Stein's method, logarithmic Sobolev and transport inequalities

M. Ledoux

University of Toulouse, France

and Institut Universitaire de France

Stein's method, logarithmic Sobolev and transport inequalities

joint work with

I. Nourdin, G. Peccati (Luxemburg)

links between

Stein's approximation method

functional and transport inequalities

Stein's method

normal approximation

central limit theorems

rates of convergence,
Edgeworth expansions, strong approximations

Poisson approximation (L. Chen), other distributions

statistics, machine learning, engineering...

logarithmic Sobolev inequalities

quantum field theory, hypercontractivity

infinite dimensional Wiener analysis

convergence to equilibrium of kinetic equations,
Markov chains, interacting particle systems

measure concentration (entropy method)

information theory, optimal transport (C. Villani)

Perelman's Ricci flow

Plan

1. logarithmic Sobolev inequalities

2. Stein's method

3. links and applications

L. Gross (1975)

A. Stam (1959), P. Federbush (1969)

L. Gross (1975)

γ standard Gaussian (probability) measure on $\mathbb{R}^d$

$$d\gamma(x) = e^{-|x|^2/2} \frac{dx}{(2\pi)^{d/2}}$$

$h > 0$ smooth, $\int_{\mathbb{R}^d} h \, d\gamma = 1$

entropy $\displaystyle\int_{\mathbb{R}^d} h \log h \, d\gamma \leq \frac{1}{2} \int_{\mathbb{R}^d} \frac{|\nabla h|^2}{h} \, d\gamma$ Fisher information

$$\int_{\mathbb{R}^d} h \log h \, d\gamma \;\leq\; \frac{1}{2} \int_{\mathbb{R}^d} \frac{|\nabla h|^2}{h} \, d\gamma, \qquad \int_{\mathbb{R}^d} h \, d\gamma = 1$$

$$\nu \ll \gamma \qquad d\nu \;=\; h \, d\gamma$$

$$\mathrm{H}\big(\nu \,|\, \gamma\big) \;\leq\; \frac{1}{2} \, \mathrm{I}\big(\nu \,|\, \gamma\big)$$

(relative) H-entropy $\qquad \mathrm{H}\big(\nu \,|\, \gamma\big) \;=\; \int_{\mathbb{R}^d} h \log h \, d\gamma$

(relative) Fisher Information $\qquad \mathrm{I}\big(\nu \,|\, \gamma\big) \;=\; \int_{\mathbb{R}^d} \frac{|\nabla h|^2}{h} \, d\gamma$

logarithmic Sobolev inequalities

hypercontractivity (integrability of Wiener chaos)

concentration inequalities (entropy method)

links with transportation cost inequalities

Herbst argument (1975)

$$\int_{\mathbb{R}^d} h \log h \, d\gamma \;\leq\; \frac{1}{2} \int_{\mathbb{R}^d} \frac{|\nabla h|^2}{h} \, d\gamma, \qquad \int_{\mathbb{R}^d} h \, d\gamma = 1$$

$\varphi : \mathbb{R}^d \to \mathbb{R}$ 1-Lipschitz, $\quad \int_{\mathbb{R}^d} \varphi \, d\gamma = 0$

$$h \;=\; \frac{e^{\lambda\varphi}}{\int_{\mathbb{R}^d} e^{\lambda\varphi} d\gamma}\,, \qquad \lambda \in \mathbb{R}$$

$$Z(\lambda) \;=\; \int_{\mathbb{R}^d} e^{\lambda\varphi} d\gamma$$

Herbst argument (1975)

$$\lambda Z'(\lambda) - Z(\lambda) \log Z(\lambda) \ \leq \ \frac{\lambda^2}{2} Z(\lambda)$$

integrate

$$Z(\lambda) = \int_{\mathbb{R}^d} e^{\lambda \varphi} d\gamma \ \leq \ e^{\lambda^2/2}$$

$\varphi : \mathbb{R}^d \to \mathbb{R}$ 1-Lipschitz

$$\gamma\big(\varphi \geq \textstyle\int_{\mathbb{R}^n} \varphi \, d\gamma + r\big) \ \leq \ e^{-r^2/2}, \quad r \geq 0$$

(dimension free) Gaussian concentration

$\mathcal{F}$ collection of functions $f : S \to \mathbb{R}$

$G(f), f \in \mathcal{F}$ centered Gaussian process

$M = \sup_{f \in \mathcal{F}} G(f)$, M Lipschitz

Gaussian concentration

$$\mathbb{P}\big(|M - m| \geq r\big) \leq 2\, e^{-r^2/2\sigma^2}, \quad r \geq 0$$

m mean or median, $\sigma^2 = \sup_{f \in \mathcal{F}} \mathbb{E}\big(G(f)^2\big)$

