ANALOGUES OF ENTROPY IN BI-FREE PROBABILITY THEORY: MICROSTATES

IAN CHARLESWORTH AND PAUL SKOUFRANIS

Abstract. In this paper, we extend the notion of microstate free entropy to the bi-free setting. In particular, using the bi-free analogue of random matrices, microstate bi-free entropy is defined. Properties essential to an entropy theory are developed, such as the behaviour of the entropy when transformations on the left variables or on the right variables are performed. In addition, the microstate bi-free entropy is demonstrated to be additive over bi-free collections provided additional regularity assumptions are included and is computed for all bi-free central limit distributions. Moreover, an orbital version of bi-free entropy is examined which provides a tighter upper bound for the subadditivity of microstate bi-free entropy and provides an alternate characterization of bi-freeness in certain settings.

1. Introduction

In a series of revolutionary papers [32–35, 37, 38], Voiculescu developed free probability analogues of the notions of entropy and Fisher's information. In particular [33] introduced a microstate notion of free entropy. In this setting 'microstates' refers to approximating the distribution of self-adjoint operators in a tracial von Neumann algebra using matrix algebras. The notion of microstate free entropy led to many important results pertaining to free group factors, such as the absence of Cartan subalgebras [34], the absence of simple maximal abelian self-adjoint algebras [11], and the free group factors being prime [12]. Alternatively, an infinitesimal version of free entropy based on derivations developed in [37] has also led to many developments.

Recently in [39] Voiculescu extended the notion of free probability to simultaneously study the left and right actions of algebras on reduced free product spaces. This so-called bi-free probability has attracted the attention of many researchers and has had numerous developments (e.g. [2, 6–8, 24–26]). The interest surrounding bi-free probability stems from the possibility of extending the techniques of free probability to solve problems pertaining to pairs of von Neumann algebras, such as a von Neumann algebra and its commutant or the tensor product of two von Neumann algebras.

One important development in bi-free probability theory was a bi-free analogue of the connection between free probability and random matrix theory exhibited in [24–26]. As microstate free entropy was motivated by the connection between free probability and random matrix theory, in this paper we use the bi-free matrix models of [24–26] to develop a notion of microstate bi-free entropy. In particular, such theory may be of interest in relation to the recent work [13] which develops a random matrix approach to the Peterson-Thom conjecture via tensors of random matrices, a topic which bi-free probability theory provides substantial information on. In our sister paper [9] a notion of non-microstate bi-free entropy is developed.

In addition to this introduction, this paper contains nine sections which are organized as follows. In Section 2 we define our microstate version of bi-free entropy (Definition 2.2). This notion of entropy only applies in the tracially bi-partite setting: that is, when the left algebra commutes with the right algebra, and the state becomes tracial when restricted to the left algebra or the right algebra. Although bi-free probability theory extends beyond the tracially bi-partite setting, many natural examples are tracially bi-partite such as pairs consisting of a type II_1 factor whose commutant is a type II_1 factor with the tracial states occurring via the same vector state from the L_2-space of some tracial von Neumann algebra. Section 2 also demonstrates this notion of microstate bi-free entropy satisfies many of the natural properties of an entropy theory.

In Section 3 we analyze how transformations affect microstate bi-free entropy. We find that when a transformation modifies only the left variables or only the right variables, microstate bi-free entropy behaves identically to how microstate free entropy behaves. However, the behaviour of microstate bi-free entropy

Date: February 4, 2021.

2010 *Mathematics Subject Classification.* 46L54, 46L53, 47B80, 94A17.

Key words and phrases. bi-free probability, entropy.

The research of the second author was supported in part by NSERC (Canada) grant RGPIN-2017-05711.

under transformations mixing left and right variables is currently unknown. This is unsurprising as such a mixing destroys the distinction of left and right variables, and so is not easy to view as a natural bi-free operation.

In Section 4 it is demonstrated, under the assumption of the existence of microstates of all orders and a limit condition, that the microstate bi-free entropy of bi-free collections is the sum of the bi-free entropies (Theorem 4.7). Assuming the existence of microstates of all orders is currently a necessity for the analogous result for free entropy, with the general case being at partially addressed in works such as [10, 18].

In Section 5 an orbital version of bi-free entropy is examined in a similar fashion to the orbital free entropy from [29]. In particular, two characterizations of orbital bi-free entropy are given and the base properties are demonstrated. Furthermore, Theorem 5.12 provides a better bound for the difference between the joint microstate bi-free entropy and the sum of the individual microstate bi-free entropies.

In Section 6 Theorem 6.1 is demonstrated, which characterizes when pairs of algebras with finite-dimensional approximants are bi-free in terms of the orbital bi-free entropy. In addition, it is shown in Corollary 6.2 that if collections of left and right operators have finite microstate bi-free entropy and the joint bi-free entropy is the sum of the individual bi-free entropies, then the collections are bi-free.

In Section 7 computations pertaining to microstate bi-free entropy are performed. In particular, the value of the microstate bi-free entropy is computed for all finite bi-free central limit distributions. This computation is non-trivial due to the same complications as in Section 3. It is worthy to note that the microstate bi-free entropy for bi-free central limit distributions has the same form as Gaussian distributions with respect to the Shannon entropy and the free central limit distributions with respect to free entropy. Furthermore, the same value is obtained for non-microstate bi-free entropy in our sister paper [9].

In Section 8 we develop the notion of microstate bi-free entropy dimension and show that for a bi-free central limit distribution pair that this dimension is equal to the dimension of the support of their joint distribution. In Section 9 we discuss generalizing this microstate version of bi-free entropy to non-bi-partite systems and the resulting complications. Finally, in Section 10, several open questions are discussed, most of which might be possible to solve from a deeper understanding of the structure of free and/or bi-free microstates.

Note it is not the intent of this paper to reprove every single fact about microstate free entropy in the bi-free setting, but rather to show that most of the base and some interesting results carry forward.

Throughout the paper, we will assume familiarity with the bi-free setting and adopt its common notation. The relevant set-up and definitions, such as "bi-free independence" and "family of pairs of faces" may be found in [39, Section 2].

2. Definition and Basic Properties

In [31] Voiculescu observed a connection between random matrix theory and free probability. Specifically it was demonstrated that the eigenvalue distribution of certain random matrices asymptotically tended to the free central limit distributions, and random matrices with independent entries tended in law to freely independent operators. However other distributions can be approximated using the eigenvalues of matrices. In an attempt to understand these approximations, Voiculescu introduced the notion of free entropy defined as follows.

Definition 2.1 ([33])**.** Let $(\mathfrak{M}, \tau)$ be a tracial von Neumann algebra and let $X_1, \ldots, X_n \in \mathfrak{M}$ be self-adjoint operators. Let $(\mathcal{M}_d, \tau_d)$ denote the tracial von Neumann algebra consisting of the $d \times d$ complex matrices with the normalized trace τ_d. We will use Tr_d to denote the unnormalized trace on $\mathcal{M}_d$ and $\mathcal{M}_d^{\mathrm{sa}}$ to denote the self-adjoint elements of $\mathcal{M}_d$.

For $M, d \in \mathbb{N}$ and $R, \epsilon > 0$, let $\Gamma_R(X_1, \ldots, X_n; M, d, \epsilon)$ denote the set of all n-tuples $(A_1, \ldots, A_n) \in (\mathcal{M}_d^{\mathrm{sa}})^n$ such that $\|A_j\| \leq R$ for all $1 \leq j \leq n$ and

$$\left|\tau(X_{i_1} \cdots X_{i_p}) - \tau_d(A_{i_1} \cdots A_{i_p})\right| < \epsilon$$

for all $i_1, \ldots, i_p \in \{1, \ldots, n\}$ and $1 \leq p \leq M$. Subsequently, if $\lambda_{d,n}$ denotes the Lebesgue measure on $(\mathcal{M}_d^{\mathrm{sa}})^n$ where $(\mathcal{M}_d^{\mathrm{sa}})^n$ is equipped with the Hilbert-Schmidt norm

$$\|(A_1, \ldots, A_n)\|_{\mathrm{HS}} = \mathrm{Tr}_d(A_1^2 + \cdots + A_n^2),$$

define

$$\begin{aligned}\chi_R(X_1,\ldots,X_n;M,d,\epsilon) &= \log\left(\lambda_{d,n}\left(\Gamma_R(X_1,\ldots,X_n;M,d,\epsilon)\right)\right)\\ \chi_R(X_1,\ldots,X_n;M,\epsilon) &= \limsup_{d\to\infty}\frac{1}{d^2}\chi_R(X_1,\ldots,X_n;M,d,\epsilon)+\frac{1}{2}n\log(d)\\ \chi_R(X_1,\ldots,X_n) &= \inf\{\chi_R(X_1,\ldots,X_n;M,\epsilon)\mid M\in\mathbb{N},\epsilon>0\},\text{ and}\\ \chi(X_1,\ldots,X_n) &= \sup_{R>0}\chi_R(X_1,\ldots,X_n).\end{aligned}$$

The quantity $\chi(X_1,\ldots,X_n)\in[-\infty,\infty)$ is called the *free entropy of* $X_1,\ldots,X_n$. The reason for the constants and various normalizations can be seen in [33] or the computations in Section 7.

As even some bi-free central limit distributions fail to be tracial (see, e.g., [6, Example 11]) we must replace microstates with a version which can approximate non-tracial distributions in order to deal with the bi-free setting. Rather than allow arbitrary non-tracial states on the matrices, though, we seek to progress in a way that recognizes the distinction between left and right variables. This leads us to the idea of microstates consisting of bounded linear maps on $\mathcal{M}_d$ given by left and right matrix multiplication operators; that is, for $A\in\mathcal{M}_d$, we define $L(A)$ and $R(A)$ to be the bounded linear maps on $\mathcal{M}_d$ defined by

$$L(A)B=AB\qquad\text{and}\qquad R(A)B=BA.$$

We then equip the bounded linear maps on $\mathcal{M}_d$ with the state $\tau_d(\cdot I_d)$ which evaluates the linear maps when applied to the identity matrix and then computes the trace of the result.

Of course, these choices force some restrictions upon us. In particular, as left matrix multiplication commutes with right matrix multiplication, we can only find microstates for so-called bi-partite families where all left variables commute with all right variables (in distribution). Furthermore, $\tau_d(\cdot I_d)$ is tracial when restricted to left multiplication operators or right multiplication operators, so we will only be able to produce microstates for distributions having this property. We shall refer to systems satisfying the above as *tracially bi-partite*, and give some indication of how to broaden this setting in Section 9.

Definition 2.2. Let $(\mathcal{A},\varphi)$ be a C*-non-commutative probability space and let $X_1,\ldots,X_n,Y_1,\ldots,Y_m$ be self-adjoint operators in $\mathcal{A}$. For $M,d\in\mathbb{N}$ and $R,\epsilon>0$, let $\Gamma_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m;M,d,\epsilon)$ denote the set of all $(n+m)$-tuples $(A_1,\ldots,A_n,B_1,\ldots,B_m)\in(\mathcal{M}_d^{\mathrm{sa}})^{n+m}$ such that $\|A_i\|,\|B_j\|\le R$ for all $1\le i\le n$ and $1\le j\le m$, such that

$$\left|\varphi(Z_{k_1}\cdots Z_{k_p})-\tau_d(C_{k_1}\cdots C_{k_p}(I_d))\right|<\epsilon$$

for all $1\le p\le M$ and $k_1,\ldots,k_p\in\{1,\ldots,n+m\}$, where

$$Z_k=\begin{cases}X_k & \text{if } k\in\{1\ldots,n\}\\ Y_{k-n} & \text{if } k\in\{n+1,\ldots,n+m\}\end{cases}\qquad\text{and}\qquad C_k=\begin{cases}L\left(A_k\right) & \text{if } k\in\{1\ldots,n\}\\ R\left(B_{k-n}\right) & \text{if } k\in\{n+1,\ldots,n+m\}\end{cases}\in B(\mathcal{M}_d).$$

With $\lambda_{d,p}$ still standing for the Lebesgue measure on $(\mathcal{M}_d^{\mathrm{sa}})^p$ equipped with the Hilbert-Schmidt norm, we successively define

$$\begin{aligned}\chi_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m;M,d,\epsilon) &= \log\left(\lambda_{d,n+m}\left(\Gamma_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m;M,d,\epsilon)\right)\right)\\ \chi_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m;M,\epsilon) &= \limsup_{d\to\infty}\frac{1}{d^2}\chi_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m;M,d,\epsilon)+\frac{1}{2}(n+m)\log(d)\\ \chi_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m) &= \inf\{\chi_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m;M,\epsilon)\mid M\in\mathbb{N},\epsilon>0\},\text{ and}\\ \chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m) &= \sup_{R>0}\chi_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m).\end{aligned}$$

The quantity $\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m)$ will be called the *microstate bi-free entropy of* $X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m$. We will see in Proposition 2.6 that $\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m)\in[-\infty,\infty)$.

Remark 2.3. By analyzing the joint distribution of $L(A_1),\ldots,L(A_n),R(B_1),\ldots,R(B_n)$ and the definition of $\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m)$, we see that $\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m)=-\infty$ unless we are in the tracially bi-partite setting. We will make this the standing assumption until Section 9 of the paper. This is a setting which includes many canonical examples and thus is of great interest. Note we will not assume that φ is tracial on $\mathcal{A}$ nor faithful on $\mathcal{A}$ as these properties need not occur in most bi-free systems (see [2] and [23] respectively).

Using the fact that the system is bi-partite, the definition of $\Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d, \epsilon)$ may be simplified slightly, as it is enough to check that only certain moments are well-approximated: indeed, for $M, d \in \mathbb{N}$ and $R, \epsilon > 0$ notice $\Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d, \epsilon)$ is the set of all $(n+m)$-tuples $(A_1, \ldots, A_n, B_1, \ldots, B_m) \in (\mathcal{M}_d^{\mathrm{sa}})^{n+m}$ such that $\|A_i\|, \|B_j\| \leq R$ for all $1 \leq i \leq n$ and $1 \leq j \leq m$, and

$$\left|\varphi(X_{i_1} \cdots X_{i_p} Y_{j_1} \cdots Y_{j_q}) - \tau_d(A_{i_1} \cdots A_{i_p} B_{j_q} \cdots B_{j_1})\right| < \epsilon$$

for all $i_1, \ldots, i_p \in \{1, \ldots, n\}$ and $j_1, \ldots, j_q \in \{1, \ldots, m\}$ with $p + q \leq M$.

Remark 2.4. It is elementary to see based on the definition of microstate bi-free entropy that if $m = 0$ then

$$\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) = \chi(X_1, \ldots, X_n),$$

whence the above notion of bi-free entropy is an extension of microstate free entropy. Further, it can be readily verified that

$$(A_1, \ldots, A_n, B_1, \ldots, B_m) \in \Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d, \epsilon)$$

if and only if

$$(B_1^t, \ldots, B_m^t, A_1^t, \ldots, A_n^t) \in \Gamma_R(Y_1, \ldots, Y_m \sqcup X_1, \ldots, X_n; M, d, \epsilon).$$

It follows that $\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) = \chi(Y_1, \ldots, Y_m \sqcup X_1, \ldots, X_n)$ as transpose preserves Lebesgue measure, and in particular when $n = 0$ we have

$$\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) = \chi(Y_1, \ldots, Y_m).$$

Proposition 2.5. *If $0 \leq p \leq n$ and $0 \leq q \leq m$ then*

$$\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) \leq \chi(X_1, \ldots, X_p \sqcup Y_1, \ldots Y_q) + \chi(X_{p+1}, \ldots, X_n \sqcup Y_{q+1}, \ldots Y_m).$$

In particular,

$$\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) \leq \chi(X_1, \ldots, X_n) + \chi(Y_1, \ldots, Y_m).$$

Proof. First note that the inequality will be demonstrated provided we can show that

$$\begin{aligned} &\chi_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d, \epsilon) \\ &\quad \leq \chi_R(X_1, \ldots, X_p \sqcup Y_1, \ldots Y_q; M, d, \epsilon) + \chi_R(X_{p+1}, \ldots, X_n \sqcup Y_{q+1}, \ldots Y_m; M, d, \epsilon) \end{aligned}$$

for all M, d, and ϵ. Since by definitions we have that

$$\begin{aligned} &\Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d, \epsilon) \\ &\quad \subseteq \Gamma_R(X_1, \ldots, X_p \sqcup Y_1, \ldots Y_q; M, d, \epsilon) \times_{\ell r} \Gamma_R(X_{p+1}, \ldots, X_n \sqcup Y_{q+1}, \ldots Y_m; M, d, \epsilon) \end{aligned}$$

where

$$(A_1, \ldots, A_p, B_1, \ldots, B_q) \times_{\ell r} (A_{p+1}, \ldots, A_n, B_{q+1}, \ldots, B_m) = (A_1, \ldots, A_n, B_1, \ldots, B_m),$$

clearly the above inequalities hold. □

These inequalities allow us to import upper bounds on entropy from the free case. In particular, we learn that the bi-free entropy never takes the value $+\infty$.

Proposition 2.6. *Let $C^2 = \varphi(X_1^2 + \cdots + X_n^2 + Y_1^2 + \cdots + Y_m^2)$. Then*

$$\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) \leq \frac{n+m}{2} \log\left(\frac{2\pi e}{n+m} C^2\right).$$

Proof. We recall that the analogous free statement was shown in [33, Proposition 2.2]. Let

$$C_X^2 = \varphi(X_1^2 + \cdots + X_n^2) \qquad \text{and} \qquad C_Y^2 = \varphi(Y_1^2 + \cdots + Y_m^2).$$

Using the above, Proposition 2.5, and the concavity of the logarithm, we obtain that

$$\begin{aligned}\chi(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m) &\le \chi(X_1,\ldots,X_n)+\chi(Y_1,\ldots,Y_m)\\ &\le \frac{1}{2}n\log\left(\frac{2\pi e}{n}C_X^2\right)+\frac{1}{2}m\log\left(\frac{2\pi e}{m}C_Y^2\right)\\ &= \frac{n+m}{2}\left(\frac{n}{n+m}\log\left(\frac{2\pi e}{n}C_X^2\right)+\frac{m}{n+m}\log\left(\frac{2\pi e}{m}C_Y^2\right)\right)\\ &\le \frac{n+m}{2}\log\left(\frac{n}{n+m}\frac{2\pi e}{n}C_X^2+\frac{m}{n+m}\frac{2\pi e}{m}C_Y^2\right)\\ &= \frac{n+m}{2}\log\left(\frac{2\pi e}{n+m}C^2\right). \qquad \square\end{aligned}$$

There is a more interesting inequality relating the microstate bi-free entropy to the microstate free entropy. In particular, the microstate bi-free entropy is bounded below by the microstate free entropy obtained by changing all of the right variables to left variables.

Theorem 2.7. *Let $\left(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m\right)$ be tracially bi-partite, self-adjoint operators in a C^*-non-commutative probability space $(\mathcal{A},\varphi)$. Suppose there exists another C^*-non-commutative probability space $(\mathcal{A}_0,\tau_0)$ and self-adjoint operators $X_1',\ldots,X_n',Y_1',\ldots,Y_m' \in \mathcal{A}_0$ such that τ_0 is tracial on $\mathcal{A}_0$ and*

$$\varphi(X_{i_1}\cdots X_{i_p}Y_{j_1}\cdots Y_{j_q}) = \tau_0(X_{i_1}'\cdots X_{i_p}'Y_{j_q}'\cdots Y_{j_1}')$$

for all $p,q\in\mathbb{N}\cup\{0\}$, $i_1,\ldots,i_p\in\{1,\ldots,n\}$, and $j_1,\ldots,j_q\in\{1,\ldots,m\}$. Then

$$\chi(X_1',\ldots,X_n',Y_1',\ldots,Y_m') \le \chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m).$$

Proof. Using the characterization from the end of Remark 2.3, we see that

$$\Gamma_R(X_1',\ldots,X_n',Y_1',\ldots,Y_m';M,d,\epsilon)\subseteq\Gamma_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m;M,d,\epsilon),$$

and hence

$$\chi(X_1',\ldots,X_n',Y_1',\ldots,Y_m') \le \chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m). \qquad \square$$

This inequality, in essence, arises because the set of bi-free microstates is defined with fewer conditions than the set of free microstates. In addition, as we need only specify certain moments for the "one-sided" family for a given pair of faces and as many of the moments can be chosen somewhat arbitrarily, Theorem 2.7 provides many possible lower bounds.

Example 2.8. Let us give an example application of Theorem 2.7. Recall that the full Fock space associated to a Hilbert space $\mathcal{H}$ is

$$\mathcal{F}(\mathcal{H}) := \bigoplus_{n\ge 0}\mathcal{H}^{\otimes n}.$$

For $e\in\mathcal{H}$, the left and right creation operators $l(e)$ and $r(e)$ act by "tensoring e on the left" and "on the right", respectively; their adjoints are the annihilation operators. The operators $l(e)+l(e)^*$ and $r(e)+r(e)^*$ have semicircular distribution with respect to the state induced by the vacuum vector $\Omega\in\mathcal{H}^{\otimes 0}\cong\mathbb{C}$. More details may be found, for example, in [21, Lecture 7] or (with the bi-free context in mind) in [39, Section 6.3].

If $e_1,e_2\in\mathcal{H}$, $c=\langle e_1,e_2\rangle\in\mathbb{R}$, and if $S_i=l(e_i)+l^*(e_i)$, $D_2=r(e_2)+r^*(e_2)$, then Theorem 2.7 implies that

$$\chi(S_1\sqcup D_2)\ge\chi(S_1,S_2).$$

Notice that if $c\in(-1,1)$ then

$$e_3 := \frac{1}{\sqrt{1-c^2}}\left(e_2-ce_1\right)$$

is a unit vector orthogonal to e_1, and so if $S_3=l(e_3)+l^*(e_3)$, then S_1 and S_3 are freely independent centred semicircular variables of variance one while

$$\begin{bmatrix}1 & 0\\ c & \sqrt{1-c^2}\end{bmatrix}\begin{bmatrix}S_1\\ S_3\end{bmatrix}=\begin{bmatrix}S_1\\ S_2\end{bmatrix}.$$

Therefore, by [32, Proposition 3.5 and Proposition 5.4] (or the analogous Proposition 3.1 in this paper), we obtain that

$$\begin{aligned}\chi(S_1 \sqcup D_2) \geq \chi(S_1, S_2) &= \chi(S_1, S_3) + \log\left|\det\begin{bmatrix}1 & 0\\ c & \sqrt{1-c^2}\end{bmatrix}\right| \\ &= \chi(S_1) + \chi(S_3) + \log(\sqrt{1-c^2}) \\ &= 2\chi(S_1) + \frac{1}{2}\log\left(1-c^2\right).\end{aligned}$$

It will be shown in Theorem 7.3 that this inequality is actually an equality.

Like with free entropy, the upper bound on the norm of microstates R can be controlled.

Proposition 2.9. *Let*

$$\rho = \max\left(\{\|X_i\| \mid 1 \leq i \leq n\} \cup \{\|Y_j\| \mid 1 \leq j \leq m\}\right).$$

If $R_2 > R_1 > \rho$, then

$$\chi_{R_2}(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) = \chi_{R_1}(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m).$$

In particular, for all $R > \rho$,

$$\chi_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) = \chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m).$$

Proof. As χ_R is an increasing function of R, it suffices to prove the first equality. The proof of said equality will be similar to that of [33, Proposition 2.4].

Fix $R_2 > R_1 > R_0 > \rho$ and define $g : [-R_2, R_2] \to [-R_1, R_1]$ to be the function which is linear on $[-R_2, -R_0]$, $[-R_0, R_0]$, and $[R_0, R_2]$, and such that $g(-R_2) = -R_1$, $g(-R_0) = -R_0$, $g(R_0) = R_0$, and $g(R_2) = R_1$. Furthermore, for $A_1, \ldots, A_n, B_1, \ldots, B_m \in \mathcal{M}_d^{\mathrm{sa}}$ with $\|A_i\| \leq R_2$ and $\|B_j\| \leq R_2$, let

$$G(A_1, \ldots, A_n, B_1, \ldots, B_m) = (g(A_1), \ldots, g(A_n), g(B_1), \ldots, g(B_m)).$$

Given $M \in \mathbb{N}$ and $\epsilon > 0$, it is not difficult to see that there exists an $M_1 \geq M$ and a $0 < \epsilon_1 < \epsilon$ such that

$$G(\Gamma_{R_2}(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M_1, d, \epsilon_1)) \subseteq \Gamma_{R_1}(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d, \epsilon)$$

for all $d \in \mathbb{N}$. Indeed for any

$$(A_1, \ldots, A_n, B_1, \ldots, B_m) \in \Gamma_{R_2}(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M_1, d, \epsilon_1)$$

we obtain that

$$|\tau_d(A_i^p)|, |\tau_d(B_j^p)| \leq \rho^p + \epsilon_1$$

for all $1 \leq i \leq n$, $1 \leq j \leq m$, and $1 \leq p \leq M_1$. Thus given $\delta > 0$, choosing M_1 large and ϵ_1 small enough yields

$$\tau_d(P_{[-R_2,-R_0]\cup[R_0,R_2]}(A_i)), \tau_d(P_{[-R_2,-R_0]\cup[R_0,R_2]}(B_j)) < \delta$$

where $P_{[-R_2,-R_0]\cup[R_0,R_2]}$ is denoting the spectral projection onto $[-R_2, -R_0] \cup [R_0, R_2]$, and thus can be selected even smaller still to make

$$\|g(A_i) - A_i\|_1, \|g(B_j) - B_j\|_1 < \delta$$

for all $1 \leq i \leq n$ and $1 \leq j \leq m$ independent of d. As M and R_2 are fixed, by selecting δ sufficiently small we obtain that the trace of any word of length at most M in $g(A_1), \ldots, g(A_n), g(B_1), \ldots, g(B_m)$ is within a function of δ, M, and R_2 which tends to 0 as δ tends to 0 to the trace of the corresponding word in $A_1, \ldots, A_n, B_1, \ldots, B_m$. Thus the claim follows.

