UNIVERSITY OF WARSAW

Functional inequalities and moment estimates

Radosław Adamczak

Bangalore Probability Seminar 2020

Gaussian concentration inequality

Theorem (Borell, Sudakov-Tsirelson, 1975)

If G is a standard Gaussian vector in $\mathbb{R}^n$ and $f\colon \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz, then for all $t > 0$,

$$\mathbb{P}(|f(G) - \mathbb{E}F(G)| \geqslant t) \leqslant 2\exp\Big(-\frac{t^2}{2L^2}\Big).$$

Gaussian concentration inequality

Theorem (Borell, Sudakov-Tsirelson, 1975)

If G is a standard Gaussian vector in $\mathbb{R}^n$ and $f\colon \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz, then for all $t > 0$,

$$\mathbb{P}(|f(G) - \mathbb{E}F(G)| \geqslant t) \leqslant 2\exp\Big(-\frac{t^2}{2L^2}\Big).$$

Remarks:

- Original formulation in terms of median, different constants.

Gaussian concentration inequality

Theorem (Borell, Sudakov-Tsirelson, 1975)

If G is a standard Gaussian vector in $\mathbb{R}^n$ and $f\colon \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz, then for all $t > 0$,

$$\mathbb{P}(|f(G) - \mathbb{E}F(G)| \geqslant t) \leqslant 2\exp\Big(-\frac{t^2}{2L^2}\Big).$$

Remarks:

- Original formulation in terms of median, different constants.
- Very useful in analysis of Gaussian processes, asymptotic convex geometry, etc.

Gaussian concentration inequality

Theorem (Borell, Sudakov-Tsirelson, 1975)

If G is a standard Gaussian vector in $\mathbb{R}^n$ and $f\colon \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz, then for all $t > 0$,

$$\mathbb{P}(|f(G) - \mathbb{E}F(G)| \geqslant t) \leqslant 2\exp\Big(-\frac{t^2}{2L^2}\Big).$$

Remarks:

- Original formulation in terms of median, different constants.
- Very useful in analysis of Gaussian processes, asymptotic convex geometry, etc.
- Linear functions show optimality.

Gaussian concentration from log-Sobolev inequalities

Definition

A probability measure μ on $\mathbb{R}^n$ satisfies the log-Sobolev inequality with constant K if for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$,

$$\text{(LSI)} \qquad \operatorname{Ent}_\mu f^2 \leqslant 2K\mathbb{E}_\mu |\nabla f|^2,$$

where for $g \geqslant 0$, $\operatorname{Ent}_\mu g = \mathbb{E}_\mu g \log(g) - \mathbb{E}_\mu g \log(\mathbb{E}_\mu g)$.

Gaussian concentration from log-Sobolev inequalities

Definition

A probability measure μ on $\mathbb{R}^n$ satisfies the log-Sobolev inequality with constant K if for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$,

$$\text{(LSI)} \qquad \operatorname{Ent}_\mu f^2 \leqslant 2K\mathbb{E}_\mu|\nabla f|^2,$$

where for $g \geqslant 0$, $\operatorname{Ent}_\mu g = \mathbb{E}_\mu g\log(g) - \mathbb{E}_\mu g\log(\mathbb{E}_\mu g)$.

- Equivalently for $X \sim \mu$

$$\operatorname{Ent} f^2(X) \leqslant 2K\mathbb{E}|\nabla f(X)|^2.$$

Gaussian concentration from log-Sobolev inequalities

Definition

A probability measure μ on $\mathbb{R}^n$ satisfies the log-Sobolev inequality with constant K if for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$,

$$\text{(LSI)} \qquad \operatorname{Ent}_\mu f^2 \leqslant 2K\mathbb{E}_\mu|\nabla f|^2,$$

where for $g \geqslant 0$, $\operatorname{Ent}_\mu g = \mathbb{E}_\mu g\log(g) - \mathbb{E}_\mu g\log(\mathbb{E}_\mu g)$.

- Equivalently for $X \sim \mu$

$$\operatorname{Ent} f^2(X) \leqslant 2K\mathbb{E}|\nabla f(X)|^2.$$

- By Jensen's ineq. $\operatorname{Ent}_\mu g \geqslant 0$ for all nonnegative functions g.

Gaussian concentration from log-Sobolev inequalities

Definition

A probability measure μ on $\mathbb{R}^n$ satisfies the log-Sobolev inequality with constant K if for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$,

$$\text{(LSI)} \qquad \operatorname{Ent}_\mu f^2 \leqslant 2K\mathbb{E}_\mu |\nabla f|^2,$$

where for $g \geqslant 0$, $\operatorname{Ent}_\mu g = \mathbb{E}_\mu g \log(g) - \mathbb{E}_\mu g \log(\mathbb{E}_\mu g)$.

- Equivalently for $X \sim \mu$
$$\operatorname{Ent} f^2(X) \leqslant 2K\mathbb{E}|\nabla f(X)|^2.$$

- By Jensen's ineq. $\operatorname{Ent}_\mu g \geqslant 0$ for all nonnegative functions g.
- (tensorization) If μ, ν satisfy LSI(K), then so does $\mu \otimes \nu$.

Theorem (Gross 1975)

If G is the standard Gaussian vector in $\mathbb{R}^n$, then G satisfies the log-Sobolev inequality with $K = 1$.

Theorem (Gross 1975)

If G is the standard Gaussian vector in $\mathbb{R}^n$, then G satisfies the log-Sobolev inequality with $K = 1$.

- The original proof
    - Establishes a discrete inequality on $\{-1, 1\}$
    - Tensorizes it to $\{-1, 1\}^n$
    - Uses CLT to pass to the Gaussian measure.

Theorem (Gross 1975)

If G is the standard Gaussian vector in $\mathbb{R}^n$, then G satisfies the log-Sobolev inequality with $K = 1$.

- The original proof
  - Establishes a discrete inequality on $\{-1, 1\}$
  - Tensorizes it to $\{-1, 1\}^n$
  - Uses CLT to pass to the Gaussian measure.
- Now many different proofs: analytic, semigroup methods, stochastic calculus...

Theorem (Gross 1975)

If G is the standard Gaussian vector in $\mathbb{R}^n$, then G satisfies the log-Sobolev inequality with $K = 1$.

- The original proof
  - Establishes a discrete inequality on $\{-1, 1\}$
  - Tensorizes it to $\{-1, 1\}^n$
  - Uses CLT to pass to the Gaussian measure.
- Now many different proofs: analytic, semigroup methods, stochastic calculus...
- Bakry-Émery: if $\mu(dx) = e^{-V} dx$ with $\nabla^2 V \geqslant K^{-1}\mathrm{Id}$, $K > 0$, then μ satisfies LSI(K)

Herbst's argument: from LSI to concentration

Theorem

If a random vector X satisfies the LSI(K) then for all L-Lipschitz $f\colon \mathbb{R}^n \to \mathbb{R}$, and $t \geqslant 0$,

$$\mathbb{P}(f(X) - \mathbb{E}f(X) \geqslant t) \leqslant \exp\Big(-\frac{t^2}{2KL^2}\Big).$$

Proof: Applying LSI to $e^{\lambda f/2}$

$$\lambda \mathbb{E} f(X) e^{\lambda f(X)} - \mathbb{E} e^{\lambda f(X)} \log \mathbb{E} e^{\lambda f(X)} \leqslant \frac{1}{2} K \lambda^2 \mathbb{E} |\nabla f(X)|^2 e^{\lambda f(X)}$$

Proof: Applying LSI to $e^{\lambda f/2}$

$$\lambda \mathbb{E} f(X) e^{\lambda f(X)} - \mathbb{E} e^{\lambda f(X)} \log \mathbb{E} e^{\lambda f(X)} \leqslant \frac{1}{2} K \lambda^2 \mathbb{E} |\nabla f(X)|^2 e^{\lambda f(X)}$$

$$\frac{\lambda \mathbb{E} f(X) e^{\lambda f(X)} - \mathbb{E} e^{\lambda f(X)} \log \mathbb{E} e^{\lambda f(X)}}{\lambda^2 \mathbb{E} e^{\lambda f(X)}} \leqslant \frac{1}{2} K L^2$$

Proof: Applying LSI to $e^{\lambda f/2}$

$$\lambda \mathbb{E} f(X) e^{\lambda f(X)} - \mathbb{E} e^{\lambda f(X)} \log \mathbb{E} e^{\lambda f(X)} \leqslant \frac{1}{2} K \lambda^2 \mathbb{E} |\nabla f(X)|^2 e^{\lambda f(X)}$$

$$\frac{\lambda \mathbb{E} f(X) e^{\lambda f(X)} - \mathbb{E} e^{\lambda f(X)} \log \mathbb{E} e^{\lambda f(X)}}{\lambda^2 \mathbb{E} e^{\lambda f(X)}} \leqslant \frac{1}{2} K L^2$$

$$\frac{d}{d\lambda} \frac{\log \mathbb{E} e^{\lambda f(X)}}{\lambda} \leqslant \frac{1}{2} K L^2$$

Proof: Applying LSI to $e^{\lambda f/2}$

$$\lambda \mathbb{E} f(X) e^{\lambda f(X)} - \mathbb{E} e^{\lambda f(X)} \log \mathbb{E} e^{\lambda f(X)} \leqslant \frac{1}{2} K \lambda^2 \mathbb{E} |\nabla f(X)|^2 e^{\lambda f(X)}$$

$$\frac{\lambda \mathbb{E} f(X) e^{\lambda f(X)} - \mathbb{E} e^{\lambda f(X)} \log \mathbb{E} e^{\lambda f(X)}}{\lambda^2 \mathbb{E} e^{\lambda f(X)}} \leqslant \frac{1}{2} K L^2$$

$$\frac{d}{d\lambda} \frac{\log \mathbb{E} e^{\lambda f(X)}}{\lambda} \leqslant \frac{1}{2} K L^2$$

Integrating

$$\log \mathbb{E} e^{\lambda (f(X) - \mathbb{E} f(X))} \leqslant \frac{1}{2} K L^2 \lambda^2.$$

