Anuran Makur
30 Oct 2017

Modified Logarithmic Sobolev Inequalities:

[based on Chapter 4 (and 2) of "Concentration of Measure and Logarithmic Sobolev Inequalities" by Michel Ledoux, 1997]

① Motivation:

Consider the product measure ν^n on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ such that ν is the (two-sided exponential) Laplace distribution: $\forall x \in \mathbb{R}$, $\nu(x) = \frac{1}{2}e^{-|x|}$.
($\mathcal{B}$: Borel σ-algebra; $\nu(x)$: pdf (with abuse of notation))

Question: Does ν^n satisfy a log-Sobolev inequality (LSI) with respect to "standard" Dirichlet form?

Suppose it does...

LSI: $\exists C > 0, \forall f \in \mathcal{A}$, $\text{Ent}_{\nu^n}(f^2) \leq 2C\, \mathbb{E}_{\nu^n}[|\nabla f|^2]$

$\mathcal{A}$ = subset of measurable functions $f: \mathbb{R}^n \to \mathbb{R}$ such that the gradient $\nabla f: \mathbb{R}^n \to \mathbb{R}^n$ exists and $\mathbb{E}_{\nu^n}[|\nabla f|^2] < \infty$

Entropy: $\text{Ent}_{\nu^n}(f^2) \triangleq \mathbb{E}_{\nu^n}[f^2 \log(f^2)] - \mathbb{E}_{\nu^n}[f^2]\log(\mathbb{E}_{\nu^n}[f^2])$

Dirichlet form: $\mathbb{E}_{\nu^n}[|\nabla f|^2] = \sum_{i=1}^n \mathbb{E}_{\nu^n}[(\partial_i f)^2]$ ($\partial_i f$: partial derivative with respect to ith coordinate)

Then, for any $F: \mathbb{R}^n \to \mathbb{R}$ smooth with $\|F\|_{\text{Lip}} \leq 1$, setting $f^2 = e^{\lambda F}$ for $\lambda \in \mathbb{R}$ and employing the Herbst argument gives the following concentration of measure inequality:

$$\forall r \geq 0, \quad \nu^n(F \geq \mathbb{E}_{\nu^n}[F] + r) \leq e^{-r^2/2C} .$$
← after optimizing over $\lambda \in \mathbb{R}$

Contradiction: Since F linear satisfies the required conditions, we see that a linear function has a Gaussian tail. However, the tail for a linear function must be exponential!

② Exponential LSI via Gaussian LSI:

Recall the Gaussian LSI: For all $f: \mathbb{R}^n \to \mathbb{R}$ with $\mathbb{E}_\gamma[|\nabla f|^2] < \infty$,

$$\text{Ent}_\gamma(f^2) \leq 2\, \mathbb{E}_\gamma[|\nabla f|^2],$$

where $\gamma \sim \mathcal{N}(0, I_n)$ is the canonical Gaussian measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.
(I_n: $n \times n$ identity matrix)

Consider (one-sided) exponential measure $\hat{\nu}$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with pdf $\hat{\nu}(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$.

It is well-known that if $Z_1, Z_2 \overset{iid}{\sim} \mathcal{N}(0,1)$, then $X = \frac{Z_1^2 + Z_2^2}{2}$ has distribution $\hat{\nu}$.

[Check: $f_{Z_1,Z_2}(z_1,z_2) = \frac{1}{2\pi}\exp\left(-\frac{1}{2}(z_1^2+z_2^2)\right) = \frac{1}{2\pi}e^{-x} = \underbrace{f_\Theta(\theta)}_{\text{Unif}[0,2\pi]} \overbrace{f_X(x)}^{\hat{\nu}(x)}$]

