(\mu_A \mid \mu)} + \sqrt{C\,\mathrm{H}(\mu_B \mid \mu)} \\ &\leq \sqrt{C\log(2)} + \sqrt{-C\log(\mu(B))} \end{aligned}$$

Marton's argument

Theorem

If μ verifies $\mathbf{T}_2(C)$, then μ verifies $\mathbf{CP}_2\left(\frac{1}{C}, t_o\right)$, with $t_o = \sqrt{C\log(2)}$

Proof.
Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$ and define $B = \mathcal{X} \setminus A^t$, $t > 0$.
Set $d\mu_A = \frac{\mathbb{1}_A}{\mu(A)}\, d\mu$ and $d\mu_B = \frac{\mathbb{1}_B}{\mu(B)}\, d\mu$.

$$\begin{aligned} t \leq W_2(\mu_A, \mu_B) &\leq W_2(\mu_A, \mu) + W_2(\mu_B, \mu) \\ &\leq \sqrt{C\,\mathrm{H}(\mu_A \mid \mu)} + \sqrt{C\,\mathrm{H}(\mu_B \mid \mu)} \\ &\leq \sqrt{C\log(2)} + \sqrt{-C\log(\mu(B))} \end{aligned}$$

Theorem

If μ verifies $\mathbf{T}_2(C)$, then μ verifies $\mathbf{CP}_2\left(\frac{1}{C}, t_o\right)$, with $t_o = \sqrt{C\log(2)}$

Proof.
Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$ and define $B = \mathcal{X} \setminus A^t$, $t > 0$.
Set $d\mu_A = \frac{\mathbb{1}_A}{\mu(A)}\, d\mu$ and $d\mu_B = \frac{\mathbb{1}_B}{\mu(B)}\, d\mu$.

$$\begin{aligned} t \leq W_2(\mu_A, \mu_B) &\leq W_2(\mu_A, \mu) + W_2(\mu_B, \mu) \\ &\leq \sqrt{C\,\mathrm{H}(\mu_A \mid \mu)} + \sqrt{C\,\mathrm{H}(\mu_B \mid \mu)} \\ &\leq \sqrt{C\log(2)} + \sqrt{-C\log(\mu(B))} \end{aligned}$$

So,

$$\mu(B) \leq \exp\left(-\frac{1}{C}\left(t - t_o\right)^2\right), \qquad \forall t \geq t_o = \sqrt{C\log(2)}.$$

Dimension free concentration

In fact a much stronger phenomenon occurs

Definition

The probability μ verifies the Gaussian dimension free concentration property $\mathbf{CP}_2^\infty(a, t_o)$, with $a, t_o \geq 0$, if for all $n \in \mathbb{N}^*$, the product measure $\mu^{\otimes n}$ verifies $\mathbf{CP}_2(a, t_o)$ on $\mathcal{X}^n$ equipped with the product distance

$$d_2(x, y) = \left[\sum_{i=1}^{n} d^2(x_i, y_i)\right]^{1/2}, \quad \forall x, y \in \mathcal{X}^n.$$

This phenomenon found many applications in Probability or Analysis in high dimensions.

$\mathbf{T}_2 \Rightarrow \mathbf{CP}_2^\infty$

Theorem (Marton-Talagrand)

If μ verifies $\mathbf{T}_2(C)$, then for all $n \in \mathbb{N}^*$, $\mu^{\otimes n}$ verifies $\mathbf{T}_2(C)$ on $(\mathcal{X}^n, d_2)$.

$\mathbf{T}_2 \Rightarrow \mathbf{CP}_2^\infty$

Theorem (Marton-Talagrand)

If μ verifies $\mathbf{T}_2(C)$, then for all $n \in \mathbb{N}^*$, $\mu^{\otimes n}$ verifies $\mathbf{T}_2(C)$ on $(\mathcal{X}^n, d_2)$.

In particular,

$$\mathbf{T}_2(C) \Rightarrow \mathbf{CP}_2^\infty(1/C, \sqrt{C\log(2)})$$

Other functional approach to dimension free concentration

- Logarithmic Sobolev inequality
  $\rightsquigarrow$ Gaussian concentration - model : $\frac{1}{Z} e^{-|x|^2}\, dx$
  Gross, Herbst, Ledoux, Bobkov-Götze...

Other functional approach to dimension free concentration

- Logarithmic Sobolev inequality

  $\rightsquigarrow$ Gaussian concentration - model : $\frac{1}{Z}e^{-|x|^2}\,dx$

  Gross, Herbst, Ledoux, Bobkov-Götze...

- Poincaré inequality

  $\rightsquigarrow$ Exponential concentration - model : $\frac{1}{Z}e^{-|x|}\,dx$

  Gromov-Milman, Bobkov-Ledoux, Bobkov-Houdré

Other functional approach to dimension free concentration

- Logarithmic Sobolev inequality
  $\rightsquigarrow$ Gaussian concentration - model : $\frac{1}{Z}e^{-|x|^2}\,dx$
  Gross, Herbst, Ledoux, Bobkov-Götze...
- Poincaré inequality
  $\rightsquigarrow$ Exponential concentration - model : $\frac{1}{Z}e^{-|x|}\,dx$
  Gromov-Milman, Bobkov-Ledoux, Bobkov-Houdré
- Latała-Oleszkiewicz inequalities
  $\rightsquigarrow$ between exponential and Gaussian - model : $\frac{1}{Z}e^{-|x|^\alpha}\,dx, \alpha \in [1,2]$
  Beckner, Latała-Oleszkiewicz, Barthe-Cattiaux-Roberto

