# Summary: Part 2

Tianpei Xie

Jan. 26th., 2023

## Contents

# 1 Logarithmic Sobolev Inequality

## 1.1 Functional Form of Logarithmic Sobolev Inequality

- From functional analysis, we have *the Sobolev inequality*,

  **Remark** (***The Sobolev Inequality***) [Evans, 2010]
  ***The Sobolev inequality*** states for smooth function $f : \mathbb{R}^n \to \mathbb{R}$ in *Sobolev space* where $n \geq 3$ and $p = \frac{2n}{n-2} > 2$

  $$\|f\|_p^2 \leq C_n \int_{\mathbb{R}^n} |\nabla f|^2 \, dx.$$

  The inequality is sharp when the constant

  $$C_n := \frac{1}{\pi n(n-2)} \left(\frac{\Gamma(n)}{\Gamma(n/2)}\right)^{2/n}$$

- **Proposition 1.1** *(**Euclidean Logarithmic Sobolev Inequality**).*
  *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function and $m$ be Lebesgue measure on $\mathbb{R}^n$, then*

  $$\begin{aligned} Ent_m(f^2) &\leq \frac{n}{2} \log\left(\frac{2}{n\pi e} \mathbb{E}_m\left[\|\nabla f\|_2^2\right]\right) \quad (1)\\ \Leftrightarrow \int f^2 \log\left(\frac{f^2}{\int f^2 dx}\right) dx &\leq \frac{n}{2} \log\left(\frac{2}{n\pi e} \int |\nabla f|^2 \, dx\right) \end{aligned}$$

- **Definition** (***Logarithmic Sobolev Inequality for General Probability Measure***).
  *A probability measure* $\mu$ on $\mathbb{R}^n$ is said to satisfy the ***logarithmic Sobolev inequality*** for some constant $C > 0$ if for any smooth function $f$

  $$\mathrm{Ent}_\mu(f^2) \leq C \, \mathbb{E}_\mu\left[\|\nabla f\|_2^2\right] \quad (2)$$

  holds for any ***continuous differentiable*** function $f : \mathbb{R}^n \to \mathbb{R}$. The left-hand side is called ***the entropy functional***, which is defined as

  $$\begin{aligned} \mathrm{Ent}(f^2) &:= \mathbb{E}_\mu\left[f^2 \log f^2\right] - \mathbb{E}_\mu\left[f^2\right] \log \mathbb{E}_\mu\left[f^2\right] \\ &= \int f^2 \log\left(\frac{f^2}{\int f^2 d\mu}\right) d\mu. \end{aligned}$$

  The right-hand side is defined as

  $$\mathbb{E}_\mu\left[\|\nabla f\|_2^2\right] = \int \|\nabla f\|_2^2 \, d\mu.$$

  Thus we can rewrite *the logarithmic Sobolev inequality* in *functional form*

  $$\int f^2 \log\left(\frac{f^2}{\int f^2 d\mu}\right) d\mu \leq C \int \|\nabla f\|_2^2 \, d\mu \quad (3)$$

- **Remark** (***Logarithmic Sobolev Inequality***)
For non-negative function $f$, we can replace $f \to \sqrt{f}$, so that *the logarithmic Sobolev inequality* becomes

$$\mathrm{Ent}_{\mu}(f) \le C \int \frac{\|\nabla f\|_2^2}{f} d\mu \tag{4}$$

- **Remark** (***Modified Logarithmic Sobolev Inequality via Convex Cost and Duality***)
For some ***convex non-negative cost*** $c : \mathbb{R}^n \to \mathbb{R}_+$, ***the convex conjugate*** of $c$ (Legendre transform of $c$) is defined as

$$c^*(x) := \sup_y \{\langle x\,,\, y\rangle - c(y)\}$$

Then we can obtain ***the modified logarithmic Sobolev inequality***

$$\mathrm{Ent}_{\mu}(f) \le \int f^2\, c^* \left(\frac{\nabla f}{f}\right) d\mu \tag{5}$$

## 1.2 Bernoulli Logarithmic Sobolev Inequality

- **Remark** (*Setting*)
Consider a ***uniformly distributed binary vector*** $Z = (Z_1\,, \ldots, Z_n)$ on the hypercube $\{-1, +1\}^n$. In other words, the components of $X$ are *independent*, *identically distributed* ***random sign (Rademacher) variables*** with $\mathbb{P}\{Z_i = -1\} = \mathbb{P}\{Z_i = +1\} = 1/2$ (i.e. *symmetric Bernoulli random variables*).

