Applications of the time derivative of the L^2-Wasserstein distance and the free entropy dissipation

Hiroaki YOSHIDA

Ochanomizu University
Tokyo, Japan

at Fields Institute   23 July, 2013

Plan of talk

1. Introduction and some definitions
2. Time derivative of the L^2-Wasserstein distance
3. Dissipations of the relative free entropy
4. Semicircular perturbations and the relative free entropy

Introduction and some definitions

Topic 1

Let μ and ν be two compactly supported probability mesures on $\mathbb{R}$.

We consider the time evolutions μ_t and ν_t ($\mu_0 = \mu$, $\nu_0 = \nu$) by the free Fokker-Planck equation.

We calculate the time derivative

$$\frac{d}{dt}\left(W_2(\mu_t, \nu_t)\right)^2$$

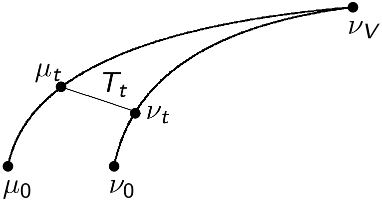
with Brenier map (the optimal transport map).

We will see

the convergence and uniqueness of the long time asymptotic stationary measure ν_V of the free Fokker-Planck equation,

and show

some inequalities between the L^2 Wasseretein distance and the relative free Fisher information.



Topic 2

We consider the time evolution μ_t $(\mu_0 = \mu)$ by the free Fokker-Planck equation.

We calculate the time derivatives (free entropy dissipations)

$$\frac{d}{dt}\Sigma(\mu_t \,|\, \nu_V) \quad \text{and} \quad \frac{d}{dt}\Phi(\mu_t \,|\, \nu_V),$$

where the reference measure ν_V is stationary.

We will see

the differential relation between the relative free entropy and the relative free Fisher information and convergence of μ_t to the stationary measure in the relative free entropy,

and show

the free logarithmic Sobolev and the free transpotation cost inequalities by time integration.

The results in the first two topics are not new, but the proofs are a little different at time integration.

Topic 3

We consider the time evolutions μ_t and ν_t by *semicircular purterbations*, which corresponds to the case of the free Fokker-Planck equation with null potential.

We calculate the time derivative of the relative free entropy

$$\frac{d}{dt}\Sigma(\mu_t \,|\, \nu_t),$$

where the reference measure ν_t is moving!

We will see

the differential relation between the relative free entropy and the relative free Fisher information,

and recover

in case of the reference measure being semicircular, the formula for the micro states free approach to the free entropy of Voiculescu by time integration.

Let $V : \mathbb{R} \to \mathbb{R}$ be an analytic function, and consider the stochastic differential equation on $N \times N$ Hermitian matrices $\mathcal{H}_N$

$$d\mathbf{X}_t = \frac{1}{\sqrt{N}} d\mathbf{B}_t - \frac{1}{2} \nabla \mathrm{Tr} V(\mathbf{X}_t) dt,$$

that is, $\mathcal{H}_N$-valued diffusion process, where $\mathbf{B}_t$ is the standard Brownian motion on $\mathcal{H}_N$ and ∇ denotes the gradient operator.

Note that $\nabla \mathrm{Tr} V(\mathbf{X}_t)$ can be written as $V'(\mathbf{X}_t)$.

Eigenvalues of the $\mathbf{X}_t$

The eigenvalues $\big(\lambda_1(t), \lambda_2(t), \dots, \lambda_N(t)\big)$ of the matrix $\mathbf{X}_t$ satisfy the SDE

$$d\lambda_i(t) = \frac{1}{\sqrt{N}} dW_i(t) + \frac{1}{N} \sum_{j:j\neq i} \frac{1}{\lambda_i(t) - \lambda_j(t)}\, dt - \frac{1}{2} V'(\lambda_i(t))\, dt$$

$$i = i, 2, \dots, N.$$

where the $W_i(t)$ are independent one dimensional Brownian motions.

The process $\boldsymbol{\lambda}_N(t) = \big(\lambda_1(t), \lambda_2(t), \dots, \lambda_N(t)\big)$ is called *the generalized Dyson Brownian motion with potential* V in physists.

This process can be modeled as an intracting N-particles (electrons) system with the logarithmic Coulomb interaction and the external potential V, and hence, the Hamiltonian of the system is given of the form

$$H(x_1, x_2, \ldots, x_N) = -\frac{1}{2}\left(\frac{1}{N}\sum_{N}^{i=1}\sum_{i\neq j}\log\left|x_i - x_j\right| - \sum_{i=1}^{N} V(x_i)\right).$$

The McKean-Vlasov equation

We assume the empirical measures of eigenvalues

$$L_N(t) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i(t)}.$$

weakly converges, as $N \to \infty$, to some limit distribution μ_t.

Then the measure μ_t is given as a weak solution of the non-linear partial differential equation:

$$\frac{d}{dt} \int_{\mathbb{R}} f(x)\, \mu_t(dx) = \frac{1}{2} \iint_{\mathbb{R}^2} \frac{\partial_x f(x) - \partial_y f(y)}{x - y}\, \mu_t(dx)\, \mu_t(dy) - \frac{1}{2} \int_{\mathbb{R}} V'(X) f'(x)\, \mu_t(dx),$$

where f is the test function in $C_b^2(\mathbb{R})$, which is called *the McKean-Vlasov equation*.