Other functional approach to dimension free concentration

- Logarithmic Sobolev inequality
  $\rightsquigarrow$ Gaussian concentration - model : $\frac{1}{Z}e^{-|x|^2}\,dx$
  Gross, Herbst, Ledoux, Bobkov-Götze. . .
- Poincaré inequality
  $\rightsquigarrow$ Exponential concentration - model : $\frac{1}{Z}e^{-|x|}\,dx$
  Gromov-Milman, Bobkov-Ledoux, Bobkov-Houdré
- Latała-Oleszkiewicz inequalities
  $\rightsquigarrow$ between exponential and Gaussian - model : $\frac{1}{Z}e^{-|x|^\alpha}\,dx, \alpha \in [1,2]$
  Beckner, Latała-Oleszkiewicz, Barthe-Cattiaux-Roberto
- Modified Logarithmic Sobolev inequalities
  $\rightsquigarrow$ Sub- and super-Gaussian - model : $\frac{1}{Z}e^{-|x|^\alpha}\,dx, \alpha \geq 1$
  Bobkov-Ledoux, Bobkov-Zegarlinski, Gentil-Guillin-Miclo, Barthe-Roberto

Other functional approach to dimension free concentration

- Logarithmic Sobolev inequality
  $\rightsquigarrow$ Gaussian concentration - model : $\frac{1}{Z}e^{-|x|^2}\,dx$
  Gross, Herbst, Ledoux, Bobkov-Götze...
- Poincaré inequality
  $\rightsquigarrow$ Exponential concentration - model : $\frac{1}{Z}e^{-|x|}\,dx$
  Gromov-Milman, Bobkov-Ledoux, Bobkov-Houdré
- Latała-Oleszkiewicz inequalities
  $\rightsquigarrow$ between exponential and Gaussian - model : $\frac{1}{Z}e^{-|x|^\alpha}\,dx, \alpha \in [1,2]$
  Beckner, Latała-Oleszkiewicz, Barthe-Cattiaux-Roberto
- Modified Logarithmic Sobolev inequalities
  $\rightsquigarrow$ Sub- and super-Gaussian - model : $\frac{1}{Z}e^{-|x|^\alpha}\,dx, \alpha \geq 1$
  Bobkov-Ledoux, Bobkov-Zegarlinski, Gentil-Guillin-Miclo, Barthe-Roberto
- Transport inequalities
  Marton, Talagrand, Otto-Villani, Bobkov-Götze, Bobkov-Gentil-Ledoux, Djellout-Guillin-Wu, Wang ...

Other functional approach to dimension free concentration

- Logarithmic Sobolev inequality
  $\rightsquigarrow$ Gaussian concentration - model : $\frac{1}{Z}e^{-|x|^2}\,dx$
  Gross, Herbst, Ledoux, Bobkov-Götze...
- Poincaré inequality
  $\rightsquigarrow$ Exponential concentration - model : $\frac{1}{Z}e^{-|x|}\,dx$
  Gromov-Milman, Bobkov-Ledoux, Bobkov-Houdré
- Latała-Oleszkiewicz inequalities
  $\rightsquigarrow$ between exponential and Gaussian - model : $\frac{1}{Z}e^{-|x|^\alpha}\,dx, \alpha \in [1,2]$
  Beckner, Latała-Oleszkiewicz, Barthe-Cattiaux-Roberto
- Modified Logarithmic Sobolev inequalities
  $\rightsquigarrow$ Sub- and super-Gaussian - model : $\frac{1}{Z}e^{-|x|^\alpha}\,dx, \alpha \geq 1$
  Bobkov-Ledoux, Bobkov-Zegarlinski, Gentil-Guillin-Miclo, Barthe-Roberto
- Transport inequalities
  Marton, Talagrand, Otto-Villani, Bobkov-Götze, Bobkov-Gentil-Ledoux, Djellout-Guillin-Wu, Wang ...
- τ Property
  Maurey, Latała-Wojtaszczyk

Other functional approach to dimension free concentration

- Logarithmic Sobolev inequality
  $\rightsquigarrow$ Gaussian concentration - model : $\frac{1}{Z}e^{-|x|^2}\,dx$
  Gross, Herbst, Ledoux, Bobkov-Götze...
- Poincaré inequality
  $\rightsquigarrow$ Exponential concentration - model : $\frac{1}{Z}e^{-|x|}\,dx$
  Gromov-Milman, Bobkov-Ledoux, Bobkov-Houdré
- Latała-Oleszkiewicz inequalities
  $\rightsquigarrow$ between exponential and Gaussian - model : $\frac{1}{Z}e^{-|x|^\alpha}\,dx, \alpha \in [1,2]$
  Beckner, Latała-Oleszkiewicz, Barthe-Cattiaux-Roberto
- Modified Logarithmic Sobolev inequalities
  $\rightsquigarrow$ Sub- and super-Gaussian - model : $\frac{1}{Z}e^{-|x|^\alpha}\,dx, \alpha \geq 1$
  Bobkov-Ledoux, Bobkov-Zegarlinski, Gentil-Guillin-Miclo, Barthe-Roberto
- Transport inequalities
  Marton, Talagrand, Otto-Villani, Bobkov-Götze, Bobkov-Gentil-Ledoux, Djellout-Guillin-Wu, Wang ...
- τ Property
  Maurey, Latała-Wojtaszczyk

A converse to Marton's Theorem

Theorem (G. 09)

A probability measure μ has the dimension free Gaussian concentration property if and only if it verifies Talagrand's $\mathbf{T}_2$ inequality.