Consider $f(x,y) = g\left(\frac{x^2+y^2}{2}\right)$. Then, $|\nabla f|^2 = \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 = \left(x\, g'\left(\frac{x^2+y^2}{2}\right)\right)^2 + \left(y\, g'\left(\frac{x^2+y^2}{2}\right)\right)^2$

$$\Rightarrow |\nabla f|^2 = 2\left(\frac{x^2+y^2}{2}\right) g'\left(\frac{x^2+y^2}{2}\right)^2$$

$$\Rightarrow \mathbb{E}_{Z_1,Z_2}[|\nabla f|^2] = 2\,\mathbb{E}_X[X g'(X)^2].$$

Hence, we have: $\underbrace{\text{Ent}_{\hat{\nu}}(g^2)}_{= \text{Ent}_{Z_1,Z_2}(f^2)} \leq \underbrace{4\mathbb{E}_{\hat{\nu}}[X g'(X)^2]}$ for every $g: \mathbb{R}_+ \to \mathbb{R}$ such that the right-hand side is finite.

↑ Consequence of Gaussian LSI

∴ For every sufficiently smooth $f: \mathbb{R}_+^n \to \mathbb{R}$, $\text{Ent}_{\hat{\nu}^n}(f^2) \leq 4\int \sum_{i=1}^n x_i |\partial_i f(x)|^2 \, d\hat{\nu}^n(x)$, via tensorization of entropy – see next section.
($\hat{\nu}^n$: product measure)

Unfortunately, this LSI does not give concentration inequality via Herbst argument. ☹

③ Tensorization of Entropy:

Given probability spaces $(\mathcal{X}_i, \mathcal{B}(\mathcal{X}_i), \mu_i)$, $1 \leq i \leq n$, let $P = \mu_1 \otimes \cdots \otimes \mu_n$ denote the product measure on the product measurable space $(\mathcal{X}_1 \times \cdots \times \mathcal{X}_n, \mathcal{B}(\mathcal{X}_1) \otimes \cdots \otimes \mathcal{B}(\mathcal{X}_n))$.

↑ product σ-algebra

Given $f: \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \rightarrow \mathbb{R}_+$ (measurable), for each $1 \leq i \leq n$, define $f_i: \mathcal{X}_i \rightarrow \mathbb{R}_+$ to be:

$$f_i(x_i) = f(x_1, \ldots, x_{i-1}, \underbrace{x_i}_{\text{only this varies}}, x_{i+1}, \ldots, x_n)$$

where $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ are fixed.

Prop: Under appropriate integrability conditions,

$$\mathrm{Ent}_P(f) \leq \sum_{i=1}^n \mathbb{E}_P\left[\mathrm{Ent}_{\mu_i}(f_i)\right].$$

← Variance has similar tensorization.

Proof: We first prove that $f: \mathcal{X} \rightarrow \mathbb{R}_+$, where $(\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu)$ is a probability space, has entropy satisfying:

[★] $\mathrm{Ent}_\mu(f) = \sup\left\{\mathbb{E}_\mu[fg] : \mathbb{E}_\mu[e^g] \leq 1\right\}$, ← variational characterization of entropy

where the supremum is over all measurable functions $g: \mathcal{X} \rightarrow \mathbb{R}$ with $\mathbb{E}_\mu[e^g] \leq 1$.

By homogeneity, assume $\mathbb{E}_\mu[f] = 1$. Recall from Young's inequality that:

↳ $\mathrm{Ent}_\mu(af) = a\,\mathrm{Ent}_\mu(f)$ for $a \geq 0$

$$\forall u \geq 0, \ \forall v \in \mathbb{R}, \quad uv \leq u\log(u) - u + e^v.$$

← Pf: Let $f(v) = e^v$ for $v \in \mathbb{R}$.

$\Rightarrow f^*(u) = \sup_{v \in \mathbb{R}} uv - f(v)$ [convex conjugate]

Let $\phi(v) = uv - e^v$, $\phi'(v) = u - e^v \overset{\text{set}}{=} 0$ $\Leftrightarrow u = e^v \ (\geq 0)$.

