Entropy-Entropy Production Inequalities

Chaoen ZHANG

Harbin Institute of Technology

The 17th Workshop on Markov Processes and Related Topics
Beijing Normal University, China.

Joint with Arnaud Guillin, Wei Liu and Liming Wu.

Outline

The kinetic theory of gases

- Gas: made of a large number of molecules colliding with each other, moving in the space.
- Maxwell and Boltzmann: describe the gas by a density function on the phase space $\mathbb{R}^d_x \times \mathbb{R}^d_v$

  $f(t,x,v)$= the probability of particles at position x and with velocity v

- The evolution: the Boltzmann equation.
- Rigorous derivation of the equation: Boltzmann-Grad limit.
- Well-posedness and regularity of solutions.
- The entropy production phenomenon: Boltzmann's H theorem.

Boltzmann's H theorem

For simplicity: assume $x \in \mathbb{T}^d$ and $V(x) = 1$ in this slide.

Theorem (Boltzmann's H theorem)

Suppose the collision kernel $B > 0$ a.e.. Assume $f = (f_t)_{t \geq 0}$ is a "nice" probability density solution of the Boltzmann equation, then Boltzmann's H functional

$$H(f) = \int f \log f \, dx dv$$

is non-increasing in time. Indeed, formally

$$\frac{d}{dt} H(f) \leq 0$$

Some mathematical problems

- Derive the equations from N interacting-particle systems:

  Boltzmann-Grad limit, mean-field limit, propagation of chaos,...

- relations between entropies and entropy production functionals:

  entropy-entropy production inequalities;

- convergence to equilibrium with constructive/realistic rates:

  transport, confinement, self-consistent, collision/diffusion.

What's next ☞

1. Motivation by an inspiring example: beyond Boltzmann's H theorem
2. More entropy producing models
3. EEP for the McKean-Vlasov equation

Boltzmann's H functional

- Clausius introduced the concept of entropy.
- Boltzmann's H functional: a statistical definition of entropy

$$H(f) = \int f \log f,$$

- The Boltzmann equation for a monatomic rarefied gas:

$$\partial_t f + \underbrace{v \cdot \nabla_x f}_{\text{transport}} - \underbrace{\nabla_x V(x) \cdot \nabla_v f}_{\text{confinement}} = \underbrace{Q(f,f)}_{\text{binary collisions}}, \quad t \geq 0$$

where

- the Boltzmann collision operator ("nonlinear jump process"?):

$$Q(f,f) = \int_{\mathbb{R}^d} \int_{S^{d-1}} (f(v')f(v'_*) - f(v)f(v_*))B(v - v_*, \sigma)\mathrm{d}\sigma \mathrm{d}v_*;$$

- (v, v_*): the velocities before(or after) collision
- (v', v'_*): the velocities after(or before) a collision

Boltzmann's H theorem

A mathematical manifestation of the second law of thermodynamics:

Theorem (Boltzmann's H theorem 1872')

Suppose the collision kernel $B = B(v - v_*, \sigma) > 0$ a.e.. Assume $f = (f_t)_{t \geq 0}$ is a "nice" probability density solution of the Boltzmann equation, then Boltzmann's H functional

$$H(f) = \int f \log f \, dx dv$$

is non-increasing in time. Indeed, at least formally

$$\frac{d}{dt} H(f) \leq 0.$$

Entropy production functional

Denote $f = f(v), f_* = f(v_*), f' = f(v'), f'_* = f(v'_*)$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}H(f) = -\frac{1}{4}\int (ff_* - f'f'_*)(\log ff_* - \log f'f'_*)B(v - v_*, \sigma)\mathrm{d}\sigma \mathrm{d}v_* \mathrm{d}v \mathrm{d}x \leq 0$$

since $(r - s)(\log r - \log s) \geq 0$ for $r, s > 0$.

- Boltzmann entropy production functional:

$$D_B(f) = \frac{1}{4}\int (ff_* - f'f'_*)(\log ff_* - \log f'f'_*)B(v - v_*, \sigma)\mathrm{d}\sigma \mathrm{d}v_* \mathrm{d}v.$$

- Identify the equilibrium: $D_B(f) = 0 \Rightarrow$

$$f(v)f(v_*) = f(v')f(v'_*) \rightsquigarrow \text{ gaussian functions.}$$

- Denote the equilibrium by f_∞.

Motivations for quantitative refinement

- The entropy production functional is nonnegative:

$$D_B(f) \geq 0$$

with equality if and only if f is a Gaussian.