Gaussian isoperimetric inequality

C. Borell, V. Sudakov, B. Tsirel'son, I. Ibragimov (1975)

M. Talagrand (1996)

$X_1, \dots, X_n$ independent in $(S, \mathcal{S})$

$\mathcal{F}$ collection of functions $f : S \to \mathbb{R}$

$$M = \sup_{f \in \mathcal{F}} \sum_{i=1}^{n} f(X_i)$$

M Lipschitz and convex

Talagrand's convex distance inequality

concentration inequalities on

$$\mathbb{P}\big(|M - m| \geq r\big), \quad r \geq 0$$

$\varphi : \mathbb{R}^d \to \mathbb{R}$ 1-Lipschitz convex

for any product probability P on $[0,1]^d$

$$P\big(|\varphi - m| \geq r\big) \leq 4\, e^{-r^2/4}, \quad r \geq 0$$

m median of φ for P

M. Talagrand (1996) isoperimetric ideas (induction)

entropy method – Herbst argument

S. Boucheron, G. Lugosi, P. Massart (2013)

$$M = \sup_{f \in \mathcal{F}} \sum_{i=1}^{n} f(X_i)$$

$$|f| \leq 1, \quad \mathbb{E}\big(f(X_i)\big) = 0, \quad f \in \mathcal{F}$$

$$\mathbb{P}\big(|M - m| \geq r\big) \leq C \exp\left(-\frac{r}{C} \log\left(1 + \frac{r}{\sigma^2 + m}\right)\right), \quad r \geq 0$$

m mean or median, $\quad \sigma^2 = \sup_{f \in \mathcal{F}} \sum_{i=1}^{n} \mathbb{E}\big(f^2(X_i)\big)$

early motivation from Probability in Banach spaces

(law of the iterated logarithm)

Otto-Villani theorem (2000)

logarithmic Sobolev inequality for μ (on $\mathbb{R}^d$)

$$\mathrm{H}\big(\nu\,|\,\mu\big) \;\leq\; \frac{C}{2}\,\mathrm{I}\big(\nu\,|\,\mu\big), \qquad \nu \ll \mu$$

implies quadratic transportation cost inequality

$$\mathrm{W}_2^2(\nu, \mu) \;\leq\; 2C\,\mathrm{H}\big(\nu\,|\,\mu\big), \qquad \nu \ll \mu$$

$$\mathrm{W}_2^2(\nu, \mu) \;=\; \inf_{\nu \leftarrow \pi \rightarrow \mu} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |x-y|^2 d\pi(x,y)$$

Kantorovich-Wasserstein distance

$\mu = \gamma$ M. Talagrand (1996)

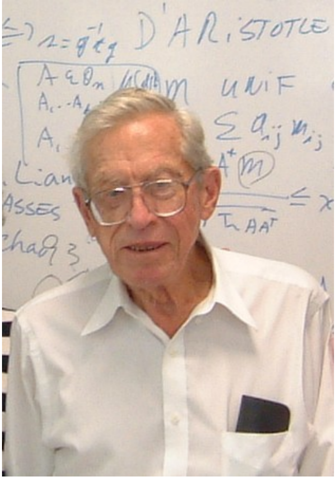
Plan

1. logarithmic Sobolev inequalities

2. Stein's method

3. links and applications

C. Stein (1972)



C. Stein (1972)

γ standard normal on $\mathbb{R}$

$$\int_{\mathbb{R}} x\,\phi\,d\gamma = \int_{\mathbb{R}} \phi'\,d\gamma, \qquad \phi : \mathbb{R} \to \mathbb{R} \quad \text{smooth}$$

characterizes γ

Stein's inequality

ν probability measure on $\mathbb{R}$

$$\|\nu - \gamma\|_{\mathrm{TV}} \leq \sup_{\|\phi\|_\infty \leq \sqrt{\pi/2},\, \|\phi'\|_\infty \leq 2} \left[\int_{\mathbb{R}} x\,\phi\,d\nu - \int_{\mathbb{R}} \phi'\,d\nu \right]$$

ν (centered) probability measure on $\mathbb{R}$

Stein's kernel for ν : $\quad x \mapsto \tau_\nu(x)$

$$\int_{\mathbb{R}} x\,\phi\,d\nu = \int_{\mathbb{R}} \tau_\nu\,\phi'\,d\nu, \qquad \phi : \mathbb{R} \to \mathbb{R} \text{ smooth}$$

γ standard normal $\qquad \tau_\gamma = 1$

$$d\nu = f\,dx$$

$$\tau_\nu(x) = \big[f(x)\big]^{-1} \int_x^\infty y f(y) dy, \qquad x \in \mathrm{supp}(f)$$

(τ_ν polynomial: Pearson class)

zero/size-biased coupling L. Goldstein, Y. Rinott, G. Reinert (1996-97)

ν (centered) probability measure on $\mathbb{R}$

Stein's kernel for ν : $\quad x \mapsto \tau_\nu(x)$

$$\int_{\mathbb{R}} x\,\phi\,d\nu = \int_{\mathbb{R}} \tau_\nu\,\phi'\,d\nu, \qquad \phi : \mathbb{R} \to \mathbb{R} \quad \text{smooth}$$