Proof: Applying LSI to $e^{\lambda f/2}$

$$\lambda \mathbb{E} f(X) e^{\lambda f(X)} - \mathbb{E} e^{\lambda f(X)} \log \mathbb{E} e^{\lambda f(X)} \leqslant \frac{1}{2} K \lambda^2 \mathbb{E} |\nabla f(X)|^2 e^{\lambda f(X)}$$

$$\frac{\lambda \mathbb{E} f(X) e^{\lambda f(X)} - \mathbb{E} e^{\lambda f(X)} \log \mathbb{E} e^{\lambda f(X)}}{\lambda^2 \mathbb{E} e^{\lambda f(X)}} \leqslant \frac{1}{2} K L^2$$

$$\frac{d}{d\lambda} \frac{\log \mathbb{E} e^{\lambda f(X)}}{\lambda} \leqslant \frac{1}{2} K L^2$$

Integrating

$$\log \mathbb{E} e^{\lambda (f(X) - \mathbb{E} f(X))} \leqslant \frac{1}{2} K L^2 \lambda^2.$$

Now, by Chebyshev's inequality

$$\mathbb{P}(f(X) - \mathbb{E} f(X) \geqslant t) \leqslant \inf_{\lambda > 0} e^{-\lambda t} \mathbb{E} e^{\lambda (f(X) - \mathbb{E} f(X))}$$

$$= \inf_{\lambda > 0} e^{-\lambda t + K L^2 \lambda^2 / 2} = e^{-\frac{t^2}{2 K L^2}}$$

Non-lipschitz functions

A typical example – Gaussian quadratic form:
$A = A^T = (a_{ij})_{i,j \leqslant n}$, $G = (g_1, \ldots, g_n)$,

$$Z = \langle AG, G \rangle = \sum_{i,j=1}^{n} a_{ij} g_i g_j.$$

Non-lipschitz functions

A typical example – Gaussian quadratic form:
$A = A^T = (a_{ij})_{i,j\leqslant n}$, $G = (g_1, \ldots, g_n)$,

$$Z = \langle AG, G\rangle = \sum_{i,j=1}^{n} a_{ij} g_i g_j.$$

We can diagonalize A in an orthonormal basis and use rotational invariance of G to get

$$Z = \sum_{i=1}^{n} \lambda_i \widetilde{g}_i^2,$$

where $\tilde{g}_i$ – i.i.d. standard Gaussian. Thus, by Bernstein's inequality

$$\mathbb{P}(|Z - \mathbb{E}Z| \geqslant t) \leqslant 2\exp\Big(-c\min\Big(\frac{t^2}{\sum_{i=1}^n \lambda_i^2}, \frac{t}{\max_i |\lambda_i|}\Big)\Big)$$
$$= 2\exp\Big(-c\min\Big(\frac{t^2}{|A|_{HS}^2}, \frac{t}{|A|_{op}}\Big)\Big).$$

Theorem (Hanson-Wright,Borell,Ledoux-Talagrand, Arcones-Giné, Latała)

If A is a symmetric matrix, then for all $t \geqslant 0$,

$$
\begin{aligned}
C^{-1} \exp\Big(-C \min\Big(\frac{t^2}{|A|_{HS}^2}, \frac{t}{|A|_{op}}\Big)\Big) \\
&\leqslant \mathbb{P}(|Z - \mathbb{E}Z| \geqslant t) \\
&\leqslant 2 \exp\Big(-c \min\Big(\frac{t^2}{|A|_{HS}^2}, \frac{t}{|A|_{op}}\Big)\Big).
\end{aligned}
$$

Theorem (Hanson-Wright,Borell,Ledoux-Talagrand, Arcones-Giné, Latała)

If A is a symmetric matrix, then for all $t \geqslant 0$,

$$
\begin{aligned}
C^{-1} \exp\Big(- C \min\Big(\frac{t^2}{|A|_{HS}^2}, \frac{t}{|A|_{op}}\Big)\Big) & \\
&\leqslant \mathbb{P}(|Z - \mathbb{E}Z| \geqslant t) \\
&\leqslant 2 \exp\Big(- c \min\Big(\frac{t^2}{|A|_{HS}^2}, \frac{t}{|A|_{op}}\Big)\Big).
\end{aligned}
$$

Remarks: there are counterparts for quadratic forms in

- $X_1, \ldots, X_n$, where X_i – independent, with subgaussian tails,
- random vectors with concentration property for convex Lipschitz functions.

Non-lipschitz functions via functional inequalities

Theorem (Aida-Stroock 1994)

If X satisfies LSI(K), then for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant \sqrt{2Kr}\|\nabla f(X)\|_r.$$

Non-lipschitz functions via functional inequalities

Theorem (Aida-Stroock 1994)

If X satisfies LSI(K), then for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant \sqrt{2Kr}\|\nabla f(X)\|_r.$$

- In the Gaussian case the result was obtained earlier by Maurey-Pisier

Non-lipschitz functions via functional inequalities

Theorem (Aida-Stroock 1994)

If X satisfies LSI(K), then for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant \sqrt{2Kr}\|\nabla f(X)\|_r.$$

- In the Gaussian case the result was obtained earlier by Maurey-Pisier
- The proof is a version of Herbst's argument, one differentiates $r \mapsto \|f(X) - \mathbb{E}f(X)\|_r^2$.

Non-lipschitz functions via functional inequalities

Theorem (Aida-Stroock 1994)

If X satisfies LSI(K), then for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant \sqrt{2Kr}\|\nabla f(X)\|_r.$$

- In the Gaussian case the result was obtained earlier by Maurey-Pisier
- The proof is a version of Herbst's argument, one differentiates $r \mapsto \|f(X) - \mathbb{E}f(X)\|_r^2$.
- It is a part of folklore that the Poincaré inequality $\mathrm{Var}(f(X)) \leqslant K\mathbb{E}|\nabla f(X)|^2$ implies

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C\sqrt{K}r\|\nabla f(X)\|_r$$

for $r \geqslant 2$.

Theorem (Aida-Stroock)

If X satisfies $\mathrm{LSI}(K)$, then for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant \sqrt{2Kr}\|\nabla f(X)\|_r.$$

By Chebyshev's ineq. $\mathbb{P}(Z \geqslant e\|Z\|_r) \leqslant e^{-r}$, so if, e.g., $\|\nabla f(X)\|_r \leqslant Lr^\alpha$, then

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geqslant t) \leqslant 2\exp\Big(-c\Big(\frac{t}{L}\Big)^{\frac{2}{1+2\alpha}}\Big).$$

Theorem (Aida-Stroock)

If X satisfies LSI(K), then for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant \sqrt{2Kr}\|\nabla f(X)\|_r.$$

By Chebyshev's ineq. $\mathbb{P}(Z \geqslant e\|Z\|_r) \leqslant e^{-r}$, so if, e.g., $\|\nabla f(X)\|_r \leqslant Lr^\alpha$, then

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geqslant t) \leqslant 2\exp\Big(-c\Big(\frac{t}{L}\Big)^{\frac{2}{1+2\alpha}}\Big).$$

General rule:

- moments $\lesssim ar^\beta$, $\Leftrightarrow$ tails $\lesssim \exp(-c(t/a)^{1/\beta})$.

Theorem (Aida-Stroock)

If X satisfies LSI(K), then for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant \sqrt{2Kr}\|\nabla f(X)\|_r.$$

By Chebyshev's ineq. $\mathbb{P}(Z \geqslant e\|Z\|_r) \leqslant e^{-r}$, so if, e.g., $\|\nabla f(X)\|_r \leqslant Lr^\alpha$, then

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geqslant t) \leqslant 2\exp\Big(-c\Big(\frac{t}{L}\Big)^{\frac{2}{1+2\alpha}}\Big).$$

General rule:

- moments $\lesssim ar^\beta$, $\Leftrightarrow$ tails $\lesssim \exp(-c(t/a)^{1/\beta})$.
- moments $\lesssim ar^\alpha + br^\beta \Leftrightarrow$ tails

$$\lesssim \exp\Big(-c\min\Big(\Big(\frac{t}{a}\Big)^{1/\alpha}, \Big(\frac{t}{b}\Big)^{1/\beta}\Big).$$

Proposition (Bednorz–Wolff–A. 2017)

If X satisfies

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant K\sqrt{r}\|\nabla f(X)\|_r,\ r \geqslant 2,$$

for all $f\colon \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^1$, then for every $f\colon \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^2$ and $r \geqslant 2$,

$$\|f(X) - \mathbb{E}f(X)\|_p \leqslant K\sqrt{r}\mathbb{E}|\nabla f(X)|_2 + K^2 r\Big\||\nabla^2 f(X)|_{op}\Big\|_r$$
$$\leqslant \sqrt{2r}K^2\Big\||\nabla^2 f(X)|_{HS}\Big\|_2 + K\sqrt{r}|\mathbb{E}_\mu \nabla f(X)|_2 + K^2 r\Big\||\nabla^2 f(X)|_{op}\Big\|_r.$$

Proposition (Bednorz–Wolff–A. 2017)

If X satisfies

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant K\sqrt{r}\|\nabla f(X)\|_r, \ r \geqslant 2,$$

for all $f\colon \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^1$, then for every $f\colon \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^2$ and $r \geqslant 2$,

$$\|f(X) - \mathbb{E}f(X)\|_p \leqslant K\sqrt{r}\mathbb{E}|\nabla f(X)|_2 + K^2 r \Big\| |\nabla^2 f(X)|_{op} \Big\|_r$$
$$\leqslant \sqrt{2r}K^2 \Big\| |\nabla^2 f(X)|_{HS} \Big\|_2 + K\sqrt{r}|\mathbb{E}_\mu \nabla f(X)|_2 + K^2 r \Big\| |\nabla^2 f(X)|_{op} \Big\|_r.$$

As a consequence, if $|\nabla^2 f|_{op} \leqslant L$, then

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geqslant t)$$
$$\leqslant 2\exp\Big(-c\min\Big(\frac{t^2}{K^4\mathbb{E}|\nabla^2 f(X)|_{HS}^2 + K^2|\mathbb{E}\nabla f(X)|_2}, \frac{t}{K^2 L}\Big)\Big).$$

$$
\begin{aligned}
&\mathbb{P}(|f(X)-\mathbb{E}f(X)|\geqslant t)\\
&\leqslant 2\exp\Big(-c\min\Big(\frac{t^2}{K^4\mathbb{E}|\nabla^2 f(X)|_{HS}^2+K^2|\mathbb{E}\nabla f(X)|_2},\frac{t}{K^2L}\Big)\Big).
\end{aligned}
$$

$$\begin{aligned}&\mathbb{P}(|f(X)-\mathbb{E}f(X)|\geqslant t)\\ &\leqslant 2\exp\Big(-c\min\Big(\frac{t^2}{K^4\mathbb{E}|\nabla^2 f(X)|_{HS}^2+K^2|\mathbb{E}\nabla f(X)|_2},\frac{t}{K^2L}\Big)\Big).\end{aligned}$$

For $f(x)=\langle Ax,x\rangle$ and $\mathbb{E}X=0$, $\mathrm{Cov}(X)=Id$ we have

$$\mathbb{E}\nabla f(X)=\mathbb{E}(2AX)=0,\ \nabla^2 f(X)=2A,$$

so we get

$$\mathbb{P}(|f(X)-\mathbb{E}f(X)|\geqslant t)\leqslant 2\exp\Big(-c'\min\Big(\frac{t^2}{K^4|A|_{HS}^2},\frac{t}{K^2|A|_{op}}\Big)\Big),$$

recovering the Hanson-Wright estimate.

