ANALYSIS & PDE

Volume 2 No. 2 2009

DIMITRI SHLYAKHTENKO

LOWER ESTIMATES ON MICROSTATES FREE ENTROPY DIMENSION

mathematical sciences publishers

ANALYSIS AND PDE
Vol. 2, No. 2, 2009

LOWER ESTIMATES ON MICROSTATES FREE ENTROPY DIMENSION

DIMITRI SHLYAKHTENKO

By proving that certain free stochastic differential equations with analytic coefficients have stationary solutions, we give a lower estimate on the microstates free entropy dimension of certain n-tuples $X_1, \dots, X_n$. In particular, we show that $\delta_0(X_1, \dots, X_n) \geq \dim_{M\bar{\otimes}M^o} V$, where $M = W^*(X_1, \dots, X_n)$ and $V = \{(\partial(X_1), \dots, \partial(X_n)) : \partial \in \mathscr{C}\}$ is the set of values of derivations $A = \mathbb{C}[X_1, \dots X_n] \to A \otimes A$ with the property that $\partial^*\partial(A) \subset A$. We show that for q sufficiently small (depending on n) and $X_1, \dots, X_n$ a q-semicircular family, $\delta_0(X_1, \dots, X_n) > 1$. In particular, for small q, q-deformed free group factors have no Cartan subalgebras. An essential tool in our analysis is a free analog of an inequality between Wasserstein distance and Fisher information introduced by Otto and Villani (and also studied in the free case by Biane and Voiculescu).

1. Introduction

We present in this paper a general technique for proving lower estimates for Voiculescu's microstates free entropy dimension δ_0. The free entropy dimension δ_0 was introduced in [Voiculescu 1994; 1996] and is a number associated to an n-tuple of self-adjoint elements $X_1, \dots, X_n$ in a tracial von Neumann algebra. This quantity has been used by various authors [Voiculescu 1996; Ge 1998; Ge and Shen 2002; Ştefan 2005; Jung 2007] to prove a number of very important results in von Neumann algebras. These results often take the form: *If* $\delta_0(X_1, \dots, X_n) > 1$, *then* $M = W^*(X_1, \dots, X_n)$ *cannot have certain decomposition properties* (for example, is non-Γ, has no Cartan subalgebras, is not a nontrivial tensor product and so on). For this reason, it is important to know if some given von Neumann algebra has a set of generators with the property that $\delta_0 > 1$. We prove that this is the case (for small values of q) for the "q-deformed free group factors" of [Bożejko and Speicher 1991].

Theorem 1. *For a fixed* N *and all* $|q| < (4N^3 + 2)^{-1}$, *the* q*-semicircular family* $X_1, \dots, X_N$ *satisfies* $\delta_0(X_1, \dots, X_N) > 1$ *and* $\delta_0(X_1, \dots, X_N) \geq N(1 - q^2N(1 - q^2N)^{-1})$.

The theorem applies for $|q| \leq 0.029$ if $N = 2$. Combined with the available results on free entropy dimension, we obtain that, in this range of values of q, the algebras $\Gamma_q(\mathbb{R}^N) = W^*(X_1, \dots, X_N)$ have no Cartan subalgebras (or, more generally, that $\Gamma_q(\mathbb{R}^N)$, when viewed as a bimodule over any of its abelian subalgebras, contain a coarse subbimodule). Theorem 1 also implies that these algebras are prime (this was already proved in [Shlyakhtenko 2004] using the techniques of [Ozawa 2004]).

The free entropy dimension δ_0 is closely related to L^2 Betti numbers [Connes and Shlyakhtenko 2005; Mineyev and Shlyakhtenko 2005] — more precisely, with Murray–von Neumann dimensions of

MSC2000: 46L54.
Keywords: free stochastic calculus, free probability, von Neumann algebras, q-semicircular elements.
Research supported by NSF grant DMS-0555680.

spaces of certain derivations. For example, the nonmicrostates free entropy dimension δ^* (which is the nonmicrostates "relative" of δ_0) is in many cases equal to L^2 Betti numbers of the underlying (nonclosed) algebra [Mineyev and Shlyakhtenko 2005; Shlyakhtenko 2006]. It is known that $\delta_0 \leq \delta^*$ and thus it is important to find lower estimates for δ_0 in terms of dimensions of spaces of derivations. To this end we prove.

Theorem 2. *Let* (A, τ) *be a finitely-generated algebra with a positive trace* τ *and generators* $X_1, \dots, X_N$, *and let* $\mathrm{Der}_c(A;\, A \otimes A)$ *denote the space of derivations from* A *to* $A \otimes A$ *which are* L^2 *closable and such that* $\partial^*\partial(X_j) \in A$. *Consider the* A,A*-bimodule*

$$V = \big\{(\delta(X_1), \dots, \delta(X_n)) : \delta \in \mathrm{Der}_c(A;\, A \otimes A)\big\} \subset (A \otimes A)^N.$$

Finally, assume that $M = W^*(A, \tau)$ *can be embedded into the ultrapower of the hyperfinite* II_1 *factor. Then*

$$\delta_0(X_1, \dots, X_n) \geq \dim_{M \bar{\otimes} M^o} \overline{V}^{L^2(A \otimes A, \tau \otimes \tau)^N}.$$

We actually prove Theorem 2 under a less restrictive assumption: we require that $\delta(X_j)$ and $\delta^*\delta(X_j)$ be "analytic" as functions of $X_1, \dots, X_N$; more precisely, there should exist noncommutative power series Ξ_j and ξ_j with sufficiently large multiradii of convergence so that $\delta(X_j) = \Xi_j(X_1, \dots, X_N)$ and $\delta^*\delta(X_j) = \xi_j(X_1, \dots, X_N)$; see Theorem 16 below for a precise statement.

This theorem is a rich source of lower estimates for δ_0. For example, if $T \in A \otimes A$, then

$$\delta : X \mapsto [X, T] = XT - TX$$

is a derivation in $\mathrm{Der}_c(A;\, A \otimes A)$. If $W^*(A)$ is diffuse, then the map

$$L^2(A \otimes A) \ni T \mapsto \big([T, X_1], \dots, [T, X_N]\big) \to L^2(A \otimes A)^N$$

is injective and thus the dimension over $M \bar{\otimes} M^o$ of its image is the same as the dimension of $L^2(A \otimes A)$, that is, 1. Hence $\dim_{M \bar{\otimes} M^o} \overline{V} \geq 1$ and so $\delta_0(X_1, \dots, X_n) \geq 1$ if $W^*(A)$ is R^ω embeddable ("hyperfinite monotonicity" in [Jung 2003b]).

If the two tuples $X_1, \dots, X_m$ and $X_{m+1}, \dots X_N$ are freely independent and each generates a diffuse von Neumann algebra, then for $T \in A \otimes A$ the derivation δ defined by $\delta(X_j) = [X_j, T]$ for $1 \leq j \leq m$ and $\delta(X_j) = 0$ for $m+1 \leq j \leq N$ is also in $\mathrm{Der}_c(A)$. Then one easily gets that $\dim_{M \bar{\otimes} M^o} \overline{V} > 1$ (indeed, V contains vectors of the form $\big([T, X_1], \dots, [T, X_m], 0, \dots, 0\big)$, $T \in L^2(A \otimes A)$, and so its closure is strictly larger than the closure of the set of all vectors $\big([T, X_1], \dots, [T, X_N]\big)$, $T \in L^2(A \otimes A)$). Thus $\delta_0(X_1, \dots, X_N) > 1$ if $W^*(A)$ is R^ω embeddable.

If $X_1, \dots, X_N$ are such that their conjugate variables [Voiculescu 1998] are polynomials, then the difference quotient derivations are in Der_c and thus $V = (A \otimes A)^N$, and so $\delta_0 = N$ (if $W^*(A)$ is R^ω embeddable).

In the case that $X_1, \dots, X_N$ are generators of the group algebra $\mathbb{C}\Gamma$ of a discrete group Γ,

$$\delta^*(X_1, \dots, X_N) = \beta_1^{(2)}(\Gamma) - \beta_0^{(2)}(\Gamma) + 1,$$

where $\beta_j^{(2)}$ are the L^2 Betti numbers of Γ (see [Lück 2002] for a definition). It is therefore natural to ask whether the same holds true for δ_0 instead of δ^* for some class of groups. If this is true, then knowing that

$\beta_1^{(2)}(\Gamma) \neq 0$ implies that $\delta_0 > 1$ and thus the group algebra has a variety of properties that we explained above (see also [Peterson 2009]).

It is clearly necessary for the equality $\delta_0 = \beta_1^{(2)} - \beta_0^{(2)} + 1$ that Γ can be embedded into the ultrapower of the hyperfinite II_1 factor (because otherwise δ_0 would be $-\infty$). In particular, one is tempted to conjecture that equality holds at least in the case when Γ is residually finite.

Theorem 2 implies a result like the one in [Brown et al. 2008]:

Theorem 3. *Assume that* Γ *is embeddable into the unitary group of the ultrapower of the hyperfinite* II_1 *factor. Then*

$$\delta_0(\Gamma) \geq \dim_{L(\Gamma)} \overline{\{c : \Gamma \to \mathbb{C}\Gamma \textit{ cocycle}\}}.$$

In particular, if Γ *belongs to the class of groups containing all groups with* $\beta_1^{(2)} = 0$ *and closed under amalgamated free products over finite subgroups, passage to finite index subgroups and finite extensions, then*

$$\delta_0(\Gamma) = \beta_1^{(2)}(\Gamma) - \beta_0^{(2)}(\Gamma) + 1.$$

Let us now describe the main idea of the present paper. Our main result states that if the von Neumann algebra $M = W^*(X_1, \dots, X_n)$ can be embedded into the ultrapower of the hyperfinite II_1 factor, then

$$\delta_0(X_1, \dots, X_n) \geq \dim_{M \bar{\otimes} M^o} V, \tag{1-1}$$

where $V = \overline{\{(\partial(X_1), \dots, \partial(X_n)) : \partial \in \mathscr{C}\}}^{L^2}$ and $\mathscr{C}$ is some class of derivations from the algebra of noncommutative polynomials $\mathbb{C}[X_1, \dots, X_n]$ to $L^2(M) \bar{\otimes} L^2(M^o)$, which will be made precise later.

The quantity $\delta_0(X_1, \dots, X_n)$ is, very roughly, a kind of Minkowski dimension ("relative" to R^ω) of the set $\mathscr{V}$ of embeddings of M into R^ω, the ultrapower of the hyperfinite II_1 factor (indeed, the set of such embeddings can be identified with the set of images under the embedding of the generators $X_1, \dots, X_n$, that is, with the set of microstates for $X_1, \dots, X_n$). On the other hand, $\dim_{M \bar{\otimes} M^o} V$ is a linear dimension (relative to $M \bar{\otimes} M^o$) of a certain vector space. If we could find an interpretation for V as a subspace of a "tangent space" to $\mathscr{V}$, then the inequality (1-1) takes the form of the inequality linking the Minkowski dimension of a manifold with the linear dimension of its tangent space. One natural proof of such an inequality would involve proving that a linear homomorphism of the tangent space to a manifold at some point can be exponentiated to a local diffeomorphism of a neighborhood of that point.

Thus an essential step in proving a lower inequality on free entropy dimension is to find an analog of such an exponential map.

This leads to the idea, given a matrix $Q_{ij} \in \left(L^2(M) \bar{\otimes} L^2(M^o)\right)^n$ of values of derivations (so that $Q_{ij} = \partial_j(X_i)$ for some n-tuple of derivation ∂_j belonging to our class $\mathscr{C}$), to try to associate to Q a one-parameter deformation α_t of a given embedding $\alpha = \alpha_0$ of M into R^ω. It turns out that there are two (related) ways to do this.

The first approach comes from the idea that we (at least in principle) know how to exponentiate derivations from an algebra to itself (the result should be a one-parameter automorphism group of the algebra). We thus try to extend $\partial = \partial_1 \oplus \cdots \oplus \partial_n$ to a derivation of a larger algebra $\mathscr{A} = \mathbb{C}[X_1, \dots, X_n, S_1, \dots, S_n]$, where $S_1, \dots, S_n$ are free from $X_1, \dots, X_n$ and form a free semicircular family. The key point is that the closure in $L^2(\mathscr{A})$ of $\operatorname{span}(MS_1M + \cdots + MS_nM)$ is isomorphic to $[L^2(M) \otimes L^2(M)]^n$. The inverse of this isomorphism takes an n-tuple $a = (a_1 \otimes b_1, \dots, a_n \otimes b_n)$ to $\sum a_j S_j b_j$, which we denote by

$a\#S$. We now define a new derivation $\tilde{\partial}$ of $\mathscr{A}$ with values in $L^2(\mathscr{A})$ by $\tilde{\partial}(X_j) = \partial(X_j)\#S$. To be able to exponentiate $\tilde{\partial}$, we need to make sure that it is antihermitian as an unbounded operator on $L^2(\mathscr{A})$, which naturally leads to the equation $\tilde{\partial}(S_j) = -\partial^*(\zeta_j)$, where $\zeta_j = (0, \dots, 1 \otimes 1, \dots, 0)$ (j-th entry nonzero). One can check that if ζ_j is in the domain of ∂^* for all j, then $\tilde{\partial}$ is a closable operator which has an antihermitian extension, and so it can be exponentiated to a one-parameter group of automorphisms α_t of $L^2(\mathscr{A})$. Unfortunately, unless we know more about the derivation ∂ (such as, for example, assuming that $\tilde{\partial}(\mathscr{A}) \subset \mathscr{A}$), we cannot prove that α_t takes $W^*(\mathscr{A})$ to $W^*(\mathscr{A})$. However, if this is the case, then we do get a one-parameter family of embeddings $\alpha_t|_M : M \to M * L(\mathbb{F}(n)) \subset R^\omega$. We explain this approach in more detail in the Appendix.

The second approach was suggested to us by A. Guionnet, to whom we are indebted for generously allowing us to publish it. The idea involves considering the free stochastic differential equation

$$dX_j(t) = \sum_i Q_{ij}(X_1(t), \dots, X_n(t))\#dS_i - \tfrac{1}{2}\xi_j(X_1(t), \dots, X_n(t)), \qquad X_j(0) = X_j, \tag{1-2}$$

where $\partial(X_j) = (Q_{1j}, \dots, Q_{nj}) \in \big(L^2(M) \mathbin{\bar{\otimes}} L^2(M^o)\big)^n$ and $\xi_j(X_1, \dots, X_n) = \partial^*\partial(X_j)$. One difficulty in even phrasing the problem is that it is not quite clear what is meant by Q_{ij} and ξ_j applied to their arguments (in the classical case, this would mean a function applied to the random variable $X(t)$). However, if this equation can be formulated and has a stationary solution $X(t)$ (namely one for which the law does not depend on t), then the map $\alpha_t : X_j \mapsto X_j(t^2)$ determines a one-parameter family of embeddings of the von Neumann algebra M into some other von Neumann algebra $\mathscr{M}$ (generated by all $X(t) : t \geq 0$). This can be carried out successfully if Q and ξ are sufficiently nice; this is this is the case, for example, when $X_1, \dots, X_n$ are q-semicircular variables, in which case Q and ξ can be taken to be analytic noncommutative power series.

