ORBITAL APPROACH TO MICROSTATE FREE ENTROPY

FUMIO HIAI [1], TAKUHO MIYAMOTO, AND YOSHIMICHI UEDA [2]

Abstract. Motivated by Voiculescu's liberation theory, we introduce the orbital free entropy $\chi_{\rm orb}$ for non-commutative self-adjoint random variables (also for "hyperfinite random multi-variables"). Besides its basic properties the relation of $\chi_{\rm orb}$ with the usual free entropy χ is shown. Moreover, the dimension counterpart $\delta_{0,{\rm orb}}$ of $\chi_{\rm orb}$ is discussed, and we obtain the relation of $\delta_{0,{\rm orb}}$ with the original free entropy dimension δ_0 with applications to δ_0 itself.

Introduction

We propose a somewhat new approach to Voiculescu's theory of free entropy (see e.g., [27] for a survey), and introduce the orbital free entropy $\chi_{\rm orb}(X_1,\dots,X_n)$. This quantity is an extension of the projection free entropy $\chi_{\rm proj}(p_1,\dots,p_n)$ studied in [11] following Voiculescu's proposal in [26]. Our essential idea is to restrict microstates for $(X_1,\dots,X_n)$ to unitary-orbital ones, that is, to use only the unitary parts of microstates with disregarding the diagonal parts under their diagonalization. We prove the exact relation (Theorem 2.6)

$$\chi(X_1,\dots,X_n)=\chi_{\rm orb}(X_1,\dots,X_n)+\sum_{i=1}^n\chi(X_i) \tag{0.1}$$

between $\chi_{\rm orb}$ and the usual free entropy χ as naturally expected from the definition. The particular formula

$$-\chi_{\rm orb}(X,Y)=-\chi(X,Y)+\chi(X)+\chi(Y)$$

resembles the expression $I(X;Y)=-H(X,Y)+H(X)+H(Y)$ of the classical mutual information in terms of the Boltzmann-Gibbs entropy $H(\cdot)$. It should be emphasized here that this expression of $I(X;Y)$ motivated Voiculescu to develop his liberation theory and introduce the mutual free information i^* in [26]. In this way, we may regard $-\chi_{\rm orb}(X,Y)$ as a kind of free analog of the mutual information and also as one possible microstate version of the mutual free information $i^*(W^*(X);W^*(Y))$. Among other properties, we prove (Theorem 3.1) that $\chi_{\rm orb}(X_1,\dots,X_n)=0$ if and only if $X_1,\dots,X_n$ are freely independent without any extra assumption. Together with the relation (0.1) this directly implies the characterization [22, 24] of freeness by the additivity of χ. The proof of the theorem is based on a certain transportation cost inequality as in the projection case in [11]. An advantage of our orbital approach is that one of equivalent definitions of $\chi_{\rm orb}$ as well as its all properties is valid even for non-commutative random multi-variables $\mathbf{X}_1,\dots,\mathbf{X}_n$ each of which generates a hyperfinite von Neumann algebra $W^*(\mathbf{X}_i)$. But the present definition obeys the essential restriction of hypefiniteness due to Jung's result [15] (or Lemma 1.2 below).

[1] Supported in part by Grant-in-Aid for Scientific Research (B)17340043.
[2] Supported in part by Grant-in-Aid for Young Scientists (B)17740096.
AMS subject classification: Primary: 46L54; secondary: 52C17, 28A78, 94A17.

Furthermore, we study the dimension counterpart $\delta_{0,\mathrm{orb}}$ of χ_{orb} for hyperfinite random multi-variables $\mathbf{X}_1, \dots, \mathbf{X}_n$. It is defined similarly to the modified free entropy dimension δ_0 with replacing the semicircular deformation by the liberation process ([26]). The $\delta_{0,\mathrm{orb}}$ enjoys properties similar to δ_0; for example, $\delta_{0,\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n) = 0$ if $\chi_{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n) > -\infty$ (in particular, this is the case if $\mathbf{X}_1, \dots, \mathbf{X}_n$ are free). Moreover, we prove the covering/packing formula of $\delta_{0,\mathrm{orb}}$ based on Jung's approach [12, 13] to δ_0, and furthermore we prove the following general formula:

$$\delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n) = \delta_{0,\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n) + \sum_{i=1}^{n} \delta_0(\mathbf{X}_i).$$

The orbital theory developed in this paper has several applications to the original free entropy dimension δ_0 itself. Among others, the most important one is the following lower semicontinuity result for δ_0: Let $\mathbf{X}_1, \dots, \mathbf{X}_n$ be hyperfinite random multi-variables, and assume, for each $1 \le i \le n$, that we have a sequence of hyperfinite random multi-variables $\mathbf{X}_i^{(k)}$ converging to $\mathbf{X}_i$ in moments. In this setup, we will see that, if $\mathbf{X}_i^{(k)} \subset W^*(\mathbf{X}_i)$ is further assumed for every $k \in \mathbb{N}$ and $1 \le i \le n$, then

$$\liminf_{k\to\infty} \delta_0(\mathbf{X}_1^{(k)} \sqcup \cdots \sqcup \mathbf{X}_n^{(k)}) \ge \delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n).$$

Note here that the lower semicontinuity of δ_0 was shown by Voiculescu in the single variable case, see [27] for the history on the semicontinuity problem of δ_0 until that time. Also, a certain related result was obtained by Jung [12, Lemma 7.3] based on his result on the δ_0 of hyperfinite algebras. However, Shlyakhtenko [18] pointed out that δ_0 is never lower semicontinuous in general. Our result is probably the first affirmative semicontinuity result for δ_0 of non-commutative nature.

Acknowledgment. This work was completed during the third-named author's stay in the Fields institute in 2007. He would like to thank the institute for hospitality and also acknowledge Professor George Elliott for inviting him to the operator algebra program.

1. Preliminaries

1.1. **Notations.** We will use the standard notations; $M_N(\mathbb{C})$ denotes the $N \times N$ matrix algebra, Tr_N stands for the usual (non-normalized) trace and $\mathrm{tr}_N := N^{-1}\mathrm{Tr}_N$ for its normalization. The operator norm of a bounded operator a is denoted by $\|a\|_\infty$. For a (tracial) state ϕ on a C^*-algebra A we write $\|a\|_{p,\phi} := \phi(|a|^p)^{1/p}$ for $a \in A$ and $1 \le p < \infty$. Let M_N^{sa} be the set of all $N \times N$ self-adjoint matrices. Let $\mathrm{U}(N)$ be the $N \times N$ unitary group and $\mathrm{T}(N)$ be its (standard) maximal torus consisting of all diagonal matrices in $\mathrm{U}(N)$ (isomorphic to the N-dimensional torus $\mathbb{T}^N$ in the obvious way). We also use the $N \times N$ special unitary group $\mathrm{SU}(N)$ in some places. The canonical quotient map from $\mathrm{U}(N)$ onto the left coset space $\mathrm{U}(N)/\mathrm{T}(N)$ is denoted by q_N^{U}. We denote by $\gamma_{\mathrm{U}(N)}$ the Haar probability measures on $\mathrm{U}(N)$ and by $\gamma_{\mathrm{U}(N)/\mathrm{T}(N)}$ the probability measure on $\mathrm{U}(N)/\mathrm{T}(N)$ induced from $\gamma_{\mathrm{U}(N)}$, that is, $\gamma_{\mathrm{U}(N)/\mathrm{T}(N)} = \gamma_{\mathrm{U}(N)} \circ (q_N^{\mathrm{U}})^{-1}$. The Haar probability measure on $\mathrm{SU}(N)$ is similarly denoted by $\gamma_{\mathrm{SU}(N)}$.

1.2. **Matricial microstate spaces.** The following measure-space isomorphism is well-known:

$$\Phi_N : \big((\mathrm{U}(N)/\mathrm{T}(N)) \times \mathbb{R}^N_{\ge}, \gamma_{\mathrm{U}(N)/\mathrm{T}(N)} \otimes \mu_N\big) \cong \big(M_N^{sa}, \Lambda_N\big), \tag{1.1}$$

where

- $\mathbb{R}^N_{\ge}$ denotes all $(x_1, \dots, x_N) \in \mathbb{R}^N$ with $x_1 \ge x_2 \ge \cdots \ge x_N$,
- $\mu_N := C_N \prod_{i<j}(x_i - x_j)^2 \prod_{i=1}^{n} dx_i$ with $C_N := \dfrac{(2\pi)^{N(N-1)/2}}{\prod_{k=1}^{N-1} k!}$,

- Λ_N is the usual Lebesgue measure on $M_N^{sa} \cong \mathbb{R}^{N^2}$.

The map Φ_N is given by the continuous surjection

$$([U], D) \in (\mathrm{U}(N)/\mathrm{T}(N)) \times \mathbb{R}^N_{\geq} \mapsto UDU^* \in M_N^{sa}$$

with identifying $(x_1, \dots, x_N) \in \mathbb{R}^N_{\geq}$ with the diagonal matrices D whose diagonal entries are $x_1, \dots, x_N$ in decreasing order from the upper diagonal corner. In what follows, we will identify an element in $\mathbb{R}^N$ with a diagonal matrix in this canonical way. The above description of M_N^{sa} will be a key of our approach.

1.3. **Technical lemmas.** Here we recall three well-known lemmas, which will be main technical ingredients in our discussions. The first lemma is just a reformulation of Voiculescu's lemma [22, Lemma 4.3] (also [7, Lemma 4.3.4]).

Lemma 1.1. *Let $1 \leq p < \infty$, $R > 0$ and $\varepsilon > 0$ be all arbitrary. Then there exist an $m \in \mathbb{N}$ and a $\delta > 0$ so that for every $N \in \mathbb{N}$ and every pair of diagonal matrices $D_1, D_2 \in \mathbb{R}^N_{\geq}$ with $\|D_2\|_\infty \leq R$, the condition*

$$\left|\mathrm{tr}_N(D_1^k) - \mathrm{tr}_N(D_2^k)\right| < \delta \quad \text{for all } 1 \leq k \leq m$$

ensures that $\|D_1 - D_2\|_{p,\mathrm{tr}_N} < \varepsilon$.

Among some generalizations of Voiculescu's lemma above, an ultimate result due to Jung [15] is the following:

Lemma 1.2. (Jung [15]) *Let M be a von Neumann algebra with a faithful normal tracial state τ, and assume that M is embeddable into the ultraproduct R^ω of the hyperfinite II_1 factor and has a finite number of self-adjoint generators, say $X_1, \dots, X_n \in M$. Let $1 \leq p < \infty$ be given. The following properties are equivalent:*

1. *M is hyperfinite.*
2. *Any two embeddings of M into R^ω are unitarily equivalent in R^ω.*
3. *For each $\varepsilon > 0$ there exist an $m \in \mathbb{N}$ and a $\delta > 0$ so that for every $N \in \mathbb{N}$, if two n-tuples $(A_1, \dots, A_n)$, $(B_1, \dots, B_n) \in (M_N^{sa})^n$ satisfy*

$$|\mathrm{tr}_N(A_{i_1} \cdots A_{i_k}) - \tau(X_{i_1} \cdots X_{i_k})| < \delta, \quad |\mathrm{tr}_N(B_{i_1} \cdots B_{i_k}) - \tau(X_{i_1} \cdots X_{i_k})| < \delta$$

for all $1 \leq i_1, \dots, i_k \leq n$ and $1 \leq k \leq m$, then there is a single unitary $U \in \mathrm{U}(N)$ such that $\|UA_iU^* - B_i\|_{p,\mathrm{tr}_N} < \varepsilon$ for $1 \leq i \leq n$.

Remark that Jung dealt with only the 2-norm $\|\cdot\|_{2,\mathrm{tr}_N}$ but his argument clearly works for any p-norm.

Let (A, ϕ) be a non-commutative probability space, and $(\Omega_i)_{i\in I}$ be a family subsets of A. Let $(A^{\star I}, \phi^{\star I})$ be the free product of copies of (A, ϕ) indexed by I, and denote by ι_i the canonical embedding of A onto the ith copy of A in $A^{\star I}$. For each $\varepsilon > 0$ and $m \in \mathbb{N}$ we will say that $(\Omega_i)_{i\in I}$ are *(m, ε)-free* in (A, ϕ) if

$$\left|\phi(a_1 \cdots a_k) - \phi^{\star I}(\iota_{i_1}(a_1) \cdots \iota_{i_k}(a_k))\right| < \varepsilon$$

for all $a_j \in \Omega_{i_j}$, $i_j \in I$ with $1 \leq j \leq k$ and $1 \leq k \leq m$. The next result due to Voiculescu [25, Corollary 2.13] is a key ingredient when dealing with the freely independent situation.

Lemma 1.3. (Voiculescu [25]) *Let $R > 0$, $\varepsilon > 0$, $\theta > 0$ and $m \in \mathbb{N}$ be given. Then there exists an $N_0 \in \mathbb{N}$ such that*

$$\gamma_{\mathrm{U}(N)}^{\otimes n}\big(\{(U_1, \dots, U_p) \in \mathrm{U}(N)^p : \{T_1^{(0)}, \dots, T_{q_0}^{(0)}\}, \{U_1T_1^{(1)}U_1^*, \dots, U_1T_{q_1}^{(1)}U_1^*\},$$
$$\dots, \{U_pT_1^{(p)}U_p^*, \dots, U_pT_{q_p}^{(p)}U_p^*\} \text{ are } (m,\varepsilon)\text{-free}\}\big) > 1 - \theta$$

whenever $N \geq N_0$ *and* $T_j^{(i)} \in M_N(\mathbb{C})$ *with* $\|T_j^{(i)}\|_\infty \leq R$ *for* $1 \leq j \leq q_i$, $1 \leq q_i \leq m$, $0 \leq i \leq p$ *and* $1 \leq p \leq m$.

2. Orbital free entropy $\chi_{\rm orb}$ and its basic properties

Throughout this section, let (M, τ) be a tracial W^*-probability space and $(X_1, \dots, X_n)$ be an n-tuple of self-adjoint random variables in (M, τ). We will use the standard notations such as the microstate set $\Gamma_R(X_1, \dots, X_n; N, m, \delta)$ appearing in the course of defining the *microstate free entropy* $\chi(X_1, \dots, X_n)$ (see [22]). We define a free entropy-like quantity as follows.

Definition 2.1. For each $\delta > 0$, $m, N \in \mathbb{N}$, $R > 0$ and $1 \leq i \leq n$, we denote by $\Delta_R(X_i; N, m, \delta)$ the set of all $N \times N$ diagonal matrices $D \in \mathbb{R}^N_\geq$ satisfying $\|D\|_\infty \leq R$ and $\big|\mathrm{tr}_N(D^k) - \tau(X_i^k)\big| < \delta$ for all $1 \leq k \leq m$, and by $\Gamma_{\mathrm{orb},R}(X_1, \dots, X_n; N, m, \delta)$ the set of all n-tuples $(U_1, \dots, U_n)$ of $N \times N$ unitary matrices such that there exists an n-tuple $(D_1, \dots, D_n)$ in $\prod_{i=1}^n \Delta_R(X_i; N, m, \delta)$ satisfying

$$\big|\mathrm{tr}_N(U_{i_1} D_{i_1} U_{i_1}^* \cdots U_{i_k} D_{i_k} U_{i_k}^*) - \tau(X_{i_1} \cdots X_{i_k})\big| < \delta$$

for all $1 \leq i_1, \dots, i_k \leq n$ with $1 \leq k \leq m$. We define

$$\chi_{\mathrm{orb},R}(X_1, \dots, X_n) := \lim_{m\to\infty, \delta \searrow 0} \limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb},R}(X_1, \dots, X_n; N, m, \delta)),$$

and

$$\chi_{\mathrm{orb}}(X_1, \dots, X_n) := \sup_{R>0} \chi_{\mathrm{orb},R}(X_1, \dots, X_n).$$

The above definition of $\Gamma_{\mathrm{orb},R}(X_1, \dots, X_n; N, m, \delta)$ clearly contains a superfluous condition. In fact, it can be rephrased more simply as the set of all $(U_1, \dots, U_n) \in \mathrm{U}(N)^n$ such that

$$(U_1 D_1 U_1^*, \dots, U_n D_n U_n^*) \in \Gamma_R(X_1, \dots, X_n; N, m, \delta) \tag{2.1}$$

for some diagonal matrices $D_1, \dots, D_n \in \mathbb{R}^N_\geq$. The map $\Phi_N : ([U], D) \mapsto UDU^*$ in (1.1) gives rise to the continuous surjection $\Phi_N^n : (\mathrm{U}(N)/\mathrm{T}(N))^n \times (\mathbb{R}^N_\geq)^n \to (M_N^{sa})^n$, which provides a measure-space isomorphism between those measure spaces. Denote by $\mathrm{pr}_N^{\mathrm{U}}$ the projection map from $(\mathrm{U}(N)/\mathrm{T}(N))^n \times (\mathbb{R}^N_\geq)^n$ onto the first n factors $(\mathrm{U}(N)/\mathrm{T}(N))^n$. It is obvious that

$$(q_N^{\mathrm{U}})^n(\Gamma_{\mathrm{orb},R}(X_1, \dots, X_n; N, m, \delta)) = \mathrm{pr}_N^{\mathrm{U}}\big((\Phi_N^n)^{-1}(\Gamma_R(X_1, \dots, X_n; N, m, \delta))\big),$$

so that $\Gamma_{\mathrm{orb},R}(X_1, \dots, X_n; N, m, \delta)$ is essentially the projection of $\Gamma_R(X_1, \dots, X_n; N, m, \delta)$ to the unitary part via matrix diagonalization.

In the following let us introduce two more definitions of χ_{orb}. The first one is a slight modification of $\chi_{\mathrm{orb}}(X_1, \dots, X_n)$, where the (operator norm) cut-off procedure is removed.

Definition 2.2. We define $\Gamma_{\mathrm{orb}}(X_1, \dots, X_n; N, m, \delta)$ to be the set of all $(U_1, \dots, U_n) \in \mathrm{U}(N)^n$ satisfying (2.1) for some $D_1, \dots, D_n \in \mathbb{R}^N_\geq$ with the microstate set $\Gamma(X_1, \dots, X_n; N, m, \delta)$ without cut-off by parameter R. Define

$$\chi_{\mathrm{orb},\infty}(X_1, \dots, X_n) := \lim_{m\to\infty, \delta \searrow 0} \limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb}}(X_1, \dots, X_n; N, m, \delta)).$$

The next definition is a natural generalization of the projection free entropy χ_{proj} introduced and studied in [11] following Voiculescu's proposal in [26, 14.2].

Definition 2.3. For each $1 \leq i \leq n$ let us first choose and fix an n-tuple $(\xi_1, \dots, \xi_n)$ of sequences $\xi_i = \{\xi_i(N)\}$ of $\xi_i(N) \in M_N^{sa}$, $N \in \mathbb{N}$, such that $\xi_i(N)$ converges to X_i in moments as $N \to \infty$ for $1 \leq i \leq n$. (Of course such sequences always exist.) We define $\Gamma_{\mathrm{orb}}(X_1, \dots, X_n : \xi_1(N), \dots, \xi_n(N); N, m, \delta)$ to be the set of all $(U_1, \dots, U_n) \in \mathrm{U}(N)^n$ such that

$$\big|\mathrm{tr}_N(U_{i_1} \xi_{i_1}(N) U_{i_1}^* \cdots U_{i_k} \xi_{i_k}(N) U_{i_k}^*) - \tau(X_{i_1} \cdots X_{i_k})\big| < \delta$$

for all $1 \le i_1, \dots, i_k \le n$ with $1 \le k \le m$, that is, the set of all $(U_1, \dots, U_n) \in \mathrm{U}(N)^n$ such that $(U_i\xi_i(N)U_i^*)_{i=1}^n$ is in $\Gamma(X_1, \dots, X_n; N, m, \delta)$. Define

$$\begin{aligned}&\chi_{\mathrm{orb}}(X_1, \dots, X_n : \xi_1, \dots, \xi_n)\\ &\quad := \lim_{m\to\infty, \delta\searrow 0} \limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb}}(X_1, \dots, X_n : \xi_1(N), \dots, \xi_n(N); N, m, \delta)).\end{aligned}$$

The next lemma asserts that all the three definitions in Definitions 2.1–2.3 are equivalent. Thus, all the quantities will be denoted by the same symbol χ_{orb}, and we call $\chi_{\mathrm{orb}}(X_1, \dots, X_n)$ the *orbital free entropy* of $(X_1, \dots, X_n)$ since the definition is based on "unitary-orbital microstates."

Lemma 2.4. *For any choice of $R \ge \max_{1\le i\le n} \|X_i\|_\infty$ and for any choice of an approximating n-tuple $(\xi_1, \dots, \xi_n)$ one has*

$$\begin{aligned}\chi_{\mathrm{orb},\infty}(X_1, \dots, X_n) &= \chi_{\mathrm{orb}}(X_1, \dots, X_n)\\ &= \chi_{\mathrm{orb},R}(X_1, \dots, X_n)\\ &= \chi_{\mathrm{orb}}(X_1, \dots, X_n : \xi_1, \dots, \xi_n).\end{aligned}$$

Proof. First, due to the invariance of $\gamma_{\mathrm{U}(N)}$ under unitary conjugation, we may and do assume that $(\xi_1, \dots, \xi_n)$ is an n-tuple of sequences $\xi_i = \{D_i(N)\}$ of diagonal matrices in $\mathbb{R}^N_\ge$. Then it is obvious that $\Gamma_{\mathrm{orb}}(X_1, \dots, X_n : D_1(N), \dots, D_n(N); N, m, \delta) \subset \Gamma_{\mathrm{orb}}(X_1, \dots, X_n; N, m, \delta)$, which implies $\chi_{\mathrm{orb}}(X_1, \dots, X_n : \xi_1, \dots, \xi_n) \le \chi_{\mathrm{orb},\infty}(X_1, \dots, X_n)$. Moreover, one can choose $D_i(N)$ so that $\|D_i(N)\|_\infty \le \|X_i\|_\infty$ for all $N \in \mathbb{N}$. In this case, whenever $R \ge R_0 := \max_{1\le i\le n} \|X_i\|_\infty$, one has

$$\Gamma_{\mathrm{orb}}(X_1, \dots, X_n : D_1(N), \dots, D_n(N); N, m, \delta) \subset \Gamma_{\mathrm{orb},R}(X_1, \dots, X_n; N, m, \delta)$$

and hence $\chi_{\mathrm{orb}}(X_1, \dots, X_n : \xi_1, \dots, \xi_n) \le \chi_{\mathrm{orb},R}(X_1, \dots, X_n)$ $(\le \chi_{\mathrm{orb},\infty}(X_1, \dots, X_n))$. Thus, it suffices to prove that for any approximating sequences $\xi_i = \{D_i(N)\}$ and for every $m \in \mathbb{N}$ and $\delta > 0$, there are an $m' \in \mathbb{N}$, a $\delta' > 0$ and an $N_0 \in \mathbb{N}$ so that

$$\Gamma_{\mathrm{orb}}(X_1, \dots, X_n; N, m', \delta') \subset \Gamma_{\mathrm{orb}}(X_1, \dots, X_n : D_1(N), \dots, D_n(N); N, m, \delta) \tag{2.2}$$

for all $N \ge N_0$. Choose a $\rho \in (0, 1)$ with $m(R_0 + 1)^{m-1}\rho < \delta/2$. By Lemma 1.1 one can find an $m' \in \mathbb{N}$ with $m' \ge 2m$, a $\delta' > 0$ with $\delta' \le \min\{1, \delta/2\}$ and an $N_0 \in \mathbb{N}$ such that for every $1 \le i \le n$ and every $D_i \in \mathbb{R}^N_\ge$ with $N \ge N_0$, if $|\mathrm{tr}_N(D_i^k) - \tau(X_i^k)| < \delta'$ for all $1 \le k \le m'$, then $\|D_i - D_i(N)\|_{m,\mathrm{tr}_N} < \rho$. Suppose $N \ge N_0$ and $(U_1, \dots, U_n)$ is in the left-hand side of (2.2) so that $(U_iD_iU_i^*)_{i=1}^n \in \Gamma(X_1, \dots, X_n; N, m', \delta')$ for some $D_1, \dots, D_n \in \mathbb{R}^N_\ge$. Since $\|D_i - D_i(N)\|_{m,\mathrm{tr}_N} < \rho$ and

$$\begin{aligned}\|D_i\|_{m,\mathrm{tr}_N} &\le \mathrm{tr}_N(D_i^{2m})^{1/2m} < (\tau(X_i^{2m}) + \delta')^{1/2m}\\ &\le (R_0^{2m} + 1)^{1/2m} \le R_0 + 1,\end{aligned}$$

we get

$$\begin{aligned}&\left|\mathrm{tr}_N(U_{i_1}D_{i_1}(N)U_{i_1}^* \cdots U_{i_k}D_{i_k}(N)U_{i_k}^*) - \tau(X_{i_1} \cdots X_{i_k})\right|\\ &\quad\le \left|\mathrm{tr}_N(U_{i_1}D_{i_1}(N)U_{i_1}^* \cdots U_{i_k}D_{i_k}(N)U_{i_k}^*) - \mathrm{tr}_N(U_{i_1}D_{i_1}U_{i_1}^* \cdots U_{i_k}D_{i_k}U_{i_k}^*)\right|\\ &\qquad + \left|\mathrm{tr}_N(U_{i_1}D_{i_1}U_{i_1}^* \cdots U_{i_k}D_{i_k}U_{i_k}^*) - \tau(X_{i_1} \cdots X_{i_k})\right|\\ &\quad\le m(R_0 + 1)^{m-1}\rho + \delta' < \delta\end{aligned}$$

for all $1 \le i_1, \dots, i_k \le n$ with $1 \le k \le m$. The above latter inequality is seen by the Hölder inequality. This implies that $(U_1, \dots, U_n)$ is in the right-hand side of (2.2). □

Some basic properties of $\chi_{\rm orb}$ are summarized in the next proposition. The properties (1)–(3) are obvious, and (4) is seen by using Definition 2.2 due to Lemma 2.4.