$\Rightarrow f^*(u) = u\log(u) - u$ for $u \geq 0$

So, $uv \leq f(v) + f^*(u)$ for all $u \geq 0, v \in \mathbb{R}$.
$= e^v + u\log(u) - u$ ▨

For any g with $\mathbb{E}_\mu[e^g] \leq 1$, we have from above:

$$\mathbb{E}_\mu[\underset{u}{f}\,\underset{v}{g}] \leq \mathbb{E}_\mu[f\log(f)] - 1 + \mathbb{E}[e^g]$$
$$\leq \mathrm{Ent}_\mu(f).$$

Setting $g = \log(f)$ [note that $\mathbb{E}_\mu[e^g] = \mathbb{E}_\mu[f] = 1$], we get the desired result [★].

Next, consider $g: \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \rightarrow \mathbb{R}$ such that $\mathbb{E}_P[e^g] \leq 1$, and define for each $1 \leq i \leq n$:

$$g^i(x_i, \ldots, x_n) = \log\left(\frac{\int e^{g(x_1, \ldots, x_n)}\, d\mu_1(x_1) \cdots d\mu_{i-1}(x_{i-1})}{\int e^{g(x_1, \ldots, x_n)}\, d\mu_1(x_1) \cdots d\mu_i(x_i)}\right).$$

Then, $\sum_{i=1}^n g^i = g - \log(\underbrace{\mathbb{E}_P[e^g]}_{\leq 1}) \geq g$, and $\mathbb{E}_{\mu_i}[e^{(g^i)_i}] = 1$.

↑ func. of x_i only

$$\Rightarrow \mathbb{E}_P[fg] \leq \sum_{i=1}^n \mathbb{E}_P[fg^i] = \sum_{i=1}^n \mathbb{E}_P\Big[\underbrace{\mathbb{E}_{\mu_i}[f_i (g^i)_i]}_{\leq \mathrm{Ent}_{\mu_i}(f_i)}\Big] \leq \sum_{i=1}^n \mathbb{E}_P\left[\mathrm{Ent}_{\mu_i}(f_i)\right] \quad \text{for } f: \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \rightarrow \mathbb{R}^+$$

← use [★] and $\mathbb{E}_{\mu_i}[e^{(g^i)_i}] = 1$

↓ use [★]

$$\therefore \mathrm{Ent}_P(f) \leq \sum_{i=1}^n \mathbb{E}_P\left[\mathrm{Ent}_{\mu_i}(f_i)\right], \text{ as desired.}$$

▨

Remark: This result allows us to prove LSI for $n = 1$ (single-letter) case and translate to general n. Moreover, the LSI constants are translated in a dimension-free manner, which makes this a useful tool for ∞-dimensional analysis.

Example: From previous section,

$$\underbrace{\mathrm{Ent}_{\hat{\mu}^n}(f^2) \leq}_{\text{tensorization}} \sum_{i=1}^n \mathbb{E}_{\hat{\mu}^n}\left[\mathrm{Ent}_{\hat{\mu}}(f_i^2)\right] \underbrace{\leq}_{\text{single-letter proof}} \sum_{i=1}^n \mathbb{E}_{\hat{\mu}^n}\left[4\int x_i \left(\frac{\partial f}{\partial x_i}\right)^2 d\hat{\mu}(x_i)\right] = 4\int \sum_{i=1}^n x_i |\partial_i f(x)|^2\, d\hat{\mu}^n(x).$$

↑ dimension-free constant

④ Talagrand's Poincaré Inequality:

Talagrand proved a certain Poincaré inequality for the Laplace distribution ν on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
The proof of this inequality offers insight on how to prove a different type of LSI for ν.
Let $S^n \triangleq \{f:\mathbb{R}^n \to \mathbb{R} \,|\, f \text{ continuous, differentiable a.e.}, \int |f| \, d\nu^n < \infty, \int |\nabla f|^2 d\nu^n < \infty, \lim_{x_i \to \pm\infty} e^{-|x_i|} f(x_1, \dots, x_n) = 0, \forall i, \forall x_j, j \neq i\}$.