The free Fokker-Planck equation

Suppose that μ_t is absolutely continuous w.r.t. the Lebesgue measure dx and denote its density by ρ_t.

Then ρ_t satisfies *the free Fokker-Planck equation*:

$$\frac{\partial}{\partial t}\rho_t = -\frac{\partial}{\partial x}\left(\rho_t\left((\mathcal{H}\rho_t)(x) - \frac{1}{2}V'(x)\right)\right).$$

Here, $(\mathcal{H}\rho_t)$ is the *Hilbert transform*$(\times\pi)$, which is defined by the principal value integral

$$(\mathcal{H}f)(x) = \mathrm{p.}v.\int_{\mathbb{R}} \frac{f(y)}{x-y}\,dy = \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{x-\varepsilon} + \int_{x+\varepsilon}^{\infty}\frac{f(y)}{x-y}\,dy\right)$$

for a probability density function f.

Characterization of the measure μ_t

In order to characterize the measure μ_t, it is enough to use the test functions $f(x) = (z-x)^{-1}$, instead of using all test function $f \in C_b^2(\mathbb{R})$ in the McKean-Vlasov equation.

Indeed, let

$$G_t(z) = \int_{\mathbb{R}} \frac{\mu_t(dx)}{z-x}$$

be the Cauchy transform of μ_t. Then we can find that $G_t(z)$ satisfies

$$\frac{\partial}{\partial t} G_t(z) = -G_t(z) \frac{\partial}{\partial z} G_t(z) - \frac{1}{2} \int_{\mathbb{R}} \frac{V'(x)}{(z-x)^2} \mu_t(dx).$$

As the special case, if we put $V(x) = \frac{x^2}{2}$, since

$$-\int_{\mathbb{R}} \frac{x}{(z-x)^2}\,\mu_t(dx) = z\,\frac{\partial}{\partial z} G_t(z) + G_t(z),$$

we obtain

$$\frac{\partial}{\partial t}\,G_t(z) = \Big(-G_t(z) + \frac{1}{2}\,z\Big)\,\frac{\partial}{\partial z} G_t(z) + \frac{1}{2}\,G_t(z),$$

which corresponds to the case of the free Ornstein-Uhlenbeck process.

The free Ornstein-Uhlenbeck process

Let $(\mathcal{M}, \tau)$ be a W^*-probability space. Let $X \in \mathcal{M}$ be a self-adjoint operator and $S \in \mathcal{M}$ be a standard semi-circular element freely independent of X in $(\mathcal{M}, \tau)$.
Then the one parameter family

$$\widetilde{X}(t) = \sqrt{e^{-t}}X + \sqrt{1 - e^{-t}}S \quad (t \geq 0)$$

is called the free Ornstein-Uhlenbeck process,

which is investigated in Biane and Speicher for the free LSI and in Biane and Voiculescu for the free TCI.

The Cauchy transform of $G_t(Z)$ of $\widetilde{X}(t)$ satisfies the above PDE.

Let V be a C^1 function on $\mathbb{R}$. The logarithmic energy with the potential V for a probability measure μ on $\mathbb{R}$ is defined by

$$\Sigma_V(\mu) = -\iint \log|x-y|\, d\mu(x)\, d\mu(y) + \int V(x)\, d\mu(x).$$

When V has a sufficient growth rate at infinity, there exists a unique minimizer $\nu_V \in \mathrm{Prob}(\mathbb{R})$ for the functional $\Sigma_V(\,\cdot\,)$, which is called the *equilibrium measure* of $\Sigma_V(\,\cdot\,)$.

Equilibrium measure

The measure ν_V has a compact support S, and is absolutely continuous with respect to the Lebesgue measure, the density g of which satisfies

$$(\mathcal{H}g)(x) = \frac{1}{2}V'(x) \quad \text{for all} \quad x \in S.$$

More precisely, there exists a constant C such that

$$\begin{aligned} V(x) &= 2\int_{\mathbb{R}} \log|x-y|\, g(y)dy + C \qquad && \text{for} \quad x \in S, \\ V(x) &\geq 2\int_{\mathbb{R}} \log|x-y|\, g(y)dy + C \qquad && \text{for} \quad x \in \mathbb{R}. \end{aligned}$$

Remark

The long-time asymptotically stationary measure of the free Fokker-Planck equation with the potential V

$$\frac{\partial}{\partial t}\rho_t = -\frac{\partial}{\partial x}\left(\rho_t\left((\mathcal{H}\rho_t)(x) - \frac{1}{2}V'(x)\right)\right),$$

is the equilibrium measure ν_V for the entropy functional $\Sigma_V(\,\cdot\,)$.

With the potential function V, we consider the functional Σ_V. Let ν_V be the equilibrium measure (unique minimizer) of Σ_V.
Then the *relative free entropy* of μ with respect to ν_V is defined by

Definition (The relative free entropy)

$$\Sigma(\mu \,|\, \nu_V) = \Sigma_V(\mu) - \Sigma_V(\nu_V).$$

The *relative free Fisher information* $\Phi(\mu \,|\, \nu_V)$ of μ with respect to ν_V is defined by

Definition (The relative free Fisher information)

$$\Phi(\mu \,|\, \nu_V) = 4 \int_{\mathbb{R}} \left((\mathcal{H}f)(x) - \frac{1}{2} V'(x)\right)^2 f(x)\, dx,$$

where f denotes the density of μ.