Let $f : \{-1, +1\}^n \to \mathbb{R}$ be a real-valued function on ***binary hypercube***. $X := f(Z)$ is an induced real-valued random variable. Define $\widetilde{Z}^{(i)} = (Z_1\,, \ldots, Z_{i-1}, Z_i', Z_{i+1}\,, \ldots, Z_n)$ be the sample $Z$ with $i$-th component replaced by an *independent copy* $Z_i'$. Since $Z, \widetilde{Z}^{(i)} \in \{-1, +1\}^n$, $\widetilde{Z}^{(i)} = (Z_1\,, \ldots, Z_{i-1}, -Z_i, Z_{i+1}\,, \ldots, Z_n)$, i.e. *the $i$-th sign is* ***flipped***. Also denote the $i$-th *Jackknife sample* as $Z_{(i)} = (Z_1\,, \ldots, Z_{i-1}, Z_{i+1}\,, \ldots, Z_n)$ by *leaving out* the $i$-th component. $\mathbb{E}_{(-i)}[X] := \mathbb{E}\left[X | Z_{(i)}\right]$.

Denote the $i$-th component of ***discrete gradient*** of $f$ as

$$\nabla_i f(z) := \frac{1}{2}\left(f(z) - f(\widetilde{z}^{(i)})\right)$$

and $\nabla f(z) = (\nabla_1 f(z)\,, \ldots, \nabla_n f(z))$

- **Proposition 1.2** *(**Logarithmic Sobolev Inequality for Rademacher Random Variables**). [Boucheron et al., 2013]*
*If $f : \{-1, +1\}^n \to \mathbb{R}$ be an arbitrary real-valued function defined on the $n$-dimensional* ***binary hypercube*** *and assume that $Z$ is* ***uniformly distributed*** *over $\{-1, +1\}^n$. Then*

$$Ent(f^2) \le \mathcal{E}(f) \tag{6}$$

$$\Leftrightarrow Ent(f^2(Z)) \le 2\mathbb{E}\left[\|\nabla f(Z)\|_2^2\right] \tag{7}$$

- **Remark** (***Logarithmic Sobolev Inequality*** $\Rightarrow$ ***Efron-Stein Inequality***). [Boucheron et al., 2013]

Note that for $f$ non-negative,

$$\mathrm{Var}(f(Z)) \le \mathrm{Ent}(f^2(Z)).$$

Thus *logarithmic Sobolev inequality* (6) implies

$$\mathrm{Var}(f(Z)) \le \mathcal{E}(f)$$

which is *the Efron-Stein inequality.*

- **Corollary 1.3** ***(Logarithmic Sobolev Inequality for Asymmetric Bernoulli Random Variables)****. [Boucheron et al., 2013]*
  *If $f : \{-1,+1\}^n \to \mathbb{R}$ be an arbitrary real-valued function and $Z = (Z_1, \ldots, Z_n) \in \{-1,+1\}^n$ with $p = \mathbb{P}\{Z_i = +1\}$. Then*

  $$\mathit{Ent}(f^2) \le c(p)\mathbb{E}\left[\|\nabla f(Z)\|_2^2\right] \tag{8}$$

  *where*

  $$c(p) = \frac{1}{1-2p}\log\frac{1-p}{p}$$

  *Note that $\lim_{p\to 1/2} c(p) = 2$.*

### 1.3 Gaussian Logarithmic Sobolev Inequality

- **Proposition 1.4** ***(Gaussian Logarithmic Sobolev Inequality)****. [Boucheron et al., 2013]*
  *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a* ***continuous differentiable*** *function and let $Z = (Z_1, \ldots, Z_n)$ be a vector of $n$* ***independent standard Gaussian*** *random variables. Then*

  $$\mathit{Ent}(f^2(Z)) \le 2\mathbb{E}\left[\|\nabla f(Z)\|_2^2\right]. \tag{9}$$