To complete the proof, it will suffice to obtain a specific lower bound on the Jacobian of G on

$$\Gamma_{R_2}(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M_1, d, \epsilon_1).$$

Let $U(d)$ denote the set of unitary elements of $\mathcal{M}_d$ and consider the change of coordinates from $\mathcal{M}_d^{\mathrm{sa}}$ to $(U(d)/\mathbb{T}) \times \{(c_1, \ldots, c_d) \in \mathbb{R}^d \mid c_1 < \cdots < c_d\}$ (where $\mathbb{T}$ is the torus of diagonal unitaries) defined by $(U, D) \mapsto U^*DU$ where $D = \mathrm{diag}(c_1, \ldots, c_d)$. This change of coordinates places the Lebesgue measure in the form

$$K\left(\prod_{1\leq i<j\leq d}(c_i - c_j)\right) d\gamma_{d,0} d\lambda_d$$

where K is a normalizing constant and $\gamma_{d,0}$ is the Haar measure on $U(d)/\mathbb{T}$. The absolute value of the Jacobian of the map $C \mapsto g(C)$ is easily seen to be

$$g'(c_1)\cdots g'(c_d) \prod_{1\leq i<j\leq d} \frac{g(c_i)-g(c_j)}{c_i-c_j}$$

when C has eigenvalues $c_1,\ldots,c_d$ and $c_k \neq \pm R_0$ for all k.

Let $\delta > 0$ be arbitrary. If M_1 is large enough and ϵ_1 is small enough, we obtain that

$$\tau_d(P_{[-R_2,-R_0]\cup[R_0,R_2]}(C)) < \delta$$

and thus we obtain

$$\left(\frac{R_1-R_0}{R_2-R_0}\right)^{d+d^2-(d(1-\delta))^2}$$

as a lower bound for the Jacobian of g on a coordinate projection of $\Gamma_{R_2}(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m; M_1, d, \epsilon_1)$. In particular a lower bound for the Jacobian of G can be obtained by taking the above lower bound for the Jacobian of g raised to the $(n+m)^{\text{th}}$ power and thus

$$\begin{aligned}\chi_{R_1}(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m; m, d, \epsilon) &\geq \chi_{R_2}(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m; M_1, d, \epsilon_1)\\ &\quad + (n+m)(d+d^2(2\delta-\delta^2))\log\left(\frac{R_1-R_0}{R_2-R_0}\right).\end{aligned}$$

Hence it follows that

$$\begin{aligned}\chi_{R_1}(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m; m, \epsilon) &\geq \chi_{R_2}(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m; M_1, \epsilon_1)\\ &\quad + (n+m)(2\delta-\delta^2)\log\left(\frac{R_1-R_0}{R_2-R_0}\right).\end{aligned}$$

Therefore, as $\delta > 0$ was arbitrary, the result follows. ☐

Remark 2.10. The proof of Proposition 2.9 can be extended further. Indeed let $R_1,\ldots,R_n,R'_1,\ldots,R'_m > 0$ and

$$\Gamma_{R_1,\ldots,R_n,R'_1,\ldots,R'_m}(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m; M, d, \epsilon)$$

be defined like $\Gamma_R(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m; M, d, \epsilon)$ where instead of $\|A_i\|, \|B_j\| \leq R$ for all i,j, we only require $\|A_i\| \leq R_i$ and $\|B_j\| \leq R'_j$ for all i,j. If we extend the notion of $\chi_R(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m)$ to $\chi_{R_1,\ldots,R_n,R'_1,\ldots,R'_m}(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m)$, then the same proof as Proposition 2.9 can be used to show that if $R_i > \|X_i\|$ and $R'_j > \|Y_j\|$ for all i,j, then

$$\chi_{R_1,\ldots,R_n,R'_1,\ldots,R'_m}(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m) = \chi(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m).$$

In fact, we note that [1] refined the techniques of [33, Proposition 2.4] to demonstrate that if one lets $R = \infty$ in the start of Definition 2.1, then the same value of the microstate free entropy is obtained. By repeating their results verbatim with the obvious modifications in our context identical to those used above in Proposition 2.9, we note that setting $R = \infty$ from the start of Definition 2.2 yields the same quantity for the microstate bi-free entropy.

On the other hand, insisting on using microstates of bounded norm allows us the following proposition.

Proposition 2.11. *Let $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$ and $\left(\left\{X_i^{(k)}\right\}_{i=1}^n, \left\{Y_j^{(k)}\right\}_{j=1}^m\right)$ for $k \in \mathbb{N}$ be tracially bi-partite tuples in a C^*-non-commutative probability space $(\mathcal{A},\varphi)$. Suppose that $\left(\left\{X_i^{(k)}\right\}_{i=1}^n, \left\{Y_j^{(k)}\right\}_{j=1}^m\right)$ converges in distribution to $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$; that is*

$$\lim_{k\to\infty} \varphi\left(X_{i_1}^{(k)}\cdots X_{i_p}^{(k)} Y_{j_1}^{(k)}\cdots Y_{j_q}^{(k)}\right) = \varphi(X_{i_1}\cdots X_{i_p} Y_{j_1}\cdots Y_{j_q})$$

for all $i_1,\ldots i_p \in \{1,\ldots,n\}$, $j_1,\ldots,j_q \in \{1,\ldots,m\}$, and $p,q \in \mathbb{N}$. Then

$$\limsup_{k\to\infty} \chi_R\left(X_1^{(k)},\ldots,X_n^{(k)} \sqcup Y_1^{(k)},\ldots,Y_m^{(k)}\right) \leq \chi_R(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m).$$

Moreover, if $\sup_{k\in\mathbb{N}}\left\|X_i^{(k)}\right\| < \infty$ *for all* $1 \leq i \leq n$ *and* $\sup_{k\in\mathbb{N}}\left\|Y_k^{(k)}\right\| < \infty$ *for all* $1 \leq j \leq m$*, then*

$$\limsup_{k\to\infty} \chi(X_1^{(k)}, \dots, X_n^{(k)} \sqcup Y_1^{(k)}, \dots, Y_m^{(k)}) \leq \chi(X_1, \dots, X_n \sqcup Y_1, \dots, Y_m).$$

Proof. Our convergence assumption tells us that all moments converge to the correct values, and so for any $M \in \mathbb{N}$ and $\epsilon > 0$ we have for large enough k that

$$\Gamma_R\left(X_1^{(k)}, \dots, X_n^{(k)} \sqcup Y_1^{(k)}, \dots, Y_m^{(k)}; M, d, \epsilon\right) \subseteq \Gamma_R(X_1, \dots, X_n \sqcup Y_1, \dots, Y_m; M, d, 2\epsilon),$$

since the sets involved see only finitely many moments. Hence for all sufficiently large k, we have

$$\chi_R\left(X_1^{(k)}, \dots, X_n^{(k)} \sqcup Y_1^{(k)}, \dots, Y_m^{(k)}; M, d, \epsilon\right) \leq \chi_R(X_1, \dots, X_n \sqcup Y_1, \dots, Y_m; M, d, 2\epsilon)$$

and passing through the appropriate limits and rescaling in d, then M, and then ϵ yields

$$\limsup_{k\to\infty} \chi_R\left(X_1^{(k)}, \dots, X_n^{(k)} \sqcup Y_1^{(k)}, \dots, Y_m^{(k)}\right) \leq \chi_R(X_1, \dots, X_n \sqcup Y_1, \dots, Y_m)$$

which is the first claimed inequality. The second inequality follows by applying Proposition 2.9. □

3. Transformations

One important property of the microstate free entropy is the ability to apply a non-commutative functional calculus to the self-adjoint operators and control the value of the free entropy. In this section, we will develop an analogue of this result for our microstate bi-free entropy. However, due to the distinction between the left and right operators, we will need to focus on transformations that modify only left variables or modify only right variables (although compositions of such transforms are allowed).

To understand the difficulty in mixing left and right variables, consider the $n = m = 1$ case. If

$$(A, B) \in \Gamma_R(X \sqcup Y; M, d, \epsilon)$$

and we wanted to consider the new pair $(X, Y + cX)$ for c sufficiently small, it is incredibly unclear whether

$$(A, B + cA) \in \Gamma_R(X \sqcup Y + cX; M', d, \epsilon')$$

as the assumptions on (A, B) yield only information about $\tau_d(A^pB^q)$ for $1 \leq p + q \leq M$ whereas we require knowledge about $\tau_d(A^p(B + cA)^q)$. The latter involves terms of the form $\tau_d(A^{i_1}B^{i_2}A^{i_3}\cdots B^{i_j})$ and direct information about these moments appears difficult to extract from knowledge of only $\tau_d(A^pB^q)$.

In order to develop our results, we recall some information from [33]. However, as the proofs are near identical, we refer the reader to [33] on most occasions.

Let $x_1, \dots, x_n$ be non-commuting indeterminates and let

$$F(x_1, \dots, x_n) = \sum_{k=1}^{\infty} \sum_{1\leq i_1,\dots,i_k\leq n} c_{i_1,\dots,i_k} x_{i_1} \cdots x_{i_k}$$

be a non-commuting power series with complex coefficients. If $R_i \geq 0$ for all $1 \leq i \leq n$, it is said that $(R_1, \dots, R_n)$ is a *multiradius of convergence of* F if

$$M(F; R_1, \dots, R_n) := \sum_{k=1}^{\infty} \sum_{1\leq i_1,\dots,i_k\leq n} |c_{i_1,\dots,i_k}| R_{i_1} \cdots R_{i_k} < \infty.$$

If $X_1, \dots, X_n$ are elements in a finite factor $(\mathfrak{M}, \tau)$ and $(\|X_1\|, \dots, \|X_n\|)$ is a multiradius of convergence of F, then $F(X_1, \dots, X_n)$ is well-defined with

$$\|F(X_1, \dots, X_n)\| \leq M(F; \|X_1\|, \dots, \|X_n\|).$$

If $(R_1, \dots, R_n)$ is a multiradius of convergence of F, then the map taking $(X_1, \dots, X_n)$ to $F(X_1, \dots, X_n)$ is an analytic function on

$$\prod_{1\leq i\leq n} \{X_i \in \mathfrak{M} \mid \|X_i\| \leq R_i\}$$

with values in $\mathfrak{M}$. If this map is denoted F, then F is differentiable with derivative denoted by DF, and the positive Jacobian of F at $(X_1, \dots, X_n)$ can be defined by

$$|\mathcal{J}|(F)(X_1, \dots, X_n) = |\det|(DF(X_1, \dots, X_n)),$$

where $|\det|$ denotes the Fuglede-Kadison determinant. Note that $DF(X_1,\ldots,X_n)$ lies in the algebra denoted in [33] by $LR(\mathfrak{M})$, which is the image in $B(\mathfrak{M})$ of the projective tensor product $\mathfrak{M}\otimes_\pi\mathfrak{M}^{\mathrm{op}}$ under the contraction $a\otimes b\mapsto L_aR_b$ (where L_a denotes left multiplication on $\mathfrak{M}$ by a and R_b denotes right multiplication on $\mathfrak{M}$ by b).

Finally, as our focus is on self-adjoint operators, we will focus on F where $F^*=F$; that is $\overline{c_{i_1,\ldots,i_k}}=c_{i_k,\ldots,i_1}$ for all k and $1\le i_1,\ldots,i_k\le n$.

Proposition 3.1. *Let $(\mathcal{A},\varphi)$ be a C^*-non-commutative probability space and let*

$$\left(\{X_i\}_{i=1}^n,\{Y_j\}_{j=1}^m\right)$$

be a tracially bi-partite collection of self-adjoint operators such that $(\mathrm{alg}(X_1,\ldots,X_n),\varphi)$ sits inside a finite factor. Let $F_1,\ldots,F_n,G_1,\ldots,G_n$ be non-commutative power series with complex coefficients such that $F_i^*=F_i$, $G_i^*=G_i$, $(\|X_1\|+\epsilon,\ldots,\|X_n\|+\epsilon)$ is a multiradius of convergence for the F_i's for some $\epsilon>0$, and

$$(M(F_1;\|X_1\|+\epsilon,\ldots,\|X_n\|+\epsilon),\ldots,M(F_n;\|X_1\|+\epsilon,\ldots,\|X_n\|+\epsilon))$$

is a multiradius of convergence for the G_j's. Assume further that

$$G_i(F_1(x_1,\ldots,x_n),\ldots,F_n(x_1,\ldots,x_n))=x_i$$

for all $1\le i\le n$. Then

$$\begin{aligned}&\chi(F_1(X_1,\ldots,X_n),\ldots,F_n(X_1,\ldots,X_n)\sqcup Y_1,\ldots,Y_n)\\&\quad\ge\log\left(|\mathcal{J}|((F_1,\ldots,F_n))(X_1,\ldots,X_n)\right)+\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m).\end{aligned}$$

Moreover, if $N_k=\|F_k(X_1,\ldots,X_n)\|$, then

$$(M(G_1;N_1+\epsilon,\ldots,N_n+\epsilon),\ldots,M(G_n;N_1+\epsilon,\ldots,N_n+\epsilon))$$

is a multiradius of convergence for the F_i's, then

$$\begin{aligned}&\chi(F_1(X_1,\ldots,X_n),\ldots,F_n(X_1,\ldots,X_n)\sqcup Y_1,\ldots,Y_n)\\&\quad=\log\left(|\mathcal{J}|((F_1,\ldots,F_n))(X_1,\ldots,X_n)\right)+\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m).\end{aligned}$$

An analogous result holds for such functions applied to the Y's instead of the X's.

Proof. First we invoke Remark 2.10. Let $\|X_i\|<R_i<\|X_i\|+\epsilon$, let $\|Y_j\|<R'_j$, and

$$M(F_i;R_1,\ldots,R_n)<\rho_i\le M(F_i;\|X_1\|+\epsilon,\ldots,\|X_n\|+\epsilon).$$

Given $M\in\mathbb{N}$, and $\epsilon>0$, there exist an $M_1\ge M$ and an $0<\epsilon_1<\epsilon$ such that the map

$$(A_1,\ldots,A_n,B_1,\ldots,B_m)\mapsto(F_1(A_1,\ldots,A_n),\ldots,F_n(A_1,\ldots,A_n),B_1,\ldots,B_m)$$

maps $\Gamma_{R_1,\ldots,R_n,R'_1,\ldots,R'_m}(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m;M_1,d,\epsilon_1)$ into

$$\Gamma_{\rho_1,\ldots,\rho_n,R'_1,\ldots,R'_m}(F_1(X_1,\ldots,X_n),\ldots,F_n(X_1,\ldots,X_n)\sqcup Y_1,\ldots Y_m;M,d,\epsilon).$$

The remainder of the proof is identical to the proof of [32, Proposition 3.5] as it simply computes how the transformation (ours being a direct sum of the one used in [32, Proposition 3.5] and the identity) modifies the microstates and thus the entropy. □

Corollary 3.2.

(1) *If $a_1,\ldots,a_n,b_1,\ldots,b_m\in\mathbb{R}$, then*

$$\chi(X_1+a_1I,\ldots,X_n+a_nI\sqcup Y_1+b_1I,\ldots,Y_m+b_mI)=\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m).$$

(2) *If $A=[a_{i,j}]\in\mathcal{M}_n$ and $B=[b_{i,j}]\in\mathcal{M}_m$ are invertible, then*

$$\chi\left(\sum_{k=1}^n a_{1,k}X_k,\ldots,\sum_{k=1}^n a_{n,k}X_k\sqcup\sum_{k=1}^m b_{1,k}Y_k,\ldots,\sum_{k=1}^m b_{m,k}Y_k\right)=\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m)+\log(|\det(A\oplus B)|).$$

(3) *If $X_1,\ldots,X_n$ are linearly dependent or $Y_1,\ldots,Y_m$ are linearly dependent, then*

$$\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m)=-\infty.$$

Proof. Parts (1) and (2) follow from Proposition 3.1. In the case of part (3), if $X_1, \ldots, X_n$ are linearly dependent then there is an $A = [a_{i,j}] \in \mathcal{M}_n$ such that $0 < |\det(A)| < 1$ and

$$\left(\sum_{k=1}^{n} a_{1,k}X_k, \ldots, \sum_{k=1}^{n} a_{n,k}X_k\right) = (X_1, \ldots, X_n).$$

Applying part (2) along with the fact that

$$\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) < \infty$$

by Proposition 2.6 yields the result. □

4. Additivity of Microstate Bi-Free Entropy

One important result for free entropy is additivity; that is, if $\{X_1, \ldots, X_p\}$ and $\{X_{p+1}, \ldots, X_n\}$ are free then

$$\chi(X_1, \ldots, X_n) = \chi(X_1, \ldots, X_p) + \chi(X_{p+1}, \ldots, X_n)$$

under certain regularity assumptions. We desire to prove a bi-free analogue of this result. Before we move to those results, we desire to analyze some limits with regards to the following concept.

Definition 4.1. A tracially bi-partite system $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$ in a C*-non-commutative probability space $(\mathcal{A}, \varphi)$ is said to have *finite-dimensional approximants* if for every $M \in \mathbb{N}$, $\epsilon > 0$, and

$$R > \max\left\{\max_{1\leq i\leq n} \|X_i\|, \max_{1\leq j\leq m} \|Y_j\|\right\},$$

there exists an $D \in \mathbb{N}$ such that $\Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d, \epsilon) \neq \emptyset$ for all $d \geq D$.

A single family of such variables $\{X_i\}_{i=1}^n$ is said to have *finite-dimensional approximants* if $(\{X_i\}_{i=1}^n, \emptyset)$ does or, equivalently by Remark 2.4, if $(\emptyset, \{X_i\}_{i=1}^n)$ does.

Remark 4.2. By repeating the same ideas as in Theorem 2.7, the existence of microstates for tracially bi-partite systems can be often deduced from knowledge of free entropy. Indeed suppose $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$ is a tracially bi-partite system in a C*-non-commutative probability space $(\mathcal{A}, \varphi)$ and that there exists another C*-non-commutative probability space $(\mathcal{A}_0, \tau_0)$ and self-adjoint operators $X_1', \ldots, X_n', Y_1', \ldots, Y_m' \in \mathcal{A}_0$ such that τ_0 is tracial on $\mathcal{A}_0$ and

$$\varphi(X_{i_1} \cdots X_{i_p} Y_{j_1} \cdots Y_{j_q}) = \tau_0(X_{i_1}' \cdots X_{i_p}' Y_{j_q}' \cdots Y_{j_1}')$$

for all $p, q \in \mathbb{N} \cup \{0\}$ and $i_1, \ldots, i_p \in \{1, \ldots, n\}$ and $j_1, \ldots, j_q \in \{1, \ldots, m\}$. As in Theorem 2.7,

$$\Gamma_R(X_1', \ldots, X_n', Y_1', \ldots, Y_m'; M, d, \epsilon) \subseteq \Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d, \epsilon).$$

Therefore if $(\{X_i'\}_{i=1}^n, \{Y_j'\}_{j=1}^m)$ have finite-dimensional approximants, then so do $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$ by Proposition 2.9. In particular, if $\chi(X_1', \ldots, X_n', Y_1', \ldots, Y_m') > -\infty$, then $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ has finite-dimensional approximants by [36, Remark 3.2].

Furthermore, if

$$\Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d_0, \epsilon) \neq \emptyset$$

for some d_0, then there exists a D such that

$$\Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d, 2\epsilon) \neq \emptyset$$

for all $d \geq D$. Indeed this follows by taking D to be a sufficiently large multiple of d_0 so that $\frac{d_0}{D}$ is sufficiently small thereby adding at most ϵ to the state estimates. Hence, as in [36, Remark 3.2], it can easily be seen that if $\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots Y_m) > -\infty$, then $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$ has finite-dimensional approximants.

In order to develop an additive result for microstate bi-free entropy, we will use the following notion.

Definition 4.3. Let $(\mathcal{A}, \varphi)$ be a C*-non-commutative probability space, let $\{\mathcal{C}_k\}_{k\in K}$ be a collection of finite subsets of $\mathcal{A}$, let $\mathcal{A}_k = \mathrm{alg}(\mathcal{C}_k)$, and let $\psi = *_{k\in K}\varphi|_{\mathcal{A}_k}$ be the unique state on $*_{k\in K}\mathcal{A}_k$ extending each $\varphi|_{\mathcal{A}_k}$ such that the $\mathcal{A}_k$ are free. Given $M \in \mathbb{N}$ and $\epsilon > 0$, it is said that $\{\mathcal{C}_k\}_{k\in K}$ are *(M, ϵ)-free in $(\mathcal{A}, \varphi)$* provided

$$|\psi(Z_1 \cdots Z_p) - \varphi(Z_1 \cdots Z_p)| < \epsilon$$

for all $Z_1, \ldots, Z_p \in \bigcup_{k\in K} \mathcal{C}_k$ and $1 \leq p \leq M$.

Given $d \in \mathbb{N}$, let $\mathcal{U}(d)$ denote the unitary group of $\mathcal{M}_d$ and let γ_d denote the normalized Haar measure on $\mathcal{U}(d)$. We recall the following result.

Lemma 4.4 ([36, Corollary 2.13])**.** *Fix $R, \epsilon, \theta > 0$ and $M \in \mathbb{N}$. Then there exists an $N \in \mathbb{N}$ such that for all $d \geq N$, $1 \leq p \leq M$, and sets $\mathcal{C}_1, \ldots, \mathcal{C}_p \subseteq \mathcal{M}_d$ of matrices bounded in norm by R, each containing no more than M elements, we have*

$$\mu_d^{\otimes p}\left(\left\{(U_1, \ldots, U_p) \in \mathcal{U}(d)^p \,\middle|\, \text{the sets } U_1^*\mathcal{C}_1U_1, \ldots, U_p^*\mathcal{C}_pU_p \text{ are } (M, \epsilon)\text{-free}\right\}\right) > 1 - \theta.$$

Our next goal is to prove a result similar to [15, Lemma 6.4.3], that "most" ways of choosing microstates for each of two bi-free pairs of faces individually produce good microstates for the joint system. This also shows why the reverse order is desirable on the right matrices in Definition 2.2 and is reminiscent of the ideas used in [7, Theorem 10.2.1] and [26, Theorem 4.13] to show that a lot related to bi-freeness in our current setting can be extrapolated from certain arrangements of freely independent variables. In order to do so, though, we need the following result.

Lemma 4.5. *Suppose that $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$ is a tracially bi-partite system in a C*-non-commutative probability space $(\mathcal{A}, \varphi)$, so that*

$$(\mathrm{alg}(X_1, \ldots, X_p), \mathrm{alg}(Y_1, \ldots, Y_q)) \qquad \text{and} \qquad (\mathrm{alg}(X_{p+1}, \ldots, X_n), \mathrm{alg}(Y_{q+1}, \ldots, Y_m))$$

are bi-free and that

$$(\{X_1, \ldots, X_p\}, \{Y_1, \ldots, Y_q\}) \qquad \text{and} \qquad (\{X_{p+1}, \ldots, X_n\}, \{Y_{q+1}, \ldots, Y_m\})$$

have finite-dimensional approximants.

Then there is a collection of bounded operators $(\{X_i'\}_{i=1}^n, \{Y_j'\}_{j=1}^m)$ in another C-non-commutative probability space $(\mathcal{A}', \varphi')$ so that for all $0 \leq p, q$, $i_1, \ldots, i_p \in \{1, \ldots, n\}$, and $j_1, \ldots, j_q \in \{1, \ldots, m\}$ we have*

$$\varphi'(X_{i_1}' \cdots X_{i_p}' Y_{j_q}' \cdots Y_{j_1}') = \varphi(X_{i_1} \cdots X_{i_p} Y_{j_1} \cdots Y_{j_q}),$$

and such that

$$\mathrm{alg}(X_1', \ldots, X_p', Y_1', \ldots, Y_q') \qquad \text{and} \qquad \mathrm{alg}(X_{p+1}', \ldots, X_n', Y_{q+1}', \ldots, Y_m')$$

are free.

Proof. Write

$$Z_1 = (\mathrm{alg}(X_1, \ldots, X_p), \mathrm{alg}(Y_1, \ldots, Y_q)) \qquad \text{and} \qquad Z_2 = (\mathrm{alg}(X_{p+1}, \ldots, X_n), \mathrm{alg}(Y_{q+1}, \ldots, Y_m)).$$

For $k = 1, 2$, as Z_k has finite dimensional approximants, by the definition of finite dimensional approximants there exists a state-preserving embedding of Z_k into $(\mathfrak{M}_k, \tau_k)$ where $\mathfrak{M}_k$ is an ultraproduct of matrix algebras and τ_k is a limit of the tracial states along an ultrafilter such that each X_i act via left matrix multiplication and each Y_j acts via right matrix multiplication.

Let $L_2(\mathfrak{M}_k)$ be the GNS construction applied to $(\mathfrak{M}_k, \tau_k)$ and let $\mathcal{H} = L_2(\mathfrak{M}_1) * L_2(\mathfrak{M}_2)$ be the reduced free product with vacuum state ψ. By the definition of bi-free independence from [39], we know that the joint distribution of Z_1 and Z_2 can be obtained by letting Z_k act on $L_2(\mathfrak{M}_k)$ in $\mathcal{H}$ as above, letting all X variables act via the left regular representation, letting all Y variables act via the right regular representation, and computing the joint distribution with respect to ψ.

