

# Lecture 4: The Entropy Methods

Tianpei Xie

Jan. 19th., 2023

## Contents

<b>1</b>	<b>Logarithmic Sobolev Inequality</b>	<b>2</b>
1.1	Bernoulli Logarithmic Sobolev Inequality . . . . .	2
1.2	Gaussian Logarithmic Sobolev Inequality . . . . .	4
1.3	Information Theory Interpretation . . . . .	6
1.4	Logarithmic Sobolev Inequality for General Probability Measures . . . . .	7
1.5	Variants of Logarithmic Sobolev Inequalities . . . . .	8
1.6	Poisson Logarithmic Sobolev Inequality . . . . .	8
1.7	Applications . . . . .	9
<b>2</b>	<b>The Entropy Methods</b>	<b>9</b>
2.1	Tensorization Property of $\Phi$ -Entropy . . . . .	9
2.2	Herbst's Argument . . . . .	9
2.3	Bounded Difference Inequality . . . . .	9
2.4	Modified Logarithmic Sobolev Inequalities . . . . .	9
2.5	Concentration of Convex Lipschitz Functions . . . . .	9
2.6	Exponential Tail Bounds for Self-Bounding Functions . . . . .	9

# 1 Logarithmic Sobolev Inequality

## 1.1 Bernoulli Logarithmic Sobolev Inequality

- **Remark (Setting)**

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

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

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

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

- **Remark (Jackknife Estimate of Variance)**

Recall that *the Jackknife estimate of variance*

$$\begin{aligned} \mathcal{E}(f) &:= \mathbb{E} \left[ \sum_{i=1}^n \left( f(Z) - \mathbb{E}_{(-i)} \left[ f(\tilde{Z}^{(i)}) \right] \right)^2 \right] \\ &= \frac{1}{2} \mathbb{E} \left[ \sum_{i=1}^n \left( f(Z) - f(\tilde{Z}^{(i)}) \right)^2 \right]. \end{aligned}$$

Using the notation of discrete gradient of  $f$ , we see that

$$\mathcal{E}(f) := 2\mathbb{E} \left[ \|\nabla f(Z)\|_2^2 \right]$$

- **Remark (Entropy Functional)**

Recall that the entropy functional for  $f$  is defined as

$$H_\Phi(f(Z)) = \text{Ent}(f) := \mathbb{E} [f(Z) \log f(Z)] - \mathbb{E} [f(Z)] \log (\mathbb{E} [f(Z)]).$$

- **Proposition 1.1 (Logarithmic Sobolev Inequality for Rademacher Random Variables).** [Boucheron et al., 2013]

If  $f : \{-1, +1\}^n \rightarrow \mathbb{R}$  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

$$\text{Ent}(f^2) \leq \mathcal{E}(f) \tag{1}$$

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

**Proof:** The key is to apply the tensorization property of  $\Phi$ -entropy. Let  $X = f(Z)$ . By tensorization property,

$$\text{Ent}(X^2) \leq \sum_{i=1}^n \mathbb{E} [\text{Ent}_{(-i)}(X^2)]$$

where  $\text{Ent}_{(-i)}(X^2) := \mathbb{E}_{(-i)} [X^2 \log X^2] - \mathbb{E}_{(-i)} [X^2] \log (\mathbb{E}_{(-i)} [X^2])$ .

It thus suffice to show that for all  $i = 1, \dots, n$ ,

$$\text{Ent}_{(-i)}(X^2) \leq \frac{1}{2} \mathbb{E}_{(-i)} \left[ \left( f(Z) - f(\tilde{Z}^{(i)}) \right)^2 \right].$$

Given any fixed realization of  $Z_{(-i)}$ ,  $X = f(Z) = \tilde{f}(Z_i)$  can only takes two different values with equal probability. Call these two values  $a$  and  $b$ . See that

$$\begin{aligned} \text{Ent}_{(-i)}(X^2) &= \frac{1}{2} a^2 \log a^2 + \frac{1}{2} b^2 \log b^2 - \frac{1}{2} (a^2 + b^2) \log \left( \frac{a^2 + b^2}{2} \right) \\ \frac{1}{2} \mathbb{E}_{(-i)} \left[ \left( f(Z) - f(\tilde{Z}^{(i)}) \right)^2 \right] &= \frac{1}{2} (a - b)^2. \end{aligned}$$

Thus we need to show

$$\frac{1}{2} a^2 \log a^2 + \frac{1}{2} b^2 \log b^2 - \frac{1}{2} (a^2 + b^2) \log \left( \frac{a^2 + b^2}{2} \right) \leq \frac{1}{2} (a - b)^2.$$

By symmetry, we may assume that  $a \geq b$ . Since  $(|a| - |b|)^2 \leq (a - b)^2$ , without loss of generality, we may further assume that  $a, b \geq 0$ .