- The density f_t is expected to converge to the equilibrium f_∞.

Question : quantitative convergence to equilibrium??

Motivations:

(A) It is expected that $f_t \to f_\infty$ very rapidly.

(B) Boltzmann's response to Zermelo's paradox.

(C) Understanding the entropy production mechanism (applications in various problems).

(D) This is a natural mathematical question...

Motivation (B1)

Zermelo's paradox: Contradiction between Botlzmann's H theorem and Poincaré recurrence theorem.

- The gas is modeled by a large number of particles moving and colliding according to Netwonian mechanics.
- Poincaré's recurrence theorem: for a Hamiltonian system

  the system will return arbitrarily close to the initial state.

- Botlzmann's H theorem $\Rightarrow$

  solutions of the Boltzmann equation will tend to the equilibirium.

- Zermelo's conclusion: Boltzmann's H theorem is physically irrelevant!

(see also in an earlier note of Poincaré)

Motivation (B2)

Boltzmann's response to Zermelo's paradox:

- The recurrence time would be too huge.
- Estimates by Boltzmann: huge even if the estimated age of the universe is taken as the time unit.
- The accuracy of Boltzmann's model describing the gas breaks down on very large time scales.

??The physical relevance of Boltzmann's H theorem??

Prove the convergence to equilibrium in a short time scale.

(i.e. Prove "H theorem" that is not only quantitative but also physically realistic.)

Beyond Boltzmann's H theorem: functional inequality?

- Cercignani's conjecture(1982'): entropy-entropy production inequality??

$$D_B(f)?? \geq \lambda \left(\int f \log f \mathrm{d}v - \int f_\infty \log f_\infty \mathrm{d}v \right) := \lambda H(f|f_\infty)$$

for f with unit mass, zero mean velocity and unit temperature, i.e.

$$f \in \mathcal{C}_{1,0,1} := \left\{ \int f(v)\mathrm{d}v = 1, \int vf(v)\mathrm{d}v = 0, \int |v|^2 f(v)\mathrm{d}v = d. \right\}$$

- Bobylev,Cercignani 1999': for a large class of collision kernels, NO such a inequality holds even for a very restricted class of functions.
- Related works: Caflisch 1980', Bobylev 1984'& 1988', Wennberg 1997'...

What if Cercignani's conjecture was true?

Consider solutions f_t to the spatially homogeneous Boltzmann equation

$$\partial_t f_t = Q(f_t, f_t),$$

then we would have

$$-\frac{\mathrm{d}}{\mathrm{d}t} H(f_t|f_\infty) = D_B(f_t) \geq \lambda H(f_t|f_\infty)$$

which would imply

$$H(f_t|f_\infty) \leq e^{-\lambda t} H(f_0|f_\infty).$$

Define the time $T(\varepsilon)$ for $\varepsilon \in (0,1)$

$$T(\varepsilon) := \inf \left\{ t > 0 : H(f_t|f_\infty) \leq \varepsilon H(f_0|f_\infty), \forall f_0 \text{ with finite entropy} \right\},$$

then it would yield

$$T(\varepsilon) \leq \frac{-\log \varepsilon}{\lambda}.$$

“Cercignani’s conjecture is sometimes ture”

$$D_B(f) = \frac{1}{4}\int (ff_* - f'f'_*)(\log ff_* - \log f'f'_*)B(v - v_*, \sigma)\mathrm{d}\sigma\mathrm{d}v_*\mathrm{d}v$$

Theorem (Villani 2003’)

Let the collision kernel B satisfy (“super hard sphere”)

$$B(v - v_*, \sigma) \geq K_B(1 + |v - v_*|^2).$$

Then for $f \in \mathcal{C}_{1,0,1}$, $f_\infty = (2\pi)^{-d/2}\exp(-\frac{1}{2}|v|^2)$,

$$D_B(f) \geq \lambda_B H(f|f_\infty).$$

where $\lambda_B > 0$ depends on an upper bound for $H(f)$.

Related works: Carlen-Carvalho 1992’, 1994’; Toscani-Villani 1999’.
Convergence to equilibrium: Desvillettes-Villani 2005’.

