Functional inequalities
and concentration of measure
Part III

Radosław Adamczak

HIM Winter School
”The Interplay between High-Dimensional Geometry and Probability”
January 2021
How can one generalize *Poin* and *LSI* to more abstract settings?
Beyond $\mathbb{R}^n$

How can one generalize *Poin* and *LSI* to more abstract settings?

- On Riemannian manifolds it is more complicated but works in a similar way as in $\mathbb{R}^n$ (one uses the Riemannian volume, and the inner product on the tangent space)
Beyond $\mathbb{R}^n$

How can one generalize *Poin* and *LSI* to more abstract settings?

- On Riemannian manifolds it is more complicated but works in a similar way as in $\mathbb{R}^n$ (one uses the Riemannian volume, and the inner product on the tangent space)

- On general ‘continuous’ metric spaces $(\mathcal{X}, d)$ one can work with the length of the gradient:

\[ |\nabla f|(x) := \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x, y)}. \]
Beyond $\mathbb{R}^n$

How can one generalize Poin and LSI to more abstract settings?

- On Riemannian manifolds it is more complicated but works in a similar way as in $\mathbb{R}^n$ (one uses the Riemannian volume, and the inner product on the tangent space)
- On general ‘continuous’ metric spaces $(\mathcal{X}, d)$ one can work with the length of the gradient:

$$|\nabla f|(x) := \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x, y)}.$$

- In ‘discrete’ situations one can also use some ad hoc lengths of gradients.
Beyond $\mathbb{R}^n$

How can one generalize Poin and LSI to more abstract settings?

- On Riemannian manifolds it is more complicated but works in a similar way as in $\mathbb{R}^n$ (one uses the Riemannian volume, and the inner product on the tangent space).
- On general ‘continuous’ metric spaces $(\mathcal{X}, d)$ one can work with the length of the gradient:

$$|
abla f|(x) := \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x, y)}.$$

- In ‘discrete’ situations one can also use some ad hoc lengths of gradients.
- Finally, one can use the language of Markov processes and Dirichlet forms.
Markov processes, generators, Dirichlet forms

**Warning:** We will disregard the questions of domains, regularity, etc., if you prefer, think of Markov processes on finite state spaces

- \((\mathcal{X}, \mathcal{F}, \mu)\) – a probability space
Markov processes, generators, Dirichlet forms

**Warning:** We will disregard the questions of domains, regularity, etc., if you prefer, think of Markov processes on finite state spaces

- \((\mathcal{X}, \mathcal{F}, \mu)\) – a probability space
- \((X_t)_{t \geq 0}\) – a reversible Markov process with transition function \(P\) and invariant measure \(\mu\)
Markov processes, generators, Dirichlet forms

**Warning:** We will disregard the questions of domains, regularity, etc., if you prefer, think of Markov processes on finite state spaces

- $(\mathcal{X}, \mathcal{F}, \mu)$ – a probability space
- $(X_t)_{t \geq 0}$ – a reversible Markov process with transition function $P$ and invariant measure $\mu$
- the corresponding semigroup of operators

$$P_t f(x) = \mathbb{E}(f(X_t)|X_0 = x) = \mathbb{E}_x f(X_t) = \int_{\mathcal{X}} f(y) P_t(x, dy).$$
Markov processes, generators, Dirichlet forms

**Warning:** We will disregard the questions of domains, regularity, etc., if you prefer, think of Markov processes on finite state spaces

- \((\mathcal{X}, \mathcal{F}, \mu)\) – a probability space
- \((X_t)_{t \geq 0}\) – a reversible Markov process with transition function \(P\) and invariant measure \(\mu\)
- the corresponding semigroup of operators

\[
P_t f(x) = \mathbb{E}(f(X_t)|X_0 = x) = \mathbb{E}_x f(X_t) = \int_{\mathcal{X}} f(y) P_t(x, dy).
\]

- \(\mathcal{L}\) – the generator of the semigroup on \(L_2(\mu)\):

\[
\mathcal{L} f = \lim_{t \to 0} \frac{P_t f - f}{t}.
\]
Markov processes, generators, Dirichlet forms

**Warning:** We will disregard the questions of domains, regularity, etc., if you prefer, think of Markov processes on finite state spaces

- $(\mathcal{X}, \mathcal{F}, \mu)$ – a probability space
- $(X_t)_{t \geq 0}$ – a reversible Markov process with transition function $P$ and invariant measure $\mu$
- the corresponding semigroup of operators

$$P_t f(x) = \mathbb{E}(f(X_t)|X_0 = x) = \mathbb{E}_x f(X_t) = \int_{\mathcal{X}} f(y) P_t(x, dy).$$