Define

$$h(a) := \frac{1}{2} a^2 \log a^2 + \frac{1}{2} b^2 \log b^2 - \frac{1}{2} (a^2 + b^2) \log \left( \frac{a^2 + b^2}{2} \right) - \frac{1}{2} (a - b)^2$$

for  $a \in [b, \infty)$ .  $h(b) = 0$ . It suffice to check that  $h'(b) = 0$  and that  $h$  is concave on  $[b, \infty)$ . Note that

$$\begin{aligned} h'(a) &= a \log a^2 + 1 - a \log \left( \frac{a^2 + b^2}{2} \right) - 1 - (a - b) \\ &= a \log \frac{2a^2}{(a^2 + b^2)} - (a - b). \end{aligned}$$

So  $h'(b) = 0$ . Moreover,

$$h''(a) = \log \frac{2a^2}{(a^2 + b^2)} + 1 - \frac{2a^2}{(a^2 + b^2)} \leq 0$$

due to inequality  $\log(x) + 1 \leq x$ . ■

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

Note that for  $f$  non-negative,

$$\text{Var}(f(Z)) \leq \text{Ent}(f^2(Z)).$$

Thus *logarithmic Sobolev inequality* (1) implies

$$\text{Var}(f(Z)) \leq \mathcal{E}(f)$$

which is the *Efron-Stein inequality*.

- **Corollary 1.2** (*Logarithmic Sobolev Inequality for Asymmetric Bernoulli Random Variables*). [Boucheron et al., 2013]  
If  $f : \{-1, +1\}^n \rightarrow \mathbb{R}$  be an arbitrary real-valued function and  $Z = (Z_1, \dots, Z_n) \in \{-1, +1\}^n$  with  $p = \mathbb{P}\{Z_i = +1\}$ . Then

$$\text{Ent}(f^2) \leq c(p) \mathbb{E} \left[ \|\nabla f(Z)\|_2^2 \right] \quad (3)$$

where

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

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

## 1.2 Gaussian Logarithmic Sobolev Inequality

- **Proposition 1.3** (*Gaussian Logarithmic Sobolev Inequality*). [Boucheron et al., 2013]  
Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a **continuous differentiable** function and let  $Z = (Z_1, \dots, Z_n)$  be a vector of  $n$  **independent standard Gaussian** random variables. Then

$$\text{Ent}(f^2(Z)) \leq 2 \mathbb{E} \left[ \|\nabla f(Z)\|_2^2 \right]. \quad (4)$$

**Proof:** We first prove for  $n = 1$ , where  $f : \mathbb{R} \rightarrow \mathbb{R}$  is continuous differentiable and  $Z$  is standard Gaussian distribution. Without loss of generality, assume that  $\mathbb{E}[f'(Z)] < \infty$  since it is trivial when  $\mathbb{E}[f'(Z)] = \infty$ . By density argument, it suffice to prove the proposition when  $f$  is twice differentiable with bounded support.

Now let  $\epsilon_1, \dots, \epsilon_n$  be independent Rademacher random variables and introduce

$$S_n := \frac{1}{\sqrt{n}} \sum_{j=1}^n \epsilon_j.$$

Note that  $\epsilon_i \in \{-1, +1\}$  with equal probability, thus

$$\begin{aligned} \mathbb{E}_{(-i)}[S_n] &= \frac{1}{2} \left[ \left( \frac{1}{\sqrt{n}} \sum_{j \neq i} \epsilon_j + \frac{1}{\sqrt{n}} \right) + \left( \frac{1}{\sqrt{n}} \sum_{j \neq i} \epsilon_j - \frac{1}{\sqrt{n}} \right) \right] \\ &= \frac{1}{2} \left[ \left( S_n + \frac{1 - \epsilon_i}{\sqrt{n}} \right) + \left( S_n - \frac{1 + \epsilon_i}{\sqrt{n}} \right) \right]. \end{aligned}$$

In the proof of Gaussian Poincaré inequality, we show that by *central limit theorem*,

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[ \sum_{i=1}^n \left| f(S_n) - f\left(S_n - \frac{2\epsilon_i}{\sqrt{n}}\right) \right|^2 \right] = 4 \mathbb{E}[(f'(Z))^2].$$

