Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2009, Vol. 45, No. 1, 239–249
DOI: 10.1214/08-AIHP164

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

A log-Sobolev type inequality for free entropy of two projections

Fumio Hiai[a,1,2] and Yoshimichi Ueda[b,1,3]

[a]*Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai 980-8579, Japan*
[b]*Graduate School of Mathematics, Kyushu University, Fukuoka 810-8560, Japan*

Received 16 November 2006; revised 8 December 2007; accepted 2 January 2008

Abstract. We prove a kind of logarithmic Sobolev inequality claiming that the mutual free Fisher information dominates the microstate free entropy adapted to projections in the case of two projections.

Résumé. Nous prouvons un genre d'inégalité de Sobolev logarithmique qui montre que l'information de Fisher libre domine l'entropie de micro-états libre adaptée aux projections dans le cas de deux projections.

MSC: Primary 46L54; secondary 94A17; 60E15

Keywords: Logarithmic Sobolev inequality; Free entropy; Mutual free Fisher information

Introduction

The aim of this paper is to add a new result to Voiculescu's liberation theory initiated in [20]. The mutual free Fisher information φ^* and the mutual free information i^* were introduced there for subalgebras unlike those quantities in [19] for random variables. More precisely, what we will prove is the inequality

$$-\chi_{\mathrm{proj}}(p,q) \le \varphi^*\bigl(\mathbb{C}p + \mathbb{C}(1-p) : \mathbb{C}q + \mathbb{C}(1-q)\bigr) \tag{0.1}$$

for projections p, q in a tracial W^*-probability space under a natural assumption. Here, χ_{proj} denotes the microstate definition of free entropy adapted to projections, which was proposed by Voiculescu [20], Section 14.2, and naturally appeared as the rate function in the large deviation principle for a pair of random projection matrices in [8]. One should remember that the original microstate free entropy χ is meaningless for projections as χ always takes $-\infty$ for them. Our subsequent paper [10] will provide several basic properties of χ_{proj} similar to those of the original χ developed in [16–18].

At least to the best of our knowledge, no further work on the liberation theory has been made after the appearance of [20]. Our present result may add a new insight to what Voiculescu discussed in [20], Sections 14.1 and 14.2, though the situation we deal with is very restricted. In fact, our proof suggests that the inequality (0.1) should be regarded as a kind of logarithmic Sobolev inequality, whose original form is an inequality giving an upper bound of the (relative) entropy by the (relative) Fisher information (or Dirichlet form) up to a constant (see, e.g. [15], Section 9.2) and whose free analogs were obtained in [3,9,12]. From this point of view our result suggests a possible relation between $-\chi_{\mathrm{proj}}$ and i^* at least for pairs of projections.

[1]Supported in part by Japan Society for the Promotion of Science, Japan–Hungary Joint Project.
[2]Supported in part by Grant-in-Aid for Scientific Research (B)17340043.
[3]Supported in part by Grant-in-Aid for Young Scientists (B)17740096.

1. Preliminaries

1.1. *Free entropy for projections*

For $N \in \mathbb{N}$ let $\mathrm{U}(N)$ be the $N \times N$ unitary group. Let $G(N,k)$ denote the set of all $N \times N$ orthogonal projection matrices of rank k, that is, $G(N,k)$ is identified with the Grassmannian manifold consisting of k-dimensional subspaces in $\mathbb{C}^N$, and it is also identified with the homogeneous space $\mathrm{U}(N)/(\mathrm{U}(k) \oplus \mathrm{U}(N-k))$. The unitarily invariant probability measure $\gamma_{G(N,k)}$ on $G(N,k)$ corresponds to the measure on $\mathrm{U}(N)/(\mathrm{U}(k) \oplus \mathrm{U}(N-k))$ induced from the Haar probability measure on $\mathrm{U}(N)$.

Let $(p_1, \dots, p_n)$ be an n-tuple of projections in a tracial W^*-probability space $(\mathcal{M}, \tau)$ with $\alpha_i := \tau(p_i)$, $1 \le i \le n$. Following Voiculescu's proposal in [20], Section 14.2, we define the *free entropy* $\chi_{\mathrm{proj}}(p_1, \dots, p_n)$ of $(p_1, \dots, p_n)$ as follows: Choose $k(N,i) \in \{0, 1, \dots, N\}$ for each $N \in \mathbb{N}$ and $1 \le i \le n$ in such a way that $k(N,i)/N \to \alpha_i$ as $N \to \infty$ for $1 \le i \le n$. For each $m \in \mathbb{N}$ and $\varepsilon > 0$, $\Gamma_{\mathrm{proj}}(p_1, \dots, p_n; k(N,1), \dots, k(N,n); N, m, \varepsilon)$ denotes the set of $(P_1, \dots, P_n) \in \prod_{i=1}^n G(N, k(N,i))$ satisfying $|\mathrm{tr}_N(P_{i_1} \cdots P_{i_r}) - \tau(p_{i_1} \cdots p_{i_r})| < \varepsilon$ for all $1 \le i_1, \dots, i_r \le n$, $1 \le r \le m$, where tr_N stands for the normalized trace on the $N \times N$ matrices. Then $\chi_{\mathrm{proj}}(p_1, \dots, p_n)$ is defined to be

$$\inf_{m \in \mathbb{N}, \varepsilon > 0} \limsup_{N \to \infty} \frac{1}{N^2} \log\left(\bigotimes_{i=1}^n \gamma_{G(N,k(N,i))} \right)\bigl(\Gamma\bigl(p_1, \dots, p_n; k(N,1), \dots, k(N,n); N, m, \varepsilon\bigr)\bigr),$$

which is known to be independent of the choices of $k(N,i)$ with $k(N,i)/N \to \alpha_i$, $1 \le i \le n$ (see [10], Proposition 1.1).

In this paper we are concerned with the case of two projections. Let (p,q) be a pair of projections in $(\mathcal{M}, \tau)$ with $\alpha := \tau(p)$ and $\beta := \tau(q)$. Set

$$\begin{aligned} &E_{11} := p \wedge q, \qquad E_{10} := p \wedge q^\perp, \qquad E_{01} := p^\perp \wedge q, \qquad E_{00} := p^\perp \wedge q^\perp, \\ &E := \mathbf{1} - (E_{00} + E_{01} + E_{10} + E_{11}) \end{aligned}$$

and $\alpha_{ij} := \tau(E_{ij})$ for $i, j = 0, 1$. Then E and E_{ij} are in the center of $\mathcal{N} := \{p, q\}''$ and $(E\mathcal{N}E, \tau|_{E\mathcal{N}E})$ is isomorphic to $L^\infty((0,1), \nu; M_2(\mathbb{C}))$, the L^∞-algebra of $M_2(\mathbb{C})$-valued functions, where ν is a measure on $(0,1)$ with $\nu((0,1)) = 1 - \sum_{i,j=0}^1 \alpha_{ij}$. Here EpE and EqE correspond to

$$t \in (0,1) \mapsto \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} t & \sqrt{t(1-t)} \\ \sqrt{t(1-t)} & 1-t \end{bmatrix},$$