- $\mathcal{L}$ – the generator of the semigroup on $L_2(\mu)$:

$$\mathcal{L} f = \lim_{t \to 0} \frac{P_t f - f}{t}.$$  

- $\mathcal{E}(f, g) = -\mathbb{E}_\mu(f \mathcal{L} g)$ – the Dirichlet form
Markov processes, generators, Dirichlet forms

**Warning:** We will disregard the questions of domains, regularity, etc., if you prefer, think of Markov processes on finite state spaces

- $(\mathcal{X}, \mathcal{F}, \mu)$ – a probability space
- $(X_t)_{t \geq 0}$ – a reversible Markov process with transition function $P$ and invariant measure $\mu$
- the corresponding semigroup of operators

\[ P_t f(x) = \mathbb{E}(f(X_t)|X_0 = x) = \mathbb{E}_x f(X_t) = \int_{\mathcal{X}} f(y) P_t(x,dy). \]

- $\mathcal{L}$ – the generator of the semigroup on $L^2(\mu)$:

\[ \mathcal{L} f = \lim_{t \to 0} \frac{P_t f - f}{t}. \]

- $\mathcal{E}(f,g) = -\mathbb{E}_\mu(f \mathcal{L} g)$ – the Dirichlet form
- $\Gamma(f,g) = \frac{1}{2}(\mathcal{L}(fg) - f \mathcal{L} g - g \mathcal{L} f)$ – carré du champ operator
Markov processes, generators, Dirichlet forms

**Warning:** We will disregard the questions of domains, regularity, etc., if you prefer, think of Markov processes on finite state spaces

- \((X, \mathcal{F}, \mu)\) – a probability space
- \((X_t)_{t \geq 0}\) – a reversible Markov process with transition function \(P\) and invariant measure \(\mu\)
- the corresponding semigroup of operators

\[ P_t f(x) = \mathbb{E}(f(X_t)|X_0 = x) = \mathbb{E}_x f(X_t) = \int_X f(y) P_t(x, dy). \]

- \(\mathcal{L}\) – the generator of the semigroup on \(L_2(\mu)\):

\[ \mathcal{L} f = \lim_{t \to 0} \frac{P_t f - f}{t}. \]

- \(\mathcal{E}(f, g) = -\mathbb{E}_\mu (f \mathcal{L} g)\) – the Dirichlet form
- \(\Gamma(f, g) = \frac{1}{2} (\mathcal{L}(fg) - f \mathcal{L} g - g \mathcal{L} f)\) – carré du champ operator
- \(\mathcal{E}(f, g) = \mathbb{E}_\mu \Gamma(f, g)\)
Example on $\mathbb{R}^n$

- $\mu(dx) = \frac{1}{Z} e^{-V(x)} dx$ for some $V: \mathbb{R}^n \to \mathbb{R}$. Define

$$\mathcal{L} f(x) = \Delta f(x) - \langle \nabla V(x), \nabla f(x) \rangle,$$

Then $\mathcal{L}$ is a generator of a diffusion with invariant measure $\mu$. 
Example on $\mathbb{R}^n$

- $\mu(dx) = \frac{1}{Z}e^{-V(x)}dx$ for some $V: \mathbb{R}^n \to \mathbb{R}$. Define

$$\mathcal{L}f(x) = \Delta f(x) - \langle \nabla V(x), \nabla f(x) \rangle,$$

Then $\mathcal{L}$ is a generator of a diffusion with invariant measure $\mu$.

In this case

$$\Gamma(f, g) = \langle \nabla f, \nabla g \rangle, \ E(f, g) = \mathbb{E}_\mu \langle \nabla f, \nabla g \rangle,$$

in particular

$$\Gamma(f, f) = |\nabla f|^2, \ E(f, f) = \mathbb{E}_\mu |\nabla f|^2.$$
Example on $\mathbb{R}^n$

- $\mu(dx) = \frac{1}{Z}e^{-V(x)}dx$ for some $V: \mathbb{R}^n \to \mathbb{R}$. Define

$$\mathcal{L}f(x) = \Delta f(x) - \langle \nabla V(x), \nabla f(x) \rangle,$$

Then $\mathcal{L}$ is a generator of a diffusion with invariant measure $\mu$.