What's next ☞

1. Motivation by an inspiring example: beyond Boltzmann's H theorem
2. More entropy producing models
3. EEP for the McKean-Vlasov equation

The Ornstein-Uhlenck process: LSI

$$\mathrm{d}X_t = \sqrt{2}\mathrm{d}B_t - X_t\mathrm{d}t$$

- The invariant measure: γ the standard Gaussian measure.
- h: The density function of X_t w.r.t γ.
- The entropy:

$$\mathrm{Ent}_\gamma(h) := \int h \log h \mathrm{d}\gamma;$$

- The entropy production functional = the relative Fisher information:

$$D_{OU}(h) = \int \frac{|\nabla h|^2}{h} \mathrm{d}\gamma.$$

Theorem (Gross's Gaussian logarithmic Sobolev inequality)

$$\int \frac{|\nabla h|^2}{h} d\gamma \geq 2 \int h \log h d\gamma.$$

The Langevin diffusion (1): Basic properties

The Langevin diffusion:

$$\begin{cases} \mathrm{d}x_t = v_t \mathrm{d}t \\ \mathrm{d}v_t = \sqrt{2}\mathrm{d}B_t - v_t \mathrm{d}t - \nabla_x V(x_t)\mathrm{d}t. \end{cases}$$

The kinetic Fokker-Planck equation:

$$\partial_t h + v \cdot \nabla_x h - \nabla_x V(x) \cdot \nabla_v h = \Delta_v h - v \cdot \nabla_v h$$

The invariant measure:

$$\mu(\mathrm{d}x, \mathrm{d}v) = \frac{1}{Z} e^{-V(x)} \cdot (2\pi)^{-\frac{Nd}{2}} e^{-\frac{|v|^2}{2}} \mathrm{d}x\mathrm{d}v := \mathrm{d}\nu(x)\mathrm{d}\gamma(v).$$

The entropy:

$$\mathrm{Ent}_\mu(h) := \int h \log h \mathrm{d}\mu$$

The Langevin diffusion (2): Entropic convergence

Theorem (Villani)

Assume that

1. *the potential* $V \in C^2(\mathbb{R}^d)$ *with* $|\nabla^2 V| \leq K$*;*
2. *the reference measure* μ *satisfies a logarithmic Sobolev inequality;*

Then there exist constant $C > 0$ *and* $\lambda > 0$*, explicitly computable, such that*

$$\int h_t \log h_t d\mu \leq Ce^{-\lambda t} \int h_0 \log h_0 d\mu.$$

Further results on entropic decay:

- Baudoin'17: local Γ calculus;
- Cattiaux-Guillin-Monmarché-Z.'19: relax the condition (1).

Related works: F.-Y. Wang, J. Wang, L.-M. Wu, X.-C. Zhang, ...

The Landau equation:

$$\partial_t f + v \cdot \nabla_x f - F(x) \cdot \nabla_v f = Q_L(f, f), \quad t \geq 0$$

The Landau entropy production functional:

$$D_L(f) = \frac{1}{2} \int ff_* \Psi(|v - v_*|) \left| \mathrm{Proj}_{(v-v_*)^\perp} \left(\nabla \log f - (\nabla \log f)_* \right) \right|^2 \mathrm{d}v_* \mathrm{d}v$$

Theorem (Desvillettes-Villani2001')

If $\Psi(|z|) \geq |z|^2$, *then*

$$D_L(f) \geq \lambda(f) H(f|f_\infty).$$

Continuous time Markov chains: MLSI

Let (K, π) be a irreducible reversible Markov chain on a finite state space.

- The Dirichlet form:

$$\mathcal{E}(f, g) = \langle (I - K)f, g \rangle$$

- Poincaré inequality:

$$\lambda_1 \mathrm{Var}_\pi(f) \le \mathcal{E}(f, f).$$

- Log Sobolev inequality:

$$\rho \mathrm{Ent}_\pi(f) \le 2\mathcal{E}(\sqrt{f}, \sqrt{f}).$$

- Modified log Sobolev inequality: (entropy-entropy production inequality)

$$\rho_0 \mathrm{Ent}_\pi(f) \le \frac{1}{2}\mathcal{E}(f, \log f).$$

Summary: Entropy producing models

- The Boltzmann equation;
- The Landau equaiton;
- The kinetic Fokker-Planck equation (the Langevin equation);
- The Ornstein-Uhlenbeck process (the log Sobolev inequalities);
- Poisson point processes;
- Random transpositions, Bernoulli-Laplace model;
- The Kac model;
- Zero range processes;
- Swendsen-Wang dynamics;
- etc.