respectively, and $\tau|_{E\mathcal{N}E}$ is represented as

$$\tau(a) = \int_0^1 \mathrm{tr}_2\bigl(a(t)\bigr)\, \mathrm{d}\nu(t)$$

for $a \in E\mathcal{N}E$ corresponding to $a(\cdot) \in L^\infty((0,1), \nu; M_2(\mathbb{C}))$. In this way, the mixed moments of (p,q) with respect to τ are determined by the data $(\nu, \{\alpha_{ij}\}_{i,j=0}^1)$. Although ν is not necessarily a probability measure, we denote $\Sigma(\nu) := \int_0^1 \int_0^1 \log|x-y|\, \mathrm{d}\nu(x)\, \mathrm{d}\nu(y)$ in the same fashion as in [16]. Furthermore we set

$$\rho := \min\{\alpha, \beta, 1-\alpha, 1-\beta\} = \frac{1}{2}\left(1 - \sum_{i,j=0}^1 \alpha_{ij}\right), \tag{1.1}$$

$$C := \rho^2 B\left(\frac{|\alpha - \beta|}{\rho}, \frac{|\alpha + \beta - 1|}{\rho}\right) \tag{1.2}$$

(meant zero if $\rho = 0$), where the function $B(s,t)$ in $s, t \ge 0$ was given in [8], Proposition 2.1. The following expression was obtained in [8] as a consequence of the large deviation principle for an independent pair of random projection matrices.

Lemma 1.1 ([8], Theorem 3.2, Proposition 3.3). *If* $\alpha_{00}\alpha_{11} = \alpha_{01}\alpha_{10} = 0$, *then*

$$\chi_{\mathrm{proj}}(p,q) = \frac{1}{4}\Sigma(\nu) + \frac{\alpha_{01}+\alpha_{10}}{2}\int_0^1 \log x \, \mathrm{d}\nu(x) + \frac{\alpha_{00}+\alpha_{11}}{2}\int_0^1 \log(1-x)\, \mathrm{d}\nu(x) - C,$$

and otherwise $\chi_{\mathrm{proj}}(p,q) = -\infty$.

Note that the condition $\alpha_{00}\alpha_{11} = \alpha_{01}\alpha_{10} = 0$ is equivalent to

$$\begin{cases} \alpha_{11} = \max\{\alpha+\beta-1, 0\}, & \alpha_{10} = \max\{\alpha-\beta, 0\}, \\ \alpha_{00} = \max\{1-\alpha-\beta, 0\}, & \alpha_{01} = \max\{\beta-\alpha, 0\}, \end{cases} \tag{1.3}$$

and in this case, $\alpha_{01}+\alpha_{10} = |\alpha-\beta|$ and $\alpha_{00}+\alpha_{11} = |\alpha+\beta-1|$.

1.2. *Mutual free Fisher information*

Let $\mathcal{A}$ and $\mathcal{B}$ be two unital $*$-subalgebras in $(\mathcal{M},\tau)$, which are assumed to be algebraically free. Let $\mathcal{A}\vee\mathcal{B}$ and $W^*(\mathcal{A}\cup\mathcal{B})$ denote the subalgebra and the von Neumann subalgebra, respectively, generated by $\mathcal{A}\cup\mathcal{B}$. Let $\delta_{\mathcal{A}:\mathcal{B}}$ be the derivation from $\mathcal{A}\vee\mathcal{B}$ into the $\mathcal{A}\vee\mathcal{B}$-bimodule $(\mathcal{A}\vee\mathcal{B})\otimes(\mathcal{A}\vee\mathcal{B})$ uniquely determined by $\delta_{\mathcal{A}:\mathcal{B}}(a) = a\otimes\mathbf{1}-\mathbf{1}\otimes a$ for $a\in\mathcal{A}$ and $\delta_{\mathcal{A}:\mathcal{B}}(b) = 0$ for $b\in\mathcal{B}$. If there is an element $\xi\in L^1(W^*(\mathcal{A}\cup\mathcal{B}))$ such that $\tau(\xi x) = (\tau\otimes\tau)(\delta_{\mathcal{A}:\mathcal{B}}(x))$ for $x\in\mathcal{A}\vee\mathcal{B}$, then ξ is called the *liberation gradient* of $(\mathcal{A},\mathcal{B})$ and denoted by $j(\mathcal{A}:\mathcal{B})$. Voiculescu [20] introduced the *mutual free Fisher information* of $\mathcal{A}$ relative to $\mathcal{B}$ by $\varphi^*(\mathcal{A}:\mathcal{B}) := \|j(\mathcal{A}:\mathcal{B})\|_2^2$ ($\|\cdot\|_2$ stands for the L^2-norm with respect to τ) if $j(\mathcal{A}:\mathcal{B})$ exists in $L^2(W^*(\mathcal{A}\cup\mathcal{B}))$; otherwise $\varphi^*(\mathcal{A}:\mathcal{B}) := +\infty$. See [20] for more about the mutual free Fisher information.

Let (p,q) be a pair of projections in $(\mathcal{M},\tau)$ and set $\mathcal{A} := \mathbb{C}p+\mathbb{C}(\mathbf{1}-p)$, $\mathcal{B} := \mathbb{C}q+\mathbb{C}(\mathbf{1}-q)$. Then the liberation gradient $j(\mathcal{A}:\mathcal{B})$ and the mutual free Fisher information $\varphi^*(\mathcal{A}:\mathcal{B})$ were computed in [20]. Here, recall that the Hilbert transform of a function f with $f(x)/(1+|x|)\in L^1(\mathbb{R},\mathrm{d}x)$ is defined to be

$$(Hf)(x) := \lim_{\varepsilon\searrow 0}(H_\varepsilon f)(x) \quad \text{with } (H_\varepsilon f)(x) := \int_{|x-t|>\varepsilon} \frac{f(t)}{x-t}\,\mathrm{d}t$$

whenever the limit exists almost everywhere. We need the following a-bit-improved version of [20], Proposition 12.7, because the original version is not applicable, in particular, to the free case when $\tau(p) = \tau(q) = 1/2$ due to the L^3-assumption with respect to $\mathrm{d}x$ rather than $\mathbf{1}_{(0,1)}(x)x(1-x)\,\mathrm{d}x$.