The definition can be slightly extended to a little more general case that for two compactly supported probability measures μ and ν on $\mathbb{R}$ with $\mu \ll \nu$.

$$\Phi(\mu \,|\, \nu) = 4 \int_{\mathbb{R}} \Big((\mathcal{H}f)(x) - (\mathcal{H}g)(x)\Big)^2 f(x)\, dx,$$

where f and g are density functions of μ and ν, respectively.

For two Borel probability measures μ and ν in $\mathbb{R}$ with finite second moment, the L^2-(or simply 2-) Wasserstein distance $W_2(\mu, \nu)$ between μ and ν is defined by

Definition (The L^2-Wasserstein distance)

$$W_2(\mu, \nu) = \sqrt{\inf_{\pi \in \Pi(\mu,\nu)} \left(\int \left|x - y\right|^2 d\pi(x, y) \right)},$$

where $\Pi(\mu, \nu)$ denotes the set of probability measures on $\mathbb{R} \times \mathbb{R}$ with marginals μ and ν, respectively

There is at least one solution $\pi \in \Pi(\mu, \nu)$ to this minimization problem.

The transport map and optimality

For two probability measures μ and ν on $\mathbb{R}$, a map $T : \mathbb{R} \longrightarrow \mathbb{R}$ defined μ-almost everywhere is said to *push* μ *forward to* ν (or to *transport map of* μ *and* ν) if for every Borel set $B \in \mathbb{R}$, $\nu\big(B\big) = \mu\big(T^{-1}(B)\big)$, which is denoted by $T\#\mu = \nu$.

A transport map T pushing μ forward to ν is said to be *optimal* if

$$W_2(\mu, \nu)^2 = \int \big|x - T(x)\big|^2 \, d\mu(x).$$

Theorem (Brenier)

If μ and ν have finite second moment, then there exists a map T such that $T\#\mu = \nu$, which realizes the optimal transport.

The free logarithmic Sobolev inequality (free LSI)

$$\Sigma(\mu \,|\, \nu) \leq \frac{1}{2K} \Phi(\mu \,|\, \nu)$$

For semicircle laws ν by Biane and Speicher, for the equilibrium measures of a strongly convex potential V by Biane via random matrix approximation.

The free transportation cost inequality (free TCI)

$$W_2(\mu \,|\, \nu) \leq \sqrt{\frac{1}{K}\,\Sigma(\mu \,|\, \nu)}$$

For semicircle laws ν by Biane and Voiculescu, for the equilibrium measures ν_V of a strongly convex potential V by Hiai, Petz, and Ueda via random matrix approximation.

Ledoux gave simpler proof of the free LSI and free TCI based on free Brunn-Minkowski inequality, and also showed the free analogue of Otto-Villani theorem.

Ledoux and Popescu showed the free LSI and the free TCI for the equilibrium measure ν_V of the strictly convex potential V by using mass transportation method and convex analysis *without* random matrix approximation. They also gave the free analogue of HWI inequality.

The free HWI inequality

$$\Sigma(\mu \,|\, \nu) \leq \sqrt{\Phi(\mu \,|\, \nu)}\, W_2(\mu \,|\, \nu) - \frac{1}{2K} W_2(\mu \,|\, \nu)^2$$

The time parameter does not appear in their transport map.

The time derivative of the L^2-Wasserstein distance

First our Situation

Let μ_t and ν_t be two solutions of the free Fokker-Planck equation with initial data μ_0 and ν_0 in $\mathcal{P}_2^c(\mathbb{R})$, and we denote by f_t and g_t the density functions of μ_t and ν_t, respectively.

Let T_t be the optimal transport map, which pushes μ_t forward to ν_t, that is. $T_t \# \mu_t = \nu_t$ for $t \geq 0$.

The map T_t^{-1} is the inverse transport of T_t, that is, $\mu_t = T_t^{-1} \# \nu_t$ for $t \geq 0$.

Time derivative of the square of the L^2-Wasserstein distance

Then the time derivative of the square of the L^2-Wasserstein distance is given as

Theorem A

$$
\begin{aligned}
&\frac{d}{dt}\left(W_2(\mu_t,\nu_t)^2\right) \\
&\quad = 2\int (x - T_t(x))\left((\mathcal{H}f_t)(x) - \frac{1}{2}V'(x)\right) f_t(x)\,dx \\
&\qquad + 2\int (x - T_t^{-1}(x))\left((\mathcal{H}g_t)(x) - \frac{1}{2}V'(x)\right) g_t(x)\,dx
\end{aligned}
$$

This formula can be obtained by applying Otto's calculus.

The following formula is frequently used in our calculation:

Lemma (The free Stein relation)

For a differentiable function η with bounded derivative, we have the formula:

$$2\int \eta(x)\,(\mathcal{H}f)(x)\,f(x)\,dx = \iint \frac{\eta(x)-\eta(y)}{x-y}\,f(x)f(y)\,dxdy.$$

It can be seen by direct calculation.

Noncommutative derivative

Let X be a self-adjoint random variable in a C^*-probability space and denote the density of the distribution μ of X by f.

Then the Hilbert transform $2\big(\mathcal{H}f\big)$ can be regarded as a free analogue of the classical score function because $||2\big(\mathcal{H}f\big)||^2$ the square of L^2 norm in $L^2(\mathbb{R}, d\mu)$ is the free Fisher information of X. $2\big(\mathcal{H}f\big)$ corresponds to the conjugate variable.