What's next ☞

1. Motivation by an inspiring example: beyond Boltzmann's H theorem
2. More entropy producing models
3. EEP for the McKean-Vlasov equation

The McKean-Vlasov equation

Question: Entropy-entropy production inequality for

$$\partial_t f = \Delta f + \nabla \cdot (f(\nabla V + \nabla W * f))$$

and the corresponding self-interacting diffusion $(X_t)_{t\geq 0}$ on $\mathbb{R}^d$:

$$\mathrm{d}X_t = \sqrt{2}\mathrm{d}B_t - \nabla V(X_t)\mathrm{d}t - \nabla W * \mathrm{law}(X_t)\mathrm{d}t,$$

where

- the confinement potential $V : \mathbb{R}^d \to \mathbb{R}$;
- the interaction potential $W : \mathbb{R}^d \to \mathbb{R}$ is even(symmetric interaction).

Convergence to equilibrium: Carrillo-McCann-Villani, Malrieu, Cattiaux-Guillin-Malrieu, Bolley-Gentil-Guillin, Eberle-Guillin-Zimmer, Guillin-Liu-Wu-Z., Liu-Wu-Z., Ren-Wang, Wang,....

Entropy and entropy production

- It is the gradient flow of the free energy

$$\mathrm{E}(f) := \int f \log f \mathrm{d}x + \int V f \mathrm{d}x + \frac{1}{2} \int W(x-y) f(x) f(y) \mathrm{d}x \mathrm{d}y$$

in the space of probability measures with the Wasserstein metric.

- We shall always assume $\mathrm{E}(f)$ admits a unique minimizer f_∞ (equilibrium) with finite free energy. Denote the relative free energy by

$$\mathrm{E}(f|f_\infty) := \mathrm{E}(f) - \mathrm{E}(f_\infty).$$

- The entropy production functional

$$D_{MV}(f) := \int |\nabla \log f + \nabla V + \nabla W * f|^2 f \mathrm{d}x.$$

- EEP inequality:

$$D_{MV}(f) \geq \lambda \mathrm{E}(f|f_\infty).$$

EEP by Carrillo-McCann-Villani'03

Theorem (Carrillo-McCann-Villani, theorem 2.1)

Assume that V is uniformly convex and

$$\nabla^2 V \geq \rho \mathrm{Id} > 0, \quad \rho > ||(\nabla^2 W)^-||_{L^\infty},$$

where $(\nabla^2 W)^-$ is the negative part of the Hessian $\nabla^2 W$. Let

$$\lambda = \rho - ||(\nabla^2 W)^-||_{L^\infty} > 0.$$

Then

1. *Existence and uniqueness of minimizer f_∞ of the free energy.*
2. *EEP inequality:*

$$\int |\nabla \log f + \nabla V + \nabla W * f|^2 f \, dx \geq 2\lambda(E(f) - E(f_\infty)).$$

There are results for the "degenerately convex interaction" as well, for

Non-convex settings

Remark: convexity of $V(x)$ is assumed in their EEP inequalities... Unlike the log Sobolev inequality, perturbation argument doesn't work well. But in many cases, linear EEP inequalities are still expected!

Question: How to prove EEP inequality in non-convex settings?

An motivating example:

$$V(x) = \beta(\frac{|x|^4}{4} - \frac{|x|^2}{2}), \quad W(x) = -\frac{\beta K}{2}|x|^2.$$

or more generally:

- $\nabla^2 W \geq -K_0 \mathrm{Id}$,
- V is "super-convex": $\nabla^2 V \geq K(|x|)\mathrm{Id}$, with

$$\lim_{R \to +\infty} K(R) = +\infty.$$

The strategy: Using interacting particle systems

- Consider the corresponding particle system with mean field interaction, $1 \le i \le N$,

$$\mathrm{d}X_i = \sqrt{2}\mathrm{d}B_t^i - \left[\nabla V(X_i) + \frac{1}{N-1}\sum_{1\le j\le N} \nabla W(X_i - X_j)\right]\mathrm{d}t$$

- and its Gibbs measure $m(\mathrm{d}x_1 \cdots \mathrm{d}x_N)$

$$m(\mathrm{d}x_1 \cdots \mathrm{d}x_N) = e^{-H_N}\mathrm{d}x_1 \cdots \mathrm{d}x_N$$

where H_N is the Hamiltonian

$$H_N(x_1, \cdots, x_N) = \sum_{1\le i\le N} V(x_i) + \frac{1}{N-1}\sum_{1\le i<j\le N} W(x_i - x_j).$$

The strategy: Using interacting particle systems+1

Under technical assumptions, as the number of particles $N \to +\infty$,

1. The mean relative entropy tends to the relative free energy (Liu-Wu'20)

$$\frac{1}{N} H(f^{\otimes N}|m) \to \mathrm{E}(f|f_\infty);$$

2. The mean Fisher information tends to $D_{MV}(f)$

$$\frac{1}{N} I(f^{\otimes N}|m) \to \int |\nabla \log f + \nabla V + \nabla W * f|^2 f \mathrm{d}x;$$

3. Then EEP inequality can be deduced from uniform LSI for the Gibbs measure m:

$$H(F|m) \leq \frac{1}{\lambda} I(F|m).$$

by taking $F = f^{\otimes N}$.