A converse to Marton's Theorem

Theorem (G. 09)

A probability measure μ has the dimension free Gaussian concentration property if and only if it verifies Talagrand's $\mathbf{T}_2$ inequality.
More precisely, μ verifies $\mathbf{T}_2(C)$ if and only if, there is t_o such that

$$\mu^n(A^t) \geq 1 - \exp\left(-\frac{1}{C}(t - t_o)^2\right), \qquad \forall n \geq 1, \quad \forall \mu^n(A) \geq 1/2, \quad \forall t \geq t_o.$$

Comments :

A converse to Marton's Theorem

Theorem (G. 09)

A probability measure μ has the dimension free Gaussian concentration property if and only if it verifies Talagrand's $\mathbf{T}_2$ inequality.
More precisely, μ verifies $\mathbf{T}_2(C)$ if and only if, there is t_o such that

$$\mu^n(A^t) \geq 1 - \exp\left(-\frac{1}{C}(t - t_o)^2\right), \qquad \forall n \geq 1, \quad \forall \mu^n(A) \geq 1/2, \quad \forall t \geq t_o.$$

Comments :

- $\mathbf{T}_2 \Rightarrow$ concentration : Marton (86), Talagrand (96).

A converse to Marton's Theorem

Theorem (G. 09)

A probability measure μ has the dimension free Gaussian concentration property if and only if it verifies Talagrand's $\mathbf{T}_2$ inequality.
More precisely, μ verifies $\mathbf{T}_2(C)$ if and only if, there is t_o such that

$$\mu^n(A^t) \geq 1 - \exp\left(-\frac{1}{C}(t - t_o)^2\right), \qquad \forall n \geq 1, \quad \forall \mu^n(A) \geq 1/2, \quad \forall t \geq t_o.$$

Comments :

- $\mathbf{T}_2 \Rightarrow$ concentration : Marton (86), Talagrand (96).
- Concentration $\Rightarrow \mathbf{T}_2$: G. (09)

A converse to Marton's Theorem

Theorem (G. 09)

A probability measure μ has the dimension free Gaussian concentration property if and only if it verifies Talagrand's $\mathbf{T}_2$ inequality.
More precisely, μ verifies $\mathbf{T}_2(C)$ if and only if, there is t_o such that

$$\mu^n(A^t) \geq 1 - \exp\left(-\frac{1}{C}(t - t_o)^2\right), \qquad \forall n \geq 1, \quad \forall \mu^n(A) \geq 1/2, \quad \forall t \geq t_o.$$

Comments :

- $\mathbf{T}_2 \Rightarrow$ concentration : Marton (86), Talagrand (96).
- Concentration $\Rightarrow \mathbf{T}_2$: G. (09)

$\rightsquigarrow$ Gozlan, *A characterization of dimension free concentration in terms of transport inequalities*, AOP (2009).

A converse to Marton's Theorem

Theorem (G. 09)

A probability measure μ has the dimension free Gaussian concentration property if and only if it verifies Talagrand's $\mathbf{T}_2$ inequality.
More precisely, μ verifies $\mathbf{T}_2(C)$ if and only if, there is t_o such that

$$\mu^n(A^t) \geq 1 - \exp\left(-\frac{1}{C}(t-t_o)^2\right), \qquad \forall n \geq 1, \quad \forall \mu^n(A) \geq 1/2, \quad \forall t \geq t_o.$$

Comments :

- $\mathbf{T}_2 \Rightarrow$ concentration : Marton (86), Talagrand (96).
- Concentration $\Rightarrow \mathbf{T}_2$: G. (09)

$\rightsquigarrow$ Gozlan, *A characterization of dimension free concentration in terms of transport inequalities*, AOP (2009).

- Use of Large deviations : G. and Léonard (07)

A converse to Marton's Theorem

Theorem (G. 09)

A probability measure μ has the dimension free Gaussian concentration property if and only if it verifies Talagrand's $\mathbf{T}_2$ inequality.
More precisely, μ verifies $\mathbf{T}_2(C)$ if and only if, there is t_o such that

$$\mu^n(A^t) \geq 1 - \exp\left(-\frac{1}{C}(t - t_o)^2\right), \qquad \forall n \geq 1, \quad \forall \mu^n(A) \geq 1/2, \quad \forall t \geq t_o.$$

Comments :

- $\mathbf{T}_2 \Rightarrow$ concentration : Marton (86), Talagrand (96).
- Concentration $\Rightarrow \mathbf{T}_2$: G. (09)

$\rightsquigarrow$ Gozlan, *A characterization of dimension free concentration in terms of transport inequalities*, AOP (2009).

- Use of Large deviations : G. and Léonard (07)

$\rightsquigarrow$ G. and Léonard, *A large deviation approach to some transportation cost inequalities*, PTRF (2007).

Sketch of proof.

The idea is to estimate the probability of the following rare event

$$\mathbb{P}(W_2(L_n,\mu) > t), \qquad t \geq 0,$$

where $L_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ and X_i an i.i.d sequence of law μ.

Sketch of proof.

The idea is to estimate the probability of the following rare event

$$\mathbb{P}(W_2(L_n,\mu) > t), \qquad t \geq 0,$$

where $L_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ and X_i an i.i.d sequence of law μ.