Higher degree Gaussian polynomials – Latała's theorem

Notation:

- $A = (a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \leqslant n}$

Higher degree Gaussian polynomials – Latała's theorem

Notation:

- $A = (a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \leqslant n}$
- $G_1 = (g_1^{(i)}, \ldots, g_n^{(i)})$ – i.i.d. standard Gaussians in $\mathbb{R}^n$

Higher degree Gaussian polynomials – Latała's theorem

Notation:

- $A = (a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \leqslant n}$
- $G_1 = (g_1^{(i)}, \ldots, g_n^{(i)})$ – i.i.d. standard Gaussians in $\mathbb{R}^n$
- $x^{(1)} \otimes \cdots \otimes x^{(d)} = (x_{i_1}^{(1)} \cdots x_{i_d}^{(d)})_{i_1,\ldots,i_d \leqslant n}$ for $x^{(j)} \in \mathbb{R}^n$

Higher degree Gaussian polynomials – Latała's theorem

Notation:

- $A = (a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \leqslant n}$
- $G_1 = (g_1^{(i)}, \ldots, g_n^{(i)})$ – i.i.d. standard Gaussians in $\mathbb{R}^n$
- $x^{(1)} \otimes \cdots \otimes x^{(d)} = (x_{i_1}^{(1)} \cdots x_{i_d}^{(d)})_{i_1,\ldots,i_d \leqslant n}$ for $x^{(j)} \in \mathbb{R}^n$
- $$\langle A, G^{\otimes d} \rangle = \sum_{i_1,\ldots,i_d=1}^{n} a_{i_1,\ldots,i_d} g_{i_1} \cdots g_{i_d},$$
  $$\langle A, G_1 \otimes \cdots \otimes G_d \rangle = \sum_{i_1,\ldots,i_d=1}^{n} a_{i_1,\ldots,i_d} g_{i_1}^{(1)} \cdots g_{i_d}^{(d)}.$$

Higher degree Gaussian polynomials – Latała's theorem

Notation:

- $A = (a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \leqslant n}$
- $G_1 = (g_1^{(i)}, \ldots, g_n^{(i)})$ – i.i.d. standard Gaussians in $\mathbb{R}^n$
- $x^{(1)} \otimes \cdots \otimes x^{(d)} = (x_{i_1}^{(1)} \cdots x_{i_d}^{(d)})_{i_1,\ldots,i_d \leqslant n}$ for $x^{(j)} \in \mathbb{R}^n$
-
$$\langle A, G^{\otimes d} \rangle = \sum_{i_1,\ldots,i_d=1}^{n} a_{i_1,\ldots,i_d} g_{i_1} \cdots g_{i_d},$$
$$\langle A, G_1 \otimes \cdots \otimes G_d \rangle = \sum_{i_1,\ldots,i_d=1}^{n} a_{i_1,\ldots,i_d} g_{i_1}^{(1)} \cdots g_{i_d}^{(d)}.$$
- Let P_d be the family of partitions of $\{1, \ldots, d\}$ into nonempty, pairwise disjoint sets
- For $\mathcal{I} = \{I_1, \ldots, I_k\} \in P_d$ define

Higher degree Gaussian polynomials – Latała's theorem

Notation:

- $A = (a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \leqslant n}$
- $G_1 = (g_1^{(i)}, \ldots, g_n^{(i)})$ – i.i.d. standard Gaussians in $\mathbb{R}^n$
- $x^{(1)} \otimes \cdots \otimes x^{(d)} = (x_{i_1}^{(1)} \cdots x_{i_d}^{(d)})_{i_1,\ldots,i_d \leqslant n}$ for $x^{(j)} \in \mathbb{R}^n$
- $$\langle A, G^{\otimes d} \rangle = \sum_{i_1,\ldots,i_d=1}^{n} a_{i_1,\ldots,i_d} g_{i_1} \cdots g_{i_d},$$
$$\langle A, G_1 \otimes \cdots \otimes G_d \rangle = \sum_{i_1,\ldots,i_d=1}^{n} a_{i_1,\ldots,i_d} g_{i_1}^{(1)} \cdots g_{i_d}^{(d)}.$$
- Let P_d be the family of partitions of $\{1, \ldots, d\}$ into nonempty, pairwise disjoint sets
- For $\mathcal{I} = \{I_1, \ldots, I_k\} \in P_d$ define

$$|A|_{\mathcal{I}} = \sup\Big\{ \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} \prod_{j=1}^{k} x_{i_{I_j}}^{(j)} \colon |(x_{i_J})|_2 \leqslant 1, j \leqslant k \Big\}$$

$$
\begin{aligned}
|(a_{ij})_{i,j\leqslant n}|_{\{1,2\}} &= \sup\Big\{\sum_{i,j\leqslant n} a_{ij}x_{ij} \colon \sum_{i,j\leqslant n} x_{ij}^2 \leqslant 1\Big\} \\
&= \sqrt{\sum_{i,j\leqslant n} a_{ij}^2} = |(a_{ij})_{i,j\leqslant n}|_{HS},
\end{aligned}
$$

$$
\begin{aligned}
|(a_{ij})_{i,j\leqslant n}|_{\{1,2\}} &= \sup\Big\{\sum_{i,j\leqslant n} a_{ij}x_{ij}\colon \sum_{i,j\leqslant n} x_{ij}^2 \leqslant 1\Big\} \\
&= \sqrt{\sum_{i,j\leqslant n} a_{ij}^2} = |(a_{ij})_{i,j\leqslant n}|_{HS},
\end{aligned}
$$

$$
\begin{aligned}
|(a_{ij})_{i,j\leqslant n}|_{\{1\}\{2\}} &= \sup\Big\{\sum_{i,j\leqslant n} a_{ij}x_iy_j\colon \sum_{i\leqslant n} x_i^2 \leqslant 1, \sum_{j\leqslant n} y_j^2 \leqslant 1\Big\} \\
&= |(a_{ij})_{i,j\leqslant n}|_{op},
\end{aligned}
$$

$$
\begin{aligned}
|(a_{ij})_{i,j\leqslant n}|_{\{1,2\}} &= \sup\Big\{\sum_{i,j\leqslant n} a_{ij}x_{ij}\colon \sum_{i,j\leqslant n} x_{ij}^2 \leqslant 1\Big\}\\
&= \sqrt{\sum_{i,j\leqslant n} a_{ij}^2} = |(a_{ij})_{i,j\leqslant n}|_{HS},
\end{aligned}
$$

$$
\begin{aligned}
|(a_{ij})_{i,j\leqslant n}|_{\{1\}\{2\}} &= \sup\Big\{\sum_{i,j\leqslant n} a_{ij}x_iy_j\colon \sum_{i\leqslant n} x_i^2 \leqslant 1, \sum_{j\leqslant n} y_j^2 \leqslant 1\Big\}\\
&= |(a_{ij})_{i,j\leqslant n}|_{op},
\end{aligned}
$$

$$
|(a_{ijk})_{i,j,k\leqslant n}|_{\{1,2\}\{3\}} = \sup\Big\{\sum_{i,j,k\leqslant n} a_{ijk}x_{ij}y_k\colon \sum_{i,j\leqslant n} x_{ij}^2 \leqslant 1, \sum_{k\leqslant n} y_k^2 \leqslant 1\Big\},
$$

$$
\begin{aligned}
|(a_{ij})_{i,j\leqslant n}|_{\{1,2\}} &= \sup\Big\{\sum_{i,j\leqslant n} a_{ij}x_{ij}\colon \sum_{i,j\leqslant n} x_{ij}^2 \leqslant 1\Big\} \\
&= \sqrt{\sum_{i,j\leqslant n} a_{ij}^2} = |(a_{ij})_{i,j\leqslant n}|_{HS},
\end{aligned}
$$

$$
\begin{aligned}
|(a_{ij})_{i,j\leqslant n}|_{\{1\}\{2\}} &= \sup\Big\{\sum_{i,j\leqslant n} a_{ij}x_iy_j\colon \sum_{i\leqslant n} x_i^2 \leqslant 1, \sum_{j\leqslant n} y_j^2 \leqslant 1\Big\} \\
&= |(a_{ij})_{i,j\leqslant n}|_{op},
\end{aligned}
$$

$$
|(a_{ijk})_{i,j,k\leqslant n}|_{\{1,2\}\{3\}} = \sup\Big\{\sum_{i,j,k\leqslant n} a_{ijk}x_{ij}y_k\colon \sum_{i,j\leqslant n} x_{ij}^2 \leqslant 1, \sum_{k\leqslant n} y_k^2 \leqslant 1\Big\},
$$

$$
\begin{aligned}
|(a_{ijk})_{i,j,k\leqslant n}|_{\{1,2,3\}} &= \sup\Big\{\sum_{i,j,k\leqslant n} a_{ijk}x_{ijk}\colon \sum_{i,j,k\leqslant n} x_{ijk}^2 \leqslant 1\Big\} \\
&= \sqrt{\sum_{i,j,k\leqslant n} a_{ijk}^2}.
\end{aligned}
$$

Theorem (Latała 2005)