In this case

$$\Gamma(f, g) = \langle \nabla f, \nabla g \rangle, \quad \mathcal{E}(f, g) = \mathbb{E}_\mu \langle \nabla f, \nabla g \rangle,$$

in particular

$$\Gamma(f, f) = |\nabla f|^2, \quad \mathcal{E}(f, f) = \mathbb{E}_\mu |\nabla f|^2.$$

For $V = \frac{1}{2}|x|^2$ we get the **Ornstein-Uhlenbeck** semigroup with

$$\mathcal{L}f = \Delta f - \langle x, \nabla f \rangle$$

in $L_2(\gamma_n)$. 
Further examples

- \( Q(\cdot)(\cdot) : \mathcal{X} \times \mathcal{F} \to \mathbb{R}_+ \) – a kernel on \( \mathcal{X} \) satisfying the detailed balance condition

\[
Q_x(dy)\mu(dx) = Q_y(dx)\mu(dy)
\]

One can define

\[
\Gamma(f,g)(x) = \frac{1}{2} \int X (f(y) - f(x))(g(y) - g(x))Q_x(dy)
\]

\[
E(f,g) = \frac{1}{2} \int X \int X (f(y) - f(x))(g(y) - g(x))Q_x(dy)\mu(dx)
\]

Thanks to reversibility,

\[
E(f,f) = E\mu \Gamma(f)
\]
Further examples

- \( Q(\cdot)(\cdot): X \times \mathcal{F} \to \mathbb{R}_+ \) – a kernel on \( X \) satisfying the detailed balance condition

\[
Q_x(dy)\mu(dx) = Q_y(dx)\mu(dy)
\]

One can define

\[
\mathcal{L}f(x) = \int_X (f(y) - f(x))Q_x(dy)
\]

Then

\[
\Gamma(f, g)(x) = \frac{1}{2} \int_X (f(y) - f(x))(g(y) - g(x))Q_x(dy),
\]

\[
\mathcal{E}(f, g) = \frac{1}{2} \int_X \int_X (f(y) - f(x))(g(y) - g(x))Q_x(dy)\mu(dx).
\]

One can also define

\[
\Gamma_+(f) = \int_X (f(x) - f(y))^2_+Q_x(dy).
\]

Thanks to reversibility,

\[
\mathcal{E}(f, f) = \mathbb{E}_\mu \Gamma_+(f).
\]
Jump processes on countable spaces

- $\mathcal{X}$ – countable
- $\mathcal{L} = (\lambda_{x,y})_{x,y \in \mathcal{X}}$, such that

$$\forall x \neq y \lambda_{xy} \geq 0, \text{ and } \sum_{y \in \mathcal{X}} \lambda_{xy} = 0.$$ 

- Thus $\lambda_{xy}$ is the intensity of jumps from $x$ to $y$
- $X_t$ stays at $x$ for exponential time with mean $1/\lambda_{xx}$ then jumps to $y$ with probability $(-\lambda_{xy}/\lambda_{xx})$. 
Jump processes on countable spaces

- $\mathcal{X}$ – countable
- $\mathcal{L} = (\lambda_{x,y})_{x,y \in \mathcal{X}}$, such that

$$\forall x \neq y \lambda_{xy} \geq 0, \text{ and } \sum_{y \in \mathcal{X}} \lambda_{xy} = 0.$$ 

Thus $\lambda_{xy}$ is the intensity of jumps from $x$ to $y$

- $X_t$ stays at $x$ for exponential time with mean $1/\lambda_{xx}$ then jumps to $y$ with probability $(-\lambda_{xy}/\lambda_{xx})$.

**Example:** Birth and death chain $\mathcal{X} = \mathbb{N}$, $\lambda_{i,i+1} = b_i$, $\lambda_{i,i-1} = d_i$, $\lambda_{ii} = -(b_i + d_i)$.

For $d_i = i/\lambda$, $b_i = 1$ ($\mathcal{M}/\mathcal{M}/\infty$-queue) we get $\mu = \text{Poisson}(\lambda)$ as the stationary measure and

$$\mathcal{L} f(i) = (f(i + 1) - f(i)) + \frac{i}{\lambda} \mathbf{1}_{\{i > 0\}}(f(i - 1) - f(i))$$

$$\mathcal{E}(f, f) = \sum_{i=0}^{\infty} (f(i + 1) - f(i))^2 \mu(i)$$
Glauber dynamics (a.k.a. Gibbs sampler)

- $\mathcal{X} = E^I$, where $I$ is a finite set
Glauber dynamics (a.k.a. Gibbs sampler)