The desired C*-non-commutative probability space $(\mathcal{A}', \varphi')$ will simply by the bounded linear maps on $\mathcal{H}$ with $\varphi' = \psi$. Moreover, to construct the desired $(\{X_i'\}_{i=1}^n, \{Y_j'\}_{j=1}^m)$, take X_i' to be the operator obtained by letting X_i act on the left of $L_2(\mathfrak{M}_k)$ and via the left regular representation on $L_2(\mathfrak{M}_k)$ in $\mathcal{H}$ where $k = 1$ if $i \leq p$ and $k = 2$ if $i > p$, and Y_j' to be the operator obtained by letting Y_j act on the left of $L_2(\mathfrak{M}_k)$ and via the left regular representation on $L_2(\mathfrak{M}_k)$ in $\mathcal{H}$ where $k = 1$ if $j \leq q$ and $k = 2$ if $j > q$. It is then clear by a definition of free independence that

$$\mathrm{alg}(X_1', \ldots, X_p', Y_1', \ldots, Y_q') \qquad \text{and} \qquad \mathrm{alg}(X_{p+1}', \ldots, X_n', Y_{q+1}', \ldots, Y_m')$$

are freely independent. Moreover, as the action of $Y_{j_q}' \cdots Y_{j_1}'$ on the vacuum vector will be the same as the action of $Y_{j_1} \cdots Y_{j_q}$ on the vacuum vector (as the former acts on the right and via the right regular representation whereas the latter acts on the left via the left regular representation - hence the reversion of the ordering), we obtain the desired moment equality. □

Lemma 4.6. *Let $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$ be a tracially bi-partite system. Suppose that for some $0 \leq p \leq n$ and $0 \leq q \leq m$ that*

$$(\mathrm{alg}(X_1, \dots, X_p), \mathrm{alg}(Y_1, \dots, Y_q)) \qquad \text{and} \qquad (\mathrm{alg}(X_{p+1}, \dots, X_n), \mathrm{alg}(Y_{q+1}, \dots, Y_m))$$

are bi-free and that

$$(\{X_1, \dots, X_p\}, \{Y_1, \dots, Y_q\}) \qquad \text{and} \qquad (\{X_{p+1}, \dots, X_n\}, \{Y_{q+1}, \dots, Y_m\})$$

have finite-dimensional approximants. Then for every $M \in \mathbb{N}$, $\epsilon > 0$, and

$$R > \max\left\{ \max_{1 \leq i \leq n} \|X_i\|, \max_{1 \leq j \leq m} \|Y_j\| \right\}$$

there exists an $\epsilon_1 > 0$ such that

$$\lim_{d \to \infty} \frac{\lambda_{d,n+m}\left(\Psi_d(M, \epsilon_1) \cap \Theta_d(M, \epsilon)\right)}{\lambda_{d,n+m}\left(\Psi_d(M, \epsilon_1)\right)} = 1$$

where $\frac{0}{0} = 1$,

$$\Psi_d(M, \epsilon_1) = \Gamma_R(X_1, \dots, X_p \sqcup Y_1, \dots, Y_q; M, d, \epsilon_1) \times_{\ell r} \Gamma_R(X_{p+1}, \dots, X_n \sqcup Y_{q+1}, \dots, Y_m; M, d, \epsilon_1),$$
$$\Theta_d(M, \epsilon) = \Gamma_R(X_1, \dots, X_n \sqcup Y_1, \dots, Y_m; M, d, \epsilon),$$

and $\times_{\ell r}$ is as defined in the proof of Proposition 2.5.

Proof. Fix $M \in \mathbb{N}$, $\epsilon > 0$, and R as described. We claim that there exists an $\epsilon_1 > 0$ such that if

$$(A_1, \dots, A_n, B_1, \dots, B_m) \in \Psi_d(M, \epsilon_1)$$

and if

$$\{A_1, \dots, A_p, B_1, \dots, B_q\} \quad \text{and} \quad \{A_{p+1}, \dots, A_n, B_{q+1}, \dots, B_m\} \qquad \text{are } (M, \epsilon_1)\text{-free},$$

then

$$(A_1, \dots, A_n, B_1, \dots, B_m) \in \Theta_d(M, \epsilon).$$

To see this, we first take a collection of operators $X_1', \dots, X_n', Y_1', \dots, Y_m'$ as in Lemma 4.5, and let R_1 be larger than the largest of their norms. Then for suitably small $\epsilon_1 > 0$, if

$$(A_1, \dots, A_n, B_1, \dots, B_m) \in \Psi_d(M, \epsilon_1)$$

and if

$$\{A_1, \dots, A_p, B_1, \dots, B_q\} \quad \text{and} \quad \{A_{p+1}, \dots, A_n, B_{q+1}, \dots, B_m\} \qquad \text{are } (M, \epsilon_1)\text{-free},$$

then for all $0 \leq p, q$ with $p + q \leq M$, $i_1, \dots i_p \in \{1, \dots, n\}$, and $j_1, \dots, j_q \in \{1, \dots, m\}$ we have that

$$\tau_d(A_{i_1} \cdots A_{i_p} B_{j_q} \cdots B_{j_1})$$

is within a multiple of ϵ_1 (involving M and R_1) of

$$\varphi'(X_{i_1}' \cdots X_{i_p}' Y_{j_q}' \cdots Y_{j_1}') = \varphi(X_{i_1} \cdots X_{i_p} Y_{j_1} \cdots Y_{j_q}),$$

thereby completing the claim.

Given $\theta > 0$, by Lemma 4.4 there exists an $N \in \mathbb{N}$ such that

$$\gamma_d\left(\left\{ U \in \mathcal{U}(d) \,\middle|\, \substack{\{A_1, \dots, A_p, B_1, \dots, B_q\} \text{ and } \{U^* A_{p+1} U, \dots, U^* A_n U, U^* B_{q+1} U, \dots, U^* B_m U\} \\ \text{are } (M, \epsilon_1)\text{-free}} \right\}\right) \geq 1 - \theta$$

for all $d \geq N$ and all $A_i, B_j \in \mathcal{M}_d^{\mathrm{sa}}$ with $\|A_i\|, \|B_j\| \leq R$ for $1 \leq i \leq n$ and $1 \leq j \leq m$.

By the assumption of finite-dimensional approximants, $\Psi_d(M, \epsilon_1)$ is non-empty for sufficiently large d. Let ν_d denote the normalized restriction of $\lambda_{d,n+m}$ to $\Psi_d(M, \epsilon_1)$. Since both $\Psi_d(M, \epsilon_1)$ and ν_d are invariant under the action of $\mathcal{U}(d)$ given by

$$(A_1, \dots, A_n, B_1, \dots, B_n) \mapsto (A_1, \dots, A_p, U^* A_{p+1} U, \dots, U^* A_n U, B_1, \dots, B_q, U^* B_{q+1} U, \dots, U^* B_m U),$$

we obtain that

$$\frac{\lambda_{d,n+m}\left(\Psi_d(M, \epsilon_1) \cap \Theta_d(M, \epsilon)\right)}{\lambda_{d,n+m}\left(\Psi_d(M, \epsilon_1)\right)}$$
$$= \int_{\Psi_d(M, \epsilon_1)} \left(\int_{\mathcal{U}(d)} 1_{\Theta_d(M, \epsilon)}(A_1, \dots, A_p, U^* A_{p+1} U, \dots, U^* A_n U, B_1, \dots, B_q, U^* B_{q+1} U, \dots, U^* B_m U) \, d\gamma(U) \right) d\nu_d.$$

By the choice of ϵ_1, we obtain for sufficiently large d that

$$\int_{\mathcal{U}(d)} 1_{\Theta_d(M,\epsilon)}(A_1,\ldots,A_p,U^*A_{p+1}U,\ldots,U^*A_nU,B_1,\ldots,B_q,U^*B_{q+1}U,\ldots,U^*B_mU)\,d\gamma(U) > 1-\theta.$$

Hence

$$\frac{\lambda_{d,n+m}\left(\Psi_d(M,\epsilon_1)\cap\Theta_d(M,\epsilon)\right)}{\lambda_{d,n+m}\left(\Psi_d(M,\epsilon_1)\right)} \geq 1-\theta$$

which completes the proof as θ was arbitrary. □

Unfortunately, at this point in trying to prove additivity of microstate bi-free entropy for bi-free collections, we reach a bit of an impasse. Either we need to know that the $\limsup_{d\to\infty}$ in Definition 2.2 is actually a limit, or we need to replaces the $\limsup_{d\to\infty}$ with a limit along an ultrafilter. Thus, for the following result, we use $\chi^\omega(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m)$ to denote the same quantity as in Definition 2.2 where $\limsup_{d\to\infty}$ is replaced with $\limsup_{d\to\omega}$ for some ω a non-principle ultrafilter on $\mathcal{N}$.

Theorem 4.7. *Let $(\{X_i\}_{i=1}^n,\{Y_j\}_{j=1}^m)$ be a tracially bi-partite system. Suppose that for some $0\leq p\leq n$ and $0\leq q\leq m$ that*

$$(\mathrm{alg}(X_1,\ldots,X_p),\mathrm{alg}(Y_1,\ldots,Y_q)) \qquad \textit{and} \qquad (\mathrm{alg}(X_{p+1},\ldots,X_n),\mathrm{alg}(Y_{q+1},\ldots,Y_m))$$

are bi-free. Suppose that the limits superior in Definition 2.2 for the bi-free entropies of $(\{X_i\}_{i=1}^p,\{Y_j\}_{j=1}^q)$ and $(\{X_i\}_{i=p+1}^n,\{Y_j\}_{j=q+1}^m)$ are attained along a common sequence of dimensions (as is the case when one of the two can be replaced by a limit). Then

$$\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m) = \chi(X_1,\ldots,X_p\sqcup Y_1,\ldots,Y_q)+\chi(X_{p+1},\ldots,X_n\sqcup Y_{q+1},\ldots,Y_m).$$

Alternatively,

$$\chi^\omega(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m) = \chi^\omega(X_1,\ldots,X_p\sqcup Y_1,\ldots,Y_q)+\chi^\omega(X_{p+1},\ldots,X_n\sqcup Y_{q+1},\ldots,Y_m).$$

Proof. By Proposition 2.5

$$\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m) \leq \chi(X_1,\ldots,X_p\sqcup Y_1,\ldots,Y_q)+\chi(X_{p+1},\ldots,X_n\sqcup Y_{q+1},\ldots,Y_m)$$

so the result is immediate if either quantity on the right hand side is $-\infty$. Thus we may assume these microstate bi-free entropies are finite (and thus have finite-dimensional approximants) and proceed with demonstrating the other inequality.

For any $M\in\mathbb{N}$, $\epsilon>0$, and

$$R > \max\left\{\max_{1\leq i\leq n}\|X_i\|,\max_{1\leq j\leq m}\|Y_j\|\right\},$$

Lemma 4.6 implies there exists an $\epsilon_1>0$ such that

$$\lim_{d\to\infty}\frac{\lambda_{d,n+m}\left(\Psi_d(M,\epsilon_1)\cap\Theta_d(M,\epsilon)\right)}{\lambda_{d,n+m}\left(\Psi_d(M,\epsilon_1)\right)} = 1,$$

where

$$\begin{aligned}\Psi_d(M,\epsilon_1) &= \Gamma_R(X_1,\ldots,X_p\sqcup Y_1,\ldots,Y_q;M,d,\epsilon_1)\times_{\ell r}\Gamma_R(X_{p+1},\ldots,X_n\sqcup Y_{q+1},\ldots,Y_m;M,d,\epsilon_1)\\ \Theta_d(M,\epsilon) &= \Gamma_R(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m;M,d,\epsilon).\end{aligned}$$

Now, assuming that the limits superior in Definition 2.2 are attained along a common subsequence,

$$\begin{aligned}
&\chi_R(X_1,\ldots,X_n \sqcup Y_1,\ldots,Y_m; M,\epsilon)\\
&= \limsup_{d\to\infty} \frac{1}{d^2}\log(\lambda_{d,n+m}(\Theta_d(M,\epsilon))) + \frac{1}{2}(n+m)\log(d)\\
&\geq \limsup_{d\to\infty} \frac{1}{d^2}\log(\lambda_{d,n+m}(\Theta_d(M,\epsilon)\cap\Psi_d(M,\epsilon_1))) + \frac{1}{2}(n+m)\log(d)\\
&= \limsup_{d\to\infty} \frac{1}{d^2}\log(\lambda_{d,n+m}(\Psi_d(M,\epsilon_1))) + \frac{1}{2}(n+m)\log(d)\\
&= \limsup_{d\to\infty} \Bigg(\frac{1}{d^2}\log(\lambda_{d,n+m}(\Gamma_R(X_1,\ldots,X_p\sqcup Y_1,\ldots,Y_q;M,d,\epsilon_1))) + \frac{1}{2}(p+q)\log(d)\\
&\qquad + \frac{1}{d^2}\log(\lambda_{d,n+m}(\Gamma_R(X_{p+1},\ldots,X_n\sqcup Y_{q+1},\ldots,Y_m;M,d,\epsilon_1))) + \frac{1}{2}(n+m-p-q)\log(d)\Bigg)\\
&= \limsup_{d\to\infty} \left(\frac{1}{d^2}\log(\lambda_{d,n+m}(\Gamma_R(X_1,\ldots,X_p\sqcup Y_1,\ldots,Y_q;M,d,\epsilon_1))) + \frac{1}{2}(p+q)\log(d)\right)\\
&\qquad + \limsup_{d\to\infty} \left(\frac{1}{d^2}\log(\lambda_{d,n+m}(\Gamma_R(X_{p+1},\ldots,X_n\sqcup Y_{q+1},\ldots,Y_m;M,d,\epsilon_1))) + \frac{1}{2}(n+m-p-q)\log(d)\right)\\
&= \chi_R(X_1,\ldots,X_p\sqcup Y_1,\ldots,Y_q;M,\epsilon_1) + \chi_R(X_{p+1},\ldots,X_n\sqcup Y_{q+1},\ldots,Y_m;M,\epsilon_1)\\
&\geq \chi_R(X_1,\ldots,X_p\sqcup Y_1,\ldots,Y_q) + \chi_R(X_{p+1},\ldots,X_n\sqcup Y_{q+1},\ldots,Y_m).
\end{aligned}$$

Here the limit superior splits across the sum because of our assumption about a common subsequence, and in the last inequality we have used the fact that $\chi_R(\cdot\sqcup\cdot;M,\epsilon)$ decreases as M increases and as ϵ decreases.

The result for χ^ω easily follows by similar arguments. □

Corollary 4.8. *Let $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$ be a tracially bi-partite system. If*

$$\mathrm{alg}(\{X_1,\ldots,X_n\}) \qquad \textit{and} \qquad \mathrm{alg}(\{Y_1,\ldots,Y_m\})$$

are classically independent and if the $\limsup_{d\to\infty}$ *in Definition 2.2 is actually a limit for $\{X_i\}_{i=1}^n$ and for $\{Y_j\}_{j=1}^m$), then*

$$\chi(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m) = \chi(X_1,\ldots,X_n) + \chi(Y_1,\ldots,Y_m).$$

Alternatively,

$$\chi^\omega(X_1,\ldots,X_n\sqcup Y_1,\ldots,Y_m) = \chi^\omega(X_1,\ldots,X_n) + \chi^\omega(Y_1,\ldots,Y_m).$$

Proof. We recall from [39] that classical independence implies the bi-freeness of $(\mathrm{alg}(\{X_1,\ldots,X_n\}),\mathbb{C})$ from $(\mathbb{C},\mathrm{alg}(\{Y_1,\ldots,Y_m\}))$. Theorem 4.7 then allows us to equate the bi-free entropy of the whole system with the sum of the bi-free entropies of the left variables and of the right variables, which by Remark 2.4 is just their free entropies. □

5. Orbital Bi-Free Entropy

In this section, we will develop a bi-free analogue of the notion of orbital free entropy, which was introduced in [14] and is deeply connected to microstate free entropy. Among other results, we use the joint orbital bi-free entropy to give a strengthened version of the subadditivity estimate from Proposition 2.5 in Theorem 5.12. The approach used here is based on that of [29] which is a close thematic fit to this paper; although it may be interesting to consider an approach similar to that of [4], we do not do so here. In fact, most proofs in this section are adaptations of those from [29].

Throughout this section, let $(\mathcal{A},\varphi)$ be a C*-non-commutative probability space and let $\ell\in\mathbb{N}$. For each $1\leq k\leq\ell$, let $\mathbf{X}_k = (X_{k,1},X_{k,2},\ldots,X_{k,n_k})$ and $\mathbf{Y}_k = (Y_{k,1},Y_{k,2},\ldots,Y_{k,m_k})$ denote, respectively, n_k- and m_k-tuples of self-adjoint operators from $\mathcal{A}$, where $n_k,m_k\geq 0$ with $n_k+m_k\geq 1$. Furthermore, we will use $\mathbf{Z}_k$ to denote the system of variables $\mathbf{X}_k\sqcup\mathbf{Y}_k$ where $\mathbf{X}_k$ are viewed as the left variables and $\mathbf{Y}_k$ are viewed as the right variables.

Let $U(d)$ denote the unitary matrices from $\mathcal{M}_d$ and let γ_d denote the Haar measure on $U(d)$. Furthermore define

$$\Phi_d : U(d)^\ell \times \left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{m_k}\right) \to \left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{m_k}\right)$$

by

$$\Phi_d((U_k)_{k=1}^\ell, (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell) = ((U_k^*\mathbf{A}_kU_k)_{k=1}^\ell, (U_k^*\mathbf{B}_kU_k)_{k=1}^\ell)$$

where for $\mathbf{A}_k = (A_{k,1}, A_{k,2}, \ldots, A_{k,n_k}) \in (M_d^{\mathrm{sa}})^{n_k}$ and $\mathbf{B}_k = (B_{k,1}, B_{k,2}, \ldots, B_{k,m_k}) \in (M_d^{\mathrm{sa}})^{m_k}$,

$$U_k^*\mathbf{A}_kU_k = (U_k^*A_{k,1}U_k, U_k^*A_{k,2}U_k, \ldots, U_k^*A_{k,n_k}U_k) \text{ and}$$
$$U_k^*\mathbf{B}_kU_k = (U_k^*B_{k,1}U_k, U_k^*B_{k,2}U_k, \ldots, U_k^*B_{k,m_k}U_k).$$

Moreover, let $\mathcal{P}\left(\left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{m_k}\right)\right)$ denote the set of all regular Borel probability measures on $\left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{m_k}\right)$.

Using

$$\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon)$$

to denote the bi-free microstates based on the self-adjoint variables contained in the left variables $\mathbf{X}_1, \ldots, \mathbf{X}_\ell$ in the order listed and the right variables $\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell$ in the order listed, we may now define the object of study in this section.

Definition 5.1. With the above notation, for each $\mu \in \mathcal{P}\left(\left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{m_k}\right)\right)$, $M, d \in \mathbb{N}$, and $R, \epsilon > 0$, let

$$\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon; \mu) = \log\left(\left(\gamma_d^{\otimes \ell} \otimes \mu\right)\left(\Phi_d^{-1}(\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon))\right)\right)$$

(with $\log(0) = -\infty$). With this we define

$$\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) = \sup_{\mu \in \mathcal{P}\left(\left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{n_k}\right)\times\left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{m_k}\right)\right)} \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon; \mu),$$
$$\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, \epsilon) = \limsup_{d\to\infty} \frac{1}{d^2}\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon),$$
$$\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = \inf\{\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, \epsilon) \mid M \in \mathbb{N}, \epsilon > 0), \text{ and}$$
$$\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = \sup_{0<R<\infty} \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell).$$

The quantity $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \in [-\infty, 0]$ will be called the orbital bi-free entropy of the collections $\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell$. Note that the fact that $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \leq 0$ is clear by definition.

Remark 5.2. Based on the definition of $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell)$, we can see that the orbital bi-free entropy is a measure of how well conjugation by unitaries preserves the bi-free microstates. We can see that the infimum over ϵ and M occurs as ϵ tends to 0 and M tends to infinity. Furthermore, it is not difficult to see that if $m_k = 0$ for all k, then $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell)$ agrees with $\chi_{\mathrm{orb}}(\mathbf{X}_1, \ldots, \mathbf{X}_\ell)$ as in [29, Definition 2.1].

If the variables in question are not tracially bipartite, we have $\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon) = \emptyset$ for appropriately antagonistic parameters, and so $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = -\infty$. We therefore continue to make the assumption that the variables are tracially bi-partite throughout this section.

Although the supremum over the probability measures portion of Definition 5.1 may seem difficult to compute with, from the theoretical standpoint it is quite natural, as we will see. However, as with [29], there are other ways to describe $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell)$ without the need to take a supremum over probability measures. To provide one such description, first we need to develop some additional notation and demonstrate a lemma that will be useful throughout the section.

Given $((\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell) \in \left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_d^{\mathrm{sa}})^{m_k}\right)$, let

$$\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell; M, d, \epsilon)$$

denote the set of all $(U_k)_{k=1}^\ell \in (U(d))^\ell$ such that

$$\Phi_d((U_k)_{k=1}^\ell, (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell) \in \Gamma_\infty(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon).$$

Note that for $\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell; M, d, \epsilon)$ to be non-empty, each $(\mathbf{A}_k, \mathbf{B}_k)$ must be good microstates for $\mathbf{X}_k \sqcup \mathbf{Y}_k$; more precisely, if $(\mathcal{M}_d^{\mathrm{sa}})_R$ denotes all elements of $\mathcal{M}_d^{\mathrm{sa}}$ of operator norm at most R and if for some k we have that $(\mathbf{A}_k, \mathbf{B}_k) \in ((\mathcal{M}_d^{\mathrm{sa}})_R)^{n_k+m_k} \setminus \Gamma_R(\mathbf{X}_k \sqcup \mathbf{Y}_k; M, d, \epsilon)$, then

$$\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell; M, d, \epsilon) = \emptyset.$$

We record the following technical observation as a lemma.

Lemma 5.3. *For every $R > 0$, the map from $\left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})_R^{n_k}\right) \times \left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})_R^{m_k}\right)$ to $\mathbb{R}$ defined by*

$$((\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell) \mapsto \gamma_d^{\otimes \ell}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell; M, d, \epsilon)\right)$$

is Borel. Furthermore, for every $\mu \in \mathcal{P}\left(\left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})^{m_k}\right)\right)$,

$$\begin{aligned}
&(\gamma_d^{\otimes \ell} \otimes \mu)\left(\Phi_d^{-1}(\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon))\right) \\
&\qquad = \int_{\left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})_R^{n_k}\right) \times \left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})_R^{m_k}\right)} \gamma_d^{\otimes \ell}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell; M, d, \epsilon)\right) d\mu\left((\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell\right) \\
&\qquad = \int_{\prod_{k=1}^\ell \Gamma_R(\mathbf{X}_k \sqcup \mathbf{Y}_k; M, d, \epsilon)} \gamma_d^{\otimes \ell}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell; M, d, \epsilon)\right) d\mu\left((\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell\right),
\end{aligned}$$

with an implicit reordering of coordinates in the second integral.

Proof. The result follows from the fact that the sets and functions involved are Borel, by the above constructions, and by Fubini's Theorem. ☐

Now we are able to demonstrate an alternate definition of the orbital bi-free entropy without the need to take a supremum over probability measures.

Proposition 5.4. *For each $M, d \in \mathbb{N}$, $\epsilon > 0$, and $R \in (0, \infty]$, let*

$$\begin{aligned}
&\tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) \\
&= \sup_{((\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell) \in \left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})_R^{n_k}\right) \times \left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})_R^{m_k}\right)} \log\left(\gamma_d^{\otimes \ell}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell; M, d, \epsilon)\right)\right) \\
&= \sup_{((\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell) \in \prod_{k=1}^\ell \Gamma_R(\mathbf{X}_k \sqcup \mathbf{Y}_k; M, d, \epsilon)} \log\left(\gamma_d^{\otimes \ell}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell; M, d, \epsilon)\right)\right),
\end{aligned}$$

Note $\tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) \in [-\infty, 0]$. Then

$$\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = \inf_{M \in \mathbb{N}, \epsilon > 0} \limsup_{d \to \infty} \frac{1}{d^2} \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon).$$

Proof. First, it is clear that the two definitions of $\tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon)$ are equivalent by the comments before Lemma 5.3. Furthermore, Lemma 5.3 implies that

$$\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon; \mu) \le \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon)$$

for any $\mu \in \mathcal{P}\left(\left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})^{m_k}\right)\right)$. Hence we clearly have

$$\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \le \inf_{M \in \mathbb{N}, \epsilon > 0} \limsup_{d \to \infty} \frac{1}{d^2} \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon).$$

To prove the reverse inequality, consider M and ϵ fixed. If $\tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) = -\infty$ for all sufficiently large d then

$$\limsup_{d \to \infty} \frac{1}{d^2} \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) \le \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, \epsilon)$$

trivially follows. Otherwise there exists an increasing sequence $(d_l)_{l \ge 1}$ such that $\tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) > -\infty$ and

$$\limsup_{d \to \infty} \frac{1}{d^2} \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) = \lim_{l \to \infty} \frac{1}{d_l^2} \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon).$$

For each $l \in \mathbb{N}$, we can choose $((\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}) \in \left(\prod_{k=1}^{\ell}(M_{d_l}^{\mathrm{sa}})_R^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_{d_l}^{\mathrm{sa}})_R^{m_k}\right)$ such that

$$\tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon) - 1 \leq \log\left(\gamma_{d_l}^{\otimes \ell}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; M, d_l, \epsilon)\right)\right).$$

Therefore, if $\delta_l \in \mathcal{P}\left(\left(\prod_{k=1}^{\ell}(M_{d_l}^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_{d_l}^{\mathrm{sa}})^{m_k}\right)\right)$ is the point-mass measure at $((\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell})$, then Lemma 5.3 implies that

$$\begin{aligned}
&\tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon) - 1 \\
&\leq \log\left(\gamma_{d_l}^{\otimes \ell}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; M, d_l, \epsilon)\right)\right) \\
&= \log\left(\int_{\left(\prod_{k=1}^{\ell}(M_{d_l}^{\mathrm{sa}})_R^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_{d_l}^{\mathrm{sa}})_R^{m_k}\right)} \gamma_{d_l}^{\otimes \ell}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; M, d_l, \epsilon)\right) d\delta_l\right) \\
&= \log\left((\gamma_{d_l}^{\otimes \ell} \otimes \delta_l)\left(\Phi_{d_l}^{-1}(\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d_l, \epsilon))\right)\right) \\
&= \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon; \delta_l) \\
&\leq \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon).
\end{aligned}$$

Hence

$$\begin{aligned}
\limsup_{d \to \infty} \frac{1}{d^2} \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) &= \lim_{l \to \infty} \frac{1}{d_l^2} \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon) \\
&\leq \limsup_{l \to \infty} \frac{1}{d_l^2}(\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon) + 1) \\
&\leq \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, \epsilon).
\end{aligned}$$

Thus the result follows. □

Remark 5.5. We note that [29] has an alternate characterization of the orbital free entropy in the case that the von Neumann algebras generated by each collection of left operators is hyperfinite. Unfortunately, the arguments for such a characterization break down in the case of orbital bi-free entropy due to the fact that we are no longer dealing with a tracial state, so Jung's lemma [19, Lemma 2.9] no longer applies. To find an analogue of this result, one would need to be able to change all right operators to left ones while still maintaining hyperfiniteness, which does not seem like a natural assumption.