Lemma 1.2. *With the same notations as in Section 1.1 assume that* $\alpha_{00}\alpha_{11} = \alpha_{01}\alpha_{10} = 0$, ν *has the density* $f := \mathrm{d}\nu/\mathrm{d}x \in L^3((0,1), x(1-x)\,\mathrm{d}x)$ *and moreover*

$$\int_0^1 \left(\frac{\alpha_{01}+\alpha_{10}}{x} + \frac{\alpha_{11}+\alpha_{00}}{1-x}\right) f(x)\,\mathrm{d}x < +\infty. \tag{1.4}$$

Define $X := pqp + (\mathbf{1}-p)(\mathbf{1}-q)(\mathbf{1}-p)$ *and*

$$\phi(x) := (Hf)(x) + \frac{\alpha_{01}+\alpha_{10}}{x} - \frac{\alpha_{00}+\alpha_{11}}{1-x}, \quad 0<x<1.$$

Then

$$j(\mathcal{A}:\mathcal{B}) = [q,p]\phi(EXE) \in L^2(\mathcal{M},\tau)$$

and hence

$$\varphi^*(\mathcal{A}:\mathcal{B}) = \int_0^1 \phi(x)^2 f(x)x(1-x)\,\mathrm{d}x < +\infty.$$

The assumption (1.4) can be reduced when $\alpha=\beta$ or $\alpha+\beta=1$. In fact, (1.4) is nothing when $\alpha=\beta=1/2$; it means $\int_0^1 x^{-1} f(x)\,\mathrm{d}x<+\infty$ when $\alpha+\beta=1$ but $\alpha\neq\beta$. All the assumptions of Lemma 1.2 are satisfied, in particular, when p and q are free (see [21], Example 3.6.7). Note ([20], Propositions 5.17 and 9.3.c) that $j(\mathcal{A}:\mathcal{B})=0$ (or equivalently $\varphi^*(\mathcal{A}:\mathcal{B})=0$) if and only if p and q are free.

In [20], Section 12, the support of ν was assumed to be an infinite set to guarantee that $\mathcal{A}$ and $\mathcal{B}$ are algebraically free. As long as $\nu\neq 0$, that is automatically satisfied from the assumption of ν having the density. In the case where $\nu=0$ so that $\rho=0$ by (1.1), it follows that $p\in\{0,\mathbf{1}\}$ or $q\in\{0,\mathbf{1}\}$; hence Lemma 1.2 trivially holds.

Proof of Lemma 1.2. Since $\int_0^1 (x(1-x))^{-1/2}\,\mathrm{d}x<+\infty$, the weighted version of M. Riesz's theorem for Hilbert transform ([11], Theorem 8) shows that there is a constant $C_w>0$ depending only on the weight function $w(x):=\mathbf{1}_{(0,1)}(x)x(1-x)$ such that for every function g (whose Hg can be defined)

$$\|Hg\|_{w,3}\le C_w\|g\|_{w,3} \tag{1.5}$$

with the weighted norm $\|g\|_{w,p}:=(\int_0^1 |g(x)|^p x(1-x)\,\mathrm{d}x)^{1/p}$ for $1\le p<\infty$, and moreover $\|H_\varepsilon g-Hg\|_{w,3}\to 0$ as $\varepsilon\searrow 0$ whenever $\|g\|_{w,3}<+\infty$. In what follows we use the same symbols as in [20], Section 12, with small exception; $p,q,\mathcal{A},\mathcal{B}$ and x,x_1,x_2 are used instead of P, Q, A, B and t,t_1,t_2, respectively. By the facts mentioned above one easily has

$$\begin{aligned}
&\int_0^1\int_0^1\left(\frac{x_1^{n+1}-x_2^{n+1}}{x_1-x_2}-\frac{x_1^n-x_2^n}{x_1-x_2}\right)\mathrm{d}\nu(x_1)\,\mathrm{d}\nu(x_2)\\
&\quad=-\lim_{\varepsilon\searrow 0}\int_{|x_1-x_2|>\varepsilon}\left(x_1^{n-1}x_1(1-x_1)\frac{f(x_2)}{x_1-x_2}f(x_1)+x_2^{n-1}x_2(1-x_2)\frac{f(x_1)}{x_2-x_1}f(x_2)\right)\mathrm{d}x_1\,\mathrm{d}x_2\\
&\quad=-2\lim_{\varepsilon\searrow 0}\int_0^1 x^{n-1}(H_\varepsilon f)(x)f(x)x(1-x)\,\mathrm{d}x\\
&\quad=-2\int_0^1 x^{n-1}(Hf)(x)f(x)x(1-x)\,\mathrm{d}x.
\end{aligned}$$

Hence the assertion of [20], Lemma 12.6, can be changed to

$$\begin{aligned}
(\tau\otimes\tau)\circ\delta_{\mathcal{B}:\mathcal{A}}\big((pq)^n\big)=\ &-\frac{1}{2}\int_0^1 x^{n-1}(Hf)(x)f(x)x(1-x)\,\mathrm{d}x\\
&+\frac{1-\alpha}{2}\int_0^1 x^{n-1}(x-1)\,\mathrm{d}\nu(x)+\frac{\alpha_{00}+\alpha_{11}}{2}\int_0^1 x^{n-1}\,\mathrm{d}\nu(x)
\end{aligned} \tag{1.6}$$

under the assumptions of Lemma 1.2. The rest of the proof goes along the same line as [20], Proposition 12.7, with replacing [20], Lemma 12.6, by (1.6). □

2. Main result

For simplicity we hereafter write $\varphi^*(p:q)$ for the mutual free Fisher information $\varphi^*(\mathbb{C}p+\mathbb{C}(\mathbf{1}-p):\mathbb{C}q+\mathbb{C}(\mathbf{1}-q))$ (see Section 1.2).

Theorem 2.1. *The inequality* $-\chi_{\mathrm{proj}}(p,q)\le\varphi^*(p:q)$ *holds under the same assumptions as in Lemma 1.2.*

The main idea of the proof is a random matrix approximation procedure based on the large deviation shown in [8]. In fact, we apply Bakry and Emery's logarithmic Sobolev inequality in [1] to random projection matrix pairs (or probability measures on the product of two Grassmannian manifolds) and pass to the scaling limit as the matrix size goes to ∞. Thus our inequality can be regarded as a kind of free probability counterpart of the logarithmic Sobolev inequality.

In the proof, we have to examine Bakry and Emery's Γ_2-criterion, which involves the Ricci curvature tensor. We thus need to compute the Ricci curvature tensor $\mathrm{Ric}(G(N,k))$ of $G(N,k)$. Let $\mathfrak{u}(N)$ be the Lie algebra of $\mathrm{U}(N)$ and regard $\mathfrak{h}(N,k) := \mathfrak{u}(k) \oplus \mathfrak{u}(N-k)$ as a Lie subalgebra of $\mathfrak{u}(N)$. The tangent space $T_P G(N,k)$ at each $P \in G(N,k)$ can be identified with $\mathfrak{g}(N,k) := \mathfrak{h}(N,k)^{\perp}$, the orthocomplement of $\mathfrak{h}(N,k)$ in $\mathfrak{u}(N)$ with respect to the Riemannian metric $\langle X, Y\rangle := \mathrm{Re}\,\mathrm{Tr}_N(XY^*)$, where Tr_N is the usual trace on $N \times N$ matrices. Choose the following complete orthonormal system of $\mathfrak{g}(N,k)$:

$$E_{ij} := \frac{1}{\sqrt{2}}(e_{ij} - e_{ji}), \qquad F_{ij} := \frac{\sqrt{-1}}{\sqrt{2}}(e_{ij} + e_{ji}) \tag{2.1}$$

with $1 \le i \le k$, $k+1 \le j \le N$. According to well-known facts on compact matrix groups and O'Neill's formula (see [6], Proposition 3.17 and Theorem 3.61, for example), the Ricci curvature tensor of $G(N,k)$ with respect to the above-mentioned Riemannian metric is computed as follows:

$$\mathrm{Ric}\big(G(N,k)\big)_P(X,X) = \sum_{1\le i\le k,\,k+1\le j\le N} \big(\big\|[X,E_{ij}]\big\|_{HS}^2 + \big\|[X,F_{ij}]\big\|_{HS}^2\big), \quad X \in \mathfrak{g}(N,k).$$