For every $r \geqslant 2$,

$$C_d^{-1} \sum_{\mathcal{I} \in P_d} r^{|\mathcal{I}|/2} |A|_{\mathcal{I}} \leqslant \|\langle A, G_1 \otimes \cdots \otimes G_d \rangle\|_r \leqslant C_d \sum_{\mathcal{I} \in P_d} r^{|\mathcal{I}|/2} |A|_{\mathcal{I}},$$

Theorem (Latała 2005)

For every $r \geqslant 2$,

$$C_d^{-1} \sum_{\mathcal{I} \in P_d} r^{|\mathcal{I}|/2} |A|_{\mathcal{I}} \leqslant \|\langle A, G_1 \otimes \cdots \otimes G_d \rangle\|_r \leqslant C_d \sum_{\mathcal{I} \in P_d} r^{|\mathcal{I}|/2} |A|_{\mathcal{I}},$$

As a consequence for all $t \geqslant 0$,

$$C_d^{-1} \exp\Big(- C_d \min_{\mathcal{I} \in P_d} \Big(\frac{t}{|A|_{\mathcal{I}}}\Big)^{\frac{2}{|\mathcal{I}|}} \Big) \leqslant \mathbb{P}(|\langle A, G_1 \otimes \cdots \otimes G_d \rangle| \geqslant t) \leqslant C_d \exp\Big(- C_d^{-1} \min_{\mathcal{I} \in P_d} \Big(\frac{t}{|A|_{\mathcal{I}}}\Big)^{\frac{2}{|\mathcal{I}|}} \Big)$$

Theorem (Latała 2005)

For every $r \geqslant 2$,

$$C_d^{-1} \sum_{\mathcal{I} \in P_d} r^{|\mathcal{I}|/2} |A|_{\mathcal{I}} \leqslant \| \langle A, G_1 \otimes \cdots \otimes G_d \rangle \|_r \leqslant C_d \sum_{\mathcal{I} \in P_d} r^{|\mathcal{I}|/2} |A|_{\mathcal{I}},$$

As a consequence for all $t \geqslant 0$,

$$C_d^{-1} \exp\Big(- C_d \min_{\mathcal{I} \in P_d} \Big(\frac{t}{|A|_{\mathcal{I}}} \Big)^{\frac{2}{|\mathcal{I}|}} \Big) \leqslant \mathbb{P}(|\langle A, G_1 \otimes \cdots \otimes G_d \rangle| \geqslant t) \leqslant C_d \exp\Big(- C_d^{-1} \min_{\mathcal{I} \in P_d} \Big(\frac{t}{|A|_{\mathcal{I}}} \Big)^{\frac{2}{|\mathcal{I}|}} \Big)$$

The same for $\langle A, G^{\otimes d} \rangle$ if A symmetric with zeros on diagonals.

Example

- $A = (a_{ijk})_{ijk \leqslant n}$ – symmetric, $a_{iik} = 0$
- $$Z = \sum_{ijk} a_{ijk} g_i g_j g_k$$

Example

- $A = (a_{ijk})_{ijk \leqslant n}$ – symmetric, $a_{iik} = 0$
- $$Z = \sum_{ijk} a_{ijk} g_i g_j g_k$$

Then

$$\|Z\|_r \simeq \sqrt{r}\|A\|_{\{1,2,3\}} + r\|A\|_{\{1,2\}\{3\}} + r^{3/2}\|A\|_{\{1\}\{2\}\{3\}}.$$

Example

- $A = (a_{ijk})_{ijk \leqslant n}$ – symmetric, $a_{iik} = 0$
- $$Z = \sum_{ijk} a_{ijk} g_i g_j g_k$$

Then

$$\|Z\|_r \simeq \sqrt{r}\|A\|_{\{1,2,3\}} + r\|A\|_{\{1,2\}\{3\}} + r^{3/2}\|A\|_{\{1\}\{2\}\{3\}}.$$

Equivalently

$$\begin{aligned}&\mathbb{P}(|Z| \geqslant t)\\ &\leqslant 2\exp\Big(-c\min\Big(\Big(\frac{t}{\|A\|_{\{1,2,3\}}}\Big)^2, \frac{t}{\|A\|_{\{1,2\}\{3\}}}, \Big(\frac{t}{\|A\|_{\{1\}\{2\}\{3\}}}\Big)^{2/3}\Big)\Big)\end{aligned}$$

and

$$\begin{aligned}&\mathbb{P}(|Z| \geqslant t)\\ &\geqslant \frac{1}{C}\exp\Big(-C\min\Big(\Big(\frac{t}{\|A\|_{\{1,2,3\}}}\Big)^2, \frac{t}{\|A\|_{\{1,2\}\{3\}}}, \Big(\frac{t}{\|A\|_{\{1\}\{2\}\{3\}}}\Big)^{2/3}\Big)\Big).\end{aligned}$$

Latała type ineq.'s for general functions & measures

Proposition (Wolff–A. 2016)

Assume that X satisfies $\|f(X) - \mathbb{E}f(X)\|_r \leqslant Kr^\gamma \|\nabla f(X)\|_r$ for $r \geqslant 2$. Let $G_1, \ldots, G_d$ be i.i.d. standard Gaussian vectors independent of X. If $f \colon \mathbb{R}^n \to \mathbb{R}$ is of class $\mathcal{C}^D$, then for all $r \geqslant 2$,

$$\begin{aligned}
&\|f(X) - \mathbb{E}f(X)\|_r \\
&\qquad \leqslant C^D K^D r^{\gamma D - D/2} \|\langle \nabla^D f(X), G_1 \otimes \cdots \otimes G_D \rangle\|_r \\
&\qquad\quad + \sum_{1 \leqslant d \leqslant D-1} C^d K^d r^{\gamma d - d/2} \|\langle \mathbb{E}_X \nabla^d f(X), G_1 \otimes \cdots \otimes G_d \rangle\|_r.
\end{aligned}$$

Proof for $D = 1$ and $D = 2$

Main observation $\|\langle G, x\rangle\|_r \simeq \sqrt{r}|x|$.

$\|f(X) - \mathbb{E}f(X)\|_r^r \leqslant K^r r^{r\gamma} \mathbb{E}|\nabla f(X)|^r \leqslant C^r K^r r^{\gamma r - r/2} \mathbb{E}|\langle \nabla f(X), G\rangle|^r.$

Proof for $D = 1$ and $D = 2$

Main observation $\|\langle G, x\rangle\|_r \simeq \sqrt{r}|x|$.

$$\|f(X) - \mathbb{E}f(X)\|_r^r \leqslant K^r r^{r\gamma} \mathbb{E}|\nabla f(X)|^r \leqslant C^r K^r r^{\gamma r - r/2} \mathbb{E}|\langle \nabla f(X), G\rangle|^r.$$

This gives the case $D = 1$.

Proof for $D = 1$ and $D = 2$

Main observation $\|\langle G, x\rangle\|_r \simeq \sqrt{r}|x|$.

$\|f(X) - \mathbb{E}f(X)\|_r^r \leqslant K^r r^{r\gamma}\mathbb{E}|\nabla f(X)|^r \leqslant C^r K^r r^{\gamma r - r/2}\mathbb{E}|\langle \nabla f(X), G\rangle|^r.$

This gives the case $D = 1$. Now $D = 2$

$$\begin{aligned}
&\|f(X) - \mathbb{E}f(X)\|_r \leqslant CKr^{\gamma-1/2}\|\langle \nabla f(X), G_1\rangle\|_r \\
\leqslant & CKr^{\gamma-1/2}\|\langle \nabla f(X), G_1\rangle - \mathbb{E}_X\langle \nabla f(X), G_1\rangle\|_r \\
&+ CKr^{\gamma-1/2}\|\langle \mathbb{E}_X \nabla f(X), G_1\rangle\|_r \\
\leqslant & C^2K^2r^{2\gamma-1}\|\langle \nabla^2 f(X), G_1 \otimes G_2\rangle\|_r + CKr^{\gamma-1/2}\|\langle \mathbb{E}_X \nabla f(X), G_1\rangle\|_r,
\end{aligned}$$

where we applied the case $D = 1$ conditionally on G_1 to the function $g(x) := \langle \nabla f(x), G_1\rangle$ and used the identity

$$\langle \nabla g(x), G_2\rangle = \langle \nabla^2 f(x), G_1 \otimes G_2\rangle.$$

Corollary (Wolff-A., 2015)

Assume that X satisfies LSI(K). Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a $\mathcal{C}^D$ function. For all $r \geqslant 2$

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C_D\Big(K^D \sum_{\mathcal{J}\in P_D} r^{\frac{\#\mathcal{J}}{2}} \Big\| \|\nabla^D f(X)\|_{\mathcal{J}} \Big\|_r + \sum_{1\leqslant d\leqslant D-1} K^d \sum_{\mathcal{J}\in P_d} r^{\frac{\#\mathcal{J}}{2}} \|\mathbb{E}\nabla^d f(X)\|_{\mathcal{J}}\Big).$$

In particular if $\nabla^D f(x)$ is uniformly bounded then for $t > 0$,

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geqslant t) \leqslant 2\exp\Big(-\frac{1}{C_D} \min\Big(\min_{\mathcal{J}\in P_D} \Big(\frac{t}{K^D \sup_{x\in\mathbb{R}^n} \|\nabla^D f(x)\|_{\mathcal{J}}} \Big)^{\frac{2}{\#\mathcal{J}}}, \min_{1\leqslant d\leqslant D-1} \min_{\mathcal{J}\in P_d} \Big(\frac{t}{K^d \|\mathbb{E}\nabla^d f(X)\|_{\mathcal{J}}} \Big)^{\frac{2}{\#\mathcal{J}}} \Big)\Big).$$

- If f is a polynomial of degree D, then $\nabla^D f$ is constant, so the inequality can be written as

$$\begin{aligned}
&\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geqslant t) \\
&\qquad \leqslant 2\exp\Big(-\frac{1}{C_D} \min_{1 \leqslant d \leqslant D} \min_{\mathcal{J} \in P_d} \Big(\frac{t}{K^d \|\mathbb{E}\nabla^d f(X)\|_{\mathcal{J}}} \Big)^{\frac{2}{\#\mathcal{J}}} \Big).
\end{aligned}$$