However, many basic properties of the orbital free entropy extend to the bi-free setting. To begin, we have the following lemma showing the independence of R.

Lemma 5.6. *Let $\rho = \max\{\|X_{i,k}\| \mid 1 \leq i \leq n_k, 1 \leq k \leq \ell\} \cup \{\|Y_{j,k}\| \mid 1 \leq j \leq m_k, 1 \leq k \leq \ell\})$. For any $R > \rho$, we have*

$$\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell)$$

including the case $R = \infty$.

Proof. Fix $R > \rho$. Clearly

$$\chi_{\mathrm{orb},\infty}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \geq \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell)$$

by definitions so it suffices to prove the other inequality. To begin, define $f : \mathbb{R} \to [-1, 1]$ by

$$f(z) = \begin{cases} z & \text{if } z \in [-1, 1] \\ -1 & \text{if } z < -1 \\ 1 & \text{if } z > 1 \end{cases},$$

and let $f_R : \mathbb{R} \to [-R, R]$ by $f_R(z) = Rf\left(\frac{z}{R}\right)$.

Fix $M \in \mathbb{N}$ and $\epsilon > 0$. Let $K = \max\{(\rho^{2M} + 1)^{\frac{1}{2M}}, R\} \geq 1$, and choose $0 < \epsilon_0 < \frac{\epsilon}{2}$ and $M' \in \mathbb{N}$ even with $M' \geq 2M$ so that

$$R\left(\left(\frac{\rho}{R}\right)^{M'} + \frac{\epsilon_0}{R^{M'}}\right)^{\frac{1}{M}} < \frac{\epsilon}{2MK^M}.$$

Let $(\mathbf{A}_k, \mathbf{B}_k) \in \Gamma_\infty(\mathbf{X}_k \sqcup \mathbf{Y}_k; M', d, \epsilon_0)$ be arbitrary. Then $\tau_d(A_{k,i}^{M''}) < \varphi(X_{k,i}^{M''}) + \epsilon_0 \leq \rho^{M''} + \epsilon_0$ for all $1 \leq i \leq n_k$ and M'' even and not greater than M'. Thus, if for $p \in [1, \infty)$, $\|\cdot\|_p$ denotes the p-norm on $\mathcal{M}_d$

with respect to τ_d, we have that $\|A_{k,i}\|_p \leq \|A_{k,i}\|_{M'} \leq K$ for all $p \leq M'$, and in particular, for all $p \leq 2M$. Thus, if $a_1, a_2, \ldots, a_d$ are the eigenvalues of $A_{k,i}$ counting multiplicities, we have for $p < M$ that

$$\begin{aligned}\|A_{k,i} - f_R(A_{k,i})\|_p \leq \|A_{k,i} - f_R(A_{k,i})\|_M &\leq R\left(\frac{1}{d}\sum_{|a_q|>R}\left|\frac{a_q}{R}\right|^M\right)^{\frac{1}{M}} \\ &\leq R\left(\frac{1}{d}\sum_{|a_q|>R}\left|\frac{a_q}{R}\right|^{M'}\right)^{\frac{1}{M}} \\ &\leq R\left(\frac{\tau_d(A_{k,i}^{M'})}{R^{M'}}\right)^{\frac{1}{M}} \\ &\leq R\left(\left(\frac{\rho}{R}\right)^{M'} + \frac{\epsilon_0}{R^{M'}}\right)^{\frac{1}{M}} < \frac{\epsilon}{2MK^M}.\end{aligned}$$

Moreover, clearly $\|f_R(A_{k,i})\|_p \leq R \leq K$ for every p, k, and i, and identical inequalities holds for the $B_{k,j}$'s.

The above implies if $(U_k)_{k=1}^\ell \in \Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; (\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell; M', d, \epsilon_0)$, then for all p, q with $p + q \leq M$, for all $k_1, \ldots, k_p, l_1, \ldots, l_q \in \{1, \ldots, \ell\}$, for all valid indices $i_1, \ldots, i_p, j_1, \ldots, j_q$, we have

$$\begin{aligned}&\Big|\tau_d\left(U_{k_1}^* f_R(A_{k_1,i_1})U_{k_1} \cdots U_{k_p}^* f_R(A_{k_p,i_p})U_{k_p} U_{l_q}^* f_R(B_{l_q,j_l})U_{l_q} \cdots U_{l_1}^* f_R(B_{l_1,j_1})U_{i_1}\right) \\ &\qquad - \varphi\left(X_{k_1,i_1} \cdots X_{k_p,i_p} Y_{l_1,j_1} \cdots Y_{l_q,j_q}\right)\Big| \\ &\leq \Big|\tau_d\left(U_{k_1}^* f_R(A_{k_1,i_1})U_{k_1} \cdots U_{k_p}^* f_R(A_{k_p,i_p})U_{k_p} U_{l_q}^* f_R(B_{l_q,j_l})U_{l_q} \cdots U_{l_1}^* f_R(B_{l_1,j_1})U_{i_1}\right) \\ &\qquad - \tau_d\left(U_{k_1}^* A_{k_1,i_1}U_{k_1} \cdots U_{k_p}^* A_{k_p,i_p}U_{k_p} U_{l_q}^* B_{l_q,j_l}U_{l_q} \cdots U_{l_1}^* B_{l_1,j_1}U_{i_1}\right)\Big| \\ &\quad + \Big|\tau_d\left(U_{k_1}^* A_{k_1,i_1}U_{k_1} \cdots U_{k_p}^* A_{k_p,i_p}U_{k_p} U_{l_q}^* B_{l_q,j_l}U_{l_q} \cdots U_{l_1}^* B_{l_1,j_1}U_{i_1}\right) - \varphi\left(X_{k_1,i_1} \cdots X_{k_p,i_p} Y_{l_1,j_1} \cdots Y_{l_q,j_q}\right)\Big| \\ &\leq \sum_{x=1}^p K^{p+q}\left\|U_{k_x}^*(f_R(A_{k_x,i_x}) - A_{k_x,i_x})U_{k_x}\right\|_{p+q} + \sum_{y=1}^q K^{p+q}\left\|U_{k_y}^*(f_R(B_{k_y,j_y}) - B_{k_y,j_y})U_{k_y}\right\|_{p+q} + \epsilon_0 \\ &\leq MK^M\frac{\epsilon}{2MK^M} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

where the second inequality if shown by the generalized Hölder's inequality for matrices. Hence

$$\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; (\mathbf{A})_{k=1}^\ell, (\mathbf{B})_{k=1}^\ell; M', d, \epsilon_0) \subseteq \Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; (\mathbf{f_R}(\mathbf{A_k}))_{k=1}^\ell, (\mathbf{f_R}(\mathbf{B_k}))_{k=1}^\ell; M, d, \epsilon)$$

thereby implying

$$\tilde{\chi}_{\mathrm{orb},\infty}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M', d, \epsilon_0) \leq \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon)$$

as $\|f_R(A_{k,i})\| \leq R$ and $\|f_R(B_{k,j})\| \leq R$ for all i, j, k. Hence Proposition 5.4 implies that

$$\chi_{\mathrm{orb},\infty}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \leq \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell).$$ □

Some basic properties of orbital bi-free entropy are readily established.

Proposition 5.7. *The following hold:*

1. $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = -\infty$ *if* $\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell$ *do not have finite-dimensional approximants.*
2. *For a single system of variables* $\mathbf{Z}$, $\chi_{\mathrm{orb}}(\mathbf{Z}) = 0$ *if* $\mathbf{Z}$ *has finite-dimensional approximants, and* $\chi_{\mathrm{orb}}(\mathbf{Z}) = -\infty$ *otherwise.*
3. $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \leq \chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_q) + \chi_{\mathrm{orb}}(\mathbf{Z}_{q+1}, \ldots, \mathbf{Z}_\ell)$ *for all* $1 \leq q < \ell$.

Proof. Note (1) follows as if $\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell$ do not have finite-dimensional approximants, then

$$\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon) = \emptyset$$

for sufficiently large M, sufficiently small ϵ, and for all d (see Remark 4.2). Note (2) follows as the $U(d)$ section of $\Phi_d^{-1}(\Gamma_R(\mathbb{Z}; M, d, \epsilon))$ is $U(d)$ if $\Gamma_R(\mathbb{Z}; M, d, \epsilon) \neq \emptyset$ and is $\emptyset$ if $\Gamma_R(\mathbb{Z}; M, d, \epsilon) = \emptyset$. Finally, (3) holds as if an ℓ-tuple of unitaries works in the first step of the definition of $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell)$, then the first q work in the first step of the definition for $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_q)$ and the remainder work in the first step of the definition for $\chi_{\mathrm{orb}}(\mathbf{Z}_{q+1}, \ldots, \mathbf{Z}_\ell)$. □

We now examine the behaviour of orbital bi-free entropy under transformations of each family of variables. Much like in Section 3, we restrict our attention to transformations which do not mix left and right variables. We nonetheless obtain a result similar to that of [29, Theorem 2.6 (6)], whose proof we adapt.

Proposition 5.8. *Suppose that $\mathbf{X}'_k \subset W^*(\mathbf{X}_k)$ and $\mathbf{Y}'_k \subset W^*(\mathbf{Y}_k)$ are tuples of self-adjoint operators, and let $\mathbf{Z}'_k$ denote the system $\mathbf{X}'_k \sqcup \mathbf{Y}'_k$ with with the $\mathbf{X}'_k$ viewed as left operators, and the $\mathbf{Y}'_k$ viewed as right operators. Then*

$$\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \leq \chi_{\mathrm{orb}}(\mathbf{Z}'_1, \ldots, \mathbf{Z}'_\ell).$$

In particular, $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_n)$ depends only on the pairs $(W^*(\mathbf{X}_k), W^*(\mathbf{Y}_k))$ of von Neumann subalgebras of $\mathcal{A}$.

Proof. Our approach will be to establish an inclusion of appropriate Γ_{orb} sets. Therefore, let $M \in \mathbb{N}$ and $\delta > 0$ be arbitrary.

Since we are working in a tracially bi-partite setting, we have by assumption that $\varphi|_{W^*(\mathbf{X}_k)}$ and $\varphi|_{W^*(\mathbf{Y}_k)}$ are tracial states. We may therefore invoke the Kaplansky density theorem to find polynomial approximations $\mathbf{P}_k$ and $\mathbf{Q}_k$ of $\mathbf{X}'_k$ and $\mathbf{Y}'_k$ in $\mathbb{C}\langle \mathbf{X}_k \rangle$ and $\mathbb{C}\langle \mathbf{Y}_k \rangle$ respectively, so that each $\|P_{kj}(\mathbf{X}_k)\| \leq \left\|X'_{kj}\right\|$ and $\|Q_{kj}(\mathbf{Y}_k)\| \leq \left\|Y'_{kj}\right\|$, and so that

$$\varphi\left(\left|P_{kj}(\mathbf{X}_k) - X'_{kj}\right|\right), \varphi\left(\left|Q_{kj}(\mathbf{Y}_k) - Y'_{kj}\right|\right) < \frac{\delta}{2ML^{M-1}}$$

for any fixed L with $1, \left\|X'_{kj}\right\|, \left\|Y'_{kj}\right\| \leq L$ for all k and j in the appropriate ranges.

As the $\mathbf{P}_k$ and $\mathbf{Q}_k$ are polynomials, there are $M' \in \mathbb{N}$ and $\delta' > 0$ so that if $\left((\mathbf{A}_k)_{k=1}^{\ell}, (\mathbf{B}_k)_{k=1}^{\ell}\right) \in \Gamma_\infty(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M', d, \delta')$, then

$$\Big|\tau_d\left(P_{k_1j_1}(\mathbf{A}_{k_1}) \cdots P_{k_sj_s}(\mathbf{A}_{k_s}) Q_{k_tj_t}(\mathbf{B}_{k_t}) \cdots Q_{k_{s+1}j_{s+1}}(\mathbf{B}_{k_{s+1}})\right) \\ - \varphi\left(P_{k_1j_1}(\mathbf{X}_{k_1}) \cdots P_{k_sj_s}(\mathbf{X}_{k_s}) Q_{k_{s+1}j_{s+1}}(\mathbf{Y}_{k_{s+1}}) \cdots Q_{k_tj_t}(\mathbf{Y}_{k_t})\right)\Big| < \frac{\delta}{2}$$

for every appropriate selection of indices with $t \leq M$. Note in particular that M', δ' may be chosen independently of d.

But now, given $\left((\mathbf{A}_k)_{k=1}^{\ell}, (\mathbf{B}_k)_{k=1}^{\ell}\right) \in \Gamma_\infty(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M', d, \delta')$, we have the estimates

$$\begin{aligned} &\left|\tau_d\left(P_{k_1j_1}(\mathbf{A}_{k_1}) \cdots P_{k_sj_s}(\mathbf{A}_{k_s}) Q_{k_tj_t}(\mathbf{B}_{k_t}) \cdots Q_{k_{s+1}j_{s+1}}(\mathbf{B}_{k_{s+1}})\right) - \varphi\left(\mathbf{X}'_{k_1j_1} \cdots \mathbf{X}'_{k_sj_s} \mathbf{Y}'_{k_{s+1}j_{s+1}} \cdots \mathbf{Y}'_{k_tj_t}\right)\right| \\ &\quad < \frac{\delta}{2} + \left|\varphi\left(P_{k_1j_1}(\mathbf{X}_{k_1}) \cdots P_{k_sj_s}(\mathbf{X}_{k_s}) Q_{k_{s+1}j_{s+1}}(\mathbf{Y}_{k_{s+1}}) \cdots Q_{k_tj_t}(\mathbf{Y}_{k_t})\right) - \varphi\left(\mathbf{X}'_{k_1j_1} \cdots \mathbf{X}'_{k_sj_s} \mathbf{Y}'_{k_{s+1}j_{s+1}} \cdots \mathbf{Y}'_{k_tj_t}\right)\right| \\ &\quad \leq \frac{\delta}{2} + tL^{t-1} \max_{j,k}\left(\varphi\left(\left|P_{kj}(\mathbf{X}_k) - X'_{kj}\right|\right), \varphi\left(\left|Q_{kj}(\mathbf{Y}_k) - Y'_{kj}\right|\right)\right) \\ &\quad < \delta \end{aligned}$$

Consequently, $((\mathbf{P}_k(\mathbf{A}_k), \mathbf{Q}_k(\mathbf{B}_k)))_k \in \Gamma_\infty(\mathbf{X}'_1, \ldots, \mathbf{X}'_\ell \sqcup \mathbf{Y}'_1, \ldots, \mathbf{Y}'_\ell; M, d, \delta)$. Since conjugating by unitaries commutes with the application of the polynomials, it follows that

$$\Gamma_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_k)_{k=1}^{\ell}, (\mathbf{B}_k)_{k=1}^{\ell}; M', d, \delta') \subseteq \Gamma_{\mathrm{orb}}(\mathbf{Z}'_1, \ldots, \mathbf{Z}'_\ell : (\mathbf{P}_k(\mathbf{A}_k))_{k=1}^{\ell}, (\mathbf{Q}_k(\mathbf{B}_k))_{k=1}^{\ell}; M, d, \delta),$$

and so

$$\tilde{\chi}_{\mathrm{orb},\infty}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M', d, \delta') \leq \tilde{\chi}_{\mathrm{orb},\infty}(\mathbf{Z}'_1, \ldots, \mathbf{Z}'_\ell; M, d, \delta).$$

The desired inequality now follows from Proposition 5.4 and Lemma 5.6. □

Like with the microstate bi-free entropy, the orbital bi-free entropy also is upper semi-continuous with respect to distributional limits.

Proposition 5.9. *If* $\mathbf{X}_1^{(l)}, \ldots, \mathbf{X}_\ell^{(l)} \sqcup \mathbf{Y}_1^{(l)}, \ldots, \mathbf{Y}_\ell^{(l)}$ *converges in distribution to* $\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell$ *as in the sense of Proposition 2.11, then*

$$\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \geq \limsup_{l\to\infty} \chi_{\mathrm{orb},R}(\mathbf{Z}_1^{(l)}, \ldots, \mathbf{Z}_\ell^{(l)})$$

for every $R > 0$ *including* $R = \infty$*. Therefore, if there exists a uniform operator norm bound of these operators over* l*, we have*

$$\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \geq \limsup_{l\to\infty} \chi_{\mathrm{orb}}(\mathbf{Z}_1^{(l)}, \ldots, \mathbf{Z}_\ell^{(l)})$$

Proof. As in the proof of Proposition 2.11, our convergence assumption tells us that all moments are converging to the correct values, and so for any $M \in \mathbb{N}$ and $\epsilon > 0$ we have for large enough l that

$$\Gamma_R\left(\mathbf{X}_1^{(l)}, \ldots, \mathbf{X}_\ell^{(l)} \sqcup \mathbf{Y}_1^{(l)}, \ldots, \mathbf{Y}_\ell^{(l)}; M, d, \epsilon\right) \subseteq \Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, 2\epsilon),$$

since the sets involved see only finitely many moments. Hence for any $\mu \in \mathcal{P}\left(\left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^\ell (M_d^{\mathrm{sa}})^{m_k}\right)\right)$, we have that

$$\begin{aligned}\chi_{\mathrm{orb},R}\left(\mathbf{Z}_1^{(l)}, \ldots, \mathbf{Z}_\ell^{(l)}; M, d, \epsilon; \mu\right) &\leq \chi_{\mathrm{orb},R}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, 2\epsilon; \mu\right) \\ &\leq \chi_{\mathrm{orb},R}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, 2\epsilon\right).\end{aligned}$$

By taking the appropriate sups, limsup, and infs, we obtain

$$\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \geq \limsup_{l\to\infty} \chi_{\mathrm{orb},R}(\mathbf{Z}_1^{(l)}, \ldots, \mathbf{Z}_\ell^{(l)})$$

for every $R > 0$ including $R = \infty$. The remaining equation then follows from Lemma 5.6. □

Of greater interest is how the orbital bi-free entropy behaves with respect to bi-free collections. In particular, the following proof uses similar ideas as those used in Lemma 4.6 and Theorem 4.7.

Theorem 5.10. *If* $\mathbf{Z}_1$ *and* $\mathbf{X}_2, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_2, \ldots, \mathbf{Y}_\ell$ *are as described above and are bi-free with respect to* φ*, then*

$$\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = \chi_{\mathrm{orb}}(\mathbf{Z}_1) + \chi_{\mathrm{orb}}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell).$$

Proof. First, suppose that $\mathbf{Z}_1$ does not have finite-dimensional approximants. Then $\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell$ also does not have finite-dimensional approximants, so the definition of the orbital bi-free entropy implies that

$$\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = -\infty = \chi_{\mathrm{orb}}(\mathbf{Z}_1) = \chi_{\mathrm{orb}}(\mathbf{Z}_1) + \chi_{\mathrm{orb}}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell).$$

Hence we may assume that $\mathbf{Z}_1$ has finite-dimensional approximants so $\chi_{\mathrm{orb}}(\mathbf{Z}_1) = 0$ by Proposition 5.7.

Next, suppose that $\chi_{\mathrm{orb}}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell) = -\infty$. Since Proposition 5.7 implies then that $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \leq -\infty$, the equation still holds. Hence we may assume that $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) > -\infty$. Furthermore, as $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) \leq \chi_{\mathrm{orb}}(\mathbf{Z}_1) + \chi_{\mathrm{orb}}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell)$ by Proposition 5.7, it suffices to prove the other inequality with $\chi_{\mathrm{orb}}(\mathbf{Z}_1) = 0$.

Fix $R > \max(\{1\} \cup \{\|X_{i,k}\| \mid 1 \leq i \leq n_k, 1 \leq k \leq \ell\} \cup \{\|Y_{j,k}\| \mid 1 \leq j \leq m_k, 1 \leq k \leq \ell\})$ and fix $M \in \mathbb{N}$ and $\epsilon > 0$. By the same argument as at the start of the proof of Lemma 4.6, there exists an $\epsilon_1 > 0$ such that if

- $(\mathbf{A}_1, \mathbf{B}_1) \in \Gamma_R(\mathbf{X}_1 \sqcup \mathbf{Y}_1; M, d, \epsilon_1)$,
- $((\mathbf{A}_k)_{k=2}^\ell, (\mathbf{B}_k)_{k=2}^\ell) \in \Gamma_R(\mathbf{X}_2, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_2, \ldots, \mathbf{Y}_\ell; M, d, \epsilon_1)$, and
- $(\mathbf{A}_1, \mathbf{B}_1)$ and $((\mathbf{A}_k)_{k=2}^\ell, (\mathbf{B}_k)_{k=2}^\ell)$ are (M, ϵ_1)-free,

then $((\mathbf{A}_k)_{k=1}^\ell, (\mathbf{B}_k)_{k=1}^\ell) \in \Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon)$.

Since by Lemma 5.6 we have that $\chi_{\mathrm{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell) = \chi_{\mathrm{orb}}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell) > -\infty$, Proposition 5.4 implies there exists an increasing sequence $(d_l)_{l\geq 1}$ such that $\tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon_1) > -\infty$ and

$$\limsup_{d\to\infty} \frac{1}{d^2} \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell; M, d, \epsilon_1) = \lim_{l\to\infty} \frac{1}{d_l^2} \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon_1).$$

For each $l \in \mathbb{N}$, choose $((\mathbf{A}_{k,l})_{k=2}^\ell, (\mathbf{B}_{k,l})_{k=2}^\ell) \in \left(\prod_{k=2}^\ell (M_{d_l}^{\mathrm{sa}})_R^{n_k}\right) \times \left(\prod_{k=2}^\ell (M_{d_l}^{\mathrm{sa}})_R^{m_k}\right)$ such that

$$-\infty < \tilde{\chi}_{\mathrm{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon_1) - 1 \leq \log\left(\gamma_{d_l}^{\otimes \ell-1}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=2}^\ell, (\mathbf{B}_{k,l})_{k=2}^\ell; M, d_l, \epsilon_1)\right)\right).$$

Note this implies $\gamma_{d_l}^{\otimes \ell-1}\left(\Gamma_{\mathrm{orb}}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=2}^\ell, (\mathbf{B}_{k,l})_{k=2}^\ell; M, d_l, \epsilon_1)\right) > 0$. Furthermore, as $\mathbf{Z}_1$ has finite-dimensional approximants, for l sufficiently large we may choose a fixed $(\mathbf{A}_{1,l}, \mathbf{B}_{1,l}) \in \Gamma_R(\mathbf{X}_1 \sqcup \mathbf{Y}_1; M, d_l, \epsilon_1)$.

To simplify notation, let

$$\Psi(M, d_l, \epsilon) = \Gamma_{\text{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; M, d_l, \epsilon)$$

$$\Theta(M, d_l, \epsilon_1) = \Gamma_{\text{orb}}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell; (\mathbf{A}_{k,l})_{k=2}^{\ell}, (\mathbf{B}_{k,l})_{k=2}^{\ell}; M, d_l, \epsilon_1)$$

$$\Omega(M, d_l, \epsilon_1) = \{(U_k)_{k=1}^{\ell} \in U(d_l)^{\ell} \mid (U_1^*\mathbf{A}_{1,l}U_1, U_1^*\mathbf{B}_{1,l}U_1), ((U_k^*\mathbf{A}_{k,l})_{k=2}^{\ell}U_k, U_k^*(\mathbf{B}_{k,l})_{k=2}^{\ell}U_k) \text{ are } (M, \epsilon_1)\text{-free}\},$$

and let μ_{d_l} be the probability measure obtained by restricting and renormalizing $\gamma^{\otimes \ell-1}$ to $\Theta(M, d_l, \epsilon_1)$.