Let us assume now that ∂ takes $\mathscr{B} = \mathbb{C}[X_1, \dots, X_n]$ to $\mathscr{B} \otimes \mathscr{B}^o$ and also $\partial^*(1 \otimes 1) \in \mathscr{B}$ (this is the case, for example, if $X_1, \dots, X_n$ have polynomial conjugate variables [Voiculescu 1998]). Then both approaches work to actually give one a stronger statement: one gets a one-parameter family of embeddings $\alpha_t : M \to R^\omega$ so that $\|\alpha_t(X_j) - (X_j + t\sum_i Q_{ij}\#S_i)\|_2 = O(t^2)$. Let us assume for the moment that $Q_{ij} = \delta_{ij} 1 \otimes 1$, so that our estimate reads

$$\|\alpha_t(X_j) - (X_j + tS_j)\|_2 = O(t^2). \tag{1-3}$$

An estimate of this kind was used as a crucial step by Otto and Villani in their work on the classical transportation cost inequality [Otto and Villani 2000, §4 Lemma 2]; a free version (for $n = 1$) is the key ingredient in the proof of free transportation cost inequality and free Wasserstein distance given in [Biane and Voiculescu 2001]. Indeed, since the law of $\alpha_t(X_j)$ is the same as X_j, one obtains after working out the error bounds an estimate on the noncommutative Wasserstein distance between the laws $\mu_{X_1,\dots,X_n}$ and $\mu_{X_1+tS_1,\dots,X_n+tS_n}$:

$$d_W(\mu_{X_1,\dots,X_n}, \mu_{X_1+tS_1,\dots,X_n+tS_n}) \leq \tfrac{1}{2}\Phi(X_1, \dots, X_n)^{1/2} t + O(t^2).$$

We now point out that this estimate is of direct relevance to a lower estimate on δ_0. Indeed, suppose that some n-tuple of $k \times k$ matrices $x_1, \dots, x_n$ has as its law approximately the law of $X_1, \dots, X_n$ (that is, $(x_1, \dots, x_n) \in \Gamma(X_1, \dots, X_n; k, l, \varepsilon)$ in the notation of [Voiculescu 1994]). Then (1-3) implies that by approximating $\alpha_t(X_j)$ with polynomials in $X_1, \dots, X_n, S_1, \dots, S_n$, one can find another n-tuple

$x_1', \dots, x_n'$ with almost the same law as $X_1, \dots, X_n$, and so that $\|x_j' - (x_j + ts_j)\| \le Ct^2$ (here $s_1, \dots, s_n$ are some matrices whose law is approximately that of $S_1, \dots, S_n$, and which are approximately free from $x_1, \dots, x_n$). But this means that if one moves along a line starting at $x_1, \dots, x_n$ in the direction of $s_1, \dots, s_n$, then the distance to the set $\Gamma(X_1, \dots, X_n; k, l, \varepsilon)$ grows quadratically. Thus this line is *tangent* to the set $\Gamma(X_1, \dots, X_n; k, l, \varepsilon)$. From this one can derive estimates relating the packing numbers of $\Gamma(X_1, \dots, X_n; k, l, \varepsilon)$ and $\Gamma(X_1 + tS_1, \dots, X_n + tS_n; k, l, \varepsilon)$ which can be converted into a lower estimate on δ_0.

In conclusion, it is worth pointing out that the main obstacle that we face in trying to extend the estimate (1-1) to larger classes of derivations is the question of existence of stationary solutions of (1-2) for more general classes of functions Q and ξ (and not, surprisingly enough, the "usual" difficulties in dealing with sets of microstates).

2. Existence of stationary solutions

2.1. *Free SDEs with analytic coefficients.* The main result of this section states that a free stochastic differential equation of the form

$$dX_t = \Xi \# dS_t - \tfrac{1}{2}\xi_t dt$$

where X_t is an N-tuple of random variables has a stationary solution, as long as the coefficients Ξ and ξ are analytic (that is, they are noncommutative power series with sufficient radii of convergence).

2.1.1. *Estimates on certain operators appearing in free Ito calculus.* Let f be a noncommutative power series in N variables. We denote by $c_f(n)$ the maximal modulus of a coefficient of a monomial of degree n in f. Thus if $f = \sum f_{i_1 \dots i_n} X_{i_1} \cdots X_{i_n}$, then $c_f(n) = \max_{i_1 \dots i_n} |f_{i_1 \dots i_n}|$. We also write

$$\phi_f(z) = \sum c_f(n) z^n.$$

Then $\phi_f(z)$ is a formal power series in z. If ρ is the radius of convergence of ϕ_f, we'll say that $R = \rho/N$ is the multiradius of convergence of f.

We also write

$$\|f\|_\rho = \sum_{n \ge 0} c_f(n) N^n \rho^n \in [0, +\infty].$$

Note that $\|f\|_\rho = \sup_{|z| \le N\rho} |\phi_f(z)|$ (since all of the coefficients in the power series $\phi_f(z)$ are real and positive).

We denote by $\mathcal{F}(R)$ the collection of all power series f for which the multiradius of convergence is at least R. In other words, we require $\|f\|_\rho < \infty$ for all $\rho < R$.

Note that $\mathcal{F}_R$ is a complete topological vector space when endowed with the topology such that $T_i \to T$ if and only if $\|T_i - T\|_\rho \to 0$ for all $\rho < R$.

Let Ψ be a noncommutative power series in N variables having the form

$$\sum f_{i_1, \dots, i_k; j_1, \dots, j_l} Y_{i_1} \cdots Y_{i_k} \otimes Y_{j_1} \cdots Y_{j_l}.$$

We call Ψ a formal noncommutative power series with values in $\mathbb{C}\langle Y_1, \dots, Y_N \rangle^{\otimes 2}$. We write $c_\Psi(m, n)$ the maximal modulus of a coefficient of a monomial of the form $Y_{i_1} \cdots Y_{i_m} \otimes Y_{j_1} \cdots Y_{j_n}$ in Ψ. We let

$\phi_\Psi(z, w) = \sum_{n,m} c_{\psi(m,n)} z^m w^n$. We put

$$\|\Psi\|_\rho = \sup_{|z|,|w|\le N\rho} |\phi_\Psi(z, w)| = \phi_\Psi(N\rho, N\rho) = \sum_{n\ge 0}\Bigl(\sum_{k+l=n} c_\Psi(k, l)\Bigr) N^n \rho^n \in [0, +\infty].$$

We denote by $\mathcal{F}'(R)$ the collection of all noncommutative power series for which $\|\Psi\|_\rho < \infty$ for all $\rho < R$.

It will be convenient to use the following notation. Let $\phi(z_1, \dots, z_n)$, $\psi(z_1, \dots, z_n)$ be two formal power series (in commuting variables). We say that $\phi \prec \psi$ if all coefficients in ϕ, ψ are real and positive, and for each $k_1, \dots, k_n$, the coefficient of $z_1^{k_1} \cdots z_n^{k_n}$ in ϕ is less than or equal to the corresponding coefficient in ψ.

If $\mathcal{M}$ is a unital Banach algebra, $Y_1, \dots, Y_N \in \mathcal{M}$ and $\|Y_j\| < \rho$ for all j, then $\|g(Y_1, \dots, Y_n)\| \le \|g\|_\rho$ whenever g is in any one of the spaces $\mathcal{F}(R)$, or $\mathcal{F}'(R)$ (here the norm $\|g(Y_1, \dots, Y_n)\|$ denotes the norm on $\mathcal{M}$ or on the projective tensor product $\mathcal{M}^{\otimes 2}$, as appropriate).

We now collect some facts about power series:

- Let $f, g \in \mathcal{F}(R)$. Then $\phi_{fg} \prec \phi_f \phi_g$. In particular, $fg \in \mathcal{F}(R)$ and $\|fg\|_\rho \le \|f\|_\rho \|g\|_\rho$.
- Let $f = \sum f_{i_1\dots i_n} X_{i_1} \cdots X_{i_n} \in \mathcal{F}(R)$ and denote by $\mathcal{D}_{ij} f$ the formal power series

$$\mathcal{D}_{ij} f = \sum_{i_1\dots i_n} \sum_{k<l} \delta_{i_k=i} \delta_{i_l=j} f_{i_1\dots i_n} X_{i_{k+1}} \cdots X_{i_{l-1}} \otimes X_{i_{l+1}} \cdots X_{i_n} X_{i_1} \cdots X_{i_{k-1}}.$$

  Since a monomial $X_{i_1} \cdots X_{i_k} \otimes X_{j_1} \cdots X_{j_r}$ could arise in the expression for $\mathcal{D}_{ij} f$ in at most $r + 1$ ways, $c_{\mathcal{D}_j f}(a, b) \le (b + 1) c_f(a + b + 2)$. Denote by $\hat{\phi}_f$ the power series

$$\hat{\phi}_f(z, w) = \sum_{n,m} (n + 1) c_f(n + m + 2) z^m w^n.$$

  Then $\phi_{\mathcal{D}_j f} \prec \hat{\phi}_f$. Since $\hat{\phi}_f(z, z) \prec \phi''_f(z)$, we conclude that

$$\|\mathcal{D}_{ij} f\|_\rho \le \sup_{|z|\le N\rho} |\phi''_f(z)|$$

  and in particular $\mathcal{D}_{ij} f \in \mathcal{F}'(R)$.
- Let $\Theta = \sum \Theta_{i_1\dots i_n; j_1\dots j_m} X_{i_1} \cdots X_{i_n} \otimes X_{j_1} \cdots X_{j_m} \in \mathcal{F}'(R)$, and let

$$\Psi = \sum \Psi_{i_1\dots i_n; j_1\dots j_m} X_{i_1} \dots X_{i_n} \otimes X_{j_1} \dots X_{j_m} \in \mathcal{F}'.$$

  Consider

$$\Psi \#_{in} \Theta = \sum \Psi_{t_1\dots t_a, s_1,\dots,s_b} \Theta_{i_1\dots i_n; j_1\dots j_m} X_{i_1} \cdots X_{i_n}\ X_{t_1} \cdots X_{t_a} \otimes X_{s_1} \cdots X_{s_b}\ X_{j_1} \cdots X_{j_m}.$$

  (In the simple case that $\Psi = A \otimes B$ and $\Theta = P \otimes Q$, where A, B, P, Q are monomials, we have $\Psi \#_{in} \Theta = PA \otimes BQ$, that is, $\#_{in}$ is the "inside" multiplication on $\mathcal{F}'(R)$). Then

$$c_{\Psi \#_{in} \Theta}(n, m) \le \sum_{k+l=n} \sum_{r+s=m} c_\Psi(k, r) c_\Theta(l, s),$$

and hence the coefficient of $z^n w^m$ in $\phi_{\Psi\#_{in}\Theta}(z, w)$ is dominated by the coefficient of $z^n w^m$ in $\phi_\Psi(z, w)\phi_\Theta(z, w)$. Consequently, $\phi_{\Psi\#_{in}\Theta} \prec \phi_\Psi\phi_\Theta$ and

$$\|\Psi\#_{in}\Theta\|_\rho \leq \|\Psi\|_\rho\|\Theta\|_\rho.$$

In particular, $\Psi\#_{in}\Theta \in \mathscr{F}'(R)$. Similar estimates and conclusion of course hold for the "outside" multiplication $\Psi\#_{out}\Theta$, defined by

$$\Psi\#_{out}\Theta = \sum \Psi_{s_1,\dots,s_b;t_1\dots t_a}\Theta_{i_1\dots i_n;j_1\dots j_m} X_{t_1}\cdots X_{t_a}\ X_{i_1}\cdots X_{i_n} \otimes X_{j_1}\cdots X_{j_m}\ X_{s_1}\cdots X_{s_b}.$$

In that case we get $\phi_{\Psi\#_{out}\Theta}(z, w) \prec \phi_\Psi(w, z)\phi_\Theta(z, w)$ and $\|\Psi\#_{out}\Theta\|_\rho \leq \|\Psi\|_\rho\|\Theta\|_\rho$.

- Let τ be a linear functional on the algebra of noncommutative polynomials in n variables, so that $|\tau(X_{i_1}\cdots X_{i_n})| \leq R_0^n$ for all n. Given $\Theta = \sum \Theta_{i_1\dots i_n;j_1\dots j_m} X_{i_1}\cdots X_{i_n} \otimes X_{j_1}\cdots X_{j_m} \in \mathscr{F}'(R)$, assume that $R_0 < R$ and consider the formal sum

$$(1\otimes\tau)(\Theta) = \sum_{n,i_1,\dots,i_n}\Biggl(\sum_{m,j_1,\dots,j_m}\Theta_{i_1\dots i_n;j_1\dots j_m}\tau(X_{j_1}\cdots X_{j_m})\Biggr)X_{i_1}\cdots X_{i_n}.$$

More precisely, we consider the formal power series in which the coefficient of $X_{i_1}\cdots X_{i_n}$ is given by the sum

$$\sum_{m,j_1,\dots,j_m}\Theta_{i_1\dots i_n;j_1\dots j_m}\tau(X_{j_1}\cdots X_{j_m}).$$

But since $|\tau(X_{j_1}\cdots X_{j_m})| \leq R_0^m$, this sum is bounded by the coefficient of z^n in the power series expansion of $\phi(z, NR_0)$ (as a function of z), and is convergent. Thus $\phi_{(1\otimes\tau)(\Theta)}(z) \prec \phi_\Theta(z, NR_0)$ and we readily see that $(1\otimes\tau)(\Theta)$ is well-defined, belongs to $\mathscr{F}(R)$, and moreover

$$\|(1\otimes\tau)(\Theta)\|_\rho \leq \|\Theta\|_\rho,$$

whenever $\rho > R_0$.

- Let $f = \sum f_{i_1\dots i_n} X_{i_1}\cdots X_{i_n} \in \mathscr{F}(R)$ and consider the j-th cyclic partial derivative [Voiculescu 1999; 2002b]

$$\mathscr{D}_j f = \sum_{i_1\dots i_n}\sum_{l=1}^{n}\delta_{i_l=j}X_{i_{l+1}}\dots X_{i_n}\cdot X_{i_1}\cdots X_{i_{l-1}}.$$

Then we see that $\phi_{\mathscr{D}_j f} \prec (\phi_f)'$ and $\mathscr{D}_j f \in \mathscr{F}(R)$.