A simple direct computation shows that the above right-hand side is $N\|X\|_{HS}^2$ so that

$$\mathrm{Ric}\big(G(N,k)\big) = N I_{2k(N-k)}. \tag{2.2}$$

Proof of Theorem 2.1. Let α, β and $(\nu, \{\alpha_{ij}\}_{i,j=0}^1)$ be as in Section 1.1 for the given pair (p,q). Since the inequality trivially holds if $\nu = 0$, assume $\nu \ne 0$ and let $\nu_1 := \nu(1)^{-1}\nu$, the normalization of ν. In addition to the assumptions of Lemma 1.2 we first assume the following (A) and (B):

(A) ν is supported in $[\delta, 1-\delta]$ for a sufficiently small $\delta > 0$ and it has the continuous density $\mathrm{d}\nu/\mathrm{d}x$.
(B) The function $Q_{\nu_1}(x) := 2\int_0^1 \log|x-y|\,\mathrm{d}\nu_1(y)$ is a well-defined C^1-function on [0, 1].

Choose C^1-functions $h_0(x)$ and $h_1(x)$ on [0, 1] such that

$$h_0(x)\begin{cases} = \log x & (\delta \le x \le 1), \\ \ge \log x & (0 \le x \le \delta), \end{cases} \qquad h_1(x)\begin{cases} = \log(1-x) & (0 \le x \le 1-\delta), \\ \ge \log(1-x) & (1-\delta \le x \le 1). \end{cases}$$

For each $N \in \mathbb{N}$ choose $k(N), l(N) \in \{1,\dots,N-1\}$ such that $k(N)/N \to \alpha$ and $l(N)/N \to \beta$ as $N \to \infty$, and set

$$\begin{aligned} n_0(N) &:= N - \min\big\{k(N), l(N)\big\}, \\ n_1(N) &:= \max\big\{k(N)+l(N)-N, 0\big\}, \\ n(N) &:= \min\big\{k(N), l(N), N-k(N), N-l(N)\big\} = N - n_0(N) - n_1(N). \end{aligned}$$

Letting

$$\psi_N(x) := \frac{n(N)}{N} Q_{\nu_1}(x) + \frac{|k(N)-l(N)|}{N} h_0(x) + \frac{|k(N)+l(N)-N|}{N} h_1(x), \quad 0 \le x \le 1,$$

we define a probability measure (regarded as a pair of $N \times N$ random projection matrices) $\lambda_N^{\psi_N}$ on $G(N,k(N)) \times G(N,l(N))$ by

$$\mathrm{d}\lambda_N^{\psi_N}(P,Q) := \frac{1}{Z_N^{\psi_N}} \exp\big(-N\mathrm{Tr}_N\big(\psi_N(PQP)\big)\big)\,\mathrm{d}\lambda_N^0(P,Q) \tag{2.3}$$

with the normalization constant $Z_N^{\psi_N}$ and the reference measure $\lambda_N^0 := \gamma_{G(N,k(N))} \otimes \gamma_{G(N,l(N))}$. When $(P,Q) \in G(N,k(N)) \times G(N,l(N))$ is distributed under λ_N^0, the joint eigenvalue distribution of PQP is known due to Ol'shanskij [13] (see also [5]). In the formulation of [8], Eq. (2.1), the eigenvalues of PQP are

$$\underbrace{0,\dots,0}_{n_0(N)\text{ times}},\ \underbrace{1,\dots,1}_{n_1(N)\text{ times}},\ x_1,\dots,x_{n(N)} \tag{2.4}$$

and the joint distribution of $(x_1, \dots, x_{n(N)}) \in [0,1]^{n(N)}$ is

$$d\tilde{\lambda}_N^0(x_1, \dots, x_{n(N)}) := \frac{1}{\widetilde{Z}_N^0} \prod_{i=1}^{n(N)} x_i^{|k(N)-l(N)|}(1-x_i)^{|k(N)+l(N)-N|} \prod_{1\le i<j\le n(N)} (x_i - x_j)^2 \prod_{i=1}^{n(N)} dx_i \tag{2.5}$$

with the normalization constant $\widetilde{Z}_N^0$. Hence it turns out that when $(P, Q) \in G(N, k(N)) \times G(N, l(N))$ is distributed under $\lambda_N^{\psi_N}$, the eigenvalues of PQP are listed as in (2.4) but the joint distribution of $(x_1, \dots, x_{n(N)}) \in [0,1]^{n(N)}$ is changed to

$$\tilde{\lambda}_N^{\psi_N}(x_1, \dots, x_{n(N)}) := \frac{1}{\widetilde{Z}_N^{\psi_N}} \exp\bigl(-\tilde{\psi}_N(x_1, \dots, x_{n(N)})\bigr) \prod_{1\le i<j\le n(N)} (x_i - x_j)^2 \prod_{i=1}^{n(N)} dx_i \tag{2.6}$$

with the new normalization constant $\widetilde{Z}_N^{\psi_N}$, where

$$\begin{aligned}\tilde{\psi}_N(x_1, \dots, x_{n(N)}) := \sum_{i=1}^{n(N)} \bigl\{ & n(N) Q_{\nu_1}(x_i) + \bigl|k(N) - l(N)\bigr|\bigl(h_0(x_i) - \log x_i\bigr) \\ & + \bigl|k(N) + l(N) - N\bigr|\bigl(h_1(x_i) - \log(1 - x_i)\bigr)\bigr\}.\end{aligned}$$

Similarly to [8], Proposition 2.1 and [7], Section 5.5, we have:

(a) The limit $C' := \lim_{N\to\infty} \frac{1}{N^2} \log \widetilde{Z}_N^{\psi_N}$ exists as well as $C = \lim_{N\to\infty} \frac{1}{N^2} \log \widetilde{Z}_N^0$ (see (1.2)).

(b) When $(x_1, \dots, x_{n(N)})$ is distributed under $\tilde{\lambda}_N^{\psi_N}$, the empirical measure $\frac{1}{n(N)} \sum_{i=1}^{n(N)} \delta_{x_i}$ satisfies the large deviation principle in the scale $1/N^2$ with the rate function

$$I(\mu) := -\rho^2 \Sigma(\mu) + \rho^2 \int_0^1 F(x)\, d\mu(x) + C' \quad \text{for } \mu \in \mathcal{M}\bigl([0,1]\bigr),$$

where $\mathcal{M}([0,1])$ is the set of probability measures on $[0,1]$, ρ is given in (1.1) and

$$F(x) := Q_{\nu_1}(x) + \frac{|\alpha - \beta|}{\rho}\bigl(h_0(x) - \log x\bigr) + \frac{|\alpha + \beta - 1|}{\rho}\bigl(h_1(x) - \log(1 - x)\bigr)$$

for $0 \le x \le 1$.