Proposition 2.5. $\chi_{\rm orb}$ *enjoys the following properties:*

(1) $\chi_{\rm orb}(X) = 0$ *for any single random variable.*
(2) $\chi_{\rm orb}(X_1, \dots, X_n) \le 0$.
(3) $\chi_{\rm orb}(X_1, \dots, X_n) \le \chi_{\rm orb}(X_1, \dots, X_k) + \chi_{\rm orb}(X_{k+1}, \dots, X_n)$ *for every* $1 \le k < n$.
(4) *If* $(X_1^{(k)}, \dots, X_n^{(k)})$, $k \in \mathbb{N}$, *are* n*-tuples of self-adjoint random variables converging to* $(X_1, \dots, X_n)$ *in the distribution sense as* $k \to \infty$, *then*

$$\chi_{\rm orb}(X_1, \dots, X_n) \ge \limsup_{k\to\infty} \chi_{\rm orb}(X_1^{(k)}, \dots, X_n^{(k)}).$$

The following exact relation between $\chi_{\rm orb}$ and the usual χ is the main result of this section.

Theorem 2.6.

$$\chi(X_1, \dots, X_n) = \chi_{\rm orb}(X_1, \dots, X_n) + \sum_{i=1}^n \chi(X_i).$$

Proof. Let $R \ge \max_{1\le i\le n} \|X_i\|_\infty$. Since

$$\begin{aligned}&(\Phi_N^n)^{-1}(\Gamma_R(X_1, \dots, X_n; N, m, \delta)) \\ &\qquad \subset (q_N^{\rm U})^n(\Gamma_{{\rm orb},R}(X_1, \dots, X_n; N, m, \delta)) \times \prod_{i=1}^n \Delta_R(X_i; N, m, \delta)\end{aligned}$$

and $\Gamma_{{\rm orb},R}(X_1, \dots, X_n; N, m, \delta)$ is invariant under the right multiplication by elements of $\mathrm{T}(N)^n$, we get by (1.1)

$$\begin{aligned}&\log \Lambda_N^{\otimes n}(\Gamma_R(X_1, \dots, X_n; N, m, \delta)) \\ &\quad \le \log \gamma_{{\rm U}(N)}^{\otimes n}(\Gamma_{{\rm orb},R}(X_1, \dots, X_n; N, m, \delta)) + \sum_{i=1}^n \log \mu_N(\Delta_R(X_i; N, m, \delta)) \\ &\quad = \log \gamma_{{\rm U}(N)}^{\otimes n}(\Gamma_{{\rm orb},R}(X_1, \dots, X_n; N, m, \delta)) + \sum_{i=1}^n \log \Lambda_N(\Gamma_R(X_i; N, m, \delta)).\end{aligned}$$

By Lemma 2.4 this immediately implies the inequality $\le$ for the required equality.

For the reverse inequality we show that for each $m \in \mathbb{N}$ and $\delta > 0$ there are an $m' \in \mathbb{N}$, a $\delta' > 0$ and an $N_0 \in \mathbb{N}$ so that

$$\begin{aligned}&(q_N^{\rm U})^n(\Gamma_{{\rm orb},R}(X_1, \dots, X_n; N, m', \delta')) \times \prod_{i=1}^n \Delta_R(X_i; N, m', \delta') \\ &\qquad \subset (\Phi_N^n)^{-1}(\Gamma_R(X_1, \dots, X_n; N, m, \delta))\end{aligned} \tag{2.3}$$

for all $N \ge N_0$. By Lemma 1.1 one can find an $m' \in \mathbb{N}$, a $\delta' \in (0, \delta/2)$ and an $N_0 \in \mathbb{N}$ such that for every $1 \le i \le n$ and every $N \ge N_0$, if $D_i, D_i' \in \Delta_R(X_i; N, m', \delta')$ then

$$\|D_i - D_i'\|_{1,{\rm tr}_N} < \frac{\delta}{2m(R+1)^{m-1}}.$$

Now suppose $N \ge N_0$ and $([U_1], \dots, [U_n], D_1, \dots, D_n)$ is in the left-hand side of (2.3). Then we have $(U_i D_i' U_i^*)_{i=1}^n \in \Gamma_R(X_1, \dots, X_n; N, m', \delta')$ for some $(D_i')_{i=1}^n \in \prod_{i=1}^n \Delta_R(X_i; N, m', \delta')$. Since

$$\left|{\rm tr}_N(U_{i_1} D_{i_1} U_{i_1}^* \cdots U_{i_k} D_{i_k} U_{i_k}^*) - \tau(X_{i_1} \cdots X_{i_k})\right|$$

$$
\begin{aligned}
&\le \left|\mathrm{tr}_N(U_{i_1}D_{i_1}U_{i_1}^*\cdots U_{i_k}D_{i_k}U_{i_k}^*) - \mathrm{tr}_N(U_{i_1}D'_{i_1}U_{i_1}^*\cdots U_{i_k}D'_{i_k}U_{i_k}^*)\right| \\
&\quad + \left|\mathrm{tr}_N(U_{i_1}D'_{i_1}U_{i_1}^*\cdots U_{i_k}D'_{i_k}U_{i_k}^*) - \tau(X_{i_1}\cdots X_{i_k})\right| \\
&\le m(R+1)^{m-1}\max_{1\le i\le n}\|D_i - D'_i\|_{1,\mathrm{tr}_N} + \delta' < \delta
\end{aligned}
$$

for all $1 \le i_1,\dots,i_k \le n$ and $1 \le k \le m$, it follows that $(U_iD_iU_i^*)_{i=1}^n \in \Gamma_R(X_1,\dots,X_n; N,m,\delta)$, proving (2.3). By Lemma 2.4 we thus obtain

$$
\begin{aligned}
&\chi_{\mathrm{orb}}(X_1,\dots,X_n) + \sum_{i=1}^n \chi(X_i) \\
&\quad \le \limsup_{N\to\infty}\frac{1}{N^2}\log\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb},R}(X_1,\dots,X_n;N,m',\delta')) \\
&\qquad + \sum_{i=1}^n\left(\lim_{N\to\infty}\frac{1}{N^2}\log\Lambda_N\left(\Gamma_R(X_i;N,m',\delta')\right) + \frac{1}{2}\log N\right) \\
&\quad \le \limsup_{N\to\infty}\left(\frac{1}{N^2}\log\Lambda_N^{\otimes n}(\Gamma_R(X_1,\dots,X_n;N,m,\delta) + \frac{n}{2}\log N\right)
\end{aligned}
$$

for every $m \in \mathbb{N}$ and $\delta > 0$. This implies inequality $\ge$ for the desired equality. (A point in the above proof is that lim sup can be replaced by lim in the definition of $\chi(X)$ in the single variable case, see [7, 5.6.2].) □

Theorem 2.6 in particular gives

$$
-\chi_{\mathrm{orb}}(X,Y) = -\chi(X,Y) + \chi(X) + \chi(Y)
$$

for two (non-commutative) self-adjoint random variables X, Y in (M,τ) with $\chi(X), \chi(Y) > -\infty$. The above expression suggests that $-\chi_{\mathrm{orb}}$ is a kind of free probability counterpart of the so-called mutual information $I(X;Y)$ for two real random variables X, Y. In fact, recall the expression

$$
I(X;Y) = -H(X,Y) + H(X) + H(Y) \tag{2.4}
$$

in terms of the Boltzmann-Gibbs entropy $H(\cdot)$, which holds as long as $H(X)$ and $H(Y)$ are finite. The following remark is another justification for the analogy between $-\chi_{\mathrm{orb}}$ and the classical mutual information.

Remark 2.7. Let $\gamma_{\mathfrak{S}_N}$ denote the uniform probability measure on the symmetric group $\mathfrak{S}_N$. Let $X_1,\dots,X_n$ be bounded real random variables on a classical probability space. For $N, m \in \mathbb{N}$ and $\delta > 0$ define $\Delta(X_1,\dots,X_n;N,m,\delta)$ to be the set of all n-tuples $(x_1,\dots,x_n)$ of vectors $x_i = (x_{i1},\dots,x_{iN})$ in $\mathbb{R}^N$ such that

$$
\left|\frac{1}{N}\sum_{j=1}^N x_{i_1j}\cdots x_{i_kj} - \mathbb{E}(X_{i_1}\cdots X_{i_k})\right| < \delta
$$

for all $1 \le i_1,\dots,i_k \le n$ and $1 \le k \le m$, where $\mathbb{E}(\cdot)$ is the expectation. Moreover, define $\Delta_{\mathrm{sym}}(X_1,\dots,X_n;N,m,\delta)$ to be the set of all n-tuples $(\sigma_1,\dots,\sigma_n)$ of $\sigma_i \in \mathfrak{S}_N$ such that $(\sigma_1(x_1),\dots,\sigma_n(x_n)) \in \Delta(X_1,\dots,X_n;N,m,\delta)$ for some $x_1,\dots,x_n \in \mathbb{R}^N_{\ge}$, where $\sigma_i(x_i) := (x_{i\sigma_i(1)},\dots,x_{i\sigma(N)})$. We then define

$$
H_{\mathrm{sym}}(X_1,\dots,X_n) := \lim_{m\to\infty,\delta\searrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\gamma_{S_N}^{\otimes n}\left(\Delta_{\mathrm{sym}}(X_1,\dots,X_n;N,m,\delta)\right).
$$

One can show that

$$H(X_1,\dots,X_n)=H_{\rm sym}(X_1,\dots,X_n)+\sum_{i=1}^n H(X_i).$$

In particular, when X and Y are real bounded random variables with $H(X),H(Y)>-\infty$, we have $I(X;Y)=-H_{\rm sym}(X,Y)$. In this way, the "classical analog" of $-\chi_{\rm orb}(X,Y)$ provides a new definition (a kind of "discretization") of the classical mutual information $I(X;Y)$. More on this idea are examined in [8].

It seems that the expression (2.4) was one of the motivations of Voiculescu to introduce the mutual free information $i^*(A_1;\dots;A_n)$ for subalgebras $A_1,\dots,A_n$ in [26] (in particular, see Introduction there). For any n-tuple of projections $(p_1,\dots,p_n)$ in a W^*-probability space, from the definition in [11] and Lemma 2.4 we notice that

$$\chi_{\rm proj}(p_1,\dots,p_n)=\chi_{\rm orb}(p_1,\dots,p_n).$$

In [10] we conjectured that $-\chi_{\rm proj}(p,q)$ coincides with the mutual free information $i^*(\mathbb{C}p+\mathbb{C}(1-p);\mathbb{C}q+\mathbb{C}(1-q))$ for two projections p,q, and gave a heuristic computation supporting it. It would be further conjectured that $-\chi_{\rm orb}(X,Y)=i^*(W^*(X);W^*(Y))$ holds for any X,Y; however this is out of scope of this paper. Here note that this is true when X,Y are freely independent (see Proposition 2.9 below). From the above point of view we are tempted to write $i(W^*(X_1);\dots;W^*(X_n)):=-\chi_{\rm orb}(X_1,\dots,X_n)$ and use the term "microstate mutual free information." However we leave the symbol i to further progress on the subject.

In view of the analogy between $-\chi_{\rm orb}$ and $I(X;Y)$ the following proposition is strongly expected.

Proposition 2.8. *$\chi_{\rm orb}(X_1,\dots,X_n)$ depends only upon $W^*(X_1),\dots,W^*(X_n)$, where $W^*(X_i)$ means the von Neumann subalgebra of M generated by X_i (and the unit $\mathbf{1}\in M$).*

Proof. Let $(X_1',\dots,X_n')$ be another n-tuple of self-adjoints in M with $W^*(X_i')=W^*(X_i)$ for $1\le i\le n$. By symmetry and Lemma 2.4 it suffices to prove that for each $m\in\mathbb{N}$ and $\delta>0$ there are an $m'\in\mathbb{N}$ and a $\delta'>0$ such that

$$\begin{aligned}\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\rm orb}(X_1,\dots,X_n:\xi_1(N),\dots,\xi_n(N);N,m',\delta'))\\ \le\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\rm orb}(X_1',\dots,X_n';N,m,\delta))\end{aligned}\tag{2.5}$$

for all $N\in\mathbb{N}$, where $\{\xi_i(N)\}$ is an approximating sequence for X_i as in Definition 2.3 for $1\le i\le n$. Let $R:=\max_{1\le i\le n}\|X_i'\|_\infty$. For each $1\le i\le n$, since $X_i'\in W^*(X_i)$, one can choose, by the Kaplansky theorem, a real polynomial $P_i(t)$ such that $\|X_i'-P_i(X_i)\|_{1,\tau}<\delta/2m(R+1)^{m-1}$ and $\|P_i(X_i)\|_\infty\le\|X_i'\|_\infty$. For each $m\in\mathbb{N}$ and $\delta>0$ one can choose an $m'\in\mathbb{N}$ and a $\delta'>0$ (depending on $P_1,\dots,P_n$ as well) such that $(A_i)_{i=1}^n\in\Gamma(X_1,\dots,X_n;N,m',\delta')$ implies $(P_i(A_i))_{i=1}^n\in\Gamma(P_1(X_1),\dots,P_n(X_n);N,m,\delta/2)$ for every $N\in\mathbb{N}$. If $(U_1,\dots,U_n)\in\Gamma_{\rm orb}(X_1,\dots,X_n:\xi_1(N),\dots,\xi_n(N);N,m',\delta')$, then we get

$$\begin{aligned}&|\mathrm{tr}_N(U_{i_1}P_{i_1}(\xi_{i_1}(N))U_{i_1}^*\cdots U_{i_k}P_{i_k}(\xi_{i_k}(N))U_{i_k}^*)-\tau(X_{i_1}'\cdots X_{i_k}')|\\ &\quad\le|\mathrm{tr}_N(P_{i_1}(U_{i_1}\xi_{i_1}(N)U_{i_1}^*)\cdots P_{i_k}(U_{i_k}\xi_{i_k}(N)U_{i_k}^*))-\tau(P_{i_1}(X_{i_1})\cdots P_{i_k}(X_{i_k}))|\\ &\qquad+|\tau(P_{i_1}(X_{i_1})\cdots P_{i_k}(X_{i_k}))-\tau(X_{i_1}'\cdots X_{i_k}')|\\ &\quad\le\frac{\delta}{2}+m(R+1)^{m-1}\max_{1\le i\le n}\|X_i'-P_i(X_i)\|_{1,\tau}<\delta\end{aligned}$$

for all $1 \le i_1, \dots, i_k \le n$ and $1 \le k \le m$. Now, for each $N \in \mathbb{N}$ and $1 \le i \le n$ write $P_i(\xi_i(N)) = V_i(N) D_i(N) V_i(N)^*$ with $D_i(N) \in \mathbb{R}^N_{\ge}$ and $V_i(N) \in \mathrm{U}(N)$. Then we have

$$\Gamma_{\mathrm{orb}}(X_1, \dots, X_n : \xi_1(N), \dots, \xi_n(N); N, m', \delta') \cdot (V_1(N), \dots, V_n(N)) \\ \subset \Gamma_{\mathrm{orb}}(X'_1, \dots, X'_n; N, m, \delta),$$

and (2.5) follows from the right invariance of the Haar measure $\gamma_{\mathrm{U}(N)}$. □

If $\chi(X_i) > -\infty$ for all $1 \le i \le n$ and $X_1, \dots, X_n$ are freely independent, then the additivity theorem [22, Proposition 5.4] and Theorem 2.6 show that $\chi_{\mathrm{orb}}(X_1, \dots, X_n) = 0$ (or the additivity of χ_{orb} in view of Proposition 2.5 (1)). The next proposition shows that this is still true even when the finiteness assumption of the $\chi(X_i)$'s is dropped.

Proposition 2.9. *If X_1 is freely independent of $X_2, \dots, X_n$, then*

$$\chi_{\mathrm{orb}}(X_1, X_2, \dots, X_n) = \chi_{\mathrm{orb}}(X_2, \dots, X_n).$$

Consequently, $\chi_{\mathrm{orb}}(X_1, \dots, X_n) = 0$ if $X_1, \dots, X_n$ are freely independent.

Proof. The proof is based on the method due to Voiculescu [25] (or Lemma 1.3) while it is easier than that for the additivity of χ. By (1) and (3) of Proposition 2.5 we may prove that

$$\chi_{\mathrm{orb}}(X_1, X_2, \dots, X_n) \ge \chi_{\mathrm{orb}}(X_2, \dots, X_n). \tag{2.6}$$

under the assumption $\chi_{\mathrm{orb}}(X_2, \dots, X_n) > -\infty$. Choose an approximating sequence $\{\xi_i(N)\}$ for X_i with $\|\xi_i(N)\|_\infty \le \|X_i\|_\infty$ for $1 \le i \le n$. For $N, m \in \mathbb{N}$ and $\delta, \delta' > 0$ set

$$\Psi(N, m, \delta) := \Gamma_{\mathrm{orb}}(X_1, \dots, X_n : \xi_1(N), \dots, \xi_n(N); N, m, \delta),$$
$$\Phi(N, m, \delta') := \Gamma_{\mathrm{orb}}(X_2, \dots, X_n : \xi_2(N), \dots, \xi_n(N); N, m, \delta'),$$

and moreover

$$\Omega(N, m, \delta') := \big\{(U_1, U_2, \dots, U_n) \in \mathrm{U}(N)^n : \{U_1 \xi_1(N) U_1^*\} \text{ and} \\ \{U_2 \xi_2(N) U_2^*, \dots, U_n \xi_n(N) U_n^*\} \text{ are } (m, \delta')\text{-free}\big\}.$$

For every $m \in \mathbb{N}$ and $\delta > 0$ one can find a $\delta' > 0$ such that if $U_1 \in \Gamma_{\mathrm{orb}}(X_1 : \xi_1(N); N, m, \delta')$ and $(U_2, \dots, U_n) \in \Phi(N, m, \delta')$ and if $(U_1, U_2, \dots, U_n)$ is in $\Omega(N, m, \delta')$, then $(U_1, \dots, U_n)$ is in $\Psi(N, m, \delta)$. Note that $\Gamma_{\mathrm{orb}}(X_1 : \xi_1(N); N, m, \delta')$ is the whole $\mathrm{U}(N)$ for sufficiently large N. Hence, by Lemma 1.3 there is an $N_0 \in \mathbb{N}$ such that $\Gamma_{\mathrm{orb}}(X_1 : \xi_1(N); N, m, \delta') = \mathrm{U}(N)$ and

$$\gamma_{\mathrm{U}(N)}\big(\{U_1 \in \mathrm{U}(N) : (U_1, U_2, \dots, U_n) \in \Omega(N, m, \delta')\}\big) \ge \frac{1}{2}$$

for all $N \ge N_0$ and every $(U_2, \dots, U_n) \in \mathrm{U}(N)^{n-1}$. From the assumption $\chi_{\mathrm{orb}}(X_2, \dots, X_n) > -\infty$, we may assume that $\gamma_{\mathrm{U}(N)}^{\otimes n-1}(\Phi(N, m, \delta')) > 0$ for all $N \ge N_0$. Hence, with the measure

$$\mu_N := \frac{1}{\gamma_{\mathrm{U}(N)}^{\otimes n-1}(\Phi(N, m, \delta'))} \gamma_{\mathrm{U}(N)}^{\otimes n-1}\big|_{\Phi(N, m, \delta')},$$

we get for every $N \ge N_0$

$$\frac{\gamma_{\mathrm{U}(N)}^{\otimes n}(\Psi(N, m, \delta))}{\gamma_{\mathrm{U}(N)}^{\otimes n-1}(\Phi(N, m, \delta'))} \\ \ge \int_{\Phi(N, m, \delta')} \left(\int_{\mathrm{U}(N)} \mathbf{1}_{\Omega(N, m, \delta')}(U_1, U_2, \dots, U_n) \, d\gamma_{\mathrm{U}(N)}(U_1) \right) d\mu_N(U_2, \dots, U_n) \ge \frac{1}{2}.$$

Therefore,

$$\limsup_{N\to\infty}\frac{1}{N^2}\log\gamma_{\mathrm{U}(N)}^{\otimes n}(\Psi(N,m,\delta))\ge\limsup_{N\to\infty}\frac{1}{N^2}\log\gamma_{\mathrm{U}(N)}^{\otimes n-1}(\Phi(N,m,\delta')),$$

which implies (2.6) thanks to Lemma 2.4. □

3. Characterization of freeness by $\chi_{\mathrm{orb}}=0$

Let $(X_1,\dots,X_n)$ be an n-tuple of self-adjoint random variables as in the preceding section. This section is devoted to proving the converse implication of the second assertion of Proposition 2.9; consequently we have the following:

Theorem 3.1. *$\chi_{\mathrm{orb}}(X_1,\dots,X_n)=0$ if and only if $X_1,\dots,X_n$ are freely independent.*

To prove the theorem, we will provide a certain transportation cost inequality similarly to the projection case in [11, §5]. In what follows we adopt the description of χ_{orb} as $\chi_{\mathrm{orb}}(X_1,\dots,X_n:\xi_1,\dots,\xi_n)$ due to Lemma 2.4. For $1\le i\le n$ let us choose and fix a sequence $\{\xi_i(N)\}$ of $\xi_i(N)\in M_N^{sa}$ such that $\|\xi_i(N)\|_\infty\le\|X_i\|_\infty$ and $\xi_i(N)\to X_i$ in moments as $N\to\infty$.

With $R:=\max_{1\le i\le n}\|X_i\|_\infty$, $C[-R,R]$ is the C^*-algebra of continuous functions on $[-R,R]$. Let $\mathcal{A}_R:=C[-R,R]^{\star n}$ be the universal free product C^*-algebra of n-copies of $C[-R,R]$ with canonical self-adjoint free generators $Z_1,\dots,Z_n$, i.e., $Z_i(t)=t$ in the ith $C[-R,R]$. We denote by $TS(\mathcal{A}_R)$ the set of all tracial states on $\mathcal{A}_R$ and by $\mathcal{P}(\mathrm{SU}(N)^n)$ the set of all probability measures on the n-fold product $\mathrm{SU}(N)^n$. For each $\lambda\in\mathcal{P}(\mathrm{SU}(N)^n)$ we associate a unique $\widehat{\lambda}\in TS(\mathcal{A}_R)$ as follows:

$$\widehat{\lambda}(h):=\int_{\mathrm{SU}(N)^n}\mathrm{tr}_N(h(U_1\xi_1(N)U_1^*,\dots,U_n\xi_n(N)U_n^*))\,d\lambda\quad\text{for } h\in\mathcal{A}_R,\tag{3.1}$$

where $h(U_1\xi_1(N)U_1^*,\dots,U_n\xi_n(N)U_n^*)$ is the image of $h\in\mathcal{A}_R$ by the $*$-homomorphism from $\mathcal{A}_R$ to $M_N(\mathbb{C})$ sending each Z_i to $U_i\xi_i(N)U_i^*$. Similarly, $\tau_{(X_1,\dots,X_n)}\in TS(\mathcal{A}_R)$ is defined by

$$\tau_{(X_1,\dots,X_n)}(h):=\tau(h(X_1,\dots,X_n))\quad\text{for } h\in\mathcal{A}_R.$$

From the trivial fact that the image of $\mathrm{SU}(N)$ by the quotient map q_N^{U} is exactly $\mathrm{U}(N)/\mathrm{T}(N)$, it is clear that no difference occurs when $\mathrm{SU}(N)$ is used in place of $\mathrm{U}(N)$ in the definition of χ_{orb} (Definitions 2.1–2.3). Letting

$$\Gamma(N,m,\delta):=\Gamma_{\mathrm{orb}}(X_1,\dots,X_n:\xi_1(N),\dots,\xi_n(N);N,m,\delta)\cap\mathrm{SU}(N)^n,$$

we thus have

$$\chi_{\mathrm{orb}}(X_1,\dots,X_n)=\lim_{m\to\infty,\delta\searrow0}\limsup_{N\to\infty}\frac{1}{N^2}\log\gamma_{\mathrm{SU}(N)}^{\otimes n}(\Gamma(N,m,\delta)).$$

Now, we can choose a subsequence $N_1<N_2<\cdots$ in such a way that

$$\chi_{\mathrm{orb}}(X_1,\dots,X_n)=\lim_{m\to\infty}\frac{1}{N_m^2}\log\gamma_{\mathrm{SU}(N_m)}^{\otimes n}(\Gamma(N_m,m,1/m)),\tag{3.2}$$

and define

$$\lambda_m:=\frac{1}{\gamma_{\mathrm{SU}(N_m)}^{\otimes n}(\Gamma_m)}\,\gamma_{\mathrm{SU}(N_m)}^{\otimes n}\big|_{\Gamma_m}\in\mathcal{P}(\mathrm{SU}(N)^n)$$

with $\Gamma_m:=\Gamma(N_m,m,1/m)$. Then the next lemma can be proven in the same way as in the proof of [9, (2.5) in p. 401].