We now combine these estimates:

Lemma 4. *Let τ as above be a linear functional on the space of noncommutative polynomials in N variables satisfying $\tau(X_{i_1}\cdots X_{i_n}) \leq R_0^n$. Let $R > R_0$ and assume that $\xi_j \in \mathscr{F}(R)$, $j = 1,\dots,N$, $\Psi = (\Psi_{ij}) \in M_{N\times N}\mathscr{F}'(R)$. For $f \in \mathscr{F}(R)$ let*

$$\mathscr{L}^{(\tau)}(f) = (1\otimes\tau)\Biggl(\sum_{ijk}\Psi_{jk}\#_{in}(\Psi_{ki}\#_{out}(\mathscr{D}_{ij}f))\Biggr) - \sum_j \tfrac{1}{2}\xi_j\mathscr{D}_j f.$$

Then $\mathcal{L}_j^{(\tau)}(f) \in \mathcal{F}(R)$ *and moreover for any* $R_0 < \rho < R$,

$$\|\mathcal{L}^{(\tau)}(f)\|_\rho \leq \sum_{ijk} \|\Psi_{jk}\|_\rho \|\Psi_{ki}\|_\rho \cdot \sup_{|z|\leq N\rho} |\phi_f''(z)| + \tfrac{1}{2}\sum_j \|\xi_j\|_\rho \sup_{|z|\leq N\rho} |\phi_f'|,$$

$$\phi_{\mathcal{L}^{(\tau)}(f)}(z) \prec \sum_{ijk} \phi_{\Psi_{jk}}(z, NR_0)\phi_{\Psi_{ki}}(NR_0, z)\hat{\phi}_f(z, NR_0) + \tfrac{1}{2}\sum_j \phi_{\xi_j}(z)\phi_f'(z),$$

where $\hat{\phi}_f(z,w) = \sum_{n,m}(n+1)c_f(n+m+2)z^m w^n$.

For ϕ a power series in $z, w_1, \dots, w_k$ with multiradius of convergence bigger than ρ and all coefficients of monomials nonnegative, let $\phi_{w_1,\dots,w_k}(z) = \phi(z, w_1, \dots, w_k)$. Set

$$Q\phi(z, w_1, \dots w_{k+1}) = \widehat{\phi_{w_1,\dots,w_k}}(z, w_{k+1}) \quad \text{and} \quad D\phi(z, w_1, \dots, w_k) = \partial_z^2\phi(z, w_1, \dots, w_k).$$

We note that $\hat{\phi}(z,z) \prec \phi''(z)$, and that Q, D and $\mathcal{D}$ are monotone for the ordering $\prec$. It follows that if κ_j, λ_j are some power series with radius of convergence bigger than ρ and positive coefficients, then for any $a_1, b_1, \dots, a_k, b_k \geq 0$ and any $R < \rho$,

$$\begin{aligned}\big[Q^{a_1}\kappa_1(z)D^{b_1}\lambda_1(z)\mu_1 Q^{a_2}\kappa_2(z)D^{b_2}\lambda_2(z)\cdots D^{b_k}\lambda_k\big]\,\big|_{z=w_1=\cdots=w_{\sum b_k}=R} \\ \leq \big[D^{a_1}\kappa_1(z)D^{b_1}\lambda_1(z)D^{a_2}\kappa_2(z)D^{b_2}\lambda_2(z)\cdots D^{b_k}\lambda_k\big]\,\big|_{z=w_1=\cdots=w_{\sum b_k}=R}.\end{aligned}$$

Now define

$$\hat{\mathcal{L}}\phi(z) = \sum_{ijk}\phi_{\psi_{jk}}(z, NR_0)\phi_{\Psi_{ki}}(NR_0, z)\phi''(z) + \tfrac{1}{2}\sum_j \phi_{\xi_j}(z)\phi'(z).$$

Then we have obtained the inequality

$$\phi_{\mathcal{L}^n f}(NR_0) \leq \hat{\mathcal{L}}^n\phi_f(NR_0),$$

which we record as:

Lemma 5. *Let* $\hat{\mathcal{L}}\phi(z) = \sum_{ijk}\phi_{\psi_{jk}}(z, NR_0)\phi_{\Psi_{ki}}(NR_0, z)\phi''(z) + \frac{1}{2}\sum_j \phi_{\xi_j}(z)\phi(z)$ *and let* τ *be a trace so that for any monomial* P, $|\tau(P)| < R_0^n$, $n = \deg P$. *Then*

$$|\tau(\mathcal{L}^n f)| \leq \hat{\mathcal{L}}^n\phi_f(NR_0).$$

2.1.2. *Analyticity of* $\partial^*\partial(X_j)$. Let us now assume that $\Xi = (\Xi_1, \dots, \Xi_N) \in \mathcal{F}'(R)$. Let $(X_1, \dots, X_N)$ be an N-tuple of self-adjoint operators in a tracial von Neumann algebra (M, τ) and assume that $\|X_j\| < R$ for all j. Let $\partial : L^2(M) \to L^2(M) \bar{\otimes} L^2(M)$ be the derivation densely defined on polynomials in $X_1, \dots, X_N$ by $\partial(X_j) = \Xi_j(X_1, \dots, X_N)$. We assume that $1 \otimes 1$ belongs to the domain of ∂^* and that there exists some $\zeta \in \mathcal{F}(R)$ so that $\partial^*(1 \otimes 1) = \zeta(X_1, \dots, X_N)$.

Lemma 6. *With the assumptions above, there exist* $\xi_j \in \mathcal{F}(R)$, $j = 1, \dots, N$, *so that*

$$\xi_j(X_1, \dots, X_N) = \partial^*\partial(X_j).$$

Proof. It follows from [Voiculescu 1998; Shlyakhtenko 1998] that under these assumptions, ∂ is closable. Moreover, for any a, b polynomials in $X_1, \dots, X_N$, $a \otimes b$ belongs to the domain of ∂^* and

$$\partial^*(a \otimes b) = a\zeta b + (1 \otimes \tau)[\partial(a)]b + a(\tau \otimes 1)[\partial(b)],$$

where $\zeta = \zeta(X_1, \dots, X_N) = \partial^*(1 \otimes 1)$.

Consider now formal power series in N variables having the form

$$\Theta = \sum \Theta_{i_1,\dots,i_k;\, j_1,\dots,j_l;t_1,\dots,t_r} Y_{i_1} \cdots Y_{i_k} \otimes Y_{j_1} \cdots Y_{j_l} \otimes Y_{t_1} \cdots Y_{t_r}.$$

We write $\phi_\Theta(z, w, v)$ for the power series whose coefficient of $z^m w^n v^k$ is equal to the maximum

$$\max\big\{|\Theta_{i_1,\dots,i_m;\, j_1,\dots,j_n;t_1,\dots,t_k}| : i_1, \dots, i_m, j_1, \dots, j_n, t_1, \dots, t_n \in \{1, \dots, N\}\big\}.$$

We denote by $\mathscr{F}''(R)$ the collection of all such power series for which ϕ_Θ has a multiradius of convergence at least NR.

Let $\mathscr{D}_1^{(s)} : \mathscr{F}'(R) \to \mathscr{F}''(R)$ be given by

$$\mathscr{D}_1^{(s)} \sum f_{i_1,\dots,i_k;\, j_1,\dots,j_l} Y_{i_1} \cdots Y_{i_k} \otimes Y_{j_1} \cdots Y_{j_l} = \sum f_{i_1,\dots,i_k;\, j_1,\dots,j_l} \sum_p \delta_{i_p=s} Y_{i_1} \cdots Y_{i_{p-1}} \otimes Y_{i_{p+1}} \cdots Y_{i_k} \otimes Y_{j_1} \cdots Y_{j_l}.$$

Then clearly $\phi_{\mathscr{D}_1^{(s)}(\Psi)}(z, z, w) \prec \partial_z \phi_\Psi(z, w)$ so that $\mathscr{D}_1^{(s)}\Psi$ indeed lies in $\mathscr{F}''(R)$ if $\Psi \in \mathscr{F}'(R)$.

Similarly, if we define for $\Psi \in \mathscr{F}'(R)$, $\Theta \in \mathscr{F}''(R)$

$$\Psi \#_{in}^{(1)} \Theta = \sum \Psi_{t_1\dots t_a, s_1,\dots,s_b} \Theta_{i_1\dots i_n;\, j_1\dots j_m;k_1\dots k_p} Y_{i_1} \cdots Y_{i_n}\ Y_{t_1} \cdots Y_{t_a} \otimes Y_{s_1} \cdots Y_{s_b}\ Y_{j_1} \cdots Y_{j_m} \otimes Y_{k_1} \cdots Y_{k_p},$$

then $\phi_{\Psi \#_{in}^{(1)} \Theta}(z, v, w) \prec \phi_\Psi(z, v) \phi_\Theta(z, v, w)$ and in particular $\Psi \#_{in}^{(1)} \Theta \in \mathscr{F}''(R)$. (Note that $\#_{in}^{(1)}$ corresponds to "multiplying around" the first tensor sign in Θ).

Finally, if τ is any linear functional so that $\tau(P) < R_0^{\deg P}$ for any monomial P and we put

$$M_2(\Psi) = \sum \Psi_{i_1,\dots,i_n;\, j_1,\dots,j_m;k_1,\dots k_p} Y_{i_1} \cdots Y_{i_n} \tau(Y_{j_1} \cdots Y_{j_m} Y_{k_1} \cdots Y_{k_p}),$$

then $\phi_{M_2(\Psi)}(z) \le \phi_\Psi(z, NR_0, NR_0)$ and in particular $M_2(\Psi) \in \mathscr{F}(R)$ once $\Psi \in \mathscr{F}''(R)$ and $R_0 < R$. In the foregoing, we'll use the trace τ of M as our functional.

So if we put

$$T_1\Theta = M_2\Big(\sum_s \Xi_s \#_{in}^{(1)} \mathscr{D}_1^{(s)}\Big),$$

then T_1 maps $\mathscr{F}'(R)$ into $\mathscr{F}(R)$.

Note that in the case that $\Theta = A \otimes B$, where A, B are monomials, $T_1\Theta = (1 \otimes \tau)(\partial(A)) \cdot B$.

One can similarly define $T_2 : \mathscr{F}'(R) \to \mathscr{F}(R)$; it will have the property that $T_2\Theta = A(\tau \otimes 1)(\partial(B))$.

Lastly, let $\zeta \in \mathscr{F}(R)$ and let $m : \mathscr{F}'(R) \to \mathscr{F}(R)$ be given by

$$m(\Theta) = \sum \Theta_{i_1,\dots,i_n;\, j_1,\dots,j_m} \zeta_{p_1,\dots,p_r} Y_{i_1} \cdots Y_{i_n} Y_{p_1,\dots p_r} Y_{j_1} \cdots Y_{j_m}.$$

Once again, $\phi_{m(\Theta)}(z) \prec \phi_\Theta(z, z)\phi_\zeta(z)$.

Let $Q(\Xi) = T_1(\Xi) + T_2(\Xi) + m(\Xi)$. We claim that $\xi = (Q(\Xi))(X_1, \dots, X_N) = \partial^*(\Xi(X_1, \dots, X_N))$.

Note that if Ξ_n is a partial sum of Ξ (say obtained as the sum of all monomials in Ξ having degree at most n), then $Q(\Xi_n)(X_1, \dots, X_N) = \partial^*(\Xi_n(X_1, \dots, X_N))$. Moreover, as $n \to \infty$, we have that $\Xi_n(X_1, \dots, X_N) \to \Xi(X_1, \dots, X_N)$ in L^2 and also $Q(\Xi_n)(X_1, \dots, X_N) \to Q(\Xi)(X_1, \dots, X_N)$ in L^2

(this can be seen by observing first that the coefficients of $Q_n(\Xi)$ converge to the coefficients of $Q(\Xi)$ and then approximating $Q(\Xi)$ and $Q(\Xi_n)$ by their partial sums).

Since ∂^* is closed, the claimed equality follows. □

2.1.3. *Existence of solutions.* Recall that a process $X_1^{(t)}, \dots, X_N^{(t)} \in (M, \tau)$ is called stationary if its law does not depend on t; that is, for any polynomial f in N noncommuting variables, $\tau\big(f(X_1^{(t)}, \dots, X_N^{(t)})\big)$ is constant.

Lemma 7. *Let $X_1^{(0)}, \dots, X_N^{(0)}$ be an N-tuple of noncommutative random variables, $R_0 > \max_j \|X_j^{(0)}\|$ and $R > R_0$. Let $\xi_j \in \mathcal{F}(R)$, $\Psi = (\Psi_{ij}) \in M_{N\times N}(\mathcal{F}'(R))$ so that $\Psi_{ij}(Z_1, \dots, Z_N)^* = \Psi_{ji}(Z_1, \dots, Z_N)$ for any self-adjoint $Z_1, \dots, Z_N$.*

Consider the free stochastic differential equation

$$dX_i(t) = \Psi(X_1(t), \dots, X_N(t))\#(dS_t^{(1)}, \dots, dS_t^{(N)}) - \tfrac{1}{2}\xi_i\big(X_1(t), \dots, X_N(t)\big)\,dt \tag{2-1}$$

with the initial condition $X_j(0) = X_j^{(0)}$, $j = 1, \dots, n$. Here $dS_t^{(1)}, \dots, dS_t^{(N)}$ is free Brownian motion, and for $Q_{kl} = \sum a_i^{kl} \otimes b_i^{kl} \in M \hat{\otimes} M$, and $Q = (Q_{kl}) \in M_{N\times N}(M \hat{\otimes} M)$, we write

$$Q\#(W_1, \dots, W_N) = \bigg(\sum_{ki} a_i^{1k} W_k b_i^{1k}, \dots, \sum a_i^{Nk} W b_i^{Nk}\bigg).$$

Let $A = W^*(X_1^{(0)}, \dots, X_N^{(0)})$ and let $\partial_j : L^2(A) \to L^2(A \bar{\otimes} A)$ be derivations densely defined on polynomials in $X_1^{(0)}, \dots, X_N^{(0)}$ and determined by

$$\partial_j(X_i^{(0)}) = \Xi_{ji}(X_1^{(0)}, \dots, X_N^{(0)}).$$

Assume that for all j, $\partial_i X_j^{(0)} \in$ domain ∂_i^* and that

$$\xi_j(X_1^{(0)}, \dots, X_N^{(0)}) = \sum_i \partial_i^* \partial_i (X_j^{(0)}).$$

Then there exists a $t_0 > 0$ and a stationary solution $X_j(t)$, $0 \le t < t_0$. This stationary solution satisfies $X_j(t) \in W^*\big(X_1, \dots, X_N, \{S_j(s) : 0 \le s \le t\}_{j=1}^N\big)$.

We note that in view of Lemma 6, we may instead assume that $1 \otimes 1 \in$ domain ∂_j^* and

$$\partial_j^*(1 \otimes 1) = \zeta_j(X_1^{(0)}, \dots, X_N^{(0)})$$

for some $\zeta_1, \dots, \zeta_N \in \mathcal{F}(R)$, since this assumption guarantees the existence of $\xi_j \in \mathcal{F}(R)$ satisfying the hypothesis of Lemma 7.

Proof. We note that, because Ψ and ξ are analytic, they are Lipschitz in their arguments.

Thus it follows from the standard Picard argument (see [Biane and Speicher 1998]) that a solution (with given initial conditions) exists, at least for all values of t lying in some small interval $[0, t_0)$, $t_0 > 0$. Now choose t_0 so that $\|X_j(t)\|_\infty \le R_0 < R$ for all $0 \le t < t_0$ (this is possible, since the solution to the SDE is locally norm-bounded).