- A first estimate (from above) is given by $\mathbf{CP}_2^\infty(a, t_o)$:

$$\mathbb{P}(W_2(L_n,\mu) > t) \leq e^{-nat^2} \quad \text{(roughly speaking)}$$

Sketch of proof.

The idea is to estimate the probability of the following rare event

$$\mathbb{P}(W_2(L_n, \mu) > t), \qquad t \geq 0,$$

where $L_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$ and X_i an i.i.d sequence of law μ.

- A first estimate (from above) is given by $\mathbf{CP}_2^\infty(a, t_o)$:

$$\mathbb{P}(W_2(L_n, \mu) > t) \leq e^{-nat^2} \quad \text{(roughly speaking)}$$

Here we use the crucial fact that $x \mapsto W_2(L_n^x, \mu)$ is $1/\sqrt{n}$ Lipschitz.

Sketch of proof.

The idea is to estimate the probability of the following rare event

$$\mathbb{P}(W_2(L_n, \mu) > t), \qquad t \geq 0,$$

where $L_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ and X_i an i.i.d sequence of law μ.

- A first estimate (from above) is given by $\mathbf{CP}_2^\infty(a, t_o)$:

$$\mathbb{P}(W_2(L_n, \mu) > t) \leq e^{-nat^2} \quad \text{(roughly speaking)}$$

Here we use the crucial fact that $x \mapsto W_2(L_n^x, \mu)$ is $1/\sqrt{n}$ Lipschitz.

- A second estimate (from below) is given by Sanov's Theorem:

$$-\inf\{H(\nu|\mu); W_2(\nu, \mu) > t\} \leq \liminf_{n \to +\infty} \frac{1}{n} \log \mathbb{P}(W_2(L_n, \mu) > t)$$

Sketch of proof.

The idea is to estimate the probability of the following rare event

$$\mathbb{P}(W_2(L_n, \mu) > t), \qquad t \geq 0,$$

where $L_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ and X_i an i.i.d sequence of law μ.

- A first estimate (from above) is given by $\mathbf{CP}_2^\infty(a, t_o)$:

$$\mathbb{P}(W_2(L_n, \mu) > t) \leq e^{-nat^2} \quad \text{(roughly speaking)}$$

Here we use the crucial fact that $x \mapsto W_2(L_n^x, \mu)$ is $1/\sqrt{n}$ Lipschitz.

- A second estimate (from below) is given by Sanov's Theorem:

$$-\inf\{H(\nu|\mu); W_2(\nu, \mu) > t\} \leq \liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(W_2(L_n, \mu) > t)$$

Comparing these two estimates gives Talagrand's inequality:

$$W_2^2(\nu, \mu) \leq \frac{1}{a} H(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Links with the Log-Sobolev inequality - Otto-Villani Theorem

Definition

The probability μ verifies the Log-Sobolev inequality $\mathbf{LSI}(C)$ if

$$\mathrm{Ent}_\mu(f^2) \leq C \int |\nabla f|^2 \, d\mu,$$

for all locally Lipschitz f, where

$$\mathrm{Ent}_\mu(g) = \int g \log(g) \, d\mu - \left(\int g \, d\mu \right) \cdot \log \left(\int g \, d\mu \right), \qquad \forall g \geq 0.$$

The following result is due to Otto and Villani (2000). Bobkov-Gentil-Ledoux proposed another proof in 2001.

Theorem

Let $(\mathcal{X}, d)$ be a complete, connected Riemannian manifold equipped with its geodesic distance, and μ be an absolutely continuous probability measure on $\mathcal{X}$. If μ verifies $\mathbf{LSI}(C)$ then it verifies $\mathbf{T}_2(C)$.

A proof based on concentration

To prove the implication

$$\mathbf{LSI}(C) \Rightarrow \mathbf{T}_2(C)$$

A proof based on concentration

To prove the implication

$$\mathbf{LSI}(C) \Rightarrow \mathbf{T}_2(C)$$

it is enough to prove that

$$\mathbf{LSI}(C) \Rightarrow \mathbf{CP}_2^\infty(1/C, t_o), \quad \text{for some } t_o.$$

A proof based on concentration

To prove the implication

$$\mathbf{LSI}(C) \Rightarrow \mathbf{T}_2(C)$$

it is enough to prove that

$$\mathbf{LSI}(C) \Rightarrow \mathbf{CP}_2^\infty(1/C, t_o), \quad \text{for some } t_o.$$

This is a well known property of the log-Sobolev inequality due to Herbst.

A proof based on concentration

To prove the implication

$$\mathbf{LSI}(C) \Rightarrow \mathbf{T}_2(C)$$

it is enough to prove that

$$\mathbf{LSI}(C) \Rightarrow \mathbf{CP}_2^\infty(1/C, t_o), \quad \text{for some } t_o.$$

This is a well known property of the log-Sobolev inequality due to Herbst.

Sketch of proof. Apply $\mathbf{LSI}(C)$

$$\mathrm{Ent}_\mu\left(g^2\right) \leq C \int |\nabla g|^2 \, d\mu$$

to $g = e^{\lambda f/2}$ with f a centered 1-Lipschitz function.

A proof based on concentration

To prove the implication

$$\mathbf{LSI}(C) \Rightarrow \mathbf{T}_2(C)$$

it is enough to prove that

$$\mathbf{LSI}(C) \Rightarrow \mathbf{CP}_2^\infty(1/C, t_o), \quad \text{for some } t_o.$$

This is a well known property of the log-Sobolev inequality due to Herbst.