for rigor

Lemma: If $\phi \in S^1$, then:

$$\int \phi \, d\nu = \phi(0) + \int \operatorname{sign}(x) \phi'(x) \, d\nu(x).$$

signum function: $\operatorname{sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$

Proof: (non-rigorous)
By integration by parts,

$$\int_{-\infty}^{\infty} \underbrace{\phi(x)}_{u} \underbrace{\tfrac{1}{2} e^{-|x|}}_{v' = \nu(x)} dx = \Big[\underbrace{\phi(x)}_{u} \underbrace{\Big(\tfrac{1}{2} + \tfrac{1}{2}\operatorname{sign}(x)\big(1 - e^{-|x|}\big)\Big)}_{v = \text{CDF of } \nu(x)} \Big]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \underbrace{\phi'(x)}_{u'} \underbrace{\Big(\tfrac{1}{2} + \tfrac{1}{2}\operatorname{sign}(x)\big(1 - e^{-|x|}\big)\Big)}_{v} dx$$

$$= \phi(\infty) - \int_{-\infty}^{\infty} \frac{\phi'(x)}{2} dx - \int_{-\infty}^{\infty} \frac{\phi'(x)\operatorname{sign}(x)}{2} dx + \int_{-\infty}^{\infty} \phi'(x) \operatorname{sign}(x) \underbrace{\tfrac{1}{2} e^{-|x|}}_{\nu(x)} dx$$

$$= \frac{\phi(\infty) + \phi(-\infty)}{2} - \int_0^{\infty} \frac{\phi'(x)}{2} dx + \int_{-\infty}^{0} \frac{\phi'(x)}{2} dx + \int_{-\infty}^{\infty} \phi'(x) \operatorname{sign}(x) \tfrac{1}{2} e^{-|x|} dx$$

$$= \phi(0) + \int_{-\infty}^{\infty} \phi'(x) \operatorname{sign}(x) \tfrac{1}{2} e^{-|x|} dx .$$

▨

Prop: (Poincaré inequality for ν) For every $f \in S^1$,

$$\underline{\mathrm{VAR}_\nu(f)} \triangleq \mathbb{E}_\nu[f^2] - \mathbb{E}_\nu[f]^2 \,\underline{\leq 4\, \mathbb{E}_\nu[f'^2]}.$$

Proof: Set $g(x) = f(x) - f(0)$. Then, we have:

$$\mathbb{E}_\nu[g^2] = \cancel{g(0)^2}^{\,0} + \int \operatorname{sign}(x) . 2 g(x) g'(x) \, d\nu(x) \qquad \text{[using Lemma]}$$

$$\leq 2\, \mathbb{E}_\nu[g^2]^{\frac{1}{2}} \, \mathbb{E}_\nu[g'^2]^{\frac{1}{2}} \qquad \text{[Cauchy-Schwarz inequality]}$$

$$\Rightarrow \mathbb{E}_\nu[g^2]^{\frac{1}{2}} \leq 2\, \mathbb{E}_\nu[g'^2]^{\frac{1}{2}}$$

$$\Rightarrow \mathbb{E}_\nu[g^2] \leq 4\, \mathbb{E}_\nu[g'^2] .$$

Hence, since $f' = g'$, $\mathrm{VAR}_\nu(f) = \mathrm{VAR}_\nu(g) \leq \mathbb{E}_\nu[g^2] \leq 4\mathbb{E}_\nu[g'^2] = 4\mathbb{E}_\nu[f'^2]$. ▨

⑤ Modified LSI for Exponential Measure:

Thm: (Modified LSI for ν) For every $0 < c < 1$ and every Lipschitz continuous function $f:\mathbb{R} \to \mathbb{R}$ such that $|f'| \leq c$ a.e.,

$$\underline{\mathrm{Ent}_\nu(e^f) \leq \frac{2}{1-c} \mathbb{E}_\nu[f'^2 e^f]}.$$

Proof: It is straightforward to check that the modified LSI is invariant to adding constants.
Assume without loss of generality that $f(0) = 0$.