γ standard normal $\qquad \tau_\gamma = 1$

Stein's discrepancy $\mathrm{S}(\nu \,|\, \gamma)$

$$\mathrm{S}^2(\nu \,|\, \gamma) = \int_{\mathbb{R}} |\tau_\nu - 1|^2 d\nu$$

Stein's inequality

$$\|\nu - \gamma\|_{\mathrm{TV}} \leq 2\,\mathrm{S}(\nu \,|\, \gamma)$$

Stein's method

normal approximation

rates of convergence in central limit theorems

convergence of Wiener chaos

Lindeberg's replacement method, random matrices

exchangeable pairs, concentration inequalities central limit theorem

$X, X_1, \dots, X_n$ iid real random variables

mean zero, variance one

$$S_n = \frac{1}{\sqrt{n}} (X_1 + \cdots + X_n)$$

$$\mathrm{S}^2\big(\mathcal{L}(S_n) \,|\, \gamma\big) \le \frac{1}{n} \mathrm{S}^2\big(\mathcal{L}(X) \,|\, \gamma\big) = \frac{1}{n} \mathrm{Var}\big(\tau_{\mathcal{L}(X)}(X)\big)$$

$$\big\| \mathcal{L}(S_n) - \gamma \big\|_{\mathrm{TV}}^2 \le 4\, \mathrm{S}^2\big(\mathcal{L}(S_n) \,|\, \gamma\big) = O\Big(\frac{1}{n}\Big)$$

Wiener multiple integrals (chaos)

multilinear Gaussian polynomials (degree k)

$$F = \sum_{i_1,\ldots,i_k=1}^{N} a_{i_1,\ldots,i_k} X_{i_1} \cdots X_{i_k}$$

$X_1, \ldots, X_N$ independent standard normals

$a_{i_1,\ldots,i_k} \in \mathbb{R}$ symmetric, vanishing on diagonals

$$\mathbb{E}(F^2) = 1$$

D. Nualart, G. Peccati (2005) fourth moment theorem

$F = F_n, \quad n \in \mathbb{N}$ k-chaos (fixed degree k)

$$N = N_n \to \infty$$

$\mathbb{E}(F_n^2) = 1$ (or $\to 1$)

F_n converges to a standard normal

if and only if

$$\mathbb{E}(F_n^4) \to 3 \qquad \left(= \int_{\mathbb{R}} x^4 d\gamma \right)$$

F Wiener functional

$$\tau_F(x) = \mathbb{E}\big(\langle DF, -D\,\mathrm{L}^{-1}F\rangle \,|\, F = x\big)$$

L Ornstein-Uhlenbeck operator, D Malliavin derivative

$$\mathrm{S}^2\big(\mathcal{L}(F)\,|\,\gamma\big) \le \frac{k-1}{3k}\,\big[\mathbb{E}(F^4) - 3\big]$$

I. Nourdin, G. Peccati (2009), I. Nourdin, J. Rosinski (2012)

multidimensional versions

J. Lindeberg (1922)

$\mathcal{X} = (X_1, \ldots, X_n)$ iid, mean zero, variance one, third moment

$\phi : \mathbb{R}^n \to \mathbb{R}$ smooth, $\quad i = 1, \ldots, n$

$$\mathbb{E}\big(\phi(\mathcal{X})\big) - \int_{\mathbb{R}^n} \phi \, d\gamma = O\bigg(\sum_{i=1}^{n} \|\partial_i^3 \phi\|_\infty \bigg)$$

Taylor expansion $\quad \mathbb{E}\big(X_i \partial_i \phi(\mathcal{X})\big) = \mathbb{E}\big(\partial_i^2 \phi(\mathcal{X})\big) + O\big(\|\partial_i^3 \phi\|_\infty\big)$

$$\mathbb{E}\big(X_i \partial_i \phi(\mathcal{X})\big) = \mathbb{E}\big(\tau(X_i) \partial_i^2 \phi(\mathcal{X})\big)$$

$$\mathbb{E}\big(\mathrm{L}\phi(\mathcal{X})\big) - \int_{\mathbb{R}^n} \mathrm{L}\phi \, d\gamma = \sum_{i=1}^{n} \mathbb{E}\big(\big[\tau(X_i) - 1\big] \partial_i^2 \phi(\mathcal{X})\big) = O\bigg(\sum_{i=1}^{n} \|\partial_i^3 \phi\|_\infty \bigg)$$

S. Chatterjee (2006)

comparison between arbitrary samples, strong embeddings

universality in random matrix theory

T. Tao, V. Vu (2010-11)

four moment theorem

L. Erdös, H. T. Yau, A. Knowles, J. Yin (2012-13)

Green's function comparison

S. Chatterjee (2005)

exchangeable pairs (X, X')

$$\varphi(X) = \mathbb{E}\big(F(X, X') \,|\, X\big) \qquad F \text{ antisymmetric}$$

$$\big|\big(\varphi(X) - \varphi(X')\big)F(X, X')\big| \leq C$$

$$\mathbb{P}\big(\varphi(X) \geq r\big) \leq 2\, e^{-r^2/2C}, \quad r \geq 0$$

S. Chatterjee, P. Dey (2009-11)

S. Ghosh, L. Goldstein (2011) size-biased couplings

graphs, matching, Gibbs measures, Ising model,

matrix concentration, machine learning....