Next, we note that if we adopt the notations of Lemma 4 and define for $f \in \mathscr{F}(R)$

$$\mathscr{L}^{(\tau)}(f) = \sum_{ijk}(1 \otimes \tau)(\Psi_{jk}\#_{in}(\Psi_{ki}\#_{out}(\mathscr{D}_{ij} f))) - \tfrac{1}{2}\sum_{j} \xi_j \mathscr{D}_j f,$$

then we have that $\mathscr{L}^{(\tau_t)} f \in \mathscr{F}(R)$ (here τ_t refers to the trace on $\mathbb{C}\langle X_1(t), \ldots, X_n(t)\rangle$ obtained by restricting the trace from the von Neumann algebra containing the process X_t for small values of t, that is, $\tau_t(P) = \tau(P(X_1(t), \ldots, X_n(t))))$. Ito calculus shows that for any $f \in \mathscr{F}(R)$,

$$\frac{d}{dt}\tau(f(X_1(t), \ldots, X_N(t)))\Big|_{t=s} = \tau_s((\mathscr{L}^{(\tau_s)} f)(X_1(s), \ldots, X_N(s))).$$

In particular, replacing f with $\mathscr{L}^{(\tau_t)} f$ and iterating gives us the equality

$$\frac{d^n}{dt^n}\tau(f(X_1(t), \ldots, X_N(t)))\Big|_{t=s} = \tau_s(((\mathscr{L}^{(\tau_s)})^n f)(X_1(s), \ldots, X_N(s))).$$

Since $\xi_j(X_1(0), \ldots, X_n(0)) = \sum_i \partial_i^* \partial_i(X_j(0))$,

$$\tau\big(\mathscr{L}^{(\tau_0)}\big(f(X_1(0), \ldots, X_N(0))\big)\big) = 0$$

for any $f \in \mathscr{F}(R)$. Applying this to f replaced with $\mathscr{L}^{(\tau_0)} f$ and iterating allows us to conclude that

$$\frac{d^n}{dt^n}\tau(f(X_1(t), \ldots, X_N(t)))\Big|_{t=0} = 0, \qquad n \geq 1.$$

Let as before

$$\hat{\mathscr{L}}\phi(z) = \sum_{ijk} \phi_{\Psi_{ik}}(z, NR_0)\phi_{\Psi_{jk}}(NR_0, z)\phi''(z) + \tfrac{1}{2}\sum_{j} \phi_{\xi_j}(z)\phi'(z) = \alpha(z)^2\phi''(z) + \beta(z)\phi'(z),$$

where $\beta(z)$ is complex-valued function and $\alpha(z)$ is a complex vector-valued analytic function, both defined on $\{z : |z| < NR\}$. Moreover, α and β are real for $z \in \mathbb{R}$.

Consider the partial differential equation

$$\partial_t \phi(x, t) = \hat{\mathscr{L}}\phi(x, t), \qquad \phi(x, 0) = \phi_f(x), \quad x \in \mathbb{R}.$$

The solution $\phi(x, t)$ will be real-analytic in time (at least for small values of t), because the equation is elliptic. By Lemma 5, we conclude that

$$\left|\partial_t^n \tau(f(X_1(t), \ldots, X_N(t)))\big|_{t=s}\right| \leq \left|(\hat{\mathscr{L}}^n \phi)(NR_0, s)\right| = \left|\partial_t^n \phi(NR_0, t)\big|_{t=s}\right|.$$

Hence, since all derivatives of $\tau(f(X_1(t)), \ldots, f(X_N(t)))$ vanish at zero,

$$\begin{aligned}
&\left|\tau(f(X_1(s), \ldots, X_N(s))) - \tau(f(X_1(0), \ldots, X_N(0)))\right| \\
&\qquad = \left|\int_0^s \cdots \int_0^s \partial_t^n \tau(f(X_1(t), \ldots, f(X_N(t))))\Big|_{t=r} (dr)^n\right| \\
&\qquad \leq C\int_0^s \cdots \int_0^s (\partial_t^n \phi(NR_0, t))\Big|_{t=r} \leq \phi(NR_0, s) - P_n(NR_0, s),
\end{aligned}$$

where P_n is the n-th Taylor polynomial of ϕ at zero. Since ϕ is real-analytic in s, the right-hand side of the equation goes to zero as $n \to \infty$ for s in some interval including zero. Thus the function $s \to \tau(f(X_1(s), \dots, X_N(s)))$ is constant and so the solution is stationary. □

We note that once the Equation (2-1) has a stationary solution on a small interval $[0, t_0)$, then it of course has a stationary solution for all time (since the same lemma applied to $X_{t_0/2}$ guarantees existence of the solution for up to $3t_0/2$ and so on). However, we will not need this here.

3. Otto–Villani type estimates

The main result of this section is an estimate on the noncommutative Biane–Voiculescu–Wasserstein distance between the law of an N-tuple of variables $X = X_1, \dots, X_N$ and the law of the N-tuple $X + \sqrt{t}\, Q\#S$, where $S = (S_1, \dots, S_N)$ is a free semicircular family, $Q \in M_{N\times N}\big(L^2(W^*(X_1, \dots, X_N)^{\otimes 2})\big)$ is a matrix, and for $Q_{ij} = \sum_k A_{ij}^{(k)} \otimes B_{ij}^{(k)}$, we denote by $Q\#S$ the N-tuple $(Y_1, \dots, Y_N)$ with

$$Y_i = \sum_j \sum_k A_{ij}^{(k)} S_j B_{ij}^{(k)}.$$

The sum defining Y_i is L^2 convergent; in fact, the L^2 norm of Y_i is the same as the L^2 norm of the element

$$\sum_j \sum_k A_{ij}^{(k)} \otimes B_{ij}^{(k)}.$$

The estimate on Wasserstein distance (Proposition 8) is obtained under the assumptions that a certain derivation, defined by $\partial(X_i) = (Q_{i1}, \dots, Q_{iN}) \in \big(L^2(W^*(X_1, \dots, X_N)^{\otimes 2}\big)^N$ is closable and satisfies certain further analyticity conditions (see below for more precise statements). Under such assumptions, the estimate states that

$$d_W(X, X + \sqrt{t}\, Q\#S) \le Ct.$$

The main use of this estimate will be to give a lower bound for the microstates free entropy dimension of $X_1, \dots, X_N$ (see Section 5).

3.1. *An Otto–Villani type estimate on Wasserstein distance via free SDEs.*

Proposition 8. *Let $\Xi \in M_{N\times N}(\mathscr{F}'(R))$, $M = W^*(X_1, \dots, X_N)$ and let $\partial_j : L^2(M) \to L^2(M \bar{\otimes} M)$ be derivations densely defined on polynomials in $X_1, \dots, X_N$ and determined by*

$$\partial_j(X_i) = \Xi_{ji}(X_1, \dots, X_N).$$

Assume that for all j, $1 \otimes 1 \in \mathrm{domain}\, \partial_i^*$ and that there exist $\zeta_1, \dots, \zeta_N \in \mathscr{F}(R)$ so that

$$\zeta_j(X_1, \dots, X_N) = \partial_j(1 \otimes 1), \quad j = 1, 2, \dots, N.$$

Then there exists a II_1 factor $\mathcal{M} \cong M * L(\mathbb{F}_\infty)$ and a $t_0 > 0$ so that for all $0 \le t < t_0$ there exists an embedding $\alpha_t : M = W^*(X_1, \dots, X_N) \to \mathcal{M}$ and a free $(0, 1)$-semicircular family $S_1, \dots, S_N \in \mathcal{M}$, free from M and satisfying the inequality

$$\big\| \alpha_t(X_j) - (X_j + \sqrt{t}\, \Xi(X_1, \dots, X_N)\#S) \big\|_2 \le Ct, \tag{3-1}$$

where C is a fixed constant. Furthermore, $\alpha_t(X_j) \in W^*(X_1, \dots, X_N, S_1, \dots, S_N, \{S'_j\}_{j=1}^{\infty})$, where $\{S'_j\}_{j=1}^{\infty}$ are a free semicircular family, free from $(X_1, \dots, X_N, S_1, \dots, S_N)$.

If A can be embedded into R^ω, so can $\mathcal{M}$.

In particular, the noncommutative Wasserstein distance of Biane–Voiculescu satisfies

$$d_W\big((X_j)_{j=1}^N,\ (X_j + \sqrt{t}\,\Xi(X_1, \dots, X_N)\#S)_{j=1}^N \big) \le Ct.$$

Proof. By Lemma 6, we can find $\xi_1, \dots, \xi_N \in \mathcal{F}(R)$ so that $\xi_j(X_1, \dots, X_N) = \partial^*\partial(X_j)$.

Let $\mathcal{M} = W^*\big(X_1, \dots, X_N, \{S_1(s), \dots, S_N(s) : 0 \le s \le t\}\big)$, where $S_j(t)$ is a free semicircular Brownian motion. Let $X_j(t)$ be a stationary solution to the SDE (2-1) (see Lemma 7). The map that takes a polynomial in $X_1, \dots, X_N$ to a polynomial in $X_1(t), \dots, X_N(t)$ preserves traces and so extends to an embedding $\alpha_t : M \to \mathcal{M}$. By the free Burkholder–Gundy inequality [Biane and Speicher 1998], it follows that for $0 \le t < t_0 < 1$

$$\|X_j(t) - X_j(0)\| \le C_1\sqrt{t} + C_2 t \le C_3\sqrt{t},$$

where $C_1 = \sup_{t<t_0} \|\Xi(X_1(t), \dots, X_N(t)\| < \infty$, $C_2 = \max_j \sup_{t<t_0} \|\xi_j(X_1, \dots, X_N(t)\|$.

Furthermore,

$$\begin{aligned} X_j(t) - X_j(0) &= \int_0^t \Xi(X_1(s), \dots, X_N(s))\#dS_j(s) - \int_0^t \xi_j(X_1(s), \dots, X_{Nn}(s))\,ds \\ &= \int_0^t \Xi(X_1(0), \dots, X_N(0))\#dS_j(s) \\ &\quad - \int_0^t \big[\Xi(X_1(0), \dots X_N(0)) - \Xi(X_1(s), \dots, X_N(s))\big]\#dS_j(s) \\ &\quad - \int_0^t \xi_j(X_1(s), \dots, X_N(s))ds. \end{aligned}$$

By the Lipschitz property of the coefficients of the SDE (2-1), we see that

$$\|\Xi(X_1(s), \dots, X_N(s)) - \Xi(X_1(0), \dots, X_N(0))\| \le K \max_j \|X_j(s) - X_j(0)\| \le K'\sqrt{s}.$$

Combining this with the estimate $\|\xi_j(X_1(t), \dots, X_N(t))\| < K''$ we may apply the free Burkholder–Gundy inequality to deduce that

$$\|X_j(t) - (X_j(0) + \Xi(X_1(0), \dots, X_N(0))\#S_j(t))\| \le \left|\int_0^t (K'\sqrt{s})^2 ds\right|^{1/2} + \left\|\int_0^t K''\,ds\right\| \le Ct.$$

Thus it is enough to notice that $\|\cdot\|_2 \le \|\cdot\|$ and to take $S_j = \frac{1}{\sqrt{t}} S_j(t)$, which is a $(0, 1)$ semicircular element.

If M is R^ω-embeddable, we may choose $\mathcal{M}$ to be R^ω-embeddable as well, since it can be chosen to be a free product of M and a free group factor.

Finally, note that $X_j(t) \in W^*\big(X_1, \dots, X_N, \{S_j(s) : 0 \le s \le t\}_{j=1}^N\big)$ by construction. But the algebra $W^*\big(\{S_j(s) : 0 \le s \le t\}\big)$ can be viewed as the algebra of the Free Gaussian functor applied to the space $L^2[0, 1]$, in such a way that $S_j(s) = S([0, s])$. Then $W^*\big(\{S_j(s) : 0 \le s \le t\}\big) \subset W^*\big(S_1, \dots, S_N, \{S'_k\}_{k \in I(j)}\big)$, where $\{S'_k : k \in I(j)\}$ are free semicircular elements corresponding to the completion of the singleton set $\{t^{-1/2}\chi_{[0,t]}\}$ to an orthonomal basis of $L^2[0, 1]$.

The estimate for the Wasserstein distance now follows if we note that the law of $(\alpha_t(X_j))_{j=1}^N$ is the same as that of $(X_j)_{j=1}^N$; thus $(X_j(t))_{j=1}^N \cup (X_j + \sqrt{t}\,\Xi\#S)_{j=1}^N$ is a particular $2N$-tuple with marginal distributions the same as those of $(X_j)_{j=1}^N$ and $(X_j + \sqrt{t}\,\Xi\#S)_{j=1}^N$, so that the estimate (3-1) becomes an estimate on the Wasserstein distance. □

Remark 9. Although we do not need this in the rest of the paper, we note that the estimate in Proposition 8 also holds in the operator norm.

We should mention that an estimate similar to the one in Proposition 8 was obtained by Biane and Voiculescu [2001] in the case $N = 1$ under the much less restrictive assumptions that $\Xi = 1 \otimes 1$ and $1 \otimes 1 \in \operatorname{domain} \partial^*$ (that is, the free Fisher information $\Phi^*(X)$ is finite). Setting $\Xi_{ij} = \delta_{ij} 1 \otimes 1$ we have proved an analog of their estimate (in the N-variable case), but under the very restrictive assumption that the conjugate variables $\partial^*(\Xi)$ are analytic functions in $X_1, \dots, X_N$. The main technical difficulty in removing this restriction lies in the question of existence of a stationary solution to (2-1) in the case of very general drifts ξ.

4. Applications to q-semicircular families

4.1. *Estimates on certain operators related to q-semicircular families.*

4.1.1. *Background on q-semicircular elements.* Let $H_{\mathbb{R}}$ be a finite-dimensional real Hilbert space, H its complexification $H = H_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}$, and let $F_q(H)$ be the q-deformed Fock space on H [Bożejko and Speicher 1991]. Thus

$$F_q(H) = \mathbb{C}\Omega \oplus \bigoplus_{n \geq 1} H^{\otimes n},$$

with the inner product given by

$$\langle \xi_1 \otimes \cdots \otimes \xi_n, \zeta_1 \otimes \cdots \otimes \zeta_m \rangle = \delta_{n=m} \sum_{\pi \in S_n} q^{i(\pi)} \prod_{j=1}^{n} \langle \xi_j, \zeta_{\pi(j)} \rangle,$$

where $i(\pi) = \#\{(i, j) : i < j \text{ and } \pi(i) > \pi(j)\}$.

We write HS for the space of Hilbert–Schmidt operators on $F_q(H)$. We denote by $\Xi \in HS$ the operator

$$\Xi = \sum q^n P_n,$$

where P_n is the orthogonal projection onto the subspace $H^{\otimes n} \subset F_q(H)$.

For $h \in H$, let $l(h) : F_q(h) \to F_q(H)$ be the creation operator, $l(h)(\xi_1 \otimes \cdots \otimes \xi_n) = h \otimes \xi_1 \otimes \cdots \otimes \xi_n$, and for $h \in H_{\mathbb{R}}$, let $s(h) = l(h) + l(h)^*$. We denote by M the von Neumann algebra $W^*(s(h) : h \in H_{\mathbb{R}})$. It is known [Ricard 2005; Śniady 2004] that M is a II_1 factor and that $\tau = \langle \cdot\Omega, \Omega \rangle$ is a faithful tracial state on M. Moreover, $F_q(H) = L^2(M, \tau)$ and $HS = L^2(M, \tau) \mathbin{\bar{\otimes}} L^2(M, \tau)$.