Lemma 3.2. *$\lim_{m\to\infty}\widehat{\lambda}_m=\tau_{(X_1,\dots,X_n)}$ in the weak* topology.*

The following is essentially a kind of reformulation of Voiculescu's asymptotic freeness result [21, 25] (also [7, §4.3]) for unitary random matrices (related to Lemma 1.3). A simple proof based on Lemma 1.3 is given for completeness.

Lemma 3.3. $\lim_{N\to\infty}\widehat{\gamma_{\mathrm{SU}(N)}^{\otimes n}} = \tau_{(X_1,\dots,X_n)}^{\mathrm{free}}$ *in the weak* topology, where* $\tau_{(X_1,\dots,X_n)}^{\mathrm{free}} := \star_{i=1}^n \tau_{X_i}$ *is the free product of the states* τ_{X_i} *on* $C[-R,R]$ *induced from the distribution measure of* X_i.

Proof. For each $m\in\mathbb{N}$ and $\delta,\theta>0$, Lemma 1.3 implies that $\gamma_{\mathrm{U}(N)}^{\otimes n}(\Omega(N,m,\delta))>1-\theta$ for all sufficiently large N, where

$$\Omega(N,m,\delta) := \{(U_1,\dots,U_n)\in\mathrm{U}(N)^n : \{U_1\xi_1(N)U_1^*\},\dots,\{U_n\xi_n(N)U_n^*\}\text{ are }(m,\delta)\text{-free}\}.$$

For any $1\le i_1,\dots,i_k\le n$ with $1\le k\le m$, notice that

$$\begin{aligned}\widehat{\gamma_{\mathrm{SU}(N)}^{\otimes n}}(Z_{i_1}\cdots Z_{i_k}) &= \int_{\mathrm{SU}(N)^n}\mathrm{tr}_N(U_{i_1}\xi_{i_1}(N)U_{i_1}^*\cdots U_{i_k}\xi_{i_k}(N)U_{i_k}^*)\,d\gamma_{\mathrm{SU}(N)}^{\otimes n}\\ &= \int_{\mathrm{U}(N)^n}\mathrm{tr}_N(U_{i_1}\xi_{i_1}(N)U_{i_1}^*\cdots U_{i_k}\xi_{i_k}(N)U_{i_k}^*)\,d\gamma_{\mathrm{U}(N)}^{\otimes n}\end{aligned}$$

and

$$\mathrm{tr}_N^{\star n}(\iota_{i_1}(\xi_{i_1}(N))\cdots\iota_{i_k}(\xi_{i_k}(N)) = \mathrm{tr}_N^{\star n}(\iota_{i_1}(U_{i_1}\xi_{i_1}(N)U_{i_1}^*)\cdots\iota_{i_k}(U_{i_k}\xi_{i_k}(N)U_{i_k}^*))$$

for all $(U_1,\dots,U_n)\in\mathrm{U}(N)^n$ (with the notations given before Lemma 1.3). Hence one can immediately estimate

$$\begin{aligned}&\Big|\widehat{\gamma_{\mathrm{SU}(N)}^{\otimes n}}(Z_{i_1}\cdots Z_{i_k}) - \mathrm{tr}_N^{\star n}(\iota_{i_1}(\xi_{i_1}(N))\cdots\iota_{i_k}(\xi_{i_k}(N)))\Big|\\ &\le\left(\int_{\Omega(N,m,\delta)}+\int_{\mathrm{U}(N)^n\setminus\Omega(N,m,\delta)}\right)\big|\mathrm{tr}_N(U_{i_1}\xi_{i_1}(N)U_{i_1}^*\cdots U_{i_k}\xi_{i_k}(N)U_{i_k}^*)\\ &\qquad\qquad - \mathrm{tr}_N^{\star n}(\iota_{i_1}(U_{i_1}\xi_{i_1}(N)U_{i_1}^*)\cdots\iota_{i_k}(U_{i_k}\xi_{i_k}(N)U_{i_k}^*))\big|\,d\gamma_{\mathrm{U}(N)}^{\otimes n}\\ &<\delta+2(R+1)^m\theta\end{aligned}$$

for every $N\ge N_0$. Since

$$\lim_{N\to\infty}\mathrm{tr}_N^{\star n}(\iota_{i_1}(\xi_{i_1}(N))\cdots\iota_{i_k}\xi_{i_k}(N))) = \tau_{(X_1,\dots,X_n)}^{\mathrm{free}}(Z_{i_1}\cdots Z_{i_k}),$$

the desired assertion follows. □

Let $W_{2,\mathrm{free}}(\tau_1,\tau_2)$ denote the free probabilistic 2-Wasserstein distance between $\tau_1,\tau_2\in TS(\mathcal{A}_R)$ introduced by Biane and Voiculescu [4] (see [9, §1.3] for a brief summary fit to our arguments). We need the next lemma comparing the free 2-Wasserstein distance with the original one (for measures) under the transformation $\lambda\in\mathcal{P}(\mathrm{SU}(N)^n)\mapsto\widehat{\lambda}\in TS(\mathcal{A}_R)$ defined in (3.1).

Lemma 3.4. *For any* $\lambda_1,\lambda_2\in\mathcal{P}(\mathrm{SU}(N)^n)$ *one has*

$$W_{2,\mathrm{free}}(\widehat{\lambda}_1,\widehat{\lambda}_2)\le\frac{2R}{\sqrt{N}}W_{2,\|\cdot\|_{\mathrm{HS}}}(\lambda_1,\lambda_2)\le\frac{2R}{\sqrt{N}}W_{2,\mathrm{geod}}(\lambda_1,\lambda_2),$$

where $W_{2,\|\cdot\|_{\mathrm{HS}}}$ *and* $W_{2,\mathrm{geod}}$ *are the* 2*-Wasserstein distances for measures with respect to the Hilbert-Schmidt norm* $\|\cdot\|_{\mathrm{HS}}$ *and the geodesic distance, respectively.*

Proof. The proof goes along the same line as that of [9, Lemma 1.3] with slight modifications in the following two points. First, let $\Pi(\lambda_1,\lambda_2)$ denote the set of all probability measures on $\mathrm{SU}(N)^n\times\mathrm{SU}(N)^n$ whose left and right marginal measures are λ_1 and λ_2, respectively. For each $\pi\in\Pi(\lambda_1,\lambda_2)$ we associate the state $\widehat{\pi}\in TS(\mathcal{A}_R\star\mathcal{A}_R)$ via (the free product of two copies of) the $*$-homomorphism sending each Z_i to $U_i\xi_i(N)U_i^*$ as above. Then one can easily observe that

$$W_{2,\mathrm{free}}(\widehat{\lambda}_1,\widehat{\lambda}_2) \le \sqrt{\int_{\mathrm{SU}(N)^n}\int_{\mathrm{SU}(N)^n}\sum_{i=1}^n \|U_i\xi_i(N)U_i^* - V_i^*\xi_i(N)V_i^*\|_{\mathrm{HS}}^2\,d\pi}$$

for any $\pi \in \Pi(\lambda_1,\lambda_2)$, where the first integration is for $(U_1,\dots,U_n)$ and the second for $(V_1,\dots,V_n)$. Secondly, we need the following elementary estimate:

$$\|U_i\xi_i(N)U_i^* - V_i\xi_i(N)V_i^*\|_{\mathrm{HS}} \le 2\|\xi_i(N)\|_\infty\|U_i - V_i\|_{\mathrm{HS}} \le 2R\|U_i - V_i\|_{\mathrm{HS}},$$

which is the reason why R appears in the desired inequality. Finally, the latter inequality is trivial because the geodesic distance majorizes the Hilbert Schmidt norm distance. □

We are now in a position to show the following transportation cost inequality. Since $W_{2,\mathrm{free}}$ is indeed a metric, this yields the implication from $\chi_{\mathrm{orb}}(X_1,\dots,X_n) = 0$ to the freeness of $X_1,\dots,X_n$, thus proving the theorem.

Proposition 3.5.

$$W_{2,\mathrm{free}}\big(\tau_{(X_1,\dots,X_n)},\tau^{\mathrm{free}}_{(X_1,\dots,X_n)}\big) \le 4\left(\max_{1\le i\le n}\|X_i\|_\infty\right)\sqrt{-\chi_{\mathrm{orb}}(X_1,\dots,X_n)}.$$

Proof. The proof is also same as that of [9, Theorem 2.2], and thus we only give an outline. Since the Ricci curvature of $\mathrm{SU}(N)^n$ (with respect to the inner product induced from $\mathrm{Re}\,\mathrm{Tr}_N$) is known to be the constant $N/2$, the transportation cost inequality

$$W_{2,\mathrm{geod}}(\lambda_m,\gamma_{\mathrm{SU}(N_m)}^{\otimes n}) \le \sqrt{\frac{4}{N_m}S(\lambda_m,\gamma_{\mathrm{SU}(N_m)}^{\otimes n})}$$

holds due to [16], where $S(\cdot\,,\cdot)$ is the relative entropy. Since

$$S(\lambda_m,\gamma_{\mathrm{SU}(N_m)}^{\otimes n}) = -\log\gamma_{\mathrm{SU}(N_m)}^{\otimes n}(\Gamma_m),$$

we have by Lemma 3.4

$$W_{2,\mathrm{free}}(\widehat{\lambda}_m,\widehat{\gamma_{\mathrm{SU}(N_m)}^{\otimes n}}) \le 4R\sqrt{-\frac{1}{N_m}\log\gamma_{\mathrm{SU}(N_m)}^{\otimes n}(\Gamma_m)}.$$

The desired inequality follows as $m\to\infty$ thanks to (3.2), Lemmas 3.2 and 3.3 together with the joint lower semicontinuity of $W_{2,\mathrm{free}}$. □

4. Generalization of χ_{orb} to hyperfinite random multi-variables

For $1\le i\le n$ let $\mathbf{X}_i = (X_{i1},\dots,X_{ir(i)})$ be a non-commutative self-adjoint random multi-variable (called a *random multi-variable* for short), which means a tuple consisting of self-adjoint random variables in (M,τ). What we want here is to generalize the orbital free entropy χ_{orb} for random variables to that for those multi-variables $\mathbf{X}_1,\dots,\mathbf{X}_n$. But there is a serious difficulty in so doing in the general setting because we have no right counterpart of the map Φ_N in (1.1) for the n-tuple space $(M_N^{sa})^n$. However, the description of χ_{orb} as $\chi_{\mathrm{orb}}(X_1,\dots,X_n:\xi_1,\dots,\xi_n)$ (see Lemma 2.4) and Jung's lemma (Lemma 1.2) allow us to define $\chi_{\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n)$ only when all $W^*(\mathbf{X}_i) := W^*(X_{i1},\dots,X_{ir(i)})$'s are hyperfinite. Throughout this section we assume that $\mathbf{X}_1,\dots,\mathbf{X}_n$ are all *hyperfinite* in this sense. Now, the definition of the orbital free entropy $\chi_{\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n)$ is similar to Definition 2.3 as follows.

Definition 4.1. For each $1 \le i \le n$ let us choose a sequence $\{\Xi_i(N)\}$ consisting of $r(i)$-tuples $\Xi_i(N) = (\xi_{i1}(N), \dots, \xi_{ir(i)}(N))$ of $\xi_{ij}(N) \in M_N^{sa}$, $N \in \mathbb{N}$, such that $\Xi_i(N)$ converges to $\mathbf{X}_i$ in the distribution sense (or in mixed moments) as $N \to \infty$. (Such a sequence always exists due to the hyperfiniteness for $\mathbf{X}_i$.) Define $\Gamma_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1(N), \dots, \Xi_n(N); N, m, \delta)$ to be the set of all n-tuples $(U_1, \dots, U_n) \in \mathrm{U}(N)^n$ such that

$$\left|\mathrm{tr}_N(U_{i_1}\xi_{i_1j_1}(N)U_{i_1}^* \cdots U_{i_k}\xi_{i_kj_k}(N)U_{i_k}^*) - \tau(X_{i_1j_1} \cdots X_{i_kj_k})\right| < \delta$$

for all $1 \le i_t \le n$, $1 \le j_t \le r(i_t)$, $1 \le t \le k$ and $1 \le k \le m$, that is, $(U_i\Xi_i(N)U_i^*)_{i=1}^n \in \Gamma(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n; N, m, \delta)$, where $U_i\Xi_i(N)U_i^*$ means $(U_i\xi_{i1}(N)U_i^*, \dots, U_i\xi_{ir(i)}U_i^*)$. Then we define

$$\begin{aligned}&\chi_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n)\\ &\quad := \lim_{m\to\infty,\delta\searrow 0} \limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1(N), \dots, \Xi_n(N); N, m, \delta)).\end{aligned}$$

If each $\mathbf{X}_i$ consists of a single random variable, then the above $\chi_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n)$ clearly coincides with the $\chi_{\rm orb}$ in §2 by definition. Moreover, the above definition is satisfactory as shown in the next lemma. The proof is similar to that of Lemma 2.4.

Lemma 4.2. *$\chi_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n)$ is independent of the choice of $(\{\Xi_1(N)\}, \dots, \{\Xi_n(N)\})$.*

Proof. Let $(\{\Xi_1'(N)\}, \dots, \{\Xi_n'(N)\})$ be another approximating n-tuple. By symmetry it suffices to show that for each $\delta > 0$ and $m \in \mathbb{N}$,

$$\begin{aligned}&\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1'(N), \dots, \Xi_n'(N); N, m, \delta/2))\\ &\qquad \le \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1(N), \dots, \Xi_n(N); N, m, \delta))\end{aligned}$$

for all sufficiently large N. Since $\Xi_i(N)$ and $\Xi_i'(N)$ converge to the same $\mathbf{X}_i$ in distribution, by Lemma 1.2 one can choose an $N_0 \in \mathbb{N}$ so that for every $N \ge N_0$ there is an n-tuple $(V_1(N), \dots, V_n(N)) \in \mathrm{U}(N)^n$ satisfying

$$\|V_i(N)\xi_{ij}(N)V_i(N)^* - \xi_{ij}'(N)\|_{m,\mathrm{tr}_N} < \frac{\delta}{2m(R+1)^{m-1}}$$

for all $1 \le j \le r(i)$ and $1 \le i \le n$, where

$$R := \sup\{\|\xi_{ij}(N)\|_{m,\mathrm{tr}_N}, \|\xi_{ij}'(N)\|_{m,\mathrm{tr}_N} : 1 \le j \le r(i),\ 1 \le i \le n,\ N \in \mathbb{N}\}\ (< +\infty).$$

If $(U_1, \dots, U_n) \in \Gamma_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1'(N), \dots, \Xi_n'(N); N, m, \delta/2)$ with $N \ge N_0$, then as in the proof of Lemma 2.4 we get

$$\begin{aligned}&\big|\mathrm{tr}_N(U_{i_1}(V_{i_1}(N)\xi_{i_1j_1}(N)V_{i_1}(N)^*)U_{i_1}^* \cdots U_{i_k}(U_{i_k}(N)\xi_{i_kj_k}(N)U_{i_k}(N)^*)U_{i_k}^*)\\ &\qquad - \tau(X_{i_1j_1} \cdots X_{i_kj_k})\big|\\ &\quad \le m(R+1)^{m-1} \max_{i,j} \|V_i(N)\xi_{ij}(N)V_i(N)^* - \xi_{ij}'(N)\|_{m,\mathrm{tr}_N} + \frac{\delta}{2}\\ &\quad < \delta\end{aligned}$$

for all $1 \le i_t \le n$, $1 \le j_t \le r(i_t)$, $1 \le t \le k$ and $1 \le k \le m$. This means that

$$\begin{aligned}&\Gamma_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1'(N), \dots, \Xi_n'(N); N, m, \delta/2)\\ &\qquad \subset \Gamma_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1(N), \dots, \Xi_n(N); N, m, \delta) \cdot (V_1(N), \dots, V_n(N))\end{aligned}$$

for all $N \ge N_0$. Hence we have the desired assertion thanks to the right invariance of $\gamma_{\mathrm{U}(N)}$. □

Remark 4.3. Jung's result [15] (or Lemma 1.2) says that the above definition of $\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)$ can work only when all the $\mathbf{X}_i$'s are hyperfinite since the microstates are not concentrated in a single (approximate) unitary orbit for a general random multi-variable. Thus, to define $\chi_{\rm orb}$ for general random multi-variables, we need an appropriate way to gather together all unitary orbits without "overlap" in each matricial level. One potential way is to use the space of unitarily equivalent classes of $*$-representations of $W^*(\mathbf{X}_i)$ in R^ω, which plays a similar role of $\mathbb{R}^N_\geq$ for random variables. Note that the restriction of diagonal matrices to $\mathbb{R}^N_\geq$ is needed in the definition of $\chi_{\rm orb}$ to avoid "overlap"; indeed, if $\mathbb{R}^N$ is used in place of $\mathbb{R}^N_\geq$, then the space of "orbital microstates" has the "overlap" coming from the symmetry of $\mathfrak{S}_N$ acting on the eigenvalue space $\mathbb{R}^N$. One more way we considered is to use a suitable fundamental domain of the diagonal action of $\mathrm{U}(N)$ on $\Gamma_R(X_{i1},\dots,X_{ir(i)};N,m,\delta)$ as a role of $\mathbb{R}^N_\geq$, but we encountered some difficulty in this approach.

Except for the relation between $\chi_{\rm orb}$ and χ (Theorem 2.6), all basic properties of $\chi_{\rm orb}$ can be extended to hyperfinite random multi-variables in the same way, which are summarized in the next proposition. Note that the assertion of Theorem 2.6 is meaningless for hyperfinite random multi-variables because both sides of the equality are $-\infty$ as long as at least one of the $\mathbf{X}_i$'s is not a single variable.

Only (4)–(8) of the proposition are somewhat non-trivial. Note that (6) is the $\chi_{\rm orb}$ counterpart of [26, Remark 9.2 (e)] while it is just a byproduct of (5). The proofs of (5), (7) and (8) are essentially same as before in the case of $\chi_{\rm orb}$ for random variables; for example, Lemma 1.2 is used in place of Lemma 1.1. We will sketch them and leave the full details to the reader.

Proposition 4.4. $\chi_{\rm orb}$ *for hyperfinite random multi-variables enjoys the following properties:*

(1) $\chi_{\rm orb}(\mathbf{X})=0$ *for any single* $\mathbf{X}$.

(2) $\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)\le 0$.

(3) $\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)\le \chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_k)+\chi_{\rm orb}(\mathbf{X}_{k+1},\dots,\mathbf{X}_n)$.

(4) *If* $\mathbf{X}_i^{(k)}=(X_{i1}^{(k)},\dots,X_{ir(i)}^{(k)})$ *are hyperfinite random multi-variables for* $1\le i\le n$ *and* $k\in\mathbb{N}$ *such that* $\mathbf{X}_1^{(k)}\sqcup\cdots\sqcup\mathbf{X}_n^{(k)}\to\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n$ *in the distribution sense as* $k\to\infty$, *then*

$$\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)\ge\limsup_{k\to\infty}\chi_{\rm orb}(\mathbf{X}_1^{(k)},\dots,\mathbf{X}_n^{(k)}).$$

(5) $\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)$ *depends only upon* $W^*(\mathbf{X}_1),\dots,W^*(\mathbf{X}_n)$*; more precisely,*

$$\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)=\chi_{\rm orb}(\mathbf{X}_1',\dots,\mathbf{X}_n')$$

for hyperfinite random multi-variables $\mathbf{X}_i$ *and* $\mathbf{X}_i'$ *with* $W^*(\mathbf{X}_i)=W^*(\mathbf{X}_i')$, $1\le i\le n$, *where the numbers of variables in* $\mathbf{X}_i$ *and in* $\mathbf{X}_i'$ *may be different.*

(6) *If* $\mathbf{Y}_1,\dots,\mathbf{Y}_n$ *are random multi-variables such that* $\mathbf{Y}_i\subset W^*(\mathbf{X}_i)$ *for* $1\le i\le n$, *then*

$$\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)\le\chi_{\rm orb}(\mathbf{Y}_1,\dots,\mathbf{Y}_n).$$

(7) *If* $\mathbf{X}_1$ *is freely independent of* $\mathbf{X}_2,\dots,\mathbf{X}_n$, *then*

$$\chi_{\rm orb}(\mathbf{X}_1,\mathbf{X}_2,\dots,\mathbf{X}_n)=\chi_{\rm orb}(\mathbf{X}_2,\dots,\mathbf{X}_n).$$

(8) $\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)=0$ *if and only if* $\mathbf{X}_1,\dots,\mathbf{X}_n$ *are freely independent.*

Proof. (4) For each $k\in\mathbb{N}$ let $(\{\Xi_1^{(k)}(N)\},\dots,\{\Xi_n^{(k)}(N)\})$ be an approximating n-tuple for $(\mathbf{X}_1^{(k)},\dots,\mathbf{X}_n^{(k)})$. For any $\alpha<\limsup_k\chi_{\rm orb}(\mathbf{X}_1^{(k)},\dots,\mathbf{X}_n^{(k)})$ one can choose a sequence $k_1<k_2<\dots$ such that $\chi_{\rm orb}(\mathbf{X}_1^{(k_m)},\dots,\mathbf{X}_n^{(k_m)})>\alpha$ and

$$|\tau(X_{i_1j_1}^{(k_m)}\cdots X_{i_\ell j_\ell}^{(k_m)})-\tau(X_{i_1j_1}\cdots X_{i_\ell j_\ell})|<\frac{1}{m}\tag{4.1}$$

for all $1 \le i_t \le n$, $1 \le j_t \le n(i_t)$, $1 \le t \le \ell$ and $1 \le \ell \le m$. Furthermore, one can find a sequence $N_1 < N_2 < \dots$ such that for every $m \in \mathbb{N}$ and $1 \le i \le n$,

$$\Xi_i^{(k_m)}(N) \in \Gamma(\mathbf{X}_i^{(k_m)}; N, m, 1/m) \quad \text{if } N \ge N_m, \tag{4.2}$$

and also

$$\frac{1}{N_m^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb}}(\mathbf{X}_1^{(k_m)}, \dots, \mathbf{X}_n^{(k_m)} : \Xi_1^{(k_m)}(N_m), \dots, \Xi_n^{(k_m)}(N_m); N_m, m, 1/m)) > \alpha. \tag{4.3}$$

For $1 \le i \le n$ define

$$\Xi_i(N) := \Xi_i^{(k_m)}(N) \quad \text{if } N_m \le N < N_{m+1},\ m \in \mathbb{N}.$$

By (4.1) and (4.2), for $1 \le i \le n$ we get $\Xi_i(N) \in \Gamma(\mathbf{X}_i; N, m, 2/m)$ if $N_m \le N < N_{m+1}$, $m \in \mathbb{N}$; hence $(\{\Xi_1(N)\}, \dots, \{\Xi_n(N)\})$ is an approximating n-tuple for $(\mathbf{X}_1, \dots, \mathbf{X}_n)$. For each $m \in \mathbb{N}$, if $(U_1, \dots, U_n)$ is in

$$\Gamma_{\mathrm{orb}}(\mathbf{X}_1^{(k_m)}, \dots, \mathbf{X}_n^{(k_m)} : \Xi_1^{(k_m)}(N_m), \dots, \Xi_n^{(k_m)}(N_m); N_m, m, 1/m),$$

then $(U_i \Xi_i(N_m) U_i^*)_{i=1}^n = (U_i \Xi_i^{(k_m)}(N_m) U_i^*)_{i=1}^n$ is in $\Gamma(\mathbf{X}_1^{(k_m)} \sqcup \dots \sqcup \mathbf{X}_n^{(k_m)}; N_m, m, 1/m)$. Since this set of microstates is included in $\Gamma(\mathbf{X}_1 \sqcup \dots \sqcup \mathbf{X}_n; N_m, m, 2/m)$ thanks to (4.1), it follows that

$$\begin{aligned} &\Gamma_{\mathrm{orb}}(\mathbf{X}_1^{(k_m)}, \dots, \mathbf{X}_n^{(k_m)} : \Xi_1^{(k_m)}(N_m), \dots, \Xi_n^{(k_m)}(N_m); N_m, m, 1/m) \\ &\qquad \subset \Gamma_{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1(N_m), \dots, \Xi_n(N_m); N_m, m, 2/m). \end{aligned}$$

Hence, by (4.3) we have

$$\frac{1}{N_m^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1(N_m), \dots, \Xi_n(N_m); N_m, m, 2/m)) > \alpha$$

for all $m \in \mathbb{N}$. This immediately implies that $\chi_{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n) \ge \alpha$, and the result follows.