On the other hands, for any *continuous uniformly bounded function*  $f$ , by *central limit theorem*,

$$\lim_{n \rightarrow \infty} \text{Ent} (f^2(S_n)) = \text{Ent}(f^2(Z))$$

The proof is then completed by invoking *the logarithmic Sobolev inequality* for *Rademacher random variables*

$$\begin{aligned} \text{Ent} (f^2(S_n)) &\leq \frac{1}{2} \mathbb{E} \left[ \sum_{i=1}^n \left| f(S_n) - f \left( S_n - \frac{2\epsilon_i}{\sqrt{n}} \right) \right|^2 \right] \\ \Rightarrow \lim_{n \rightarrow \infty} \text{Ent} (f^2(S_n)) &\leq \frac{1}{2} \lim_{n \rightarrow \infty} \mathbb{E} \left[ \sum_{i=1}^n \left| f(S_n) - f \left( S_n - \frac{2\epsilon_i}{\sqrt{n}} \right) \right|^2 \right] \\ &\Rightarrow \text{Ent}(f^2(Z)) \leq 2\mathbb{E} [(f'(Z))^2]. \end{aligned}$$

The extension of the result to dimension  $n \geq 1$  follows easily from *the sub-additivity of entropy* which states that

$$\text{Ent}(f^2) \leq \sum_{i=1}^n \mathbb{E} [\mathbb{E}_{(-i)} [f^2(Z) \log f^2(Z)] - \mathbb{E}_{(-i)} [f^2(Z)] \log \mathbb{E}_{(-i)} [f^2(Z)]]$$

where  $\mathbb{E}_{(-i)} [\cdot]$  denotes the integration with respect to  $i$ -th variable  $Z_i$  only. Thus by induction, for all  $i$

$$\mathbb{E}_{(-i)} [f^2(Z) \log f^2(Z)] - \mathbb{E}_{(-i)} [f^2(Z)] \log \mathbb{E}_{(-i)} [f^2(Z)] \leq 2\mathbb{E}_{(-i)} [(\partial_i f(Z))^2].$$

Thus

$$\text{Ent}(f^2) \leq 2\mathbb{E} \left[ \mathbb{E}_{(-i)} \left[ \sum_{i=1}^n (\partial_i f(Z))^2 \right] \right] = 2\mathbb{E} [\|\nabla f(Z)\|_2^2]. \quad \blacksquare$$

- **Remark (*Dimension Free Property*).**

The *Gaussian logarithmic Sobolev inequality* has a constant  $C = 2$  that is ***independent of dimension***  $n$ :

$$\mathbb{E}_\mu [f^2] \leq 2\mathbb{E}_\mu [\|\nabla f\|_2^2].$$

This *dimension-free property* is related to the ***concentration of Gaussian measure***  $\mu$ . As a consequence, this inequality can be extended to functions of *Gaussian measure* on ***infinite dimensional space***, such as Gibbs measure, *Gaussian process* etc.

- **Remark (*Equivalent Form of Gaussian Logarithmic Sobolev Inequality*)**

Assume  $f : \mathbb{R}^n \rightarrow (0, \infty)$  and  $\int_{\mathbb{R}^n} f d\mu = 1$  under Gaussian measure  $\mu$ . Substituting  $f \rightarrow \sqrt{f}$ , the *logarithmic Sobolev inequality* becomes

$$\text{Ent}_\mu(f) = \int f \log f d\mu \leq \frac{1}{2} \int \frac{\|\nabla f\|_2^2}{f} d\mu \quad (5)$$

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

Recall that *the Gaussian Poincaré inequality*

$$\text{Var}(f(Z)) \leq \mathbb{E} \left[ \|\nabla f(Z)\|_2^2 \right]$$

Since

$$(1+t) \log(1+t) = t + \frac{t^2}{2} + o(t^2)$$

as  $t \rightarrow 0$ , we can get for Gaussian measures,

$$\text{Ent}_\mu(1 + \epsilon h) = \frac{\epsilon^2}{2} \text{Var}_\mu(h) + o(\epsilon^2).$$

Similarly,

$$\int \frac{\|\nabla(1 + \epsilon h)\|_2^2}{1 + \epsilon h} d\mu = \epsilon^2 \int \|\nabla h\|_2^2 d\mu + o(\epsilon^2).$$

Thus from *the Gaussian logarithmic Sobolev inequality*,

$$\begin{aligned} \text{Ent}_\mu(1 + \epsilon h) &\leq \frac{1}{2} \int \frac{\|\nabla(1 + \epsilon h)\|_2^2}{1 + \epsilon h} d\mu \\ \Leftrightarrow \frac{\epsilon^2}{2} \text{Var}_\mu(h) + o(\epsilon^2) &\leq \frac{\epsilon^2}{2} \int \|\nabla h\|_2^2 d\mu + o(\epsilon^2) \\ \Leftrightarrow \text{Var}(f(Z)) &\leq \mathbb{E} \left[ \|\nabla f(Z)\|_2^2 \right] \quad \text{as } \epsilon \rightarrow 0. \end{aligned}$$

Thus *the Gaussian logarithmic Sobolev inequality* implies *the Gaussian Poincaré inequality*.