Moreover the difference quotient

$$D\eta = \frac{\eta(x) - \eta(y)}{x - y}$$

is regarded as the non-commutative derivative.

Compare to the classical Stein relation

The previous Lemma formula can be read as the free counterpart of the classical Stein relation:

Classical Stein relation

$$E_X\left(\eta(X)\,\rho_X(X)\right) = -E_X\left(\eta'(X)\right),$$

where ρ_X is the classical score function and E_X stands for the expectation with respect to a classical random variable X.

We should also note that the sign of the free analogue of the score function is opposite to the classical one.

Let μ_t and ν_t be the same as in Theorem A.

Theorem B

We assume that the potential function V is uniformly K-convex with a positive constant K, namely, $V''(x) \geq K > 0 \quad \text{for} \quad x \in \mathbb{R}$. Then we obtain

$$\frac{d}{dt}\Big(W_2(\mu_t, \nu_t)^2\Big) \leq -K\, W_2(\mu_t, \nu_t)^2.$$

Thus we have the exponential decay

$$W_2(\mu_t, \nu_t) \leq e^{-(K/2)\,t}\, W_2(\mu_0, \nu_0),$$

which means the uniqueness of the long-time asymptotic stationary measure of the free FP equation.

Split the formula in Theorem A into two parts:

$$\frac{d}{dt}\Big(W_2(\mu_t,\nu_t)^2\Big)$$
$$= -\underbrace{\left(\int \big(x - T_t(x)\big)\, V'(x)\, f_t(x)\, dx + \int \big(x - T_t^{-1}(x)\big)\, V'(x)\, g_t(x)\, dx\right)}_{\text{(I)}}$$
$$-\underbrace{\left(2\int \big(T_t(x) - x\big)\big(\mathcal{H}f_t\big)(x)\, f_t(x)\, dx + 2\int \big(T_t^{-1}(x) - x\big)\big(\mathcal{H}g_t\big)(x)\, g_t(x)\, dx\right)}_{\text{(II)}}.$$

The uniform K-convexity for the potential function V of that $V''(x) \geq K > 0$ for any $x \in \mathbb{R}$ yields the inequality

$$\left(x - T_t(x)\right)\left(V'(x) - V'(T_t(x))\right) \geq K \left|x - T_t(x)\right|^2.$$

Using Taylor expansion, we get

$$V(x) = V(y) + V'(y)(x - y) + \frac{1}{2} V''(x + \theta y) \left|x - y\right|^2$$

for any $x, y \in \mathbb{R}$, where $\theta \in (0, 1)$.

Applying this inequality, the first part can be estimated as

$$\begin{aligned}(\mathrm{I}) &= \int \left(x - T_t(x)\right) V'(x)\, f_t(x)\, dx + \int \left(T_t(x) - x\right) V'(T_t(x))\, f_t(x)\, dx \\ &= \int \left(x - T_t(x)\right) \left(V'(x) - V'(T_t(x))\right) f_t(x)\, dx \\ &\geq \int K \left|x - T_t(x)\right|^2 f_t(x)\, dx = K\, W_2(\mu_t, \nu_t)^2,\end{aligned}$$

On the second part:
Using the free analogue of Stein relation, we obtain

$$\begin{aligned}
2\int &\left(T_t(x) - x\right) \left(\mathcal{H}f_t\right)(x)\, f_t(x)\, dx \\
&= \iint \left(\frac{T_t(x) - T_t(y)}{x - y} - 1\right) f_t(y)\, f_t(x)\, dx\, dy, \\
2\int &\left(T_t^{-1}(x) - x\right) \left(\mathcal{H}g_t\right)(x)\, g_t(x)\, dx \\
&= \iint \left(\frac{T_t^{-1}(x) - T_t^{-1}(y)}{x - y} - 1\right) g_t(y)\, g_t(x)\, dx\, dy \\
&= \iint \left(\frac{x - y}{T_t(x) - T_t(y)} - 1\right) f_t(y)\, f_t(x)\, dx\, dy,
\end{aligned}$$

where at the last equality we have applied the optimal transport map T_t to both of the integral variables x and y.

Hence, we have

$$(\mathrm{II}) = \iint \left(\frac{T_t(x) - T_t(y)}{x-y} + \frac{x-y}{T_t(x)-T_t(y)} - 2 \right) f_t(y)\, f_t(x)\, dx\, dy$$

$$= \iint \left(\sqrt{\frac{T_t(x) - T_t(y)}{x-y}} - \sqrt{\frac{x-y}{T_t(x)-T_t(y)}} \right)^2 f_t(y)\, f_t(x)\, dx\, dy \geq 0,$$

where $\frac{T_t(x) - T_t(y)}{x-y} \geq 0$ since T_t is monotone.