### 1.4 Modified Logarithmic Sobolev Inequalities

- **Proposition 1.5** ***(A Modified Logarithmic Sobolev Inequalities for Moment Generating Function)*** *[Boucheron et al., 2013]*
  *Consider independent random variables $Z_1, \ldots, Z_n$ taking values in $\mathcal{X}$, a real-valued function $f : \mathcal{X}^n \to \mathbb{R}$ and the random variable $X = f(Z_1, \ldots, Z_n)$. Also denote $Z_{(-i)} = (Z_1, \ldots, Z_{i-1}, Z_{i+1}, \ldots, Z_n)$ and $X_{(-i)} = f_i(Z_{(-i)})$ where $f_i : \mathcal{X}^{n-1} \to \mathbb{R}$ is an arbitrary function. Let $\phi(x) = e^x - x - 1$. Then for all $\lambda \in \mathbb{R}$,*

  $$\mathit{Ent}(e^{\lambda X}) := \mathbb{E}\left[\lambda X e^{\lambda X}\right] - \mathbb{E}\left[e^{\lambda X}\right]\log\mathbb{E}\left[e^{\lambda X}\right] \le \sum_{i=1}^{n}\mathbb{E}\left[e^{\lambda X}\phi(-\lambda(X - X_{(-i)}))\right] \tag{10}$$

- **Proposition 1.6** ***(Symmetrized Modified Logarithmic Sobolev Inequalities)*** *[Boucheron et al., 2013]*

*Consider independent random variables $Z_1, \dots, Z_n$ taking values in $\mathcal{X}$, a real-valued function $f : \mathcal{X}^n \to \mathbb{R}$ and the random variable $X = f(Z_1, \dots, Z_n)$. Also denote $\widetilde{X}^{(i)} = f(Z_1, \dots, Z_{i-1}, Z_i', Z_{i+1}, \dots, Z_n)$. Let $\phi(x) = e^x - x - 1$. Then for all $\lambda \in \mathbb{R}$,*

$$\lambda \mathbb{E}\left[X e^{\lambda X}\right] - \mathbb{E}\left[e^{\lambda X}\right] \log \mathbb{E}\left[e^{\lambda X}\right] \le \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda X} \phi(-\lambda(X - \widetilde{X}^{(i)}))\right] \tag{11}$$

*Moreover, denoting $\tau(x) = x(e^x - 1)$, for all $\lambda \in \mathbb{R}$,*

$$\lambda \mathbb{E}\left[X e^{\lambda X}\right] - \mathbb{E}\left[e^{\lambda X}\right] \log \mathbb{E}\left[e^{\lambda X}\right] \le \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda X} \tau(-\lambda(X - \widetilde{X}^{(i)})_+)\right],$$

$$\lambda \mathbb{E}\left[X e^{\lambda X}\right] - \mathbb{E}\left[e^{\lambda X}\right] \log \mathbb{E}\left[e^{\lambda X}\right] \le \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda X} \tau(\lambda(\widetilde{X}^{(i)} - X)_-)\right].$$

# 2 Isoperimetric Inequalities and Concentration of Measure

## 2.1 Brunn-Minkowski Inequality

- **Definition** (***Minkowski Sum of Sets***)
  Consider sets $A, B \subseteq \mathbb{R}^n$ and define ***the Minkowski sum*** of $A$ and $B$ as the set of all vectors in $\mathbb{R}^n$ formed by sums of elements of $A$ and $B$:

  $$A + B := \{x + y : x \in A, y \in B\}$$

  Similarly, for $c \in \mathbb{R}$, let $cA = \{cx : x \in A\}$. Denote by $\mathrm{Vol}(A)$ the ***Lebesgue measure*** of a *(measurable) set $A \subset \mathbb{R}^n$*.