ν (centered) probability measure on $\mathbb{R}$

Stein's kernel for ν : $x \mapsto \tau_\nu(x)$

$$\int_{\mathbb{R}} x\,\phi\,d\nu = \int_{\mathbb{R}} \tau_\nu\,\phi'\,d\nu, \qquad \phi : \mathbb{R} \to \mathbb{R} \text{ smooth}$$

γ standard normal $\qquad \tau_\gamma = 1$

Stein's discrepancy $\mathrm{S}(\nu\,|\,\gamma)$

$$\mathrm{S}^2(\nu\,|\,\gamma) = \int_{\mathbb{R}} |\tau_\nu - 1|^2 d\nu$$

Stein's inequality

$$\|\nu - \gamma\|_{\mathrm{TV}} \leq 2\,\mathrm{S}(\nu\,|\,\gamma)$$

ν (centered) probability measure on $\mathbb{R}^d$

(existence) Stein's kernel (matrix) for ν : $x \mapsto \tau_\nu(x) = \big(\tau_\nu^{ij}(x)\big)_{1 \leq i,j \leq d}$

$$\int_{\mathbb{R}^d} x \, \phi \, d\nu = \int_{\mathbb{R}^d} \tau_\nu \, \nabla \phi \, d\nu, \qquad \phi : \mathbb{R}^d \to \mathbb{R} \quad \text{smooth}$$

Stein's discrepancy $\mathrm{S}(\nu \,|\, \gamma)$

$$\mathrm{S}^2\big(\nu \,|\, \gamma\big) = \int_{\mathbb{R}^d} \|\tau_\nu - \mathrm{Id}\|_{\mathrm{HS}}^2 \, d\nu$$

no Stein's inequality in general

Stein's inequality (on $\mathbb{R}$)

$$\|\nu - \gamma\|_{\mathrm{TV}} \leq 2\,\mathrm{S}(\nu \,|\, \gamma)$$

stronger distance in entropy

ν probability measure on $\mathbb{R}^d$, $\nu \ll \gamma$ $d\nu = h\,d\gamma$

relative H-entropy $\mathrm{H}(\nu \,|\, \gamma) = \int_{\mathbb{R}^d} h \log h \, d\gamma$

Pinsker's inequality

$$\|\nu - \gamma\|_{\mathrm{TV}}^2 \leq \frac{1}{2}\,\mathrm{H}(\nu \,|\, \gamma)$$

I. Nourdin, G. Peccati, Y. Swan (2013)

$(F_n)_{n \in \mathbb{N}}$ sequence of Wiener chaos, fixed degree

$$\mathrm{H}\big(\mathcal{L}(F_n)\,|\,\gamma\big) \to 0 \qquad \text{as} \qquad \mathrm{S}\big(\mathcal{L}(F_n)\,|\,\gamma\big) \to 0$$

(fourth moment theorem $\mathrm{S}(\mathcal{L}(F_n)\,|\,\gamma) \to 0$)

Plan

1. logarithmic Sobolev inequalities

2. Stein's method

3. links and applications

γ standard Gaussian measure on $\mathbb{R}^d$, $\quad d\nu = h\,d\gamma$

logarithmic Sobolev inequality

$$\mathrm{H}(\nu \,|\, \gamma) \leq \frac{1}{2}\,\mathrm{I}(\nu \,|\, \gamma)$$

(relative) H-entropy $\quad \mathrm{H}(\nu \,|\, \gamma) = \int_{\mathbb{R}^d} h \log h \, d\gamma$

(relative) Fisher Information $\quad \mathrm{I}(\nu \,|\, \gamma) = \int_{\mathbb{R}^d} \frac{|\nabla h|^2}{h} \, d\gamma$

(relative) Stein discrepancy

$$\mathrm{S}^2(\nu \,|\, \gamma) = \int_{\mathbb{R}^d} \|\tau_\nu - \mathrm{Id}\|_{\mathrm{HS}}^2 \, d\nu$$

new HSI inequality

H-entropy-Stein-Information

$$\mathrm{H}(\nu \,|\, \gamma) \;\leq\; \frac{1}{2}\,\mathrm{S}^2(\nu \,|\, \gamma)\log\left(1 + \frac{\mathrm{I}(\nu \,|\, \gamma)}{\mathrm{S}^2(\nu \,|\, \gamma)}\right), \qquad \nu \ll \gamma$$

$\log(1+x) \leq x$ improves upon the logarithmic Sobolev inequality

potential towards concentration inequalities new HSI inequality

H-entropy-Stein-Information

$$\mathrm{H}\big(\nu\,|\,\gamma\big) \;\leq\; \frac{1}{2}\,\mathrm{S}^2\big(\nu\,|\,\gamma\big)\log\left(1+\frac{\mathrm{I}(\nu\,|\,\gamma)}{\mathrm{S}^2(\nu\,|\,\gamma)}\right), \qquad \nu \ll \gamma$$

entropic convergence

if $\mathrm{S}(\nu_n\,|\,\gamma) \to 0$ and $\mathrm{I}(\nu_n\,|\,\gamma)$ bounded, then