(c) ν_1 is a unique minimizer of I with $I(\nu_1) = 0$.

The last assertion follows from [14], Theorems I.1.3 and I.3.1, because by the construction of h_0 and h_1 we get

$$Q_{\nu_1}(x) \begin{cases} = F(x) & \text{if } x \in [\delta, 1-\delta] \supset \operatorname{supp} \nu_1, \\ \le F(x) & \text{for } x \in [0,1]. \end{cases}$$

Furthermore the above large deviation yields:

(d) The mean eigenvalue distribution $\hat{\lambda}_N^{\psi_N} := \int_{[0,1]^{n(N)}} \frac{1}{n(N)} \sum_{i=1}^{n(N)} \delta_{x_i}\, d\tilde{\lambda}_N^{\psi_N}(x_1, \dots, x_{n(N)})$ weakly converges to ν_1 as $N \to \infty$.

Since the Riemannian manifold $G(N, k(N)) \times G(N, l(N))$ has the volume measure λ_N^0 and its Ricci curvature tensor is $N I_{2k(N-k)+2l(N-l)}$ by (2.2), the classical logarithmic Sobolev inequality due to Bakry and Emery [1] implies that

$$S\bigl(\lambda_N^{\psi_N}, \lambda_N^0\bigr) \le \frac{1}{2N} \int_{G(N,k(N))\times G(N,l(N))} \left\| \nabla \log \frac{d\lambda_N^{\psi_N}}{d\lambda_N^0} \right\|_{HS}^2 d\lambda_N^{\psi_N}, \tag{2.7}$$

where the left-hand side is the relative entropy of $\lambda_N^{\psi_N}$ with respect to λ_N^0 and the gradient $\nabla\log(\mathrm{d}\lambda_N^{\psi_N}/\mathrm{d}\lambda_N^0)(P,Q)$ is considered in $\mathfrak{g}(N,k(N))\oplus\mathfrak{g}(N,l(N))$ via the natural identification $T_{(P,Q)}G(N,k(N))\times G(N,l(N))=\mathfrak{g}(N,k(N))\oplus\mathfrak{g}(N,l(N))$. By (2.5) and (2.6) notice that

$$\frac{\mathrm{d}\lambda_N^{\psi_N}}{\mathrm{d}\lambda_N^0}(P,Q)=\frac{1}{Z_N^{\psi_N}}\exp\bigl(-N\mathrm{Tr}_N\bigl(\psi_N(PQP)\bigr)\bigr)=\frac{\widetilde{Z}_N^0}{\widetilde{Z}_N^{\psi_N}}\exp\left(-N\sum_{i=1}^{n(N)}\psi_N(x_i)\right)$$

for $(P,Q)\in G(N,k(N))\times G(N,l(N))$ and for the eigenvalues $(x_1,\dots,x_{n(N)})$ of PQP except $n_0(N)$ zeros and $n_1(N)$ ones (see (2.4)). Hence we get

$$\begin{aligned}S\bigl(\lambda_N^{\psi_N},\lambda_N^0\bigr)&=\int_{G(N,k(N))\times G(N,l(N))}\log\frac{\mathrm{d}\lambda_N^{\psi_N}}{\mathrm{d}\lambda_N^0}(P,Q)\,\mathrm{d}\lambda_N^{\psi_N}(P,Q)\\&=\log\widetilde{Z}_N^0-\log\widetilde{Z}_N^{\psi_N}-Nn(N)\int_0^1\psi_N(x)\,\mathrm{d}\hat{\lambda}_N^{\psi_N}(x).\end{aligned}$$

Since $\psi_N(x)$ converges to $\rho Q_{\nu_1}(x)+|\alpha-\beta|h_0(x)+|\alpha+\beta-1|h_1(x)$ uniformly on [0, 1], it follows from (a) and (d) above that

$$\lim_{N\to\infty}\frac{1}{N^2}S\bigl(\lambda_N^{\psi_N},\lambda_N^0\bigr)=C-C'-\rho\int_0^1\bigl(\rho Q_{\nu_1}(x)+|\alpha-\beta|h_0(x)+|\alpha+\beta-1|h_1(x)\bigr)\,\mathrm{d}\nu_1(x).$$

Since (c) gives

$$-C'=-\rho^2\Sigma(\nu_1)+\rho^2\int_0^1F(x)\,\mathrm{d}\nu_1(x)=-\rho^2\Sigma(\nu_1)+\rho^2\int_0^1Q_{\nu_1}(x)\,\mathrm{d}\nu_1(x),$$

we have

$$\lim_{N\to\infty}\frac{1}{N^2}S\bigl(\lambda_N^{\psi_N},\lambda_N^0\bigr)=-\chi_{\mathrm{proj}}(p,q)\tag{2.8}$$

thanks to Lemma 1.1 and (1.3) together with $\nu_1=(2\rho)^{-1}\nu$ (see [8], Eq. (3.4)).

On the other hand, since $\nabla\log(\mathrm{d}\lambda_N^{\psi_N}/\mathrm{d}\lambda_N^0)(P,Q)=-N\nabla(\mathrm{Tr}_N(\psi_N(PQP)))$, one can compute

$$\left\|\nabla\log\frac{\mathrm{d}\lambda_N^{\psi_N}}{\mathrm{d}\lambda_N^0}(P,Q)\right\|_{HS}^2=4N^2\mathrm{Tr}_N\bigl(\bigl(\psi_N'(PQP)\bigr)^2PQP(I-PQP)\bigr),$$

whose proof will be given as Lemma 2.2 below for completeness. Therefore,

$$\begin{aligned}&\int_{G(N,k(N))\times G(N,l(N))}\left\|\nabla\log\frac{\mathrm{d}\lambda_N^{\psi_N}}{\mathrm{d}\lambda_N^0}(P,Q)\right\|_{HS}^2\mathrm{d}\lambda_N^{\psi_N}(P,Q)\\&\quad=4N^2\int_{[0,1]^{n(N)}}\sum_{i=1}^{n(N)}\bigl(\psi_N'(x_i)\bigr)^2x_i(1-x_i)\,\mathrm{d}\tilde{\lambda}_N^{\psi_N}(x_1,\dots,x_{n(N)})\\&\quad=4N^2n(N)\int_0^1\bigl(\psi_N'(x)\bigr)^2x(1-x)\,\mathrm{d}\hat{\lambda}_N^{\psi_N}(x)\\&\quad=4n(N)\int_0^1\bigl(n(N)Q_{\nu_1}'(x)+\bigl|k(N)-l(N)\bigr|h_0'(x)+\bigl|k(N)+l(N)-1\bigr|h_1'(x)\bigr)^2x(1-x)\,\mathrm{d}\hat{\lambda}_N^{\psi_N}(x),\end{aligned}$$