Sketch of proof. Apply $\mathbf{LSI}(C)$

$$\mathrm{Ent}_\mu\left(g^2\right) \leq C \int |\nabla g|^2 \, d\mu$$

to $g = e^{\lambda f/2}$ with f a centered 1-Lipschitz function. After some elementary calculations, this yields the following bound:

$$\int e^{\lambda f} \, d\mu \ \leq e^{C\lambda^2/4}, \qquad \forall \lambda \geq 0.$$

A proof based on concentration

To prove the implication

$$\mathbf{LSI}(C) \Rightarrow \mathbf{T}_2(C)$$

it is enough to prove that

$$\mathbf{LSI}(C) \Rightarrow \mathbf{CP}_2^{\infty}(1/C, t_o), \quad \text{for some } t_o.$$

This is a well known property of the log-Sobolev inequality due to Herbst.

Sketch of proof. Apply $\mathbf{LSI}(C)$

$$\mathrm{Ent}_\mu\ (g^2) \leq C \int |\nabla g|^2 \, d\mu$$

to $g = e^{\lambda f/2}$ with f a centered 1-Lipschitz function. After some elementary calculations, this yields the following bound:

$$\int e^{\lambda f} \, d\mu \ \leq e^{C\lambda^2/4}, \qquad \forall \lambda \geq 0.$$

This implies

$$\mu\ (f \geq t) \leq e^{-t^2/C}, \qquad \forall t \geq 0.$$

A proof based on concentration

To prove the implication

$$\mathbf{LSI}(C) \Rightarrow \mathbf{T}_2(C)$$

it is enough to prove that

$$\mathbf{LSI}(C) \Rightarrow \mathbf{CP}_2^{\infty}(1/C, t_o), \quad \text{for some } t_o.$$

This is a well known property of the log-Sobolev inequality due to Herbst.

Sketch of proof. Apply $\mathbf{LSI}(C)$

$$\mathrm{Ent}_{\mu^n}(g^2) \leq C \int |\nabla g|^2 \, d\mu^n$$

to $g = e^{\lambda f/2}$ with f a centered 1-Lipschitz function. After some elementary calculations, this yields the following bound:

$$\int e^{\lambda f} \, d\mu^n \leq e^{C\lambda^2/4}, \qquad \forall \lambda \geq 0.$$

This implies

$$\mu^n(f \geq t) \leq e^{-t^2/C}, \qquad \forall t \geq 0.$$

Inf-Convolution Log-Sobolev inequality

LSI:

$$\mathrm{Ent}_\mu(g^2) \leq C \int |\nabla g|^2 \, d\mu, \quad \forall g.$$

New gradient:

Inf-Convolution Log-Sobolev inequality

LSI:

$$\mathrm{Ent}_\mu(e^f) \leq \frac{C}{4} \int |\nabla f|^2 e^f \, d\mu, \quad \forall f.$$

New gradient:

Inf-Convolution Log-Sobolev inequality

LSI:

$$\mathrm{Ent}_\mu(e^f) \leq \frac{C}{4} \int |\nabla f|^2 e^f \, d\mu, \quad \forall f.$$

New gradient:

Inf-Convolution Log-Sobolev inequality

LSI:

$$\mathrm{Ent}_\mu(e^f) \leq \frac{C}{4} \int |\nabla f|^2 e^f \, d\mu, \quad \forall f.$$

New gradient:

$$|\nabla f|^2(x) \quad \longleftrightarrow \quad \begin{cases} f(x) - Q_\lambda f(x) \\ \text{where} \\ Q_\lambda f(x) = \inf_{y \in \mathcal{X}} \{ f(y) + \frac{1}{2\lambda} d^2(x,y) \} \end{cases}$$

Inf-Convolution Log-Sobolev inequality

LSI:

$$\mathrm{Ent}_\mu(e^f) \leq \frac{C}{4} \int |\nabla f|^2 e^f \, d\mu, \quad \forall f.$$

New gradient:

$$|\nabla f|^2(x) \quad \longleftrightarrow \quad \left\{ \begin{array}{l} f(x) - Q_\lambda f(x) \\ \text{where} \\ Q_\lambda f(x) = \inf_{y \in \mathcal{X}} \{ f(y) + \frac{1}{2\lambda} d^2(x,y) \} \end{array} \right.$$

Definition

The probability μ is said to verify the inf-convolution log-Sobolev inequality with constants A and λ if

$$\mathrm{Ent}_\mu(e^f) \leq A \int (f - Q_\lambda f) e^f \, d\mu, \qquad \forall f.$$

$\rightsquigarrow$ joint work with C. Roberto and P-M Samson.

Inf-Convolution Log-Sobolev inequality

Theorem (G-Roberto-Samson (2011))

Let μ be a probability on some polish space $(\mathcal{X}, d)$; the following statements are equivalent:

(1) There is some constant C such that μ verifies $\mathbf{T}_2(C)$.
(2) There are constants $A, \lambda > 0$ such that μ verifies the following inf-convolution log-Sobolev inequality:

$$\mathrm{Ent}_\mu(e^f) \leq A \int (f - Q_\lambda f) e^f \, d\mu, \qquad \forall f,$$

where $Q_\lambda f(x) = \inf_{y \in \mathcal{X}} \left\{ f(y) + \lambda d^2(x, y) \right\}$.