Consequently,

$$\frac{d}{dt}\left(W_2(\mu_t, \nu_t)^2 \right) = -(\mathrm{I}) - (\mathrm{II}) \leq -K\, W_2(\mu_t, \nu_t)^2.$$

□

The derivative of of the L^2-Wasserstein distance

We choose particularly the stationary measure ν_V of the free Fokker-Planck equation as the initial datum ν_0, then it holds, of course, that $g_t(x)dx = d\nu_V(x)$ and

$$\big(\mathcal{H}g_t\big)(x) = \frac{1}{2}V'(x) \quad \text{for} \quad t \geq 0.$$

Thus we have immediately

Theorem A'

$$\frac{d}{dt}\Big(W_2(\mu_t, \nu_V)^2\Big) = 2\int (x - T_t(x))\,\Big(\big(\mathcal{H}f_t\big)(x) - \frac{1}{2}V'(x)\Big)\, f_t(x)\, dx.$$

Moreover, we have

Theorem B'

If V is uniformly K convex that such $V''(x) \geq K > 0$ for $x \in \mathbb{R}$, then we obtain

$$\frac{d}{dt}\left(W_2(\mu_t, \nu_V)^2\right) \leq -K\, W_2(\mu_t, \nu_V)^2.$$

This can be read that

$$W_2(\mu_t, \nu_V) \leq e^{-(K/2)t}\, W_2\left(\mu_0, \nu_V\right) \text{ for } t \geq 0,$$

which implies that μ_t converges to ν_V in the L^2-Wasserstein distance with exponential rate $K/2$.

The relative free Fisher information

For simplicity, we put

$$J_t(x) = \big(\mathcal{H} f_t\big)(x) - \frac{1}{2} V'(x).$$

Then the derivative formula is written as

$$\frac{d}{dt}\left(W_2(\mu_t, \nu_V)^2\right) = 2 \int (x - T_t(x))\, J_t(x)\, d\mu_t(x)$$

and the relative free information is given as

$$\Phi(\mu_t \,|\nu_V) = 4 \int J_t(x)^2 \, d\mu_t(x).$$

$W_2 \, \Phi$ inequalities

In our second situation, we have the following inequalities:

Proposition C

$$(C.1) \qquad \frac{d}{dt}\left(W_2(\mu_t, \nu_V)^2\right) \geq -W_2(\mu_t, \nu_V)\sqrt{\Phi(\mu_t \,|\nu_V\,)},$$

$$(C.2) \qquad \frac{d}{dt}\left(W_2(\mu_t, \nu_V)^2\right) \geq -\frac{1}{K}\,\Phi(\mu_t \,|\nu_V\,).$$

Proof of Prop. C

The inequality (C.1) is a simple application of the Cauthy-Schwarz inequality on $L^2(\mathbb{R}, d\mu_t)$.

$$
\begin{aligned}
\left| \frac{d}{dt} \left(W_2(\mu_t, \nu_V)^2 \right) \right| &= \left| 2 \int (x - T_t(x)) \, J_t(x) \, d\mu_t(x) \right| \\
&\leq \sqrt{\int \left| x - T_t(x) \right|^2 d\mu_t(x)} \sqrt{4 \int J_t(X)^2 \, d\mu_t(x)} \\
&= W_2(\mu_t, \nu_V) \sqrt{\Phi(\mu_t \,|\nu_V\,)},
\end{aligned}
$$

where we should note that $\frac{d}{dt} \left(W_2(\mu_t, \nu_V)^2 \right) \leq 0$.

We know by Theorem B' that

$$\frac{d}{dt}\left(W_2(\mu_t,\nu_V)^2\right) \leq -K\, W_2(\mu_t,\nu_V)^2.$$

Combine it with

$$(C.1) \qquad \frac{d}{dt}\left(W_2(\mu_t,\nu_V)^2\right) \geq -W_2(\mu_t,\nu_V)\,\sqrt{\Phi(\mu_t\,|\nu_V\,)},$$

we obtain

$$W_2(\mu_t,\nu_V) \leq \frac{1}{K}\sqrt{\Phi(\mu_t\,|\nu_V\,)}.$$

Substitute this into (C.1) again, it follows

$$(C.2) \qquad \frac{d}{dt}\left(W_2(\mu_t,\nu_V)^2\right) \geq -\frac{1}{K}\,\Phi(\mu_t\,|\nu_V\,).$$

As we will find later that the inequality

$$(C.2) \qquad \frac{d}{dt}\left(W_2(\mu_t,\nu_V)^2\right) \geq -\frac{1}{K}\,\Phi\left(\mu_t\,|\nu_V\,\right)$$

yields the free TCI.

For this purpose, we will see the dissipation formulas of the free entropy

The dissipations of the relative free entropy

Let ν_V be the stationary measure of the free Fokker-Planck equation with the C^2 potential V, and let μ_t be the time-evolution by the free Fokker-Planck equation starting from a compactly supported probability measure μ_0.

Proposition D

$$\frac{d}{dt}\Sigma(\mu_t \mid \nu_V) = -\frac{1}{2}\Phi(\mu_t \mid \nu_V). \tag{D.1}$$

This formula is well-known since the work of Biane and Speicher.

Since $\Sigma_V(\nu_V)$ does not depend on t,

$$\frac{d}{dt}\Sigma(\mu_t \,|\, \nu_V) = -\frac{d}{dt}\iint \log|x-y|\, d\mu_t(x)d\mu_t(y) + \frac{d}{dt}\int V(x)\, d\mu_t(x).$$

The calculation is routine that by using the free Fokker-Planck equation, we exchange the time derivative to the space one and apply integration by parts.

Dissipation of the entropy dissipation

Here we shall consider $\frac{d^2}{dt^2}\Sigma\left(\mu_t \,\middle|\, \nu_V\right)$, which is nothing but the timie derivative $\frac{d}{dt}\Phi\left(\mu_t \,\middle|\, \nu_V\right)$.

This is also obtained by the result of the entropy dissipation formula in Carrillo, McCann, and Villani by Otto's calculus.