- If f is a polynomial of degree D, then $\nabla^D f$ is constant, so the inequality can be written as

$$\begin{aligned}&\mathbb{P}(|f(X)-\mathbb{E}f(X)|\geqslant t)\\&\qquad\leqslant 2\exp\Big(-\frac{1}{C_D}\min_{1\leqslant d\leqslant D}\min_{\mathcal{J}\in P_d}\Big(\frac{t}{K^d\|\mathbb{E}\nabla^d f(X)\|_{\mathcal{J}}}\Big)^{\frac{2}{\#\mathcal{J}}}\Big).\end{aligned}$$

- In the Gaussian setting the inequality for polynomials can be reversed (up to constants).
- The inequality for polynomials generalizes to independent subgaussian variables (Wolff–A., 2015)

- If f is a polynomial of degree D, then $\nabla^D f$ is constant, so the inequality can be written as

$$
\begin{aligned}
&\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geqslant t) \\
&\qquad \leqslant 2\exp\Big(-\frac{1}{C_D}\min_{1\leqslant d\leqslant D}\min_{\mathcal{J}\in P_d}\Big(\frac{t}{K^d\|\mathbb{E}\nabla^d f(X)\|_{\mathcal{J}}}\Big)^{\frac{2}{\#\mathcal{J}}}\Big).
\end{aligned}
$$

- In the Gaussian setting the inequality for polynomials can be reversed (up to constants).
- The inequality for polynomials generalizes to independent subgaussian variables (Wolff–A., 2015)
- There are generalizations to Gaussian polynomials with coefficients in certain classes of Banach spaces (Latała–Meller–A. 2020).

Definition

Let $\Psi\colon [0,\infty) \to [0,\infty]$ be a Young function. We will say that μ on $\mathbb{R}^n$ satisfies a Ψ-modified log-Sobolev inequality if for every $f\colon \mathbb{R}^n \to (0,\infty)$

$$\operatorname{Ent}_\mu f^2 \leqslant K \mathbb{E}_\mu \sum_{i=1}^{n} \Psi\Big(\frac{|\partial_i f|}{f}\Big) f^2$$

Definition

Let $\Psi\colon [0,\infty) \to [0,\infty]$ be a Young function. We will say that μ on $\mathbb{R}^n$ satisfies a Ψ-modified log-Sobolev inequality if for every $f\colon \mathbb{R}^n \to (0,\infty)$

$$\operatorname{Ent}_\mu f^2 \leqslant K\mathbb{E}_\mu \sum_{i=1}^{n} \Psi\Big(\frac{|\partial_i f|}{f}\Big) f^2$$

Remarks

- Reduces to LSI for $\Psi(x) = x^2$,

Definition

Let $\Psi\colon [0,\infty) \to [0,\infty]$ be a Young function. We will say that μ on $\mathbb{R}^n$ satisfies a Ψ-modified log-Sobolev inequality if for every $f\colon \mathbb{R}^n \to (0,\infty)$

$$\operatorname{Ent}_\mu f^2 \leqslant K\mathbb{E}_\mu \sum_{i=1}^{n} \Psi\Big(\frac{|\partial_i f|}{f}\Big) f^2$$

Remarks

- Reduces to LSI for $\Psi(x) = x^2$,
- introduced by Bobkov-Ledoux for $\Psi(x) = x^2$ if $|x| < 1$, $f(x) = \infty$ otherwise,

Definition

Let $\Psi\colon [0,\infty) \to [0,\infty]$ be a Young function. We will say that μ on $\mathbb{R}^n$ satisfies a Ψ-modified log-Sobolev inequality if for every $f\colon \mathbb{R}^n \to (0,\infty)$

$$\operatorname{Ent}_\mu f^2 \leqslant K\mathbb{E}_\mu \sum_{i=1}^{n} \Psi\Big(\frac{|\partial_i f|}{f}\Big) f^2$$

Remarks

- Reduces to LSI for $\Psi(x) = x^2$,
- introduced by Bobkov-Ledoux for $\Psi(x) = x^2$ if $|x| < 1$, $f(x) = \infty$ otherwise,
- in full generality introduced by Gentil-Guillin-Miclo, studied e.g, by Barthe-Roberto, Barthe-Kolesnikov, Barthe-Strzelecki,

Definition

Let $\Psi\colon [0,\infty) \to [0,\infty]$ be a Young function. We will say that μ on $\mathbb{R}^n$ satisfies a Ψ-modified log-Sobolev inequality if for every $f\colon \mathbb{R}^n \to (0,\infty)$

$$\mathrm{Ent}_\mu f^2 \leqslant K\mathbb{E}_\mu \sum_{i=1}^n \Psi\Big(\frac{|\partial_i f|}{f}\Big) f^2$$

Remarks

- Reduces to LSI for $\Psi(x) = x^2$,
- introduced by Bobkov-Ledoux for $\Psi(x) = x^2$ if $|x| < 1$, $f(x) = \infty$ otherwise,
- in full generality introduced by Gentil-Guillin-Miclo, studied e.g, by Barthe-Roberto, Barthe-Kolesnikov, Barthe-Strzelecki,
- shares many properties of LSI, e.g., tensorization,

Definition

Let $\Psi\colon [0,\infty) \to [0,\infty]$ be a Young function. We will say that μ on $\mathbb{R}^n$ satisfies a Ψ-modified log-Sobolev inequality if for every $f\colon \mathbb{R}^n \to (0,\infty)$

$$\operatorname{Ent}_\mu f^2 \leqslant K\mathbb{E}_\mu \sum_{i=1}^n \Psi\Big(\frac{|\partial_i f|}{f}\Big) f^2$$

Remarks

- Reduces to LSI for $\Psi(x) = x^2$,
- introduced by Bobkov-Ledoux for $\Psi(x) = x^2$ if $|x| < 1$, $f(x) = \infty$ otherwise,
- in full generality introduced by Gentil-Guillin-Miclo, studied e.g, by Barthe-Roberto, Barthe-Kolesnikov, Barthe-Strzelecki,
- shares many properties of LSI, e.g., tensorization,
- implies concentration with various profiles, expressed in terms of various norms of gradients of f.

For a Young function Ψ define a norm $|\cdot|_{\Psi_r}$ on $\mathbb{R}^n$ as

$$|x|_{\Psi_r} = \inf\left\{a > 0\colon \sum_{i=1}^{n} \Psi\Big(\frac{r|x_i|}{a}\Big) \leqslant r\right\}$$

For a Young function Ψ define a norm $|\cdot|_{\Psi_r}$ on $\mathbb{R}^n$ as

$$|x|_{\Psi_r} = \inf\Big\{a > 0\colon \sum_{i=1}^{n} \Psi\Big(\frac{r|x_i|}{a}\Big) \leqslant r\Big\}$$

Examples

- if $\Psi(x) = x^2$, then $|x|_{\Psi_r} = \sqrt{r}|x|_2$

For a Young function Ψ define a norm $|\cdot|_{\Psi_r}$ on $\mathbb{R}^n$ as

$$|x|_{\Psi_r} = \inf\Big\{a > 0\colon \sum_{i=1}^{n} \Psi\Big(\frac{r|x_i|}{a}\Big) \leqslant r\Big\}$$

Examples

- if $\Psi(x) = x^2$, then $|x|_{\Psi_r} = \sqrt{r}|x|_2$
- if $\Psi(x) \simeq \max(x^2, x^\alpha)$ for $\alpha > 2$, then

$$|x|_{\Psi_r} \simeq \sqrt{r}|x|_2 + r^{1/\alpha^*}|x|_\alpha \lesssim r^{1/\alpha^*}|x|_2$$

Theorem (Bednorz-Wolff-A. 2017)

Assume that there are $1 < \alpha \leqslant 2 \leqslant \beta < \infty$ and $R \geqslant 1$ such that for $x > 0, t > 1$,

$$R^{-1} t^{\alpha} \leqslant \frac{\Psi(tx)}{\Psi(x)} \leqslant R t^{\beta}.$$

If X satisfies the Ψ-modified LSI with constant K, then for every smooth $f \colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$,

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C_{K,R,\alpha,\beta} \Big\| |\nabla f(X)|_{\Psi_r} \Big\|_r$$

Example: $X = (X_1, \ldots, X_n)$ where X_i–i.i.d. with density $\frac{1}{Z}e^{-|x|^\alpha}$, $\alpha \in (1, 2]$. Then

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C_\alpha\Big(\sqrt{r}\Big\||\nabla f(X)|_2\Big\|_r + r^{1/\alpha}\Big\||\nabla f(X)|_{\alpha^*}\Big\|_r\Big)$$

Example: $X = (X_1, \ldots, X_n)$ where X_i–i.i.d. with density $\frac{1}{Z}e^{-|x|^\alpha}$, $\alpha \in (1, 2]$. Then

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C_\alpha\Big(\sqrt{r}\Big\||\nabla f(X)|_2\Big\|_r + r^{1/\alpha}\Big\||\nabla f(X)|_{\alpha^*}\Big\|_r\Big)$$

As a consequence if $a = \sup_x |\nabla f(x)|_2$, $b = \sup_x |\nabla f(x)|_{\alpha^*}$, then

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geqslant t) \leqslant 2\exp\Big(-c_\alpha \min\Big(\frac{t^2}{a^2}, \frac{t^\alpha}{b^\alpha}\Big)\Big)$$

Example: $X = (X_1, \dots, X_n)$ where X_i–i.i.d. with density $\frac{1}{Z} e^{-|x|^\alpha}$, $\alpha \in (1, 2]$. Then

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C_\alpha \Big(\sqrt{r} \Big\| |\nabla f(X)|_2 \Big\|_r + r^{1/\alpha} \Big\| |\nabla f(X)|_{\alpha^*} \Big\|_r \Big)$$

As a consequence if $a = \sup_x |\nabla f(x)|_2$, $b = \sup_x |\nabla f(x)|_{\alpha^*}$, then

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geqslant t) \leqslant 2 \exp\Big(- c_\alpha \min\Big(\frac{t^2}{a^2}, \frac{t^\alpha}{b^\alpha} \Big) \Big)$$

Other tail decays if one assumes growth conditions on gradients.