- $\mathcal{X} = E^I$, where $I$ is a finite set
- $\mathcal{X}$-valued $X = (X_i)_{i \in I}$ with law $\mu$

In words, after an exponential time we pick up a coordinate at random and replace it by its conditionally independent copy given the value of the other coordinates.
Glauber dynamics (a.k.a. Gibbs sampler)

- $\mathcal{X} = E^I$, where $I$ is a finite set
- $\mathcal{X}$-valued $X = (X_i)_{i \in I}$ with law $\mu$
- For $i \in I$,

$$\mu_i(\cdot | x) := \mathbb{P}(X_i \in \cdot \mid X\{i\}_c = x\{i\}_c)$$
Glauber dynamics (a.k.a. Gibbs sampler)

- $\mathcal{X} = E^I$, where $I$ is a finite set
- $\mathcal{X}$-valued $X = (X_i)_{i \in I}$ with law $\mu$
- For $i \in I$,

$$\mu_i(\cdot|x) := \mathbb{P}(X_i \in \cdot | X\{i\}_c = x\{i\}_c)$$

- The Glauber dynamics is given by generator

$$\mathcal{L} f(x) = \sum_{i \in I} \int_E (f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) - f(x)) \mu_i(dy|x),$$

- In words, after an exponential time we pick up a coordinate at random and replace it by its conditionally independent copy given the value of the other coordinates.
Poincaré and entropic inequalities for Markov chains

Given a Dirichlet form $\mathcal{E}$ and the corresponding invariant measure we will say that

- the Poincaré inequality with constant $C$ holds if for all admissible $f : \mathcal{X} \to \mathbb{R}$,

$$\text{Var}_\mu(f) \leq C \mathcal{E}(f, f)$$

- the log-Sobolev inequality with constant $C$ holds if for all admissible $f : \mathcal{X} \to \mathbb{R}$,

$$\text{Ent}_\mu(f^2) \leq 2C \mathcal{E}(f, f)$$

- the entropic inequality with constant $C$ holds if for all admissible $f : \mathcal{X} \to (0, \infty)$,

$$\text{Ent}_\mu(f) \leq \frac{C}{2} \mathcal{E}(f, \log f)$$
Problems with the chain rule

For $\mathcal{E}(f, f) = \mathbb{E}_\mu |\nabla f|^2$ the entropic and log-Sobolev inequality are equivalent. Indeed if $LSI(C)$ holds then

$$\text{Ent}_\mu f = \text{Ent}_\mu (\sqrt{f})^2 \leq 2C\mathbb{E}_\mu |\nabla \sqrt{f}|^2 = \frac{C}{2} \mathbb{E}_\mu \frac{|\nabla f|^2}{f} = \frac{C}{2} \mathbb{E}_\mu \langle \nabla f, \nabla \log f \rangle.$$ 

and if $EI(C)$ holds, then

$$\text{Ent}_\mu f^2 \leq \frac{C}{2} \mathbb{E}_\mu \langle \nabla f^2, \nabla \log f^2 \rangle = 2C\mathbb{E}_\mu |\nabla f|^2.$$
Problems with the chain rule

For $\mathcal{E}(f, f) = \mathbb{E}_\mu |\nabla f|^2$ the entropic and log-Sobolev inequality are equivalent. Indeed if $LSI(C)$ holds then

$$\text{Ent}_\mu f = \text{Ent}_\mu (\sqrt{f})^2 \leq 2C \mathbb{E}_\mu |\nabla \sqrt{f}|^2$$

$$= \frac{C}{2} \mathbb{E}_\mu \frac{|\nabla f|^2}{f} = \frac{C}{2} \mathbb{E}_\mu \langle \nabla f, \nabla \log f \rangle.$$

and if $EI(C)$ holds, then

$$\text{Ent}_\mu f^2 \leq \frac{C}{2} \mathbb{E}_\mu \langle \nabla f^2, \nabla \log f^2 \rangle = 2C \mathbb{E}_\mu |\nabla f|^2.$$