$$\mathrm{H}\big(\nu_n\,|\,\gamma\big) \;\to\; 0$$

entropic central limit theorem

$X, X_1, \ldots, X_n$ iid real random variables, mean zero, variance one

$$S_n = \frac{1}{\sqrt{n}}\,(X_1 + \cdots + X_n)$$

J. Linnik (1959), A. Barron (1986)

$$\mathrm{H}\big(\mathcal{L}(S_n)\,|\,\gamma\big) \to 0$$

entropic central limit theorem

$X, X_1, \ldots, X_n$ iid real random variables, mean zero, variance one

$$S_n = \frac{1}{\sqrt{n}} (X_1 + \cdots + X_n)$$

$$\mathrm{S}^2\big(\mathcal{L}(S_n) \,|\, \gamma\big) \leq \frac{1}{n} \mathrm{Var}\big(\tau_{\mathcal{L}(X)}(X)\big)$$

Stam's inequality $\quad \mathrm{I}\big(\mathcal{L}(S_n) \,|\, \gamma\big) \leq \mathrm{I}\big(\mathcal{L}(X) \,|\, \gamma\big) < \infty$

HSI inequality

$$\mathrm{H}\big(\nu\,|\,\gamma\big) \;\leq\; \frac{1}{2}\,\mathrm{S}^2\big(\nu\,|\,\gamma\big)\log\left(1+\frac{\mathrm{I}(\nu\,|\,\gamma)}{\mathrm{S}^2(\nu\,|\,\gamma)}\right), \qquad \nu \ll \gamma$$

H-entropy	$\mathrm{H}(\nu\,|\,\gamma)$
Fisher Information	$\mathrm{I}(\nu\,|\,\gamma)$
Stein discrepancy	$\mathrm{S}(\nu\,|\,\gamma)$

entropic central limit theorem

$X, X_1, \ldots, X_n$ iid real random variables, mean zero, variance one

$$S_n = \frac{1}{\sqrt{n}} (X_1 + \cdots + X_n)$$

$$\mathrm{S}^2\big(\mathcal{L}(S_n) \,|\, \gamma\big) \le \frac{1}{n} \mathrm{Var}\big(\tau_{\mathcal{L}(X)}(X)\big)$$

Stam's inequality $\quad \mathrm{I}\big(\mathcal{L}(S_n) \,|\, \gamma\big) \le \mathrm{I}\big(\mathcal{L}(X) \,|\, \gamma\big) < \infty$

HSI inequality $\quad \mathrm{H}\big(\mathcal{L}(S_n) \,|\, \gamma\big) = O\bigg(\frac{\log n}{n}\bigg)$

optimal $O(\frac{1}{n})$ under fourth moment on X

S. Bobkov, G. Chistyakov, F. Götze (2013-14)

$\varphi : \mathbb{R}^d \to \mathbb{R}$ 1-Lipschitz, $\int_{\mathbb{R}^d} \varphi \, d\gamma = 0$

$$\gamma(\varphi \geq r) \ \leq \ e^{-r^2/2}, \quad r \geq 0$$

Gaussian concentration

equivalent (up to numerical constants)

$$\left(\int_{\mathbb{R}^d} |\varphi|^p d\gamma \right)^{1/p} \ \leq \ C \sqrt{p} \,, \qquad p \geq 1$$

moment growth: concentration rate

ν (centered) probability measure on $\mathbb{R}^d$

$\varphi : \mathbb{R}^d \to \mathbb{R}$ 1-Lipschitz, $\int_{\mathbb{R}^d} \varphi \, d\nu = 0$

moment growth in $p \geq 2$, $C > 0$ numerical

$$\bigg(\int_{\mathbb{R}^d} |\varphi|^p d\nu \bigg)^{1/p} \leq C \bigg[\mathrm{S}_p(\nu \,|\, \gamma) + \sqrt{p} \bigg(\int_{\mathbb{R}^d} \|\tau_\nu\|_{\mathrm{Op}}^{p/2} \, d\nu \bigg)^{1/p} \bigg]$$

$$\mathrm{S}_p(\nu \,|\, \gamma) = \bigg(\int_{\mathbb{R}^d} \|\tau_\nu - \mathrm{Id}\|_{\mathrm{HS}}^p \, d\nu \bigg)^{1/p}$$

$$S_n = \frac{1}{\sqrt{n}} (X_1 + \cdots + X_n)$$

$X, X_1, \ldots, X_n$ iid mean zero, covariance identity in $\mathbb{R}^d$

$\varphi : \mathbb{R}^d \to \mathbb{R}$ 1-Lipschitz

$$\mathbb{P}\Big(\Big| \varphi(S_n) - \mathbb{E}\big(\varphi(S_n)\big) \Big| \geq r \Big) \leq C\, e^{-r^2/C}$$

$$0 \leq r \leq r_n \to \infty$$

according to the growth in p of $\mathrm{S}_p(\nu \,|\, \gamma)$

F. Otto, C. Villani (2000)

$$\mathrm{H}\big(\nu \,|\, \gamma\big) \,\leq\, \mathrm{W}_2(\nu,\gamma)\sqrt{\mathrm{I}\big(\nu \,|\, \gamma\big)} - \frac{1}{2}\,\mathrm{W}_2^2(\nu,\gamma)$$