(5) Let $\mathbf{X}_i = (X_{i1}, \dots, X_{ir(i)})$ and $\mathbf{X}'_i = (X'_{i1}, \dots, X'_{ir'(i)})$ be as stated in the proposition, and choose their approximating n-tuples $(\{\Xi_1(N)\}, \dots, \{\Xi_n(N)\})$ and $(\{\Xi'_1(N)\}, \dots, \{\Xi'_n(N)\})$, respectively, with $\Xi_i(N) = (\xi_{ij}(N))_{j=1}^{r(i)}$ and $\Xi'_i(N) = (\xi'_{ij}(N))_{j=1}^{r'(i)}$. We may assume that $\|\xi_{ij}(N)\|_\infty, \|\xi'_{ij'}(N)\|_\infty \le \max_{i,j,j'}\{\|X_{ij}\|_\infty, \|X'_{ij'}\|_\infty\}$ for all i, j, j' and N. Now, it suffices to prove that for each $m \in \mathbb{N}$ and $\delta > 0$ there are an $m' \in \mathbb{N}$, a $\delta' > 0$ and an $N_0 \in \mathbb{N}$ such that

$$\begin{aligned} &\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m', \delta')) \\ &\qquad \le \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb}}((\mathbf{X}'_i)_{i=1}^n : (\Xi'_i(N))_{i=1}^n; N, m, \delta)) \end{aligned} \tag{4.4}$$

for all $N \ge N_0$. The proof is essentially same as that of (2.5) but more complicated since the right-hand side of (4.4) contains $\Xi'_i(N)$ differently from (2.5). The Kaplansky density theorem enables us to choose non-commutative self-adjoint polynomials P_{ij} of $r(i)$ indeterminates for $1 \le j \le r'(i)$, $1 \le i \le n$ such that $\|P_{ij}(\mathbf{X}_i)\|_\infty \le R$ and $\|X'_{ij} - P_{ij}(\mathbf{X}_i)\|_{1,\tau}$ is arbitrarily small; hence $\mathbf{X}'_1 \sqcup \dots \sqcup \mathbf{X}'_n$ is arbitrarily approximated by $(P_{1j}(\mathbf{X}_1))_{j=1}^{r'(1)} \sqcup \dots \sqcup (P_{nj}(\mathbf{X}_n))_{j=1}^{r'(n)}$ in distribution. Since $\Xi'_i(N) \to \mathbf{X}'_i$ and $(P_{ij}(\Xi_i(N)))_{j=1}^{r'(i)} \to (P_{ij}(\mathbf{X}_i))_{j=1}^{r'(i)}$ in distribution as $N \to \infty$, by Lemma 1.2 one can find an $N_0 \in \mathbb{N}$ such that for every $N \ge N_0$ and $1 \le i \le n$ there exists a $V_i(N) \in \mathrm{U}(N)$ for which $\|P_{ij}(\Xi_i(N)) - V_i(N)\xi'_{ij}(N)V_i(N)^*\|_{2,\mathrm{tr}_N}$ is arbitrarily small for $1 \le j \le r'(i)$. Then one can choose an $m' \in \mathbb{N}$ and a $\delta' > 0$ such that

$$\begin{aligned} &\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m', \delta') \cdot (V_i(N))_{i=1}^n \\ &\qquad \subset \Gamma_{\mathrm{orb}}((\mathbf{X}'_i)_{i=1}^n : (\Xi'_i(N))_{i=1}^n; N, m, \delta) \end{aligned}$$

for all $N \ge N_0$, implying (4.4).

(6) Letting $\mathbf{X}_i' := \mathbf{X}_i \sqcup \mathbf{Y}_i$ for $1 \le i \le n$ we have by (5)

$$\chi_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n) = \chi_{\rm orb}(\mathbf{X}_1', \dots, \mathbf{X}_n') \le \chi_{\rm orb}(\mathbf{Y}_1, \dots, \mathbf{Y}_n),$$

since the latter inequality is obvious by definition.

(7) The proof is completely same as that of Proposition 2.9; just replace X_i, $\xi_i(N)$ by $\mathbf{X}_i$, $\Xi_i(N)$. See also Proposition 4.7 for its generalization.

(8) The assertions (1) and (7) show that the freeness implies $\chi_{\rm orb} = 0$. The converse is proven by extending the transportation cost inequality in Proposition 3.5 to hyperfinite random multi-variables. The proof is same as before, so only a few remarks are mentioned here. Set $R := \max_{i,j} \|X_{ij}\|_\infty$ and let $\mathcal{A}_R$ be the universal free product of $r(1) + \cdots + r(n)$ copies of $C[-R, R]$ with canonical generators Z_{ij} for $1 \le j \le r(i)$, $1 \le i \le n$. By the $*$-homomorphism sending each Z_{ij} to X_{ij} we obtain $\tau_{(\mathbf{X}_1,\dots,\mathbf{X}_n)} \in TS(\mathcal{A}_R)$ as in §3. Also, for every $\lambda \in \mathcal{P}(\mathrm{SU}(N)^n)$ we associate $\widehat{\lambda} \in TS(\mathcal{A}_R)$ in the same manner as in §3 by the integral over the unitary orbit $\{(U_1\Xi_1(N)U_1^*, \dots, U_n\Xi_n(N)U_n^*) : (U_1, \dots, U_n) \in \mathrm{SU}(N)^n\}$ with respect to λ. Then, the counterparts of Lemmas 3.2 and 3.3 are proven exactly in the same way. Indeed, applying Lemma 1.3 to

$$\{(U_1, \dots, U_n) \in \mathrm{U}(N)^n : U_1\Xi_1(N)U_1^*, \dots, U_n\Xi_n(N)U_n^* \text{ are } (m, \delta)\text{-free}\}$$

one can show that $\lim_{N\to\infty} \widehat{\gamma_{\mathrm{SU}(N)}^{\otimes n}} = \tau^{\rm free}_{(\mathbf{X}_1,\dots,\mathbf{X}_n)}$ weakly*, where $\tau^{\rm free}_{(\mathbf{X}_1,\dots,\mathbf{X}_n)} \in TS(\mathcal{A}_R)$ is the free product of the states $\tau_{\mathbf{X}_i}$ on $C^*(Z_{ij}, \dots, Z_{ir(i)})$ induced from the original τ on M via the $*$-homomorphism sending Z_{ij} to X_{ij} for $1 \le j \le r(i)$. With these the same argument as before proves

$$W_{2,\rm free}(\tau_{(\mathbf{X}_1,\dots,\mathbf{X}_n)}, \tau^{\rm free}_{(\mathbf{X}_1,\dots,\mathbf{X}_n)}) \le 4R\sqrt{-\chi_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n)},$$

from which we get the conclusion. □

Next, we introduce the $\chi_{\rm orb}(\cdots : \mathbf{v})$ in the presence of unitary random variables, which will be necessary in the next section.

First, let us recall the Γ-set of microstates approximating $\mathbf{X} = (X_1, \dots, X_n)$ in the presence of unitary random variables. In addition to $\mathbf{X}$ let $\mathbf{v} = (v_1, \dots, v_\ell)$ be an ℓ-tuple of unitary random variables in (M, τ). For $N, m \in \mathbb{N}$ and $\delta > 0$ we denote by $\Gamma(\mathbf{X}, \mathbf{v}; N, m, \delta)$ the set of all $(A_1, \dots, A_n, V_1, \dots, V_\ell)$ in $(M_N^{sa})^n \times \mathrm{U}(N)^\ell$ such that

$$|\mathrm{tr}_N(h(A_1, \dots, A_n, V_1, \dots, V_\ell)) - \tau(h(\mathbf{X}, \mathbf{v}))| < \delta$$

for all $*$-monomials h of $n + \ell$ indeterminates of degree not greater than m, and by $\Gamma(\mathbf{X} : \mathbf{v}; N, m, \delta)$ the set of all $(A_1, \dots, A_n) \in (M_N^{sa})^n$ such that $(A_1, \dots, A_n, V_1, \dots, V_\ell) \in \Gamma(\mathbf{X}, \mathbf{v}; N, m, \delta)$ for some $(V_1, \dots, V_\ell) \in \mathrm{U}(N)^\ell$.

Definition 4.5. For $1 \le i \le n$ choose a microstate sequence $\Xi_i(N) = (\xi_{i1}(N), \dots, \xi_{ir(i)}(N))$ in $(M_N^{sa})^{r(i)}$, $N \in \mathbb{N}$, such that $\Xi_i(N)$ converges to $\mathbf{X}_i$ in the distribution sense as $N \to \infty$. Moreover, let $\mathbf{v} = (v_1, \dots, v_\ell)$ be unitary random variables in (M, τ). For $N, m \in \mathbb{N}$ and $\delta > 0$ define $\Gamma_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1(N), \dots, \Xi_n(N) : \mathbf{v}; N, m, \delta)$ to be the set of all $(U_1, \dots, U_n) \in \mathrm{U}(N)^n$ such that $(U_i\Xi_i(N)U_i^*)_{i=1}^n$ is in $\Gamma(\mathbf{X}_1, \dots, \mathbf{X}_n : \mathbf{v}; N, m, \delta)$. Then we define the *orbital free entropy* of $(\mathbf{X}_1, \dots, \mathbf{X}_n)$ in the presence of $\mathbf{v}$ by

$$\begin{aligned}
&\chi_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \mathbf{v}) \\
&:= \lim_{m\to\infty,\,\delta\searrow 0} \limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}\big(\Gamma_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \Xi_1(N), \dots, \Xi_n(N) : \mathbf{v}; N, m, \delta)\big).
\end{aligned}$$

Similarly to Lemma 4.2 the above definition of $\chi_{\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n : \mathbf{v})$ is independent of the choice of an approximating n-tuple $(\{\Xi_1(N)\}, \dots, \{\Xi_n(N)\})$.

The next proposition can be regarded as the $\chi_{\rm orb}$-counterpart of [26, Proposition 10.4]. In what follows, $\chi_u(\cdots)$ means the free entropy of unitary random variables (see [7, §6.5]).

Proposition 4.6. *Let $\mathbf{v} = (v_1,\dots,v_n)$ be a freely independent n-tuple of unitary random variables with $\chi_u(v_i) > -\infty$ for all $1 \le i \le n$. If $\mathbf{X}_1,\dots,\mathbf{X}_n$ are freely independent of $\mathbf{v}$, then*

$$\begin{aligned}\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n) &\le \chi_{\rm orb}(v_1\mathbf{X}_1v_1^*,\dots,v_n\mathbf{X}_nv_n^* : \mathbf{v})\\ &\le \chi_{\rm orb}(v_1\mathbf{X}_1v_1^*,\dots,v_n\mathbf{X}_nv_n^*).\end{aligned}$$

In particular, when the above $\mathbf{X}_1,\dots,\mathbf{X}_n$ are single self-adjoint random variables $X_1,\dots,X_n$, one has

$$\chi(X_1,\dots,X_n) \le \chi(v_1X_1v_1^*,\dots,v_nX_nv_n^*).$$

Proof. The latter assertion follows immediately from the first thanks to Theorem 2.6. For the first assertion it is enough to prove only the first inequality. Choose $\Xi_i(N)$ as in Definition 4.5 with $\|\Xi_i(N)\|_\infty \le \|\mathbf{X}_i\|_\infty$ for $1 \le i \le n$ and $N \in \mathbb{N}$, where $\|\Xi_i(N)\|_\infty := \max_{1\le j\le r(i)} \|\xi_{ij}\|_\infty$ and $\|\mathbf{X}_i\|_\infty := \max_{1\le j\le r(i)} \|X_{ij}\|_\infty$. For $N, m \in \mathbb{N}$ and $\delta, \rho > 0$ we write for short

$$\begin{aligned}\Phi(N,m,\delta) &:= \Gamma_{\rm orb}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N,m,\delta),\\ \widehat{\Psi}(N,m,\rho) &:= \Gamma(v_1\mathbf{X}_1v_1^* \sqcup \cdots \sqcup v_n\mathbf{X}_nv_n^*, \mathbf{v}; N,m,\rho),\\ \widetilde{\Psi}(N,m,\rho) &:= \Gamma(v_1\mathbf{X}_1v_1^* \sqcup \cdots \sqcup v_n\mathbf{X}_nv_n^* : \mathbf{v}; N,m,\rho),\\ \Psi(N,m,\rho) &:= \Gamma_{\rm orb}((v_i\mathbf{X}_iv_i^*)_{i=1}^n : (\Xi_i(N))_{i=1}^n : \mathbf{v}; N,m,\rho).\end{aligned}$$

We define two probability measures μ_N and ν_N on $\mathrm{U}(N)^n$ by

$$\begin{aligned}\mu_N &:= \frac{1}{\gamma_{\mathrm{U}(N)}^{\otimes n}(\Phi(N,m,\delta))}\,\gamma_{\mathrm{U}(N)}^{\otimes n}\big|_{\Phi(N,m,\delta)},\\ \nu_N &:= \frac{1}{\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma(\mathbf{v};N,2m,\delta))}\,\gamma_{\mathrm{U}(N)}^{\otimes n}\big|_{\Gamma(\mathbf{v};N,2m,\delta)},\end{aligned}$$

where $\Gamma(\mathbf{v};N,2m,\delta)$ is the Γ-set of unitary microstates in $\mathrm{U}(N)^n$ approximating $\mathbf{v}$ (see [7, §6.5]). Here we may and do assume that $\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n) > -\infty$ so that μ_N is well-defined for all sufficiently large $N \in \mathbb{N}$. Also, note that ν_N is well-defined for all sufficiently large $N \in \mathbb{N}$ thanks to the assumption of free independence for $\mathbf{v}$. Furthermore, define

$$\begin{aligned}\Omega(N,3m,\delta) := \big\{&(U_1,\dots,U_n,V_1,\dots,V_n) \in \mathrm{U}(N)^n \times \mathrm{U}(N)^n :\\ &(U_i\Xi_i(N)U_i^*)_{i=1}^n \text{ and } (V_i)_{i=1}^n \text{ are } (3m,\delta)\text{-free}\big\}.\end{aligned}$$

For every $m \in \mathbb{N}$ and $\rho > 0$ one can choose a $\delta > 0$ such that if $\mathbf{U} = (U_1,\dots,U_n) \in \Phi(N,m,\delta)$, $\mathbf{V} = (V_1,\dots,V_n) \in \Gamma(\mathbf{v};N,2m,\delta)$ and $(\mathbf{U},\mathbf{V}) \in \Omega(N,3m,\delta)$, then $((V_iU_i\Xi_i(N)U_i^*V_i^*)_{i=1}^n, \mathbf{V}) \in \widehat{\Psi}(N,m,\rho)$. Since $\mu_N\otimes\nu_N$ is invariant under the $\mathrm{U}(N)$-action given by $(\mathbf{U},\mathbf{V}) \mapsto (\mathbf{U}, W\mathbf{V}W^*)$, $W\mathbf{V}W^* := (WV_iW^*)_{i=1}^n$, for $W \in \mathrm{U}(N)$, it follows from Lemma 1.3 (see the proof of [25, Corollary 2.14]) that $(\mu_N \otimes \nu_N)(\Omega(N,3m,\delta)) \ge 1/2$ whenever N is large enough (depending only on m,δ). For each such N one can choose a $\mathbf{V} \in \Gamma(\mathbf{v};N,2m,\delta)$ such that

$$\frac{1}{2} \le \mu_N(\Omega(N,3m,\delta : \mathbf{V})) = \frac{\gamma_{\mathrm{U}(N)}^{\otimes n}(\Phi(N,m,\delta) \cap \Omega(N,3m,\delta : \mathbf{V}))}{\gamma_{\mathrm{U}(N)}^{\otimes n}(\Phi(N,m,\delta))}, \tag{4.5}$$

where $\Omega(N,3m,\delta : \mathbf{V}) := \{\mathbf{U} \in \mathrm{U}(N)^n : (\mathbf{U},\mathbf{V}) \in \Omega(N,3m,\delta)\}$. From the above choice of δ we have

$$\big\{(V_iU_i\Xi_i(N)U_i^*V_i^*)_{i=1}^n : \mathbf{U} \in \Phi(N,m,\delta) \cap \Omega(N,3m,\delta : \mathbf{V})\big\} \subset \widetilde{\Psi}(N,m,\rho),$$

that is,

$$\big\{\mathbf{VU} = (V_iU_i)_{i=1}^n : \mathbf{U} \in \Phi(N,m,\delta) \cap \Omega(N,3m,\delta : \mathbf{V})\big\} \subset \Psi(N,m,\rho).$$

Thanks to the left invariance of $\gamma_{\mathrm{U}(N)}$, this and (4.5) imply that

$$\frac{1}{2}\gamma_{\mathrm{U}(N)}^{\otimes n}(\Phi(N,m,\delta)) \le \gamma_{\mathrm{U}(N)}^{\otimes n}(\Psi(N,m,\rho)).$$

Therefore,

$$\chi_{\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n) \le \limsup_{N\to\infty} \frac{1}{N^2}\log\gamma_{\mathrm{U}(N)}^{\otimes n}(\Psi(N,m,\rho)),$$

implying the required inequality. □

The next proposition is exactly the χ_{orb}-counterpart of [25, Theorem 3.8].

Proposition 4.7. *Let $\mathbf{v} = (v_1,\dots,v_n)$ be unitary random variables. If $(\mathbf{X}_1, v_1)$ is freely independent of $\mathbf{X}_2,\dots,\mathbf{X}_n$ and $v_2,\dots,v_n$, then*

$$\chi_{\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n : \mathbf{v}) = \chi_{\mathrm{orb}}(\mathbf{X}_1 : v_1) + \chi_{\mathrm{orb}}(\mathbf{X}_2,\dots,\mathbf{X}_n : v_2,\dots,v_n)$$

whenever $\mathbf{X}_1$ is regular in the presence of v_1, that is, replacing the lim sup *as $N\to\infty$ by* lim inf *gives the same value in the definition $\chi_{\mathrm{orb}}(\mathbf{X}_1 : v_1)$.*

Proof. Since the subadditivity

$$\begin{aligned}&\chi_{\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n : v_1,\dots,v_n)\\ &\quad\le \chi_{\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_k : v_1,\dots,v_k) + \chi_{\mathrm{orb}}(\mathbf{X}_{k+1},\dots,\mathbf{X}_n : v_{k+1},\dots,v_n)\end{aligned} \tag{4.6}$$

is obvious by definition, it suffices to show inequality $\ge$ for the required equality. We can assume that $\chi_{\mathrm{orb}}(\mathbf{X}_1 : v_1) > -\infty$ and $\chi_{\mathrm{orb}}(\mathbf{X}_2,\dots,\mathbf{X}_n : v_2,\dots,v_n) > -\infty$. We choose $\Xi_i(N)$ as in the previous proof and for each $N, m \in \mathbb{N}$ and $\delta,\rho > 0$ write

$$\begin{aligned}\Phi(N,m,\delta) &:= \Gamma_{\mathrm{orb}}(\mathbf{X}_1 : \Xi_1(N) : v_1; N,m,\delta)\\ &\qquad\times \Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=2}^n : (\Xi_i(N))_{i=2}^n : (v_i)_{i=2}^n; N,m,\delta),\\ \Psi(N,m,\rho) &:= \Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n : \mathbf{v}; N,m,\rho).\end{aligned}$$

The assumption guarantees that $\Phi(N,m,\delta)$ is not of $\gamma_{\mathrm{U}(N)}^{\otimes n}$-measure 0 for all N large enough. We will prove that for each $m \in \mathbb{N}$ and $\rho > 0$ there is a $\delta > 0$ such that

$$\frac{\gamma_{\mathrm{U}(N)}^{\otimes n}(\Psi(N,m,\rho)\cap\Phi(N,m,\delta))}{\gamma_{\mathrm{U}(N)}^{\otimes n}(\Phi(N,m,\delta))} \ge \frac{1}{2} \tag{4.7}$$

for all sufficiently large N. The proof is similar to that of [25, Lemma 3.5]. First, note that $\Gamma_{\mathrm{orb}}(\mathbf{X}_1 : \Xi_1(N) : v_1; N,m,\delta)$ is invariant under the left action $U_1 \mapsto UU_1$ for $U \in \mathrm{U}(N)$. Hence the probability measure

$$\mu_N := \frac{1}{\gamma_{\mathrm{U}(N)}^{\otimes n}(\Phi(N,m,\delta))}\,\gamma_{\mathrm{U}(N)}^{\otimes n}\big|_{\Phi(N,m,\delta)}$$

is invariant under the same action of $\mathrm{U}(N)$ to the only first component. Next, for any $m \in \mathbb{N}$ and $\rho > 0$, one can choose a $\delta > 0$ so that if $(\mathbf{A}_1, V_1) \in \Gamma(\mathbf{X}_1, v_1; N,m,\delta)$ with $\|\mathbf{A}_1\|_\infty \le \|\mathbf{X}_1\|_\infty$ and $((\mathbf{A}_i)_{i=2}^n, (V_i)_{i=2}^n) \in \Gamma((\mathbf{X}_i)_{i=2}^n, (v_i)_{i=2}^n; N,m,\delta)$ with $\|\mathbf{A}_i\|_\infty \le \|\mathbf{X}_i\|_\infty$ and if $(\mathbf{A}_1, V_1)$ and $((\mathbf{A}_i)_{i=2}^n, (V_i)_{i=2}^n)$ are (m,δ)-free, then $((\mathbf{A}_i)_{i=1}^n, (V_i)_{i=1}^n) \in \Gamma((\mathbf{X}_i)_{i=1}^n, (v_i)_{i=1}^n; N,m,\rho))$. Lemma 1.3 implies that

$$\gamma_{\mathrm{U}(N)}(\{U \in \mathrm{U}(N) : (U\mathbf{A}_1U^*, UV_1U^*) \text{ and } ((\mathbf{A}_i)_{i=2}^n, (V_i)_{i=2}^n) \text{ are } (m,\delta)\text{-free}\}) \ge \frac{1}{2}$$

for every $\mathbf{A}_i$ and V_i as above whenever N is sufficiently large (depending only on m, δ). Then it follows that

$$\gamma_{\mathrm{U}(N)}\big(\{U \in \mathrm{U}(N) : (UU_1, (U_i)_{i=2}^n) \in \Psi(N, m, \rho)\}\big) \geq \frac{1}{2}$$

for all $(U_1, (U_i)_{i=2}^n) \in \Phi(N, m, \delta)$ whenever N is sufficiently large. This implies that

$$\mu_N(\Psi(N, m, \rho)) = \int_{\mathrm{U}(N)^n} \left(\int_{\mathrm{U}(N)} \mathbf{1}_{\Psi(N,m,\rho)}(UU_1, (U_i)_{i=2}^n)\, d\gamma_{\mathrm{U}(N)}(U) \right) d\mu_N \geq \frac{1}{2},$$

implying (4.7). Therefore, we obtain

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}(\Psi(N, m, \rho)) \\
&\quad \geq \limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}(\Phi(N, m, \delta)) \\
&\quad \geq \liminf_{N\to\infty} \frac{1}{N^2} \log \Gamma_{\mathrm{orb}}(\mathbf{X}_1 : \Xi_1(N) : v_1; N, m, \delta) \\
&\qquad + \limsup_{N\to\infty} \frac{1}{N^2} \log \Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=2}^n : (\Xi_i(N))_{i=2}^n : (v_i)_{i=2}^n; N, m, \delta),
\end{aligned}$$

and the desired inequality follows thanks to the regularity assumption of $\mathbf{X}_1$ (in the presence of v_1). □

5. Orbital free entropy dimension

The microstate free entropy dimension δ and its modified one δ_0 due to Voiculescu [22, 23] are defined for self-adjoint random variables based on the microstate free entropy χ and the semicircular deformation

$$(X_1 + \varepsilon S_1, \ldots, X_n + \varepsilon S_n), \qquad \varepsilon > 0, \tag{5.1}$$

where $(S_1, \ldots, S_n)$ is a free semicircular system freely independent of given self-adjoint random variables $X_1, \ldots, X_n$. In this section we will introduce the orbital version $\delta_{0,\mathrm{orb}}$ of δ_0 (and also δ_{orb} of δ), or in other words the dimension counterpart of the orbital free entropy χ_{orb} discussed in the previous sections. Our essential idea to define $\delta_{0,\mathrm{orb}}$ is to replace χ by χ_{orb} and more importantly the semicircular deformation (5.1) by the so-called liberation process

$$(v_1(t)X_1v_1(t)^*, \ldots, v_n(t)X_nv_n(t)^*), \qquad t > 0, \tag{5.2}$$

introduced by Voiculescu [26], where $(v_1(t), \ldots, v_n(t))$ is a free n-tuple of multiplicative free unitary Brownian motions (see [1]) freely independent of the X_i's. The idea to use the liberation process goes back to our attempt to define the dimension counterpart of χ_{proj}; note that the space of projections with fixed traces is not closed under the semicircular deformation (5.1) while it is under the liberation process (5.2).