This is not the case for general Dirichlet forms. For instance the Poisson variable $X$ on $\mathbb{N}$ satisfies $EI$ for

$$\mathcal{E}(f) = \mathbb{E}(f(X + 1) - f(X))^2$$

(Wu, Dai Pra–Paganoni–Posta) but does not satisfy $LSI$. Still we have

$$LSI(C) \Rightarrow EI(C) \Rightarrow Poinc(C).$$
More on Dirichlet forms

\[
\mathcal{E}(f, g) = -\mathbb{E}_\mu f \mathcal{L} g = -\mathbb{E}_\mu f(X_0) \lim_{t \to 0} \frac{\mathbb{E}(g(X_t)|X_0) - g(X_0)}{t}
\]

\[
= \lim_{t \to 0} \frac{1}{t} \left( - \mathbb{E}_\mu f(X_0) g(X_t) + \mathbb{E}_\mu f(X_0) g(X_0) \right)
\]

\[
= \lim_{t \to 0} \frac{1}{2t} \mathbb{E}_\mu \left( f(X_0) g(X_0) + f(X_t) g(X_t) - f(X_0) g(X_t) - f(X_t) g(X_0) \right)
\]

\[
= \lim_{t \to 0} \frac{1}{2t} \mathbb{E}_\mu (f(X_t) - f(X_0))(g(X_t) - g(X_0))
\]

\[
= \lim_{t \to 0} \frac{1}{2t} \int_X \int_X (f(y) - f(x))(g(y) - g(x)) P_t(x, dy) \mu(dy).
\]
More on Dirichlet forms

\[ \mathcal{E}(f, g) = -\mathbb{E}_\mu f \mathcal{L} g = -\mathbb{E}_\mu f(X_0) \lim_{t \to 0} \frac{\mathbb{E}(g(X_t) | X_0) - g(X_0)}{t} \]

\[ = \lim_{t \to 0} \frac{1}{t} \left( -\mathbb{E}_\mu f(X_0) g(X_t) + \mathbb{E}_\mu f(X_0) g(X_0) \right) \]

\[ = \lim_{t \to 0} \frac{1}{2t} \mathbb{E}_\mu \left( f(X_0) g(X_0) + f(X_t) g(X_t) - f(X_0) g(X_t) - f(X_t) g(X_0) \right) \]

\[ = \lim_{t \to 0} \frac{1}{2t} \mathbb{E}_\mu (f(X_t) - f(X_0))(g(X_t) - g(X_0)) \]

\[ = \lim_{t \to 0} \frac{1}{2t} \int_{\mathcal{X}} \int_{\mathcal{X}} (f(y) - f(x))(g(y) - g(x)) P_t(x, dy) \mu(dy). \]

**Consequence** If \( f, g, h, k : \mathcal{X} \to \mathbb{R} \) satisfy a pointwise inequality

\[ (f(x) - f(y))(g(x) - g(y)) \leq (h(x) - h(y))(k(x) - k(y)), \]

then

\[ \mathcal{E}(f, g) \leq \mathcal{E}(h, k). \]
Proposition

$$LSI(C) \Rightarrow EI(C) \Rightarrow Poinc(C).$$
Proposition

\[ LSI(C) \Rightarrow EI(C) \Rightarrow Poinc(C). \]

**Sketch of proof:** Applying \( LSI(C) \) to \( \sqrt{f} \) we obtain

\[ \text{Ent}_\mu f \leq 2C \mathcal{E}(\sqrt{f}, \sqrt{f}). \]
Proposition

\[ LSI(C) \Rightarrow EI(C) \Rightarrow Poinc(C). \]

**Sketch of proof:** Applying \( LSI(C) \) to \( \sqrt{f} \) we obtain

\[ \text{Ent}_\mu f \leq 2C \mathcal{E}(\sqrt{f}, \sqrt{f}). \]

Now, the pointwise inequality

\[ (\sqrt{a} - \sqrt{b})^2 \leq \frac{1}{4} (a - b)(\log a - \log b) \]

gives

\[ \mathcal{E}(\sqrt{f}, \sqrt{f}) \leq \frac{1}{4} \mathcal{E}(f, \log f), \]

yielding \( EI(C) \).
Proposition

\[ LSI(C) \Rightarrow EI(C) \Rightarrow Poinc(C). \]

**Sketch of proof:** Applying \( LSI(C) \) to \( \sqrt{f} \) we obtain

\[ \text{Ent}_\mu f \leq 2C \mathcal{E}(\sqrt{f}, \sqrt{f}). \]

Now, the pointwise inequality

\[ (\sqrt{a} - \sqrt{b})^2 \leq \frac{1}{4} (a - b)(\log a - \log b) \]

gives

\[ \mathcal{E}(\sqrt{f}, \sqrt{f}) \leq \frac{1}{4} \mathcal{E}(f, \log f), \]

yielding \( EI(C) \).