Fix an orthonormal basis $\{h_i\}_{i=1}^N \subset H_{\mathbb{R}}$ and let $X_i = s(h_i)$. Thus $M = W^*(X_1, \dots, X_N)$, $N = \dim H_{\mathbb{R}}$.

Lemma 10 [Shlyakhtenko 2004]. *For $j = 1, \dots, N$, let $\partial_j : \mathbb{C}[X_1, \dots, X_N] \to HS$ be the derivation given by $\partial_j(X_i) = \delta_{i=j}\Xi$. Let $\partial : \mathbb{C}[X_1, \dots, X_N] \to HS^N$ be given by $\partial = \partial_1 \oplus \cdots \oplus \partial_N$ and regard ∂ as an unbounded operator densely defined on $L^2(M)$. Then*:

(i) *∂ is closable.*

(ii) *If we denote by* Z_j *the vector* $0 \oplus \cdots \oplus P_\Omega \oplus \cdots \oplus 0 \in HS^N$ (*nonzero entry in* j*-th place*, P_Ω *is the orthogonal projection onto* $\mathbb{C}\Omega \in F_q(H)$), *then* Z_j *is in the domain of* ∂^* *and* $\partial^*(Z_j) = h_j$.

As a consequence of (ii), if we let ∂ be as in the lemma, we have $\xi_j = \partial^*(Z_j) \in \mathbb{C}[X_1, \cdots, X_N] \subset \mathcal{F}(R)$ for any R.

4.1.2. Ξ *as an analytic function of* $X_1, \ldots, X_n$. We now claim that for small values of q, the element $\Xi \in L^2(M)^{\otimes 2}$ defined in Lemma 10 can be thought of as an analytic function of the variables $X_1, \ldots, X_N$. Recall that $h_i \in H$ is a fixed orthonormal basis and $X_j = s(h_j)$, $j = 1, \ldots, N$ thus form a q-semicircular family.

Lemma 11. *Let* $W_{i_1,\ldots,i_n}$ *be noncommutative polynomials so that*

$$W_{i_1,\ldots,i_n}(X_1, \ldots, X_N)\Omega = h_{i_1} \otimes \cdots \otimes h_{i_n}.$$

Then the degree of $W_{i_1,\ldots,i_n}$ *is* n, *and the maximal absolute value* $c_k^{(n)}$ *of a coefficient of a monomial* $X_{j_1} \cdots X_{j_k}$, $k \le n$, *in* $W_{i_1,\ldots,i_n}$ *satisfies*

$$c_k^{(n)} \le 2^{n-k} \left(\frac{1}{1 - |q|} \right)^{n-k}.$$

Furthermore, $\|W_{i_1,\ldots,i_n}\|^2_{L^2(M)} \le 2^n (1 - |q|)^{-n}$.

Proof. Clearly, $c_n^{(n)} = 1$. Moreover (compare [Effros and Popa 2003])

$$W_{i_1,\ldots,i_n} = X_{i_1} W_{i_2,\ldots,i_n} - \sum_{j \ge 2} q^{j-2} \delta_{i_1 = i_j} W_{i_2,\ldots,\hat{i_j},\ldots,i_n}$$

(where $\hat{\ }$ denotes omission). So the degree of $W_{i_1,\ldots,i_n}$ is n and the coefficient c_n of a monomial of degree k in $W_{i_1,\ldots,i_n}$ is at most the sum of a coefficient of a degree $k-1$ monomial in $W_{i_2,\ldots,i_n}$ and $\sum_{j \ge 2} q^{j-2} |k_j|$, where k_j is a coefficient of a degree k monomial in $W_{i_2,\ldots,\hat{i}_j,\ldots,i_n}$. By induction, we see that

$$\begin{aligned}
c_k^{(n)} &\le c_{k-1}^{(n-1)} + \sum_{j \ge 2}^{n} |q|^{j-2} c_k^{(n-2)} \\
&\le 2^{n-k-2} \Big(\frac{1}{1-|q|} \Big)^{n-k} + 2^{n-k-2} \Big(\frac{1}{1-|q|} \Big)^{n-k-2} \sum_{j \ge 0} |q|^j \\
&= 2^{n-k-2} \left[\Big(\frac{1}{1-|q|} \Big)^{n-k} + \Big(\frac{1}{1-|q|} \Big)^{n-k-2} \frac{1}{1-|q|} \right] \\
&\le 2^{n-k-2} \cdot 2 \Big(\frac{1}{1-|q|} \Big)^{n-k} \le 2^{n-k} \Big(\frac{1}{1-|q|} \Big)^{n-k}.
\end{aligned}$$

as claimed.

The upper estimate on $\|W_{i_1,\ldots,i_n}\|^2_{L^2(M)}$ follows in a similar way. □

Lemma 12. *Let $\{\xi_k : k \in K\}$ be a finite set of vectors in an inner product space V. Let Γ be the matrix $\Gamma_{k,l} = \langle \xi_k, \xi_l \rangle$. Assume that Γ is invertible and let $B = \Gamma^{-1/2}$. Then the vectors*

$$\zeta_l = \sum_k B_{k,l}\xi_k$$

form an orthonormal basis for the span of $\{\xi_k : k \in K\}$. Moreover, if λ denotes the smallest eigenvalue of Γ, then $|B_{k,l}| \le \lambda^{-1/2}$ for each k, l.

Proof. We have, using the fact that B is symmetric and $B\Gamma B = I$: $\langle \zeta_l, \zeta_{l'} \rangle = \langle \sum_{k,k'} B_{k,l}\xi_k, B_{k',l'}\xi_{k'} \rangle = \sum_{k,k'} B_{k,l}B_{k',l'}\Gamma_{k,k'} = (B\Gamma B)_{l,l'} = \delta_{l=l'}$. □

Lemma 13. *There exist noncommutative polynomials $p_{i_1,\dots,i_n}$ in $X_1, \dots, X_N$ so that the vectors*

$$\{p_{i_1,\dots,i_n}(X_1, \dots, X_n)\Omega\}_{i_1,\dots,i_n=1}^N$$

are orthonormal and have the same span as $\{W_{i_1,\dots,i_n}\}_{i_1,\dots,i_n=1}^N$.

Moreover, these can be chosen so that $p_{i_1,\dots,i_n}$ is a polynomial of degree at most n and the coefficient of each degree k monomial in p is at most $(1-2|q|)^{-n/2}(2N)^n(1-|q|)^k 2^{-k}$.

Proof. Consider the inner product matrix

$$\Gamma_n = [\langle W_{i_1,\dots,i_n}, W_{j_1,\dots,j_n}\rangle]_{i_1,\dots,i_n,\ j_1,\dots,j_n=1}^N.$$

Dykema and Nica [1993, Lemma 3.1] proved that one has the following recursive formula for Γ_n. Consider an N^n-dimensional vector space W with orthonormal basis $e_{i_1,\dots,i_n}$, $i_1, \dots, i_n \in \{1, \dots, N\}$, and consider the unitary representation π_n of the symmetric group S_n given by $\sigma \cdot e_{i_1,\dots,i_n} = e_{i_{\sigma(1)},\dots,i_{\sigma(n)}}$. Denote by $(1 \to j)$ the action (via π_n) of the permutation that sends 1 to j, k to $k-1$ for $2 \le k \le j$, and l to l for $l > j$ on W. Let $M_n = \sum_{j=1}^n q^{j-1}(1 \to j) \in \mathrm{End}(W)$. Then Γ_1 is the identity $N \times N$ matrix, and

$$\Gamma_n = (1 \otimes \Gamma_{n-1})M_n,$$

where $1 \otimes \Gamma_n$ acts on the basis element $e_{j_1,\dots,j_n}$ by sending it to $\sum_{k_2,\dots,k_n}(\Gamma_{n-1})_{j_2,\dots,j_n,\ k_2,\dots,k_n} e_{j_1,k_2,\dots,k_n}$ and Γ acts on the basis elements by sending $e_{j_1,\dots,j_n}$ to $\sum_{k_1,\dots,k_n}(\Gamma_n)_{j_1,\dots,j_n,\ k_1,\dots,k_n} e_{k_1,\dots,k_n}$. They then proceeded to prove that the operator M_n is invertible and derive a bound for its inverse in the course of proving [Dykema and Nica 1993, Lemma 4.1]. Combining this bound and the recursive formula for Γ_n yields the following lower estimate for the smallest eigenvalue of Γ_n:

$$\begin{aligned} c_n &= \left(\frac{1}{1-|q|}\prod_{k=1}^{\infty}\frac{1-|q|^k}{1+|q|^k}\right)^n = \left(\frac{1}{1-|q|}\sum_{k=-\infty}^{\infty}(-1)^k|q|^{k^2}\right)^n \\ &\ge \left(\frac{1}{1-|q|}\left(1-\sum_{k\ge 0}|q|^{k^2}\right)\right)^n \ge \frac{1}{(1-|q|)^{n/2}}\left(1-\sum_{k\ge 1}|q|^k\right)^n \\ &\ge \left(\frac{1}{1-|q|}\left(1-\frac{|q|}{1-|q|}\right)\right)^n = \left(\frac{1-2|q|}{(1-|q|)^2}\right)^n. \end{aligned}$$

Thus if we set $B = \Gamma_n^{-1/2}$, then all entries of B are bounded from above by $c_n^{-1/2}$. Thus if we apply the previous lemma with $K = \{1, \dots, N\}^n$ to the vectors $\xi_{i_1,\dots,i_n} = W_{i_1,\dots,i_n}\Omega$, we obtain that the vectors

$$\zeta_i = \sum_{j \in K} B_{j,i}\xi_j, \quad i \in K$$

form an orthonormal basis for the subspace of the Fock space spanned by tensors of length n.

Now for $i = (i_1, \dots, i_n) \in K$, let

$$p_i(X_1, \dots, X_N) = \sum_{j \in K} B_{j,i} W_j(X_1, \dots, X_N).$$

Then $\zeta_i = p_i(X_1, \dots, X_N)\Omega$ are orthonormal and (because the vacuum vector is separating), the polynomials $\{p_i : i \in K\}$ have the same span as $\{W_i : i \in K\}$.

Furthermore, if a is the coefficient of a degree k monomial r in p_i, then a is a sum of at most N^n terms, each of the form (the coefficient of r in W_j)$B_{j,i}$. Using Lemma 11, we therefore obtain the estimate

$$|a| \le N^n c_n^{-1/2} 2^{n-k}(1-|q|)^{-(n-k)} = \left(\frac{2N}{(1-2|q|)^{1/2}}\right)^n 2^{-k}(1-|q|)^k. \qquad \square$$

We now use the terminology of Section 2.1.1 in dealing with noncommutative power series.

Let $R_0 = 2(1-|q|)^{-1} \ge 2(1-q)^{-1} \ge \|X_j\|$. Then if $\alpha > 1$, $p = p_{i_1,\dots,i_n}$ is as in Lemma 13, and ϕ_p is as in Section 2.1.1, then the coefficient of z^k, $k \le n$ in ϕ_p is bounded by

$$\left(\frac{2N}{(1-2|q|)^{1/2}}\right)^n R_0^{-k} \le \left(\frac{2N\alpha}{(1-2|q|)^{1/2}}\right)^n (\alpha N R_0)^{-k}.$$

In particular for any $\rho < \alpha R_0$,

$$\|p_{i_1,\dots,i_n}\|_\rho \le \left(\frac{2N\alpha}{(1-2|q|)^{1/2}}\right)^n \sum_{k=0}^n (\alpha N R_0)^{-k} N^k \rho^k \le \left(\frac{2N\alpha}{(1-2|q|)^{1/2}}\right)^n \frac{1}{1-\rho/(\alpha R_0)}.$$

Lemma 14. *Let q be such that $|q| < (4N^3+2)^{-1}$. Then:*

(a) *The formula*

$$\Xi(Y_1, \dots, Y_N) = \sum_n q^n \sum_{i_1,\dots,i_n} p_{i_1,\dots,i_n}(Y_1, \dots, Y_N) \otimes p_{i_1,\dots,i_n}(Y_1, \dots, Y_N)$$

defines a noncommutative power series with values in $\mathbb{C}\langle Y_1, \dots, Y_N\rangle^{\otimes 2}$ with radius of convergence strictly bigger than the norm of a q-semicircular element, $\|X_j\| \le 2(1-q)^{-1}$.

(b) *If $X_1, \dots, X_N$ are q-semicircular elements and Ξ is as in Lemma 10, then $\Xi = \Xi(X_1, \dots, X_N)$ (convergence in Hilbert–Schmidt norm, identifying HS with $L^2(M) \bar{\otimes} L^2(M)$).*

Proof. Clearly,

$$\|p_{i_1,\dots,i_n} \otimes p_{i_1,\dots,i_n}\|_\rho \le \|p_{i_1,\dots,i_n}\|_\rho^2 \le \left(\frac{2N\alpha}{(1-2|q|)^{1/2}}\right)^{2n} \frac{1}{(1-\rho/(\alpha R_0))^2} = K_\rho \left(\frac{4N^2\alpha^2}{1-2|q|}\right)^n$$

for any $\rho < \alpha R_0$, where $R_0 = 2(1-|q|)^{-1} \ge \|X_j\|$.

Thus

$$\|\Xi\|_\rho \le K_\rho \sum_n \left(\frac{4N^2\alpha^2}{1-2|q|}\right)^n |q|^n N^n \le K_\rho \sum_n \left(\frac{4N^3\alpha|q|}{1-2|q|}\right)^n,$$

which is finite as long as $\rho < \alpha R_0$ and the fraction in the sum in the right is less than 1. Thus as long as $4N^3|q| < 1-2|q|$, that is, $|q| < (4N^3+2)^{-1}$, we can choose some $\alpha > 1$ so that the series defining Ξ has a radius of convergence of at least $\alpha R_0 > \|X_j\|$.

For part (b), we note that because $\|\cdot\|_{L^2(M)} \le \|\cdot\|_M$ and because of the definition of the projective tensor product, we see that

$$\|\cdot\|_{HS} \le \|\cdot\|_{M\hat{\otimes}M}$$

on $M\hat{\otimes}M$. Thus convergence in the projective norm on $M\hat{\otimes}M$ guarantees convergence in Hilbert–Schmidt norm. Furthermore, by definition of orthogonal projection onto a space,

$$\Xi = \sum q^n P_n,$$

where $P_n = \sum_{i_1,\dots,i_n} p_{i_1,\dots,i_n} \otimes p_{i_1,\dots,i_n} = \Xi^{(n)}(X_1,\dots,X_N)$ are the partial sums of $\Xi(X_1,\dots,X_N)$ (here we again identify HS and $L^2 \bar{\otimes} L^2$). Hence $\Xi = \Xi(X_1,\dots,X_N)$. □

5. An estimate on free entropy dimension

We now show how an estimate of the form (1-3) can be used to prove a lower bound for the free entropy dimension δ_0.