### 1.3 Information Theory Interpretation

- **Remark** (*Information Interpretation of Gaussian Logarithmic Sobolev Inequality*)

Let  $\nu, \mu$  be two probability measures on  $(\mathcal{X}, \mathcal{F})$  and  $\nu \ll \mu$ . Define  $f := \frac{d\nu}{d\mu}$  be the Radon-Nikodym derivative of  $\nu$  with respect to  $\mu$  (i.e  $f$  is the probability density function of  $\nu$  with respect to  $\mu$ ). Then the entropy becomes *the relative entropy*

$$\text{Ent}_\mu(f) := \mathbb{E}_\mu[f \log f] = \text{KL}(\nu \parallel \mu)$$

since  $\mathbb{E}_\mu[f] = \int_{\mathcal{X}^n} f d\mu = 1$ .

On the other hand, *the (relative) Fisher information* is defined as

$$\begin{aligned} I(\nu \parallel \mu) &:= \mathbb{E}_\nu \left[ \|\nabla \log f\|_2^2 \right] \\ &= \int \left\| \frac{\nabla f}{f} \right\|_2^2 d\nu = \int \frac{\|\nabla f\|_2^2}{f^2} d\nu \\ &= \int \frac{\|\nabla f\|_2^2}{f} d\mu \end{aligned}$$

Thus *the information interpretation* of the Gaussian logarithmic Sobolev inequality is

$$\text{KL}(\nu \parallel \mu) \leq \frac{1}{2} I(\nu \parallel \mu) \tag{6}$$

Note that *the Fisher information metric* is *the Riemannian metric* induced by the relative entropy.

## 1.4 Logarithmic Sobolev Inequality for General Probability Measures

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

**Remark** (*The Sobolev Inequality*) [Evans, 2010]

*The Sobolev inequality* states for smooth function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  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 1.4** (*Euclidean Logarithmic Sobolev Inequality*).

Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a smooth function and  $m$  be Lebesgue measure on  $\mathbb{R}^n$ , then

$$\begin{aligned} \text{Ent}_m(f^2) &\leq \frac{n}{2} \log \left( \frac{2}{n\pi e} \mathbb{E}_m [\|\nabla f\|_2^2] \right) \\ \Leftrightarrow \int f^2 \log \left( \frac{f^2}{\int f^2 dx} \right) dx &\leq \frac{n}{2} \log \left( \frac{2}{n\pi e} \int |\nabla f|^2 dx \right) \end{aligned} \quad (7)$$

- **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  $f$

$$\text{Ent}_\mu(f^2) \leq C \mathbb{E}_\mu [\|\nabla f\|_2^2] \quad (8)$$

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

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

The right-hand side is defined as

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

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

$$\int f^2 \log \left( \frac{f^2}{\int f^2 d\mu} \right) d\mu \leq C \int \|\nabla f\|_2^2 d\mu \quad (9)$$

- **Remark** (*Logarithmic Sobolev Inequality*)

For non-negative function  $f$ , we can replace  $f \rightarrow \sqrt{f}$ , so that *the logarithmic Sobolev inequality* becomes

$$\text{Ent}_\mu(f) \leq \frac{1}{2} \int \frac{\|\nabla f\|_2^2}{f} d\mu \quad (10)$$

- **Remark (Modified Logarithmic Sobolev Inequality via Convex Cost and Duality)**  
For some **convex non-negative cost**  $c : \mathbb{R}^n \rightarrow \mathbb{R}_+$ , 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**

$$\text{Ent}_\mu(f) \leq \int f^2 c^* \left( \frac{\nabla f}{f} \right) d\mu \quad (11)$$

## 1.5 Variants of Logarithmic Sobolev Inequalities

- **Proposition 1.5 (A Modified Logarithmic Sobolev Inequalities for Moment Generating Function)** [Boucheron et al., 2013]