Throughout the rest of this section, let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be hyperfinite random multi-variables in (M, τ) as treated in §4.

Definition 5.1. Let $\mathbf{v}(t) = (v_1(t), \ldots, v_n(t))$, $t \geq 0$, be a freely independent n-tuple of multiplicative free unitary Brownian motions (see [1]) with $v_i(0) = \mathbf{1}$ chosen to be freely independent of $\mathbf{X}_1, \ldots, \mathbf{X}_n$. (We may always assume that such extra variables exist in (M, τ).) Write $v_i(t)\mathbf{X}_iv_i(t)^* := (v_i(t)X_{i1}v_i(t)^*, \ldots, v_i(t)X_{ir(i)}v_i(t)^*)$ and define the *modified orbital free entropy dimension* of $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ by

$$\delta_{0,\mathrm{orb}}(\mathbf{X}_1, \ldots, \mathbf{X}_n) := \limsup_{\varepsilon \searrow 0} \frac{\chi_{\mathrm{orb}}(v_1(\varepsilon)\mathbf{X}_1v_1(\varepsilon)^*, \ldots, v_n(\varepsilon)\mathbf{X}_nv_n(\varepsilon)^* : \mathbf{v}(\varepsilon))}{|\log \varepsilon^{1/2}|}.$$

One may also define the orbital free entropy dimension $\delta_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)$ in the same manner by using $\chi_{\rm orb}(v_1(\varepsilon)\mathbf{X}_1v_1(\varepsilon)^*,\dots,v_n(\varepsilon)\mathbf{X}_nv_n(\varepsilon)^*)$ without the presence of $\mathbf{v}(\varepsilon)$. However, we will deal with only $\delta_{0,{\rm orb}}$ in this paper.

Remark 5.2. Let $(X_1,\dots,X_n)$ be an n-tuple of self-adjoint random variables and $\mathbf{v}$ a tuple of unitary random variables in (M,τ). The proof of Theorem 2.6 can be slightly modified to obtain

$$\chi(X_1,\dots,X_n:\mathbf{v})=\chi_{\rm orb}(X_1,\dots,X_n:\mathbf{v})+\sum_{i=1}^n\chi(X_i).$$

Applying this to $(v_1(\varepsilon)X_1v_1(\varepsilon)^*,\dots,v_n(\varepsilon)X_nv_n(\varepsilon)^*)$ and $\mathbf{v}(\varepsilon)$ yields

$$\begin{aligned}&\chi(v_1(\varepsilon)X_1v_1(\varepsilon)^*,\dots,v_n(\varepsilon)X_nv_n(\varepsilon)^*:\mathbf{v}(\varepsilon))\\&\quad=\chi_{\rm orb}(v_1(\varepsilon)X_1v_1(\varepsilon)^*,\dots,v_n(\varepsilon)X_nv_n(\varepsilon)^*:\mathbf{v}(\varepsilon))+\sum_{i=1}^n\chi(X_i).\end{aligned}$$

Consequently, if $\chi(X_i)>-\infty$ for all $1\le i\le n$, then we have

$$\delta_{0,{\rm orb}}(X_1,\dots,X_n)=\limsup_{\varepsilon\searrow0}\frac{\chi(v_1(\varepsilon)X_1v_1(\varepsilon)^*,\dots,v_n(\varepsilon)X_nv_n(\varepsilon)^*:\mathbf{v}(\varepsilon))}{|\log\varepsilon^{1/2}|}.$$

This formula might serve as the definition of $\delta_{0,{\rm orb}}$ for random variables $(X_1,\dots,X_n)$ such that $\chi(X_i)>-\infty$ for $1\le i\le n$. However, it does not make sense for hyperfinite random multi-variables $(\mathbf{X}_1,\dots,\mathbf{X}_n)$ since $\chi(v_1(\varepsilon)\mathbf{X}_1v_1(\varepsilon)^*\sqcup\cdots\sqcup v_n(\varepsilon)\mathbf{X}_nv_n(\varepsilon)^*:\mathbf{v}(\varepsilon))=-\infty$ as long as at least one of the $\mathbf{X}_i$'s is not a single variable.

The next proposition summarizes properties of $\delta_{0,{\rm orb}}$; (1)–(3) are rather obvious. The assertion (4) says that $\delta_{0,{\rm orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n)$ can be regarded as the (modified) orbital free entropy dimension of the hyperfinite subalgebras $W^*(\mathbf{X}_1),\dots,W^*(\mathbf{X}_n)$. Note that (6) is the orbital counterpart of [23, Proposition 6.10]. Also, note that (7) is the $\delta_{0,{\rm orb}}$-counterpart of Proposition 2.9, which slightly strengthens the second assertion of (6).

Proposition 5.3. *$\delta_{0,{\rm orb}}$ for hyperfinite random multi-variables enjoys the following properties*:

(1) $\delta_{0,{\rm orb}}(\mathbf{X})=0$ *for a single multi-variable* $\mathbf{X}$.
(2) $\delta_{0,{\rm orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n)\le0$.
(3) $\delta_{0,{\rm orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n)\le\delta_{0,{\rm orb}}(\mathbf{X}_1,\dots,\mathbf{X}_k)+\delta_{0,{\rm orb}}(\mathbf{X}_{k+1},\dots,\mathbf{X}_n)$ *for every* $1\le k<n$.
(4) $\delta_{0,{\rm orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n)$ *depends only upon* $W^*(\mathbf{X}_1),\dots,W^*(\mathbf{X}_n)$.
(5) *If* $\mathbf{Y}_1,\dots,\mathbf{Y}_n$ *are random multi-variables such that* $\mathbf{Y}_i\subset W^*(\mathbf{X}_i)$ *for* $1\le i\le n$, *then*
$$\delta_{0,{\rm orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n)\le\delta_{0,{\rm orb}}(\mathbf{Y}_1,\dots,\mathbf{Y}_n).$$
(6) *If* $\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)>-\infty$, *then* $\delta_{0,{\rm orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n)=0$. *In particular,* $\delta_{0,{\rm orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n)=0$ *if* $\mathbf{X}_1,\dots,\mathbf{X}_n$ *are freely independent.*
(7) *If* $\mathbf{X}_1$ *is freely independent of* $\mathbf{X}_2,\dots,\mathbf{X}_n$, *then*
$$\delta_{0,{\rm orb}}(\mathbf{X}_1,\mathbf{X}_2\dots,\mathbf{X}_n)=\delta_{0,{\rm orb}}(\mathbf{X}_2,\dots,\mathbf{X}_n).$$

Proof. Since $\chi_{\rm orb}(\mathbf{X})=0$ for a single $\mathbf{X}$, (1) is contained in (6). (2) is trivial since $\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n:\mathbf{v})\le0$ for any $\mathbf{X}_1,\dots,\mathbf{X}_n$ and $\mathbf{v}$. (3) follows from the subadditivity (4.6).

(4) For $1\le i\le n$ let $\mathbf{X}_i'=(X_{i1}',\dots,X_{ir'(i)}')$ be another random multi-variable with $W^*(\mathbf{X}_i')=W^*(\mathbf{X}_i)$. To show the assertion, it suffices to prove the equality of the modified orbital free entropies

$$\chi_{\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n:\mathbf{v})=\chi_{\rm orb}(\mathbf{X}_1',\dots,\mathbf{X}_n':\mathbf{v})$$

in the presence of unitary random variables $\mathbf{v}$. But the proofs of Propositions 2.8 and 4.4 (5) can be easily modified to prove this, so the details are omitted.

(5) follows immediately from (4) as in the proof of Proposition 4.4 (6).

(6) Since $\chi_u(v_i(\varepsilon)) > -\infty$ for every $\varepsilon > 0$ (see e.g., [26, Proposition 10.10]), Proposition 4.6 shows that

$$\chi_{\mathrm{orb}}(v_1(\varepsilon)\mathbf{X}_1 v_1(\varepsilon)^*, \dots, v_n(\varepsilon)\mathbf{X}_n v_n(\varepsilon)^* : \mathbf{v}(\varepsilon)) \ge \chi_{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n)$$

for every $\varepsilon > 0$, from which the desired assertion immediately follows. The latter assertion follows from Proposition 4.4 (8).

(7) The proof of Proposition 4.6 shows that for every $m \in \mathbb{N}$ and $\rho > 0$ there is a $\delta > 0$ such that

$$\begin{aligned}&\frac{1}{2}\gamma_{\mathrm{U}(N)}\big(\Gamma_{\mathrm{orb}}(\mathbf{X}_1 : \Xi_1(N); N, m, \delta)\big)\\ &\qquad \le \gamma_{\mathrm{U}(N)}\big(\Gamma(v_1(\varepsilon)\mathbf{X}_1 v_1(\varepsilon)^* : \Xi_1(N) : v_1(\varepsilon); N, m, \rho)\big)\end{aligned}$$

for all sufficiently large N. Since $\Gamma_{\mathrm{orb}}(\mathbf{X}_1 : \Xi_1(N); N, m, \delta)$ is the whole $\mathrm{U}(N)$ whenever N is large enough, $v_1(\varepsilon)\mathbf{X}_1 v_1(\varepsilon)^*$ is regular in the presence of $v_1(\varepsilon)$ as in Proposition 4.7 and $\chi_{\mathrm{orb}}(v_1(\varepsilon)\mathbf{X}_1 v_1(\varepsilon)^* : v_1(\varepsilon)) = 0$ for every $\varepsilon > 0$. Therefore, Proposition 4.7 shows that

$$\begin{aligned}&\chi_{\mathrm{orb}}(v_1(\varepsilon)\mathbf{X}_1 v(\varepsilon)^*, v_2(\varepsilon)\mathbf{X}_2 v_2(\varepsilon)^*, \dots, v_n(\varepsilon)\mathbf{X}_n v_n(\varepsilon)^* : \mathbf{v}(\varepsilon))\\ &\quad = \chi_{\mathrm{orb}}(v_1(\varepsilon)\mathbf{X}_1 v(\varepsilon)^*, \dots, v_n(\varepsilon)\mathbf{X}_n v_n(\varepsilon)^* : v_2(\varepsilon), \dots, v_n(\varepsilon))\end{aligned}$$

for every $\varepsilon > 0$, which immediately implies the required equality. □

Now, we examine how Jung's covering/packing approach [12, 13] to δ_0 works for $\delta_{0,\mathrm{orb}}$ introduced above. First, let us recall the notions of covering/packing numbers. Let $(\mathcal{X}, d)$ be a Polish space and $\Gamma \subset \mathcal{X}$. Consider Γ as a metric space with the restriction of d on Γ. For each $\varepsilon > 0$ we denote by $K_\varepsilon(\Gamma)$ the minimum number of open ε-balls covering Γ, and by $P_\varepsilon(\Gamma)$ the maximum number of elements in a family of mutually disjoint open ε-balls in Γ, where ε-balls in Γ are taken as subsets of Γ. Those numbers will sometimes be denoted by $K_\varepsilon(\Gamma, d)$ and $P_\varepsilon(\Gamma, d)$ to emphasize the metric d. A subset $\{x_s : s \in S\}$ of Γ is called an *ε-net* of Γ if the open ε-balls centered at x_s, $s \in S$, cover Γ, and also an *ε-separated set* of Γ if the ε-balls centered at x_s, $s \in S$, are mutually disjoint. This definition is slightly different from that in [20] but consistent with the definition of packing numbers used here. Moreover, $\mathcal{N}_\varepsilon(\Gamma)$ stands for the open ε-neighborhood of Γ. Remark that $P_\varepsilon(\Gamma) \ge K_{2\varepsilon}(\Gamma) \ge P_{4\varepsilon}(\Gamma)$ holds in general, and thus if a lower/upper estimate for either $K_\varepsilon(\Gamma)$ or $P_\varepsilon(\Gamma)$ was proven, then the essentially same estimate for the other would immediately follow.

On the space $(M_N^{sa})^n$ ($\cong \mathbb{R}^{nN^2}$) we consider the metric d_2 induced from the Hilbert-Schmidt norm with respect to tr_N.

Definition 5.4. Let $\mathbf{X}_1, \dots, \mathbf{X}_n$ and $\{\Xi_1(N)\}, \dots, \{\Xi_n(N)\}$ be as in Definition 4.5. Define the *orbital fractal free entropy dimension* of $(\mathbf{X}_1, \dots, \mathbf{X}_n)$ by

$$\delta_{1,\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n) := \limsup_{\varepsilon \searrow 0} \frac{\mathbb{K}_\varepsilon^{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n)}{|\log \varepsilon|} - n = \limsup_{\varepsilon \searrow 0} \frac{\mathbb{P}_\varepsilon^{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n)}{|\log \varepsilon|} - n,$$

where

$$\mathbb{K}_\varepsilon^{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n) := \lim_{m \to \infty, \delta \searrow 0} \limsup_{N \to \infty} \frac{1}{N^2} \log K_\varepsilon\big(\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m, \delta)\big)$$

and $\mathbb{P}_\varepsilon^{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n)$ is similar with P_ε in place of K_ε. Indeed, it is seen from the proof of Lemma 4.2 that the definitions of $\mathbb{K}_\varepsilon^{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n)$, $\mathbb{P}_\varepsilon^{\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n)$ and hence $\delta_{1,\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n)$ are independent of the choice of $(\{\Xi_1(N)\}, \dots, \{\Xi_n(N)\})$.

Let us then prove the equality $\delta_{0,\mathrm{orb}} = \delta_{1,\mathrm{orb}}$. Indeed, the subtraction by n in the above definition of $\delta_{1,\mathrm{orb}}$ is necessary to get this equality. To do so we need a lemma, which says that $\mathbf{v}(t)$ is regular; namely, we have the same value if lim sup is replaced by lim inf in the definition of $\chi_u(\mathbf{v}(t))$ (see [7, §6.5]). Its proof is essentially same as in the case of self-adjoint variables (the large deviation principle in [7, 5.4.10] might be important).

Lemma 5.5. *Let $\mathbf{v}(t)$, $t \geq 0$, be as in Definition* 5.1. *Then for every $t \geq 0$,*

$$\lim_{m\to\infty,\delta\searrow 0} \liminf_{N\to\infty} \frac{1}{N^2} \log \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma(\mathbf{v}(t); N, m, \delta)) = \chi_u(\mathbf{v}(t)) = \sum_{i=1}^{n} \chi_u(v_i(t)).$$

Proposition 5.6.

$$\delta_{0,\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n) = \delta_{1,\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n).$$

Proof. The idea of the proof is similar to that in [13]. First, by [1, Lemma 8] there is a constant $K > 0$ such that $\|v_i(t) - \mathbf{1}\|_\infty \leq K t^{1/2}$ for all $0 \leq t \leq 1$. In what follows let $C := K^2 + 2$, and let $N, m \in \mathbb{N}$ and $\varepsilon, \delta > 0$ be arbitrary with restriction $\delta < \varepsilon \leq 1$. Also let $\Xi_i(N)$ be as in Definition 4.5.

First let us prove the inequality $\geq$. One can choose a $2(Cn\varepsilon)^{1/2}$-separated subset $\{\mathbf{U}_{Ns} = (U_{Nsi})_{i=1}^n : s \in S_N\}$ of $\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m, \delta)$ with

$$|S_N| = P_{2(Cn\varepsilon)^{1/2}}\big(\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m, \delta)\big). \tag{5.3}$$

(See the remark above Definition 5.4 for the terminology of "ε-separated sets".) Define two probability measures μ_N, ν_N on $\mathrm{U}(N)^n$ by

$$\mu_N := \frac{1}{|S_N|} \sum_{s\in S_N} \delta_{\mathbf{U}_{Ns}} \quad (\delta_{\mathbf{U}_{Ns}} \text{ is the Dirac measure at } \mathbf{U}_{Ns} \in \mathrm{U}(N)^n),$$
$$\nu_N := \frac{1}{\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma(\mathbf{v}(\varepsilon); N, m, \delta))} \gamma_{\mathrm{U}(N)}^{\otimes n}\big|_{\Gamma(\mathbf{v}(\varepsilon); N, m, \delta)}.$$

Write $\mathbf{U} = (U_i)_{i=1}^n \in \mathrm{U}(N)^n$ etc., and set

$$\Omega(N, 3m, \delta) := \big\{(\mathbf{U}, \mathbf{V}) \in \mathrm{U}(N)^n \times \mathrm{U}(N)^n : (U_i \Xi_i(N) U_i^*)_{i=1}^n \text{ and } \mathbf{V} \text{ are } (3m, \delta)\text{-free}\big\},$$
$$\Phi_s(N, 3m, \delta) := \big\{\mathbf{V} \in \Gamma(\mathbf{v}(\varepsilon); N, m, \delta) : (U_{Nsi} \Xi_i(N) U_{Nsi}^*)_{i=1}^n \text{ and } \mathbf{V} \text{ are } (3m, \delta)\text{-free}\big\}$$

for $s \in S_N$. Since $\mu_N \otimes \nu_N$ is invariant under the $\mathrm{U}(N)$-action $(\mathbf{U}, \mathbf{V}) \mapsto (\mathbf{U}, W\mathbf{V}W^*)$ for $W \in \mathrm{U}(N)$, by Lemma 1.3 (as [25, Corollary 2.14]) we have

$$\frac{1}{2} \leq (\mu_N \otimes \nu_N)(\Omega(N, 3m, \delta)) = \frac{1}{|S_N|} \sum_{s\in S_N} \nu_N(\Phi_s(N, 3m, \delta))$$

so that

$$\sum_{s\in S_N} \gamma_{\mathrm{U}(N)}^{\otimes n}(\Phi_s(N, 3m, \delta)) \geq \frac{1}{2} |S_N| \gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma(\mathbf{v}(\varepsilon); N, m, \delta)) \tag{5.4}$$

whenever N is large enough. For every $\mathbf{V} = (V_1, \dots, V_n) \in \Gamma(\mathbf{v}(\varepsilon); N, m, \delta)$ we get

$$\begin{aligned} \|V_i U_{Nsi} - U_{Nsi}\|_{2,\mathrm{tr}_N}^2 &= \|V_i - I\|_{2,\mathrm{tr}_N}^2 = \mathrm{tr}_N(2I - V_i - V_i^*) \\ &\leq \tau(2\mathbf{1} - v_i(\varepsilon) - v_i(\varepsilon)^*) + 2\delta \\ &< \|v_i(\varepsilon) - \mathbf{1}\|_\infty^2 + 2\varepsilon \leq C\varepsilon \end{aligned} \tag{5.5}$$

so that $d_2(\mathbf{V}\mathbf{U}_{Ns},\mathbf{U}_{Ns}) < (Cn\varepsilon)^{1/2}$. Hence it follows that $\Phi_s(N,3m,\delta)\mathbf{U}_{Ns} := \big\{\mathbf{V}\mathbf{U}_{Ns} : \mathbf{V} \in \Phi_s(N,3m,\delta)\big\}$, $s \in S_N$, are mutually disjoint. Furthermore, it is seen that for any $\rho > 0$ we have

$$\bigsqcup_{s\in S_N} \Phi_s(N,3m,\delta)\mathbf{U}_{Ns} \subset \Gamma_{\mathrm{orb}}((v_i(\varepsilon)\mathbf{X}_i v_i(\varepsilon)^*)_{i=1}^n : (\Xi_i(N))_{i=1}^n : \mathbf{v}(\varepsilon); N, m, \rho) \tag{5.6}$$

if a sufficiently small $\delta \in (0,\varepsilon)$ was chosen. By (5.6), (5.4) and (5.3) we have

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{1}{N^2}\log\gamma_{\mathrm{U}(N)}^{\otimes n}\big(\Gamma_{\mathrm{orb}}((v_i(\varepsilon)\mathbf{X}_i v_i(\varepsilon)^*)_{i=1}^n : (\Xi_i(N))_{i=1}^n : \mathbf{v}(\varepsilon); N, m, \rho)\big) \\
&\qquad \ge \limsup_{N\to\infty} \frac{1}{N^2}\log\sum_{s\in S_N}\gamma_{\mathrm{U}(N)}^{\otimes n}(\Phi_s(N,3m,\delta)) \\
&\qquad \ge \limsup_{N\to\infty} \frac{1}{N^2}\log\bigg(\frac{1}{2}|S_N|\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma(\mathbf{v}(\varepsilon);N,m,\delta))\bigg) \\
&\qquad \ge \mathbb{P}_{2(Cn\varepsilon)^{1/2}}(\mathbf{X}_1,\dots,\mathbf{X}_n) + \sum_{i=1}^n \chi_u(v_i(\varepsilon))
\end{aligned}$$

thanks to Lemma 5.5. Note here that $\lim_{\varepsilon\searrow 0}\chi_u(v_i(\varepsilon))/|\log\varepsilon^{1/2}| = -1$ can be easily derived from Voiculescu's computation [22, Proposition 6.3] based on [1, Lemma 8] and [26, Proposition 1.6] since the spectrum of $v_i(\varepsilon)$ is concentrated in a very small arc around 1 for all sufficiently small $\varepsilon > 0$ (also see [17, Proposition 6.1]). Hence the above estimate implies the required inequality.

Next let us prove the inequality $\le$. Let $\mathbf{U} = (U_i)_{i=1}^n \in \Gamma_{\mathrm{orb}}((v_i(\varepsilon)\mathbf{X}_i v_i(\varepsilon)^*)_{i=1}^n : (\Xi_i(N))_{i=1}^n : \mathbf{v}(\varepsilon); N, 3m, \delta)$, which is accompanied by another $\mathbf{V} = (V_i)_{i=1}^n \in \Gamma(\mathbf{v}(\varepsilon); N, 3m, \delta)$ by definition. One easily observes that $\mathbf{V}^*\mathbf{U} = (V_i^*U_i)_{i=1}^n$ is in $\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m, \delta)$. As similar to (5.5) we have $d_2(\mathbf{U},\mathbf{V}^*\mathbf{U}) < (Cn\varepsilon)^{1/2}$. Hence $\Gamma_{\mathrm{orb}}((v_i(\varepsilon)\mathbf{X}_i v_i(\varepsilon)^*)_{i=1}^n : (\Xi_i(N))_{i=1}^n : \mathbf{v}(\varepsilon); N, 3m, \delta)$ is included in $\mathcal{N}_{(Cn\varepsilon)^{1/2}}(\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m, \delta))$. Now choose an $\varepsilon^{\mathbf{1/2}}$-net $\{\mathbf{U}'_{Ns} : s \in S'_N\}$ of $\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m, \delta)$ with

$$|S'_N| = K_{\varepsilon^{1/2}}(\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m, \delta)).$$

Then $\Gamma_{\mathrm{orb}}(((v_i(\varepsilon)\mathbf{X}_i(v_i(\varepsilon)^*)_{i=1}^n : (\Xi_i(N))_{i=1}^n : \mathbf{v}(\varepsilon); N, 3m, \delta)$ is clearly included in the union of the $((Cn)^{1/2}+1)\varepsilon^{1/2}$-balls $B_{((Cn)^{1/2}+1)\varepsilon^{1/2}}(\mathbf{U}'_{Ns})$ centered at $\mathbf{U}'_{Ns}$, $s \in S'_N$. By using the packing number estimate of $\mathrm{U}(N)$ due to Szarek [19] one easily sees that there is a constant $C' > 0$ independent of N so that

$$\gamma_{\mathrm{U}(N)}^{\otimes n}(B_{((Cn)^{1/2}+1)\varepsilon^{1/2}}(\mathbf{U}'_{Ns})) \le (C'((Cn)^{1/2}+1)\varepsilon^{1/2})^{nN^2}$$

as long as $\varepsilon > 0$ is small enough. Therefore we get

$$\begin{aligned}
&\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb}}((v_i(\varepsilon)\mathbf{X}_i v_i(\varepsilon)^*)_{i=1}^n : (\Xi_i(N))_{i=1}^n : \mathbf{v}(\varepsilon); N, 3m, \delta)) \\
&\qquad \le |S'_N|(C'((Cn)^{1/2}+1)\varepsilon^{1/2})^{nN^2},
\end{aligned}$$

and hence

$$\begin{aligned}
&\frac{1}{N^2}\log\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb}}((v_i(\varepsilon)\mathbf{X}_i v_i(\varepsilon)^*)_{i=1}^n : (\Xi_i(N))_{i=1}^n : \mathbf{v}(\varepsilon); N, 3m, \delta)) \\
&\qquad \le \frac{1}{N^2}\log K_{\varepsilon^{1/2}}(\Gamma_{\mathrm{orb}}((\mathbf{X}_i)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m, \delta)) \\
&\qquad\qquad + n\log\varepsilon^{1/2} + n\log(C'((Cn)^{1/2}+1)).
\end{aligned}$$

Taking $\lim_{m\to\infty,\delta\searrow 0}\limsup_{N\to\infty}$ of both sides, we have

$$\begin{aligned}&\chi_{\rm orb}(v_1(\varepsilon)\mathbf{X}_1 v_1(\varepsilon)^*,\dots,v_n(\varepsilon)\mathbf{X}_n v_n(\varepsilon)^*:\mathbf{v}(\varepsilon))\\ &\quad\le \mathbb{K}^{\rm orb}_{\varepsilon^{1/2}}(\mathbf{X}_1,\dots,\mathbf{X}_n)+n\log\varepsilon^{1/2}+n\log(C'((Cn)^{1/2}+1)),\end{aligned}$$

from which the desired inequality immediately follows. □

Remark 5.7. The role of multiplicative free unitary Brownian motions $v_i(t)$ is not so essential in the above proof of Proposition 5.6. In fact, besides the free independence assumption, we used only the facts that $\|v_i(t)-\mathbf{1}\|_\infty \le Kt^{1/2}$ for small $t\ge 0$ and that $\lim_{\varepsilon\searrow 0}\chi_u(v_i(\varepsilon))/|\log\varepsilon^{1/2}| = -1$, while Lemma 5.5 is valid for general freely independent unitary random variables. Consequently, we notice that the definition $\delta_{0,\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)$ in Definition 5.1 is equivalent when $\mathbf{v}(t)$ is replaced by, for example, $(e^{\sqrt{-1}\sqrt{t}h_1},\dots,e^{\sqrt{-1}\sqrt{t}h_n})$, where $h_1,\dots,h_n$ are freely independent self-adjoint random variables with $\chi(h_i)>-\infty$ for $1\le i\le n$ chosen to be freely independent of $\mathbf{X}_1,\dots,\mathbf{X}_n$. The situation is similar to the case δ_0 shown in [13].