Kantorovich-Wasserstein distance

$$\mathrm{W}_2^2(\nu,\gamma) \,=\, \inf_{\nu\leftarrow\pi\rightarrow\gamma} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |x-y|^2 d\pi(x,y)$$

covers both the logarithmic Sobolev inequality

and Talagrand's transportation cost inequality

$$\mathrm{W}_2^2(\nu,\gamma) \,\leq 2\,\mathrm{H}\big(\nu \,|\, \gamma\big)$$

Talagrand's inequality

$$\mathrm{W}_2^2(\nu,\gamma) \leq 2\,\mathrm{H}(\nu\,|\,\gamma)$$

$\nu << \gamma$ (centered) probability measure on $\mathbb{R}^d$

WSH inequality

$$\mathrm{W}_2(\nu,\gamma) \leq \mathrm{S}(\nu\,|\,\gamma)\arccos\left(e^{-\frac{\mathrm{H}(\nu\,|\,\gamma)}{\mathrm{S}^2(\nu\,|\,\gamma)}}\right)$$

$$\arccos(e^{-r}) \leq \sqrt{2r}$$

Stein's inequality (in $\mathbb{R}^d$) $\quad \mathrm{W}_2(\nu,\gamma) \leq \mathrm{S}(\nu\,|\,\gamma)$

HSI inequality

$$\mathrm{H}\big(\nu\,|\,\gamma\big) \;\leq\; \frac{1}{2}\,\mathrm{S}^2\big(\nu\,|\,\gamma\big)\log\left(1+\frac{\mathrm{I}(\nu\,|\,\gamma)}{\mathrm{S}^2(\nu\,|\,\gamma)}\right), \qquad \nu \ll \gamma$$

H-entropy	$\mathrm{H}(\nu\,|\,\gamma)$
Fisher Information	$\mathrm{I}(\nu\,|\,\gamma)$
Stein discrepancy	$\mathrm{S}(\nu\,|\,\gamma)$

Ornstein-Uhlenbeck semigroup $(P_t)_{t \geq 0}$

$$P_t f(x) = \int_{\mathbb{R}^d} f\big(e^{-t}x + \sqrt{1 - e^{-2t}}\, y\big) d\gamma(y)$$

$$d\nu = h\, d\gamma, \quad d\nu_t = P_t h\, d\gamma \qquad (\nu_0 = \nu, \ \ \nu_\infty = \gamma)$$

de Bruijin's formula $\quad \mathrm{H}\big(\nu \,|\, \gamma\big) = \int_0^\infty \mathrm{I}\big(\nu_t \,|\, \gamma\big) dt$

$$\mathrm{I}\big(\nu_t \,|\, \gamma\big) \leq e^{-2t}\, \mathrm{I}\big(\nu \,|\, \gamma\big)$$

Bakry-Émery criterion

classical logarithmic Sobolev inequality

$$\mathrm{H}(\nu \,|\, \gamma) \;=\; \int_0^\infty \mathrm{I}(\nu_t \,|\, \gamma)\, dt$$

classical $\quad \mathrm{I}(\nu_t \,|\, \gamma) \;\leq\; e^{-2t}\, \mathrm{I}(\nu \,|\, \gamma)$

new main ingredient $\quad \mathrm{I}(\nu_t \,|\, \gamma) \;\leq\; \dfrac{e^{-4t}}{1 - e^{-2t}}\, \mathrm{S}^2(\nu \,|\, \gamma)$

(information theoretic) representation of $\mathrm{I}(\nu_t \,|\, \gamma)$

$$\frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \Big[\big(\tau_\nu(x) - \mathrm{Id}\big) y \cdot \nabla v_t \big(e^{-t} x + \sqrt{1 - e^{-2t}}\, y\big) \Big] d\nu(x) d\gamma(y)$$

$$v_t = \log P_t h$$

optimize small $t > 0$ and large $t > 0$

$$\mathrm{H}(\nu \,|\, \mu) \leq \frac{1}{2}\, \mathrm{S}^2(\nu \,|\, \mu) \log\left(1 + \frac{\mathrm{I}(\nu \,|\, \mu)}{\mathrm{S}^2(\nu \,|\, \mu)}\right)$$

μ gamma, beta distributions

multidimensional

families of log-concave distributions μ

Markov Triple (E, μ, Γ)

(abstract Wiener space)

Markov Triple (E, μ, Γ)

Markov (diffusion) operator L with state space E

μ invariant and symmetric probability measure

Γ bilinear gradient operator (carré du champ)

$$\Gamma(f,g) = \frac{1}{2}\big[\mathrm{L}(fg) - f\,\mathrm{L}g - g\,\mathrm{L}f\big], \qquad f,g \in \mathcal{A}$$

integration by parts

$$\int_E f(-\mathrm{L}g)\,d\mu = \int_E \Gamma(f,g)d\mu$$

second order differential operator on $E = \mathbb{R}^d$

$$\mathrm{L} = \sum_{i,j=1}^{d} a^{ij}\,\partial_{ij}^2 f + \sum_{i=1}^{d} b^i\,\partial_i f$$

$$\Gamma(f,g) = \sum_{i,j=1}^{d} a^{ij}\,\partial_i f\,\partial_j g$$

$$\int_{\mathbb{R}^d} \mathrm{L} f\,d\mu = 0$$

example: Ornstein-Uhlenbeck operator

$$\mathrm{L} f = \Delta f - x \cdot \nabla f = \sum_{i,j=1}^{d} \partial_{ij}^2 f - \sum_{i=1}^{d} x_i\,\partial_i f$$