[continued.]

Proof continued:

Since $\forall u \geq 0$, $u\log(u) \geq u-1$, we have:

$$\mathrm{Ent}_\nu(e^f) = \mathbb{E}_\nu[fe^f] - \underbrace{\mathbb{E}_\nu[e^f]}_{=u}\log(\mathbb{E}_\nu[e^f]) \leq \mathbb{E}_\nu[fe^f] - \mathbb{E}_\nu[e^f] + 1 = \mathbb{E}_\nu[fe^f - e^f + 1].$$

Moreover, as $|f'| \leq c$ a.e., e^f, fe^f, f^2e^f are all in S'. ← straightforward to check.

$$\mathbb{E}_\nu[fe^f - e^f + 1] = \underbrace{f(0)e^{f(0)} - e^{f(0)} + 1}_{=0} + \int \mathrm{sign}(x)\left[f'(x)f(x)e^{f(x)} + \cancel{f'(x)e^{f(x)}} - \cancel{f'(x)e^{f(x)}}\right] d\nu(x)$$

$$\overset{[*]}{=} \int \mathrm{sign}(x) f'(x) f(x) e^{f(x)}\, d\nu(x) \qquad \text{[using Lemma]}$$

$$\mathbb{E}_\nu[f^2e^f] = \underbrace{f(0)^2e^{f(0)}}_{=0} + \int \mathrm{sign}(x)\left[f'(x)f(x)^2e^{f(x)} + 2f'(x)f(x)e^{f(x)}\right] d\nu(x)$$

$$= 2\int \underbrace{\mathrm{sign}(x)f'(x)f(x)e^{f(x)}}_{\mathrm{sign}(x)f'(x)e^{f(x)/2}\cdot f(x)e^{f(x)/2}}\, d\nu(x) + \int \underbrace{\mathrm{sign}(x)f'(x)f(x)^2e^{f(x)}}_{\leq c\, f(x)^2 e^{f(x)}}\, d\nu(x) \qquad \text{[using Lemma]}$$

By Cauchy-Schwarz inequality,

$$\mathbb{E}_\nu[f^2e^f] \leq 2\,\mathbb{E}_\nu[f'^2e^f]^{\frac{1}{2}}\mathbb{E}_\nu[f^2e^f]^{\frac{1}{2}} + c\,\mathbb{E}_\nu[f^2e^f]$$

$$\Rightarrow (1-c)\,\mathbb{E}_\nu[f^2e^f]^{\frac{1}{2}} \leq 2\,\mathbb{E}_\nu[f'^2e^f]^{\frac{1}{2}}$$

$$\Rightarrow \mathbb{E}_\nu[f^2e^f]^{\frac{1}{2}} \underset{[*]}{\leq} \left(\frac{2}{1-c}\right)\mathbb{E}_\nu[f'^2e^f]^{\frac{1}{2}}.$$

Finally, we have:

$$\mathrm{Ent}_\nu(e^f) \leq \mathbb{E}_\nu[fe^f - e^f + 1] \leq \int \left|f'(x)e^{f(x)/2}\right|\left|f(x)e^{f(x)/2}\right| d\nu(x) \qquad \text{[triangle inequality on [*]]}$$

$$\leq \mathbb{E}_\nu[f'^2e^f]^{\frac{1}{2}}\,\mathbb{E}_\nu[f^2e^f]^{\frac{1}{2}} \qquad \text{[Cauchy-Schwarz inequality]}$$

$$\leq \left(\frac{2}{1-c}\right)\mathbb{E}_\nu[f'^2e^f]. \qquad \text{[using [*]]}$$

■

Modified LSI for product measure ν^n:

Using the tensorization of entropy, for every smooth enough $F: \mathbb{R}^n \to \mathbb{R}$ such that $\max_{1\leq i\leq n} |\partial_i F| \leq 1$ a.e. and every $\lambda \in \mathbb{R}$ with $|\lambda| \leq c < 1$, we have:

$$\mathrm{Ent}_{\nu^n}(e^{\lambda F}) \leq \frac{2\lambda^2}{1-c}\,\mathbb{E}_{\nu^n}\left[\sum_{i=1}^n (\partial_i F)^2 e^{\lambda F}\right] = \frac{2\lambda^2}{1-c}\,\mathbb{E}_{\nu^n}\left[|\nabla F|^2 e^{\lambda F}\right].$$

comes from $\partial_i(\lambda F) = \lambda(\partial_i F)$

In particular, setting $c = \frac{1}{2}$ gives:

$$\mathrm{Ent}_{\nu^n}(e^{\lambda F}) \leq 4\lambda^2\,\mathbb{E}_{\nu^n}\left[|\nabla F|^2 e^{\lambda F}\right].$$

We may use this to prove Talagrand's concentration inequality for ν^n.

⑥ Talagrand's Concentration Inequality for Exponential Measure:

Assume that we have $F : \mathbb{R}^n \to \mathbb{R}$ smooth enough so that $\max\limits_{1 \leq i \leq n} |\partial_i F| \leq 1$ a.e. and $|\nabla F|^2 = \sum_{i=1}^{n} (\partial_i F)^2 \leq \alpha^2$ a.e. . For every λ with $|\lambda| \leq \frac{1}{2}$, we have:

$$(\star) \quad \mathrm{Ent}_{\nu^n}(e^{\lambda F}) \leq 4\lambda^2 \, \mathbb{E}_{\nu^n}\left[|\nabla F|^2 e^{\lambda F}\right] \leq 4\lambda^2 \alpha^2 \, \mathbb{E}_{\nu^n}\left[e^{\lambda F}\right]. \quad \text{[using modified LSI]}$$

We can now use the Herbst argument. Define $H(\lambda) \triangleq \mathbb{E}_{\nu^n}[e^{\lambda F}]$ and $K(\lambda) \triangleq \frac{\log(H(\lambda))}{\lambda}$.
↖natural log.

$$\mathrm{Ent}_{\nu^n}(e^{\lambda F}) = \mathbb{E}_{\nu^n}\left[\lambda F e^{\lambda F}\right] - \mathbb{E}_{\nu^n}[e^{\lambda F}] \log\left(\mathbb{E}_{\nu^n}[e^{\lambda F}]\right)$$

$$= \lambda H'(\lambda) - H(\lambda)\log(H(\lambda)) \quad \text{[swap } \mathbb{E}[\cdot] \text{ \& derivative using DCT]}$$

$$\Rightarrow \lambda H'(\lambda) - H(\lambda)\log(H(\lambda)) \leq 4\lambda^2\alpha^2 H(\lambda) \quad \text{[using } (\star)\text{]}$$

$$\Rightarrow \underbrace{\frac{H'(\lambda)}{\lambda H(\lambda)} - \frac{\log(H(\lambda))}{\lambda^2}} \leq 4\alpha^2$$

$$\Rightarrow K'(\lambda) \leq 4\alpha^2$$

$$\Rightarrow K(\lambda) = \underbrace{K(0)}_{= \mathbb{E}_{\nu^n}[F]} + \int_0^\lambda K'(t)\,dt \leq \mathbb{E}_{\nu^n}[F] + \int_0^\lambda 4\alpha^2\,dt = \mathbb{E}_{\nu^n}[F] + 4\alpha^2\lambda$$

$$\Rightarrow H(\lambda) = \mathbb{E}_{\nu^n}[e^{\lambda F}] \leq \exp\left(\mathbb{E}_{\nu^n}[F]\lambda + 4\alpha^2\lambda^2\right) \quad \text{for } 0 \leq \lambda \leq \tfrac{1}{2}.$$