Inf-Convolution Log-Sobolev inequality

Theorem (G-Roberto-Samson (2011))

Let μ be a probability on some polish space $(\mathcal{X}, d)$; the following statements are equivalent:

(1) There is some constant C such that μ verifies $\mathbf{T}_2(C)$.
(2) There are constants $A, \lambda > 0$ such that μ verifies the following inf-convolution log-Sobolev inequality:

$$\mathrm{Ent}_\mu(e^f) \leq A \int (f - Q_\lambda f) e^f \, d\mu, \qquad \forall f,$$

where $Q_\lambda f(x) = \inf_{y \in \mathcal{X}} \left\{ f(y) + \lambda d^2(x,y) \right\}$.

There is a precise relation between C, A and λ.

Inf-Convolution Log-Sobolev inequality

Theorem (G-Roberto-Samson (2011))

Let μ be a probability on some polish space $(\mathcal{X}, d)$; the following statements are equivalent:

(1) There is some constant C such that μ verifies $\mathbf{T}_2(C)$.
(2) There are constants $A, \lambda > 0$ such that μ verifies the following inf-convolution log-Sobolev inequality:

$$\mathrm{Ent}_\mu(e^f) \leq A \int (f - Q_\lambda f) e^f \, d\mu, \qquad \forall f,$$

where $Q_\lambda f(x) = \inf_{y \in \mathcal{X}} \left\{ f(y) + \lambda d^2(x, y) \right\}$.

There is a precise relation between C, A and λ.

The proof of (2) $\Rightarrow$ (1) uses the same arguments as the proof of Otto-Villani Theorem (tensorization + sophisticated Herbst argument)

Perturbation of Talagrand's inequality.

Theorem (GRS 2011)

If μ verifies $\mathbf{T}_2(C)$ and if $\bar{\mu}$ is a probability such that

$$\bar{\mu}(dx) = e^{\varphi(x)}\,\mu(dx),$$

where $\varphi : \mathcal{X} \to \mathbf{R}$ is a bounded function, then $\bar{\mu}$ verifies $\mathbf{T}_2(\overline{C})$, with

$$\overline{C} = \kappa e^{\mathrm{Osc}\,(\varphi)} C, \qquad \text{where } \mathrm{Osc}\,(\varphi) = \sup \varphi - \inf \varphi.$$

Perturbation of Talagrand's inequality.

Theorem (GRS 2011)

If μ verifies $\mathbf{T}_2(C)$ and if $\bar{\mu}$ is a probability such that

$$\bar{\mu}(dx) = e^{\varphi(x)}\,\mu(dx),$$

where $\varphi : \mathcal{X} \to \mathbf{R}$ is a bounded function, then $\bar{\mu}$ verifies $\mathbf{T}_2(\overline{C})$, with

$$\overline{C} = \kappa e^{\mathrm{Osc}\,(\varphi)} C, \qquad \text{where } \mathrm{Osc}\,(\varphi) = \sup \varphi - \inf \varphi.$$

The same conclusion holds (with $\kappa = 1$) for the Log-Sobolev or Poincaré inequality and all their variants (Holley-Stroock perturbation lemma).

Perturbation of Talagrand's inequality.

Theorem (GRS 2011)

If μ verifies $\mathbf{T}_2(C)$ and if $\bar{\mu}$ is a probability such that

$$\bar{\mu}(dx) = e^{\varphi(x)}\,\mu(dx),$$

where $\varphi : \mathcal{X} \to \mathbf{R}$ is a bounded function, then $\bar{\mu}$ verifies $\mathbf{T}_2(\overline{C})$, with

$$\overline{C} = \kappa e^{\mathrm{Osc}\,(\varphi)} C, \qquad \text{where } \mathrm{Osc}\,(\varphi) = \sup \varphi - \inf \varphi.$$

The same conclusion holds (with $\kappa = 1$) for the Log-Sobolev or Poincaré inequality and all their variants (Holley-Stroock perturbation lemma).

Proof of the Theorem. We use the equivalence between $\mathbf{T}_2$ and the inf-convolution log-Sobolev inequality and we apply the Holley-Stroock perturbation lemma in its classical form.

Thank you for your attention !

Link with the usual gradient for semi convex functions

If f is K-semi convex on $\mathbf{R}^k$, then, by definition,

$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x) - \frac{K}{2}|y - x|_2^2$$

Link with the usual gradient for semi convex functions

If f is K-semi convex on $\mathbf{R}^k$, then, by definition,

$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x) - \frac{K}{2}|y - x|_2^2$$

Therefore, if $\lambda < 1/K$

$$Q_\lambda f(x) = \inf_y \left\{ f(y) + \frac{1}{2\lambda}|x - y|_2^2 \right\}$$

Link with the usual gradient for semi convex functions

If f is K-semi convex on $\mathbf{R}^k$, then, by definition,

$$f(y) \geq f(x) + \nabla f(x) \cdot (y-x) - \frac{K}{2}|y-x|_2^2$$

Therefore, if $\lambda < 1/K$

$$\begin{aligned} Q_\lambda f(x) &= \inf_y \left\{ f(y) + \frac{1}{2\lambda}|x-y|_2^2 \right\} \\ &\geq \inf_y \left\{ f(x) + \nabla f(x) \cdot (y-x) + \frac{1}{2}\left(\frac{1}{\lambda} - K\right)|x-y|_2^2 \right\} \end{aligned}$$