Recall from [Voiculescu 1996; 1994] that if $X_1,\dots,X_n \in (M,\tau)$ is an n-tuple of self-adjoint elements, then the set of microstates $\Gamma_R(X_1,\dots,X_n; l,k,\varepsilon)$ is defined by

$$\Gamma_R(X_1,\dots,X_n; l,k,\varepsilon) = \Big\{(x_1,\dots,x_n) \in (M^{sa}_{k\times k})^n : \|x_j\| < R \text{ and } \big|\tau(p(X_1,\dots,X_n)) - (1/k)\operatorname{Tr}(p(x_1,\dots,x_n))\big| < \varepsilon \text{ for any monomial } p \text{ of degree } \le l\Big\}.$$

If R is omitted, the value $R=\infty$ is understood. The free entropy is defined by

$$\chi(X_1,\dots,X_n) = \sup_R \inf_{l,\varepsilon} \limsup_{k\to\infty} \left(\frac{1}{k^2}\log \operatorname{Vol}(\Gamma_R(X_1,\dots,X_n; l,k,\varepsilon)) + \frac{n}{2}\log k\right).$$

The set of microstates for $X_1,\dots,X_n$ in the presence of $Y_1,\dots,Y_m$ is defined by

$$\Gamma_R(X_1,\dots,X_n : Y_1,\dots,Y_m; l,k,\varepsilon) = \big\{(x_1,\dots,x_n) : \exists (y_1,\dots,y_m) \text{ s.t. } (x_1,\dots,x_n,y_1,\dots,y_m) \in \Gamma_R(X_1,\dots,X_n,Y_1,\dots,Y_m; l,k,\varepsilon)\big\}.$$

The corresponding free entropy in the presence is then defined as by

$$\chi(X_1,\dots,X_n : Y_1,\dots,Y_m) = \sup_R \inf_{l,\varepsilon} \limsup_{k\to\infty} \left(\frac{1}{k^2}\log \operatorname{Vol}(\Gamma_R(X_1,\dots,X_n : Y_1,\dots,Y_m; l,k,\varepsilon)) + \frac{n}{2}\log k\right).$$

The $\sup_R$ is attained [Belinschi and Bercovici 2003]; in fact, once $R > \max_{i,j}\{\|X_i\|, \|Y_j\|\}$, we have

$$\chi(X_1,\dots,X_n : Y_1,\dots,Y_m) = \chi_R(X_1,\dots,X_n : Y_1,\dots,Y_m).$$

Finally, the free entropy dimension δ_0 is defined by

$$\delta_0(X_1,\dots,X_n) = n + \limsup_{t\to 0} \frac{\chi(X_1+\sqrt{t}S_1,\dots,X_n+\sqrt{t}S_n : S_1,\dots,S_n)}{|\log t|},$$

where $S_1,\dots,S_n$ are a free semicircular family, free from $X_1,\dots,X_n$. Equivalently [Jung 2003a] one sets

$$\mathbb{K}_\delta(X_1,\dots,X_n) = \inf_{\varepsilon,l} \limsup_{k\to\infty} \frac{1}{k^2} \log K_\delta(\Gamma_\infty(X_1,\dots,X_n;k,l,\varepsilon)),$$

where $K_\delta(X)$ is the covering number of a set X (the minimal number of δ-balls needed to cover X). Then

$$\delta_0(X_1,\dots,X_n) = \limsup_{t\to 0} \frac{\mathbb{K}_t(X_1,\dots,X_n)}{|\log t|}.$$

Lemma 15. *Assume that $X_1,\dots,X_n \in (M,\tau)$, $T_{jk} \in W^*(X_1,\dots,X_n)\bar\otimes W^*(X_1,\dots,X_n)^{op}$ are given. Set $S_j^T = \sum_k T_{jk}\#S_k$. Let $\eta = \dim_{M\bar\otimes M^o}\Big(\overline{\operatorname{span} MS_1^T M + \cdots + MS_n^T M}^{L^2(M\bar\otimes M^o)}\Big)$.*

Then there exists a constant K depending only on T so that for all $R>0$, $\alpha>0$, $t>0$, there are $\varepsilon'>0$, $l'>0$, and $k'>0$ so that for all $0<\varepsilon<\varepsilon'$, $k>k'$, and $l>l'$, and any $(x_1,\dots,x_n)\in\Gamma(X_1,\dots,X_n;k,l,\varepsilon)$ the set

$$\begin{aligned}\Gamma_R(tS_1^{I-T},\dots,tS_n^{I-T} \mid (x_1,\dots,x_n) : S_1,\dots,S_n;k,l,\varepsilon) = \\ \big\{(y_1,\dots,y_n) : \exists(s_1,\dots,s_n) \text{ s.t. } (y_1,\dots,y_n,x_1,\dots,x_n,s_1,\dots,s_n) \in \\ \Gamma_R(tS_1^{I-T},\dots,tS_n^{I-T},X_1,\dots,X_n,S_1,\dots,S_n;k,l,\varepsilon)\big\}\end{aligned}$$

can be covered by $(K/t)^{(n-\eta+\alpha)k^2}$ balls of radius t^2.

Proof. By considering the diffeomorphism of $(M_{k\times k}^{sa})^n$ given by $(a_1,\dots,a_n)\mapsto((1/t)a_1,\dots,(1/t)a_n)$, we may reduce the statement to showing that the set

$$\Gamma_R(S_1^{I-T},\dots,S_n^{I-T} \mid (x_1,\dots,x_n) : S_1,\dots S_n;k,l,\varepsilon)$$

can be covered by $(C/t)^{(n-\eta+\alpha)k^2}$ balls of radius t.

Note that η is the Murray–von Neumann dimension over $M\bar\otimes M^o$ of the image of the map

$$(\zeta_1,\dots,\zeta_n)\mapsto(\zeta_1^T,\dots,\zeta_n^T),$$

where $\zeta_j \in L^2(M)\bar\otimes L^2(M)$, $M = W^*(X_1,\dots,X_n)$. Thus if we view T as a matrix in $M_{n\times n}(M\bar\otimes M^o)$, then $\tau\otimes\tau\otimes\operatorname{Tr}(E_{\{0\}}((I-T)^*(I-T))) = \eta$ (here E_X denotes the spectral projection corresponding to the set $X\subset\mathbb{R}$).

Fix $\alpha>0$.

Then there exists $Q\in M_{n\times n}(\mathbb{C}[X_1,\dots,X_n]^{\otimes 2})$ depending only on t so that $\|Q_{ij}-(I-T)_{ij}\|_2 < t/(2n)$ (here we view Q as a matrix whose entries are noncommutative functions in n indeterminates; the entries of Q are in the space $\mathscr{F}'(\infty)$ in the notation of Section 2.1.1).

Set $S_j^Q = \sum_k Q_{jk} \# S_k$. Then

$$\| S_j^{Q(X_1,\dots,X_n)} - S_j^{I-T} \| < \frac{t}{2}.$$

In particular, $\| S_j^{Q(X_1,\dots,X_n)} - S_j^{I-T} \|_2 < t/2$. We may moreover choose Q (again, depending only on t) so that

$$\tau \otimes \tau \otimes \mathrm{Tr}(E_{[0,t/2[}(Q^*Q)^{1/2}(X_1,\dots,X_n)) \geq \tau \otimes \tau \otimes \mathrm{Tr}(E_{\{0\}}(I-T)^*(I-T)) = \eta - \tfrac{1}{2}\alpha.$$

Thus for l sufficiently large and $\varepsilon > 0$ sufficiently small, we have that if

$$(y_1,\dots,y_n) \in \Gamma_R(S_1^{I-T},\dots,S_n^{I-T} | (x_1,\dots,x_n) : S_1,\dots,S_n; k, l, \varepsilon),$$

then there exist $s_1,\dots,s_n$ such that

$$(s_1,\dots,s_n,x_1,\dots,x_n) \in \Gamma_R(S_1,\dots,S_n,X_1,\dots,X_n; k, l, \varepsilon)$$

and

$$\| s_j^{Q(x_1,\dots,x_n)} - y_j \|_2 < t.$$

By approximating the characteristic function $\chi_{[0,t/2]}$ with polynomials on the interval $[0, \| Q(x_1,\dots,x_n) \|]$ (which is compact, since $\|x_j\| < R$), we may moreover assume that l is large enough and ε is small enough that

$$\frac{1}{k^2} \mathrm{Tr} \otimes \mathrm{Tr} \otimes \mathrm{Tr}\big(E_{[0,t/2]}(Q^*Q)^{1/2}(x_1,\dots,x_n)\big) \geq \eta - \alpha.$$

Denote by ϕ the map

$$(s_1,\dots,s_n) \mapsto \big(s_1^{Q(x_1,\dots,x_n)},\dots,s_n^{Q(x_1,\dots,x_n)}\big).$$

Let $R_1 = \max_j \| S_j^{I-T} \|_2 + 1$. Assume that $\varepsilon < 1$. Then $\phi : (M_{k\times k}^{sa})^n \to (M_{k\times k}^{sa})^n$ is a linear map, and since $\|s_j\|_2^2 \leq 1 + \varepsilon < 2$, we have the inclusion

$$\Gamma_R(S_1^{I-T},\dots,S_n^{I-T} | (x_1,\dots,x_n) : S_1,\dots,S_n; k, l, \varepsilon) \subset N_t(\phi(B(2)) \cap B(R_1)),$$

where $B(R)$ the a ball of radius R in $(M_{k\times k}^{sa})^n$ (endowed with the L^2 norm) and N_t denotes a t-neighborhood.

The matrix of ϕ is precisely the matrix $Q(x_1,\dots,x_n) \in M_{n\times n}(M_{k\times k})^{\otimes 2}$.

Let β be such that $\beta n k^2$ eigenvalues of $(\phi^*\phi)^{1/2}$ are less than R_0. Then the t-covering number of $\phi(B(2)) \cap B(R_1)$ is at most

$$\left(\frac{R_1}{t}\right)^{(1-\beta)nk^2} \left(\frac{2R_0}{t}\right)^{\beta n k^2}.$$

Let $R_0 = t/2$, so $\beta = (\eta - \alpha)/n$. We conclude that the t-covering number of

$$\Gamma_R\big(S_1^{I-T},\dots,S_n^{I-T} \mid (x_1,\dots,x_n) : S_1,\dots,S_n; k, l, \varepsilon\big)$$

is at most $(K/t)^{(n-\eta+\alpha)k^2}$, for some constant K depending only on R_1, which itself depends only on T. $\square$

Theorem 16. *Assume that $X_1, \dots, X_n \in (M, \tau)$, $S_1, \dots, S_n, \{S_j : j \in J\}$ is a free semicircular family, free from M, $T_{jk} \in W^*(X_1, \dots, X_n) \bar{\otimes} W^*(X_1, \dots, X_n)^{op}$ are given, and that for each $t > 0$ there exist $Y_j^{(t)} \in W^*(X_1, \dots, X_n, S_1, \dots, S_n, \{S'_j\}_{j\in J})$ so that:*

- *the joint law of $(Y_1^{(t)}, \dots, Y_n^{(t)})$ is the same as that of $(X_1, \dots, X_n)$,*
- *if we set $S_j^T = \sum_k T_{jk}\#S_k$ and $Z_j^{(t)} = X_j + tS_j^T$, then for some $t_0 > 0$ and some constant $C < \infty$ independent of t, we have $\|Z_j^{(t)} - Y_j^{(t)}\|_2 \le Ct^2$ for all $t < t_0$.*

Let $M = W^*(X_1, \dots, X_n)$ and let

$$\eta = \dim_{M\bar{\otimes}M^o}\Big(\overline{\operatorname{span} MS_1^T M + \cdots + MS_n^T M}^{L^2}\Big).$$

Assume finally that $W^*(X_1, \dots, X_n)$ embeds into R^ω. Then $\delta_0(X_1, \dots, X_n) \ge \eta$.

Proof. Let $T : (M \bar{\otimes} M^o)^n \to (M \bar{\otimes} M^o)^n$ be the linear map given by

$$T(Y_1, \dots, Y_n) = \Big(\sum_k T_{1k}\#Y_k, \dots, \sum_k T_{nk}\#Y_k\Big)$$

(here, as before, we identify $(M \bar{\otimes} M^o)^n$ with the linear span of $MS_1M + \cdots + MS_nM$ via the map $(T_1, \dots, T_n) \mapsto (S^{T_1}, \dots, S^{T_n})$). Then η is the Murray–von Neumann dimension of the image of T, and consequently

$$\eta = n - \dim_{M\bar{\otimes}M^o} \ker T.$$

Let t be fixed.

Since $Y_j^{(t)}$ can be approximated by noncommutative polynomials in $X_1, \dots, X_n$, $S_1, \dots, S_n$ and $\{S'_j : j \in J\}$, for any k_0, ε_0, l_0 sufficiently large we may find $k > k_0$, $l > l_0$, $\varepsilon < \varepsilon_0$ and $J_0 \subset J$ finite so that whenever

$$(z_1, \dots, z_n) \in \Gamma_R\big(X_1 + tS_1^T, \dots, X_n + tS_n^T : S_1, \dots, S_n, \{S'_j\}_{j\in J_0}; k, l, \varepsilon\big),$$

there exists

$$(y_1, \dots, y_n) \in \Gamma_R(X_1, \dots, X_n; k, l_0, \varepsilon_0)$$

so that

$$\|y_j - z_j\|_2 \le Ct^2. \tag{5-1}$$

For a set $X \subset (M_{k\times k}^{sa})^n$ we'll write K_r for its covering number by balls of radius r.

Consider a covering of $\Gamma_R(X_1 + tS_1, \dots, X_n + tS_n : S, \dots, S_n, \{S'_j\}_{j\in J_0}; k, l, \varepsilon)$ by balls of radius $(C + 2)t^2$ constructed as follows.