The main result of this section is the following exact relation between $\delta_{0,\rm orb}$ and the usual δ_0.

Theorem 5.8.

$$\delta_0(\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n)=\delta_{0,\rm orb}(\mathbf{X}_1,\dots,\mathbf{X}_n)+\sum_{i=1}^n\delta_0(\mathbf{X}_i).$$

The rest of this section is devoted to the proof of the theorem. We will prove the part "≤" first and next "≥." The latter is more involved than the former.

Let $\Xi_i(N)$ be chosen for $\mathbf{X}_i$, $1\le i\le n$, as in Definition 4.5. For $N,m\in\mathbb{N}$ and $\delta>0$ define

$$\begin{aligned}&\Gamma(\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n:(\Xi_i(N))_{i=1}^n;N,m,\delta)\\ &\quad:=\Gamma(\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n;N,m,\delta)\cap\{(U_i\Xi_i(N)U_i^*)_{i=1}^n:(U_1,\dots,U_n)\in\mathrm{U}(N)^n\},\end{aligned}\tag{5.7}$$

where $U_i\Xi_i(N)U_i^*$ is as in Definition 4.1. We need the following simple lemma.

Lemma 5.9.

$$\begin{aligned}&\delta_0(\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n)\\ &\quad=\limsup_{\varepsilon\searrow 0}\frac{1}{|\log\varepsilon|}\left(\lim_{m\to\infty,\delta\searrow 0}\limsup_{N\to\infty}\frac{1}{N^2}\log K_\varepsilon\big(\Gamma(\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n:(\Xi_i(N))_{i=1}^n;N,m,\delta)\big)\right).\end{aligned}$$

The same formula holds also when K_ε is replaced by P_ε.

Proof. Thanks to Jung's covering/packing approach [13] to δ_0 with additional remarks [6, p. 455] and [14, Lemma 2.2], it suffices to show that for every $m\in\mathbb{N}$ and $\varepsilon,\delta>0$ there are an $m_0\in\mathbb{N}$ and a $\delta_0>0$ such that

$$\Gamma(\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n;N,m_0,\delta_0)\subset\mathcal{N}_\varepsilon\big(\Gamma(\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n:(\Xi_i(N))_{i=1}^n;N,m,\delta)\big)$$

for all sufficiently large N. But this can be easily verified by Lemma 1.2. □

For a while fix an arbitrary $1\le i\le n$. In what follows we assume that $W^*(\mathbf{X}_i)$ has both diffuse and atomic parts since this case is most involved and needs all the ingredients of the proof. Let us decompose

$$W^*(\mathbf{X}_i)=\bigoplus_{k=0}^{s(i)}M_{ik}$$

with $s(i) \in \mathbb{N} \cup \{\infty\}$ such that M_{i0} is diffuse and $M_{ik} \cong M_{m_{ik}}(\mathbb{C})$ for $k \geq 1$, and denote by p_{ik} the central support projection of M_{ik}, $k \geq 0$. By Jung's result [12] one has

$$\delta_0(\mathbf{X}_i) = 1 - \sum_{k=1}^{s(i)} \frac{\tau(p_{ik})^2}{m_{ik}^2}. \tag{5.8}$$

We choose and fix a matrix unit system $\{e_{\alpha\beta}^{(ik)} : 1 \leq \alpha, \beta \leq m_{ik}\}$ of $M_{ik} \cong M_{m_{ik}}(\mathbb{C})$ for each $k \geq 1$. Let $\ell \in \mathbb{N}$ be arbitrary, and write $q_i^{(\ell)} := \mathbf{1} - \sum_{k=0}^{s(i)\wedge\ell} p_{ik}$ with $s(i) \wedge \ell := \min\{s(i), \ell\}$. We can choose, for any sufficiently large $N \in \mathbb{N}$, positive integers $n_{ik}^{(\ell)}(N)$ and $m_{ik}^{(\ell)}(N)$ for $0 \leq k \leq s(i) \wedge \ell$ such that $n_{ik}^{(\ell)}(N) = m_{ik} m_{ik}^{(\ell)}(N)$, $k \geq 1$, and

$$\begin{aligned} &\sum_{k=0}^{s(i)\wedge\ell} n_{ik}^{(\ell)}(N) \leq N, \\ &\lim_{N\to\infty} \frac{n_{ik}^{(\ell)}(N)}{N} = \tau(p_{ik}), \qquad 0 \leq k \leq s(i) \wedge \ell. \end{aligned} \tag{5.9}$$

Moreover, choose orthogonal projections $P_{ik}^{(\ell)}(N) \in M_N(\mathbb{C})$ of rank $n_{ik}^{(\ell)}(N)$ for $0 \leq k \leq s(i) \wedge \ell$ so that we can identify

$$P_{ik}^{(\ell)}(N) M_N(\mathbb{C}) P_{ik}^{(\ell)}(N) = M_{m_{ik}}(\mathbb{C}) \otimes M_{m_{ik}^{(\ell)}(N)}(\mathbb{C}), \qquad 1 \leq k \leq s(i) \wedge \ell.$$

Under this identification, we set

$$\eta_{\alpha\beta}^{(ik\ell)}(N) := e_{\alpha\beta}^{(ik)} \otimes I_{m_{ik}^{(\ell)}(N)} \in M_{m_{ik}}(\mathbb{C}) \otimes M_{m_{ik}^{(\ell)}(N)}(\mathbb{C}),$$

for $1 \leq \alpha, \beta \leq m_{ik}$ and $1 \leq k \leq s(i) \wedge \ell$. Also we set

$$Q_i^{(\ell)}(N) := I - \sum_{k=0}^{s(i)\wedge\ell} P_{ik}^{(\ell)}(N).$$

Now let us start the proof of the inequality $\leq$. For each $\ell \in \mathbb{N}$ fixed, we enlarge the given multi-variables $\mathbf{X}_i = (X_{i1}, \ldots, X_{ir(i)})$, $1 \leq i \leq n$, as follows:

$$\mathbf{X}_i^{(\ell)} := \mathbf{X}_i \sqcup \bigsqcup_{k=1}^{s(i)\wedge\ell} \left(e_{11}^{(ik)}, \ldots, e_{m_{ik}m_{ik}}^{(ik)}, \sum_{\alpha,\beta=1}^{m_{ik}} e_{\alpha\beta}^{(ik)} \right), \qquad 1 \leq i \leq n.$$

Since $W^*(\mathbf{X}_i^{(\ell)}) = W^*(\mathbf{X}_i)$, Proposition 5.3 (4) gives

$$\delta_{0,\mathrm{orb}}(\mathbf{X}_1, \ldots, \mathbf{X}_n) = \delta_{0,\mathrm{orb}}(\mathbf{X}_1^{(\ell)}, \ldots, \mathbf{X}_n^{(\ell)}) \tag{5.10}$$

Moreover, since $\mathbf{X}_i \subset \mathbf{X}_i^{(\ell)}$, by [25, Theorem 4.3] one has

$$\delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n) \leq \delta_0(\mathbf{X}_1^{(\ell)} \sqcup \cdots \sqcup \mathbf{X}_n^{(\ell)}). \tag{5.11}$$

The next lemma is plain to show by the use of Lemma 1.2.

Lemma 5.10. *For each $1 \leq i \leq n$ and for any sufficiently large $N \in \mathbb{N}$, one can find microstates $\xi_{ij}^{(\ell)}(N) \in M_N^{sa}$, $1 \leq j \leq r(i)$, in such a way that $\|\xi_{ij}^{(\ell)}(N)\|_\infty \leq \|\mathbf{X}_i\|_\infty$, the $\xi_{ij}^{(\ell)}(N)$'s are contained in*

$$P_{i0}^{(\ell)}(N) M_N(\mathbb{C}) P_{i0}^{(\ell)}(N) \oplus \left[\bigoplus_{k=1}^{s(i)\wedge\ell} M_{m_{ik}}(\mathbb{C}) \otimes \mathbb{C} I_{m_{ik}^{(\ell)}(N)} \right] \oplus Q_i^{(\ell)}(N) M_N(\mathbb{C}) Q_i^{(\ell)}(N),$$

and moreover

$$\Xi_i^{(\ell)}(N) := \big(\xi_{i1}^{(\ell)}(N), \dots, \xi_{ir(i)}^{(\ell)}(N)\big) \sqcup \bigsqcup_{k=1}^{s(i)\wedge\ell} \left(\eta_{11}^{(ik\ell)}(N), \dots, \eta_{m_{ik}m_{ik}}^{(ik\ell)}(N), \sum_{\alpha,\beta=1}^{m_{ik}} \eta_{\alpha\beta}^{(ik\ell)}(N)\right)$$

converges in the distribution sense to $\mathbf{X}_i^{(\ell)}$ *as* $N \to \infty$.

Proof. For each $1 \le i \le n$, from hyperfiniteness one can choose an approximating sequence of microstates

$$(\xi_1(N), \dots, \xi_{r(i)}(N)) \sqcup (P(N)) \sqcup \bigsqcup_{k=1}^{s(i)\wedge\ell} \Big(\zeta_{11}^{(k)}(N), \dots, \zeta_{m_{ik}m_{ik}}^{(k)}(N), \zeta^{(k)}(N)\Big) \sqcup (Q(N))$$

for

$$(X_{i1}, \dots, X_{ir(i)}) \sqcup (p_{i0}) \sqcup \bigsqcup_{k=1}^{s(i)\wedge\ell} \left(e_{11}^{(ik)}, \dots, e_{m_{ik}m_{ik}}^{(ik)}, \sum_{\alpha,\beta=1}^{m_{ik}} e_{\alpha\beta}^{(ik)}\right) \sqcup (q_i^{(\ell)})$$

with the norm condition $\|\xi_j(N)\|_\infty \le \|X_{ij}\|_\infty$, $1 \le j \le r(i)$. Compare these $P(N)$, $Q(N)$, $\zeta_{\alpha\alpha}^{(k)}(N)$'s and $\zeta^{(k)}(N)$'s with $P_{i0}^{(\ell)}(N)$, $Q_i^{(\ell)}(N)$, $\eta_{\alpha\alpha}^{(ik\ell)}(N)$'s and $\sum_{\alpha,\beta=1}^{m_{ik}} \eta_{\alpha\beta}^{(ik\ell)}(N)$'s in mixed moments. By Lemma 1.2, for sufficiently large N we then get unitaries U_N which intertwine the two families approximately in the sense that $\|U_N P(N) U_N^* - P_{i0}^{(\ell)}(N)\|_{m,\mathrm{tr}_N}$ etc. go to 0 as $N \to \infty$ for all $m \in \mathbb{N}$. Then one can get a new approximating sequence of microstates for $X_{ij}p_{i0}$, $X_{ij}q_i^{(\ell)}$ $(1 \le j \le r(i))$ and $e_{11}^{(ik)}, \dots, e_{m_{ik}m_{ik}}^{(ik)}$, $\sum_{\alpha,\beta=1}^{m_{ik}} e_{\alpha\beta}^{(ik)}$ $(1 \le k \le s(i) \wedge \ell)$ by sending the $\xi_j(N)$'s via $\mathrm{Ad}\, U_N$ and cutting with $P_{i0}^{(\ell)}(N)$ or $Q_i^{(\ell)}(N)$ such that the part of those corresponding to $e_{11}^{(ik)}, \dots, e_{m_{ik}m_{ik}}^{(ik)}$, $\sum_{\alpha,\beta=1}^{m_{ik}} e_{\alpha\beta}^{(ik)}$ $(1 \le k \le s(i) \wedge \ell)$ are exactly $\eta_{11}^{(ik\ell)}(N), \dots, \eta_{m_{ik}m_{ik}}^{(ik\ell)}(N)$, $\sum_{\alpha,\beta=1}^{m_{ik}} \eta_{\alpha\beta}^{(ik\ell)}(N)$ $(1 \le k \le s(i) \wedge \ell)$. Then the desired microstates can easily be made from those. □

Remark that the commutant of $\Xi_i^{(\ell)}(N)$ includes

$$\mathbb{C}P_{i0}^{(\ell)}(N) \oplus \left[\bigoplus_{k=1}^{s(i)\wedge\ell} \mathbb{C}I_{m_{ik}} \otimes M_{m_{ik}^{(\ell)}(N)}(\mathbb{C})\right] \oplus \mathbb{C}Q_i^{(\ell)}(N).$$

We denote by $\mathrm{U}_i^{(\ell)}(N)$ the unitary group of this algebra, i.e.,

$$\mathrm{U}_i^{(\ell)}(N) := \mathbb{T}P_{i0}^{(\ell)}(N) \oplus \left[\bigoplus_{k=1}^{s(i)\wedge\ell} I_{m_{ik}} \otimes \mathrm{U}(m_{ik}^{(\ell)}(N))\right] \oplus \mathbb{T}Q_i^{(\ell)}(N). \tag{5.12}$$

We then have

$$\lim_{N\to\infty} \frac{1}{N^2} \dim_{\mathbb{R}} \mathrm{U}_i^{(\ell)}(N) = \lim_{N\to\infty} \frac{2 + \sum_{k=1}^{s(i)\wedge\ell} m_{ik}^{(\ell)}(N)^2}{N^2} = \sum_{k=1}^{s(i)\wedge\ell} \frac{\tau(p_{ik})^2}{m_{ik}^2}. \tag{5.13}$$

Consider the embedding

$$\Psi_N^{(\ell)} : ([U_i])_{i=1}^n \in \prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N) \mapsto (U_i \Xi_i^{(\ell)}(N) U_i^*)_{i=1}^n \in (M_N^{sa})^{n(\ell)},$$

where $n(\ell)$ is the sum of the numbers of variables in $\mathbf{X}_i^{(\ell)}$ for $1 \le i \le n$ and $[U_i]$ denotes the coset determined by $U_i \in \mathrm{U}(N)$. We introduce the "embedding" metric $d_{2,E}$ on the homogeneous space $\prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)$ by

$$d_{2,E}(([U_i])_{i=1}^n, ([V_i])_{i=1}^n) := \|\Psi_N^{(\ell)}(([U_i])_{i=1}^n) - \Psi_N^{(\ell)}(([V_i])_{i=1}^n)\|_{2,\mathrm{tr}_N}$$

for $([U_i])_{i=1}^n, ([V_i])_{i=1}^n \in \prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)$ with $U_i, V_i \in \mathrm{U}(N)$, $1 \le i \le n$. Another natural metric on $\prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)$ is the quotient metric $d_{2,Q}$ induced from d_2 on $\mathrm{U}(N)$ by

$$d_{2,Q}([U_i]_{i=1}^n, [V_i]_{i=1}^n) := \inf\left\{ d_2((U_i)_{i=1}^n, (V_iW_i)_{i=1}^n) : (W_i)_{i=1}^n \in \prod_{i=1}^n \mathrm{U}_i^{(\ell)}(N) \right\}.$$

It is plain to see that

$$d_{2,E}(([U_i])_{i=1}^n, ([V_i])_{i=1}^n) \le C_\ell \, d_{2,Q}(([U_i])_{i=1}^n, ([V_i])_{i=1}^n) \tag{5.14}$$

for all $([U_i])_{i=1}^n, ([V_i])_{i=1}^n \in \prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)$, where

$$C_\ell := 2n(\ell) \max\{\|\mathbf{X}_i\|_\infty, \, m_{ik} : 1 \le k \le s(i) \wedge \ell, \, 1 \le i \le n\}.$$

Viewing $\prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)$ as $\mathrm{U}(N)^n / \prod_{i=1}^n \mathrm{U}_i^{(\ell)}(N)$ we have the canonical quotient map

$$\Phi_N^{(\ell)} : \mathrm{U}(N)^n \to \prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N).$$

Now let $\varepsilon > 0$ be arbitrary, and note that the ε-covering number of $\Gamma(\mathbf{X}_1^{(\ell)} \sqcup \cdots \sqcup \mathbf{X}_n^{(\ell)} : (\Xi_i^{(\ell)}(N))_{i=1}^n; N, m, \delta)$ with respect to d_2 is equal to that of $\Phi_N^{(\ell)}(\Gamma_{\mathrm{orb}}((\mathbf{X}_i^{(\ell)})_{i=1}^n : (\Xi_i^{(\ell)}(N))_{i=1}^n; N, m, \delta))$ with respect to $d_{2,E}$ since $\Psi_N^{(\ell)}$ isometrically maps the latter set to the former. By (5.14) we hence get

$$\begin{aligned} &K_\varepsilon(\Gamma(\mathbf{X}_1^{(\ell)} \sqcup \cdots \sqcup \mathbf{X}_n^{(\ell)} : (\Xi_i^{(\ell)}(N))_{i=1}^n; N, m, \delta)) \\ &\quad \le K_{C_\ell\varepsilon}(\Phi_N^{(\ell)}(\Gamma_{\mathrm{orb}}((\mathbf{X}_i^{(\ell)})_{i=1}^n : (\Xi_i^{(\ell)}(N))_{i=1}^n; N, m, \delta))), \end{aligned} \tag{5.15}$$

where the above right-hand side is counted with respect to $d_{2,Q}$. Here we need the following simple (probably known) fact on the packing/covering numbers in homogeneous spaces. For the convenience of the reader we give it with a proof.

Lemma 5.11. (cf. [19, Lemma 6]) *Let G be a compact group with a bi-invariant metric d, and H be its closed subgroup. Let $\pi : G \to G/H$ be the canonical quotient map sending $g \in G$ to the coset gH, and equip G/H with the quotient metric $d_Q(g_1H, g_2H) := \min\{d(g_1, g_2h) : h \in H\}$. Then, for any $\Gamma \subset G$ with $\pi^{-1}(\pi(\Gamma)) = \Gamma$ and for every $\varepsilon > 0$, one has*

$$K_\varepsilon(\Gamma) \ge K_\varepsilon(H) \cdot P_{2\varepsilon}(\pi(\Gamma)) \ge K_\varepsilon(H) \cdot K_{4\varepsilon}(\pi(\Gamma)).$$

Proof. The ball centered at x of radius r in a metric space is denoted by $B_r(x)$. One can choose an ε-net $\{g_i : i \in I\}$ of Γ with cardinality $|I| = K_\varepsilon(\Gamma)$, and a 2ε-separated set $\{g'_j : j \in J\}$ of $\pi(\Gamma)$ with cardinality $|J| = P_{2\varepsilon}$. Let $I_j := \{i \in I : B_\varepsilon(g_i) \cap g'_jH \ne \emptyset\}$ for $j \in J$; then it is clear that $I_{j_1} \cap I_{j_2} = \emptyset$ if $j_1 \ne j_2$. On the other hand, $\{{g'_j}^{-1}g_i : i \in I_j\}$ gives an ε-net of H so that $|I_j| \ge K_\varepsilon(H)$ for all $j \in J$. Hence $K_\varepsilon(\Gamma) = |I| \ge \sum_{j\in J} |I_j| \ge K_\varepsilon(H) \cdot P_{2\varepsilon}(\pi(\Gamma))$. Then the assertion follows thanks to the obvious relation between covering and packing numbers. □

By the above lemma and (5.15) we have

$$K_\varepsilon(\Gamma(\mathbf{X}_1^{(\ell)} \sqcup \cdots \sqcup \mathbf{X}_n^{(\ell)} : (\Xi_i^{(\ell)}(N))_{i=1}^n; N, m, \delta))$$

$$\leq K_{(C_\ell/4)\varepsilon}(\Gamma_{\rm orb}((\mathbf{X}_i^{(\ell)})_{i=1}^n : (\Xi_i^{(\ell)}(N))_{i=1}^n; N, m, \delta)) \cdot K_{(C_\ell/4)\varepsilon}\left(\prod_{i=1}^n \mathrm{U}_i^{(\ell)}(N)\right)^{-1}. \tag{5.16}$$

Identify

$$\prod_{i=1}^n \mathrm{U}_i^{(\ell)}(N) = \prod_{i=1}^n \left(\mathbb{T} \times \left(\prod_{k=1}^{s(i)\wedge\ell} \mathrm{U}(m_{ik}^{(\ell)}(N))\right) \times \mathbb{T}\right)$$

in the obvious way, and consider the ℓ^∞-product metric induced from $\|\cdot\|_{2,\mathrm{tr}_{m_{ik}^{(\ell)}(N)}}$ on $\mathrm{U}(m_{ik}^{(\ell)}(N))$. Since the m_{ik}'s as well as ℓ are independent of N in (5.12), this metric is clearly equivalent to the original metric d_2 (restricted on $\prod_{i=1}^n \mathrm{U}_i^{(\ell)}(N)$) uniformly in all sufficiently large N. Then, by [20, Theorem 7] with the help of [19, Lemma 5], there is a constant $C'_\ell > 0$ independent of N such that

$$K_{(C_\ell/4)\varepsilon}\left(\prod_{i=1}^n \mathrm{U}_i^{(\ell)}(N)\right) \geq \left(\frac{C'_\ell}{(C_\ell/4)\varepsilon}\right)^{\sum_{i=1}^n \dim_{\mathbb{R}} \mathrm{U}_i^{(\ell)}(N)} \tag{5.17}$$

as long as $\varepsilon > 0$ is small enough. By using (5.16), (5.17) and (5.13) we thus get

$$\begin{aligned}
&\lim_{m\to\infty,\delta\searrow 0} \limsup_{N\to\infty} \frac{1}{N^2} \log K_\varepsilon(\Gamma(\mathbf{X}_1^{(\ell)} \sqcup \cdots \sqcup \mathbf{X}_n^{(\ell)} : (\Xi_i^{(\ell)}(N))_{i=1}^n; N, m, \delta)) \\
&\quad \leq \mathbb{K}^{\rm orb}_{C_\ell\varepsilon}(\mathbf{X}_1^{(\ell)}, \dots, \mathbf{X}_n^{(\ell)}) + \left(\log\left(\frac{C_\ell}{4C'_\ell}\right) + \log\varepsilon\right) \sum_{i=1}^n \sum_{k=1}^{s(i)\wedge\ell} \frac{\tau(p_{ik})^2}{m_{ik}^2}.
\end{aligned}$$

By Lemma 5.9, (5.11), Proposition 5.6 and (5.10), this implies that

$$\delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n) \leq \delta_{0,\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n) + n - \sum_{i=1}^n \sum_{k=1}^{s(i)\wedge\ell} \frac{\tau(p_{ik})^2}{m_{ik}^2}$$

Since ℓ is arbitrary, we get the inequality $\leq$ in Theorem 5.8 thanks to (5.8).