and thus by (d) we have

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{2N^3}\int_{G(N,k(N))\times G(N,l(N))}\left\|\nabla\log\frac{\mathrm{d}\lambda_N^{\psi_N}}{\mathrm{d}\lambda_N^0}\right\|_{HS}^2\mathrm{d}\lambda_N^{\psi_N}\\
&\quad=2\rho\int_0^1\left(\rho Q'_{\nu_1}(x)+|\alpha-\beta|h'_0(x)+|\alpha+\beta-1|h'_1(x)\right)^2x(1-x)\,\mathrm{d}\nu_1(x)\\
&\quad=\int_0^1\left(\rho Q'_{\nu_1}(x)+\frac{|\alpha-\beta|}{x}-\frac{|\alpha+\beta-1|}{1-x}\right)^2x(1-x)\,\mathrm{d}\nu(x)\\
&\quad=\varphi^*(p:q)
\end{aligned}\tag{2.9}$$

by Lemma 1.2, since $\nu_1=(2\rho)^{-1}\nu$ so that $\rho Q'_{\nu_1}(x)=(Hf)(x)$ with $f:=\mathrm{d}\nu/\mathrm{d}x$. Combining (2.7)–(2.9) yields the desired inequality under (A) and (B).

Next let us remove (A) and (B). First suppose that (A) is still satisfied but (B) is not. For each $\varepsilon>0$ choose a non-negative C^∞-function ψ_ε supported in $[-\varepsilon,\varepsilon]$ with $\int\psi_\varepsilon(x)\,\mathrm{d}x=1$. Let $f_\varepsilon:=f*\psi_\varepsilon$ for $f:=\mathrm{d}\nu/\mathrm{d}x$ and define $\mathrm{d}\nu_\varepsilon(x):=f_\varepsilon(x)\,\mathrm{d}x$; then ν_ε is a measure supported in a closed proper subinterval of $(0,1)$ with $\nu_\varepsilon((0,1))=1-\sum_{i,j=0}^1\alpha_{ij}$ whenever ε is small enough. Let $(p_\varepsilon,q_\varepsilon)$ be a pair of projections in some $(\mathcal{M},\tau)$ corresponding to the representing data $(\nu_\varepsilon,\{\alpha_{ij}\}_{i,j=0}^1)$. (Such a pair can be constructed via the GNS representation associated with the tracial state determined by $(\nu_\varepsilon,\{\alpha_{ij}\}_{i,j=0}^1)$; see [8], Section 3.) Since (A) and (B) are satisfied for ν_ε, we get $-\chi_{\mathrm{proj}}(p_\varepsilon,q_\varepsilon)\le\varphi^*(p_\varepsilon:q_\varepsilon)$. Since $\|f_\varepsilon-f\|_{w,3}\to0$ and $\|Hf_\varepsilon-Hf\|_{w,3}\to0$ as $\varepsilon\searrow0$ (see the proof of Lemma 1.2 for the weighted norm $\|\cdot\|_{w,3}$), the Hölder inequality together with (1.5) implies that

$$\int_0^1\left((Hf_\varepsilon)(x)\right)^2f_\varepsilon(x)x(1-x)\,\mathrm{d}x\longrightarrow\int_0^1\left((Hf)(x)\right)^2f(x)x(1-x)\,\mathrm{d}x\tag{2.10}$$

and hence $\lim_{\varepsilon\searrow0}\varphi^*(p_\varepsilon:q_\varepsilon)=\varphi^*(p:q)$. Since $\Sigma(\mu)$ is weakly upper semicontinuous in μ (see, e.g. [7], Proposition 5.3.2), we also have $-\chi_{\mathrm{proj}}(p,q)\le\liminf_{\varepsilon\searrow0}(-\chi_{\mathrm{proj}}(p_\varepsilon,q_\varepsilon))$ so that $-\chi_{\mathrm{proj}}(p,q)\le\varphi^*(p:q)$.

Finally suppose only the assumptions stated in Lemma 1.2. For $\delta>0$ set

$$\mathrm{d}\nu_\delta(s):=\frac{1-\sum_{i,j=0}^1\alpha_{ij}}{\nu([\delta,1-\delta])}\mathbf{1}_{[\delta,1-\delta]}(x)\,\mathrm{d}\nu(x)$$

and let (p_δ,q_δ) be a pair of projections corresponding to $(\nu_\delta,\{\alpha_{ij}\}_{i,j=0}^1)$. Let us denote the density of ν_δ by f_δ; then it is immediate to see that $\|f_\delta-f\|_{w,3}\to0$. To show that $\varphi^*(p_\delta:q_\delta)\to\varphi^*(p:q)$ as $\delta\searrow0$, it suffices to prove the following convergences as $\delta\searrow0$:

$$\int_0^1\left((Hf_\delta)(x)\right)^2f_\delta(x)x(1-x)\,\mathrm{d}x\longrightarrow\int_0^1\left((Hf)(x)\right)^2f(x)x(1-x)\,\mathrm{d}x,\tag{2.11}$$

$$\int_0^1(Hf_\delta)(x)x^{-1}f_\delta(x)\,x(1-x)\,\mathrm{d}x\longrightarrow\int_0^1(Hf)(x)x^{-1}f(x)\,x(1-x)\,\mathrm{d}x,\tag{2.12}$$

$$\int_0^1(Hf_\delta)(x)(1-x)^{-1}f_\delta(x)\,x(1-x)\,\mathrm{d}x\longrightarrow\int_0^1(Hf)(x)(1-x)^{-1}f(x)\,x(1-x)\,\mathrm{d}x,\tag{2.13}$$

$$\int_0^1x^{-2}f_\delta(x)\,x(1-x)\,\mathrm{d}x\longrightarrow\int_0^1x^{-2}f(x)\,x(1-x)\,\mathrm{d}x,\tag{2.14}$$

$$\int_0^1(1-x)^{-2}f_\delta(x)\,x(1-x)\,\mathrm{d}x\longrightarrow\int_0^1(1-x)^{-2}f(x)\,x(1-x)\,\mathrm{d}x.\tag{2.15}$$

Remark here that (2.12) and (2.14) are unnecessary when $\alpha_{01}+\alpha_{10}=|\alpha-\beta|=0$, and so are (2.13) and (2.15) when $\alpha_{11}+\alpha_{00}=|\alpha+\beta-1|=0$. The convergence (2.11) is verified as (2.10) above. Also, (2.14) and (2.15) immediately follow from the hypothesis (1.4). Since (2.12) and (2.13) are similarly shown, let us prove only the former here. Thus we should assume $\alpha \neq \beta$, and (1.4) means $\int_0^1 x^{-1} f(x)\,\mathrm{d}x < +\infty$. By the Hölder inequality together with (1.5) one can estimate

$$\begin{aligned}\left\|(Hf_\delta)x^{-1}f_\delta - (Hf)x^{-1}f\right\|_{w,1} &\le C_w\|f_\delta - f\|_{w,3}\cdot\left\|x^{-1}f_\delta^{1/2}\right\|_{w,2}\cdot\|f_\delta\|_{w,3}^{1/2}\\ &\quad + C_w\|f\|_{w,3}\cdot\left\|x^{-1}|f_\delta - f|^{1/2}\right\|_{w,2}\cdot\|f_\delta - f\|_{w,3}^{1/2}.\end{aligned}$$