Notice that by the choice of ϵ_1 and the conditions defining the sets in question, we have

$$(U(d_l) \times \Theta(M, d_l, \epsilon_1)) \cap \Omega(M, d_l, \epsilon_1) \subseteq \Psi(M, d_l, \epsilon).$$

By Lemma 4.4 (with $p = 1$) there exists a $D_0 \in \mathbb{N}$ such that

$$\gamma_{d_l}(\{U_1 \in U(d_l) \mid (U_k)_{k=1}^{\ell} \in \Omega(M, d_l, \epsilon_1)\}) > \frac{1}{2}$$

for every $d_l \geq D_0$ and every $(U_k)_{k=2}^{\ell} \in U(d_l)^{\ell-1}$. Hence for all $d_l \geq D_0$ we have that

$$\begin{aligned}
\frac{\gamma_{d_l}^{\otimes \ell}(\Psi(M, d_l, \epsilon))}{\gamma_{d_l}^{\otimes \ell-1}(\Theta(M, d_l, \epsilon_1))} &\geq (\gamma_{d_l} \otimes \mu_{d_l})(\Psi(M, d_l, \epsilon)) \\
&\geq (\gamma_{d_l} \otimes \mu_{d_l})((U(d_l) \times \Theta(M, d_l, \epsilon_1)) \cap \Omega(M, d_\ell, \epsilon_1)) \\
&= \int_{\Theta(M, d_l, \epsilon_1)} \gamma_{d_l}(\{U_1 \in U(d_l) \mid (U_k)_{k=1}^{\ell} \in \Omega(M, d_\ell, \epsilon_1)\})\, d\mu_{d_l}((U_k)_{k=2}^{\ell}) > \frac{1}{2}
\end{aligned}$$

by Fubini's Theorem. Hence

$$\begin{aligned}
\tilde{\chi}_{\text{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon_1) &\leq 1 + \log\left(\gamma_{d_l}^{\otimes \ell-1}(\Theta(M, d_l, \epsilon_1))\right) \\
&< 1 + \log(2) + \log\left(\gamma_{d_l}^{\otimes \ell}(\Psi(M, d_l, \epsilon))\right) \\
&\leq 1 + \log(2) + \tilde{\chi}_{\text{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon)
\end{aligned}$$

whenever $d_l \geq D_0$. Thus as $\tilde{\chi}_{\text{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell; M, d, \epsilon_1)$ from Proposition 5.4 decreases as M increases and as ϵ_1 decreases, we have

$$\begin{aligned}
\chi_{\text{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell) &\leq \limsup_{d \to \infty} \frac{1}{d^2} \tilde{\chi}_{\text{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell; M, d, \epsilon_1) \\
&= \lim_{l \to \infty} \frac{1}{d_l^2} \tilde{\chi}_{\text{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon_1) \\
&\leq \limsup_{l \to \infty} \frac{1}{d_l^2} (1 + \log(2) + \tilde{\chi}_{\text{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d_l, \epsilon)) \\
&\leq \limsup_{d \to \infty} \frac{1}{d^2} \tilde{\chi}_{\text{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon).
\end{aligned}$$

Hence Proposition 5.4 implies that $\chi_{\text{orb},R}(\mathbf{Z}_2, \ldots, \mathbf{Z}_\ell) \leq \chi_{\text{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell)$. □

We are finally able to compute the orbital bi-free entropy of certain collections. In particular, in the following case the orbital bi-free entropy is maximized.

Corollary 5.11. *If $\mathbf{Z}_1$, $\mathbf{Z}_2$, ..., $\mathbf{Z}_\ell$ are bi-free with respect to φ and individually have finite-dimensional approximants, then $\chi_{\text{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = 0$.*

Proof. This immediately follows from Theorem 5.10 and part (2) of Proposition 5.7. □

To finish off this section, we note an improvement to the subadditivity result for microstate bi-free entropy. In particular, the following gives us a smaller upper bound for the joint microstate bi-free entropy in terms of the individual microstate bi-free entropies.

Theorem 5.12. *With the notation used throughout this section, we have that*

$$\chi(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell) \leq \chi_{\text{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) + \sum_{k=1}^{\ell} \chi(\mathbf{Z}_k)$$

Proof. First, if $\chi(\mathbf{Z}_k) = -\infty$ for some k, then the result follows from Proposition 2.5. Hence we may assume that $\chi(\mathbf{Z}_k) > -\infty$ for all k and thus $\mathbf{Z}_k$ has finite-dimensional approximants for all k by Remark 4.2.

Fix $R > \max(\{1\} \cup \{\|X_{i,k}\| \mid 1 \leq i \leq n_k, 1 \leq k \leq \ell\} \cup \{\|Y_{j,k}\| \mid 1 \leq j \leq m_k, 1 \leq k \leq \ell\})$, $M \in \mathbb{N}$, and $\epsilon > 0$. As $\mathbf{Z}_k$ has finite-dimensional approximants for all k, there exists an $D_0 \in \mathbb{N}$ such that $\Gamma_R(\mathbf{Z}_k; M, d, \epsilon) \neq \emptyset$ for all $d \geq D_0$ and $1 \leq k \leq \ell$.

Define $\sigma : \prod_{k=1}^{\ell} ((\mathcal{M}_d^{\mathrm{sa}})^{n_k} \times (\mathcal{M}_d^{\mathrm{sa}})^{m_k}) \to \left(\prod_{k=1}^{\ell} (M_d^{\mathrm{sa}})^{n_k}\right) \times \left(\prod_{k=1}^{\ell} (M_d^{\mathrm{sa}})^{m_k}\right)$ by

$$\sigma\left(((\mathbf{A}_k, \mathbf{B}_k))_{k=1}^{\ell}\right) = ((\mathbf{A}_k)_{k=1}^{\ell}, (\mathbf{B}_k)_{k=1}^{\ell}).$$

Since each $\Gamma_R(\mathbf{Z}_k; M, d, \epsilon)$ is non-empty and open, we know the Lebesgue measure of $\Gamma_R(\mathbf{Z}_k; M, d, \epsilon)$ is non-zero. Therefore, as σ preserves the Lebesgue measure $\lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}$ under the natural isomorphism with the domain and co-domain, we have that $\sigma\left(\prod_{k=1}^{\ell} \Gamma_R(\mathbf{Z}_k; M, d, \epsilon)\right)$ has positive Lebesgue measure. Let $\nu_R(M, d, \epsilon)$ denote the probability measure obtained by renormalizing $\lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}$ after restricting to $\sigma\left(\prod_{k=1}^{\ell} \Gamma_R(\mathbf{Z}_k; M, d, \epsilon)\right)$ when $d \geq D_0$; that is

$$\nu_R(M, d, \epsilon) = \frac{1}{\prod_{k=1}^{\ell} \lambda_d^{\otimes n_k+m_k}(\Gamma_R(\mathbf{Z}_k; M, d, \epsilon))} \, \lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}\Big|_{\sigma\left(\prod_{k=1}^{\ell} \Gamma_R(\mathbf{Z}_k; M, d, \epsilon)\right)}.$$

Since

$$\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon) \subseteq \sigma\left(\prod_{k=1}^{\ell} \Gamma_R(\mathbf{Z}_k; M, d, \epsilon)\right),$$

we have by Definition 5.1 that for all $d \geq D_0$

$$\begin{aligned}
\chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) &\geq \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon; \nu_R(M, d, \epsilon)) \\
&= \log\left(\left(\gamma_d^{\otimes \ell} \otimes \lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}\right)\left(\Phi_d^{-1}(\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon))\right)\right) \\
&\quad - \sum_{k=1}^{\ell} \log\left(\lambda_d^{\otimes n_k+m_k}\left(\Gamma_R(\mathbf{Z}_k; M, d, \epsilon)\right)\right).
\end{aligned}$$

Thus

$$\begin{aligned}
&\log\left(\left(\gamma_d^{\otimes \ell} \otimes \lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}\right)\left(\Phi_d^{-1}(\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon))\right)\right) \\
&\quad \leq \chi_{\mathrm{orb},R}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; M, d, \epsilon) + \sum_{k=1}^{\ell} \log\left(\lambda_d^{\otimes n_k+m_k}\left(\Gamma_R(\mathbf{Z}_k; M, d, \epsilon)\right)\right)
\end{aligned}$$

for sufficiently large d for every $M \in \mathbb{N}$ and $\epsilon > 0$.

For a fixed $(U_k)_{k=1}^{\ell} \in U(d)^{\ell}$, notice that the corresponding section of $\Phi_d^{-1}(\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon))$, namely

$$\{((\mathbf{A}_k)_{k=1}^{\ell}, (\mathbf{B}_k)_{k=1}^{\ell}) \mid \Phi_d((U_k)_{k=1}^{\ell}, (\mathbf{A}_k)_{k=1}^{\ell}, (\mathbf{B}_k)_{k=1}^{\ell}) \in \Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon)\},$$

is exactly

$$\Phi_d\left((U_k^*)_{k=1}^{\ell}, \Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon)\right).$$

Hence Fubini's theorem and the fact that Lebesgue measure is unitarily-invariant together imply that

$$\begin{aligned}
&\left(\gamma_d^{\otimes \ell} \otimes \lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}\right)\left(\Phi_d^{-1}(\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon)\right) \\
&= \int_{U(d)^{\ell}} \lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}\left(\Phi_d\left((U_k^*)_{k=1}^{\ell}, \Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon)\right)\right) d\gamma_d^{\otimes \ell} \\
&= \int_{U(d)^{\ell}} \lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}\left(\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon)\right) d\gamma_d^{\otimes \ell} \\
&= \lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}\left(\Gamma_R(\mathbf{X}_1, \ldots, \mathbf{X}_\ell \sqcup \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell; M, d, \epsilon)\right).
\end{aligned}$$

Hence

$$\log\left(\lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}\left(\Gamma_R(\mathbf{X}_1,\ldots,\mathbf{X}_\ell\sqcup\mathbf{Y}_1,\ldots,\mathbf{Y}_\ell;M,d,\epsilon)\right)\right)$$
$$\leq \chi_{\mathrm{orb},R}(\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell;M,d,\epsilon)+\sum_{k=1}^{\ell}\log\left(\lambda_d^{\otimes n_k+m_k}\left(\Gamma_R(\mathbf{Z}_k;M,d,\epsilon)\right)\right)$$

so

$$\frac{1}{d^2}\log\left(\lambda_d^{\otimes n_1+\cdots+n_\ell+m_1+\cdots+m_\ell}\left(\Gamma_R(\mathbf{X}_1,\ldots,\mathbf{X}_\ell\sqcup\mathbf{Y}_1,\ldots,\mathbf{Y}_\ell;M,d,\epsilon)\right)\right)+\frac{1}{2}\left(\sum_{k=1}^{\ell}n_k+m_k\right)\log(d)$$
$$\leq \frac{1}{d^2}\chi_{\mathrm{orb},R}(\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell;M,d,\epsilon)+\sum_{k=1}^{\ell}\frac{1}{d^2}\log\left(\lambda_d^{\otimes n_k+m_k}\left(\Gamma_R(\mathbf{Z}_k;M,d,\epsilon)\right)\right)+\frac{1}{2}(n_k+m_k)\log(d).$$

Now, taking the appropriate limits, the result follows. ☐

We note that inequality in Theorem 5.12 need not be an equality. Indeed [30] shows that the inequality can be strict in the free setting.

6. A Characterization of Bi-Freeness

The goal of this section is to develop another characterization of bi-freeness for specific tracially bi-partite systems. To be specific, using the same notation as Section 5, the main goal of this section is to prove the following, a bi-free version of [14, Theorem 3.1].

Theorem 6.1. *Let $\mathbf{Z}_1$, $\mathbf{Z}_2$, …, $\mathbf{Z}_\ell$ be such that*

$$\left(\bigcup_{k=1}^{\ell}\{X_{k,i}\}_{i=1}^{n_k},\bigcup_{k=1}^{\ell}\{Y_{k,j}\}_{j=1}^{m_k}\right)$$

is a tracially bi-partite system. Then $\mathbf{Z}_1$, $\mathbf{Z}_2$, …, $\mathbf{Z}_\ell$ are bi-free and individually have finite-dimensional approximants if and only if $\chi_{\mathrm{orb}}(\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell)=0$.

Of course, the only if direction immediately follows from Corollary 5.11, so the work of this section is to establish the other direction. Moreover, since finite orbital bi-free entropy immediately implies the existence of microstates, our real task is to deduce bi-free independence. Before we get to that, we point out the following corollary, which follows immediately by combining Theorem 6.1 with the results of Section 5.

Corollary 6.2. *Let $\mathbf{Z}_1$, $\mathbf{Z}_2$, …, $\mathbf{Z}_\ell$ be such that $\chi(\mathbf{Z}_k)>-\infty$ for all $1\leq k\leq\ell$. Suppose further that*

$$\left(\bigcup_{k=1}^{\ell}\{X_{k,i}\}_{i=1}^{n_k},\bigcup_{k=1}^{\ell}\{Y_{k,j}\}_{j=1}^{m_k}\right)$$

is a tracially bi-partite system. If

$$\chi(\mathbf{X}_1,\ldots,\mathbf{X}_\ell\sqcup\mathbf{Y}_1,\ldots,\mathbf{Y}_\ell)=\sum_{k=1}^{\ell}\chi(\mathbf{Z}_k),$$

then $\mathbf{Z}_1$, $\mathbf{Z}_2$, …, $\mathbf{Z}_\ell$ are bi-free.

Proof. As $\chi(\mathbf{Z}_k)>-\infty$ for all $1\leq k\leq\ell$, we know from Remark 4.2 that $\mathbf{Z}_1$, $\mathbf{Z}_2$, …, $\mathbf{Z}_\ell$ individually have finite-dimensional approximants. Furthermore, the assumption of additivity of the microstate bi-free entropy implies that

$$\chi(\mathbf{X}_1,\ldots,\mathbf{X}_\ell\sqcup\mathbf{Y}_1,\ldots,\mathbf{Y}_\ell)>-\infty.$$

By Theorem 5.12 along with the assumption, we know that

$$\chi(\mathbf{X}_1,\ldots,\mathbf{X}_\ell\sqcup\mathbf{Y}_1,\ldots,\mathbf{Y}_\ell)\leq\chi_{\mathrm{orb}}(\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell)+\sum_{k=1}^{\ell}\chi(\mathbf{Z}_k)$$
$$=\chi_{\mathrm{orb}}(\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell)+\chi(\mathbf{X}_1,\ldots,\mathbf{X}_\ell\sqcup\mathbf{Y}_1,\ldots,\mathbf{Y}_\ell).$$

Thus $\chi_{\mathrm{orb}}(\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell)\geq 0$. However, as $\chi_{\mathrm{orb}}(\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell)\leq 0$ by definition, we obtain that $\chi_{\mathrm{orb}}(\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell)=0$. Hence Theorem 6.1 implies that $\mathbf{Z}_1$, $\mathbf{Z}_2$, …, $\mathbf{Z}_\ell$ are bi-free. ☐

The proof of Theorem 6.1 will follow immediately from Proposition 6.10, which is a Talagrand-like inequality. Thus we devote the rest of the section to developing the necessary framework to state and prove this proposition. We first need an analogue of the free Wasserstein metric from [5] for the following objects.

Definition 6.3. A quadruple $(\mathcal{A}, \mathcal{L}, \mathcal{R}, \varphi)$ is said to be a *left-right, tracially bi-partite, C*-non-commutative probability space* if $(\mathcal{A}, \varphi)$ is a C*-non-commutative probability space, $\mathcal{L}$ and $\mathcal{R}$ are unital C*-subalgebras of $\mathcal{A}$ that commute with one another, and φ is tracial when restricted to $\mathcal{L}$ and when restricted to $\mathcal{R}$.

By saying a tracially bi-partite system $\left(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m\right)$ is in a left-right, tracially bi-partite, C*-non-commutative probability space $(\mathcal{A}, \mathcal{L}, \mathcal{R}, \varphi)$, we mean $\{X_i\}_{i=1}^n \subseteq \mathcal{L}$ and $\{Y_j\}_{j=1}^m \subseteq \mathcal{R}$. Note any tracially bi-partite system can be realized in a left-right, tracially bi-partite, C*-non-commutative probability space.

Definition 6.4. Let $\left(\{X_{i,1}\}_{i=1}^n, \{Y_{j,1}\}_{j=1}^m\right)$ and $\left(\{X_{i,2}\}_{i=1}^n, \{Y_{j,2}\}_{j=1}^m\right)$ be tracially bi-partite systems in left-right, tracially bi-partite, C*-non-commutative probability spaces $(\mathcal{A}_1, \mathcal{L}_1, \mathcal{R}_1, \varphi_1)$ and $(\mathcal{A}_2, \mathcal{L}_2, \mathcal{R}_2, \varphi_2)$ respectively. We define

$$W_2\left(\left(\{X_{i,1}\}_{i=1}^n, \{Y_{j,1}\}_{j=1}^m\right), \left(\{X_{i,2}\}_{i=1}^n, \{Y_{j,2}\}_{j=1}^m\right)\right)$$

to be infimum of

$$\left(\sum_{i=1}^n \left\|X'_{i,1} - X'_{i,2}\right\|_2^2 + \sum_{j=1}^m \left\|Y'_{j,1} - Y'_{j,2}\right\|_2^2\right)^{\frac{1}{2}}$$

over all tracially bi-partite systems $\left(\{X'_{i,1}\}_{i=1}^n, \{Y'_{j,1}\}_{j=1}^m\right)$ and $\left(\{X'_{i,2}\}_{i=1}^n, \{Y'_{j,2}\}_{j=1}^m\right)$ in a left-right, tracially bi-partite, C*-non-commutative probability spaces $(\mathcal{A}, \mathcal{L}, \mathcal{R}, \varphi)$ such that $\left(\{X_{i,k}\}_{i=1}^n, \{Y_{j,k}\}_{j=1}^m\right)$ and $\left(\{X'_{i,k}\}_{i=1}^n, \{Y'_{j,k}\}_{j=1}^m\right)$ have the same $*$-distributions and individual operator norms for $k = 1, 2$, where $\|\cdot\|_2$ denotes the 2-seminorm with respect to φ (note we may only have a seminorm as we are not restricting ourselves to faithful states).

Remark 6.5. It is natural and necessary to ask whether one can find a left-right, tracially bi-partite, C*-non-commutative probability spaces $(\mathcal{A}, \mathcal{L}, \mathcal{R}, \varphi)$ as described in Definition 6.4 so that the infimum is over a non-empty set. This is indeed the case by considering reduced free products. If one takes the reduced free product Hilbert space $(\mathcal{A}_1, \varphi_1) * (\mathcal{A}_2, \varphi_2)$, we can let $\mathcal{L}_1$ and $\mathcal{L}_2$ act via the left regular representation on $\mathcal{A}_1$ and $\mathcal{A}_2$ respectively, and let $\mathcal{R}_1$ and $\mathcal{R}_2$ act via the right regular representation on $\mathcal{A}_1$ and $\mathcal{A}_2$ respectively. These representations are φ-preserving $*$-homomorphism and thus preserve distributions and the operator norms. Furthermore, the C*-algebra $\mathcal{L}$ generated by the images of $\mathcal{L}_1$ and $\mathcal{L}_2$ commutes with the C*-algebra $\mathcal{R}$ generated by the images of $\mathcal{R}_1$ and $\mathcal{R}_2$. Finally, the reduced free product state is tracial on $\mathcal{L}$ and is tracial on $\mathcal{R}$ by properties of the reduced free product (i.e. the free case). Of course, this is one reason why the states in a left-right, tracially bi-partite, C*-non-commutative probability space need not be faithful as the work of [23] shows we would be greatly restricting the systems we can study in that the bi-free product of faithful states need not be faithful.

Using Definition 6.4, we can consider a similar definition for 'nice' states.

Definition 6.6. Let $\mathcal{A}$ be a C*-algebra and let $\mathcal{L}$ and $\mathcal{R}$ be unital subalgebras of $\mathcal{A}$ that commute with one another. Suppose that $\mathcal{A}$ is generated by $\mathcal{L}$ and $\mathcal{R}$, which in turn are generated by prescribed sets $\{X_i\}_{i=1}^n$ and $\{Y_j\}_{j=1}^m$ respectively.

Let $\mathcal{CS}(\mathcal{A}, \mathcal{L}, \mathcal{R})$ denote the set of all states (positive unital linear functionals of norm one) that are tracial when restricted to $\mathcal{L}$ and are tracial when restricted to $\mathcal{R}$. We define

$$W_2(\varphi_1, \varphi_2) = W_2\left(\left(\{X_{i,1}\}_{i=1}^n, \{Y_{j,1}\}_{j=1}^m\right), \left(\{X_{i,2}\}_{i=1}^n, \{Y_{j,2}\}_{j=1}^m\right)\right)$$

where $\left(\{X_{i,k}\}_{i=1}^n, \{Y_{j,k}\}_{j=1}^m\right)$ denote $\left(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m\right)$ in $(\mathcal{A}, \mathcal{L}, \mathcal{R}, \varphi_k)$ for $k = 1, 2$. Note that W_2 depends on the choice of generating set, but we leave this implicit.

Given a state $\varphi \in \mathcal{CS}(\mathcal{A}, \mathcal{L}, \mathcal{R})$, let us denote by $\varphi^{\text{bi-free}}_{\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell} \in \mathcal{CS}(\mathcal{A}, \mathcal{L}, \mathcal{R})$ the state whose marginals on $\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell$ agree with those of φ, so that the families $\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell$ are bi-free.

Like the free Wasserstein metric from [5], the function W_2 has some nice properties.

Proposition 6.7. *The bi-free analogue of the Wasserstein metric is a semimetric on the collection of tracially bi-partite systems with equal numbers of left variables and equal numbers of right variables, and is a semimetric $\mathcal{CS}(\mathcal{A}, \mathcal{L}, \mathcal{R})$.*

Proof. The reasons that

$$W_2\left(\left(\{X_{i,1}\}_{i=1}^n, \{Y_{j,1}\}_{j=1}^m\right), \left(\{X_{i,2}\}_{i=1}^n, \{Y_{j,2}\}_{j=1}^m\right)\right) = 0$$

implies $\left(\{X_{i,1}\}_{i=1}^n, \{Y_{j,1}\}_{j=1}^m\right)$ and $\left(\{X_{i,2}\}_{i=1}^n, \{Y_{j,2}\}_{j=1}^m\right)$ have the same distribution and the reasons that $W_2(\varphi_1, \varphi_2) = 0$ implies $\varphi_1 = \varphi_2$ both follow from the facts that the operator norms of the representations in Definition 6.4 are bounded, the left and right algebras commute with each other, all linear functionals considered are states, the traciality of the states on the individual left and right algebras, and the definition of W_2. Indeed, for an example computation, with terms as in Definition 6.4 (where all operator are self-adjoint), notice that

$$\begin{aligned}
\left|\varphi(X'_{1,1}X'_{2,1}Y'_{1,1}) - \varphi(X'_{1,2}X'_{2,1}Y'_{1,1})\right| &\leq \varphi(1)\varphi\left(Y'_{1,1}X'_{2,1}(X'_{1,1} - X'_{1,2})(X'_{1,1} - X'_{1,2})X'_{2,1}Y'_{1,1}\right)^{\frac{1}{2}} \\
&= \varphi\left(X'_{2,1}(X'_{1,1} - X'_{1,2})Y'_{1,1}Y'_{1,1}(X'_{1,1} - X'_{1,2})X'_{2,1}\right)^{\frac{1}{2}} \\
&\leq \left\|Y'_{1,1}\right\| \varphi\left(X'_{2,1}(X'_{1,1} - X'_{1,2})(X'_{1,1} - X'_{1,2})X'_{2,1}\right)^{\frac{1}{2}} \\
&= \left\|Y'_{1,1}\right\| \varphi\left((X'_{1,1} - X'_{1,2})X'_{2,1}X'_{2,1}(X'_{1,1} - X'_{1,2})\right)^{\frac{1}{2}} \\
&\leq \left\|Y'_{1,1}\right\| \left\|X'_{2,1}\right\| \left\|X'_{1,1} - X'_{1,2}\right\|_2
\end{aligned}$$

where the first equality is left-right commutation, and the second equality is traciality on the left. Using telescoping sums along with the bounds on the operator norms, the fact that W_2 is 0 and thus we can find $\left(\{X'_{i,1}\}_{i=1}^n, \{Y'_{j,1}\}_{j=1}^m\right)$ and $\left(\{X'_{i,2}\}_{i=1}^n, \{Y'_{j,2}\}_{j=1}^m\right)$ as in Definition 6.4 with arbitrarily small 2-seminorms, we can show the difference in the distribution of any monomial is as small as we desire and thus equal. The remaining properties of a semimetric are trivial to verify. □

Remark 6.8. Unfortunately we do not know whether or not W_2 is a metric. The problem with trying to repeat the proof of [5] is that there is no current bi-free product that enables one to amalgamate the left operators over one subalgebra and the right operators over another non-isomorphic subalgebra; that is, [7] amalgamates over a copy of an algebra contained in both the left and right operators. This creates a problem with trying to use the bi-free product construction from Remark 6.5 to take two different pairs and construct a left-right, tracially bi-partite, C*-non-commutative probability space containing all three in a way that the both pairs are identified in the appropriate way. In particular, positivity becomes an issue.

It would also be nice to generalize the above to non-bi-partite systems. However, as we are dealing with seminorms, it does appear difficult to even get a semimetric considering the current proof of Proposition 6.7.

Fortunately for the discussions in this paper, Proposition 6.7 along with the following result are enough.