By Chernoff bound,

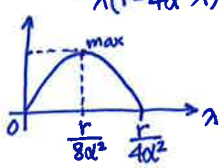
$$\nu^n\left(F \geq \mathbb{E}_{\nu^n}[F] + r\right) \leq \frac{\mathbb{E}_{\nu^n}[e^{\lambda F}]}{\exp(\mathbb{E}_{\nu^n}[F]\lambda + r\lambda)} \leq \underbrace{\exp\left(-r\lambda + 4\alpha^2\lambda^2\right)}, \quad \forall r \geq 0$$

same as Gaussian concentration proof until here

for $0 \leq \lambda \leq \frac{1}{2}$. Optimizing this bound over λ gives two cases:

$$\max_{0 \leq \lambda \leq \frac{1}{2}} \underbrace{r\lambda - 4\alpha^2\lambda^2}_{= \lambda(r - 4\alpha^2\lambda)} = \begin{cases} r\left(\frac{r}{8\alpha^2}\right) - 4\alpha^2\left(\frac{r}{8\alpha^2}\right)^2, & \text{if } \frac{r}{8\alpha^2} \leq \frac{1}{2} \\ r\left(\frac{1}{2}\right) - 4\alpha^2\left(\frac{1}{2}\right)^2, & \text{if } \frac{r}{8\alpha^2} > \frac{1}{2} \end{cases} = \begin{cases} \frac{r^2}{16\alpha^2}, & \text{if } r \leq 4\alpha^2 \\ \underbrace{\frac{r}{2} - \alpha^2}_{\geq \frac{r}{4}}, & \text{if } r > 4\alpha^2 \end{cases}.$$

$\geq \frac{r}{4}$ (because $\frac{r}{4} \leq \frac{r}{2} - \alpha^2 \Leftrightarrow r \leq 2r - 4\alpha^2 \Leftrightarrow r \geq 4\alpha^2$)



$$\Rightarrow \nu^n\left(F \geq \mathbb{E}_{\nu^n}[F] + r\right) \leq \begin{cases} \exp\left(-\frac{r^2}{16\alpha^2}\right), & \text{if } r \leq 4\alpha^2 \quad \leftarrow \text{Gaussian bound for small } r \\ \exp\left(-\frac{r}{4}\right), & \text{if } r > 4\alpha^2 \quad \leftarrow \text{Exponential bound for large } r \end{cases}$$

$$= \begin{cases} r^2/4\alpha^2, & \text{if } r^2/4\alpha^2 \leq r \Leftrightarrow r \leq 4\alpha^2 \\ r, & \text{if } r > 4\alpha^2 \end{cases}$$

$$\therefore \nu^n\left(F \geq \mathbb{E}_{\nu^n}[F] + r\right) \leq \exp\left(-\frac{1}{4}\min\left(r, \frac{r^2}{4\alpha^2}\right)\right) \quad \text{for every } r \geq 0.$$

By homogeneity, we get for every $F : \mathbb{R}^n \to \mathbb{R}$ smooth enough so that $|\nabla F|^2 \leq \alpha^2$ a.e. and $\max\limits_{1 \leq i \leq n} |\partial_i F| \leq \beta$ a.e.

$$\forall r \geq 0, \quad \nu^n\left(F \geq \mathbb{E}_{\nu^n}[F] + r\right) \leq \exp\left(-\frac{1}{16}\min\left(\frac{r}{\beta}, \frac{r^2}{\alpha^2}\right)\right).$$

↖Talagrand concentration inequality

⑦ General Modified LSI:

Let $(X, \mathcal{B}(X), \mu)$ be a probability space, and let $\mathcal{A}$ be a set of functions from X to $\mathbb{R}$ such that the ensuing objects are well-defined.
Let Γ be a "gradient" operator on $\mathcal{A}$ such that $\Gamma(f) \geq 0$ and $\Gamma(\lambda f) = \lambda^2 \Gamma(f)$ for all $f \in \mathcal{A}, \lambda \in \mathbb{R}$.
↳ like $|\nabla f|^2$