Link with the usual gradient for semi convex functions

If f is K-semi convex on $\mathbf{R}^k$, then, by definition,

$$f(y) \geq f(x) + \nabla f(x) \cdot (y-x) - \frac{K}{2}|y-x|_2^2$$

Therefore, if $\lambda < 1/K$

$$\begin{aligned} Q_\lambda f(x) &= \inf_y \left\{ f(y) + \frac{1}{2\lambda}|x-y|_2^2 \right\} \\ &\geq \inf_y \left\{ f(x) + \nabla f(x) \cdot (y-x) + \frac{1}{2}\left(\frac{1}{\lambda} - K\right)|x-y|_2^2 \right\} \\ &= f(x) - \frac{1}{2\left(\frac{1}{\lambda} - K\right)} |\nabla f|^2(x). \end{aligned}$$

Link with the usual gradient for semi convex functions

If f is K-semi convex on $\mathbf{R}^k$, then, by definition,

$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x) - \frac{K}{2}|y - x|_2^2$$

Therefore, if $\lambda < 1/K$

$$\begin{aligned} Q_\lambda f(x) &= \inf_y \left\{ f(y) + \frac{1}{2\lambda}|x - y|_2^2 \right\} \\ &\geq \inf_y \left\{ f(x) + \nabla f(x) \cdot (y - x) + \frac{1}{2}\left(\frac{1}{\lambda} - K\right)|x - y|_2^2 \right\} \\ &= f(x) - \frac{1}{2\left(\frac{1}{\lambda} - K\right)}|\nabla f|^2(x). \end{aligned}$$

So,

$$f(x) - Q_\lambda f(x) \leq \frac{1}{2\left(\frac{1}{\lambda} - K\right)}|\nabla f|^2(x)$$

and so the inf convolution log-Sobolev inequality implies a restricted log-Sobolev inequality...

Restricted Log-Sobolev inequality

Theorem (GRS-2010)

Let μ be a probability measure on $\mathbf{R}^k$; the following propositions are equivalent:
(1) There is $C_1 > 0$ such that μ verifies $\mathbf{T}_2(C_1)$.
(2) There is $C_2 > 0$ such that for all $0 \leq K < \frac{2}{C}$ and all K-semi-convex $f : \mathbf{R}^k \to \mathbf{R}$,

$$\mathrm{Ent}_\mu(e^f) \leq \frac{C}{(1 - \frac{KC}{2})^2} \int |\nabla f|^2 e^f \, d\mu.$$

Restricted Log-Sobolev inequality

Theorem (GRS-2010)

Let μ be a probability measure on $\mathbf{R}^k$; the following propositions are equivalent:
(1) There is $C_1 > 0$ such that μ verifies $\mathbf{T}_2(C_1)$.
(2) There is $C_2 > 0$ such that for all $0 \leq K < \frac{2}{C}$ and all K-semi-convex $f : \mathbf{R}^k \to \mathbf{R}$,

$$\mathrm{Ent}_\mu(e^f) \leq \frac{C}{\left(1 - \frac{KC}{2}\right)^2} \int |\nabla f|^2 e^f \, d\mu.$$

The constants C_1 and C_2 are related in the the following way:

$$(1) \Rightarrow (2) \text{ with } C_2 = C_1.$$

$$(2) \Rightarrow (1) \text{ with } C_1 = 9C_2.$$

Restricted Log-Sobolev inequality

Theorem (GRS-2010)

Let μ be a probability measure on $\mathbf{R}^k$; the following propositions are equivalent:
(1) There is $C_1 > 0$ such that μ verifies $\mathbf{T}_2(C_1)$.
(2) There is $C_2 > 0$ such that for all $0 \leq K < \frac{2}{C}$ and all K-semi-convex $f : \mathbf{R}^k \to \mathbf{R}$,

$$\mathrm{Ent}_\mu(e^f) \leq \frac{C}{(1 - \frac{KC}{2})^2} \int |\nabla f|^2 e^f \, d\mu.$$

The constants C_1 and C_2 are related in the the following way:

$$(1) \Rightarrow (2) \text{ with } C_2 = C_1.$$

$$(2) \Rightarrow (1) \text{ with } C_1 = 9C_2.$$

$\rightsquigarrow$ G., Roberto, Samson *A new characterization of Talagrand's transport-entropy inequalities and applications*, AOP (2011).

Proof - $\mathbf{T}_2 \Rightarrow$ inf-convolution **LSI**

Let f be a function on $\mathcal{X}$ and define

$$d\nu_f = \frac{e^f}{\int e^f \, d\mu} d\mu.$$

Proof - $\mathbf{T}_2 \Rightarrow$ inf-convolution **LSI**

Let f be a function on $\mathcal{X}$ and define

$$d\nu_f = \frac{e^f}{\int e^f\, d\mu} d\mu.$$

Then

$$\mathrm{H}(\nu_f \mid \mu) = \int \log\left(\frac{e^f}{\int e^f\, d\mu}\right) d\nu_f$$

Proof - $\mathbf{T}_2 \Rightarrow$ inf-convolution **LSI**

Let f be a function on $\mathcal{X}$ and define

$$d\nu_f = \frac{e^f}{\int e^f \, d\mu} d\mu.$$

Then

$$\begin{aligned} \mathrm{H}(\nu_f \mid \mu) &= \int \log \left(\frac{e^f}{\int e^f \, d\mu} \right) \, d\nu_f \\ &= \int f \, d\nu_f - \log \int e^f \, d\mu \end{aligned}$$