- **Theorem 2.1** ***(The Prékopa-Leindler Inequality)****. [Boucheron et al., 2013, Wainwright, 2019]*
  *Let $\lambda \in (0,1)$, and let $f, g, h : \mathbb{R}^n \to [0, \infty)$ be* ***non-negative measurable functions*** *such that for all $x, y \in \mathbb{R}^n$,*

  $$h\left(\lambda x + (1-\lambda) y\right) \ge f(x)^{\lambda} g(y)^{1-\lambda}.$$

  *Then*

  $$\int_{\mathbb{R}^n} h(x) dx \ge \left(\int_{\mathbb{R}^n} f(x) dx\right)^{\lambda} \left(\int_{\mathbb{R}^n} g(x) dx\right)^{1-\lambda}. \tag{12}$$

- **Corollary 2.2** ***(Weaker Brunn-Minkowski Inequality)*** *[Boucheron et al., 2013, Wainwright, 2019]*
  *Let $A, B \subset \mathbb{R}^n$ be* ***non-empty compact sets****. Then for all $\lambda \in [0,1]$,*

  $$Vol\left(\lambda A + (1-\lambda) B\right) \ge Vol(A)^{\lambda}\, Vol(B)^{1-\lambda}. \tag{13}$$

- **Theorem 2.3** ***(Brunn-Minkowski Inequality)*** *[Boucheron et al., 2013, Vershynin, 2018, Wainwright, 2019]*
  *Let $A, B \subset \mathbb{R}^n$ be* ***non-empty compact sets****. Then for all $\lambda \in [0,1]$,*

  $$Vol\left(\lambda A + (1-\lambda) B\right)^{\frac{1}{n}} \ge \lambda\, Vol(A)^{\frac{1}{n}} + (1-\lambda)\, Vol(B)^{\frac{1}{n}}. \tag{14}$$

## 2.2 Classical Isoperimetric Problem on Euclidean Space $\mathbb{R}^n$

- **Definition** (***Blowup of Sets***)
  For any $t > 0$, and any (measurable) sets $A \subset \mathbb{R}^n$, ***the $t$-blowup (or, $t$-enlargement)*** *of* $A$ is defined by
  $$A_t := \{x \in \mathbb{R}^n : d(x, A) < t\} = A + t\,B$$
  where $B = \{x \in \mathbb{R}^n : d(0, x) < 1\}$ is an *open unit ball* and $d(x, A) = \inf_{y \in A} d(x, y)$.

- **Definition** (***Surface Area of Sets***)
  let $A \subset \mathbb{R}^n$ be a measurable set and denote by $\mathrm{Vol}(A)$ its *Lebesgue measure. The* ***surface area*** *of* $A$ is defined by
  $$\mathrm{Vol}(\partial A) = \lim_{t \to 0} \frac{\mathrm{Vol}(A_t) - \mathrm{Vol}(A)}{t}.$$
  provided that the limit exists. Here $A_t$ denotes *the $t$-blowup* of $A$.

- **Remark** (***Isoperimetry Theorem***)
  The classical isoperimetric theorem in $\mathbb{R}^n$ states that, among all sets with ***a given volume***, ***the Euclidean unit ball minimizes the surface area***. This theoerm can be formally stated as below:

- **Theorem 2.4** ***(Isoperimetry Theorem)*** *[Boucheron et al., 2013, Vershynin, 2018, Wainwright, 2019]*
  *Let $A \subset \mathbb{R}^n$ be such that Vol($A$) = Vol($B$) where $B := \{x \in \mathbb{R}^n : d(0, x) < 1\}$ is an unit ball. Then for any $t > 0$,*
  $$Vol(A_t) \ge Vol(B_t) \tag{15}$$
  *Moreover, if Vol($\partial A$) exists, then*
  $$Vol(\partial A) \ge Vol(\partial B). \tag{16}$$