Question

Assume $X = (X_1, \ldots, X_n)$, where X_i – i.i.d. with density $\frac{1}{2}e^{-|x|}$. Do we have moment estimates of the form

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C\Big(\sqrt{p}\Big\||\nabla f(X)|_2\Big\|_r + r\Big\||\nabla f(X)|_\infty\Big\|_r\Big)?$$

Question

Assume $X = (X_1, \ldots, X_n)$, where X_i – i.i.d. with density $\frac{1}{2}e^{-|x|}$. Do we have moment estimates of the form

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C\Big(\sqrt{p}\Big\||\nabla f(X)|_2\Big\|_r + r\Big\||\nabla f(X)|_\infty\Big\|_r\Big)?$$

Positive answer would allow for a transference principle from polynomials as in the Gaussian case.

A more abstract setting

- $(\mathcal{X}, \mathcal{F}, \mu)$ – a probability space
- $(X_t)_{t \geqslant 0}$ – a revers. Markov process with inv. measure μ
- L – the generator of the corresponding semigroup $(P_t)_{t \geqslant 0}$ of operators on $L_2(\mu)$
- $\mathcal{E}(f, g) = -\mathbb{E}_\mu(fLg)$ – the Dirichlet form
- $\Gamma(f, g) = \frac{1}{2}(L(fg) - fLg - gLf)$ – carré du champ operator
- $\mathcal{E}(f, g) = \mathbb{E}_\mu \Gamma(f, g)$

A more abstract setting

- $(\mathcal{X}, \mathcal{F}, \mu)$ – a probability space
- $(X_t)_{t\geqslant 0}$ – a revers. Markov process with inv. measure μ
- L – the generator of the corresponding semigroup $(P_t)_{t\geqslant 0}$ of operators on $L_2(\mu)$
- $\mathcal{E}(f,g) = -\mathbb{E}_\mu(fLg)$ – the Dirichlet form
- $\Gamma(f,g) = \frac{1}{2}(L(fg) - fLg - gLf)$ – carré du champ operator
- $\mathcal{E}(f,g) = \mathbb{E}_\mu\Gamma(f,g)$

Warning: For this talk I will disregard the question of domains.

Examples:

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

$$Lf(x) = \Delta f(x) - \langle \nabla V(x), \nabla f(x) \rangle,$$
$$\Gamma(f,g) = \langle \nabla f, \nabla g \rangle,\ \mathcal{E}(f,g) = \mathbb{E}_\mu \langle \nabla f, \nabla g \rangle.$$

Examples:

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

$$Lf(x) = \Delta f(x) - \langle \nabla V(x), \nabla f(x) \rangle,$$
$$\Gamma(f,g) = \langle \nabla f, \nabla g \rangle,\ \mathcal{E}(f,g) = \mathbb{E}_\mu \langle \nabla f, \nabla g \rangle.$$

- $Q_x(\cdot)$ – a kernel on $\mathcal{X}$ satisfying the detailed balance condition

$$Q_x(dy)\mu(dx) = Q_y(dx)\mu(dy)$$

Examples:

- $\mu(dx) = \frac{1}{Z} e^{-V(x)} dx$ for some $V\colon \mathbb{R}^n \to \mathbb{R}$. Then
$$\begin{aligned} Lf(x) &= \Delta f(x) - \langle \nabla V(x), \nabla f(x) \rangle, \\ \Gamma(f,g) &= \langle \nabla f, \nabla g \rangle,\ \mathcal{E}(f,g) = \mathbb{E}_\mu \langle \nabla f, \nabla g \rangle. \end{aligned}$$

- $Q_x(\cdot)$ – a kernel on $\mathcal{X}$ satisfying the detailed balance condition
$$Q_x(dy)\mu(dx) = Q_y(dx)\mu(dy)$$
One can define
$$Lf(x) = \int_{\mathcal{X}} (f(y) - f(x)) Q_x(dy)$$
Then
$$\begin{aligned} \Gamma(f,g)(x) &= \frac{1}{2} \int_{\mathcal{X}} (f(y) - f(x))(g(y) - g(x)) Q_x(dy), \\ \mathcal{E}(f,g) &= \frac{1}{2} \int_{\mathcal{X}} \int_{\mathcal{X}} (f(y) - f(x))(g(y) - g(x)) Q_x(dy)\mu(dx). \end{aligned}$$

Definition

Let us introduce a family of inequalities

- the log-Sobolev inequality:

$$\text{(LSI)} \qquad \operatorname{Ent}_\mu f^2 \leqslant K\, \mathcal{E}(f,f)$$

Definition

Let us introduce a family of inequalities

- the log-Sobolev inequality:

$$\text{(LSI)} \qquad \operatorname{Ent}_\mu f^2 \leqslant K\,\mathcal{E}(f,f)$$

- modified log-Sobolev inequality:

$$\text{(mLSI)} \qquad \operatorname{Ent}_\mu f \leqslant K\,\mathcal{E}(f,\log f)$$

  for positive f.

Definition

Let us introduce a family of inequalities

- the log-Sobolev inequality:

  (LSI) $$\operatorname{Ent}_\mu f^2 \leqslant K\,\mathcal{E}(f,f)$$

- modified log-Sobolev inequality:

  (mLSI) $$\operatorname{Ent}_\mu f \leqslant K\,\mathcal{E}(f,\log f)$$

  for positive f.

- LSI is responsible for hypercontractivity of the semigroup,

Definition

Let us introduce a family of inequalities

- the log-Sobolev inequality:

  (LSI) $$\mathrm{Ent}_\mu f^2 \leqslant K\,\mathcal{E}(f,f)$$

- modified log-Sobolev inequality:

  (mLSI) $$\mathrm{Ent}_\mu f \leqslant K\,\mathcal{E}(f,\log f)$$

  for positive f.

- LSI is responsible for hypercontractivity of the semigroup,
- mLSI is responsible for exponential decay of $\mathrm{Ent}(P_t f)$ (convergence to equilibrium in relative entropy).

Definition

Let us introduce a family of inequalities

- the log-Sobolev inequality:

$$\text{(LSI)} \qquad \mathrm{Ent}_\mu f^2 \leqslant K\,\mathcal{E}(f,f)$$

- modified log-Sobolev inequality:

$$\text{(mLSI)} \qquad \mathrm{Ent}_\mu f \leqslant K\,\mathcal{E}(f,\log f)$$

  for positive f.

- LSI is responsible for hypercontractivity of the semigroup,
- mLSI is responsible for exponential decay of $\mathrm{Ent}(P_t f)$ (convergence to equilibrium in relative entropy).
- For diffusions (under chain rule) they are equivalent

Definition

Let us introduce a family of inequalities

- the log-Sobolev inequality:

$$\text{(LSI)} \qquad \mathrm{Ent}_\mu f^2 \leqslant K\,\mathcal{E}(f,f)$$

- modified log-Sobolev inequality:

$$\text{(mLSI)} \qquad \mathrm{Ent}_\mu f \leqslant K\,\mathcal{E}(f,\log f)$$

  for positive f.

- LSI is responsible for hypercontractivity of the semigroup,
- mLSI is responsible for exponential decay of $\mathrm{Ent}(P_t f)$ (convergence to equilibrium in relative entropy).
- For diffusions (under chain rule) they are equivalent
- In general LSI is strictly stronger

Theorem (Aida-Stroock – abstract version)

If μ satisfies $LSI(K)$, then for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$

$$\|f - \mathbb{E}_\mu f\|_r \leqslant \sqrt{2Kr}\left\|\sqrt{\Gamma(f,f)}\right\|_r.$$

Theorem (Aida-Stroock – abstract version)

If μ satisfies $LSI(K)$, then for all smooth $f\colon \mathbb{R}^n \to \mathbb{R}$ and $r \geqslant 2$

$$\|f - \mathbb{E}_\mu f\|_r \leqslant \sqrt{2Kr}\left\|\sqrt{\Gamma(f,f)}\right\|_r.$$

Natural question: Can one obtain a similar inequality under $mLSI$?

BBLM inequalities

Theorem (Boucheron-Bousquet-Lugosi-Massart, 2005)

Let $X = (X_1, \dots, X_n)$, where X_i are independent random variables. For any function f and $r \geqslant 2$,

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C\sqrt{r}\Big\|\Big(\sum_{i=1}^{n} \mathbb{E}\Big((f(X) - f(X^{(i)}))^2 \Big| X\Big)\Big)^{1/2}\Big\|_r,$$

BBLM inequalities

Theorem (Boucheron-Bousquet-Lugosi-Massart, 2005)

Let $X = (X_1, \dots, X_n)$, where X_i are independent random variables. For any function f and $r \geqslant 2$,

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C\sqrt{r}\Big\|\Big(\sum_{i=1}^{n} \mathbb{E}\Big((f(X) - f(X^{(i)}))^2 \Big| X\Big)\Big)^{1/2}\Big\|_r,$$

where $X^{(i)} = (X_1, \dots, X_{i-1}, \tilde{X}_i, X_{i+1}, \dots, X_n)$ and $\tilde{X}_i$'s are independent copies of X_i's.

BBLM inequalities

Theorem (Boucheron-Bousquet-Lugosi-Massart, 2005)

Let $X = (X_1, \dots, X_n)$, where X_i are independent random variables. For any function f and $r \geqslant 2$,

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C\sqrt{r}\left\|\left(\sum_{i=1}^{n} \mathbb{E}\left((f(X) - f(X^{(i)}))^2 \middle| X\right)\right)^{1/2}\right\|_r,$$

where $X^{(i)} = (X_1, \dots, X_{i-1}, \tilde{X}_i, X_{i+1}, \dots, X_n)$ and $\tilde{X}_i$'s are independent copies of X_i's.

- This generalizes (up to constant) the Efron-Stein ineq. $(r = 2)$

BBLM inequalities

Theorem (Boucheron-Bousquet-Lugosi-Massart, 2005)

Let $X = (X_1, \ldots, X_n)$, where X_i are independent random variables. For any function f and $r \geqslant 2$,

$$\|f(X) - \mathbb{E}f(X)\|_r \leqslant C\sqrt{r}\Big\|\Big(\sum_{i=1}^{n}\mathbb{E}\Big((f(X) - f(X^{(i)}))^2\Big|X\Big)\Big)^{1/2}\Big\|_r,$$

where $X^{(i)} = (X_1, \ldots, X_{i-1}, \tilde{X}_i, X_{i+1}, \ldots, X_n)$ and $\tilde{X}_i$'s are independent copies of X_i's.