Proof - $\mathbf{T}_2 \Rightarrow$ inf-convolution **LSI**

Let f be a function on $\mathcal{X}$ and define

$$d\nu_f = \frac{e^f}{\int e^f \, d\mu} d\mu.$$

Then

$$\begin{aligned} \mathrm{H}(\nu_f \mid \mu) &= \int \log \left(\frac{e^f}{\int e^f \, d\mu} \right) d\nu_f \\ &= \int f \, d\nu_f - \log \int e^f \, d\mu \\ &\leq \int f \, d\nu_f - \int f \, d\mu, \qquad \text{(Jensen)} \end{aligned}$$

Let f be a function on $\mathcal{X}$ and define

$$d\nu_f = \frac{e^f}{\int e^f \, d\mu} d\mu.$$

Then

$$\begin{aligned} H(\nu_f \mid \mu) &= \int \log \left(\frac{e^f}{\int e^f \, d\mu} \right) d\nu_f \\ &= \int f \, d\nu_f - \log \int e^f \, d\mu \\ &\leq \int f \, d\nu_f - \int f \, d\mu, \qquad \text{(Jensen)} \end{aligned}$$

If π is an optimal transport plan between ν_f and μ, then

$$H(\nu_f \mid \mu) \leq \iint f(x) - f(y) \, \pi(dxdy).$$

Proof - $\mathbf{T}_2 \Rightarrow$ inf-convolution **LSI**

$$\mathrm{H}(\nu_f \mid \mu) \leq \iint f(x) - f(y)\, \pi(dxdy).$$

Proof - $\mathbf{T}_2 \Rightarrow$ inf-convolution **LSI**

$$\mathrm{H}(\nu_f \mid \mu) \leq \iint f(x) - f(y)\,\pi(dxdy).$$

But

$$f(y) \geq Q_\lambda f(x) - \frac{1}{2\lambda} d^2(x,y),$$

so

$$\mathrm{H}(\nu_f \mid \mu) \leq \iint f(x) - Q_\lambda f(x)\,\pi(dxdy) + \frac{1}{2\lambda}\iint d^2(x,y)\,\pi(dxdy)$$

Proof - $\mathbf{T}_2 \Rightarrow$ inf-convolution **LSI**

$$\mathrm{H}(\nu_f \mid \mu) \leq \iint f(x) - f(y)\,\pi(dxdy).$$

But

$$f(y) \geq Q_\lambda f(x) - \frac{1}{2\lambda} d^2(x,y),$$

so

$$\begin{aligned}\mathrm{H}(\nu_f \mid \mu) &\leq \iint f(x) - Q_\lambda f(x)\,\pi(dxdy) + \frac{1}{2\lambda}\iint d^2(x,y)\,\pi(dxdy)\\ &= \int f(x) - Q_\lambda f(x)\,\nu_f(dx) + \frac{1}{2\lambda}\mathcal{T}_2(\nu_f,\mu)\end{aligned}$$

Proof - $\mathbf{T}_2 \Rightarrow$ inf-convolution **LSI**

$$\mathrm{H}(\nu_f \mid \mu) \leq \iint f(x) - f(y)\,\pi(dxdy).$$

But

$$f(y) \geq Q_\lambda f(x) - \frac{1}{2\lambda} d^2(x,y),$$

so

$$\begin{aligned}\mathrm{H}(\nu_f \mid \mu) &\leq \iint f(x) - Q_\lambda f(x)\,\pi(dxdy) + \frac{1}{2\lambda}\iint d^2(x,y)\,\pi(dxdy) \\ &= \int f(x) - Q_\lambda f(x)\,\nu_f(dx) + \frac{1}{2\lambda}\mathcal{T}_2(\nu_f,\mu)\end{aligned}$$

Since μ verifies $\mathbf{T}_2(C)$, one gets

$$\mathrm{H}(\nu_f \mid \mu) \leq \int f(x) - Q_\lambda f(x)\,\nu_f(dx) + \frac{C}{2\lambda}\mathrm{H}(\nu_f \mid \mu).$$

Proof - $\mathbf{T}_2 \Rightarrow$ inf-convolution **LSI**

$$\mathrm{H}(\nu_f \mid \mu) \leq \iint f(x) - f(y)\,\pi(dxdy).$$

But

$$f(y) \geq Q_\lambda f(x) - \frac{1}{2\lambda} d^2(x,y),$$

so

$$\begin{aligned}\mathrm{H}(\nu_f \mid \mu) &\leq \iint f(x) - Q_\lambda f(x)\,\pi(dxdy) + \frac{1}{2\lambda}\iint d^2(x,y)\,\pi(dxdy)\\ &= \int f(x) - Q_\lambda f(x)\,\nu_f(dx) + \frac{1}{2\lambda}\mathcal{T}_2(\nu_f,\mu)\end{aligned}$$

Since μ verifies $\mathbf{T}_2(C)$, one gets

$$\mathrm{H}(\nu_f \mid \mu) \leq \int f(x) - Q_\lambda f(x)\,\nu_f(dx) + \frac{C}{2\lambda}\mathrm{H}(\nu_f \mid \mu).$$

and so for all $\lambda > C/2$, it holds

$$\mathrm{H}(\nu_f \mid \mu) \leq \frac{1}{1 - \frac{C}{2\lambda}} \int f(x) - Q_\lambda f(x)\,\nu_f(dx).$$