Proposition 6.9. *Given sequences $(\varphi_{1,k})_{k\geq 1}$ and $(\varphi_{2,k})_{k\geq 1}$ in $\mathcal{CS}(\mathcal{A}, \mathcal{L}, \mathcal{R})$ that converge weak* to φ_1 and φ_2 in $\mathcal{CS}(\mathcal{A}, \mathcal{L}, \mathcal{R})$ respectively, we have*

$$\liminf_{k\to\infty} W_2(\varphi_{1,k}, \varphi_{2,k}) \geq W_2(\varphi_1, \varphi_2).$$

Similarly, suppose $\left(\left(\{X_{i,1,k}\}_{i=1}^n, \{Y_{j,1,k}\}_{j=1}^m\right)\right)_{k\geq 1}$ and $\left(\left(\{X_{i,2,k}\}_{i=1}^n, \{Y_{j,2,k}\}_{j=1}^m\right)\right)_{k\geq 1}$ are tracially bi-partite systems in left-right, tracially bi-partite, C*-non-commutative probability spaces $(\mathcal{A}_1, \mathcal{L}_1, \mathcal{R}_1, \varphi_1)$ and $(\mathcal{A}_2, \mathcal{L}_2, \mathcal{R}_2, \varphi_2)$ respectively that converge in distributions to $\left(\{X_{i,1}\}_{i=1}^n, \{Y_{j,1}\}_{j=1}^m\right)$ and $\left(\{X_{i,2}\}_{i=1}^n, \{Y_{j,2}\}_{j=1}^m\right)$ in $(\mathcal{A}_1, \mathcal{L}_1, \mathcal{R}_1, \varphi_1)$ and $(\mathcal{A}_2, \mathcal{L}_2, \mathcal{R}_2, \varphi_2)$ respectively and for which there is a uniform bound on all operator norms of all operators. Then

$$\begin{aligned}
\liminf_{k\to\infty}\ & W_2\left(\left(\{X_{i,1,k}\}_{i=1}^n, \{Y_{j,1,k}\}_{j=1}^m\right), \left(\{X_{i,2,k}\}_{i=1}^n, \{Y_{j,2,k}\}_{j=1}^m\right)\right) \\
&\geq W_2\left(\left(\{X_{i,1}\}_{i=1}^n, \{Y_{j,1}\}_{j=1}^m\right), \left(\{X_{i,2}\}_{i=1}^n, \{Y_{j,2}\}_{j=1}^m\right)\right).
\end{aligned}$$

Proof. The result follows by considering Definitions 6.6 and 6.4 taking weak*-limits and a compactness argument, much like the free case in [5, Proposition 1.4]. □

With the analogue of the Wasserstein metric, we are now able to state our Talagrand-like inequality. This is an adaptation to the (multivariate) bi-free setting of [14, Proposition 3.5], and our proof draws inspiration from theirs (and their proof, in turn, draws inspiration from [16]). The proof will require several lemmas, which are analogues of [14, Lemmas 3.2, 3.3, 3.4].

Proposition 6.10. *Let* $\mathbf{Z}_1$, $\mathbf{Z}_2$, ..., $\mathbf{Z}_\ell$ *be such that*

$$\left(\bigcup_{k=1}^{\ell}\{X_{k,i}\}_{i=1}^{n_k}, \bigcup_{k=1}^{\ell}\{Y_{k,j}\}_{j=1}^{m_k}\right)$$

is a tracially bi-partite system, and suppose that $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) > -\infty$. *Then*

$$W_2(\varphi_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell}, \varphi^{\text{bi-free}}_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell}) \leq 4R_0\sqrt{n+m}\sqrt{-\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell)}$$

where $n = \max_{1\leq k\leq \ell} n_k$, $m = \max_{1\leq k\leq\ell} m_k$, *and*

$$R_0 = \max(\{\|X_{i,k}\| \mid 1 \leq i \leq n_k, 1 \leq k \leq \ell\} \cup \{\|Y_{j,k}\| \mid 1 \leq j \leq m_k, 1 \leq k \leq \ell\}).$$

Proof. Let $R > R_0$. By Proposition 5.4 (and Lemma lem:orbital-R-doesnt-matter), we can chose an increasing sequence $(d_l)_{l\geq 1}$ of natural numbers such that

$$\tilde{\chi}_{\mathrm{orb},R}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; l, d_l, \frac{1}{l}\right) > -\infty$$

for all $l \in \mathbb{N}$ and

$$\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = \lim_{l\to\infty}\frac{1}{d_l^2}\tilde{\chi}_{\mathrm{orb},R}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; l, d_l, \frac{1}{l}\right).$$

For each $l \in \mathbb{N}$, choose $((\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}) \in \left(\prod_{k=1}^{\ell}(M_{d_l}^{\mathrm{sa}})_R^{n_k}\right) \times \left(\prod_{k=1}^{\ell}(M_{d_l}^{\mathrm{sa}})_R^{m_k}\right)$ such that

$$-\infty < \tilde{\chi}_{\mathrm{orb},R}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; l, d_l, \frac{1}{l}\right) - 1 \leq \log\left(\gamma_{d_l}^{\otimes\ell}\left(\Gamma_{\mathrm{orb}}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; l, d_l, \frac{1}{l}\right)\right)\right).$$

Note this implies $\gamma_{d_l}^{\otimes\ell}\left(\Gamma_{\mathrm{orb}}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; l, d_l, \frac{1}{l}\right)\right) > 0$.

Let $SU(d)$ denote the special unitary group of $\mathcal{M}_d$, let $\mathbb{T}_d := U(d) \cap \mathbb{C}I_d = \mathbb{T}I_d$ denote the set of scalar unitaries, and let $\gamma_{d,s}$ denote the Haar measure on $SU(d)$. We want to work with $SU(d)$ instead of $U(d)$ here for technical reasons. Indeed this is possible as we note that if

$$(U_k)_{k=1}^{\ell} \in \Gamma_{\mathrm{orb}}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; l, d_l, \frac{1}{l}\right)$$

then

$$(V_kU_k)_{k=1}^{\ell} \in \Gamma_{\mathrm{orb}}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; l, d_l, \frac{1}{l}\right)$$

for all $(V_k)_{k=1}^{\ell} \in \mathbb{T}_d^{\ell}$. Hence it immediately follows that if

$$\Gamma_l = SU(d) \cap \Gamma_{\mathrm{orb}}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; l, d_l, \frac{1}{l}\right)$$

then

$$\gamma_{d_l}^{\otimes\ell}\left(\Gamma_{\mathrm{orb}}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; l, d_l, \frac{1}{l}\right)\right) = \gamma_{d_l,s}^{\otimes\ell}(\Gamma_l)$$

(so $\tilde{\chi}_{\mathrm{orb},R}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell; l, d_l, \frac{1}{l}\right) \leq 1 + \log\left(\gamma_{d_l,s}^{\otimes\ell}(\Gamma_l)\right)$).

Let $C[-R, R]$ denote the C*-algebra of continuous functions on $[-R, R]$ and let

$$\mathcal{B}_R = \left(*_{k=1}^{\ell}C[-R, R]^{*n_k}\right) \otimes_{\max} \left(*_{k=1}^{\ell}C[-R, R]^{*m_k}\right)$$

where $*$ denotes the universal free product of C*-algebras. Thus, by properties of the universal free product C*-algebra and the maximal tensor product, there exists a homomorphism $\pi : \mathcal{B}_R \to \mathcal{A}$ such that $\pi(x_{k,i}) = X_{k,i}$ and $\pi(y_{k,j}) = Y_{k,j}$ where $x_{k,i}$ is the identity function on $C[-R, R]$ in the k^{th} term of $*_{k=1}^{\ell}C[-R, R]^{*n_k} \subseteq \mathcal{B}_R$ and the i^{th} term of $C[-R, R]^{*n_k}$, and $y_{k,j}$ is the identity function on $C[-R, R]$ in the k^{th} term of $*_{k=1}^{\ell}C[-R, R]^{*m_k}$ and the j^{th} term of $C[-R, R]^{*m_k}$. Consequently, if $\varphi_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell} = \varphi \circ \pi$, $\mathcal{L} = \left(*_{k=1}^{\ell}C[-R, R]^{*n_k}\right) \otimes 1$, and $\mathcal{R} = 1 \otimes \left(*_{k=1}^{\ell}C[-R, R]^{*m_k}\right)$, then $\mathcal{L}$ and $\mathcal{R}$ are C*-subalgebras of $\mathcal{B}_R$ that commute with each other and $\varphi_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell} \in \mathcal{CS}(\mathcal{B}_R, \mathcal{L}, \mathcal{R})$.

For a fixed l, for each probability measure μ on $SU(d_l)^{\ell}$ we will define $\hat{\mu} \in \mathcal{CS}(\mathcal{B}_R, \mathcal{L}, \mathcal{R})$ as follows. For each $(U_k)_{k=1}^{\ell} \in SU(d)^{\ell}$ note there exists a *-homomorphism $\pi_{(U_k)_{k=1}^{\ell}}$ from $\mathcal{B}_R$ to $\mathcal{B}(\mathcal{M}_{d_l})$ that sends $x_{k,i}$ to left multiplication by $U_k^* A_{k,i,l} U_k$ and sends $y_{k,j}$ to right multiplication by $U_k^* B_{k,j,l} U_k$. We then desire to define

$$\hat{\mu}(Z) = \int_{SU(d_l)^{\otimes \ell}} \tau_{d_\ell}(\pi_{(U_k)_{k=1}^{\ell}}(Z) I_{d_l}) \, d\mu.$$

The fact that $\hat{\mu} \in \mathcal{CS}(\mathcal{B}_R, \mathcal{L}, \mathcal{R})$ follows as $\pi_{(U_k)_{k=1}^{\ell}}$ is a representation and τ_{d_ℓ} is a trace.

With the above in hand, we need several technical lemmas on the weak*-convergence of certain elements of $\mathcal{CS}(\mathcal{B}_R, \mathcal{L}, \mathcal{R})$.

Lemma 6.11. *Let $\mu_l = \frac{1}{\gamma_{d_l,s}^{\otimes \ell}(\Gamma_l)} \gamma_{d_l,s}^{\otimes \ell}\Big|_{\Gamma_l}$. Then the weak* limit of $(\widehat{\mu_l})_{l \geq 1}$ is $\varphi_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell}$.*

Proof. Notice for any $z = x_{k_1,i_1} \cdots x_{k_p,i_p} \otimes y_{l_1,j_1} \cdots y_{l_q,j_q} \in \mathcal{B}_R$ with $p + q \leq l$ that

$$\begin{aligned}\widehat{\mu_l}(z) - \varphi_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell}(z) = & \frac{1}{\gamma_{d_l,s}^{\otimes \ell}(\Gamma_l)} \int_{\Gamma_l} \tau_{d_l}\left(U_{k_1}^* A_{k_1,i_1,l} U_{k_1} \cdots U_{k_p}^* A_{k_p,i_p,l} U_{k_p} U_{l_q}^* B_{l_q,i_q,l} U_{l_q} \cdots U_{l_1}^* B_{l_1,i_1,l} U_{l_1}\right) d(\gamma_{d_l,s}^{\otimes \ell}) \\ & - \varphi(X_{k_1,i_1} \cdots X_{k_p,i_p} Y_{l_1,j_1} \cdots Y_{l_q,j_q})\end{aligned}$$

which is at most $\frac{1}{l}$ in absolute value by the definition of Γ_l. Thus $\widehat{\mu_l}(z)$ tends to $\varphi_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell}(z)$ as l tends to infinity for any $z \in \mathcal{B}_R$ thereby completing the proof. ☐

Lemma 6.12. *The weak* limit of $(\widehat{\gamma_{d_l,s}^{\otimes \ell}})_{l \geq 1}$ is $\varphi^{\text{bi-free}}_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell}$.*

Proof. For each $M \in \mathbb{N}$ and $\epsilon, \theta > 0$, Lemma 4.4 implies if

$$\Omega(M, d_l, \epsilon) = \{(U_k)_{k=1}^{\ell} \in U(d_l)^{\ell} \mid (U_1^* \mathbf{A}_{1,l} U_1, U_1^* \mathbf{B}_{1,l} U_1), \ldots, (U_\ell^* \mathbf{A}_{\ell,l} U_1, U_1^* \mathbf{B}_{\ell,l} U_1) \text{ are } (M,\epsilon)\text{-free}\}$$

then $\gamma^{\otimes \ell}(\Omega(M, d_\ell, \epsilon)) > 1 - \theta$ for d_l sufficiently large.

Let $\tau_{d_l}^{**\ell}$ denote the state on $\mathcal{B}_R$ obtained as follows: take the reduced free product of ℓ-copies of $\mathcal{B}(\mathcal{M}_{d_l})$ with respect to $z \mapsto \tau_{d_l}(z 1_d)$, and constructing the *-homomorphism π on $\mathcal{B}_R$ that sends $x_{k,i}$ to the left regular representation on the k^{th} copy of $\mathcal{B}(\mathcal{M}_{d_l})$ acting by left multiplication by $A_{k,i,l}$ and sends $y_{k,j}$ to the right regular representation on the k^{th} copy of $\mathcal{B}(\mathcal{M}_{d_l})$ acting by right multiplication by $B_{k,j,l}$. Then $\tau_{d_l}^{**\ell}$ is the vacuum state on the reduced free product composed with π. That is, $\tau_{d_l}^{**\ell}$ is the distribution so that $\{(\mathbf{A}_{k,l}, \mathbf{B}_{k,l})\}_{k=1}^{\ell}$ are bi-free with respect to the left-right matrix multiplication actions of $(\mathbf{A}_{k,l}, \mathbf{B}_{k,l})$ on $\mathcal{M}_{d_l}$. Note this distribution does not change if $\{(\mathbf{A}_{k,l}, \mathbf{B}_{k,l})\}_{k=1}^{\ell}$ is replaced with $\{(U_k^* \mathbf{A}_{k,l} U_k, U_k^* \mathbf{B}_{k,l} U_k)\}_{k=1}^{\ell}$.

Clearly $(\tau_{d_l}^{**\ell})_{l \geq 1}$ converges weak* to $\varphi^{\text{bi-free}}_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell}$ as

$$\gamma_{d_l}^{\otimes \ell}\left(\Gamma_{\text{orb}}\left(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell : (\mathbf{A}_{k,l})_{k=1}^{\ell}, (\mathbf{B}_{k,l})_{k=1}^{\ell}; l, d_l, \frac{1}{l}\right)\right) > 0.$$

Therefore, since for all $z = x_{k_1,i_1} \cdots x_{k_p,i_p} \otimes y_{l_1,j_1} \cdots y_{l_q,j_q} \in \mathcal{B}_R$ with $p + q \leq M$ we have

$$\begin{aligned}\widehat{\gamma_{d_l,s}^{\otimes \ell}}(z) &= \int_{SU(d_l)^{\otimes \ell}} \tau_{d_l}\left(U_{k_1}^* A_{k_1,i_1,l} U_{k_1} \cdots U_{k_p}^* A_{k_p,i_p,l} U_{k_p} U_{l_q}^* B_{l_q,i_q,l} U_{l_q} \cdots U_{l_1}^* B_{l_1,i_1,l} U_{l_1}\right) d\gamma_{d_l,s}^{\otimes \ell} \\ &= \int_{U(d_l)^{\otimes \ell}} \tau_{d_l}\left(U_{k_1}^* A_{k_1,i_1,l} U_{k_1} \cdots U_{k_p}^* A_{k_p,i_p,l} U_{k_p} U_{l_q}^* B_{l_q,i_q,l} U_{l_q} \cdots U_{l_1}^* B_{l_1,i_1,l} U_{l_1}\right) d\gamma_{d_l}^{\otimes \ell},\end{aligned}$$

we have that for sufficiently large l that

$$\begin{aligned}&\left|\widehat{\gamma_{d_l,s}^{\otimes \ell}}(z) - \tau_{d_l}^{**\ell}(z)\right| \\ &\leq \int_{\Omega(M,d_l,\epsilon)} \left|\tau_{d_l}\left(U_{k_1}^* A_{k_1,i_1,l} U_{k_1} \cdots U_{k_p}^* A_{k_p,i_p,l} U_{k_p} U_{l_q}^* B_{l_q,i_q,l} U_{l_q} \cdots U_{l_1}^* B_{l_1,i_1,l} U_{l_1}\right) - \tau_{d_l}^{**\ell}(z)\right| d\gamma_{d_l}^{\otimes \ell} \\ &\quad + \int_{U(d_l)^{\ell} \setminus \Omega(M,d_l,\epsilon)} \left|\tau_{d_l}\left(U_{k_1}^* A_{k_1,i_1,l} U_{k_1} \cdots U_{k_p}^* A_{k_p,i_p,l} U_{k_p} U_{l_q}^* B_{l_q,i_q,l} U_{l_q} \cdots U_{l_1}^* B_{l_1,i_1,l} U_{l_1}\right) - \tau_{d_l}^{**\ell}(z)\right| d\gamma_{d_l}^{\otimes \ell} \\ &\leq \epsilon + 2(R+1)^M \theta\end{aligned}$$

where the first inequality follows from (M, ϵ)-freeness (which gives the correct approximation of $\tau_{d_l}^{**\ell}(z)$ by the same arguments at the beginning of Lemma 4.6) and the second inequality follows from operator norm estimates and our bound on $\gamma^{\otimes \ell}(\Omega(M, d_\ell, \epsilon))$. As ϵ and θ can be made sufficiently small for any such M, we have that $(\widehat{\gamma_{d_l,s}^{\otimes \ell}})_{l \geq 1}$ and $(\tau_{d_l}^{**\ell})_{l \geq 1}$ have the same weak*-limit thereby completing the lemma. ☐

Now we need to know that the operation of taking a probability measure on $SU(d_l)^{\otimes \ell}$ and producing an element of $\mathcal{CS}(\mathcal{B}_R, \mathcal{L}, \mathcal{R})$ is well-behaved.

Lemma 6.13. *For any probability measures μ_1 and μ_2 on $SU(d_l)^\ell$, we have that*

$$W_2(\widehat{\mu_1}, \widehat{\mu_2}) \leq \frac{2R\sqrt{n+m}}{\sqrt{d_l}} W_{2, \|\cdot\|_{HS}}(\mu_1, \mu_2) \leq \frac{2R\sqrt{n+m}}{\sqrt{d_l}} W_{2, \|\cdot\|_{geod}}(\mu_1, \mu_2)$$

where $n = \max_{1 \leq k \leq \ell} n_k$, $m = \max_{1 \leq k \leq \ell} m_k$, and $W_{2, \|\cdot\|_{HS}}$ and $W_{2, \|\cdot\|_{geod}}$ are the 2-Wasserstein distances for measures with respect to the Hilbert-Schmidt norm $\|\cdot\|_{HS}$ and for the geodesic distance, respectively; here these are with respect to the inner product induced by the unnormalized *trace.*

Proof. The proof goes along the same lines as [14, Lemma 3.4]. First, let $\Pi(\mu_1, \mu_2)$ denote the set of all probability measures on $SU(d_l)^\ell \times SU(d_l)^\ell$ whose left- and right- marginal measures are μ_1 and μ_2 respectively. For each $\mu \in \Pi(\mu_1, \mu_2)$ we associate a state $\widehat{\mu}$ on

$$\left(*_{k=1}^\ell C[-R,R]^{*n_k}\right) * \left(*_{k=1}^\ell C[-R,R]^{*n_k}\right) \otimes_{\max} \left(*_{k=1}^\ell C[-R,R]^{*m_k}\right) * \left(*_{k=1}^\ell C[-R,R]^{*m_k}\right)$$

as described above (i.e. for each $((U_k)_{k=1}^\ell, (V_k)_{k=1}^\ell) \in SU(d_\ell)^\ell \times SU(d_\ell)^\ell$, for $1 \leq k \leq \ell$ we send $x_{k,i}$ to left multiplication by $U_k^* A_{k,i,l} U_k$ and $y_{k,j}$ to right multiplication by $U_k^* B_{k,j,l} U_k$, and for $\ell + 1 \leq k \leq 2\ell$ we send $x_{k,i}$ to left multiplication by $V_k^* A_{k,i,l} V_k$ and sends $y_{k,j}$ to right multiplication by $V_k^* B_{k,j,l} V_k$). By the definition of W_2 this immediately implies

$$W_2(\widehat{\mu_1}, \widehat{\mu_2}) \leq \sqrt{\int_{SU(d_l)^\ell} \int_{SU(d_l)^\ell} \sum_{k=1}^\ell \sum_{i=1}^{n_k} \|U_k^* A_{k,i,l} U_k - V_k^* A_{k,i,l} V_k\|_{\mathrm{HS}}^2 + \sum_{j=1}^{m_k} \|U_k^* B_{k,j,l} U_k - V_k^* B_{k,j,l} V_k\|_{\mathrm{HS}}^2 \, d\mu}$$

for any $\mu \in \Pi(\mu_1, \mu_2)$ where the first integration is with respect to $(U_k)_{k=1}^\ell$ and the second is with respect to $(V_k)_{k=1}^\ell$. Thus as

$$\|U_k^* A_{k,i,l} U_k - V_k^* A_{k,i,l} V_k\|_{\mathrm{HS}}^2 \leq 4R^2 \|U_k - V_k\|_{\mathrm{HS}}^2$$

with a similar inequality for the B-terms, we obtain the first inequality by the definition of the Wasserstein distances for measures with respect to the Hilbert-Schmidt norm.

Finally, the second inequality is trivial because the geodesic distance majorizes the Hilbert-Schmidt norm distance. ☐

The remainder of the proof of Proposition 6.10 is near identical to [14, Proposition 3.5]. Indeed since the Ricci curvature of $SU(d_l)^\ell$ (with respect to the inner product induced by the real part of the unnormalized trace) is known to be constant and equal to $\frac{d_l}{2}$, the transportation cost inequality

$$W_{2,\mathrm{geod}}\left(\mu_l, \gamma_{d_l,s}^{\otimes \ell}\right) \leq \sqrt{\frac{4}{d_l} S\left(\mu_l, \gamma_{d_l,s}^{\otimes \ell}\right)}$$

holds by [22], where μ_l is as in Lemma 6.11, $S\left(\mu_l, \gamma_{d_l,s}^{\otimes \ell}\right)$ denotes the relative entropy of μ_l with respect to $\gamma_{d_l,s}^{\otimes \ell}$ and thus

$$S\left(\mu_l, \gamma_{d_l,s}^{\otimes \ell}\right) = -\log\left(\gamma_{d_l,s}^{\otimes \ell}(\Gamma_l)\right)$$

by the definitions. By Lemma 6.13, we obtain that

$$W_2\left(\widehat{\mu}_l, \widehat{\gamma_{d_l,s}^{\otimes \ell}}\right) \leq 4R\sqrt{n+m}\sqrt{-\frac{1}{d_l^2}\log\left(\gamma_{d_l,s}^{\otimes \ell}(\Gamma_l)\right)}.$$

Hence Proposition 6.9, Lemma 6.11, and Lemma 6.12 yield the result by taking l to infinity and R to R_0. ☐

For completeness, we record the proof of Theorem 6.1.

Proof of Theorem 6.1. As $\chi_{\mathrm{orb}}(\mathbf{Z}_1, \ldots, \mathbf{Z}_\ell) = 0$, we obtain that

$$W_2(\varphi_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell}, \varphi_{\mathbf{Z}_1,\ldots,\mathbf{Z}_\ell}^{\text{bi-free}}) = 0,$$

thereby showing that $\mathbf{Z}_1$, $\mathbf{Z}_2$, …, $\mathbf{Z}_\ell$ are bi-free by Proposition 6.7. ☐

7. Calculating Microstate Entropy

In this section, we will compute the microstate bi-free entropy of several collections. We begin with the cases where there is a 'linear dependence in distribution'.