Theorem E

Under the same situation in Proposition D, we have

$$\frac{d}{dt}\Phi\big(\mu_t \,\big|\, \nu_V\big) = -4\int V''(x)\,J_t(x)^2\,d\mu(x) - 4\iint\left(\frac{J_t(x)-J_t(y)}{x-y}\right)^2 d\mu(x)\,d\mu(y),$$

where $J_t(x) = \big(\mathcal{H}f_t\big)(x) - \frac{1}{2}V'(x)$.

Theorem F

If the C^2 potential function V is uniformly K convex that $V''(x) \geq K > 0$ for $x \in \mathbb{R}$, then we obtain

$$\frac{d}{dt}\Phi\left(\mu_t \,\middle|\, \nu_V\right) \leq -K\,\Phi\left(\mu_t \,\middle|\, \nu_V\right),$$

Hence, we have

$$\Phi\left(\mu_t \,\middle|\, \nu_V\right) \leq e^{-Kt}\Phi\left(\mu_0 \,\middle|\, \nu_V\right) \quad \text{for} \quad t \geq 0,$$

which implies that μ_t converges to the equilibrium ν_V in the relative Fisher information with exponential rate K.

By the uniform K-convexity $V''(x) \geq K > 0$, it follows that

$$\begin{aligned}\frac{d}{dt}\Phi\left(\mu_t \,\middle|\, \nu_V\right) &\leq -4\int V''(x)J_t(x)^2 d\mu_t(x)\\ &\leq -4K\int J_t(x)^2 d\mu_t(x) = -K\,\Phi\left(\mu_t \,\middle|\, \nu\right).\end{aligned}$$

Together with

$$(D.1) \qquad \frac{d}{dt}\Sigma\left(\mu_t \,\middle|\, \nu_V\right) = -\frac{1}{2}\Phi\left(\mu_t \,\middle|\, \nu_V\right),$$

we have the following inequality on the derivatives.

The inequality for the fee LSI

Corollary G

$$\frac{d}{dt}\Sigma\left(\mu_t \,\middle|\, \nu_V\right) \geq \frac{1}{2K}\frac{d}{dt}\Phi\left(\mu_t \,\middle|\, \nu_V\right) \ \text{ for } \ t \geq 0.$$

As we will see this inequality yields the free LSI by the time integration.

But we need the convergence of μ_t to the equilibrium ν_V in the relative free entropy.

As the same in the classical case, the exponential decay of the relative free entropy $\Sigma\left(\mu_t \,\middle|\, \nu_V\right)$ and the free LSI are equivalent.

Here we will show the convergence of $\Sigma\left(\mu_t \,\middle|\, \nu_V\right)$ to 0 without relying the free LSI.

Convergence to the equilibrium in the relative free entropy

Theorem H

In our second situation with uniform K-convexity of V that $V''(x) \geq K > 0$ for $x \in \mathbb{R}$, we obtain

$$(H.1) \qquad \Sigma(\mu_t \,|\, \nu_V) \leq -\frac{d}{dt}\left(W_2(\mu_t, \nu_V)^2\right) - \frac{K}{2} W_2(\mu_t, \nu_V)^2$$

This inequality implies the convergence of μ_t to the equilibrium ν_V in the relative free entropy.

Proof of Th. H

The proof is similar to one for the free HWI of Ledoux and Popescu that we will use the optimal transport map and apply the free Stein relation, because the free HWI can be obtained as the direct consequence of this inequality.

The key points are as follows:

Key 1

Using the transport map T_t such that $T_t\#\mu_t = \nu$, the relative free entropy $\Sigma\big(\mu_t \,\big|\, \nu_V\big)$ can be reformulated as

$$\Sigma\big(\mu_t \,\big|\, \nu_V\big) = -\iint \log|x-y|\, d\mu_t(x)\, d\mu_t(y) + \int V(x)\, d\mu_t(x)$$

$$+ \iint \log\big|T_t(x) - T_t(y)\big|\, d\mu_t(x)\, d\mu_t(y) - \int V\big(T_t(x)\big)\, d\mu_t(x)$$

$$= \iint \log \frac{T_t(x) - T_t(y)}{x-y}\, d\mu_t(y)\, d\mu_t(y) - \int \Big(V\big(T_t(x)\big) - V\big(x\big)\Big)\, d\mu_t(x).$$

Key 2

Using the free Stein relation, the time derivative $\frac{d}{dt}W_2(\mu_t, \nu_V)^2$ can be reformulated as

$$\frac{d}{dt}\left(W_2(\mu_t, \nu_V)^2\right) = 2\int (x - T_t(x))\left((\mathcal{H}f_t)(x) - \frac{1}{2}V'(x)\right) f_t(x)\,dx$$

$$= -\iint \left(\frac{T_t(x) - T_t(y)}{x - y} - 1\right) d\mu_t(x)\,d\mu_t(y)$$

$$+ \int \left(T_t(x) - x\right) V'(x)\,d\mu_t(x).$$

We now know two inequalities:

$$(H.1) \qquad \Sigma(\mu_t \,|\, \nu_V) \leq -\frac{d}{dt}\left(W_2(\mu_t, \nu_V)^2\right) - \frac{K}{2} W_2(\mu_t, \nu_V)^2$$

$$(C.1) \qquad \frac{d}{dt}\left(W_2(\mu_t, \nu_V)^2\right) \geq -W_2(\mu_t, \nu_V)\sqrt{\Phi(\mu_t \,|\nu_V\,)},$$

Hence we obtain

The free HWI inequality

$$\Sigma(\mu_t \,|\, \nu_V) \leq W_2(\mu_t, \nu_V)\sqrt{\Phi(\mu_t \,|\nu_V\,)} - \frac{K}{2} W_2(\mu_t, \nu_V)^2,$$

Exponential decay of $\Sigma(\mu_t \mid \nu)$

μ_t converges to the equilibrium

in the L^2-Wasserstein distance with exponential rate $K/2$,
in the relative free Fisher information with exponential rate K.