The second implication follows by applying \( EI \) to \( (1 + \varepsilon f) \) and Taylor’s expansion for \( \varepsilon \to 0 \) as in the ‘smooth’ case.
Concentration

Let $X$ be distributed according to $\mu$.

**Proposition**

If $Poinc(C)$ holds then for all functions $f : \mathcal{X} \to \mathbb{R}$ with $\sqrt{\Gamma(f, f)} \leq L$ and all $t \geq 0$,

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geq t) \leq 2 \exp \left( - \frac{t}{2 \sqrt{CL}} \right).$$
Concentration

Let $X$ be distributed according to $\mu$.

**Proposition**

If $Poinc(C)$ holds then for all functions $f : \mathcal{X} \to \mathbb{R}$ with $\sqrt{\Gamma(f, f)} \leq L$ and all $t \geq 0$,

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geq t) \leq 2 \exp \left( - \frac{t}{2\sqrt{CL}} \right).$$

If $EI(C)$ holds then for $f$ as above

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geq t) \leq \exp \left( - \frac{t^2}{4CL^2} \right).$$
Sketch of proof for the EI case
We start with a (non-rigorous) observation:

\[ \Gamma(f, f)(x) = \frac{1}{2} \left( \mathcal{L}(f^2)(x) - 2g(x)\mathcal{L}f(x) \right) \]

\[ = \lim_{t \to 0} \frac{1}{2t} \int_{\mathcal{X}} \left( f^2(y) - f^2(x) - 2f(x)(f(y) - f(x)) \right) P_t(x, dy) \]

\[ = \lim_{t \to 0} \frac{1}{2t} \int_{\mathcal{X}} (f(y) - f(x))^2 P_t(x, dy). \]
Sketch of proof for the EI case
We start with a (non-rigorous) observation:

\[
\Gamma(f, f)(x) = \frac{1}{2}\left(\mathcal{L}(f^2)(x) - 2g(x)\mathcal{L}f(x)\right) \\
= \lim_{t \to 0} \frac{1}{2t} \int_X \left(f^2(y) - f^2(x) - 2f(x)(f(y) - f(x))\right)P_t(x, dy) \\
= \lim_{t \to 0} \frac{1}{2t} \int_X (f(y) - f(x))^2 P_t(x, dy).
\]

Applying EI to \(e^{\lambda f}\), \(\lambda > 0\), we get

\[
\lambda \mathbb{E}_\mu f e^{\lambda f} - \mathbb{E}_\mu e^{\lambda f} \log \mathbb{E}_\mu e^{\lambda f} \leq \frac{C}{2} \mathcal{E}(e^{\lambda f}, \lambda f) \\
= \frac{C\lambda}{2} \lim_{t \to 0} \int_X \int_X \frac{1}{t}(e^{\lambda f(x)} - e^{\lambda f(y)})(f(x) - f(y)) + P_t(x, dy)\mu(dx) \\
\leq \frac{C\lambda}{2} \int_X \lambda e^{\lambda f(x)} \lim_{t \to 0} \int_X \frac{1}{t}(f(x) - f(y))^2 P_t(x, dy)\mu(dx) \\
= C\lambda^2 \mathbb{E}_\mu e^{\lambda f} \Gamma(f, f) \leq C\lambda^2 L^2 \mathbb{E}_\mu e^{\lambda f}.
\]

This is the starting point for the Herbst argument.
Kernel case

Proposition

Assume $EI(C')$. If for all $x$,

$$\Gamma_+(f) = \int_X (f(x) - f(y))^2 Q_x(dy) \leq L^2,$$

then for all $t \geq 0$,

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geq t) \leq \exp \left( - \frac{t^2}{2CL^2} \right).$$
Inequalities for product measures

Let $Y$ and $Y'$ be i.i.d. random variables. Then

$$\text{Var } f(Y) = \frac{1}{2} \mathbb{E}(f(Y) - f(Y'))^2 = \mathbb{E}(f(Y) - f(Y'))^2_+.$$
Inequalities for product measures

Let $Y$ and $Y'$ be i.i.d. random variables. Then

$$\text{Var } f(Y) = \frac{1}{2} \mathbb{E}((f(Y) - f(Y'))^2 = \mathbb{E}(f(Y) - f(Y'))^2_+.$$ 