γ invariant measure

$$\mathrm{L} = \sum_{i,j=1}^{d} a^{ij}\,\partial_{ij}^2 f + \sum_{i=1}^{d} b^i\,\partial_i f = \big\langle a, \mathrm{Hess}(f)\big\rangle_{\mathrm{HS}} + b \cdot \nabla f$$

μ invariant measure $\quad \big(\int_{\mathbb{R}^d} \mathrm{L}f\, d\mu = 0 \big)$

Stein kernel τ_ν

$$- \int_{\mathbb{R}^d} b \cdot \nabla f\, d\nu = \int_{\mathbb{R}^d} \big\langle \tau_\nu, \mathrm{Hess}(f)\big\rangle_{\mathrm{HS}}\, d\nu$$

Stein discrepancy

$$\mathrm{S}^2(\nu \,|\, \mu) = \int_{\mathbb{R}^d} \big\| a^{-\frac{1}{2}} \tau_\nu a^{-\frac{1}{2}} - \mathrm{Id} \big\|_{\mathrm{HS}}^2\, d\nu$$

approximation of diffusions A. Barbour (1990), F. Götze (1991)

Markov Triple (E, μ, Γ)

Markov (diffusion) operator L with state space E

μ invariant and symmetric probability measure

Γ bilinear gradient operator (carré du champ)

$$\Gamma(f,g) = \frac{1}{2}\big[\mathrm{L}(fg) - f\,\mathrm{L}g - g\,\mathrm{L}f\big], \qquad f,g \in \mathcal{A}$$

logarithmic Sobolev inequality

Γ_2 iterated gradient operator D.Bakry, M. Émery (1985)

$$\Gamma_2(f,g) = \frac{1}{2}\big[\mathrm{L}\,\Gamma(f,g) - \Gamma(f,\mathrm{L}g) - \Gamma(g,\mathrm{L}f)\big], \qquad f,g \in \mathcal{A}$$

$$P_t = e^{t\,\mathrm{L}},\ t \geq 0, \qquad d\nu_t = P_t h d\mu$$

$$\mathrm{H}(\nu \mid \mu) = \int_0^\infty \mathrm{I}(\nu_t \mid \mu) dt$$

classical $\qquad \mathrm{I}(\nu_t \mid \mu) \leq e^{-2\rho t}\, \mathrm{I}(\nu \mid \mu)$

Bakry-Émery Γ_2 criterion

$$\Gamma_2(f) = \Gamma_2(f,f) \geq \rho\, \Gamma(f,f) = \Gamma(f), \quad f \in \mathcal{A}$$

$$\frac{d}{dt}\, \mathrm{I}(\nu_t \mid \mu) = -2 \int_E P_t h\, \Gamma_2(\log P_t h) d\mu$$

$$\leq -2\rho \int_E P_t h\, \Gamma(\log P_t h) d\mu = -2\rho\, \mathrm{I}(\nu_t \mid \mu)$$

$$\mathrm{I}\big(\nu_t \,|\, \mu\big) \;\leq\; \frac{c}{e^{2ct}-1}\,\mathrm{S}^2\big(\nu \,|\, \mu\big)$$

new feature: Γ_3 operator

$$\Gamma_3(f) \;\geq\; c\,\Gamma_2(f), \quad f \in \mathcal{A}$$

$$\Gamma_3(f,g) \;=\; \tfrac{1}{2}\big[\mathrm{L}\,\Gamma_2(f,g) - \Gamma_2(f,\mathrm{L}g) - \Gamma_2(g,\mathrm{L}f)\big], \qquad f,g \in \mathcal{A}$$

control of Hessians (Stein's kernel)

$$P_t\big(\Gamma(h)\big) \;\geq\; \frac{e^{2ct}-1}{c}\,\big\|a^{\frac{1}{2}}\,\mathrm{Hess}(P_t h)a^{\frac{1}{2}}\big\|_{\mathrm{HS}}^2$$

Ornstein-Uhlenbeck operator $\mathrm{L}f = \Delta f - x \cdot \nabla f$ on $\mathbb{R}^d$

$$\Gamma(f) = |\nabla f|^2$$
$$\Gamma_2(f) = |\nabla^2 f|^2 + |\nabla f|^2$$
$$\Gamma_3(f) = |\nabla^3 f|^2 + 3|\nabla^2 f|^2 + 2|\nabla f|^2$$