Hence with the help of the free HWI inequality, we have that

μ_t converges to the equilibrium
in the relative free entropy with exponential rate K.

We know the inequality by Cor. G that

$$\frac{d}{dt}\Sigma(\mu_t \,|\, \nu_V) \geq \frac{1}{2K}\frac{d}{dt}\Phi(\mu_t \,|\, \nu_V) \quad \text{for} \quad t \geq 0.$$

Taking the time integration from 0 to ∞, we have

The free LSI

$$\Sigma(\mu_0 \,|\, \nu_V) \leq \frac{1}{2K}\Phi(\mu_0 \,|\, \nu_V).$$

As we noted $\lim_{t\to\infty}\Phi(\mu_t \,|\nu_V) = 0$ and $\lim_{t\to\infty}\Sigma(\mu_t \,|\nu_V) = 0$.

We know the inequality and the equality by (C.2) and (D,1) that

$$\begin{cases} \dfrac{d}{dt}\Big(W_2(\mu_t,\nu_V)^2\Big) \geq -\dfrac{1}{K}\,\Phi\big(\mu_t\,|\nu_V\,\big), \\ \dfrac{d}{dt}\Sigma\big(\mu_t\,\big|\,\nu_V\big) = -\dfrac{1}{2}\Phi\big(\mu_t\,\big|\,\nu_V\big). \end{cases}$$

Thus we obtain the inequality

$$\frac{d}{dt}\Big(W_2(\mu_t,\nu_V)^2\Big) \geq \frac{2}{K}\,\frac{d}{dt}\Sigma\big(\mu_t\,|\nu_V\,\big).$$

The free TCI

Taking the time integration from 0 to ∞, we have

$$W_2(\mu, \nu_V)^2 \leq \frac{2}{K} \Sigma(\mu \,|\nu_V\,)$$

and, hence,

The free TCI

$$W_2(\mu, \nu_V) \leq \sqrt{\frac{2}{K} \Sigma(\mu \,|\nu_V\,)}$$

As we noted $\lim_{t \to \infty} W_2\left(\mu_t, \nu_V\,\right)^2 = 0$ and $\lim_{t \to \infty} \Sigma\left(\mu_t \,|\nu_V\,\right) = 0$.

Semicircular perturbations and the relative free entropy

In the classical case

For probability measures μ and ν ($\mu \ll \nu$) with the densities f and g, respectively, the relative classical entropy (Kullback-Leibler divergence) $D(\mu \,|\, \nu)$ is defined as

$$D(\mu \,|\, \nu) = \int \Big(\log f(x) - \log g(x) \Big) f(x) \mathrm{d}x,$$

and the relative classical Fisher information $I(\mu \,|\, \nu)$ is defined as

$$I(\mu \,|\, \nu) = \int \Big(\{ \log f(x) \}' - \{ \log g(x) \}' \Big)^2 f(x) \mathrm{d}x.$$

We consider the gaussian perturbations μ_t and ν_t of μ and ν, and take the derivation of t.

The gaussian perturbation

Let X be a random variable, which has the distribution μ and let Z be a standard gaussian random variable.

We consider the random variable $X + \sqrt{t}Z$ and denote the corresponding distribution by μ_t, which is called the gaussian perturbation of μ.

Let μ_t and ν_t be the gaussian perturbations of μ and ν, respectively. Then it holds that

Proposition

$$\frac{d}{dt} D(\mu_t \,|\, \nu_t) = -\frac{1}{2} I\big(\mu_t \,|\, \nu_t\big).$$

Verdú (2010) showed the above identity via the estimation theory, Hirata, Nemoto, and Y. (2012) gave another more direct proof by using the heat equation and integration by parts.

Here, let us see the free version of this identity.

Assumption

We consider two compactly supported probability measures μ and ν on $\mathbb{R}$ such that $\mu \ll \nu$.

We assume that μ and ν are absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ and have the density functions p and q, respectively.

After this, we shall pay attention to the case where

$$\mathrm{Supp}(p) \subset \mathrm{Supp}(q).$$

Let X be a self-adjoint random variable in a W^*-probability space $(\mathcal{M}, \tau)$ with the distribution μ, and S be a (standard) $(0,1)$-semicircular element in $(\mathcal{M}, \tau)$ freely independent of X.

Let μ_t be the distribution of semicircular perturbed random variable $X + \sqrt{t}S$, that is, the free convolution $\mu_t = \mu \boxplus w_{0,t}$, where $w_{0,t}$ is the centered semicircular law of variance t.

Theorem

The derivative formula of the free case can be give as

$$\frac{d}{dt}\Sigma(\mu_t \mid \nu_t) = -\frac{1}{2}\Phi(\mu_t|\nu_t) + \Psi(\mu_t|\nu_t).$$

Here

$$\Psi(\mu_t|\nu_t) = -2\int_{S(p)} (\mathcal{H}p)\,(\mathcal{H}q)\,p\,dx + \int_{S(p)} (\mathcal{H}q)^2\,p\,dx \\ -2\int_{S(q)} (\mathcal{H}q)^2\,q\,dx + \int_{S(p)} \pi^2\,q^2 p\,dx.$$

There is an extra term $\Psi(\mu_t|\nu_t)$ in general.