Let us turn to the proof of the inequality $\geq$. Keep an arbitrary $\ell \in \mathbb{N}$. Since M_{i0} is diffuse, one can choose $p_{i01}^{(\ell)}, \dots, p_{i0\ell}^{(\ell)} \in M_{i0}$ in such a way that $p_{i01}^{(\ell)} + \cdots + p_{i0\ell}^{(\ell)} = p_{i0}$ and $\tau(p_{i01}^{(\ell)}) = \cdots = \tau(p_{i0\ell}^{(\ell)})$ $(= \tau(p_{i0})/\ell)$. For $1 \leq i \leq n$ consider a new hyperfinite random multi-variables $\mathbf{Y}_i^{(\ell)}$ in $W^*(\mathbf{X}_i)$ given by

$$\mathbf{Y}_i^{(\ell)} := (p_{i01}^{(\ell)}, \dots, p_{i0\ell}^{(\ell)}) \sqcup \bigsqcup_{k=1}^{s(i)\wedge\ell} \left(e_{11}^{(ik)}, \dots, e_{m_{ik}m_{ik}}^{(ik)}, \sum_{\alpha,\beta=1}^{m_{ik}} e_{\alpha\beta}^{(ik)}\right) \sqcup (q_i^{(\ell)}).$$

Then, similarly to (5.10) one has

$$\delta_{0,\rm orb}(\mathbf{X}_1, \dots, \mathbf{X}_n) = \delta_{0,\rm orb}(\mathbf{X}_1 \sqcup \mathbf{Y}_1^{(\ell)}, \dots, \mathbf{X}_n \sqcup \mathbf{Y}_n^{(\ell)}). \tag{5.18}$$

On the other hand, one can apply [14, Corollary 4.2], a corollary of the so-called hyperfinite inequality due to Jung, to obtain the equality

$$\delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n) = \delta_0(\mathbf{X}_1 \sqcup \mathbf{Y}_1^{(\ell)} \sqcup \cdots \sqcup \mathbf{X}_n \sqcup \mathbf{Y}_n^{(\ell)}) \tag{5.19}$$

unlike the previous (5.11). Let $n_{ik}^{(\ell)}(N)$, $P_{ik}^{(\ell)}(N)$ $(0 \leq k \leq s(i)\wedge\ell)$, $m_{ik}^{(\ell)}(N)$ $(1 \leq k \leq s(i)\wedge\ell)$ and $Q_i^{(\ell)}(N)$ be as in the proof of the part "$\leq$." Moreover, for any sufficiently large N, one can choose $n_{i01}^{(\ell)}(N), \dots, n_{i0\ell}^{(\ell)}(N) \in \mathbb{N}$ so that

$$n_{i01}^{(\ell)}(N) + \cdots + n_{i0\ell}^{(\ell)}(N) = n_{i0}^{(\ell)}(N),$$

$$\lim_{N\to\infty}\frac{n_{i0k}^{(\ell)}(N)}{N}=\frac{\tau(p_{i0})}{\ell},\qquad 1\le k\le \ell. \tag{5.20}$$

Then, choose orthogonal projections $P_{i0k}^{(\ell)}(N)\in M_N(\mathbb{C})$ of rank $n_{i0k}^{(\ell)}(N)$ for $0\le k\le\ell$ such that $\sum_{k=1}^{\ell}P_{i0k}^{(\ell)}(N)=P_{i0}^{(\ell)}(N)$. A special approximating microstates for $\mathbf{Y}_i^{(\ell)}$ is given by

$$\begin{aligned}\widehat{\Xi}_i^{(\ell)} &:= (P_{i01}^{(\ell)}(N),\dots,P_{i0\ell}^{(\ell)}(N))\\ &\qquad \sqcup \bigsqcup_{k=1}^{s(i)\wedge\ell}\left(\eta_{11}^{(ik\ell)}(N),\dots,\eta_{m_{ik},m_{ik}}^{(ik\ell)}(N),\sum_{\alpha,\beta=1}^{m_{ik}}\eta_{\alpha\beta}^{(ik\ell)}(N)\right)\sqcup(Q_i^{(\ell)}(N)).\end{aligned}$$

The next lemma is proven in the same way as Lemma 5.10, so the details are left to the reader.

Lemma 5.12. *For $1\le i\le n$ and for any sufficiently large N, one can find microstates $\Xi_i^{(\ell)}(N)$ for $\mathbf{X}_i$ in such a way that $\|\Xi_i^{(\ell)}(N)\|_\infty\le\|\mathbf{X}_i\|_\infty$ and $\Xi_i^{(\ell)}(N)\sqcup\widehat{\Xi}_i^{(\ell)}(N)$ converges to $\mathbf{X}_i\sqcup\mathbf{Y}_i^{(\ell)}$ in the distribution sense as $N\to\infty$.*

We denote by $\mathrm{U}_i^{(\ell)}(N)$ the unitary group of the commutant of $\Xi_i^{(\ell)}(N)\sqcup\widehat{\Xi}_i^{(\ell)}(N)$ and by $\widehat{\mathrm{U}}_i^{(\ell)}(N)$ that of the commutant of $\widehat{\Xi}_i^{(\ell)}$. Since the commutant of $\widehat{\Xi}_i^{(\ell)}$ is

$$\begin{aligned}&\bigoplus_{k=1}^{\ell}P_{i0k}^{(\ell)}(N)M_N(\mathbb{C})P_{i0k}^{(\ell)}(N)\\ &\qquad\oplus\left[\bigoplus_{k=1}^{s(i)\wedge\ell}\mathbb{C}I_{m_{ik}}\otimes M_{m_{ik}^{(\ell)}(N)}(\mathbb{C})\right]\oplus Q_i^{(\ell)}(N)M_N(\mathbb{C})Q_i^{(\ell)}(N),\end{aligned}$$

the real dimension of $\widehat{\mathrm{U}}_i^{(\ell)}(N)$ is

$$\dim_{\mathbb{R}}\widehat{\mathrm{U}}_i^{(\ell)}(N)=\sum_{k=1}^{\ell}n_{i0k}^{(\ell)}(N)^2+\sum_{k=1}^{s(i)\wedge\ell}m_{ik}^{(\ell)}(N)^2+\left(N-n_{i0}^{(\ell)}(N)-\sum_{k=1}^{s(i)\wedge\ell}n_{ik}^{(\ell)}(N)\right)^2$$

so that by (5.9) and (5.20) we have

$$\lim_{N\to\infty}\frac{N^2-\dim_{\mathbb{R}}\widehat{\mathrm{U}}_i^{(t\ell)}(N)}{N^2}=1-\frac{\tau(p_{i0})^2}{\ell}-\sum_{k=1}^{s(i)\wedge\ell}\frac{\tau(p_{ik})^2}{m_{ik}^2}-\tau(q_i^{(\ell)})^2. \tag{5.21}$$

Introduce the embeddings

$$\begin{aligned}\Psi_N^{(\ell)}&:([U_i])_{i=1}^n\in\prod_{i=1}^n\mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)\mapsto\big(U_i(\Xi_i^{(\ell)}(N)\sqcup\widehat{\Xi}_i^{(\ell)}(N))U_i^*\big)_{i=1}^n\in(M_N^{sa})^{n(\ell)},\\ \widehat{\Psi}_N^{(\ell)}&:(\langle U_i\rangle)_{i=1}^n\in\prod_{i=1}^n\mathrm{U}(N)/\widehat{\mathrm{U}}_i^{(\ell)}(N)\mapsto\big(U_i(\widehat{\Xi}_i^{(\ell)}(N))U_i^*\big)_{i=1}^n\in(M_N^{sa})^{\widehat{n}(\ell)},\end{aligned}$$

where $n(\ell)$ is the sum of the numbers of variables in $\mathbf{X}_i\sqcup\mathbf{Y}_i^{(\ell)}$, $1\le i\le n$, and $\widehat{n}(\ell)$ is that of variables in $\mathbf{Y}_i^{(\ell)}$, $1\le i\le n$. Moreover, we introduce the "embedding" metric $\widehat{d}_{2,E}$ in terms of $\widehat{\Psi}_N^{(\ell)}$ and the quotient metric $\widehat{d}_{2,Q}$ on the homogeneous space $\prod_{i=1}^n\mathrm{U}(N)/\widehat{\mathrm{U}}_i^{(\ell)}(N)$ in the same way as in the proof of the part "$\le$." The "embedding" metric $d_{2,E}$ in terms of $\Psi_N^{(\ell)}$ is also introduced on $\prod_{i=1}^n\mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)$. From the trivial inclusions $\mathrm{U}_i^{(\ell)}(N)\subset\widehat{\mathrm{U}}_i^{(\ell)}(N)$ and

$\widehat{\Xi}_i^{(\ell)}(N) \subset \Xi_i^{(\ell)}(N) \sqcup \widehat{\Xi}_i^{(\ell)}(N)$, we have the well-defined surjective map

$$([U_i])_{i=1}^n \in \prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N) \mapsto (\langle U_i\rangle)_{i=1}^n \in \prod_{i=1}^n \mathrm{U}(N)/\widehat{\mathrm{U}}_i^{(\ell)}(N)$$

so that

$$\widehat{d}_{2,E}((\langle U_i\rangle)_{i=1}^n, (\langle V_i\rangle)_{i=1}^n) \le d_{2,E}(([U_i])_{i=1}^n, ([V_i])_{i=1}^n). \tag{5.22}$$

The next lemma is essentially [12, Lemma 5.4]. In fact, the argument there works well when $\bigoplus_{k=1}^{\ell} \mathbb{C}p_{i0k}^{(\ell)} \oplus \bigoplus_{k=1}^{s(i)\wedge\ell} M_{ik} \oplus \mathbb{C}q_i^{(\ell)}$ and $\Xi_i^{(\ell)}(N)$ here play the roles of M and π there. Thus a chosen constant C_ℓ depends only on m_{ik}, $1 \le k \le s(i)\wedge\ell$, $1 \le i \le n$, as well as ℓ.

Lemma 5.13. *There is a constant $C_\ell > 0$ independent of N such that*

$$\widehat{d}_{2,Q}((\langle U_i\rangle)_{i=1}^n, (\langle V_i\rangle)_{i=1}^n) \le C_\ell\, \widehat{d}_{2,E}((\langle U_i\rangle)_{i=1}^n, (\langle V_i\rangle)_{i=1}^n)$$

for all $(\langle U_i\rangle)_{i=1}^n, (\langle V_i\rangle)_{i=1}^n \in \prod_{i=1}^n \mathrm{U}(N)/\widehat{\mathrm{U}}_i^{(\ell)}(N)$.

Viewing $\prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)$ as $\mathrm{U}(N)^n/\prod_{i=1}^n \mathrm{U}_i^{(\ell)}(N)$ we have the canonical quotient map

$$\Phi_N^{(\ell)} : \mathrm{U}(N)^n \to \prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N),$$

and denote by $\mu_N^{(\ell)}$ the left-invariant probability measure on $\prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)$ induced from $\gamma_{\mathrm{U}(N)}^{\otimes n}$. In what follows, let $\varepsilon, \delta > 0$ be arbitrary with restriction $\delta < \varepsilon$, and for $N, m \in \mathbb{N}$ we write for short

$$\begin{aligned}
\Gamma_{\mathrm{orb}}(\varepsilon, N, 3m, \delta) &:= \Gamma_{\mathrm{orb}}((v_i(\varepsilon)(\mathbf{X}_i \sqcup \mathbf{Y}_i^{(\ell)})v_i(\varepsilon)^*)_{i=1}^n : \\
&\qquad\qquad (\Xi_i^{(\ell)}(N) \sqcup \widehat{\Xi}_i^{(\ell)}(N))_{i=1}^n : \mathbf{v}(\varepsilon); N, 3m, \delta), \\
\Gamma_{\mathrm{orb}}(N, m, \delta) &:= \Gamma_{\mathrm{orb}}((\mathbf{X}_i \sqcup \mathbf{Y}_i^{(\ell)})_{i=1}^n : (\Xi_i^{(\ell)}(N) \sqcup \widehat{\Xi}_i^{(\ell)}(N))_{i=1}^n; N, m, \delta), \\
\Gamma(N, m, \delta) &:= \Gamma(\mathbf{X}_1 \sqcup \mathbf{Y}_1^{(\ell)} \sqcup \cdots \sqcup \mathbf{X}_n \sqcup \mathbf{Y}_n^{(\ell)} : (\Xi_i^{(\ell)}(N) \sqcup \widehat{\Xi}_n^{(\ell)}(N))_{i=1}^n : N, m, \delta)
\end{aligned}$$

(see (5.7)). The following inequality is trivial:

$$\gamma_{\mathrm{U}(N)}^{\otimes n}(\Gamma_{\mathrm{orb}}(\varepsilon, N, 3m, \delta)) \le \mu_N^{(\ell)}(\Phi_N^{(\ell)}(\Gamma_{\mathrm{orb}}(\varepsilon, N, 3m, \delta))). \tag{5.23}$$

Assume $(U_i)_{i=1}^n \in \Gamma_{\mathrm{orb}}(\varepsilon, N, 3m, \delta)$; then there is a $(V_i)_{i=1}^n \in \Gamma(\mathbf{v}(\varepsilon); N, 3m, \delta)$ such that $(V_i^* U_i)_{i=1}^n \in \Gamma_{\mathrm{orb}}(N, m, \delta)$ and hence for any $\xi \in \Xi_i^{(\ell)}(N) \sqcup \widehat{\Xi}_i^{(\ell)}(N)$ we have

$$\|U_i \xi U_i^* - V_i^* U_i \xi U_i^* V_i\|_{2,\mathrm{tr}_N} \le 2R(C\varepsilon)^{1/2}$$

so that

$$d_{2,E}(([U_i])_{i=1}^n, ([V_i^* U_i])_{i=1}^n) \le 2R(n(\ell)C\varepsilon)^{1/2},$$

where $R := \max\{\|\mathbf{X}_i \sqcup \mathbf{Y}_i^{(\ell)}\|_\infty, m_{ik} : 1 \le k \le s(i)\wedge\ell, 1 \le i \le n\}$ and $C > 0$ is the same constant as in the proof of Proposition 5.6. Therefore we get

$$\Phi_N^{(\ell)}(\Gamma_{\mathrm{orb}}(\varepsilon, N, 3m, \delta)) \subset \mathcal{N}_{L\varepsilon^{1/2}}(\Phi_N^{(\ell)}(\Gamma_{\mathrm{orb}}(N, m, \delta))),$$

where the right-hand side is the open $L\varepsilon^{1/2}$-neighborhood of $\Gamma_{\mathrm{orb}}(N, m, \delta)$ with respect to the metric $d_{2,E}$ and $L := 2R(n(\ell)C)^{1/2} + 1$. Here note that the $\varepsilon^{1/2}$-covering number of $\Phi_N^{(\ell)}(\Gamma_{\mathrm{orb}}(N, m, \delta))$ with respect to $d_{2,E}$ is equal to $K_{\varepsilon^{1/2}}(\Gamma(N, m, \delta))$ with respect to d_2 (as noted just above (5.16)). Hence the above inclusion immediately implies that

$$\mu_N^{(\ell)}(\Phi_N^{(\ell)}(\Gamma_{\mathrm{orb}}(\varepsilon, N, 3m, \delta))) \le K_{\varepsilon^{1/2}}(\Gamma(N, m, \delta)) \cdot \mu_N^{(\ell)}(\mathrm{Ball}((L+1)\varepsilon^{1/2}, d_{2,E})), \tag{5.24}$$

where $\mathrm{Ball}((L+1)\varepsilon^{1/2}, d_{2,E})$ stands for the $(L+1)\varepsilon^{1/2}$-ball in $\prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N)$ with respect to $d_{2,E}$. The measure of this ball can be estimated from above by packing numbers as follows:

$$\begin{aligned}\mu_N^{(\ell)}(\mathrm{Ball}((L+1)\varepsilon^{1/2}, d_{2,E}))^{-1} &\geq P_{(L+1)\varepsilon^{1/2}}\left(\prod_{i=1}^n \mathrm{U}(N)/\mathrm{U}_i^{(\ell)}(N), d_{2,E}\right) \\ &\geq P_{(L+1)\varepsilon^{1/2}}\left(\prod_{i=1}^n \mathrm{U}(N)/\widehat{\mathrm{U}}_i^{(\ell)}(N), \widehat{d}_{2,E}\right) \\ &\geq P_{C_\ell(L+1)\varepsilon^{1/2}}\left(\prod_{i=1}^n \mathrm{U}(N)/\widehat{\mathrm{U}}_i^{(\ell)}(N), \widehat{d}_{2,Q}\right).\end{aligned}$$

The second and the third inequalities in the above follow from (5.22) and Lemma 5.13, respectively. Furthermore, the packing estimate due to Jung [12, Lemma 5.2 and §8] (based on [20]) says that there is a constant $C'_\ell > 0$ independent of N such that

$$P_{C_\ell(L+1)\varepsilon^{1/2}}\left(\prod_{i=1}^n \mathrm{U}(N)/\widehat{\mathrm{U}}_i^{(\ell)}(N), \widehat{d}_{2,Q}\right) \geq \left(\frac{C'_\ell}{C_\ell(L+1)\varepsilon^{1/2}}\right)^{nN^2 - \sum_{i=1}^n \dim_{\mathbb{R}} \widehat{U}_i^{(\ell)}(N)}$$

as long as $\varepsilon > 0$ is small enough. Therefore we get

$$\mu_N^{(\ell)}(\mathrm{Ball}((L+1)\varepsilon^{1/2}, d_{2,E})) \leq \left(\frac{C_\ell(L+1)\varepsilon^{1/2}}{C'_\ell}\right)^{nN^2 - \sum_{i=1}^n \dim_{\mathbb{R}} \widehat{U}_i^{(\ell)}(N)} \tag{5.25}$$

for all sufficiently small $\varepsilon > 0$.

Combining (5.23)–(5.25) and (5.21) altogether implies that

$$\begin{aligned}&\limsup_{N\to\infty} \frac{1}{N^2} \log \gamma_{\mathrm{U}(N)^n}(\Gamma_{\mathrm{orb}}(\varepsilon, N, 3m, \delta)) \\ &\leq \limsup_{N\to\infty} \frac{1}{N^2} \log K_{\varepsilon^{1/2}}(\Gamma(N, m, \delta)) \\ &\qquad + \sum_{i=1}^n \left(1 - \frac{\tau(p_{i0})^2}{\ell} - \sum_{k=1}^{s(i)\wedge\ell} \frac{\tau(p_{ik})^2}{m_{ik}^2} - \tau(q_i^{(\ell)})^2\right) \cdot \left(\log \varepsilon^{1/2} + \log \frac{C_\ell(L+1)}{C'_\ell}\right)\end{aligned}$$

whenever $\varepsilon > 0$ is sufficiently small. Take $\lim_{m\to\infty,\delta\searrow 0}$ and then $\limsup_{\varepsilon\searrow 0}$ after dividing by $|\log \varepsilon^{1/2}|$; then by Lemma 5.9, (5.18), Proposition 5.6 and (5.19) we have

$$\delta_{0,\mathrm{orb}}(\mathbf{X}_1, \dots, \mathbf{X}_n) \leq \delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n) - \sum_{i=1}^n \left(1 - \frac{\tau(p_{i0})^2}{\ell} - \sum_{k=1}^{s(i)\wedge\ell} \frac{\tau(p_{ik})^2}{m_{ik}^2} - \tau(q_i^{(\ell)})^2\right).$$

Hence the inequality $\geq$ in Theorem 5.8 follows by taking $\ell \to \infty$ thanks to (5.8).

6. Applications

6.1. **Immediate corollaries.** The next corollary is immediate from Theorem 5.8 and (5)–(7) of Proposition 5.3.

Corollary 6.1. *Let $\mathbf{X}_1, \dots, \mathbf{X}_n$ be hyperfinite random multi-variables in (M, τ).*

(1) *If $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ are random multi-variables such that $\mathbf{Y}_i \subset W^*(\mathbf{X}_i)$ for $1 \leq i \leq n$, then*

$$\delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n) - \sum_{i=1}^n \delta_0(\mathbf{X}_i) \leq \delta_0(\mathbf{Y}_1 \sqcup \cdots \sqcup \mathbf{Y}_n) - \sum_{i=1}^n \delta_0(\mathbf{Y}_i).$$

In addition, if $W^*(\mathbf{Y}_i)$'s are all diffuse, then $\delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n) \leq \delta_0(\mathbf{Y}_1 \sqcup \cdots \sqcup \mathbf{Y}_n)$.

(2) *If* $\chi_{\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n) > -\infty$*, then* $\delta_0(\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n) = \sum_{i=1}^n \delta_0(\mathbf{X}_i)$*. In particular, this is the case if* $\mathbf{X}_1,\dots,\mathbf{X}_n$ *are freely independent.*

(3) *Let* $\mathbf{X}$ *be a hyperfinite random multi-variable and* $\mathbf{Y}$ *a general random multi-variable. If* $\mathbf{X}$ *is freely independent of* $\mathbf{Y}$*, then* $\delta_0(\mathbf{X}\sqcup\mathbf{Y}) = \delta_0(\mathbf{X}) + \delta_0(\mathbf{Y})$*.*

The following is an immediate corollary of Theorem 5.8 too. But we note that it can also be shown by a direct method of estimating the covering numbers of orbital microstate spaces.

Corollary 6.2. (General upper bound of $\delta_{0,\mathrm{orb}}$) *For any* n*-tuple of hyperfinite self-adjoint multi-variables* $\mathbf{X}_1,\dots,\mathbf{X}_n$*,*

$$\delta_{0,\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n) \le -(n-1)\,\delta_0\big(W^*(\mathbf{X}_1)\cap\cdots\cap W^*(\mathbf{X}_n)\big),$$

and equality holds either when the $\mathbf{X}_k$*'s are the same or when the* $\mathbf{X}_k$*'s are freely independent with amalgamation over their common subalgebra* $W^*(\mathbf{X}_1)\cap\cdots\cap W^*(\mathbf{X}_n)$*. Here,* $\delta_0\big(W^*(\mathbf{X}_1)\cap\cdots\cap W^*(\mathbf{X}_n)\big)$ *denotes the unique value of* $\delta_0(\mathbf{X})$ *with* $W^*(\mathbf{X}) = W^*(\mathbf{X}_1)\cap\cdots\cap W^*(\mathbf{X}_n)$ *due to* [12].

Proof. By Proposition 5.3 (5) we have

$$\delta_{0,\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n) \le \delta_{0,\mathrm{orb}}(\mathbf{X},\dots,\mathbf{X}),$$

and the right-hand side is $-(n-1)\,\delta_0(\mathbf{X})$ by Theorem 5.8. The second equality condition follows from [3]. □

6.2. **Liberation process vs δ_0.** Let $X_1,\dots,X_n$ be an n-tuple of single self-adjoint random variables and $S_1,\dots,S_n$ a standard semicircular system freely independent of $X_1,\dots,X_n$. Concerning the condition of $X_1,\dots,X_n$ having f.d.a. (i.e., finite-dimensional approximants), it is known (see [7, 7.3.9] and [13]) that the following are all equivalent:

- $X_1,\dots,X_n$ has f.d.a. (i.e., finite-dimensional approximants);
- $\chi(X_1+\varepsilon S_1,\dots,X_n+\varepsilon S_n) > -\infty$ for all $\varepsilon>0$;
- $\delta(X_1,\dots,X_n)\ge 0$ (also $\delta_0(X_1,\dots,X_n)\ge 0$);
- $\delta(X_1+\varepsilon S_1,\dots,X_n+\varepsilon S_n) = n$ (also $\delta_0(X_1+\varepsilon S_1,\dots,X_n+\varepsilon S_n) = n$) for all $\varepsilon>0$.

The next proposition gives similar equivalent conditions in terms of the orbital theory.

Proposition 6.3. *Let* $\mathbf{X}_1,\dots,\mathbf{X}_n$ *be hyperfinite random multi-variables, and* $\mathbf{v}(t) = (v_i(t))_{i=1}^n$*,* $t\ge 0$*, be a freely independent* n*-tuple of free unitary Brownian motions that is freely independent of* $\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n$*. Then the following conditions are equivalent:*

(i) $\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n$ *has f.d.a.*
(ii) $\chi_{\mathrm{orb}}(v_1(\varepsilon)\mathbf{X}_1v_1(\varepsilon)^*,\dots,v_n(\varepsilon)\mathbf{X}_nv_n(\varepsilon)^*) > -\infty$ *for all* $\varepsilon>0$*.*
(iii) $\delta_{0,\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n) > -\infty$*.*
(iv) $\delta_{0,\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n) \ge -n$*.*
(v) $\delta_{0,\mathrm{orb}}(v_1(\varepsilon)\mathbf{X}_1v_1(\varepsilon)^*,\dots,v_n(\varepsilon)\mathbf{X}_nv_n(\varepsilon)^*) = 0$ *for all* $\varepsilon>0$*.*

Proof. (i) ⇒ (iv). Note that (i) implies $\delta_0(\mathbf{X}_1\sqcup\cdots\sqcup\mathbf{X}_n)\ge 0$ as mentioned above. Hence $\delta_{0,\mathrm{orb}}(\mathbf{X}_1,\dots,\mathbf{X}_n) \ge -\sum_{i=1}^n \delta_0(\mathbf{X}_i) \ge -n$ by Theorem 5.8.

(iv) ⇒ (iii) is trivial.