Lemma 7.1. *Let $(\mathcal{A},\varphi)$ be a C^*-non-commutative probability space and let $X, Y \in \mathcal{A}$ be self-adjoint such that $\varphi(X) = \varphi(Y) = 0$ and $\varphi(X^2) = \varphi(Y^2) = \varphi(XY) = 1$. Then*

$$\chi(X \sqcup Y) = -\infty.$$

Proof. Fix $R > \max\{\|X\|, \|Y\|\}$. Notice that it suffices to show that $\chi_R(X \sqcup Y; 2, \epsilon) \to -\infty$ as $\epsilon \to 0$. Towards this end, notice that for any $1 > \epsilon > 0$ and $d \in \mathbb{N}$,

$$\Gamma_R(X \sqcup Y; 2, d, \epsilon) \subseteq \left\{(A, B) \in (M_d^{\mathrm{sa}})^2 \mid \tau_d(A^2), \tau_d(B^2), \tau(AB) \in (1-\epsilon, 1+\epsilon)\right\}.$$

Recall, however, that the Lebesgue measure used is normalized based on the inner product given by the *unnormalized* trace:

$$\langle A, B\rangle_{\mathcal{M}_d^{\mathrm{sa}}} = d\tau_d(B^*A) = \mathrm{Tr}_d(B^*A) = \langle A, B\rangle_{\mathbb{R}^{d^2}}.$$

The three conditions on the set above then become $\|A\|^2_{\mathbb{R}^{d^2}}, \|B\|^2_{\mathbb{R}^{d^2}}, \langle A, B\rangle_{\mathbb{R}^{d^2}} \in (d(1-\epsilon), d(1+\epsilon))$. These restrictions allow us to deduce a bound on the angle $\theta_{A,B}$ between any A and B in the set:

$$\cos\theta_{A,B} = \frac{\langle A, B\rangle}{\|A\|\,\|B\|} \geq \frac{1-\epsilon}{1+\epsilon} \qquad \text{whence} \qquad \tan\theta_{A,B} = \frac{\sqrt{1-\cos^2\theta_{A,B}}}{\cos\theta_{A,B}} \leq \frac{\sqrt{1-\left(\frac{1-\epsilon}{1+\epsilon}\right)^2}}{\frac{1-\epsilon}{1+\epsilon}} = \frac{2\sqrt{\epsilon}}{1-\epsilon}.$$

Consequently B must lie in the cone from the origin in the direction of A with height $\sqrt{d(1+\epsilon)}$ and radius at its base $\sqrt{d(1+\epsilon)}\frac{2\sqrt{\epsilon}}{1-\epsilon}$. Letting $C(A, d, \epsilon)$ represent this cone, we have

$$\Gamma_R(X \sqcup Y; 2, d, \epsilon) \subseteq \left\{(A, B) \in \mathbb{R}^{2d^2} \mid \|A\|^2 \leq d(1+\epsilon), B \in C(A, d, \epsilon)\right\}.$$

Since volume of the cone $C(A, d, \epsilon)$ does not depend on A, the volume of the set on the right hand side is the product of that of the ball $\mathcal{B}\left(d^2, \sqrt{d(1+\epsilon)}\right)$ of radius $\sqrt{d(1+\epsilon)}$ in dimension d^2, and that of any cone $C(A, d, \epsilon)$. Fortunately, it is known that the volume of the n-ball of radius R is

$$\lambda_n(\mathcal{B}(n, R)) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}R^n;$$

see, e.g., [27, VIII.13], although this has been known at least since [20] (we were not able to find an earlier reference, though it is very likely one exists). In particular,

$$\lambda_{d^2}\left(\mathcal{B}\left(n, \sqrt{d(1+\epsilon)}\right)\right) = \frac{\pi^{\frac{d^2}{2}}}{\Gamma\left(\frac{d^2}{2}+1\right)}(d(1+\epsilon))^{\frac{d^2}{2}}.$$

We now may compute the volume of $C(A, d, \epsilon)$ by integrating along its axis:

$$\begin{aligned}
\lambda_{d^2}(C(A, d, \epsilon)) &= \int_0^{\sqrt{d(1+\epsilon)}} \lambda_{d^2-1}\left(\mathcal{B}\left(d^2-1, R\frac{2\sqrt{\epsilon}}{1-\epsilon}\right)\right) dR \\
&= \int_0^{\sqrt{d(1+\epsilon)}} \frac{\pi^{\frac{d^2-1}{2}}}{\Gamma\left(\frac{d^2-1}{2}+1\right)}\left(R\frac{2\sqrt{\epsilon}}{1-\epsilon}\right)^{d^2-1} dR \\
&= \frac{1}{d^2}(d(1+\epsilon))^{\frac{d^2}{2}} \frac{\pi^{\frac{d^2-1}{2}}}{\Gamma\left(\frac{d^2-1}{2}+1\right)}\left(\frac{2\sqrt{\epsilon}}{1-\epsilon}\right)^{d^2-1}.
\end{aligned}$$

Finally, we recall that Stirling's formula allows us to make the estimate that for large $z > 0$, $\frac{1}{z}\log\Gamma(z) = \log z + \mathcal{O}(1)$. This allows us to make the following estimate:

$$\begin{aligned}
\frac{1}{d^2}\chi_R(X \sqcup Y; 2, d, \epsilon) &\leq \log(d) - \frac{1}{d^2}\log\Gamma\left(\frac{d^2}{2}+1\right) - \frac{1}{d^2}\log\Gamma\left(\frac{d^2-1}{2}+1\right) + \frac{d^2-1}{2d^2}\log\epsilon + \mathcal{O}_{d,\epsilon}(1) \\
&= -\log(d) + \frac{d^2-1}{2d^2}\log\epsilon + \mathcal{O}_{d,\epsilon}(1).
\end{aligned}$$

Thus $\chi_R(X \sqcup Y; 2, \epsilon) \leq \frac{1}{2} \log \epsilon + \mathcal{O}_\epsilon(1)$ so sending $\epsilon \to 0$ yields $\chi(X \sqcup Y) = -\infty$. □

Using the above, we can prove the following which, when combined with Corollary 3.2, completely determines the microstate bi-free entropy of a tracially bi-partite system with a linear dependence in distribution.

Theorem 7.2. *Let $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$ be a tracially bi-partite system in a C^*-non-commutative probability space $(\mathcal{A}, \varphi)$. If there exists an $X \in \mathrm{span}\{X_1, \ldots, X_n\}$ and a $Y \in \mathrm{span}\{Y_1, \ldots, Y_n\}$ such that $1 = \varphi(X^2) = \varphi(XY) = \varphi(Y^2)$ (e.g. $X_1, \ldots, X_n$ linearly independent, $Y_1, \ldots, Y_m$ linearly independent, yet $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ linearly dependent in distribution), then*

$$\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) = -\infty.$$

Proof. If $X_1, \ldots, X_n$ or $Y_1, \ldots, Y_n$ are linearly dependent, then the result follows from Corollary 3.2. Otherwise there exists an $i \in \{1, \ldots, n\}$ and a $j \in \{1, \ldots, m\}$ such that $\{X_1, \ldots, X_n\}$ and $\{X_1, \ldots, X_{i-1}, X, X_{i+1}, \ldots, X_n\}$ are bases for the same subspace of $\mathcal{A}$, and $\{Y_1, \ldots, Y_m\}$ and $\{Y_1, \ldots, Y_{j-1}, Y, Y_{j+1}, \ldots, Y_m\}$ are bases for the same subspace of $\mathcal{A}$. By Corollary 3.2 there exists a $C \in \mathbb{R}$ such that

$$\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) = C + \chi(X_1, \ldots, X_{i-1}, X, X_{i+1}, \ldots, X_n \sqcup Y_1, \ldots, Y_{j-1}, Y, Y_{j+1}, \ldots, Y_m).$$

As

$$\begin{aligned} &\chi(X_1, \ldots, X_{i-1}, X, X_{i+1}, \ldots, X_n \sqcup Y_1, \ldots, Y_{j-1}, Y, Y_{j+1}, \ldots, Y_m) \\ &\quad \leq \chi(X \sqcup Y) + \chi(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n \sqcup Y_1, \ldots, Y_{j-1}, Y_{j+1}, \ldots, Y_m) \end{aligned}$$

by Proposition 2.5 and as $\chi(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n \sqcup Y_1, \ldots, Y_{j-1}, Y_{j+1}, \ldots, Y_m) < \infty$ by Proposition 2.6, Lemma 7.1 yields $\chi(X \sqcup Y) = -\infty$ and the result. □

Next we investigate the bi-free entropy of bi-free central limit distributions. Since we are only able to apply transformations to the left variables and the right variables separately, we cannot directly remove correlations between left and right semicircular variables. We therefore start with the case of two variables.

Theorem 7.3. *Let $(\mathcal{A}, \varphi)$ be a C^*-non-commutative probability space and let (S_ℓ, S_r) be a centred, self-adjoint bi-free central limit distribution in $\mathcal{A}$ in which each variable is of variance one. If $c = \varphi(S_\ell S_r) \in [-1, 1]$, then*

$$\chi(S_\ell \sqcup S_r) = \log(2\pi e) + \frac{1}{2}\log(1 - c^2).$$

Furthermore, the $\limsup_{d\to\infty}$ when computing $\chi(S_\ell \sqcup S_r)$ is actually a $\lim_{d\to\infty}$.

Proof. By Example 2.8, we see that

$$\chi(S_\ell \sqcup S_r) \geq \log(2\pi e) + \frac{1}{2}\log(1 - c^2)$$

as the free entropy of a single semicircular operator of variance one is $\frac{1}{2}\log(2\pi e)$. Furthermore, we claim this inequality holds if we use the $\liminf_{d\to\infty}$ in place of $\limsup_{d\to\infty}$ for $\chi(S_\ell \sqcup S_r)$. To see this, notice Example 2.8 holds for the $\liminf_{d\to\infty}$ version since both [32, Proposition 3.5 and Proposition 5.4] and Theorem 2.7 do as well. Therefore, since the free entropy of a single semicircular operator agrees with the $\liminf_{d\to\infty}$ variety, the claim is complete.

For the other direction, we will apply some volume arguments. Note the case $c = \pm 1$ follows from Lemma 7.1, so we will assume $|c| < 1$.

For each $R > 2$ and $M \in \mathbb{N}$ with $M \geq 2$, notice that $\Gamma_R(S_\ell \sqcup S_r; M, d, \epsilon)$ is contained in

$$\Psi := \left\{ (A_1, A_2) \in (\mathcal{M}_d^{\mathrm{sa}})^2 \,\middle|\, 1 - \epsilon \leq \tau_d(A_k^2) \leq 1 + \epsilon, c - \epsilon \leq \tau_d(A_1 A_2) \leq c + \epsilon \right\}.$$

We desire an estimate on the Lebesgue measure of Ψ.

Recall we view $(\mathcal{M}_d^{\mathrm{sa}})^2 \cong (\mathbb{R}^{d^2})^2$ as Hilbert spaces where for $A, B \in \mathcal{M}_d^{\mathrm{sa}}$ we have

$$\langle A, B\rangle_{\mathcal{M}_d^{\mathrm{sa}}} = d\tau_d(B^*A) = \mathrm{Tr}(B^*A) = \langle A, B\rangle_{\mathbb{R}^{d^2}}.$$

Hence

$$\Psi \cong \left\{ (A_1, A_2) \in (\mathbb{R}^{d^2})^2 \,\middle|\, \sqrt{d(1-\epsilon)} \leq \|A_k\|_2 \leq \sqrt{d(1+\epsilon)}, d(c - \epsilon) \leq \langle A_1, A_2\rangle_{\mathbb{R}^{d^2}} \leq d(c + \epsilon) \right\}.$$

Consider the map $\Theta : (\mathbb{R}^{d^2})^2 \to (\mathbb{R}^{d^2})^2$ defined by

$$\Theta(A_1, A_2) = \left(A_1, -\frac{c}{\sqrt{1-c^2}}A_1 + \frac{1}{\sqrt{1-c^2}}A_2\right).$$

Clearly Θ is a direct sum of d^2 copies of the matrix

$$Q = \begin{bmatrix} 1 & 0 \\ -\frac{c}{\sqrt{1-c^2}} & \frac{1}{\sqrt{1-c^2}} \end{bmatrix}$$

via a specific choice of orthonormal basis of $\mathbb{R}^{d^2}$. Hence the Jacobian of Θ is also a direct sum of d^2 copies of Q and thus

$$\mathrm{Vol}(\Psi) = \frac{1}{\det(\mathcal{J}(\Theta))}\mathrm{Vol}(\Theta(\Psi)) = (1-c^2)^{\frac{d^2}{2}}\mathrm{Vol}(\Theta(\Psi)).$$

To obtain an upper bound for the volume of $\Theta(\Psi)$, we claim that

$$\Theta(\Psi) \subseteq \left\{(B_1, B_2) \in (\mathbb{R}^{d^2})^2 \,\middle|\, \|B_k\|_2 \leq \sqrt{d\left(1 + \epsilon\frac{(1+|c|)^2}{1-c^2}\right)}\right\}.$$

To see this, fix $(A_1, A_2) \in \Psi$ and let $(B_1, B_2) = \Theta(A_1, A_2)$. Then $B_1 = A_1$ so

$$\|B_1\|_2 \leq \sqrt{d(1+\epsilon)} \leq \sqrt{d\left(1 + \epsilon\frac{(1+|c|)^2}{1-c^2}\right)}.$$

Next notice that

$$\begin{aligned}
\|B_2\|_2^2 &= \left\langle -\frac{c}{\sqrt{1-c^2}}A_1 + \frac{1}{\sqrt{1-c^2}}A_2, -\frac{c}{\sqrt{1-c^2}}A_1 + \frac{1}{\sqrt{1-c^2}}A_2\right\rangle_{\mathbb{R}^{d^2}} \\
&= \frac{1}{1-c^2}\left(c^2\langle A_1, A_1\rangle - 2c\langle A_1, A_2\rangle + \langle A_2, A_2\rangle\right) \\
&\leq \frac{1}{1-c^2}\left(d(1+\epsilon)c^2 - 2dc^2 + 2d|c|\epsilon + d(1+\epsilon)\right) \\
&= \frac{d}{1-c^2}\left((1-c^2) + \epsilon(1 + 2|c| + c^2)\right) \\
&= d\left(1 + \epsilon\frac{(1+|c|)^2}{1-c^2}\right).
\end{aligned}$$

Hence the claim is complete.

Using the above and the fact that $\Theta(\Psi)$ is contained in the product of two d^2-dimensional balls of radius $\sqrt{d\left(1+\epsilon\frac{(1+|c|)^2}{1-c^2}\right)}$, we obtain that

$$\begin{aligned}
&\lambda_{d,2}(\Gamma_R(S_\ell \sqcup S_r; M, d, \epsilon)) \\
&\leq \mathrm{Vol}(\Psi) \\
&\leq (1-c^2)^{\frac{d^2}{2}}\mathrm{Vol}(\Theta(\Psi)) \\
&\leq ((1-c^2)^{\frac{d^2}{2}}\left(\frac{\pi^{\frac{d^2}{2}}}{\Gamma\left(\frac{d^2}{2}+1\right)}\left(d\left(1+\epsilon\frac{(1+|c|)^2}{1-c^2}\right)\right)^{\frac{d^2}{2}}\right)^2.
\end{aligned}$$

Hence, via an application of Stirling's formula, we obtain that

$$\begin{aligned}
&\chi_R(S_\ell \sqcup S_r; M, \epsilon) \\
&\leq \limsup_{d\to\infty} \frac{1}{2}\log(1-c^2) + \log(\pi) - 2\frac{1}{d^2}\log\left(\Gamma\left(\frac{d^2}{2}+1\right)\right) + \log(d) + \log\left(1+\epsilon\frac{(1+|c|)^2}{1-c^2}\right) + \frac{2}{2}\log(d) \\
&\leq \limsup_{d\to\infty} \frac{1}{2}\log(1-c^2) + \log(\pi) + \log(2e) + \log\left(1+\epsilon\frac{(1+|c|)^2}{1-c^2}\right).
\end{aligned}$$

Therefore
$$\chi(S_\ell \sqcup S_r) \leq \log(2\pi e) + \frac{1}{2}\log(1-c^2)$$
completing the claim. □

Combining all of the results of this paper, we obtain the following.

Theorem 7.4. *Let $(\{S_k\}_{k=1}^n, \{S_k\}_{k=n+1}^{n+m})$ be a centred self-adjoint bi-free central limit distribution with respect to φ with $\varphi(S_k^2) = 1$ for all k. Recall that the joint distribution is completely determined by the positive matrix*
$$A = [a_{i,j}] = [\varphi(S_iS_j)] \in \mathcal{M}_n(\mathbb{R}).$$
Then
$$\chi(S_1, \ldots, S_n \sqcup S_{n+1}, \ldots, S_{n+m}) = \frac{n+m}{2}\log(2\pi e) + \frac{1}{2}\log\left(\det(A)\right).$$

Proof. Note that if A is not invertible then either $\{S_k\}_{k=1}^n$ are linearly dependent (in distribution), $\{S_k\}_{k=n+1}^{n+m}$ are linearly dependent (in distribution), or the hypotheses of Theorem 7.2 are satisfied. Hence, by Corollary 3.2, the result holds if A is not invertible. Thus we will suppose that A is invertible.

Recall that we can view $(\{S_k\}_{k=1}^n, \{S_k\}_{k=n+1}^{n+m})$ as left and right semicircular operators acting on a real Fock space. In particular for $k \in \{1, \ldots, n\}$ we can write
$$S_k = l(e_k) + l^*(e_k)$$
and for $k \in \{n+1, \ldots, n+m\}$ we can write
$$S_k = r(e_k) + r^*(e_k)$$
where $\{e_k\}_{k=1}^{n+m} \in \mathcal{H}$ are unit vectors. Note
$$A = [\langle e_i, e_j\rangle]$$
so we obtain that $\{e_k\}_{k=1}^{n+m}$ is linearly independent.

We now discuss how modifications to $\{e_k\}_{k=1}^n$ and modifications to $\{e_k\}_{k=n+1}^{n+m}$ modify the bi-free entropy and the covariance matrix. Suppose $Q = [q_{i,j}] \in \mathcal{M}_n(\mathbb{R})$ and $R = [r_{i,j}] \in \mathcal{M}_m(\mathbb{R})$ are invertible. If for each $k \in \{1, \ldots, n\}$ we define
$$e_k' = \sum_{i=1}^n q_{k,i}e_i$$
and for each $k \in \{n+1, \ldots, n+m\}$ we define
$$e_k' = \sum_{j=1}^m r_{k,j}e_{j+n}$$
then $\{e_k'\}_{k=1}^{n+m}$ is linearly independent,
$$\begin{aligned}
&\chi(l(e_1') + l^*(e_1'), \ldots, l(e_n') + l^*(e_n') \sqcup r(e_{n+1}') + r^*(e_{n+1}'), \ldots, r(e_{n+m}) + r^*(e_{n+m})) \\
&= \chi\left(\sum_{i=1}^n q_{1,i}S_i, \ldots, \sum_{i=1}^n q_{n,i}S_i \sqcup \sum_{j=1}^m r_{1,j}S_{j+n}, \ldots, \sum_{j=1}^m r_{m,j}S_{j+n}\right) \\
&= \chi(S_1, \ldots, S_n \sqcup S_{n+1}, \ldots, S_{n+m}) + \log(|\det(Q)|) + \log(|\det(R)|)
\end{aligned}$$
by Corollary 3.2, and
$$[\langle e_i', e_j'\rangle] = (Q \oplus R)[\langle e_i, e_j\rangle](Q \oplus R)^*.$$
Thus
$$\frac{1}{2}\log(\left|\det([\langle e_i', e_j'\rangle])\right|) = \frac{1}{2}\log(|\det([\langle e_i, e_j\rangle])|) + \log(|\det(Q)|) + \log(|\det(R)|).$$
Therefore, as both sides of the claimed formula
$$\chi(S_1, \ldots, S_n \sqcup S_{n+1}, \ldots, S_{n+m}) = \frac{n+m}{2}\log(2\pi e) + \frac{1}{2}\log\left(\det(A)\right)$$
are preserved under such operations, we will apply such operations until we arrive at a case we can deduce from previous results.

First, as applying the Gram-Schmidt Orthogonalization Process to $\{e_k\}_{k=1}^n$ and to $\{e_k\}_{k=n+1}^{n+m}$ produces such matrices Q and R due to linear independence of $\{e_k\}_{k=1}^{n+m}$, we may assume that $\{e_k\}_{k=1}^n$ is orthonormal and $\{e_k\}_{k=n+1}^{n+m}$ is orthonormal. In this case

$$A = \begin{bmatrix} I_n & B \\ B^* & I_m \end{bmatrix}$$

where B is an $n \times m$ matrix with real entries. Let us assume that $m > n$ (the other case being similar). It follows that there are $m - n$ columns of B that are linear combinations of the other n columns of B. Let $\{j_1, \ldots, j_n\}$ denote the indices of these other n columns of B. Notice since $\{e_k\}_{k=n+1}^{n+m}$ is linearly independent set of m vectors that we can replace e_k where $k \geq n+1$ and $k \neq j_q$ for all q with $e_k - \sum_{q=1}^n c_{k,q} e_{j_q}$ (where the $c_{k,q}$ are chosen based on how column k of B is a linear combination of columns $j_1, \ldots, j_n$) so that $\{e_k\}_{k=n+1}^{n+m}$ remains a linearly independent set and so that $\langle e_k, e_p \rangle = 0$ for all $k \geq n+1$ with $k \neq j_q$ for all q, and all $p \leq n$. Subsequently, if we apply the Gram-Schmidt Orthogonalization Process first to the modified e_k for $k \neq j_q$ for all q, and then the remainder of the e_k, and if we then permute the order of the resulting vectors, the resulting change of basis matrix can then, with the above arguments, be used so that we may assume

$$A = \begin{bmatrix} I_n & C & 0_{n,m-n} \\ C^* & I_n & 0_{n,m-n} \\ 0_{m-n,n} & 0_{m-n,n} & I_{m-n} \end{bmatrix}$$

where C is an $n \times n$ matrix with real entries.

Recall, by the Singular Value Decomposition, we can write $C = UDV$ where $U, V \in \mathcal{M}_n(\mathbb{R})$ are unitary matrices and $D = \mathrm{diag}(d_1, \ldots, d_n)$ is a diagonal matrix. By using $Q = U$ and $R = V^* \oplus I_{m-n}$, we reduce to the case where

$$A = \begin{bmatrix} I_n & D & 0_{n,m-n} \\ D^* & I_n & 0_{n,m-n} \\ 0_{m-n,n} & 0_{m-n,n} & I_{m-n} \end{bmatrix}.$$

Notice in this case that the determinant of A is $\prod_{k=1}^n (1 - d_k^2)$. Furthermore, in this case, we obtain that

$$(\{S_1\}, \{S_{n+1}\}), (\{S_2\}, \{S_{n+2}\}), \ldots, (\{S_n\}, \{S_{2n}\}), (\emptyset, \{S_{2n+1}\}), \ldots, (\emptyset, \{S_{n+m}\})$$

are bi-freely independent. Therefore, as pairs of semicirculars have finite-dimensional approximants and as Theorem 7.3 implies the $\limsup_{d\to\infty}$ for pairs of semicirculars is actually a limit, Theorem 4.7 implies that

$$\chi(S_1, \ldots, S_n \sqcup S_{n+1}, \ldots, S_{n+m}) = \sum_{k=1}^n \chi(S_k \sqcup S_{k+n}) + \sum_{j=2n+1}^{n+m} \chi(\sqcup S_j) = \sum_{k=1}^n \chi(S_k \sqcup S_{k+n}) + \sum_{j=2n+1}^{n+m} \chi(S_j).$$

(Here we denote by $\chi(\sqcup S_j)$ As Theorem 7.3 implies that

$$\chi(S_k \sqcup S_{k+n}) = \log(2\pi e) + \frac{1}{2}\log(1 - d_k^2),$$

and as we know

$$\chi(S_j) = \frac{1}{2}\log(2\pi e),$$

we obtain that

$$\begin{aligned} \chi(S_1, \ldots, S_n \sqcup S_{n+1}, \ldots, S_{n+m}) &= \frac{n+m}{2}\log(2\pi e) + \frac{1}{2}\sum_{k=1}^n \log(1 - d_k^2) \\ &= \frac{n+m}{2}\log(2\pi e) + \frac{1}{2}\log(|\det(A)|). \end{aligned}$$ □

Remark 7.5. Note that Theorem 7.4 includes the free case (i.e. when $m = 0$). However, the proof for the microstate free entropy of free central limit distributions is substantially easier as one may apply transformations to all of the variables. The bi-free proof is more difficult as Section 3 did not demonstrate the ability to mix left and right variables. Still it is not surprising that we get the same result as the free case seeing as, asymptotically, almost all matrices are microstates for semicircular operators so it is simply a matter of angles. One would expect other random variables which may have more complicated microstate sets could lead to different behaviours for which the above angle arguments would not apply.

8. Microstate Bi-Free Entropy Dimension

For the sake of completeness, we briefly study the microstate bi-free entropy dimension. Unfortunately, we do not know the correct bi-free generalizations of the known von Neumann algebra implications of free entropy dimension.

Definition 8.1. Let $(\mathcal{A}, \varphi)$ be a C*-non-commutative probability space and let $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ be self-adjoint operators in $\mathcal{A}$. The *n-left, m-right, microstate bi-free entropy dimension* is defined by

$$\delta(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) = n + m + \limsup_{\epsilon \to 0^+} \frac{\chi(X_1 + \sqrt{\epsilon}S_1, \ldots, X_n + \sqrt{\epsilon}S_n \sqcup Y_1 + \sqrt{\epsilon}T_1, \ldots, Y_m + \sqrt{\epsilon}T_m)}{|\log(\sqrt{\epsilon})|}$$

where $\{(S_i, I)\}_{i=1}^n \cup \{(I, T_j)\}_{j=1}^m$ is a bi-free central limit distribution of semicircular operators with variances 1 and covariances 0 that is bi-free from $(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m)$.

It is elementary to see based on bi-freeness that the self-adjoint operators $(\{X_i + \sqrt{\epsilon}S_i\}_{i=1}^n, \{Y_j + \sqrt{\epsilon}T_i\}_{j=1}^m)$ still form a tracially bi-partite collection and thus $\delta(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m)$ is well-defined. In addition, a few basis properties of free entropy dimension carry-forward to the bi-free setting.

Proposition 8.2. *If $0 \leq p \leq n$ and $0 \leq q \leq m$ then*

$$\delta(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) \leq \delta(X_1, \ldots, X_p \sqcup Y_1, \ldots Y_q) + \delta(X_{p+1}, \ldots, X_n \sqcup Y_{q+1}, \ldots Y_m).$$

In particular,

$$\delta(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) \leq \delta(X_1, \ldots, X_n) + \delta(Y_1, \ldots, Y_m).$$

Proof. This result immediately follows from Definition 8.1 and Proposition 2.5. ☐

Remark 8.3. Under sufficiently strong assumptions, the bi-free entropy dimension can be shown to be additive across bi-free families. One always has the estimate

$$\delta(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) \leq \delta(X_1, \ldots, X_p \sqcup Y_1, \ldots, Y_q) + \delta(X_{p+1}, \ldots, X_n \sqcup Y_{q+1}, \ldots, Y_m)$$

by applying the same inequality for χ and splitting the limits superior. The other direction, however, requires knowing χ is additive for the perturbed variables (up to $o\left(|\log \epsilon|\right)$) along some sequence of ϵ's witnessing the limit superior. This could also be ensured by taking limits along appropriate ultrafilters for $(0, 1]$ and $\mathcal{N}$ in the definitions of δ and χ, respectively, in place of the limsups. This approach is the same as that proposed by Voiculescu in the free setting in [36, Remark 4.8].

What is most interesting about microstate bi-free entropy dimension is its value of bi-free central limit distributions.