Theorem

If $\Psi(\mu_t|\nu_t) = 0$ for $t > 0$ then we have the integral representation of the relative free entropy,

$$\Sigma(\mu\,|\,\nu) = \frac{1}{2}\int_0^\infty \Phi(\mu_t|\nu_t)dt.$$

Since the relative free entropy is invariant under dilations and the dilation $D_{\frac{1}{\sqrt{t}}}(\mu_t) \longrightarrow w_{0,1}$ weakly as $t \to \infty$, where $w_{0,1}$ is the standard semicircle law, we have

$$\lim_{t\to\infty} \Sigma(\mu_t\,|\,\nu_t) = \lim_{t\to\infty} \Sigma\big(D_{\frac{1}{\sqrt{t}}}(\mu_t)\,|\,D_{\frac{1}{\sqrt{t}}}(\nu_t)\big) = \Sigma\big(w_{0,1}\,|\,w_{0,1}\big) = 0.$$

Remark

For the centered semicircle law ν, we can find that for any probability measure μ with $\mathrm{Supp}(\mu) \subset \mathrm{Supp}(\nu)$, $\Psi(\mu_t|\nu_t) = 0$ for $t > 0$.

Especially, we take the standard semicircle law $w_{0,1}$ as the reference measure ν and Let μ be a standarized measure.

Then ν_t becomes $w_{0,t}$, the centered semicircular distribution of variance t, and we have

$$\Phi(\mu_t \mid \nu_t) = \Phi(\mu_t) - \frac{1}{1+t}.$$

Thus by the integral formula above, we obtain the relative free entropy

$$\Sigma(\mu \mid w_{0,1}) = \frac{1}{2}\int_0^\infty \left(\Phi(\mu_t) - \frac{1}{1+t}\right) dt.$$

Formula for the maicrostate free approach

For the free entropy $\chi(\mu)$, using the non-semicircularity equality (semicircular is maximum),

$$\frac{1}{2}\log(2\pi e) - \chi(\mu) = \Sigma\left(\mu \,|\, w_{0,1}\right),$$

we obtain the integral representation of the free entropy for a probability measure of unit variance

$$\begin{aligned}\chi(\mu) &= -\Sigma\left(\mu \,|\, w_{0,1}\right) + \frac{1}{2}\log(2\pi e) \\ &= \frac{1}{2}\int_0^\infty \left(\frac{1}{1+t} - \Phi(\mu_t)\right) dt + \frac{1}{2}\log(2\pi e),\end{aligned}$$

which is consistent with Voiculescu's formula of the micro states free approach to the free entropy.

The key of proof

Let ν_t be the semicircular perturbation of ν, and let $V(x,t)$ be the potential function, for which the equilibrium measure of $\Sigma_{V(x,t)}$ is given by ν_t.

Then we should be careful to calculate the derivative

$$\frac{d}{dt}\int_{S(p)} V(x,t)p(x,t)\,dx$$

by taking a count of the inclusion of supports $S(p) \subset S(q)$.

Key Lemma

Lemma

$$\frac{d}{dt}\int_{S(p)} V(x,t)\,p(x,t)\,dx - \frac{d}{dt}\int_{S(q)} V(x,t)\,q(x,t)\,dx$$
$$= \int_{S(p)} \left(\pi^2\,q(x,t)^2 - \left((\mathcal{H}q)(x,t)\right)^2\right) p(x,t)\,dx$$
$$+\,2\int_{S(p)} (\mathcal{H}p)(x,t)\,(\mathcal{H}q)(x,t)\,p(x,t)\,dx - 4\int_{S(q)} \left((\mathcal{H}q)(x,t)\right)^2 q(x,t)\,dx.$$

The complex Burgers equation

From the complex Burgers equation for semicircular purterbation, we have

$$\begin{cases} \frac{\partial}{\partial t}(\mathcal{H}\rho_t)(x) = \frac{1}{2}\frac{\partial}{\partial x}\left(\left(\pi\,\rho_t(x)\right)^2 - \left((\mathcal{H}\rho_t)(x)\right)^2\right), \\ \frac{\partial}{\partial t}\,\rho_t(x) = -\frac{\partial}{\partial x}\left((\mathcal{H}\rho_t)(x)\,\rho_t(x)\right), \end{cases}$$

for $x \in \mathrm{Supp}(\rho_t)$ and $t \geq 0$.

In the proof for key Lemma, we have to use not only the lower equation (the free heat equation) but both of them.

Concluding remarks

Topic 1 and 2 are completely the same in parallel to the classical case (both of methods and results).

Topic 3 is a little different compare to the classical case. But at least in the case where the reference measure is semicircular, it goes in parallel to the classical case.

Refereces

[Yo] H. Yoshida, Remark on the free relative entropy and the free Fisher information distance, *Inf. Dimens. Anal. Quant. Probab. Rel. Top.* **13**, (2010) 243–255.

[HNY] M. Hirata, A.Nemoto, H. Yoshida, An integral representation of the relative entropy, *Entropy* **14**, (2012) 1469-1477.

Thank you very much for your attention.