- This generalizes (up to constant) the Efron-Stein ineq. $(r = 2)$
- Many applications in geometry, statistics, combinatorics.

Definition

Let $p \in (1,2]$. We say that Beckner inequality is satisfied if for $f\colon \mathcal{X} \to [0,\infty)$,

$$\mathbb{E}_\mu f^p - (\mathbb{E}_\mu f)^p \leqslant B_p \frac{p}{2}\, \mathcal{E}(f, f^{p-1}).$$

Definition

Let $p \in (1,2]$. We say that Beckner inequality is satisfied if for $f\colon \mathcal{X} \to [0,\infty)$,

$$\mathbb{E}_\mu f^p - (\mathbb{E}_\mu f)^p \leqslant B_p \frac{p}{2}\, \mathcal{E}(f, f^{p-1}).$$

- for each individual p, Bec_p is equivalent (up to constants depending on p) to the Poincaré inequality $\mathrm{Var}_\mu(f) \leqslant K\, \mathcal{E}(f,f)$.

Definition

Let $p \in (1,2]$. We say that Beckner inequality is satisfied if for $f\colon \mathcal{X} \to [0,\infty)$,

$$\mathbb{E}_\mu f^p - (\mathbb{E}_\mu f)^p \leqslant B_p \frac{p}{2}\, \mathcal{E}(f, f^{p-1}).$$

- for each individual p, Bec_p is equivalent (up to constants depending on p) to the Poincaré inequality $\mathrm{Var}_\mu(f) \leqslant K\, \mathcal{E}(f,f)$.
- If $\sup_p B_p < \infty$ then we can infer mLSI (divide by $p-1$ and pass with p to 1).

Proposition (Polaczyk-Strzelecki-A. 2020)

Assume that for all $p \in (1,2]$ the inequality Bec_p holds with $B_p \leqslant \frac{K}{(p-1)^s}$ for some $s, K \geqslant 0$. Then for all $r \geqslant 2$,

$$\|f - \mathbb{E}_\mu f\|_r \leqslant C K^{1/2} r^{\frac{s+1}{2}} \left\| \sqrt{\Gamma(f)} \right\|_r$$

Proposition (Polaczyk-Strzelecki-A. 2020)

Assume that for all $p \in (1,2]$ the inequality Bec_p holds with $B_p \leqslant \frac{K}{(p-1)^s}$ for some $s, K \geqslant 0$. Then for all $r \geqslant 2$,

$$\|f - \mathbb{E}_\mu f\|_r \leqslant CK^{1/2} r^{\frac{s+1}{2}} \left\| \sqrt{\Gamma(f)} \right\|_r$$

The proof follows closely the argument by Boucheron et al., rewritten in an abstract form.

Proposition (Polaczyk-Strzelecki-A. 2020)

Assume that for all $p \in (1,2]$ the inequality Bec_p holds with $B_p \leqslant \frac{K}{(p-1)^s}$ for some $s, K \geqslant 0$. Then for all $r \geqslant 2$,

$$\|f - \mathbb{E}_\mu f\|_r \leqslant C K^{1/2} r^{\frac{s+1}{2}} \left\| \sqrt{\Gamma(f)} \right\|_r$$

The proof follows closely the argument by Boucheron et al., rewritten in an abstract form.

Theorem (Polaczyk-Strzelecki-A. 2020)

If $\mathrm{mLSI}(K)$ is satisfied then for all $p \in (1,2]$ the inequality Bec_p holds with $B_p \leqslant 6K$.

Proposition (Polaczyk-Strzelecki-A. 2020)

Assume that for all $p \in (1,2]$ the inequality Bec_p holds with $B_p \leqslant \frac{K}{(p-1)^s}$ for some $s, K \geqslant 0$. Then for all $r \geqslant 2$,

$$\|f - \mathbb{E}_\mu f\|_r \leqslant C K^{1/2} r^{\frac{s+1}{2}} \left\| \sqrt{\Gamma(f)} \right\|_r$$

The proof follows closely the argument by Boucheron et al., rewritten in an abstract form.

Theorem (Polaczyk-Strzelecki-A. 2020)

If $\mathrm{mLSI}(K)$ is satisfied then for all $p \in (1,2]$ the inequality Bec_p holds with $B_p \leqslant 6K$.

In fact the optimal constants satisfy $2K^{\mathrm{opt}} = \lim_{p \to 1^+} B_p^{\mathrm{opt}}$ and $B_2^{\mathrm{opt}} = 2K^{\mathrm{opt}}$ but we don't know if one can replace 6 by 2 in general.

Corollary

If mLSI(K) holds then for every $r \geqslant 2$,

$$\|f - \mathbb{E}_\mu f\|_r \leqslant C\sqrt{Kr}\left\|\sqrt{\Gamma(f)}\right\|_r$$

Corollary

If mLSI(K) holds then for every $r \geqslant 2$,

$$\|f - \mathbb{E}_\mu f\|_r \leqslant C\sqrt{Kr}\left\|\sqrt{\Gamma(f)}\right\|_r$$

Remark In the kernel case we can also obtain estimates of the form

$$\|(f - \mathbb{E}_\mu f)_+\|_r \leqslant C\sqrt{Kr}\left\|\sqrt{\Gamma_+(f)}\right\|_r,$$

where

$$\Gamma_+(f) = \int_{\mathcal{X}} (f(x) - f(y))_+^2 Q_x(dy).$$

(BBLM also had bounds of this type in the product case).

Applications

- Moment inequalities for Cauchy-type measures, for measures of the form $e^{-|x|_\alpha^\alpha}$, and on manifolds (based on Beckner-type inequalities due to Bakry, Gentil, and Scheffer, Latała-Oleszkiewicz and Wang)
- Inequalities for Glauber dynamics: Ising model, Exponential Random Graphs (mLSI by Marton, Götze-Sambale-Sinulis)
- Symmetric group (mLSI by Gao-Quastel, mLSI, Beckner by Bobkov-Tetali): Hoeffding statistics and their suprema
- Measures with Stochastic covering property (mLSI by Hermon-Salez)
- Zero range processes (mLSI by Hermon-Salez)
- the Poisson space (mLSI by Wu)

Examples: Glauber dynamics.

- $\mathcal{X} = E^I$, where I is a finite set

Examples: Glauber dynamics.

- $\mathcal{X} = E^I$, where I is a finite set
- $\mathcal{X}$-valued X with law μ

Examples: Glauber dynamics.

- $\mathcal{X} = E^I$, where I is a finite set
- $\mathcal{X}$-valued X with law μ
- for $i \in I$,

$$\mu_i(\cdot|x) := \mathbb{P}(X_i \in \cdot \,|X_{\{i\}^c} = x_{\{i\}^c})$$

Examples: Glauber dynamics.

- $\mathcal{X} = E^I$, where I is a finite set
- $\mathcal{X}$-valued X with law μ
- for $i \in I$,

$$\mu_i(\cdot|x) := \mathbb{P}(X_i \in \cdot \, | X_{\{i\}^c} = x_{\{i\}^c})$$

- The Glauber dynamic is given by generator

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

- 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.

Corollary

Let X_i', $i \in I$ be a r.v. such that

$$\mathbb{P}(X_i' \in \cdot \,|X = x) = \mu_i(\cdot|x)$$

and $X^i = (X_1, \dots, X_{i-1}, X_i', X_{i+1}, \dots, X_n)$. If the Glauber dynamics satisfies mLSI(K) then for any $f\colon \mathcal{X} \to \mathbb{R}$ and $p \geqslant 2$,

$$\begin{aligned}
&\|f(X) - \mathbb{E}f(X)\|_r \\
&\qquad \leqslant C\sqrt{Kr}\Big\|\Big(\sum_{i\in I}\mathbb{E}((f(X) - f(X^i))^2|X)\Big)^{1/2}\Big\|_r \\
&\qquad\qquad \leqslant C\sqrt{Kr}\Big\|\Big(\sum_{i\in I}(f(X) - f(X^i))^2\Big)^{1/2}\Big\|_r,
\end{aligned}$$

Corollary

Let X_i', $i \in I$ be a r.v. such that

$$\mathbb{P}(X_i' \in \cdot \,|X = x) = \mu_i(\cdot|x)$$

and $X^i = (X_1, \ldots, X_{i-1}, X_i', X_{i+1}, \ldots, X_n)$. If the Glauber dynamics satisfies mLSI(K) then for any $f\colon \mathcal{X} \to \mathbb{R}$ and $p \geqslant 2$,

$$\begin{aligned}\|f(X) - \mathbb{E}f(X)\|_r \\ &\leqslant C\sqrt{Kr}\Big\|\Big(\sum_{i\in I}\mathbb{E}((f(X) - f(X^i))^2|X)\Big)^{1/2}\Big\|_r \\ &\leqslant C\sqrt{Kr}\Big\|\Big(\sum_{i\in I}(f(X) - f(X^i))^2\Big)^{1/2}\Big\|_r,\end{aligned}$$

Remark: Recovers (up to constants) the BBLM result since mLSI holds for all product measures.

Tetrahedral polynomials

Recall: A multivariate polynomial is tetrahedral if it is affine in each variable.

Tetrahedral polynomials

Recall: A multivariate polynomial is tetrahedral if it is affine in each variable.

Corollary (Kotowski-Polaczyk-Strzelecki-A. 2019, 2020)

Let X be a random vector with values in $[-1,1]^n$ satisfying mlSI(K). Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a tetrahedral polynomial of degree d. Then for any $t > 0$,

$$\mathbb{P}\Big(\Big|f(X) - \mathbb{E}f(X)\Big| \geqslant t\Big) \leqslant 2\exp\Big(-\frac{1}{C_d}\min_{1\leqslant k\leqslant d}\min_{\mathcal{J}\in P_k}\Big(\frac{t}{K^{k/2}\|\mathbb{E}\nabla^k f(X)\|_{\mathcal{J}}}\Big)^{2/|\mathcal{J}|}\Big).$$

Tetrahedral polynomials

Recall: A multivariate polynomial is tetrahedral if it is affine in each variable.