Together with the tensorization of the variance this gives

**Theorem (Efron-Stein inequality)**

Let $X = (X_1, \ldots, X_n)$, where $X_i$ are independent r.v.’s and for $i \leq n$ let

$$X^{(i)} = (X_1, \ldots, X_{i-1}, X'_i, X_{i+1}, \ldots, X_n)$$

where $(X'_1, \ldots, X'_n)$ is an independent copy of $X$. Then

$$\text{Var } f(X) \leq \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}(f(X) - f(X^{(i)}))^2.$$
Inequalities for product measures

Similarly,

\[
\text{Ent } f(Y) = \mathbb{E} f(Y) \log f(Y) - \mathbb{E} f(Y) \log \mathbb{E} f(Y) \\
\leq \mathbb{E} f(Y) \log f(Y) - \mathbb{E} f(Y) \mathbb{E} \log f(Y) \\
= \frac{1}{2} \mathbb{E} (f(Y) - f(Y'))(\log f(Y) - f(Y')).
\]
Inequalities for product measures

Similarly,

\[ \text{Ent } f(Y) = \mathbb{E}f(Y) \log f(Y) - \mathbb{E}f(Y) \log \mathbb{E}f(Y) \]
\[ \leq \mathbb{E}f(Y) \log f(Y) - \mathbb{E}f(Y) \mathbb{E} \log f(Y) \]
\[ = \frac{1}{2} \mathbb{E}(f(Y) - f(Y'))(\log f(Y) - f(Y')). \]

Theorem (Ledoux, Boucheron-Bousquet-Lugosi-Massart)

Let \( X = (X_1, \ldots, X_n) \), where \( X_i \) are independent r.v.’s and for \( i \leq n \) let

\[ X^{(i)} = (X_1, \ldots, X_{i-1}, X'_i, X_{i+1}, \ldots, X_n) \]

where \((X'_1, \ldots, X'_i)\) is an independent copy of \( X \). Then

\[ \text{Ent } f(X) \leq \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}(f(X) - f(X^{(i)}))(\log f(X) - \log f(X^{(i)})). \]
Inequalities for product measures

The above two inequalities can be rewritten as

**Theorem**

If $\mu$ is a product measure then the corresponding Glauber dynamics satisfies $Poinc(1)$ and $EI(2)$. 

What about $LSI$?

One can easily show that $LSI$ for the Glauber dynamics may hold only if the measure is 'finitely supported'. In the product case the constants depend on the smallest atom of one-dimensional marginals.
The above two inequalities can be rewritten as

**Theorem**

If \( \mu \) is a product measure then the corresponding Glauber dynamics satisfies \( Poinc(1) \) and \( EI(2) \).

**What about LSI?**

One can easily show that LSI for the Glauber dynamics may hold only if the measure is ‘finitely supported’. In the product case the constants depend on the smallest atom of one-dimensional marginals.
An application

Let $X = (X_1, \ldots, X_n)$ with $X_i$ - independent, $X_i \in [-1, 1]$. Let $f : [-1, 1]^n \rightarrow \mathbb{R}$ be an $L$-Lipschitz, differentiable, \textit{separately convex} function.
An application

Let $X = (X_1, \ldots, X_n)$ with $X_i$ - independent, $X_i \in [-1, 1]$. Let $f: [-1, 1]^n \to \mathbb{R}$ be an $L$-Lipschitz, differentiable, separately convex function. Then

$$
\Gamma_+(f) = \sum_{i=1}^{n} (f(X) - f(X^{(i)}))^2_{+} \leq \sum_{i=1}^{n} (\partial_i f(X))^2 (X_i - X'_i)^2
$$

$$
\leq 4|\nabla f(X)|^2 \leq 4L^2.
$$
An application

Let $X = (X_1, \ldots, X_n)$ with $X_i$ - independent, $X_i \in [-1, 1]$. Let $f: [-1, 1]^n \to \mathbb{R}$ be an $L$-Lipschitz, differentiable, separately convex function. Then

$$\Gamma_+(f) = \sum_{i=1}^{n} (f(X) - f(X^{(i)}))^2_+ \leq \sum_{i=1}^{n} (\partial_i f(X))^2 (X_i - X'_i)^2 \leq 4|\nabla f(X)|^2 \leq 4L^2.$$ 

Theorem (Ledoux)

For all $t \geq 0$,

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geq t) \leq \exp \left(- \frac{t^2}{8L^2}\right).$$
An application