- **Example** (***Concentration of Lebesgue Measure in $\mathbb{R}^n$ and Isoperimetric Inequality***)
  Note that the volume of a *$t$-ball in* $\mathbb{R}^n$ is
  $$\mathrm{Vol}(tB) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2} + 1\right)} t^n \equiv c_n t^n$$
  Thus the radius of ball $B$ with the same volume of $A$ is
  $$r := \left(\frac{\mathrm{Vol}(A)}{c_n}\right)^{\frac{1}{n}}.$$
  ***The classical isoperimetric inequality*** states that
  $$\begin{aligned} \mathrm{Vol}(A_t)) &\ge \left((r + t)\mathrm{Vol}(B)^{1/n}\right)^n \\ \Leftrightarrow \mathrm{Vol}(A_t) &\ge c_n \left(\left(\frac{\mathrm{Vol}(A)}{c_n}\right)^{\frac{1}{n}} + t\right)^n \\ \Leftrightarrow \left(\frac{\mathrm{Vol}(A_t)}{c_n}\right)^{\frac{1}{n}} &\ge \left(\frac{\mathrm{Vol}(A)}{c_n}\right)^{\frac{1}{n}} + t \end{aligned} \tag{17}$$

- **Definition** (***Isoperimetric Function of Probability Measure***)
  Define ***the isoperimetric function*** of *the Lebesgue measure space* $(\mathbb{R}^n, \mu)$ as

$$\lambda(u) := \left(\frac{u}{c_n}\right)^{\frac{1}{n}}$$

so *the classical isoperimetric inequality* is equivalent to *the concentration of Lebesgue measure*

$$\lambda\left(\mu(A_t)\right) \geq \lambda\left(\mu(A)\right) + t.$$

## 2.3 Isoperimetric Problem on Unit Sphere

- **Definition** (***Spherical Cap and its $t$-Blowup***)
  Let $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}$ be the $(n-1)$-*dimensional* ***unit sphere***. The ***intersection*** of a ***half-space*** and $\mathbb{S}^{n-1}$ is called a ***spherical cap***. In particular, for some $y \in \mathbb{R}^n$, denote the associated spherical cap as

$$H_y := \left\{x \in \mathbb{S}^{n-1} : \langle x\,,\, y\rangle \leq 0\right\}$$

With some simple geometry, it can be shown that its $t$-*blowup* corresponds to the set

$$H_y^t := \left\{x \in \mathbb{S}^{n-1} : \langle x\,,\, y\rangle < \sin(t)\right\}$$

- **Theorem 2.5** *(**Isoperimetry Theorem on Unit Sphere**) [Boucheron et al., 2013, Vershynin, 2018, Wainwright, 2019]*
  *Let $A$ be a subset of the sphere $\mathbb{S}^{n-1}$, and let $\sigma$ denote the **normalized area** on that sphere. Let $t > 0$. Then, among all sets $A \subset \mathbb{S}^{n-1}$ with given area $\sigma(A)$, the **spherical caps minimize the area of the neighborhood** $\sigma(A_t)$, where*

$$A_t := \left\{x \in \mathbb{S}^{n-1} : \exists y \in A \text{ such that } \|x - y\| < t\right\}$$

- **Remark** Define a *metric* $\rho$ on sphere $\mathbb{S}^{n-1}$ as

$$\rho(x, y) := \arccos(\langle x\,,\, y\rangle)$$

  Thus $(\mathbb{S}^{n-1}, \rho)$ is a ***metric space***. Let $\mathbb{P}$ be uniform distribution on $\mathbb{S}^{n-1}$ so that $((\mathbb{S}^{n-1}, \rho), \mathbb{P})$ is a probability space.

- **Proposition 2.6** *(**Isoperimetric Inequalities for Uniform Distribution over Sphere**) [Boucheron et al., 2013, Vershynin, 2018, Wainwright, 2019]*
  *Let $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}$ be the $(n-1)$-dimensional **unit sphere**. For any $t \in [0, 1]$,*

$$\alpha_{\mathbb{S}^{n-1}}(t) \leq c\exp\left(-\frac{nt^2}{2}\right) \tag{18}$$

  *for some constant $c$.*

### 2.4 Concentration via Isoperimetric Inequalities

- **Definition** (***Isoperimetry Problem***) [Boucheron et al., 2013]
  Given a ***metric space*** $\mathcal{X}$ with corresponding *distance* $d$, consider ***the measure space*** formed by $\mathcal{X}$ , *the $\sigma$-algebra* of all ***Borel sets** of* $\mathcal{X}$, and a probability measure $\mathbb{P}$. Let $X$ be a *random variable* taking values in $\mathcal{X}$, distributed according to $\mathbb{P}$.