Def: We say that μ satisfies a modified LSI with respect to Γ if there is a function $B: \mathbb{R}_+ \to \mathbb{R}_+$ such that for all $f \in \mathcal{A}$ with $\|\Gamma(f)\|_\infty^{\frac{1}{2}} \leq \lambda$, (↑ sup-norm)

$$\mathrm{Ent}_\mu(e^f) \leq B(\lambda)\, \mathbb{E}_\mu[\Gamma(f) e^f].$$

Examples:

1. (Exponential measure ν ↑ on $\mathbb{R}$) $\Gamma(f) = f'^2$, $B(\lambda) = \frac{2}{1-c}$ (← constant for small λ) for $0 \leq \lambda \leq c$ for some $c < 1$.
2. (Gaussian measure $\gamma = \mathcal{N}(0,1)$ ↑ on $\mathbb{R}$) $\Gamma(f) = f'^2$, $B(\lambda) = \frac{1}{2}$ for $\lambda \geq 0$.

Gaussian LSI: $\forall f: \mathbb{R} \to \mathbb{R}$ with $\mathbb{E}_\gamma[f'^2] < \infty$, $\mathrm{Ent}_\gamma(f^2) \leq 2\, \mathbb{E}_\gamma[f'^2]$
$\Longrightarrow$ For $F: \mathbb{R} \to \mathbb{R}$ smooth, set $f^2 = e^F$: $\mathrm{Ent}_\gamma(e^F) \leq \frac{1}{2} \mathbb{E}_\gamma[F'^2 e^F]$.

New Feature:

LSI tensorizes in terms of the Lipschitz bound, but modified LSI tensorizes in terms of two parameters!

Given probability spaces $(X_i, \mathcal{B}(X_i), \mu_i)$, $1 \leq i \leq n$, let $P = \mu_1 \otimes \cdots \otimes \mu_n$ denote the product measure on $(X_1 \times \cdots \times X_n, \mathcal{B}(X_1) \otimes \cdots \otimes \mathcal{B}(X_n))$. Let Γ_i be the "gradient" operators on the corresponding spaces of functions $\mathcal{A}_i$ from X_i to $\mathbb{R}$. Finally, let $\mathcal{A}$ be the space of functions from $X_1 \times \cdots \times X_n$ to $\mathbb{R}$ such that for any $f \in \mathcal{A}$, $f_i \in \mathcal{A}_i$.

The tensorization of entropy yields:

Prop: Assume for every $f \in \mathcal{A}_i$ such that $\|\Gamma_i(f)\|_\infty^{\frac{1}{2}} \leq \lambda$,

$$\mathrm{Ent}_{\mu_i}(e^f) \leq B(\lambda)\, \mathbb{E}_{\mu_i}[\Gamma_i(f) e^f] \quad \text{[modified LSI in each space]}$$

for $i = 1, \ldots, n$. Then, for every $f \in \mathcal{A}$ such that $\max_{1 \leq i \leq n} \|\Gamma_i(f_i)\|_\infty^{\frac{1}{2}} \leq \lambda$,

$$\mathrm{Ent}_P(e^f) \leq B(\lambda)\, \mathbb{E}_P\left[\sum_{i=1}^n \Gamma_i(f_i) e^f\right]. \quad \left(\leq B(\lambda) \left\|\sum_{i=1}^n \Gamma_i(f_i)\right\|_\infty \mathbb{E}_P[e^f]\right)$$

So, this modified LSI tensorizes in terms of two parameters:

$$\max_{1 \leq i \leq n} \|\Gamma_i(f_i)\|_\infty^{\frac{1}{2}} \quad \text{and} \quad \left\|\sum_{i=1}^n \Gamma_i(f_i)\right\|_\infty.$$

↑ like $\max_{1 \leq i \leq n} |\partial_i f|$ ↑ like $|\nabla f|^2$

Depending on the structure of $B(\lambda)$, the corresponding concentration inequality can have one or two "behaviours" (eg: min). If B is bounded for small λ, then we get inequalities like that of ν^n, and if B is bounded for all $\lambda \geq 0$, then we get inequalities like that of the Gaussian measure.

[THE END]