Let $X = (X_1, \ldots, X_n)$ with $X_i$ - independent, $X_i \in [-1, 1]$. Let $f: [-1,1]^n \to \mathbb{R}$ be an $L$-Lipschitz, differentiable, separately convex function. Then

$$\Gamma_+(f) = \sum_{i=1}^{n} (f(X) - f(X^{(i)}))^2 \leq \sum_{i=1}^{n} (\partial_i f(X))^2 (X_i - X'_i)^2 \leq 4|\nabla f(X)|^2 \leq 4L^2.$$ 

Theorem (Ledoux)

For all $t \geq 0$,

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geq t) \leq \exp \left( - \frac{t^2}{8L^2} \right).$$

Remark: For convex functions Talagrand proved inequalities for both the upper and lower tail. Bounding the lower term with entropy methods is possible but seems quite challenging.
Non-product examples

In the non-product case one can prove functional inequalities using Dobrushin type conditions, e.g., we have

**Theorem (Marton, Götze–Sambale–Sinulis)**

Assume that $\mathcal{X} = E^I$, where $E$ is finite and $\mu$ has full support. Define $A_{ii} = 0$ and for $i \neq j$

$$A_{ij} = \sup_{x\{j\}^c = y\{j\}^c} \|\mu_i(\cdot| x) - \mu_i(\cdot| y)\|_{TV}$$

and $\alpha = 1 - |A|_{op}$. Define also

$$\beta = \inf_{i \in I} \min_{x \in \mathcal{X}} \mu_i(y_i | y_{i}^c)$$

Then the Glauber dynamics satisfies $EI(\alpha^{-2} \beta^{-1})$ and $LSI(2\alpha^{-2} \beta^{-1} \log(\beta^{-1}) / \log 2)$
Non-product examples

Ising model on finite graphs

- $\mathcal{X} = \{+1, -1\}^n$ (spins)
Non-product examples

Ising model on finite graphs

- $\mathcal{X} = \{+1, -1\}^n$ (spins)
- $J = (J_{ij})_{i,j=1}^n$ – a symmetric matrix $J_{ii} = 0$ (interactions)
Non-product examples

Ising model on finite graphs

- $\mathcal{X} = \{+1, -1\}^n$ (spins)
- $J = (J_{ij})_{i,j=1}^n$ - a symmetric matrix $J_{ii} = 0$ (interactions)
- $h \in \mathbb{R}^n$ (external field)
Non-product examples

Ising model on finite graphs

- $\mathcal{X} = \{+1, -1\}^n$ (spins)
- $J = (J_{ij})_{i,j=1}^n$ – a symmetric matrix $J_{ii} = 0$ (interactions)
- $h \in \mathbb{R}^n$ (external field)

\[
\mu(\{\varepsilon\}) = \frac{1}{Z} \exp \left( \frac{1}{2} \sum_{i,j=1}^n J_{ij} \varepsilon_i \varepsilon_j - \sum_{i=1}^n h_i \varepsilon_i \right),
\]
Non-product examples

Ising model on finite graphs

- $\mathcal{X} = \{+1, -1\}^n$ (spins)
- $J = (J_{ij})_{i,j=1}^n$ - a symmetric matrix $J_{ii} = 0$ (interactions)
- $h \in \mathbb{R}^n$ (external field)

$$
\mu(\{\varepsilon\}) = \frac{1}{Z} \exp\left(\frac{1}{2} \sum_{i,j=1}^n J_{ij} \varepsilon_i \varepsilon_j - \sum_{i=1}^n h_i \varepsilon_i\right),
$$

- Götze-Sambale-Sinulis:

$$
\alpha \geq 1 - \max_{i \leq n} \sum_{j \leq n} |J_{ij}|, \quad \beta \geq ce^{-\|h\|_{\infty}}.
$$
Non-product examples

Ising model on finite graphs

- $\mathcal{X} = \{+1, -1\}^n$ (spins)
- $J = (J_{ij})_{i,j=1}^n$ – a symmetric matrix $J_{ii} = 0$ (interactions)
- $h \in \mathbb{R}^n$ (external field)

$$
\mu(\{\varepsilon\}) = \frac{1}{Z} \exp \left( \frac{1}{2} \sum_{i,j=1}^n J_{ij}\varepsilon_i\varepsilon_j - \sum_{i=1}^n h_i\varepsilon_i \right),
$$

- Götze-Sambale-Sinulis:
  $$
  \alpha \geq 1 - \max_{i \leq n} \sum_{j \leq n} |J_{ij}|, \quad \beta \geq ce^{-\|h\|_{\infty}}.
  $$

Other examples: exponential random graphs, hardcore models.
Thank you