Zegarlinski's theorem for the Gibbs measure

$$\mathrm{d}m(x_1, \cdots, x_N) = \exp\left\{ \sum_{i=1}^{N} V(x_i) + \frac{1}{N-1} \sum_{1\le i<j\le N} W(x_i - x_j) \right\} \mathrm{d}x_1 \cdots \mathrm{d}x_N$$

Notations: the conditional measure knowing $(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N)$,

$$\mathrm{d}m_i(x_i) = \frac{1}{Z_i} \exp\left\{ -V(x_i) - \frac{1}{N-1} \sum_{j:\ j\neq i} W(x_i, x_j) \right\} \mathrm{d}x_i.$$

Theorem (Zegarlinski'92)

(Z1) *uniform LSI for all m_i's*

(Z2) *Zegarlinski's condition on interdependence*

implies LSI for the Gibbs measure.

Beyond convexity: Lipschitzian spectral gap condition

Define for $r > 0$

$$b(r) := \sup - \left\langle \frac{x-y}{|x-y|}, (\nabla V(x) - \nabla V(y)) + (\nabla_x W(x-z) - \nabla_x W(y-z)) \right.$$

where the supremum runs over $x, y, z \in \mathbb{R}^d$ with $|x-y| = r$.

Assumption (L1)

Suppose that the following Lipschitzian constant

$$c_{Lip} := \frac{1}{4}\int_0^\infty \exp\left\{\frac{1}{4}\int_0^s b(u)\mathrm{d}u\right\} s\mathrm{d}s < +\infty$$

Lemma (Wu'09)

Suppose (L1), then the conditional measure $m_i(dx_i)$ satisfies a Poincaré inequality with uniform constant c_{Lip}.

By this we are able to verify Zegarlinski's condition (Z2).

Log-Sobolev inequalities uniform in N

Theorem (Guillin-Liu-Wu-Z. '22)

Assume (L1) and

1. *for some constant* $\rho_{\mathrm{LS}} > 0$*, the conditional measures* m_i *on* $\mathbb{R}^d$ *satisfy the log-Sobolev inequality :*

$$2\rho_{\mathrm{LS}} H(f|m_i) \le I(f|m_i), \quad f \in C_b^1(\mathbb{R}^d)$$

   for all i *and* $(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N)$ *;*

2. *(Dobrushin-Zegarlinski condition)*

$$\gamma_0 := c_{Lip,m} \sup_{x,y\in\mathbb{R}^d, |z|=1} |\nabla^2_{x,y} W(x,y)z| < 1.$$

then m *satisfies*

$$2\rho_{\mathrm{LS}}(1-\gamma_0)^2 H(f|m) \le I(f|m), \quad f \in C_b^1(\mathbb{R}^{dN}).$$

The strategy recalled

Under technical assumptions, as the number of particles $N \to +\infty$,

1. The mean relative entropy tends to the relative free energy (Liu-Wu'20)

$$\frac{1}{N} H(f^{\otimes N}|m) \to \mathrm{E}(f|f_\infty);$$

2. The mean Fisher information tends to $D_{MV}(f)$

$$\frac{1}{N} I(f^{\otimes N}|m) \to \int |\nabla \log f + \nabla V + \nabla W * f|^2 f \mathrm{d}x;$$

3. Then EEP inequality is deduced from uniform LSI for the Gibbs measure m_N:

$$H(F|m) \leq \frac{1}{\lambda} I(F|m_N).$$

by taking $F = f^{\otimes N}$.

The motivating example recalled

Question: How to prove EEP inequality in non-convex settings?
An motivating example:

$$V(x) = \beta(\frac{|x|^4}{4} - \frac{|x|^2}{2}), \; W(x) = -\frac{\beta K}{2}|x|^2.$$

or more generally:

- $\nabla^2 W \geq -K_0 \mathrm{Id}$,
- V is "super-convex": $\nabla^2 V \geq K(|x|)\mathrm{Id}$, with

$$\lim_{R \to +\infty} K(R) = +\infty.$$

Thank you for your attention!!