Bochner's formula for

Δ Laplacian on (M, g) Riemannian manifold

$$\Gamma_2(f) = |\nabla^2 f|^2 + \mathrm{Ric}_g(\nabla f, \nabla f)$$

$$\Gamma_2(f) \geq \rho\,\Gamma(f) \iff \mathrm{Ric}_g \geq \rho$$

$$\mathrm{H}(\nu \,|\, \mu) \;\leq\; C\,\mathrm{S}^2(\nu \,|\, \mu)\,\Psi\left(\frac{C\,\mathrm{I}(\nu \,|\, \mu)}{\mathrm{S}^2(\nu \,|\, \mu)}\right)$$

$$\Psi(r) \;=\; 1 + \log r, \quad r \geq 1$$

μ gamma, beta distributions

multidimensional

families of log-concave distributions μ

discrete models (Poisson) under investigation

Markov Triple (E, μ, Γ) (abstract Wiener space)

$\mathrm{I}(\nu \,|\, \gamma)$ difficult to control in general

$F = (F_1, \ldots, F_d) : E \to \mathbb{R}^d$ with law $\mathcal{L}(F)$

$$\mathrm{H}\big(\mathcal{L}(F) \,|\, \gamma\big) \leq C_F \, \mathrm{S}^2\big(\mathcal{L}(F) \,|\, \gamma\big) \, \Psi\left(\frac{C_F}{\mathrm{S}^2(\mathcal{L}(F) \,|\, \gamma)}\right)$$

$$\Psi(r) = 1 + \log r, \quad r \geq 1$$

$C_F > 0$ depend on integrability of F, $\Gamma(F_i, F_j)$

and inverse of the determinant of $(\Gamma(F_i, F_j))_{1 \leq i,j \leq d}$

(Malliavin calculus)

towards entropic convergence via HSI

$\mathrm{I}(\nu \,|\, \gamma)$ difficult to control in general

Wiener chaos or multilinear polynomial

$$F = \sum_{i_1,\ldots,i_k=1}^{N} a_{i_1,\ldots,i_k} X_{i_1} \cdots X_{i_k}$$

$X_1, \ldots, X_N$ independent standard normals

$a_{i_1,\ldots,i_k} \in \mathbb{R}$ symmetric, vanishing on diagonals

law $\mathcal{L}(F)$ of F? Fisher information $\mathrm{I}\big(\mathcal{L}(F) \,|\, \gamma\big)$?

I. Nourdin, G. Peccati, Y. Swan (2013)

$(F_n)_{n\in\mathbb{N}}$ sequence of Wiener chaos, fixed degree

$\mathrm{H}\big(\mathcal{L}(F_n)\,|\,\gamma\big) \to 0$ as $\mathrm{S}\big(\mathcal{L}(F_n)\,|\,\gamma\big) \to 0$

(fourth moment theorem $\mathrm{S}(\mathcal{L}(F_n)\,|\,\gamma) \to 0$)

Markov Triple (E, μ, Γ) (abstract Wiener space)

$F = (F_1, \ldots, F_d) : E \to \mathbb{R}^d$ with law $\mathcal{L}(F)$

$$\mathrm{H}\big(\mathcal{L}(F) \,|\, \gamma\big) \leq C_F \, \mathrm{S}^2\big(\mathcal{L}(F) \,|\, \gamma\big) \, \Psi\left(\frac{C_F}{\mathrm{S}^2(\mathcal{L}(F) \,|\, \gamma)}\right)$$

$$\Psi(r) = 1 + \log r, \quad r \geq 1$$

$C_F > 0$ depend on integrability of F, $\Gamma(F_i, F_j)$

and inverse of the determinant of $(\Gamma(F_i, F_j))_{1 \leq i,j \leq d}$

(Malliavin calculus)

$$\mathrm{H}\big(\mathcal{L}(F)\,|\,\gamma\big) \;\leq\; C_\alpha\,\mathrm{S}^2\big(\mathcal{L}(F)\,|\,\gamma\big)\,\Psi\left(\frac{2(A_F + d(B_F+1))}{\mathrm{S}^2(\mathcal{L}(F)\,|\,\gamma)}\right)$$

$A_F < \infty$ moment assumptions

$$B_F \;=\; \int_E \frac{1}{\det(\widetilde{\Gamma})^\alpha}\,d\mu, \qquad \alpha > 0$$

$$\widetilde{\Gamma} \;=\; \big(\Gamma(F_i, F_j)\big)_{1 \leq i,j \leq d}$$

(Malliavin matrix)

$$B_F \;=\; \int_E \frac{1}{\det(\widetilde{\Gamma})^\alpha}\, d\mu, \qquad \alpha > 0$$

Gaussian vector chaos $F = (F_1, \ldots, F_d)$

$$\Gamma(F_i, F_j) \;=\; \langle DF_i, DF_j \rangle_{\mathfrak{H}}$$

$\mathcal{L}(F)$ density: $\mathbb{E}(\det(\widetilde{\Gamma})) > 0$

$$\mathbb{P}\big(\det(\widetilde{\Gamma}) \leq \lambda\big) \;\leq\; cN\lambda^{1/N}\, \mathbb{E}\big(\det(\widetilde{\Gamma})\big)^{-1/N}, \quad \lambda > 0$$

N degrees of the F_i's

A. Carbery, J. Wright (2001)

log-concave models

Thank you for your attention