Corollary (Kotowski-Polaczyk-Strzelecki-A. 2019, 2020)

Let X be a random vector with values in $[-1,1]^n$ satisfying mlSI(K). Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a tetrahedral polynomial of degree d. Then for any $t > 0$,

$$\mathbb{P}\Big(\Big|f(X) - \mathbb{E}f(X)\Big| \geqslant t\Big) \\ \leqslant 2\exp\Big(-\frac{1}{C_d}\min_{1\leqslant k\leqslant d}\min_{\mathcal{J}\in P_k}\Big(\frac{t}{K^{k/2}\|\mathbb{E}\nabla^k f(X)\|_{\mathcal{J}}}\Big)^{2/|\mathcal{J}|}\Big).$$

Remark If $\mathcal{X} \subseteq \{-1,0,1\}^n$ then every polynomial can be written in a tetrahedral form.

Theorem (Marton, Götze-Sambale-Sinulis)

Assume that $\mathcal{X}$ is finite and μ has full support. Define

$$\begin{aligned} A_{ij} \\ = \sup_{x_{\{j\}^c}=y_{\{j\}^c}} \|\mathcal{L}(X_i|X_{\{i\}^c}=x_{\{i\}^c}) - \mathcal{L}(X_i|X_{\{i\}^c}=y_{\{i\}^c})\|_{TV} \end{aligned}$$

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 mLSI(K) holds with $K \leqslant \alpha^{-2}\beta^{-1}$.

Ising model on finite graphs

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

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)

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)

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)

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

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)
- $$\mu(\{\varepsilon\}) = \frac{1}{Z} \exp\Big(\frac{1}{2} \sum_{i,j=1}^n J_{ij}\varepsilon_i\varepsilon_j - \sum_{i=1}^n h_i\varepsilon_i\Big),$$
- Götze-Sambale-Sinulis:
$$\alpha \geqslant 1 - \max_{i \leqslant n} \sum_{j \leqslant n} |J_{ij}|, \quad \beta \geqslant ce^{-\|h\|_\infty}.$$

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)
- $$\mu(\{\varepsilon\}) = \frac{1}{Z} \exp\Big(\frac{1}{2} \sum_{i,j=1}^n J_{ij}\varepsilon_i\varepsilon_j - \sum_{i=1}^n h_i\varepsilon_i\Big),$$
- Götze-Sambale-Sinulis:
$$\alpha \geqslant 1 - \max_{i \leqslant n} \sum_{j \leqslant n} |J_{ij}|, \quad \beta \geqslant ce^{-\|h\|_\infty}.$$
- thus we get inequalities for polynomials: applications to testing Ising models.

Triangle counts in exponential random graphs

- $I_n = \{(i, j) \in [n]^2 \colon i < j\}$

Triangle counts in exponential random graphs

- $I_n = \{(i,j) \in [n]^2 \colon i < j\}$
- $\mathcal{X}_n = \{0,1\}^{I_n}$ – simple graphs on n-vertices

Triangle counts in exponential random graphs

- $I_n = \{(i, j) \in [n]^2 \colon i < j\}$
- $\mathcal{X}_n = \{0, 1\}^{I_n}$ – simple graphs on n-vertices
- $\gamma = (\gamma_1, \ldots, \gamma_s) \in \mathbb{R}^s$

Triangle counts in exponential random graphs

- $I_n = \{(i, j) \in [n]^2 \colon i < j\}$
- $\mathcal{X}_n = \{0, 1\}^{I_n}$ – simple graphs on n-vertices
- $\gamma = (\gamma_1, \dots, \gamma_s) \in \mathbb{R}^s$
- $G_i = (V_i, E_i)$, $i = 1, \dots, s$ – simple connected graphs (convention $G_1 = K_2$)

Triangle counts in exponential random graphs

- $I_n = \{(i,j) \in [n]^2 \colon i < j\}$
- $\mathcal{X}_n = \{0,1\}^{I_n}$ – simple graphs on n-vertices
- $\gamma = (\gamma_1, \dots, \gamma_s) \in \mathbb{R}^s$
- $G_i = (V_i, E_i)$, $i = 1, \dots, s$ – simple connected graphs (convention $G_1 = K_2$)
- $\mu \sim e^{-H_\gamma}$ with

$$H_\gamma(x) = n^2 \sum_{i=1}^{s} \gamma_i \frac{N_{G_i}(x)}{n^{|V_i|}},$$

where $N_H(G)$ is the number of copies of H in G.

Triangle counts in exponential random graphs

- $I_n = \{(i,j) \in [n]^2 \colon i < j\}$
- $\mathcal{X}_n = \{0,1\}^{I_n}$ – simple graphs on n-vertices
- $\gamma = (\gamma_1, \ldots, \gamma_s) \in \mathbb{R}^s$
- $G_i = (V_i, E_i)$, $i = 1, \ldots, s$ – simple connected graphs (convention $G_1 = K_2$)
- $\mu \sim e^{-H_\gamma}$ with

$$H_\gamma(x) = n^2 \sum_{i=1}^{s} \gamma_i \frac{N_{G_i}(x)}{n^{|V_i|}},$$

  where $N_H(G)$ is the number of copies of H in G.
- Sambale-Sinulis: if

$$\delta := \frac{1}{2} \sum_{i=2}^{s} |\gamma_i||E_i|(|E_i| - 1) < 1,$$

  then $\alpha \geqslant 1 - \delta$ and $\beta \geqslant ce^{-2|\gamma_1|}$.

Corollary (Sambale–Sinulis, Polaczyk–Strzelecki–A.)

Let T be the number of triangles in an exponential random graph with $\delta < 1$. Set $K = (1-\delta)^{-2}e^{2|\gamma_1|}$. Then for $t > 0$,

$$\mathbb{P}(|T - \mathbb{E}T| \geqslant t) \leqslant$$

$$2\exp\Big(-\frac{1}{C}\frac{t^2}{n^3(K^3 + K^2A^2) + n^4KB^2} \wedge \frac{t}{\sqrt{n}K^{3/2} + nKA} \wedge \frac{t^{2/3}}{K}\Big)\Big),$$

where A is a probability of finding an edge and B is a probability of finding a *cherry* at a fixed place in a graph.

Corollary (Sambale–Sinulis, Polaczyk–Strzelecki–A.)

Let T be the number of triangles in an exponential random graph with $\delta < 1$. Set $K = (1-\delta)^{-2}e^{2|\gamma_1|}$. Then for $t > 0$,

$$\mathbb{P}(|T - \mathbb{E}T| \geqslant t) \leqslant$$
$$2\exp\Big(-\frac{1}{C}\frac{t^2}{n^3(K^3 + K^2A^2) + n^4KB^2} \wedge \frac{t}{\sqrt{n}K^{3/2} + nKA} \wedge \frac{t^{2/3}}{K}\Big)\Big),$$

where A is a probability of finding an edge and B is a probability of finding a *cherry* at a fixed place in a graph.

Remark: The subgraph-count problem in the Erdős–Rényi case has long history: Kim–Vu, Janson–Ruciński–Oleszkiewicz, Chatterjee, Wolff-A., DeMarco–Kahn, Chatterjee–Dembo, Lubetzky–Zhao, Šileikis–Warnke

The Poisson space

- η – a Poisson process on $\mathcal{X}$ with a σ-finite intensity λ.
- Add-one and remove-one gradients

$$D_x^+ f(\eta) = f(\eta + \delta_x) - f(\eta), \; D_x^- f(\eta) = (f(\eta) - f(\eta - \delta_x)) \mathbb{1}_{\{x \in \eta\}}$$

- The Dirichlet form is related to the add-one gradient, i.e.,

$$\mathcal{E}(f, g) = \mathbb{E} \int_{\mathcal{X}} (D_x^+ f)(D_x^+ g) \lambda(dx)$$

- Wu proved mLSI(1)

Translation into our setting by Mecke's formula:

$$\Gamma(f) = \frac{1}{2}\int_{\mathcal{X}} (D_x^- f(\eta))^2 \eta(dx) + \frac{1}{2}\int_{\mathcal{X}} (D_x^+ f(\eta))^2 \lambda(dx)$$

and

$$\Gamma_+(f) \quad = \quad \frac{1}{2}\int_{\mathcal{X}} (D_x^- f(\eta))_+^2 \eta(dx) \; + \; \frac{1}{2}\int_{\mathcal{X}} (D_x^+ f(\eta))_-^2 \lambda(dx)$$

Corollary (Polaczyk-Strzelecki-A., 2020)

For $r \geqslant 2$,

$$\|f - \mathbb{E}f\|_r \leqslant C\sqrt{r}\left\|\sqrt{\Gamma(f)}\right\|_r$$

and

$$\|(f - \mathbb{E}F)_+\|_r \leqslant C\sqrt{r}\left\|\sqrt{\Gamma_+(f)}\right\|_r.$$

- Obtained independently by Sambale–Sinulis–Thäle

- Obtained independently by Sambale–Sinulis–Thäle
- Applies, e.g., to Poisson U-statistics (e.g., subgraph counts in geometric random graphs) or suprema of Poisson integrals,

- Obtained independently by Sambale–Sinulis–Thäle
- Applies, e.g., to Poisson U-statistics (e.g., subgraph counts in geometric random graphs) or suprema of Poisson integrals,
- Many concentration inequalities for the Poisson space were obtained by Bachmann–Peccati–Reitzner, Reynaud-Bouret, and other authors.

- Obtained independently by Sambale–Sinulis–Thäle
- Applies, e.g., to Poisson U-statistics (e.g., subgraph counts in geometric random graphs) or suprema of Poisson integrals,
- Many concentration inequalities for the Poisson space were obtained by Bachmann–Peccati–Reitzner, Reynaud-Bouret, and other authors.
- For geometric examples we have, inequalities of similar strength follow from their approach, based directly on mLSI and some self-bounding properties.

- Obtained independently by Sambale–Sinulis–Thäle
- Applies, e.g., to Poisson U-statistics (e.g., subgraph counts in geometric random graphs) or suprema of Poisson integrals,
- Many concentration inequalities for the Poisson space were obtained by Bachmann–Peccati–Reitzner, Reynaud-Bouret, and other authors.
- For geometric examples we have, inequalities of similar strength follow from their approach, based directly on mLSI and some self-bounding properties.
- Question: can one find natural geometric examples with heavier tails in which one is forced to work with moments?

Thank you