Theorem 8.4. *Let $(\{S_k\}_{k=1}^n, \{S_k\}_{k=n+1}^{n+m})$ be a centred self-adjoint bi-free central limit distribution with respect to φ with $\varphi(S_k^2) = 1$ for all k. Recall that the joint distribution is completely determined by the positive matrix*

$$A = [a_{i,j}] = [\varphi(S_i S_j)] \in \mathcal{M}_n(\mathbb{R}).$$

Then

$$\delta(S_1, \ldots, S_n \sqcup S_{n+1}, \ldots, S_{n+m}) = \mathrm{rank}(A).$$

Proof. Let $(\{T_k\}_{k=1}^n, \{T_k\}_{k=n+1}^{n+m})$ be a centred self-adjoint bi-free central limit distribution with respect to φ with

$$\varphi(T_i T_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

If we define $Z_{k,\epsilon} = S_k + \sqrt{\epsilon}T_k$ for all $1 \leq k \leq n + m$, then $(\{Z_k\}_{k=1}^n, \{Z_k\}_{k=n+1}^{n+m})$ is a centred self-adjoint bi-free central limit distribution with respect to φ with

$$\varphi(Z_{i,\epsilon} Z_{j,\epsilon}) = \begin{cases} 1 + \epsilon & \text{if } i = j \\ \varphi(S_i S_j) & \text{if } i \neq j \end{cases}$$

and

$$\delta(S_1, \ldots, S_n \sqcup S_{n+1}, \ldots, S_{n+m}) = n + m + \limsup_{\epsilon \to 0+} \frac{\chi(Z_{1,\epsilon}, \ldots, Z_{n,\epsilon} \sqcup Z_{n+1,\epsilon}, \ldots, Z_{n+m,\epsilon})}{|\log(\sqrt{\epsilon})|}.$$

By applying Corollary 3.2 and Theorem 7.4, we see that

$$\begin{aligned}
&\chi(Z_{1,\epsilon},\dots,Z_{n,\epsilon}\sqcup Z_{n+1,\epsilon},\dots,Z_{n+m,\epsilon})\\
&=(n+m)\log(\sqrt{1+\epsilon})+\chi\left(\frac{1}{\sqrt{1+\epsilon}}Z_{1,\epsilon},\dots,\frac{1}{\sqrt{1+\epsilon}}Z_{n,\epsilon}\sqcup\frac{1}{\sqrt{1+\epsilon}}Z_{n+1,\epsilon},\dots,\frac{1}{\sqrt{1+\epsilon}}Z_{n+m,\epsilon}\right)\\
&=\frac{n+m}{2}\log(1+\epsilon)+\frac{n+m}{2}\log(2\pi e)+\frac{1}{2}\log\left(\det\left(\left(1-\frac{1}{1+\epsilon}\right)I_{n+m}+\frac{1}{1+\epsilon}A\right)\right)\\
&=\frac{n+m}{2}\log(2\pi e)+\frac{1}{2}\log\left(\det\left(\epsilon I_{n+m}+A\right)\right).
\end{aligned}$$

As A is a positive matrix and thus diagonalizable, we know that

$$\det\left(\epsilon I_{n+m}+A\right)=\epsilon^{\mathrm{nullity}(A)}p(\epsilon)$$

where p is a polynomial of degree $\mathrm{rank}(A)$ with real coefficients that does not vanish at 0. Consequently, we obtain that

$$\begin{aligned}
&\delta(S_1,\dots,S_n\sqcup S_{n+1},\dots,S_{n+m})\\
&=n+m+\limsup_{\epsilon\to0^+}\frac{\frac{n+m}{2}\log(2\pi e)+\frac{1}{2}\log(\epsilon^{\mathrm{nullity}(A)}p(\epsilon))}{|\log(\sqrt{\epsilon})|}\\
&=n+m+\limsup_{\epsilon\to0^+}\frac{\frac{n+m}{2}\log(2\pi e)+\frac{1}{2}\mathrm{nullity}(A)\log(\epsilon)+\frac{1}{2}\log(p(\epsilon))}{|\log(\sqrt{\epsilon})|}\\
&=n+m-\mathrm{nullity}(A)=\mathrm{rank}(A)
\end{aligned}$$

as desired. ☐

Remark 8.5. Let (S,T) be a bi-free central limit distribution with variances 1 and covariance $c\in[-1,1]$. Hence

$$\delta(S\sqcup T)=\begin{cases}2 & \text{if } c\neq\pm1\\ 1 & \text{if } c=\pm1\end{cases}.$$

In particular, the support of the joint distribution of (S,T) has dimension $\delta^*(S\sqcup T)$: indeed, if $c\neq\pm1$ then (S,T) has joint distribution with support $[-2,2]^2\subset\mathbb{R}^2$ by [17], while otherwise it is supported on the line $y=cx$. This adds validation to the name "bi-free microstate entropy dimension".

9. Microstate Bi-Free Entropy for Non-Bi-Partite Systems

In the section, we will discuss our notion of microstate bi-free entropy to non-bi-partite systems where further complications arise. To do this, we will find it useful to take an approach from operator-valued bi-free probability. We refer the reader to [7] rather than reintroduce the entire setting here.

Let $(\mathcal{C},\varphi)$ be a non-commutative probability space and let B be a unital algebra. Then $\mathcal{C}\otimes B$ can be viewed as a B-B-bimodule where

$$b\cdot(a\otimes b')=a\otimes bb',\qquad\text{and}\qquad(a\otimes b')\cdot b=a\otimes b'b$$

for $b,b'\in B$ and $a\in\mathcal{C}$. Let us denote by L_b and R_b the left and right actions of b above. If $p_B:\mathcal{C}\otimes B\to B$ is defined by

$$p_B(a\otimes b)=\varphi(a)b,$$

then $\mathcal{L}(\mathcal{C}\otimes B)$ is a B-B-non-commutative probability space with left and right B-operators L_b and R_b respectively and expectation $E:\mathcal{L}(\mathcal{C}\otimes B)\to B$ defined by

$$E(Z)=p_B(Z(1_{\mathcal{C}}\otimes1_B))$$

for all $Z\in\mathcal{L}(\mathcal{C}\otimes B)$. Let $\mathcal{L}(\mathcal{C}\otimes B)_\ell$ denote all elements of $\mathcal{L}(\mathcal{C}\otimes B)$ that commute with elements of $\{R_b\mid b\in B\}$ and let $\mathcal{L}(\mathcal{C}\otimes B)_r$ denote all elements of $\mathcal{L}(\mathcal{C}\otimes B)$ that commute with elements of $\{L_b\mid b\in B\}$. Therefore, if $X,Y\in\mathcal{C}$ and $b\in B$, we can define $L(X\otimes b)\in\mathcal{L}(\mathcal{C}\otimes B)_\ell$ and $R(Y\otimes b)\in\mathcal{L}(\mathcal{C}\otimes B)_r$ via

$$L(X\otimes b)(a\otimes b')=Xa\otimes bb'\qquad\text{and}\qquad R(Y\otimes b)(a\otimes b')=Ya\otimes b'b.$$

for all $a\in\mathcal{C}$ and $b'\in B$.

Our current approach to matricial microstates has been to find matrices in $\mathcal{M}_d$ for which the moments of the appropriate left or right multiplication operators computed against $\tau_d(\cdot I_d)$ have been approximately correct. Since $\mathcal{L}(\mathcal{M}_d) \cong \mathcal{M}_d \otimes \mathcal{M}_d^{\mathrm{op}}$ via $L(A)R(B) \mapsto A \otimes B^{\mathrm{op}}$, we may view this in the above setting with $\mathcal{C} = \mathbb{C}$. If we replace $\mathbb{C}$ by some larger matrix algebra, we introduce non-commutativity between the left and the right approximants.

Indeed for fixed $d_1, d_2 \in \mathbb{N}$, if we identify $\mathcal{L}(\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}) \cong \mathcal{L}(\mathcal{M}_{d_1}) \otimes \mathcal{L}(\mathcal{M}_{d_2}) \cong \mathcal{L}(\mathcal{M}_{d_1}) \otimes \mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2}^{\mathrm{op}}$, we find

$$\mathcal{L}(\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2})_\ell \cong \mathcal{L}(\mathcal{M}_{d_1}) \otimes \mathcal{M}_{d_2} \otimes \mathbb{C} \qquad \text{and} \qquad \mathcal{L}(\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2})_r \cong \mathcal{L}(\mathcal{M}_{d_1}) \otimes \mathbb{C} \otimes \mathcal{M}_{d_2}^{\mathrm{op}}.$$

In particular, the pair of faces $(L(\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}), R(\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}))$ in $\mathcal{L}(\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2})$ is isomorphic to the pair of faces

$$\left((\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} \otimes \mathbb{C}, \mathcal{M}_{d_1} \otimes \mathbb{C} \otimes \mathcal{M}_{d_2}^{\mathrm{op}})\right)$$

in $\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2}^{\mathrm{op}}$. Note that the state becomes $\tau_{d_1} \otimes (\tau_{d_2} \circ m)$, where $m(B_1 \otimes B_2^{\mathrm{op}}) = B_1 B_2$ is the multiplication map. These faces are each as measure spaces isomorphic to $\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$. We have for $A \otimes B \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$, $L(A \otimes B) = A \otimes B \otimes I_{d_2}$ while $R(A \otimes B) = A \otimes I_{d_2} \otimes B^{\mathrm{op}}$.

Using the above constructions, we postulate the following generalization of our microstate bi-free entropy to the non-tracially bi-partite setting. Let $(\mathcal{A}, \varphi)$ be a C*-non-commutative probability space and let $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ be self-adjoint operators in $\mathcal{A}$, where we will consider $X_1, \ldots, X_n$ as left variables and $Y_1, \ldots, Y_m$ as right variables. We desire to approximate $X_1, \ldots, X_n$ by $A_1, \ldots, A_n \in (\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} \otimes \mathbb{C})^{\mathrm{sa}}$, and $Y_1, \ldots, Y_m$ by $B_1, \ldots, B_m \in (\mathcal{M}_{d_1} \otimes \mathbb{C} \otimes \mathcal{M}_{d_2}^{\mathrm{op}})^{\mathrm{sa}}$. For $M, d \in \mathbb{N}$ and $R, \epsilon > 0$, let $\Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d_1, d_2, \epsilon)$ denote the set of all $(n+m)$-tuples $(A_1, \ldots, A_n, B_1, \ldots, B_m) \in ((M_{d_1} \otimes M_{d_2})^{\mathrm{sa}})^{n+m}$ such that $\|A_i\|, \|B_j\| \le R$ for all $1 \le i \le n$ and $1 \le j \le m$, and

$$\left|\varphi(Z_{k_1} \cdots Z_{k_p}) - \tau_{d_1} \otimes (\tau_{d_2} \circ m)(C_{k_1} \cdots C_{k_p})\right| < \epsilon$$

for all $i_1, \ldots, i_p \in \{1, \ldots, n+m\}$ and $1 \le p \le M$ where

$$Z_k = \begin{cases} X_k & \text{if } k \in \{1 \ldots, n\} \\ Y_{k-n} & \text{if } k \in \{n+1, \ldots, n+m\} \end{cases} \qquad \text{and} \qquad C_k = \begin{cases} L(A_k) & \text{if } k \in \{1 \ldots, n\} \\ R(B_{k-n}) & \text{if } k \in \{n+1, \ldots, n+m\} \end{cases}.$$

Definition 9.1. Using the above notation, if $\lambda_{d_1 d_2, n+m}$ denotes the Lebesgue measure on $(\mathcal{M}_{d_1 d_2}^{\mathrm{sa}})^{n+m}$, define

$$\begin{aligned}
\chi_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d_1, d_2, \epsilon) &= \log\left(\lambda_{d_1 d_2, n+m}\left(\Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d_1, d_2, \epsilon)\right)\right) \\
\chi_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, \epsilon) &= \limsup_{(d_1, d_2) \to \infty} \frac{1}{(d_1 d_2)^2} \chi_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d_1, d_2, \epsilon) \\
&\qquad + \frac{1}{2}(n+m)\log(d_1 d_2) \\
\chi_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) &= \inf\{\chi_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, \epsilon) \mid M \in \mathbb{N}, \epsilon > 0\} \\
\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) &= \sup_{R > 0} \chi_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m).
\end{aligned}$$

The quantity $\chi(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m) \in [-\infty, \infty)$ will be called the *microstate bi-free entropy of* $X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m$.

Remark 9.2. Of course, one must specify what is meant by $\limsup_{(d_1, d_2) \to \infty}$. There are many possible definitions (i.e. $d_1 + d_2 \ge K$ for sufficiently large K or $\min\{d_1, d_2\} \ge K$ for sufficiently large K). It may even be possible that if d_2 is sufficiently large, then there is no difference using $d_1 = 2$ or $d_1 > 2$. Of course, the real question is, "How do the sets $\Gamma_R(X_1, \ldots, X_n \sqcup Y_1, \ldots, Y_m; M, d_1, d_2, \epsilon)$ behave as d_1 and d_2 vary?"

Remark 9.3. We have seen above that the $d_1 = 1$ case may only model tracially bi-partite systems. However, adding the flexibility that $d_1 > 1$ appears to reduce these restrictions. Specifically, the only obvious restriction is that $\tau_{d_1} \otimes (\tau_{d_2} \circ m)$ is self-adjoint so we may only approximate distributions of left and right operators with respect to self-adjoint states. This is not a cumbersome restriction since we are already assuming that the operators and φ are self-adjoint.

Remark 9.4. We note that many results in this paper may be simply extended to apply to Definition 9.1; specifically Proposition 2.5, Proposition 2.6, Proposition 2.9, Proposition 3.1, and computations like those in Theorem 7.3. (Note, though, that the computations in Theorem 7.3 provided an upper bound on the entropy of a semicircular system, while the lower bound came from Theorem 2.7 which does not have an analogue in this setting.) One point of interest is there is less of a connection between these bi-free microstates and known free microstates (e.g. the argument used in the proof of Theorem 2.7 and Theorem 4.7 are no longer clear). Of course knowledge that lim sup can be replaced with lim inf in the definition of microstate free entropy (Definition 2.1) immediately implies the quantities in Definition 9.1 agree with those in Definition 2.1 when $n = 0$ or $m = 0$. However, when $n, m > 1$, it is not clear how to obtain these generalized bi-free microstates from free microstates of the left variables and microstates of the right variables.

Remark 9.5. We have made the choice to find approximants in the algebra $\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2}^{\mathrm{op}}$. One may consider replacing $\mathcal{M}_{d_1}$ by some other algebra – possibly of infinite dimension – to allow more flexibility. While it then becomes easier to find approximants, it becomes less clear how to treat the measure of the set of approximants. Nonetheless, [25] has argued that this is the correct constructs for the bi-free analogue of random matrices and thus the correct construct for bi-free microstates.

10. Open Questions

We conclude this paper with several important and intriguing questions raised in this paper in addition to the question of whether results in bi-free entropy may be applied to obtain results pertaining to von Neumann algebras.

To begin, as we are dealing with tracially bi-partite systems, one of the most natural questions is the following.

Question 10.1. *Given a tracially bi-partite family of operators $\left(\{X_i\}_{i=1}^n, \{Y_j\}_{j=1}^m\right)$, is there always a single-sided version as in Theorem 2.7 for which the stated inequality is an equality? If not, does taking a supremum over all systems which may stand on the left hand side lead to equality?*

One can produce examples by making a "poor choice" where the inequality is strict: for example, if (X, Y) is a pair of classically independent semi-circular operators (of non-zero variance), letting X' and Y' in the parlance of that theorem merely be X and Y themselves, the hypotheses of the theorem are satisfied and

$$-\infty = \chi(X, Y) < \chi(X \sqcup Y).$$

The answer to Question 10.1 is affirmative for the bi-free central limit distributions and for independent distributions. A general answer to Question 10.1 would be of interest as it directly relates the free and bi-free non-microstate entropies and could answer the following.

Question 10.2. *Is there an analogue of Proposition 3.1 where the transformation can intermingle left and right variables simultaneously?*

Of course Question 10.2 would be of interest as it would provide a greater flexibility in handling this entropy theory. However, there have been no instances in bi-free probability where right operators can intermingle with left operators and the resulting operators still behave like left operators.

Question 10.1 also relates to the following question.

Question 10.3. *Let (X, Y) be a bi-partite pair with joint distribution μ. Is there an integration formula involving μ to compute $\chi(X \sqcup Y)$?*

Question 10.3 arises from the integration formula established in [33]: if X is a self-adjoint operator with distribution μ, then the free entropy of X is

$$\chi(X) = \frac{1}{2}\log(2\pi) + \frac{3}{4} + \int_{\mathbb{R}}\int_{\mathbb{R}} \log|s - t|\, d\mu(s)\, d\mu(t).$$

To determine $\chi(X \sqcup Y)$ for a tracially bi-partite pair (X, Y), one must understand the microstates $(A, B) \in (\mathcal{M}_d(\mathbb{C})^{\mathrm{sa}})^2$ that are good approximants for (X, Y). If $(A', B') \in (\mathcal{M}_d(\mathbb{C})^{\mathrm{sa}})^2$ is another microstate that is a good approximation of (X, Y), then [32] implies that $\|A' - A\|_2$ and $\|B' - B\|_2$ are small. Therefore, for any $n, m \in \mathbb{N}$ and any unitary $U \in \mathcal{M}_d(\mathbb{C})$ we have that $\|A^n U^* B^m U - (A')^n U^* (B')^m U\|_2$ is small in norm (as the operator norm of microstates will be bounded by some R). Therefore, an understanding of microstates of the pair (X, Y) can be reduced to understanding the vector-valued random variable on the unitary group of $\mathcal{M}_d(\mathbb{C})$ defined by

$$U \mapsto (A^{n_1} U^* B^{m_1} U, \ldots, A^{n_k} U^* B^{m_k} U)$$

for every $k \in \mathbb{N}$ and every distinct $(n_1, m_1), \ldots, (n_k, m_k) \in \mathbb{N}^2$. When $k = 1$, the characteristic function of this random variable may be computable using the Harish-Chandra-Itzykson-Zuber integral formula, but deriving the necessary information from the characteristic function to describe the microstate bi-free entropy appears difficult.

Of course, an affirmative answer to both Questions 10.1 and 10.3 would enable the computation of the microstate free entropy of certain pairs of self-adjoint operators via an integration formula. Thus we do not expect an affirmative answer to both Questions 10.1 and 10.3.

Other natural questions pertaining to this microstate bi-free entropy are

Question 10.4. *Does Theorem 4.7 hold without the explicit assumption of the existence of a nice subsequence?*

which clearly will follow from

Question 10.5. *Can* $\limsup$ *be replaced with* $\liminf$ *in Definition 2.2?*

As these questions have been extremely difficult even in the free setting, we presume they will have equal if not greater difficulty in the bi-free setting. Another natural question to extend to the bi-free setting is the following.

Question 10.6. *When does the microstate bi-free entropy from* [9] *agree with the above non-microstate bi-free entropy for tracially bi-partite collections?*

In the free setting, [3] first showed that the microstate free entropy is always less than the non-microstate free entropy. Thus perhaps a good starting point would be a bi-free version of [3]. Much progress was made towards the converse in [10, 18].

Finally, as most of this paper deals only with the tracially bi-partite setting, we ask the following.

Question 10.7. *Are the quantities in Definition 9.1 finite when $n, m > 0$? Furthermore, does Definition 9.1 agree with Definition 2.2 for tracially bi-partite systems?*

An answer to Question 10.7 would enable us to extend the notion of microstate bi-free entropy to non-bi-partite systems thereby allowing a richer theory and demonstrating the notions in this paper are the correct extensions of microstate free entropy to the bi-free setting.

Acknowledgements

The authors would like to thank Dimitri Shlyakhtenko for many conversations and ideas that led to the computations in Section 7. They are also grateful to the anonymous referees, who offered numerous recommendations for improving the results of this paper, including the suggestion that we investigate Proposition 5.8, and a slight strengthening of Theorem 4.7.

References

[1] S.T. Belinschi and H. Bercovivi, *A Property of Free Entropy*, Pac. J. Math **211** (2003), no. 1, 35-40.
[2] S. Belinschi, H. Bercovici, Y. Gu, and P. Skoufranis, *Analytic subordination for bi-free convolution*, J. Funct. Anal. **275** (2018), no. 4, 926-966.
[3] P. Biane, M. Capitaine, and A. Guionnet, *Large deviation bounds for matrix Brownian motion*, Invent. math. **152** (2003), no. 2, 433-459.
[4] P. Biane and Y. Dabrowski, *Concavification of free entropy*, Adv. Math. **234** (2013), 667-696.
[5] P. Biane and D. Voiculescu, *A free probability analogue of the Wasserstein metric on the trace-state space*, GAFA **11** (2001), no. 6, 1125-1138.
[6] I. Charlesworth, *An alternating moment condition for bi-freeness*, Adv. Math. **346** (2019), no. 13, 546-568.
[7] I. Charlesworth, B. Nelson, and P. Skoufranis, *Combinatorics of Bi-Free Probability with Amalgamation*, Comm. Math. Phys. **338** (2015), no. 2, 801-847.
[8] I. Charlesworth, B. Nelson, and P. Skoufranis, *On two-faced families of non-commutative random variables*, Canad. J. Math. **67** (2015), no. 6, 1290-1325.
[9] I. Charlesworth and P. Skoufranis, *Analogues of Entropy in Bi-Free Probability Theory: Non-Microstate* (2017), 46 pp., available at `preprint`.
[10] Y. Dabrowski, *A Laplace Principle for Hermitian Brownian Motion and Free Entropy I: the convex functional case* (2017), 83 pp., available at `arXiv:1604.06420`.

[11] L. Ge, *Applications of free entropy to finite von Neumann algebras*, Am. J. Math (1997), 467-485.
[12] L. Ge, *Applications of free entropy to finite von Neumann algebras, II*, Ann. Math **147** (1998), no. 1, 143-157.
[13] B. Hayes, *A random matrix approach to the Peterson-Thom conjecture* (2020), 45 pp., available at `arXiv:2008.12287`.
[14] F. Hiai, T. Miyamoto, and Y. Ueda, *Orbital approach to microstate free entropy*, Int. J. Math **20** (2009), no. 2, 227-273.
[15] F. Hiai and D. Petz, *The semicircle law, free random variables and entropy* **77** (2006).
[16] F. Hiai and Y. Ueda, *Free transportation cost inequalities for noncommutative multi-variables*, Infinite Dimensional Analysis, Quantum Probability and Related Topics **9** (2006), no. 03, 391–412.
[17] H.-W. Huang and J.-C. Wang, *Analytic Aspects of Bi-Free Partial R-Transforms*, J. Funct. Anal. **271** (2016), no. 4, 922-957.
[18] D. Jekel, *An elementary approach to free entropy theory for convex potentials*, APDE **13** (2020), no. 8, 2289-2374.
[19] K. Jung, *Amenability, tubularity, and embeddings into* $\mathcal{R}^\omega$, Math. Ann. **228** (2007), 241-248.
[20] P. G. Lejeune-Dirichlet, *Sur une nouvelle Méthode pour la détermination des Intégrales multiples*, Journal de Mathématiques Pures et Appliquées **4** (1839), 164-168.
[21] Alexandru Nica and Roland Speicher, *Lectures on the combinatorics of free probability*, Vol. 13, Cambridge University Press, 2006.
[22] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361-400.
[23] C. Ramsey, *Faithfulness of bi-free product states*, Proc. Amer. Math. Soc. **146** (2018), 5279-5288.
[24] P. Skoufranis, *Independences and Partial R-Transforms in Bi-Free Probability*, Ann. Inst. Henri Poincaré Probab. Stat. **52** (2016), no. 3, 1437-1473.
[25] P. Skoufranis, *On Operator-Valued Bi-Free Distributions*, Adv. Math. **303** (2016), 638-715.
[26] P. Skoufranis, *Some Bi-Matrix Models for Bi-Free Limit Distributions*, Indiana Univ. Math. J. **66** (2017), no. 5, 1755-1795.
[27] D. M. Y. Sommerville, *An introduction to the geometry of n dimensions*, Dover Publications, Inc., New York, 1958. MR0100239
[28] Roland Speicher, *Multiplicative functions on the lattice of non-crossing partitions and free convolution*, Mathematische Annalen **298** (1994), no. 1, 611–628.
[29] Y. Ueda, *Orbital Free Entropy, Revisited*, Indiana Univ. Math. J. **63** (2014), no. 2, 551-577.
[30] Y. Ueda, *A remark on orbital free entropy*, Arch. Math. **108** (2017), 629-638.
[31] D. V. Voiculescu, *Limit laws for random matrices and free products*, Invent. Math. **104** (1991), no. 1, 201-220.
[32] D. V. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory, I*, Comm. Math. Phys. **155** (1993), no. 1, 71-92.
[33] D. V. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory, II*, Invent. Math. **118** (1994), no. 1, 411-440.
[34] D. V. 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), no. 1, 172-199.
[35] D. V. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory, IV: Maximum entropy and freeness*, Free Probability Theory **12** (1997), 293-302.
[36] D. V. Voiculescu, *A strengthened asymptotic freeness result for random matrices with applications to free entropy*, Internat. Math. Res. Notices **1998** (1998), no. 1, 41-63.
[37] D. V. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory, V: Noncommutative Hilbert transforms*, Invent. Math. **132** (1998), no. 1, 189-227.
[38] D. V. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory: VI. Liberation and Mutual Free Information*, Adv. Math. **146** (1999), no. 2, 101-166.
[39] D. V. Voiculescu, *Free probability for pairs of faces I*, Comm. Math. Phys. **332** (2014), 955-980.

Department of Mathematics, University of California, Berkeley, California, 94720, USA
Email address: `ilc@math.berkeley.edu`

Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada
Email address: `pskoufra@yorku.ca`