(iii) ⇒ (ii). Condition (iii) implies that there is a sequence $\varepsilon_k \searrow 0$ such that

$$\begin{aligned}&\chi_{\mathrm{orb}}(v_1(\varepsilon_k)\mathbf{X}_1v_1(\varepsilon_k)^*,\dots,v_n(\varepsilon_k)\mathbf{X}_nv_n(\varepsilon_k)^*)\\ &\qquad\ge \chi_{\mathrm{orb}}(v_1(\varepsilon_k)\mathbf{X}_1v_1(\varepsilon_k)^*,\dots,v_n(\varepsilon_k)\mathbf{X}_nv_n(\varepsilon_k)^* : \mathbf{v}(\varepsilon_k)) > -\infty.\end{aligned}$$

Hence (ii) follows if we see that $\chi_{\mathrm{orb}}(v_1(t)\mathbf{X}_1v_1(t)^*,\dots,v_n(t)\mathbf{X}_nv_n(t)^*)$ is increasing in $t\ge 0$. To show this, let $t>s\ge 0$ and set $w_i(t,s) := v_i(t)v_i(s)^*$, $\mathbf{Y}_i := v_i(s)\mathbf{X}_iv_i(s)^*$. Note [1] that

$w_i(t,s)$ has the same distribution as $v_i(t-s)$ and is freely independent of $\mathbf{Y}_1 \sqcup \cdots \sqcup \mathbf{Y}_n$. Hence we have

$$\begin{aligned}\chi_{\mathrm{orb}}&(v_1(t)\mathbf{X}_1 v_1(t)^*, \dots, v_n(t)\mathbf{X}_n v_n(t)^*)\\ &= \chi_{\mathrm{orb}}(w_1(t,s)\mathbf{Y}_1 w_1(t,s)^*, \dots, w_n(t,s)\mathbf{Y}_n w_n(t,s)^*)\\ &\geq \chi_{\mathrm{orb}}(\mathbf{Y}_1, \dots, \mathbf{Y}_n)\end{aligned}$$

thanks to Proposition 4.6.

(ii) $\Rightarrow$ (v) is Proposition 5.3 (6).

(v) $\Rightarrow$ (i). From (v) there is a sequence $\varepsilon_k \searrow 0$ such that

$$\chi_{\mathrm{orb}}(v_1(\varepsilon_k)\mathbf{X}_1 v_1(\varepsilon_k)^*, \dots, v_n(\varepsilon_k)\mathbf{X}_n v_n(\varepsilon_k)^*) > -\infty.$$

Hence, for every $m \in \mathbb{N}$, $\delta > 0$ and $k \in \mathbb{N}$ we have

$$\Gamma_{\mathrm{orb}}((v_i(\varepsilon_k)\mathbf{X}_i v_i(\varepsilon_k)^*)_{i=1}^n : (\Xi_i(N))_{i=1}^n; N, m, \delta/2) \neq \emptyset$$

for sufficiently large N, where $\Xi_i(N) = (\xi_{i1}(N), \dots, \xi_{ir(i)}(N))$, $1 \leq i \leq n$, are chosen as in Definition 4.5 with $\|\xi_{ij}(N)\|_\infty \leq \|X_{ij}\|_\infty$. Since $v_i(\varepsilon_k)X_{ij}v_i(\varepsilon_k) \to X_{ij}$ strongly as $k \to \infty$ for $1 \leq j \leq r(i)$ with $\mathbf{X}_i = (X_{i1}, \dots, X_{ir(i)})$, $1 \leq i \leq n$, this implies that $\Gamma_R(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n; N, m, \delta) \neq \emptyset$ for every $R \geq \max\{\|X_{ij}\|_\infty : 1 \leq j \leq r(i), 1 \leq i \leq n\}$ if N is sufficiently large, so (i) follows. □

The next corollary is immediate from Theorem 5.8 and the above proposition.

Corollary 6.4. *If $\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n$ has f.d.a., then*

$$\delta_0(v_1(\varepsilon)\mathbf{X}_1 v_1(\varepsilon)^* \sqcup \cdots \sqcup v_n(\varepsilon)\mathbf{X}_n v_n(\varepsilon)^*) = \sum_{i=1}^n \delta_0(\mathbf{X}_i)$$

for every $\varepsilon > 0$. In particular, δ_0 is discontinuous at 0 in the liberation process.

It might be worth mentioning that the above corollary provides another route toward Brown's observation [2] based on the liberation process instead of the semicircular deformation.

6.3. **Lower semicontinuity for δ_0.** The next lemma partially strengthens [12, Lemma 7.3].

Lemma 6.5. *Let $\mathbf{X} = (X_1, \dots, X_r)$ be hyperfinite self-adjoint multi-variables. If $\mathbf{X}^{(k)} = (X_1^{(k)}, \dots, X_r^{(k)})$ is a sequence of (not necessarily hyperfinite) random multi-variables converging to $\mathbf{X}$ in mixed moments as $k \to \infty$, then one has*

$$\liminf_{k\to\infty} \delta_0(\mathbf{X}^{(k)}) \geq \delta_0(\mathbf{X}).$$

Proof. As same as in the proof of Theorem 5.8 we assume that $W^*(\mathbf{X})$ has both diffuse and atomic parts, and decompose

$$W^*(\mathbf{X}) = \bigoplus_{j=0}^s M_j$$

with possibly $s = \infty$, where M_0 is diffuse and $M_j \cong M_{m_j}(\mathbb{C})$ for $j \geq 1$. Let p_j be the central support projection of M_j. Fix an arbitrary $\ell \in \mathbb{N}$, and one can choose projections $p_{01}^{(\ell)}, \dots, p_{0\ell}^{(\ell)} \in M_0$ in such a way that $p_{01}^{(\ell)} + \cdots + p_{0\ell}^{(\ell)} = p_0$ and $\tau(p_{01}^{(\ell)}) = \cdots = \tau(p_{0\ell}^{(\ell)})$. Define a finite-dimensional subalgebra $\mathcal{P}_\ell \subset W^*(\mathbf{X})$ by

$$\mathcal{P}_\ell := \bigoplus_{i=1}^\ell \mathbb{C}p_{0i}^{(\ell)} \oplus \bigoplus_{j=1}^{s\wedge\ell} M_j \oplus \mathbb{C}q^{(\ell)},$$

with $q^{(\ell)} := \mathbf{1} - \sum_{j=0}^{s\wedge\ell} p_j$. Here we choose and fix a matrix unit system $\{e_{\alpha\beta}^{(j)}\}_{1\le\alpha,\beta\le m_j}$ of $M_j \cong M_{m_j}(\mathbb{C})$ for $1 \le j \le s\wedge\ell$. For any element e chosen from $p_{0i}^{(\ell)}$ $(1\le i\le\ell)$, $e_{\alpha\beta}^{(j)}$ $(1\le\alpha,\beta\le m_j,\ 1\le j\le s\wedge\ell)$ and q_ℓ (altogether forming a matrix unit system of $\mathcal{P}_\ell$) and for any $\delta>0$, one can choose a non-commutative polynomial P of r indeterminates such that $\|e - P(X_1,\dots,X_r)\|_{\tau,2} < \delta$. Then $\|e - P(X_1^{(k)},\dots,X_r^{(k)})\|_{\tau,2} < \delta$ for all sufficiently large k. This shows that, for every $\varepsilon>0$, one has $\mathcal{P}_\ell \subset_\delta W^*(\mathbf{X}^{(k)})$ for all k large enough. (See [5] for the notation "$\subset_\delta$".) Hence, by [5, Lemma 4] we see that, for every $\ell\in\mathbb{N}$ and every $\varepsilon>0$, there is a $k_0\in\mathbb{N}$ such that for every $k\ge k_0$ one can find a new matrix unit system (depending on k) $\hat p_{0i}$ $(1\le i\le\ell)$, $\hat e_{\alpha\beta}^{(j)}$ $(1\le\alpha,\beta\le m_j,\ 1\le j\le s\wedge\ell)$ and $\hat q$ inside $W^*(\mathbf{X}^{(k)})$ satisfying

$$\sum_{i=1}^{\ell}\hat p_{0i} + \sum_{j=1}^{s\wedge\ell}\sum_{\alpha=1}^{m_j}\hat e_{\alpha\alpha}^{(j)} + \hat q = \mathbf{1}$$

and

$$\begin{aligned}
&\|\hat p_{0i} - p_{0i}^{(\ell)}\|_{2,\tau} < \varepsilon, \quad 1\le i\le\ell,\\
&\|\hat e_{\alpha1}^{(j)} - e_{\alpha1}^{(j)}\|_{2,\tau} < \varepsilon, \quad 1\le\alpha\le m_j,\ 1\le j\le s\wedge\ell,\\
&\|\hat q - q^{(\ell)}\|_{2,\tau} < \varepsilon.
\end{aligned}$$

For every $k\ge k_0$, since $W^*(\mathbf{X}^{(k)})$ has a subalgebra (depending on k)

$$\bigoplus_{i=1}^{\ell}\mathbb{C}\hat p_{0i} \oplus \bigoplus_{j=1}^{s\wedge\ell}\mathrm{Alg}\big(\{\hat e_{\alpha\beta}^{(j)} : 1\le\alpha,\beta\le m_j\}\big) \oplus \mathbb{C}\hat q,$$

it follows from [12, Corollary 7.2] that

$$\begin{aligned}
\delta_0(\mathbf{X}^{(k)}) &\ge 1 - \sum_{i=1}^{\ell}\tau(\hat p_{0i})^2 - \sum_{j=1}^{s\wedge\ell}\frac{\tau\big(\sum_{\alpha=1}^{m_j}\hat e_{\alpha\alpha}^{(j)}\big)^2}{m_j^2} - \tau(\hat q)^2\\
&\ge 1 - \ell\left(\frac{\tau(p_0)}{\ell}+\varepsilon\right)^2 - \sum_{j=0}^{s\wedge\ell}\frac{(\tau(p_j)+2m_j\varepsilon)^2}{m_j^2} - \left(\tau\left(\mathbf{1}-\sum_{j=0}^{s\wedge\ell}p_j\right)+\varepsilon\right)^2,
\end{aligned} \tag{6.1}$$

since

$$\begin{aligned}
\left\|\sum_{\alpha=1}^{m_j}\hat e_{\alpha\alpha}^{(j)} - p_j\right\|_{2,\tau} &\le \sum_{\alpha=1}^{m_j}\|\hat e_{\alpha1}^{(j)}\hat e_{1\alpha}^{(j)} - e_{\alpha1}^{(j)}e_{1\alpha}^{(j)}\|_{2,\tau}\\
&\le \sum_{\alpha=1}^{m_j}\Big\{\|(\hat e_{\alpha1}^{(j)} - e_{\alpha1}^{(j)})\hat e_{1\alpha}^{(j)}\|_{2,\tau} + \|e_{\alpha1}^{(j)}(\hat e_{1\alpha}^{(j)} - e_{1\alpha}^{(j)})\|_{2,\tau}\Big\}\\
&\le 2m_j\varepsilon.
\end{aligned}$$

Since the value of (6.1) converges to $\delta_0(\mathbf{X}) = 1 - \sum_{j=1}^{s}\tau(p_j)^2/m_j^2$ as $\varepsilon\searrow 0$ and then $\ell\to\infty$, we get the desired assertion. □

Remark 6.6. If $\mathbf{X}^{(k)} \subset W^*(\mathbf{X})$ for all k is further assumed in the above lemma, the consequence becomes $\lim_{k\to\infty}\delta_0(\mathbf{X}^{(k)}) = \delta_0(\mathbf{X})$ due to [12, Corollary 7.2].

Proposition 6.7. (Lower semicontinuity of δ_0 under convergence inside hyperfinite subalgebras) *Let* $\mathbf{X}_i = (X_{i1},\dots,X_{ir(i)})$, $1\le i\le n$, *be hyperfinite random multi-variables. For each*

$1 \le i \le n$ assume that $\mathbf{X}_i^{(k)} = (X_{i1}^{(k)}, \dots, X_{ir(i)}^{(k)}) \subset W^*(\mathbf{X}_i)$ is a sequence of hyperfinite random multi-variables converging to $\mathbf{X}_i$ in mixed moments as $k \to \infty$. Then one has

$$\liminf_{k\to\infty} \delta_0(\mathbf{X}_1^{(k)} \sqcup \cdots \sqcup \mathbf{X}_n^{(k)}) \ge \delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n).$$

Proof. We have

$$\begin{aligned}
\liminf_{k\to\infty} \delta_0(\mathbf{X}_1^{(t)} \sqcup \cdots \sqcup \mathbf{X}_n^{(k)}) &= \liminf_{k\to\infty} \left(\delta_{\mathrm{orb},0}(\mathbf{X}_1^{(k)}, \dots, \mathbf{X}_n^{(k)}) + \sum_{k=1}^n \delta_0(\mathbf{X}_k^{(k)}) \right) \\
&\ge \liminf_{k\to\infty} \left(\delta_{\mathrm{orb},0}(\mathbf{X}_1, \dots, \mathbf{X}_n) + \sum_{i=1}^n \delta_0(\mathbf{X}_i^{(k)}) \right) \\
&\ge \delta_{\mathrm{orb},0}(\mathbf{X}_1, \dots, \mathbf{X}_n) + \sum_{i=1}^n \liminf_{k\to\infty} \delta_0(\mathbf{X}_i^{(k)}) \\
&\ge \delta_{\mathrm{orb},0}(\mathbf{X}_1, \dots, \mathbf{X}_n) + \sum_{i=1}^n \delta_0(\mathbf{X}_i) \\
&= \delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n),
\end{aligned}$$

where the first and the last equalities are due to Theorem 5.8, the second inequality is Proposition 5.3 (5) and the last inequality is Lemma 6.5. □

Remark 6.8. It is worth pointing out that the invariance of δ_0 for hyperfinite von Neumann algebras, a corollary of Jung's result [12], can also be shown from Lemma 6.5 and the argument in [22, Remark 6.13] (together with [25, Theorem 4.3 and Corollary 4.5]). Similarly, Proposition 6.7 and the same pattern of argument show that if $\mathbf{X}_1, \dots, \mathbf{X}_n$ and $\mathbf{X}'_1, \dots, \mathbf{X}'_n$ are hyperfinite random multi-variables and $W^*(\mathbf{X}_i) = W^*(\mathbf{X}'_i)$ for $1 \le i \le n$, then $\delta_0(\mathbf{X}_1 \sqcup \cdots \sqcup \mathbf{X}_n) = \delta_0(\mathbf{X}'_1 \sqcup \cdots \sqcup \mathbf{X}'_n)$ holds. In fact, this is a "part" of [14, Corollary 4.2] and also a consequence of Theorem 5.8 and Proposition 5.3 (4). However, those two results of Jung [12, 14] were crucially used in the proofs of Lemma 6.5 and Theorem 5.8 while it is desirable to prove Theorem 5.8 without the use of the hyperfinite inequality in [14]. Our discussion on lower semicontinuity for δ_0 should be regarded as a kind of "converse argument" of [22, Remark 6.13], thus suggesting that a kind of lower semicontinuity for δ_0 is essentially equivalent to the affirmative solution of "the entropy dimension problem" in [27, §2.6].

6.4. **A few remarks on definitions of δ_0 and $\delta_{0,\mathrm{orb}}$.** In the course of proving Theorem 5.8 we examined several ideas and observed some small facts concerning the free entropy dimension δ_0 and Voiculescu's liberation process, which may be of independent interest. Here we give a brief summary of them.

Let $X_1, \dots, X_n$ be non-commutative self-adjoint random variables in a tracial W^*-probability space, and we choose a freely independent n-tuple of free multiplicative unitary Brownian motions $\mathbf{v}(t) = (v_k(t))_{k=1}^n$, $t \ge 0$. The next proposition can be shown directly by an argument as in the part "≤" of Proposition 5.6 applied to χ rather than χ_{orb} while we will derive it simply from Theorem 5.8.

Proposition 6.9. *With the above assumption,*

$$n + \limsup_{\varepsilon\searrow 0} \frac{\chi(v_1(\varepsilon)X_1v_1(\varepsilon)^*, \dots, v_n(\varepsilon)X_nv_n(\varepsilon)^* : \mathbf{v}(\varepsilon))}{|\log \varepsilon^{1/2}|} \le \delta_0(X_1, \dots, X_n).$$

Furthermore, if $\chi(X_i) > -\infty$ for all $1 \le i \le n$, then equality holds and $\delta_0(X_1, \dots, X_n) = \delta_{0,\mathrm{orb}}(X_1, \dots, X_n) + n$.

Proof. One has nothing to do when $\chi(X_i) = -\infty$ for some i since the χ in the left-hand side always becomes $-\infty$. Thus we may and do assume $\chi(X_i) > -\infty$ for all $1 \le i \le n$. By Remark 5.2 the left-hand side is nothing but $n + \delta_{0,\mathrm{orb}}(X_1, \dots, X_n)$, and hence the assertion follows from Theorem 5.8 thanks to $\delta_0(X_i) = 1$ (see [22, Proposition 6.3] and [23, Corollary 6.7]). □

Hence, no difference occurs even if the semicircular deformation is replaced by the liberation process in the definition of δ_0 when all $\chi(X_i)$ are finite. Furthermore, we can prove the following:

Proposition 6.10. *If $\mathbf{S} = (S_i)_{i=1}^n$ is a free semicircular system freely independent of the other random variables, then*

$$\delta_0(X_1, \dots, X_n) = n + \limsup_{\varepsilon \searrow 0} \frac{\chi(v_1(\varepsilon)(X_1 + \varepsilon^{1/2}S_1)v_1(\varepsilon)^*, \dots, v_n(\varepsilon)(X_n + \varepsilon^{1/2}S_n)v_n(\varepsilon)^* : \mathbf{S} \sqcup \mathbf{v}(\varepsilon))}{|\log \varepsilon^{1/2}|}.$$

Proof. (Sketch) Fix $0 < \delta < \varepsilon \le 1$ arbitrarily and let $R \ge \max_{1\le i\le n} \|X_i\|_\infty + 2$. We write

$$\Gamma(\varepsilon, N, 3m, \delta/2^m) := \Gamma_R((v_i(\varepsilon)(X_i + \varepsilon^{1/2}S_i)v_i(\varepsilon)^*)_{i=1}^n : \mathbf{S} \sqcup \mathbf{v}(\varepsilon); N, 3m, \delta/2^m),$$
$$\Gamma(N, m, \delta) := \Gamma_R((X_i)_{i=1}^n; N, m, \delta).$$

By an essentially same argument as in the part "$\le$" of Proposition 5.6 one can prove that $\Gamma(\varepsilon, N, 3m, \delta/2^m)$ is included in $\mathcal{N}_{L\varepsilon^{1/2}}(\Gamma(N, m, \delta))$ with respect to d_2 (the metric induced from tr_N) with some constant $L > 0$. Then one has

$$\Lambda_N^{\otimes n}(\Gamma(t, N, 3m, \gamma/2^m)) \le K_{\varepsilon^{1/2}}(\Gamma(N, m, \gamma)) \cdot \Lambda_N^{\otimes n}(\mathrm{Ball}((L+1)\varepsilon^{1/2}, d_2).$$

From this one can derive the inequality $\ge$ for the desired equality similarly to the part "$\le$" of Proposition 5.6. On the other hand, we can modify the proof of Proposition 4.6 to show the inequality

$$\begin{aligned}\chi(v_1(\varepsilon)(X_1 + \varepsilon^{1/2}S_1)v_1(\varepsilon)^*, \dots, v_n(\varepsilon)(X_n + \varepsilon^{1/2}S_n)v_n(\varepsilon)^* : \mathbf{S} \sqcup \mathbf{v}(t)) \\ \ge \chi(X_1 + \varepsilon^{1/2}S_1, \dots, X_n + \varepsilon^{1/2}S_n : \mathbf{S}),\end{aligned}$$

which gives the reverse inequality. □

Proposition 6.11. *With the same assumption as Proposition 6.10,*

$$\delta_{0,\mathrm{orb}}(X_1, \dots, X_n) = \limsup_{\varepsilon \searrow 0} \frac{\chi_{\mathrm{orb}}(v_1(\varepsilon)(X_1 + \varepsilon^{1/2}S_1)v_1(\varepsilon)^*, \dots, v_n(\varepsilon)(X_n + \varepsilon^{1/2}S_n)v_n(t)^* : \mathbf{S} \sqcup \mathbf{v}(\varepsilon))}{|\log \varepsilon^{1/2}|}.$$

Proof. (Sketch) The proofs of Lemma 2.4 and Theorem 2.6 can slightly be modified to show the equality

$$\begin{aligned}&\chi(v_1(\varepsilon)(X_1 + \varepsilon^{1/2}S_1)v_1(\varepsilon)^*, \dots, v_n(\varepsilon)(X_n + \varepsilon^{1/2}S_n)v_n(\varepsilon)^* : \mathbf{S} \sqcup \mathbf{v}(\varepsilon)) \\ &\quad = \chi_{\mathrm{orb}}(v_1(\varepsilon)(X_1 + \varepsilon^{1/2}S_1)v_1(\varepsilon)^*, \dots, v_n(\varepsilon)(X_n + \varepsilon^{1/2}S_n)v_n(\varepsilon)^* : \mathbf{S} \sqcup \mathbf{v}(\varepsilon)) \\ &\qquad + \sum_{i=1}^n \chi(X_i + \varepsilon^{1/2}S_i).\end{aligned}$$

Hence the desired formula follows from Theorem 5.8 and Proposition 6.10 thanks to $\delta_0(X_i) = 1 + \lim_{\varepsilon\searrow 0} \chi(X_i + \varepsilon^{1/2}S_i)/|\log \varepsilon^{1/2}|$ by [22, Proposition 6.3]. □

We were interested in the formulas in Propositions 6.10 and 6.11 because those together immediately imply Theorem 5.8 for self-adjoint variables $X_1, \ldots, X_n$. But the direct proof of Proposition 6.11 seems difficult since the orbital theory does not fit well in additive operations like the semicircular deformation.

References

[1] P. Biane, Free brownian motion, free stochastic calculus and random matrices, in *Free Probability Theory*, D.V. Voiculescu (ed.), Fields Inst. Commun. **12**, Amer. Math. Soc., 1997, pp. 1–19.

[2] N. Brown, Finite free entropy and free group factors, *Int. Math. Res. Not.* **2005**, no. 28, 1709–1715.

[3] N.P. Brown, K.J. Dykema and K. Jung, Free entropy dimension in amalgamated free products, with an appendix by W. Lück, math.OA/0609080.

[4] Ph. Biane and D. Voiculescu, A free probability analogue of the Wasserstein metric on the trace-state space, *Geom. Funct. Anal.* **11** (2001), 1125–1138.

[5] A. Connes and E. Størmer, Entropy for automorphisms of II_1 von Neumann algebras, *Acta Math.* **134** (1975), 289–306.

[6] M. Dostál and D. Hadwin, An alternative to free entropy for free group factors, *Acta math. Sinica, English Series* **19** (2003), 419–472.

[7] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*, Mathematical Surveys and Monographs, Vol. 77, Amer. Math. Soc., Providence, 2000.

[8] F. Hiai and D. Petz, A new approach to mutual information, In *Noncommutative Harmonic Analysis with Applications to Probability*, Banach Center Publ. **78**, Polish Acad. Sci., 2007, to appear.

[9] F. Hiai and Y. Ueda, Free transformation cost inequalities for noncommutative multi-variables, *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, **9** (2006), 391–412.

[10] F. Hiai and Y. Ueda, A log-Sobolev type inequality for free entropy of two projections, preprint, 2006, math.OA/0601171.

[11] F. Hiai and Y. Ueda, Notes on microstate free entropy of projections, preprint, 2006, math.OA/0605633.

[12] K. Jung, The free entropy dimension of hyperfinite von Neumann algebras, *Trans. Amer. Math. Soc.* **355** (2003), 5053–5089.

[13] K. Jung, A free entropy dimension lemma, *Pacific J. Math.* **211** (2003), 265–271.

[14] K. Jung, A hyperfinite inequality for free entropy dimension, *Proc. Amer. Math. Soc.* **134** (2006), 2099–2108.

[15] K. Jung, Amenability, tubularity, and embeddings into R^ω, preprint, 2005, math.OA/0506108.

[16] F. Otto and C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* **173** (2000), 361–400.

[17] D. Shlyakhtenko, Free Fisher information with respect to a completely positive map and cost of equivalence relations, *Comm. Math. Phys.* **218** (2001), 133-152.

[18] D. Shlyakhtenko, Remarks on free entropy dimension, in *Operator Algebras: The Abel Symposium 2004*, 249–257, Abel Symp., **1**, Springer, Berlin, 2006.

[19] S. J. Szarek, Nets of Grassmann manifold and orthogonal group, in *Proceedings of Research Workshop on Banach Space Theory*, Univ. Iowa, Iowa City, Iowa, 1982, pp. 169–185.

[20] S. J. Szarek, Metric entropy of homogeneous spaces, in *Quantum Probability*, Banach Center Publ. **43**, Polish Acad. Sci., 1998, pp. 395–410.

[21] D. Voiculescu, Limit laws for random matrices and free products, *Invent. Math.* **104** (1991), 201–220.

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

[23] D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory III: The absence of Cartan subalgebras, *Geom. Funct. Anal.* **6** (1996), 172–199.

[24] 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*, D.V. Voiculescu (ed.), Fields Inst. Commun. **12**, Amer. Math. Soc., 1997, pp. 293–302.

[25] D. Voiculescu, A strengthened asymptotic freeness result for random matrices with applications to free entropy, *Int. Math. Res. Not.* **1998**, 41–63.

[26] 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.

[27] D. Voiculescu, Free entropy, *Bull. London Math. Soc.* **34** (2002), 257–278.

Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai 980-8579, Japan

Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai 980-8579, Japan

Graduate School of Mathematics, Kyushu University, Fukuoka 810-8560, Japan