First, let $(B_\alpha)_{\alpha\in I}$ be a covering of $\Gamma_R\big(X_1 + tS_1^T, \dots, X_n + tS_n^T : S_1, \dots, S_n, \{S'_j\}_{j\in J_0}; k, l_0, \varepsilon_0\big)$ by balls of radius $(C + 1)t^2$. Because of (5-1), we may assume that

$$|I| \le K_{t^2}(\Gamma_R(X_1, \dots, X_n; k, l, \varepsilon)).$$

Next, for each $\alpha \in I$, let $(x_1^{(\alpha)}, \dots, x_n^{(\alpha)}) \in B_\alpha$ be the center of B_α. Consider a covering $(C_\beta^{(\alpha)} : \beta \in J_\alpha)$ of $\Gamma_R\big(tS_1^{I-T}, \dots, tS_n^{I-T} \mid (x_1^{(\alpha)}, \dots, x_n^{(\alpha)}) : S_1, \dots, S_n; k, l, \varepsilon\big)$ by balls of radius t^2. By Lemma 15, this covering can be chosen to contain $|J_\alpha| \leq (K/t)^{n-\eta'}$ balls, for any $\eta' < \eta$. Thus the sets

$$(B_\alpha + C_\beta^{(\alpha)} : \alpha \in I, \beta \in J_\alpha),$$

each of which is contained in a ball of radius at most $(C+2)t^2$, cover the set

$$\Gamma_R\big(X_1 + tS_1, \dots, X_n + tS_n : S_1, \dots, S_n; k, l_0, \varepsilon_0\big).$$

The cardinality of this covering is at most

$$f(t^2, k) \leq |I| \cdot \sup_\alpha |J_\alpha| \leq K_{t^2}(\Gamma_R(X_1, \dots, X_n; k, l, \varepsilon) \cdot (Kt)^{\eta'-n}.$$

It follows that if we denote by $V(R, d)$ the volume of a ball of radius R in $\mathbb{R}^d$, we find that

$$\operatorname{Vol}(\Gamma_R(X_1 + tS_1, \dots, X_n + tS_n : S_1, \dots, S_n, \{S_j'\}_{j\in J_0})) \leq f(t^2, k) \cdot V((C+2)t^2, nk^2),$$

so that if we denote by $\mathbb{K}_{t^2}(X_1, \dots, X_n)$ the expression

$$\inf_{\varepsilon,l} \limsup_{k\to\infty} \frac{1}{k^2} \log K_{t^2}(\Gamma(X_1, \dots, X_n; k, l, \varepsilon))$$

and set $C' = \log(C+2)$, we obtain the inequality

$$\begin{aligned}
\inf_{\epsilon,l} \limsup_{k\to\infty} \frac{1}{k^2} &\log \operatorname{Vol}\Gamma_R(X_1 + tS_1, \dots, X_n + tS_n : S_1, \dots, S_n, \{S_j'\}_{j\in J_0}; k, l, \varepsilon) \\
&\leq \limsup\nolimits_{k\to\infty} \log f(t^2, k) + 2n \log t + \log(C+2) \\
&\leq \mathbb{K}_{t^2}(X_1, \dots, X_n) + (\eta' - n) \log Kt + 2n \log t + C' \\
&= \mathbb{K}_{t^2}(X_1, \dots, X_n) + (\eta' + n) \log t + (\eta' - n) \log K + C'.
\end{aligned}$$

By the freeness of $\{S_j'\}_{j\in J}$ and $\{S_1, \dots, S_n, X_1, \dots, X_n\}$, the lim sup on the right-hand side remains the same if we take $J_0 = \varnothing$. Thus

$$\chi_R(X_1 + tS_1, \dots, X_n + tS_n : S_1, \dots, S_n) \leq \mathbb{K}_{t^2}(X_1, \dots, X_n) + (\eta' + n) \log t + C''.$$

If we divide both sides by $|\log t|$ and add n to both sides of the resulting inequality, we obtain

$$\begin{aligned}
n + \frac{\chi_R(X_1 + tS_1, \dots, X_n + tS_n : S_1, \dots, S_n)}{|\log t|} &\leq \frac{\mathbb{K}_{t^2}(X_1, \dots, X_n)}{|\log t|} + (\eta' + n)\frac{\log t}{|\log t|} + n \\
&= 2\frac{\mathbb{K}_{t^2}(X_1, \dots, X_n)}{|\log t^2|} + (\eta' + n)\frac{\log t}{|\log t|} + n.
\end{aligned}$$

Taking $\sup_R$ and $\limsup_{t\to 0}$ and noticing that $\log t < 0$ for $t < 1$, we get the inequality

$$\delta_0(X_1, \dots, X_n) \leq 2\delta_0(X_1, \dots, X_n) - (\eta + n) + n = 2\delta_0(X_1, \dots, X_n) - \eta'.$$

Solving this inequality for $\delta_0(X_1, \dots, X_n)$ gives finally

$$\delta_0(X_1, \dots, X_n) \geq \eta'.$$

Since $\eta' < \eta$ was arbitrary, we conclude that $\delta_0(X_1, \dots, X_n) \geq \eta$ as claimed. $\square$

Corollary 17. *Let (A, τ) be a finitely-generated algebra with a positive trace τ and generators $X_1, \dots, X_n$, and let $\mathrm{Der}_a(A; A \otimes A)$ denote the space of derivations from A to $L^2(A \otimes A, \tau \otimes \tau)$ which are L^2 closable and so that for some $\Xi_j \in \mathscr{F}'(R)$, $\xi \in \mathscr{F}(R)$, $R > \max_j \|X_j\|$, $\partial^*(1 \otimes 1) = \xi(X_1, \dots, X_n)$ and $\partial(X_j) = \Xi_j(X_1, \dots, X_n)$. Consider the A,A-bimodule*

$$V = \big\{(\delta(X_1), \dots, \delta(X_n)) : \delta \in \mathrm{Der}_a(A; A \otimes A)\big\} \subset L^2(A \otimes A, \tau \otimes \tau)^n.$$

Assume finally that $M = W^*(A, \tau) \subset R^\omega$. Then

$$\delta_0(X_1, \dots, X_n) \geq \dim_{M \bar{\otimes} M^o} \overline{V}^{L^2(A \otimes A, \tau \otimes \tau)^n}.$$

Proof. Let $P : L^2(A \otimes A, \tau \otimes \tau)^n \to \overline{V}$ be the orthogonal projection, and set $v_j = P(0, \dots, 1 \otimes 1, \dots, 0)$ with $1 \otimes 1$ in the j-th position. Let $v_j^{(k)} = (v_{1j}^{(k)}, \dots, v_{nj}^{(k)}) \in L^2(A \otimes A)^n$ be vectors approximating v_j and having the property that the derivations defined by $\delta(X_j) = v_{ij}^{(k)}$ lie in Der_a. Then

$$\eta_k = \dim_{M \bar{\otimes} M^o} \overline{\operatorname{span} A v_1^{(k)} A + \cdots + A v_n^{(k)} A} \to \dim_{M \bar{\otimes} M^o} \overline{V}$$

as $k \to \infty$. Now for each k, the derivations $\delta_j : A \to L^2(A \otimes A)$ so that $\delta_j(X_i) = v_{ij}^{(k)}$ belong to Der_a. Applying Lemma 6 and Proposition 8 to $T_{ij} = v_{ij}^{(k)}$ and combining the conclusion with Theorem 16 gives

$$\delta_0(X_1, \dots, X_n) \geq \eta_k.$$

Taking $k \to \infty$ we get

$$\delta_0(X_1, \dots, X_n) \geq \dim_{M \bar{\otimes} M^o} V,$$

as claimed. □

Corollary 18. *For a fixed N, and all $|q| < (4N^3 + 2)^{-1}$, the q-semicircular family $X_1, \dots, X_N$ satisfies*

$$\delta_0(X_1, \dots, X_N) > 1 \text{ and } \delta_0(X_1, \dots, X_N) \geq N\left(1 - \frac{q^2 N}{1 - q^2 N}\right).$$

In particular, $M = W^*(X_1, \dots, X_N)$ has no Cartan subalgebra. Moreover, for any abelian subalgebra $\mathscr{A} \subset M$, $L^2(M)$, as an $\mathscr{A}$,$\mathscr{A}$-bimodule, contains a copy of the coarse correspondence.

Proof. Let ∂_i be a derivation as in Lemma 10; thus $\partial_i(X_j) = \delta_{i=j}\,\Xi$, as defined in Lemma 10. Then for $|q| < (4N^3 + 2)^{-1}$, Lemma 14 shows that $\partial_i \in \mathrm{Der}_a$. Then Theorem 16 implies that

$$\delta_0(X_1, \dots, X_N) \geq \dim_{M \bar{\otimes} M^o} \overline{\sum M \Xi_i M},$$

$M = W^*(X_1, \dots, X_n)$. It is known [Shlyakhtenko 2004] that for $q^2 < 1/N$ (which is the case if we make the assumptions about q as in the hypothesis of the corollary), this dimension is strictly bigger than 1 and is no less than $N(1 - q^2 N(1 - q^2 N)^{-1})$.

The facts about M follow from [Voiculescu 1996]. □

For $N = 2$, $(4N^3 + 2)^{-1} = 1/34$. Thus the theorem applies for $0 \leq q \leq 1/34 = 0.029\dots$. Our estimate also shows that $\delta_0(X_1, \dots, X_N) \to N$ as $q \to 0$.

Corollary 19. *Let Γ be a discrete group generated by $g_1, \dots, g_n$, and let $V \subset C^1(\Gamma, \ell^2\Gamma)$ be the subset consisting of cocycles valued in $\mathbb{C}\Gamma \subset \ell^2\Gamma$. If the group von Neumann algebra of Γ can be embedded into the ultrapower of the hyperfinite II_1 factor (for example, if the group is sofic), then*

$$\delta_0(\mathbb{C}\Gamma) \geq \dim_{L(\Gamma)} \overline{V}.$$

Proof. Any such cocycle gives rise to a derivation into $\mathbb{C}\Gamma^{\otimes 2}$ by the formula

$$\partial(\gamma) = c(\gamma) \otimes \gamma^{-1}.$$

Then $\partial^*\partial(\gamma) = \|c(\gamma)\|_2^2\gamma \in \mathbb{C}\Gamma$. Moreover, the bimodule generated by values of these derivations on any generators of $\mathbb{C}\Gamma$ has the same dimension over $L(\Gamma)\bar{\otimes} L(\Gamma)$ as $\dim_{L(\Gamma)} \bar{V}$. □

For certain R^ω embeddable groups (for example, free groups, amenable groups, residually finite groups with property T, more generally embeddable groups with first L^2 Betti number $\beta_1^{(2)} = 0$, as well as groups obtained from these by taking amalgamated free products over finite subgroups and passing to finite index subgroups and finite extensions), V is actually dense in the set of ℓ^2 1-cocycles. Indeed, this is the case if all ℓ^2 derivations are inner (that is, $\beta_1^{(2)}(\Gamma) = 0$). Moreover, it follows from the Mayer–Vietoris exact sequence that amalgamated free products over finite subgroups retain the property that V is dense in the space of ℓ^2 cocycles. Moreover, this property is also clearly preserved by passing to finite-index subgroups and finite extensions. So it follows that for such groups Γ, $\delta_0(\Gamma) = \beta_1^{(2)}(\Gamma) + \beta_0^{(2)}(\Gamma) - 1$ (compare [Brown et al. 2008]).

It is likely that the techniques of the present paper could be extended to prove the following:

Conjecture 20. *Let Γ be a group generated by $g_1, \dots, g_n$ and assume that $L(\Gamma)$ can be embedded into R^ω. Let $V \subset \ell^2(\Gamma)^n$ be the subspace $\{(c(g_1), \dots, c(g_n)) : c : \Gamma \to \ell^2(\Gamma)$ 1-cocycle$\}$. Let $P_V : \ell^2(\Gamma)^n \to V$ be the orthogonal projection, so that $P_V \in M_{n\times n}(R(\Gamma))$, where $R(\Gamma)$ is the von Neumann algebra generated by the right regular representation of the group.*

Let $\mathcal{A} \subset R(\Gamma)$ be the closure of $\mathbb{C}\Gamma \subset R(\Gamma)$ under holomorphic functional calculus, and let $P_a \in \mathcal{A}$ be any projection so that $P_a \leq P_V$. Then $\delta_0(\Gamma) \geq \mathrm{Tr}_{M_{n\times n}} \otimes \tau_{R(\Gamma)}(P_a)$.

Note that with the notations of the Conjecture, $\mathrm{Tr}_{M_{n\times n}} \otimes \tau_{R(\Gamma)}(P_V) = \beta_1^{(2)}(\Gamma) - \beta_0^{(2)}(\Gamma) + 1 = \delta^*(\Gamma)$.

It should be noted that the restriction on the values of the cocycles ($\mathbb{C}\Gamma$ rather than $\ell^2\Gamma$) comes from the difficulty in the extending the results of Proposition 8 to the case of nonanalytic Ξ (though the term $\partial^*\partial(\gamma)$ continues to be a polynomial even in the case that the cocycle is valued in $\ell^2(\Gamma)$ rather than $\mathbb{C}\Gamma$).

Appendix: Otto–Villani type estimates via exponentiation of derivations

Let $M = W^*(X_1, \dots, X_N)$, where $X_1, \dots, X_N$ are self-adjoint.

Let us denote by ζ_j the vector $(0, \dots, 0, 1 \otimes 1, 0, \dots, 0) \in [L^2(M, \tau)^{\otimes 2}]^N$ (the only nonzero entry is in the j-th position). One can realize a free semicircular family of cardinality N on the space

$$H = L^2(M, \tau) \oplus \bigoplus_{k\geq 1} \big[(L^2(M, \tau) \bar{\otimes} L^2(M, \tau))^{\oplus N}\big]^{\otimes_M k}.$$

using creation and annihilation operators $S_i = L_i + L_i^*$, where

$$L_i \xi = \zeta_i \otimes_M \xi.$$

Then for $\zeta \in W^*(M) \bar{\otimes} W^*(M)$, the notation S_ζ makes sense with $S_{\zeta_i} = S_i$, $aS_\zeta b + b^* S_\zeta a^* = S_{a\zeta b + b^*\zeta a^*}$ and $\|S_\zeta\|_2 = \|\zeta\|_2$.

Let $A = \mathrm{Alg}(X_1, \dots, X_N)$. For $a, b \in A \otimes A$ and $j = 1, \dots, N$ write

$$(a \otimes b)\#S = aSb.$$

Proposition 21. *Let $\partial : A \to V_0 = [W^*(M, \tau) \bar{\otimes} W^*(M, \tau)]^{\oplus N} \subset V = [L^2(M, \tau) \bar{\otimes} L^2(M, \tau)]^{\oplus N}$ be a derivation. We assume that for each j, ζ_j is in the domain of $\partial^* : V \to L^2(M, \tau)$ and that $\partial(a^*) = (\partial(a))^*$, where $* : L^2(M) \bar{\otimes} L^2(M)$ is the involution $(a \otimes b)^* = b^* \otimes a^*$. Let $S_1, S_2, \dots$ be semicircular elements, free from M.*

Assume that $\partial(A) \subset (A \otimes A)^{\oplus N}$ and also that $\partial^*(1 \otimes 1) \in A$.

Then there exists a one-parameter group α_t of automorphisms of $M * W^*(S_1, \dots, S_N) \cong M * L(\mathbb{F}_N)$ so that $A \cup \{S_j : 1 \le j \le N\}$ are analytic for α_t and

$$\frac{d}{dt}\alpha_t(a)\Big|_{t=0} = S_{\partial(a)} \quad \text{for all } a \in A, \qquad \frac{d}{dt}\alpha_t(S_j)\Big|_{t=0} = -\partial^*(\zeta_j) \quad \text{for } j = 1, 2, \dots.$$

In particular,

$$\alpha_t(a) \cdot 1 = \Big(a - \frac{t^2}{2}\partial^*(\partial(a))\Big) + t\partial(a) - \frac{t^2}{2}(1 \otimes \partial + \partial \otimes 1)(\partial(a)) \in H.$$

Proof. Let B be the algebra generated by A and $S_1, \dots, S_N$ in $\mathcal{M} = W^*(A, \tau) * L(\mathbb{F}_N)$.