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

$$\lambda \mathbb{E} [X e^{\lambda X}] - \mathbb{E} [e^{\lambda X}] \log \mathbb{E} [e^{\lambda X}] \leq \sum_{i=1}^n \mathbb{E} [e^{\lambda X} \phi(-\lambda(X - X_{(-i)}))] \quad (12)$$

- **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  $f : \mathcal{X}^n \rightarrow \mathbb{R}$  and the random variable  $X = f(Z_1, \dots, Z_n)$ . Also denote  $\tilde{X}^{(i)} = f(Z_1, \dots, Z_{i-1}, Z'_i, Z_{i+1}, \dots, Z_n)$ . Let  $\phi(x) = e^x - x - 1$ . Then for all  $\lambda \in \mathbb{R}$ ,

$$\lambda \mathbb{E} [X e^{\lambda X}] - \mathbb{E} [e^{\lambda X}] \log \mathbb{E} [e^{\lambda X}] \leq \sum_{i=1}^n \mathbb{E} [e^{\lambda X} \phi(-\lambda(X - \tilde{X}^{(i)}))] \quad (13)$$

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

$$\begin{aligned} \lambda \mathbb{E} [X e^{\lambda X}] - \mathbb{E} [e^{\lambda X}] \log \mathbb{E} [e^{\lambda X}] &\leq \sum_{i=1}^n \mathbb{E} [e^{\lambda X} \tau(-\lambda(X - \tilde{X}^{(i)})_+)] , \\ \lambda \mathbb{E} [X e^{\lambda X}] - \mathbb{E} [e^{\lambda X}] \log \mathbb{E} [e^{\lambda X}] &\leq \sum_{i=1}^n \mathbb{E} [e^{\lambda X} \tau(\lambda(\tilde{X}^{(i)} - X)_-)] . \end{aligned}$$

## 1.6 Poisson Logarithmic Sobolev Inequality

- **Proposition 1.7 (Modified Logarithmic Sobolev Inequality for Bernoulli Random Variable).** [Boucheron et al., 2013]

Let  $f : \{0, 1\} \rightarrow (0, \infty)$  be a **non-negative** real-valued function defined on the binary set  $\{0, 1\}$ . Define **the discrete derivative** of  $f$  at  $x \in \{0, 1\}$  by

$$\nabla f := f(1 - x) - f(x).$$



Let  $X$  be a Bernoulli random variable with parameter  $p \in (0, 1)$  (i.e.  $\mathbb{P}\{X = 1\} = p$ ). Then

$$\text{Ent}(f(X)) \leq (p(1-p))\mathbb{E}[\nabla f(X)\nabla \log f(X)]. \quad (14)$$

and

$$\text{Ent}(f(X)) \leq (p(1-p))\mathbb{E}\left[\frac{|\nabla f(X)|^2}{f(X)}\right]. \quad (15)$$

- **Proposition 1.8 (Poisson Logarithmic Sobolev Inequality).** [Boucheron et al., 2013]  
Let  $f : \mathbb{N} \rightarrow (0, \infty)$  be a **non-negative** real-valued function defined on the set of non-negative integers  $\mathbb{N}$ . Define **the discrete derivative** of  $f$  at  $x \in \mathbb{N}$  by

$$\nabla f := f(x+1) - f(x).$$

Let  $X$  be a Poisson random variable. Then

$$\text{Ent}(f(X)) \leq (\mathbb{E}[X])\mathbb{E}[\nabla f(X)\nabla \log f(X)]. \quad (16)$$

and

$$\text{Ent}(f(X)) \leq (\mathbb{E}[X])\mathbb{E}\left[\frac{|\nabla f(X)|^2}{f(X)}\right]. \quad (17)$$

## 1.7 Applications

## 2 The Entropy Methods

### 2.1 Tensorization Property of $\Phi$ -Entropy

- **Remark** Recall that the  $\Phi$ -entropy for  $\Phi(x) = x \log(x)$  as

$$H_\Phi(X) = \text{Ent}(X) := \mathbb{E}[X \log X] - \mathbb{E}[X] \log(\mathbb{E}[X]).$$

The variational formulation of  $H_\Phi(X)$  is

$$\text{Ent}(X) = \sup_T \{X(\log(T) - \log(\mathbb{E}[T]))\}$$

•

### 2.2 Herbst's Argument

### 2.3 Bounded Difference Inequality

### 2.4 Modified Logarithmic Sobolev Inequalities

### 2.5 Concentration of Convex Lipschitz Functions

### 2.6 Exponential Tail Bounds for Self-Bounding Functions

## 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.