  ***The isoperimetric problem*** in this case is the following: given $p \in (0,1)$ and $t > 0$, ***determine the sets*** $A$ with $\mathbb{P}[X \in A] \geq p$ for which *the measure*

  $$\mathbb{P}[d(X,A) \geq t]$$

  is ***maximal***.

- **Remark** (***Isoperimetric Inequalities***)
  Even though the exact solution is only known in a few special cases, useful *bounds* for $\mathbb{P}[d(X,A) \geq t]$ can be derived under remarkably general circumstances. *Such bounds are usually referred to as* ***isoperimetric inequalities***.

- **Definition** (***Concentration Function***) [Boucheron et al., 2013, Wainwright, 2019]
  ***The concentration function*** $\alpha : [0,\infty) \to \mathbb{R}_{+}$ associated with ***metric measure space*** $((\mathcal{X},d),\mathbb{P})$ is given by

  $$\alpha_{\mathbb{P},(\mathcal{X},d)}(t) := \sup_{A \subset \mathcal{X}:\, \mathbb{P}(A) \geq \frac{1}{2}} \mathbb{P}[d(X,A) \geq t] = \sup_{A \subset \mathcal{X}:\, \mathbb{P}(A) \geq \frac{1}{2}} \mathbb{P}(A_t^c)$$

  where $A_t := A + tB = \{x \in \mathcal{X} : d(x,A) < t\}$ is *the $t$-blowup* of $A \subset \mathcal{X}$. We simply denote it as $\alpha(t)$.

  Thus the optimal $A^*$ for isoperimetry problem is the one that attains the $\alpha(t) = \mathbb{P}(A_t^c)$.

- **Theorem 2.7** ***(Levy's Inequalities)****[Boucheron et al., 2013, Wainwright, 2019]*
  *For any Lipschitz function $f : \mathcal{X} \to \mathbb{R}$,*

  $$\begin{aligned} &\mathbb{P}\{f(X) \geq Med(f(X)) + t\} \leq \alpha_{\mathbb{P}}(t) \\ &\mathbb{P}\{f(X) \leq Med(f(X)) - t\} \leq \alpha_{\mathbb{P}}(t). \end{aligned} \tag{19}$$

  *where $Med(f(X))$ is* ***the median*** *of $f(X)$, i.e.*

  $$\mathbb{P}\{f(X) \leq Med(f(X)\} \geq \frac{1}{2}, \quad \textit{and} \quad \mathbb{P}\{f(X) \geq Med(f(X)\} \geq \frac{1}{2}.$$

### 2.5 Convex Distance Inequality

### 2.6 Concentration of Convex Lipschitz Functions

## 3 Concentration of Gaussian Measure

### 3.1 Gaussian Isoperimetric Theorem and Gaussian Concentration Theorem

### 3.2 Lipschitz Functions of Gaussian Variables

### 3.3 Gaussian Logarithmic Sobolev Inequality

### 3.4 Gaussian Transportation Inequality

### 3.5 Gaussian Hypercontractivity

### 3.6 Suprema of Gaussian Process

## 4 Concentration of Bernoulli Measure on the Binary Hypercube

### 4.1 Edge Isoperimetric Inequality on the Binary Hypercube

### 4.2 Bobkov's Inequality

### 4.3 Vertex Isoperimetric Inequality on the Binary Hypercube

### 4.4 Hypercontractivity: The Bonami-Beckner Inequality

### 4.5 Influence Function

### 4.6 Monotone Sets

### 4.7 Thresholding Phenonemon

## References

Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

Lawrence C Evans. *Partial differential equations*, volume 19. American Mathematical Soc., 2010.

Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.