Note that

$$\left\|x^{-1}f_\delta^{1/2}\right\|_{w,2}^2 \le \int_0^1 x^{-1}f_\delta(x)\,\mathrm{d}x \longrightarrow \int_0^1 x^{-1}f(x)\,\mathrm{d}x,$$

$$\left\|x^{-1}|f_\delta - f|^{1/2}\right\|_{w,2}^2 \le \int_0^1 x^{-1}\left|f_\delta(x) - f(x)\right|\mathrm{d}x \longrightarrow 0$$

as $\delta \searrow 0$, since $\int_0^1 x^{-1} f(x)\,\mathrm{d}x < +\infty$. These apparently imply (2.12) thanks to $f \in L^3((0,1), x(1-x)\,\mathrm{d}x)$ and $\|f_\delta - f\|_{w,3} \longrightarrow 0$.

Moreover, since $-\log x < x^{-1}$ near 0 and $-\log(1-x) < (1-x)^{-1}$ near 1, the hypothesis (1.4) implies that

$$\int_0^1 (-\log x) f_\delta(x)\,\mathrm{d}x \longrightarrow \int_0^1 (-\log x) f(x)\,\mathrm{d}x < +\infty,$$

$$\int_0^1 \bigl(-\log(1-x)\bigr) f_\delta(x)\,\mathrm{d}x \longrightarrow \int_0^1 \bigl(-\log(1-x)\bigr) f(x)\,\mathrm{d}x < +\infty$$

as $\delta \searrow 0$ (whenever those are needed) so that $-\chi_{\mathrm{proj}}(p,q) \le \liminf_{\delta\searrow 0}(-\chi_{\mathrm{proj}}(p_\delta, q_\delta))$. Hence the proof is completed. □

Lemma 2.2. *Let ψ be a C^1-function on $[0,1]$ and define $\Psi(P,Q) := \mathrm{Tr}_N(\psi(PQP))$ for $(P,Q) \in G(N,k) \times G(N,l)$. Then*

$$\left\|\nabla\Psi(P,Q)\right\|_{HS}^2 = 4\mathrm{Tr}_N\bigl(\bigl(\psi'(PQP)\bigr)^2 PQP(I - PQP)\bigr)$$

holds for every $(P,Q) \in G(N,k) \times G(N,l)$.

Proof. Write $(X_r)_{r=1}^{2k(N-k)}$ for the orthonormal basis of $\mathfrak{g}(N,k)$ given in (2.1) and also $(Y_s)_{s=1}^{2l(N-l)}$ for that of $\mathfrak{g}(N,l)$. For each $(P,Q) = (UP_N(k)U^*, VP_N(l)V^*)$ in $G(N,k) \times G(N,l)$, a local normal coordinate at (P,Q) is given by the mapping

$$(X,Y) \in \mathfrak{g}(N,k) \oplus \mathfrak{g}(N,l) \mapsto \bigl(Ue^X P_N(k)e^{-X}U^*, Ve^Y P_N(l)e^{-Y}V^*\bigr) \in G(N,k) \times G(N,l).$$

By a direct computation using this coordinate, one can compute

$$\begin{aligned}\nabla\Psi(P,Q) &= \sum_r \bigl\langle U^*QPf'(PQP)PU - U^*Pf'(PQP)PQU, X_r\bigr\rangle X_r\\ &\quad + \sum_s \bigl\langle V^*Pf'(PQP)PQV - V^*QPf'(PQP)PV, Y_s\bigr\rangle Y_s\end{aligned}$$

so that

$$\begin{aligned}\left\|\nabla\Psi(P,Q)\right\|_{HS}^2 &= 2\bigl\|P_k(N)U^*Pf'(PQP)PQU\bigl(I - P_k(N)\bigr)\bigr\|_{HS}^2\\ &\quad + 2\bigl\|P_l(N)V^*QPf'(PQP)PV\bigl(I - P_l(N)\bigr)\bigr\|_{HS}^2\\ &= 4\mathrm{Tr}_N\bigl(\bigl(f'(PQP)\bigr)^2 PQP(I - PQP)\bigr).\end{aligned}$$

□

3. Remarks

3.1. *Classical vs. free probabilistic mutual information*

The classical mutual information of two random variables, say X, Y, is usually formulated to be

$$I(X;Y) := S(\mu_{(X,Y)}, \mu_X \otimes \mu_Y) = \int_{\mathcal{X}\times\mathcal{X}} \log \frac{d\mu_{(X,Y)}}{d(\mu_X \otimes \mu_Y)}(x, y)\, d\mu_{(X,Y)}(x, y),$$

where μ_X, μ_Y are the distribution measures of X, Y on the phase space $\mathcal{X}$ and $\mu_{(X,Y)}$ the joint distribution of (X, Y) on $\mathcal{X} \times \mathcal{X}$. The mutual information is in turn written as $I(X;Y) = H(X) + H(Y) - H(X, Y)$ as long as all the Shannon–Gibbs entropies $H(X)$, $H(Y)$ and $H(X, Y)$ are finite. This is nothing but what Voiculescu mentioned in [20] as an initial motivation of his introduction of the liberation theory. Let us now apply the definition of $I(X;Y)$ to our random matrix model $(P(N), Q(N))$ of a given pair (p, q) of projections, and then the proof of Theorem 2.1 (see (2.8)) shows that

$$-\chi_{\mathrm{proj}}(p, q) = \lim_{N\to\infty} \frac{1}{N^2} S\big(\lambda_N^{\psi_N}, \lambda_N^0\big) = \lim_{N\to\infty} \frac{1}{N^2} I\big(P(N);Q(N)\big).$$

3.2. $-\chi_{\mathrm{proj}} = i^*$ *for two projections*

We simply write $i^*(p : q)$ for the *free mutual information* $i^*(\mathbb{C}p + \mathbb{C}(1 - p) : \mathbb{C}q + \mathbb{C}(1 - q))$ introduced in [20], and its "heuristic definition" should be "$\chi_{\mathrm{proj}}(p) + \chi_{\mathrm{proj}}(q) - \chi_{\mathrm{proj}}(p, q)$" on the analogy of classical theory. However the actual definition is completely different based on the so-called liberation process so that in view of $\chi_{\mathrm{proj}}(p) = \chi_{\mathrm{proj}}(q) = 0$ it may be particularly interesting to examine whether $i^*(p : q)$ coincides with $-\chi_{\mathrm{proj}}(p, q)$ or not. In fact, the inequality in Theorem 2.1 is a kind of logarithmic Sobolev inequality and its right-hand side (the Dirichlet form part) is a "derivative" of $i^*(p : q)$, which also gives us a strong reason for that question. The equality $-\chi_{\mathrm{proj}}(p, q) = i^*(p : q)$ is technically similar to $\chi(X) = \chi^*(X)$ in the single variable case shown in [19], though the former is much more involved. (Here it should be noted that $\chi \le \chi^*$ holds in general [4].) At the moment we can give only a heuristic argument for the question. Let $(\nu_t, \{\alpha_{ij}(t)\}_{i,j=0}^1)$ be the representing data of the liberation process $(p(t) := u(t)pu(t)^*, q)$ started at (p, q) with a free unitary Brownian motion $\{u(t)\}_{t\ge 0}$ (see [2]) freely independent of (p, q). Remark that the $\alpha_{ij}(t)$'s are constant in t due to [20], Lemma 12.5. Assume that ν_t satisfies the same properties as in Lemma 1.2 for every $t \ge 0$. By [20], Corollary 5.7, one has