Let $P_j : V \to L^2(A \otimes A)$ be the j-th coordinate projection, and let $\partial_j : A \to A \otimes A$ be given by $\partial_j = P_j \circ \partial$.

Let $V_1, \dots, V_N \in B$ be given by

$$V_j = \sum_k \partial_k(X_j)\#S_k = S_{\partial(X_j)}, \qquad j = 1, \dots, N.$$

Let $V_{N+1}, \dots, V_{2N} \in B$ be given by

$$V_{N+k} = -\partial_k^*(1 \otimes 1) = -\partial^*(\zeta_k), \qquad k = 1, \dots, N.$$

Then $(V_1, \dots, V_{2N}) \in B \subset L^2(B, \tau)$ is a noncommutative vector field in the sense of [Voiculescu 2002a]. It is routine to check that this vector field is orthogonal to the cyclic gradient space.

We now use [Voiculescu 2002a] to deduce that there exists a one-parameter automorphism group α_t of $\mathcal{M} = W^*(B, \tau)$ such that

$$\frac{d}{dt}\alpha_t(X_j)\Big|_{t=0} = V_j \quad \text{for } j = 1, \dots, N, \qquad \frac{d}{dt}\alpha_t(S_k)\Big|_{t=0} = V_{N+k} \quad \text{for } k = 1, \dots, N,$$

and moreover that all elements in B are analytic for α_t. In particular, we see that

$$\begin{aligned}
\frac{d}{dt}\alpha_t(X_j)\Big|_{t=0} &= S_{\partial(X_j)}, \\
\frac{d^2}{dt^2}\alpha_t(X_j)\Big|_{t=0} &\cdot 1 = \delta(S_{\partial(X_j)}) = -\partial^*(\partial(X_j)) - (1 \otimes \partial + \partial \otimes 1)(\partial(X_j)),
\end{aligned}$$

as claimed. $\square$

Example 22. We give three examples in which the automorphisms α_t can be explicitly constructed. The first is the case that $X_1, \dots, X_N$ is a free semicircular system and $\partial(X_j) = (0, \dots, 1 \otimes 1, \dots 0)$ (that is, $\partial = \oplus \partial_j$, where ∂_j are the difference quotient derivations of [Voiculescu 1998]). In this case, the automorphism α_t is given by

$$\alpha_t(X_j) = (\cos t)X_j + (\sin t)S_j, \qquad \alpha_t(S_j) = -(\sin t)X_j + (\cos t)S_j.$$

Another case is that of a general N-tuple $X_1, \dots, X_N$ and ∂ an inner derivation given by $\partial(X) = [X, T]$, for $[T_j]_{j=1}^N = [-T_j^*]_{j=1}^N \in [M \bar{\otimes} M^o]^N$. Put $z = \sum T_j \# S_j$. Then α_t is an inner automorphism given by $\alpha_t(Y) = \exp(izt) Y \exp(-izt)$. Lastly, assume that $M = M_1 * M_2$ and the derivations ∂_j are determined by $\partial_j|_{M_1} = 0$, $\partial_j|_{M_2}(x) = [x, T_j]$ for some $T_j \in M \bar{\otimes} M^o$. Then again put $z = \sum T_j \# S_j$. The automorphism α_t is then given by $\alpha_t(Y) = \exp(izt) Y \exp(-izt)$. In particular, $\alpha_t|_{M_1} = \mathrm{id}$ and $\alpha_t|_{M_2}$ is given by conjugation by unitaries $\exp(izt)$ which are free from M_1 and M_2.

Proposition 21 can be used to give another proof to the Otto–Villani type estimates (Proposition 8) in the case of polynomial coefficients, using the following standard lemma:

Lemma 23. *Let $\beta_t : (M, \tau) \to (M, \tau)$ be a one-parameter group of automorphisms so that $\tau \circ \beta_t = \tau$. Let $X \in M$ be an element so that $t \mapsto \beta_t(X)$ is twice-differentiable. Finally let*

$$Z = \frac{d}{dt}\beta_t(X)\Big|_{t=0}, \qquad \xi = \frac{d^2}{dt^2}\beta_t(X)\Big|_{t=0}.$$

Then, for all t,

$$\|\beta_t(X) - (X + tZ)\|_2 \leq \frac{t^2}{2}\|\xi\|_2.$$

Corollary 24. *Assume that $X_1, \dots, X_N \in A$ and $\partial_1, \dots \partial_N : A \to A \otimes A$ are derivations, so that $\partial_j^*(1 \otimes 1) \in A$. Then we have the following estimate for the free Wasserstein distance*:

$$d_W\big((X_1, \dots, X_N), (X_1 + \sqrt{t}\sum_k \partial_k(X_1)\# S_k, \dots, X_N + \sqrt{t}\sum_k \partial_k(X_N)\# S_k)\big) \leq Ct,$$

where C is the constant given by

$$C = \tfrac{1}{2}\bigg(\sum_j \|\partial^*\partial(X_j)\|^2_{L^2(A)} + \|(1 \otimes \partial + \partial \otimes 1)(\partial(X_j))\|^2_{[L^2(A)\otimes L^2(A)\otimes L^2(A)]^{N^2}}\bigg)^{1/2},$$

where $\partial : A \to [L^2(A) \otimes L^2(A)]^N$ is the derivation $\partial = \partial_1 \oplus \cdots \oplus \partial_N$.

In the specific case of the difference quotient derivations determined by $\partial_k(X_j) = \delta_{kj} 1 \otimes 1$, we have

$$d_W\big((X_1, \dots, X_N), (X_1 + \sqrt{t}S_1, \dots, X_N + \sqrt{t}S_N)\big) \leq \frac{t}{2}\Phi^*(X_1, \dots, X_N)^{1/2}.$$

Proof. Let α_t be the one-parameter group of automorphisms as in Proposition 21. We note that

$$\bigg(\sum_j \|\alpha_{\sqrt{t}}(X_j) - (X_j + \sqrt{t}\sum_k \partial_k(X_j)\# S_k)\|_2^2\bigg)^{1/2} \leq Ct$$

in view of Lemma 23 and the expression for $\alpha_t''(X_j)$. On the other hand, $(\alpha_{\sqrt{t}}(X_1), \dots, \alpha_{\sqrt{t}}(X_N))$ has the same law as $(X_1, \dots, X_N)$, since $\alpha_{\sqrt{t}}$ is a $*$-homomorphism. It follows that

$$d_W\big(X_1, \dots, X_N, (X_1 + \sqrt{t} \sum_k \partial_k(X_1)\#S_k, \dots, X_N + \sqrt{t} \sum_k \partial_k(X_N)\#S_k)\big)$$
$$= d_W\big(\alpha_{\sqrt{t}}(X_1), \dots, \alpha_{\sqrt{t}}(X_N), (X_1 + \sqrt{t} \sum_k \partial_k(X_1)\#S_k, \dots, X_N + \sqrt{t} \sum_k \partial_k(X_N)\#S_k)\big) \le Ct.$$

In the case of the difference quotient derivations, we have:

$$\sum_k \partial_k(X_j)\#S_k = S_j, \quad (1 \otimes \partial + \partial \otimes 1)(\partial(X_j)) = (1 \otimes \partial + \partial \otimes 1)(1 \otimes 1) = 0, \quad \partial^* \partial(X_j) = \partial_j^*(1 \otimes 1).$$

Thus

$$C = \frac{1}{2} \Big(\sum_j \|\partial_j^*(1 \otimes 1)\|_2^2 \Big)^{1/2} = \frac{1}{2} \Phi^*(X_1, \dots, X_N)^{1/2}$$

as claimed. □

Acknowledgment

The author is grateful to A. Guionnet for suggesting the idea of using stationary solutions to free SDEs as an alternative form of "exponentiating" derivations, and to (patiently) explaining to him about stochastic differential equations. The author also thanks D.-V. Voiculescu for a number of comments and suggestions. I am also very grateful to the referees for a number of suggestions and improvements to the paper.

References

[Belinschi and Bercovici 2003] S. T. Belinschi and H. Bercovici, "A property of free entropy", *Pacific J. Math.* **211**:1 (2003), 35–40. MR 2004i:46097 Zbl 1058.46043

[Biane and Speicher 1998] P. Biane and R. Speicher, "Stochastic calculus with respect to free Brownian motion and analysis on Wigner space", *Probab. Theory Related Fields* **112**:3 (1998), 373–409. MR 99i:60108 Zbl 0919.60056

[Biane and Voiculescu 2001] P. Biane and D. Voiculescu, "A free probability analogue of the Wasserstein metric on the trace-state space", *Geom. Funct. Anal.* **11**:6 (2001), 1125–1138. MR 2003d:46087 Zbl 1020.46020

[Bożejko and Speicher 1991] M. Bożejko and R. Speicher, "An example of a generalized Brownian motion", *Comm. Math. Phys.* **137**:3 (1991), 519–531. MR 92m:46096 Zbl 0722.60033

[Brown et al. 2008] N. P. Brown, K. J. Dykema, and K. Jung, "Free entropy dimension in amalgamated free products", *Proc. Lond. Math. Soc.* (3) **97**:2 (2008), 339–367. MR MR2439665 Zbl 1158.46045 arXiv math.OA/0609080

[Connes and Shlyakhtenko 2005] A. Connes and D. Shlyakhtenko, "L^2-homology for von Neumann algebras", *J. Reine Angew. Math.* **586** (2005), 125–168. MR 2007b:46104 Zbl 1083.46034 arXiv math.OA/0309343

[Dykema and Nica 1993] K. Dykema and A. Nica, "On the Fock representation of the q-commutation relations", *J. Reine Angew. Math.* **440** (1993), 201–212. MR 94e:46117 Zbl 0767.46038

[Effros and Popa 2003] E. G. Effros and M. Popa, "Feynman diagrams and Wick products associated with q-Fock space", *Proc. Natl. Acad. Sci. USA* **100**:15 (2003), 8629–8633. MR 2004h:81099 Zbl 1064.81045

[Ge 1998] L. Ge, "Applications of free entropy to finite von Neumann algebras, II", *Ann. of Math.* (2) **147**:1 (1998), 143–157. MR 99c:46068 Zbl 0924.46050

[Ge and Shen 2002] L. Ge and J. Shen, "On free entropy dimension of finite von Neumann algebras", *Geom. Funct. Anal.* **12**:3 (2002), 546–566. MR 2003f:46109 Zbl 1040.46043

[Jung 2003a] K. Jung, "A free entropy dimension lemma", *Pacific J. Math.* **211**:2 (2003), 265–271. MR 2004k:46115 Zbl 1058.46044 arXiv math.OA/0207149

[Jung 2003b] K. Jung, "The free entropy dimension of hyperfinite von Neumann algebras", *Trans. Amer. Math. Soc.* **355**:12 (2003), 5053–5089. MR 2004f:46080 Zbl 1028.46096

[Jung 2007] K. Jung, "Strongly 1-bounded von Neumann algebras", *Geom. Funct. Anal.* **17**:4 (2007), 1180–1200. MR 2009i:46115 Zbl 1146.46034

[Lück 2002] W. Lück, *L^2-invariants: theory and applications to geometry and K-theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics. **44**, Springer, Berlin, 2002. MR 2003m:58033 Zbl 1009.55001

[Mineyev and Shlyakhtenko 2005] I. Mineyev and D. Shlyakhtenko, "Non-microstates free entropy dimension for groups", *Geom. Funct. Anal.* **15**:2 (2005), 476–490. MR 2006e:46073 Zbl 1094.46039

[Otto and Villani 2000] F. Otto and C. Villani, "Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality", *J. Funct. Anal.* **173**:2 (2000), 361–400. MR 2001k:58076 Zbl 0985.58019

[Ozawa 2004] N. Ozawa, "Solid von Neumann algebras", *Acta Math.* **192**:1 (2004), 111–117. MR 2005e:46115 Zbl 1072.46040

[Peterson 2009] J. Peterson, "L^2-rigidity in von Neumann algebras", *Invent. Math.* **175**:2 (2009), 417–433. Zbl 05530522

[Ricard 2005] É. Ricard, "Factoriality of q-Gaussian von Neumann algebras", *Comm. Math. Phys.* **257**:3 (2005), 659–665. MR 2006e:46069 Zbl 1079.81038

[Shlyakhtenko 1998] D. Shlyakhtenko, "Some applications of freeness with amalgamation", *J. Reine Angew. Math.* **500** (1998), 191–212. MR 99j:46079 Zbl 0926.46046

[Shlyakhtenko 2004] D. Shlyakhtenko, "Some estimates for non-microstates free entropy dimension with applications to q-semicircular families", *Int. Math. Res. Not.* **51** (2004), 2757–2772. MR 2006e:46074 Zbl 1075.46055

[Shlyakhtenko 2006] D. Shlyakhtenko, "Remarks on free entropy dimension", pp. 249–257 in *Operator Algebras: The Abel Symposium 2004*, Abel Symp. **1**, Springer, Berlin, 2006. MR 2007i:46065 Zbl 1118.46058 arXiv math.OA/0504062

[Śniady 2004] P. Śniady, "Factoriality of Bożejko–Speicher von Neumann algebras", *Comm. Math. Phys.* **246**:3 (2004), 561–567. MR 2005d:46130 Zbl 1064.46047

[Ştefan 2005] M. B. Ştefan, "Indecomposability of free group factors over nonprime subfactors and abelian subalgebras", *Pacific J. Math.* **219**:2 (2005), 365–390. MR 2009e:46060 Zbl 1098.46050

[Voiculescu 1994] D. Voiculescu, "The analogues of entropy and of Fisher's information measure in free probability theory, II", *Invent. Math.* **118**:3 (1994), 411–440. MR 96a:46117 Zbl 0820.60001

[Voiculescu 1996] D. Voiculescu, "The analogues of entropy and of Fisher's information measure in free probability theory, III: The absence of Cartan subalgebras", *Geom. Funct. Anal.* **6**:1 (1996), 172–199. MR 96m:46119 Zbl 0856.60012

[Voiculescu 1998] D. Voiculescu, "The analogues of entropy and of Fisher's information measure in free probability theory. V. Noncommutative Hilbert transforms", *Invent. Math.* **132**:1 (1998), 189–227. MR 99d:46087 Zbl 0930.46053

[Voiculescu 1999] D. Voiculescu, "The analogues of entropy and of Fisher's information measure in free probability theory, VI: Liberation and mutual free information", *Adv. Math.* **146**:2 (1999), 101–166. MR 2001a:46064 Zbl 0956.46045

[Voiculescu 2002a] D. Voiculescu, "Cyclomorphy", *Int. Math. Res. Not.* **6** (2002), 299–332. MR 2003a:46092 Zbl 1027.46089

[Voiculescu 2002b] D. Voiculescu, "Free entropy", *Bull. London Math. Soc.* **34**:3 (2002), 257–278. MR 2003c:46077 Zbl 1036.46051

Received 10 Jan 2008. Revised 10 Nov 2008. Accepted 24 Mar 2009.

DIMITRI SHLYAKHTENKO: `shlyakht@math.ucla.edu`
Department of Mathematics, University of California, Los Angeles, Los Angeles, CA 90095, United States