$$\tau\big(\big(p(t+\varepsilon)qp(t+\varepsilon)\big)^m\big) = \tau\big(\big(p(t)qp(t)\big)^m\big) + \frac{m\varepsilon}{2}\tau\big(\big[J_t, p(t)\big]\big(qp(t)q\big)^{m-1}\big) + O\big(\varepsilon^2\big)$$

with the liberation gradient $J_t := j(\mathbb{C}p(t) + \mathbb{C}(\mathbf{1} - p(t)) : \mathbb{C}q + \mathbb{C}(\mathbf{1} - q))$, and hence letting $d\nu_t(x) = f_t(x)\,dx$ one can derive

$$\frac{\partial}{\partial t} f_t(x) = -\frac{\partial}{\partial x}\left[x(1-x)f_t(x)\left((Hf_t)(x) + \frac{\alpha_{01}+\alpha_{10}}{x} - \frac{\alpha_{00}+\alpha_{11}}{1-x}\right)\right] \tag{3.1}$$

under the additional assumption that $f_t(x)$ is smooth in (t, x). The differential formula $\frac{d}{dt}\chi_{\mathrm{proj}}(p(t), q) = \varphi^*(p(t) : q)$ shows up after a rather heuristic computation by using (3.1). Then one can show $\lim_{t\to\infty} \chi_{\mathrm{proj}}(p(t), q) = 0$ by using Theorem 2.1 and [20], Proposition 10.11.c, so that the desired $-\chi_{\mathrm{proj}}(p, q) = i^*(p : q)$ follows. The insufficient points in the derivation outlined above are: (i) we need to prove that all ν_t, $t > 0$, automatically satisfy the same properties as in Lemma 1.2 whenever ν_0 is assumed to satisfy; (ii) we need to prove that $f_t(x)$ is smooth in (t, x); and finally (iii) we have to give a rigorous derivation of $\frac{d}{dt}\chi_{\mathrm{proj}}(p(t), q) = \varphi^*(p(t) : q)$ from (3.1).

Acknowledgment

The authors thank Professor Masaki Izumi for fruitful discussions. They also thank the referees for comments.

References

[1] D. Bakry and M. Emery. Diffusion hypercontractives. *Séminaire Probabilités XIX* 177–206. Lecture Notes in Math. **1123**. Springer, Berlin, 1985. MR0889476

[2] P. Biane. Free Brownian motion, free stochastic calculus and random matrices. In *Free Probability Theory* 1–19. D. V. Voiculescu (Ed.). Fields Inst. Commun. **12**. Amer. Math. Soc. Providence, RI, 1997. MR1426833

[3] P. Biane. Logarithmic Sobolev inequalities, matrix models and free entropy. *Acta Math. Sinica* **19** (2003) 497–506. MR2014030

[4] P. Biane, M. Capitaine and A. Guionnet. Large deviation bounds for matrix Brownian motion. *Invent. Math.* **152** (2003) 433–459. MR1975007

[5] B. Collins. Product of random projections, Jacobi ensembles and universality problems arising from free probability. *Probab. Theory Related Fields* **133** (2005) 315–344. MR2198015

[6] S. Gallot, D. Hulin and J. Lafontaine. *Riemannian Geometry*, 2nd edition. Universitext, Springer, Berlin, 1990. MR1083149

[7] F. Hiai and D. Petz. *The Semicircle Law, Free Random Variables and Entropy*. Amer. Math. Soc., Providence, RI, 2000. MR1746976

[8] F. Hiai and D. Petz. Large deviations for functions of two random projection matrices. *Acta Sci. Math. (Szeged)* **72** (2006) 581–609. MR2289756

[9] F. Hiai, D. Petz and Y. Ueda. Free logarithmic Sobolev inequality on the unit circle. *Canad. Math. Bull*. **49** (2006) 389–406. MR2252261

[10] F. Hiai and Y. Ueda. Notes on microstate free entropy of projections. *Publ. Res. Inst. Math. Sci.* **44** (2008), 49–89.

[11] R. Hunt, B. Muckenhoupt and R. Wheeden. Weighted norm inequalities for the conjugate function and Hilbert transform. *Trans. Amer. Math. Soc.* **176** (1973) 227–251. MR0312139

[12] M. Ledoux. A (one-dimensional) free Brunn–Minkowski inequality. *C. R. Math. Acad. Sci. Paris* **340** (2005) 301–304. MR2121895

[13] G. I. Ol'shanskij. Unitary representations of infinite dimensional pairs (g, k) and the formalism of R. Howe. In *Representation of Lie Groups and Related Topics* 269–463. A. M. Vershik and D. P. Zhelobenko (Eds). *Adv. Stud. Contemp. Math.* **7**. Gordon and Breach, New York, 1990. MR1104279

[14] E. B. Saff and V. Totik. *Logarithmic Potentials with External Fields*. Springer, Berlin, 1997. MR1485778

[15] C. Villani. *Topics in Optimal Transportation*. Amer. Math. Soc., Providence, RI, 2003. MR1964483

[16] D. Voiculescu. The analogues of entropy and of Fisher's information measure in free probability theory, I. *Comm. Math. Phys.* **155** (1993) 71–92. MR1228526

[17] D. Voiculescu. The analogues of entropy and of Fisher's information measure in free probability theory, II. *Invent. Math.* **118** (1994) 411–440. MR1296352

[18] D. Voiculescu. The analogues of entropy and of Fisher's information measure in free probability theory, IV: Maximum entropy and freeness. In *Free Probability Theory* 293–302. D. V. Voiculescu (Ed.). Fields Inst. Commun. **12**. Amer. Math. Soc., Providence, RI, 1997. MR1426847

[19] D. Voiculescu. The analogues of entropy and of Fisher's information measure in free probability theory, V: Noncommutative Hilbert transforms. *Invent. Math.* **132** (1998) 189–227. MR1618636

[20] D. Voiculescu. The analogue of entropy and of Fisher's information measure in free probability theory VI: Liberation and mutual free information. *Adv. Math.* **146** (1999) 101–166. MR1711843

[21] D. V. Voiculescu, K. J. Dykema and A. Nica. *Free Random Variables*. Amer. Math. Soc., Providence, RI, 1992. MR1217253